APPLICATIONS OF CLUSTERING ANALYSIS TO SIGNAL PROCESSING PROBLEMS

WING-KEUNG SIM

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF PHILOSOPHY

IN THE DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

THE CHINESE UNIVERSITY OF HONG KONG

1999

Abstract

Clustering analysis is the generic name for a wide variety of procedures that can create an unsupervised classification. The objective of clustering is to sort the waveform data or feature extracted data into groups which have a high natural association among the data members. This dissertation is concerned with the application of clustering to signal processing problems.

The first application studied was the problem of reconstructing electrophyiological neural spike data. The complex networks of neural cells are studied by physiologists who record spike signals using a single microelectrode. Such an electrode will record the spike activities of several neurons and the mixed spike trains recordings need to be demultiplexed using a clustering analysis methodology. A quantitative comparison of clustering techniques in term of accuracy and computational requirements was made.

The second application was the evaluation of headphone sound quality. A fast and reliable technique to identify the best sounding headphones and discover the most important measurable parameters was developed. Clustering analysis was applied in such a way that similar sounding headphones were clustered into a manageable number of group and samples were selected from each groups for subsequent measurement and subjective listening tests. Finally, the statistical relationship between perceived sound quality and measured parameters was discovered.

摘要

“類聚分析”是一种無需要管理訓練的分類方法。它的算法目的是要把有圖案狀態的數据重新排列成几組有高度內部連系的數据組。這論文主題是要將類聚分析方法應用于信號處理的問題中。

第一個應用問題是要重新建立于生理學實驗室內收集的腦神經細胞信號。生理學家利用單一微電子探測器收集腦神經細胞信號，然后再用這些信號數据去進行复雜腦神經細胞网絡研究。單一微電子探測器會于同一時間內收集多于一個腦神經細胞所釋放出的信號。類聚分析方法能夠將些混集多個腦神經細胞信號重新排列成獨立腦神經細胞信號。這頁論文將不同的類聚分析方法在分類腦神經細胞信號應用上的精确性和運算速度作深入和全面的比較。

第二個應用問題是要評估耳机的音質。這頁論文提供了一個有效率和可靠的方法去分類出不同音質的耳机以及尋找出最重要的量度數据。類聚分析方法應用于從大量耳机中選出最有代表示性的樣品作爲隨后測量和傾听測驗的樣本。最后，本論文利用統計學方法尋找和表達耳机音質与耳机量度數据的關系。

Acknowledgements

I would like to acknowledge the kind support of my supervisor, Dr. Philip Leong, for his innovative idea of using neural network in electrophysiological spike discrimination and conducting statistical measure in headphone evaluation. His continuous support and advice is invaluable to the success of these projects. Without him, timely and successful completion of the projects would not be possible.

Regarding the online electrophysiological spike discrimination project, I would like to thank Dr. Paul Martin of the Department of Physiology, University of Sydney for providing the data used in the experiments as well as many useful discussions regarding spike sorting. I would also like to thank Prof. Lai-Wan Chan of the Chinese University of Hong Kong for discussions on this work and suggestions. The research on online spike discrimination was supported by the Chinese University of Hong Kong direct grant.

Regarding the sound quality measurement in headphones project, I would like to thank the generous assistance from Globe Audio Company for providing the technical support and test equipment and Dr. Dallas Lam for his help in headphones measurement and sample collection. I would like to especially thank Dr. Yiu-San Moon for his help in conducting listening tests.

Contents

List of Figures

List of Tables

CHAPTER 1

Introductions

1.1 Motivation & Aims

The aim of this thesis was to verify the applicability of clustering analysis methods in some ill-posed signal processing problems in which the optimal solution is still unknown. Signal processing problems often involve classifying data from a pool of observations which in most cases contain much variability and noise. Clustering analysis can often be used in the analysis of such signals, either as a preprocessing step or to directly obtain information from the signals. In this dissertation, applications of clustering to two signal processing problems are presented.

The first problem was the demultiplexing of mixed neural spike signals recordings into separate spike trains. It is an important problem for physiologists who study the complex assembly of biological neural network systems. Clustering was performed on the spike signals in order to separate the contributions from different neurons. The commonly employed clustering techniques include template matching, K-means and artificial neuron networks. In this thesis, a quantitative comparison of different clustering algorithms for this problem was performed.

Another interesting signal processing problem is concerned with quality control in headphone production. Headphone sound quality evaluation is a difficult problem since whether a headphone sounds 'good' or not is very subjective. Clustering was applied to classify similar

sounding headphones into a manageable number of groups so that representative samples could be selected for subsequent listening tests and measurements.

1.2 Contributions

Techniques of clustering analysis were verified for their ability to solve practical signal processing problems, namely the electrophysiological spike discrimination and headphone sound quality control problems.

Multi-neuron extracellular recordings were collected from nervous system of monkey using single microelectrode. Four online clustering methods, namely K-means, simple competitive learning (SCL), ART2 and online template matching, were tested their ability to demultiplex these recordings in term of accurately and reliably. Rigorous comparisons of the clustering techniques was given based on their robustness to noise and ability to classify real spike data. Also, the feature extraction process using principal component analysis (PCA) was studied for its ability in improving accuracy and reliability of classifications. This is the first study in which a rigorous comparison of a number of standard clustering techniques was compared in this problem domain. It is also the first study to quantify the effects of PCA.

A method was studied for identifying correlation between subjective listening tests and measurable parameters of mass produced headphones. The method can identify parameters which directly affect the subjective sound quality of headphones. This makes it possible to for a computer to measure the sound quality of headphones. Clustering techniques were applied so that a manageable size of representative headphone samples could be selected for listening tests. To the best of my knowledge, this is the first published technique for applying clustering analysis to help finding a direct link between measurable parameters and subjective sound quality.

1.3 Structure of Thesis

Firstly an introduction to the electrophysiological spike discrimination problem is presented in chapter 2. This chapter begins with an introduction to the cellular physiology, shows

mechanism which cause action potentials and describes the recording of multi-neuron spike signals using a single microelectrode. Reviews of previous work on demultiplexing mixed spike signal are also presented.

In chapter 3, the problem commonly finding correlation of perceived headphone sound quality with physical parameters is described. A description of headphone characterization is also presented.

Chapter 4 explains the principle and implementation detail of principal component analysis as well as the two categories of clustering methods namely traditional clustering methods (template matching and K-means) and unsupervised neural learning (simple competitive learning and ART2).

The experimental setups for spike discrimination and headphone evaluation are presented in chapter 5. All of the experimental results are included in chapter 6. The performance of different clustering techniques in classifying real spike data is presented, together with results of correlating measurable parameters to headphone sound quality.

In the final chapter, a summary of the work in this thesis and directions for future work is presented.

CHAPTER 2

Electrophysiological Spike Discrimination

2.1 Introduction

This chapter begins with an introduction to cellular physiology and the mechanisms responsible for the spiking of neurons. The current practice of physiologists in recording multi-neuron signals using single microelectrode is then introduced followed by the process of demultiplexing multi-neuron data into separate wave trains. Finally, an overview of previous work is given.

2.2 Cellular Physiology

This section provides an introduction to the cellular physiology. This description is presented in a very informal way with an aim to provide only enough material to achieve enough understanding of ideas in this thesis and explain some of the basic terminology. There are several good introductory textbooks on this topic (Jennett S., 1989; Vander A. J et. al., 1970).

2.2.1 Action Potential

The basic processing unit of the nervous system is called a neuron. There is a complex network inside the brain and spinal cord of animals. The neurons communicate via electric signals in the form of action potentials. All the action potentials starting from a particular neuron are identical with the same amplitude, the same duration and propagation velocity. An action potential is generated when some chemical component concentrations inside the cell reach a predominantly excitatory level. For instance, the normal electrical potential difference

of a cell is about -70mV (inner cell potential relative to the outer cell potential). The chemical movement between a cell and its surrounding is controlled by the semi-permeable cell membrane which is more permeable to K^+ ions than to Na^+ ions. The inner K^+ potential is higher than the outside and K^+ ions tend to leak out. Na^+ ions are also far from equilibrium but its inward movement toward the cell is hindered by the low permeability of the cell membrane.

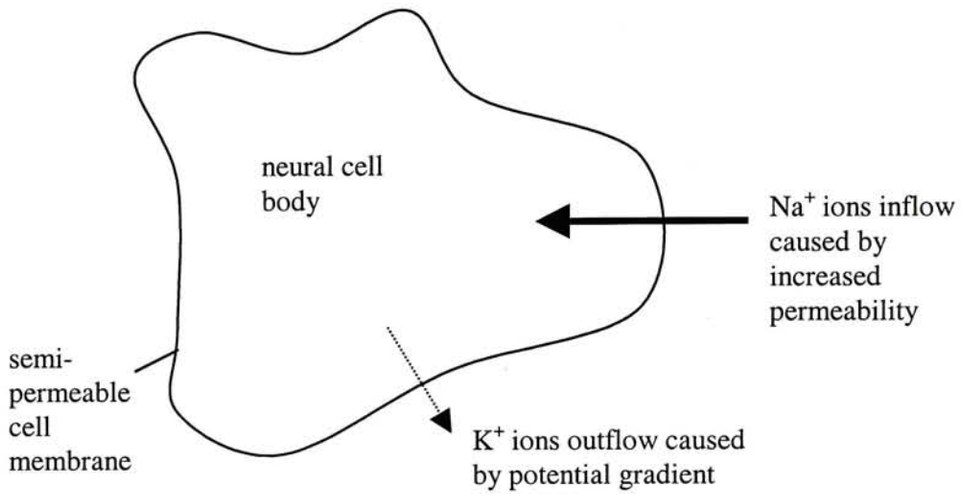

Figure 2-1: Neural cell body showing K^+ and Na^+ ions movement during depolarization

Given this background, a sudden increase of permeability to Na^+ ions by the cell membrane (caused by an external excitatory stimulus) can cause a positive swing of charge potential inside the cell, which is known as a depolarization (see Figure 2-1). When the degree of depolarization reaches a critical level, the cell will respond with an action potential (see Figure 2-2). After reaching a maximum potential, the permeability to Na^+ decreases sharply and the permeability to K^+ becomes dominant again and the charge potential returns to the normal state.

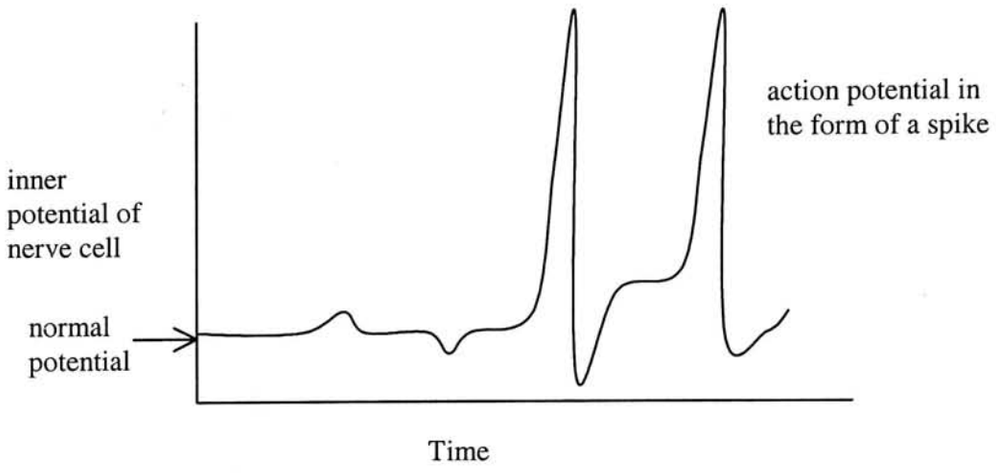

Figure 2-2: Change of inner cell potential and spiking action potential

2.2.2 Recording of Spikes Activities

Spikes coming out of a nerve cell can be recorded by measuring the electrical change alongside or inside the nerve fibers. Some mammalian cells are large enough for spikes recording by inserting a microelectrode inside the cell body. This is called an intracellular recording (Figure 2-3).

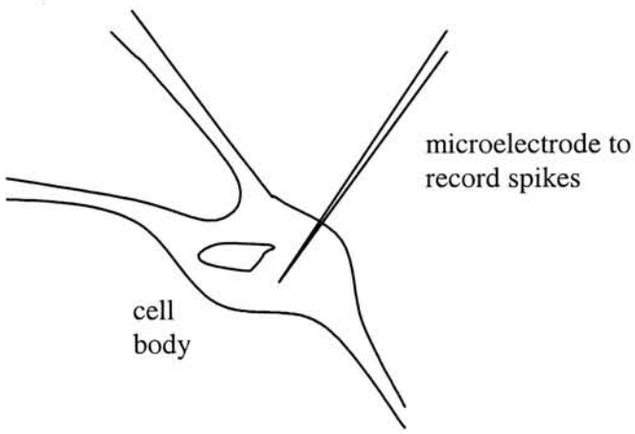

Figure 2-3: Intracellular Recording

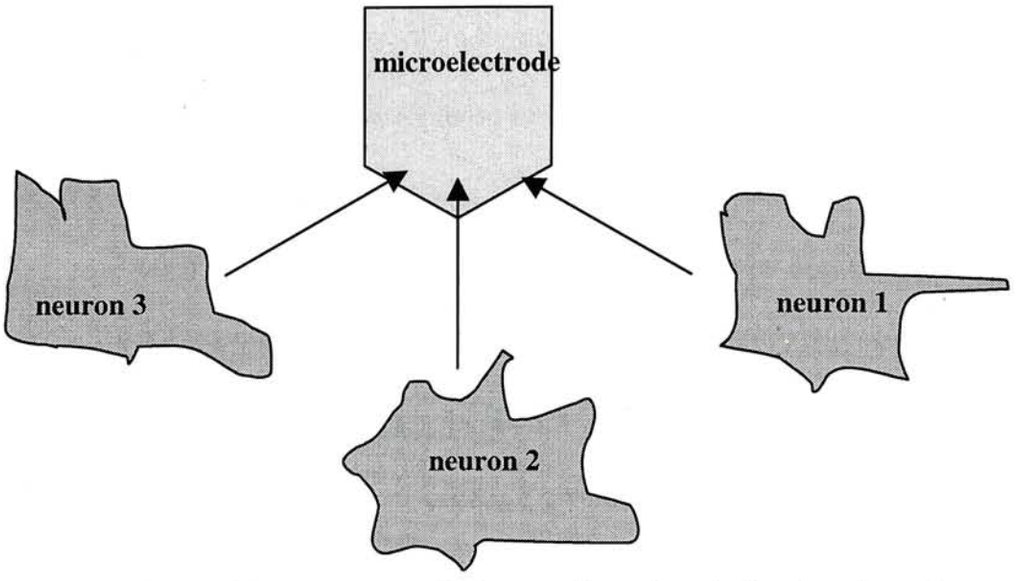

Figure 2-4: Multi-neuron extracellular recording using single microelectrode

The spike activities can also be recorded with electrodes placed outside the cells. This is called an extracellular recording (Figure 2-4). Spikes from nearby neurons so recorded have to be distinguished from background neuronal activities and noises. Spiking activities from one or in most cases several nearby neurons may reach the electrode simultaneously. Spike trains recorded in this manner are called multi-neuron extracellular recordings.

Extracellular recordings of neuronal spiking activity are becoming increasingly important in studies of multi-neuronal activity. An increasing number of laboratories are starting to conduct multi-neuron experiments in which extracellular recordings are made simultaneously from up to 30 individual neurons (Schmidt, 1984b). This kind of experiment provides a direct method to determine the basic properties and functions of neuronal assemblies in the brain.

The neurons to be studied in this kind of experiments are generally close together, probably within 100• m of distance (Schmidt, 1984b). Several distinct spike trains may be recorded by a single microelectrode. Several neuronal signals from different neurons may enter the electrode simultaneously. Each neuronal signal exhibits different spike shapes. Based on these factors, it is possible to recognize spike signals originating from different neurons even though they are measured from the same electrode.

2.2.3 Demultiplexing of Multi-Neuron Recordings

If one decides to study the activity of individual neurons and possible interactions, some technique must be employed to separate these multi-unit records into single-unit activities. By demultiplexing such recordings adequately, these recordings allow the activities of multiple closely spaced neurons to be recorded from a single electrode as opposed to the usual case where only a single neuron can be monitored. This extra information may be an important tool for gaining improved understanding of the interactions between neurons. Also, the yield of experiments can be improved greatly since a given electrode can record several individual spike trains.

However, all the above-mentioned improvements can be realized only when a reliable and computationally efficient waveform-sorting algorithm is found. It had been extremely difficult in the past to measure several closely spaced neurons with the use of several single neuron electrodes. Waveform classification is by now the only solution to monitor neuronal activities of several spatially contiguous neurons.

2.3 Application of Clustering for Mixed Spikes Train Separation

The most critical problem in using multi-neuron extracellular recordings is that of demultiplexing individual neuronal activities from the recorded spike trains. There are many approaches using both software and hardware as discussed in Section 2.4 of this chapter. This section presents the basic design principles to discriminate spike signals and discussed the applicability of clustering analysis techniques.

2.3.1 Design Principles for Spike Discrimination Procedures

Ideally a solution to this spike-sorting problem should have the following features:

- real-time - use an efficient online computation algorithm to facilitate operation in a real time environment

- accurate - discard the occasional corrupted spike but form new templates for genuine spikes of a new class
- unsupervised - operate with minimal human intervention
- adaptive - adapt reference patterns (templates) dynamically to track changes in the shape of a spike over time due to movement of the electrodes or changes in the membrane potential.
- cost and convenience – should run on standard personal computers so no special hardware is required.

In this thesis, different clustering techniques were employed to demultiplex the spike signals in a hope to find the best algorithm satisfying the above design principles.

2.3.2 Clustering Analysis

Clustering analysis is the generic name for a wide variety of procedures that can create an unsupervised classification (see e.g. Everitt, 1974). Clustering analysis methods empirically form groups which have high similarity. Its objective is to sort the waveform data into groups such that the degree of natural association is high among members of the same group and low between members of different groups.

Clustering techniques have been widely used in many engineering problems where large data sets are difficult to handle and so methods of summarizing and extracting relevant information are important. The mixed spike train demultiplexing procedure can be accomplished by clustering the different spikes by shape, each different shape assumed to be associated with a different neuron.

When clustering data units, the similarity measure of individuals is usually expressed as Euclidean distance. Most traditional clustering methods (like template matching and K-means) have operational interpretations in this similarity measure (which is the Euclidean measure) while others are rather less intuitive, like artificial neural networks (like simple

competitive learning and ART2). Refer to chapter 4 for detailed description of the above-mentioned clustering methods.

A substantial practical problem in performing a clustering analysis is to decide on the correct number of clusters in the data set. Some clustering methods give a configuration for every number of clusters from one cluster (the entire data set) up to the number of entities (each cluster has only one data member) (see Mirkin B., 1996). Other algorithms decide automatically a best fitting number of clusters (like ART2). Some algorithms (the best example is K-means) begin with a user-selected number of groups and then find the best fitting clustering structures with it.

2.3.3 Comparison of Clustering Techniques

One of the goal of this thesis is to present a comprehensive comparison of current clustering algorithms feasible for the sorting of spike trains from multi-neuron. Based on the above design criteria, four clustering techniques (namely online template-matching, K-means, simple competitive learning and ART2) were chosen to be compared in our experiment for their feasibility and suitability in sorting out wave trains from the mixed extracellular recordings. The comparison is based on computational cost and accuracy.

2.4 Literature Review

This section is aimed at provide a throughout review of previous work on discriminating spike data using different clustering approaches and implementation hardware.

2.4.1 Template Spike Matching

Many different techniques have been proposed for the automatic waveform classification of such signals (see Schmidt (1984b) for a review). The earliest techniques discriminated between waveforms based on the amplitude of several selected points (Schmidt, 1984a). Computer based methods have focussed on template matching in which templates which are representative of single neuron spikes are constructed and spikes classified by comparison with the templates (Schmidt, 1984b). The experimenter selects a spike as a standard or

averages a number of spikes from the database to form the standard spike for comparison. The computer then compares this standard to each spike in the record and calculate the degree of similarity by making out the weighted mean square difference over all the sampling point of the spike. Those spikes showing high similarity with the standard spike are classified as coming from the same neuron. For spikes vary significantly from standard spikes and appear with considerable frequency, new standard spikes are selected to represent their family in the data set. This process will be repeated until as many units as possible are classified. Both online and batch (offline), adaptive and non-adaptive methods have been proposed for producing templates. The algorithms used have included generating new templates when no good match with existing templates are found (Millecchia et al., 1978), the K-means clustering algorithm (Salganicoff et al., 1988, Sarna et al., 1988), linear filters (Gozani and Miller, 1994) and hierarchical clustering (Fee et al., 1996).

2.4.2 Reduced Feature Matching

All the data points (it can be up to 64 data points per spike) of the spikes may not be necessary for reliable spike sorting. One technique for reducing a spike to a smaller set of features is called principal component analysis. This technique constructs a set of orthogonal basis functions or principal components to represent the original spike with the least mean-square error. Each spike can be represented by the first 2 or 3 principal components of the waveforms. Marks (1965) described the way of using a computer preprocessor that used a 3-channel transversal filter to generate the first 3 components of the spike waveform. Other reduced feature matching techniques include amplitude separation (Wyss, U.R. and Handwerker, 1971) in which the spike is sorted in terms of peak amplitude. Fourier analysis (Bessou, P. and Perl, E.R., 1969) is another feature reduction technique which classifies the spike into different categories according to the first few (usually 8) Fourier components of sine series of each spike waveforms. Some techniques make use of additional spike properties like spike area and rms value along with peak amplitude to improve sorting accuracy. Among this variety of techniques, many are computationally expensive. Most of the previous

implementations of the system are analog in nature while some of them employ simple digital pre-processor before going through the processing stage in computer. The new-coming generation of spike sorting system will most probably rely increasingly on high performance digital computer systems and hence be digital in nature.

2.4.3 Artificial Neural Networks

Various different types of artificial neural networks (ANNs) have also been successfully applied to this problem including backpropagation networks (Gozani and Miller, 1994), Kohonen Self Organizing Maps (SOM) (Öhberg et. al., 1996) and ART2 networks (Oghalai et. al., 1994). Many of these algorithms employ principal component analysis (Schmidt et al., 1984) as a preprocessing step, serving to reduce the dimensionality of the patterns to be clustered as well as extract only the most salient features of the waveform. Some of ANN algorithms, like backpropagation, employ a supervised learning rule, which requires an offline training stage with a selection of training data set prior the online classification stage (Jansen R. F., 1990). SOM and ART2 algorithms employ self-organizing association learning rule and in an online fashion.

2.4.4 Hardware Implementation

The earliest spike sorting systems were implemented in hardware to meet real-time constraints and are analog in nature, however, their accuracy was limited and they did not adapt to changing spike patterns. A real time digital system circa 1995 (Gadicke and Albus, 1995) used two AT&T DSP32C floating point digital signal processing chips per channel to demultiplex up to 8 spikes in real-time. In the past when the speed of microprocessors in desktop computers was not fast enough, online computer based methods for spike sorting have relied on custom high-speed digital pre-processors to share part of computations so that the real-time requirements could be met. When high-speed computers became available to the neurophysiologist, numerous complex computational algorithms emerged. Recently, real-time implementations based solely on standard personal computers (Öhberg et al., 1996) and workstations (Oghalai et al., 1994) have been described. Note that a standard modern desktop

or laptop computer (e.g. a 450 MHz Pentium II processor) has much more computing power than a DSP32C. High performance computer becomes the main computing hardware for spike sorting and the increasing computing resource enhances the application of more sophisticated sorting algorithms (like ANN) in the tradeoff of speed to accuracy.

All of the previously proposed techniques were found to have their own limitation for clustering spikes. They also vary in terms of classification accuracy and computational requirements.

2.5 Summary

This chapter introduced basic terminology in cellular physiology and discussed basic concepts in multi-neuron electrophysiological experiments. The application of clustering techniques to demultiplex the mixed multi-neuron spike trains was presented. Finally, relevant previous research work was presented to conclude this chapter.

Chapter 3

Correlation of Perceived Headphone Sound Quality with Physical Parameters

3.1 Introduction

In this chapter, the problem of general sound quality evaluation and difficulties specifically related to mass production headphones are presented. Important physical parameters for characterization of headphones, including frequency response characteristics, harmonic distortion and driver voice-coil parameters are introduced and a brief review of the technique of statistical correlation is given at the end of the chapter.

3.2 Sound Quality Evaluation

The perceived sound quality of audio equipment and its relation to various physical properties of audio systems has long been a subject for much discussion and debate (Garbrielsson A. et al., 1985 and 1990; Tsujimoto K., 1986; Nakayama Y., 1988; Yajima M., 1997; Thurmond B. et. al., 1992 and Tannaka Y., 1990). The sound quality of an audio system is a very subjective measure not easily characterized by quantitative measurements. The psycho-acoustical properties of sound-reproducing system like loudspeakers and headphones can be studied using subjective listening tests in which human subjects are invited to select the best sounding headphones solely based on their subjective judgements. The main psychological preference

among common listeners to the evaluation of sound quality can be determined by calculating the statistical correlation between the listening test results and measurable parameters.

The perceived sound quality of sound-reproducing system like loudspeakers, headphones and hearing aids is multidimensional (Garbrielsson A. et. al., 1990). It depends on a number of perceptual dimensions like clarity, fullness, spaciousness, brightness versus dullness, softness versus sharpness, absence of extraneous sounds. One method is to use humans to assess the perceived sound quality. From (Garbrielsson A. et al., 1985 and 1990), experienced listeners were instructed to rate the sound quality using perceptual scales and evaluative scales (fidelity, pleasantness). The possibility of defining a reliable and valid scale of sound quality rating was demonstrated.

The complex relations between the perceived quality and various physical properties of audio systems are still largely unresolved. Among all the physical variables, measurement of frequency response is most commonly employed design technique. For audio system engineers, the flatness and extent of frequency response is usually considered to be critical to good sounding performance. There have been many studies on the importance of frequency response characteristics (Charles S. and Gengel R. W., 1969; Moller H. et. al., 1995; Harrison J., 1996; Kates J. M., 1984; Harwood H. D., 1976; Tannaka Y. et. al., 1989 and Garbrielsson A. et. al., 1990). Garbrielsson et. al. (1990) showed that largely different frequency responses can affect all perceived sound quality dimensions in his designed listening. A criterion for evaluating loudspeaker performance based on frequency response measure and auditory central spectrum model was presented by Kates (1984) and the audibility of phase-frequency effect in loudspeakers was studies by Harwood (1976). To date, solid evidence of the ability of frequency response to discriminate between good and poor sounding audio equipment has not been found. More research is required to show how the frequency response can affect their corresponding overall psycho-acoustic performance.

Other commonly used design criteria include Thiele-Small parameters and harmonic distortion. The most well known technique is frequency response control of loudspeaker

systems using Thiele-Small parameters for vented box (Small, 1973) and closed box (Small, 1972) loudspeakers. Following Small's work, Leach (1979) presented a technique to determine Thiele-Small parameters from the specifications of moving-coil drivers, that is the system cutoff frequencies, volume of the cavity behind driver, driver area and system electrical impedance. Leach's work further enhanced design of loudspeaker system from the specification of moving coil drivers. There are also design techniques based on minimization of distortion effects. Birkett and Goubran (1996) proposed a method to compensate for nonlinear elements in loudspeaker systems by improving the acoustic echo canellation at high frequencies using a nonlinear adaptive filter. Other approaches include using a second order Volerra compensators, a feedforward nonlinear digital inverse circuit, for reduction of second order harmonics (Schurer H. et. al., 1995) and distortion-canceling differential amplifier circuit (Goto T. et. al., 1975).

It is very difficult to implement an efficient and reliable quality control procedure for audio equipment. It has in the past been done by quality control personnel listening to test signals. The sound quality evaluation technique by listening tests has the shortcoming that human ear perception has great variation among different people. Fatigue and emotional change can also alter the judgement result. Therefore, reliance on single person's judgement in this quality evaluation process is unreliable. Multiple test subjects (subject number most possibly in order of tens) are needed when conducting listening tests to average out individual variation among the test group, achieving a more reliable judgement than a single quality controller. Techniques to select samples precisely and highly simplify the evaluation process are necessary to optimize the production process.

A factory may produce millions of headphones annually so the headphone evaluation process must be highly efficient and automatic with minimum human intervention. An automatic quality control system can speed up the feedback mechanism of the production quality status and other relevant information to the production line. If significant correlation between

subjective listening test result and some physical parameters can be found, this discovered knowledge can greatly improve the yield rate in further production.

3.3 Headphone Characterization

This section describes the basic terminology used in characterizing sound quality of headphones. It is aimed at providing enough information to understand the ideas of this thesis and many simplifications have been made in the description.

3.3.1 Frequency Response

Frequency response parameters are most frequently used when comparing the sounding quality of different audio equipment. The amplitude, phase, delay and transient response characteristics are interrelated and this means that changing one of the parameters can affect the others. For audio system engineers, the flatness and extent of the frequency responses is usually considered to be critical to good psycho-acoustic performance.

Frequency response is most commonly measured by performing a Fourier transform (Kreyzig, 1980) on the impulse response of the system (S.R.L. A., 1998).

3.3.2 Harmonic Distortion

Harmonic distortion parameter is a direct function of a system's non-linearity. It is one of the most important parameters of audio amplifiers. For a pure linear system, an input single-tone signal will give rise to a linearly generated output. The output signal is of the same frequency as the input signal.

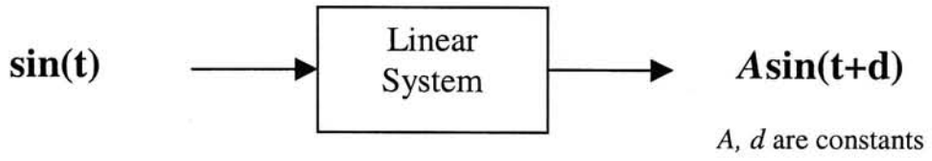

Figure 3-1: Linear system

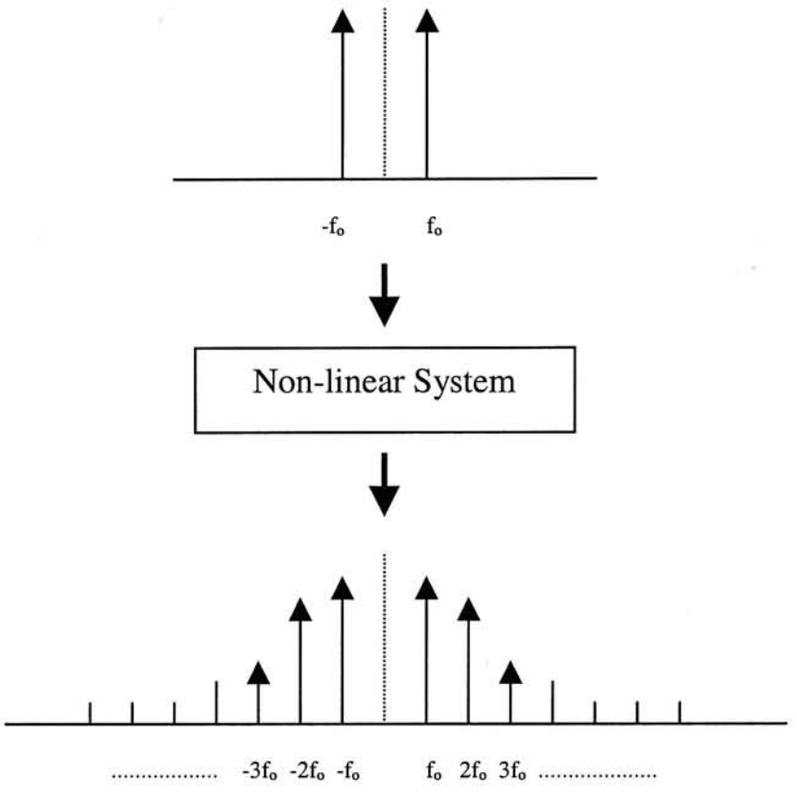

Figure 3-2: Non-linear system

Non-linearity gives rise to output signal components at multiple frequencies of the input. The frequency component at the first multiple ($2f_o$) of the input signal frequency (f_o) is called the second harmonic while the component at the second multiple ($3f_o$) is known as the third harmonic. The power of harmonic component with a unit power input signal is a measure of system distortion. The total harmonic distortion (THD) is the sum of the distortion signals at all harmonic frequencies.

3.3.3 Voice-Coil Driver Parameters

An electrical equivalent circuit of a voice-coil driver is shown in Figure 3.3. R_E is the dc resistance of headphone voice coil, R_g is the output resistance of power source, C_{MES} is the electrical capacitance due to the drive mass, L_{CES} is the electrical inductance due to driver compliance and R_{ES} is the electrical resistance due to the driver suspension.

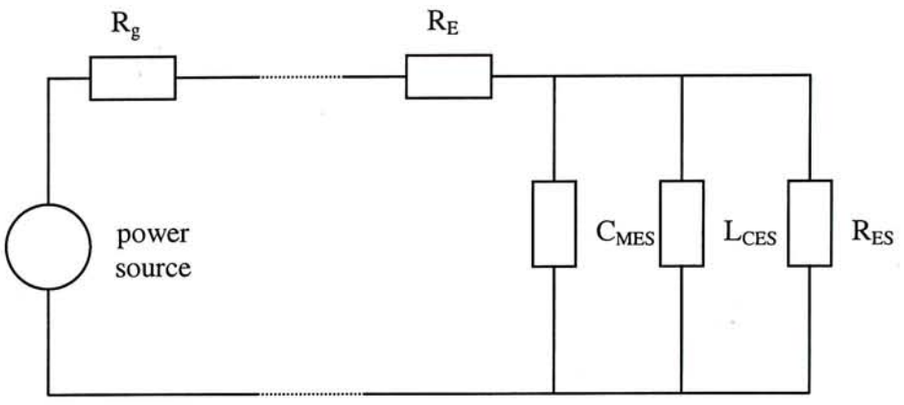

Figure 3-3: Electrical equivalent circuit of voice-coil driver (Small, 1972)

The resonant frequency, F_s, of a voice-coil driver is the frequency at which the impedance-frequency characteristic function (a second-order function) of the voice-coil driver in free air has the maximum magnitude and zero phase (Small, 1971). It is also the frequency at which the maximum power is output from the driver. Figure 3.4 shows a typical amplitude-frequency plot of a driver. The acoustic output remains relatively constant at frequencies higher than F_s. The acoustic output reaches a maximum value at F_s. At frequencies below F_s, the output power goes down sharply.

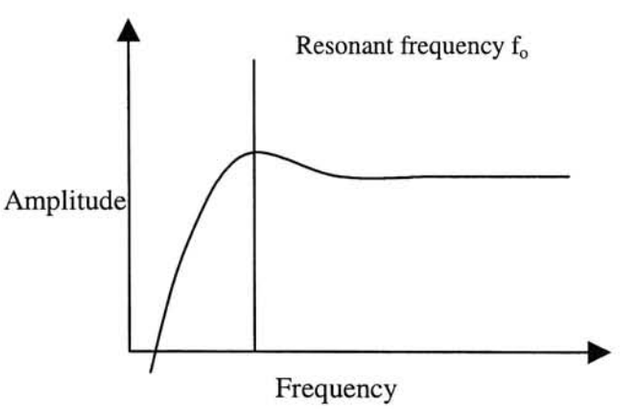

Figure 3-4: Resonant frequency in amplitude-frequency plot

The ratio of reactive energy to resistive (or damping) energy in an electrical system is referred to as Q. Driver voice-coil is partially characterized by its Q value in the electrical equivalent circuit. High-Q voice-coils have relatively large amounts of reactive energy and have higher tendency to ring. They tend to have heavy moving mass and stiff suspensions and relatively

small voice-coil driver. The reverse is true for low-Q drivers. Low-Q drivers tend to be damped and sound tight. They have lighter cones, soft suspensions, and magnetically strong driver.

Q_{ES} is the ratio of voice-coil dc resistance to the reflected motional reactance of the driver (Small, 1972) at F_s. It refers to the total Q of voice-coil driver at frequency F_s considering only electrical resistance in the driver electrical equivalent circuit and is given by (Small, 1972):

$$Q_{ES} = 2\bullet\ F_S C_{MES} R_E$$

Q_{MS} is ratio of driver electrical equivalent frictional resistance to the reflected motional reactance of the driver at F_s. It refers to the total Q of driver at frequency F_s considering only non-electrical resistance in the driver electrical equivalent circuit and is given by (Small, 1972):

$$Q_{MS} = 2\bullet\ F_S\ C_{MES} R_{ES}$$

Q_{TS} is ratio of driver total electrical equivalent resistance (both R_E and R_{ES} in Figure 3.3) to the reflected motional reactance of the driver at F_s. It refers to the total Q of driver at frequency F_s considering all system resistance in the driver electrical equivalent circuit and is given by (Small, 1972):

$$Q_{TS} = \frac{Q_{ES} Q_{MS}}{Q_{ES} + Q_{MS}}$$

The measurement of driver voice-coil parameters (R_E, F_s, Q_{es} , Q_{ms} , Q_{ts} , L_{1k} and L_{10k}) was based on derivation from impedance-frequency response of the driver (Small, 1972). For details of the measurement system, please refer to CLIO operator's manual (A.S.R.L., 1998).

3.4 Statistical Correlation Measurement

This section presents the basic concept of statistical correlation relevant to this thesis.

3.4.1 Correlation Coefficient

When two variables are closely related, like the height and weight of a human, they are statistically correlated. The correlation coefficient r is a single figure of merit to measure the

degree of statistical correlation between two random variables. The correlation coefficient r between two sets of variable, X and Y is given by the ratio:

$$r = \frac{\mathrm{cov}(x, y)}{\sqrt{\mathrm{var}(x).\mathrm{var}(y)}}$$

where x and y are samples of variable X and Y, and n is the number of samples of variable X and Y, and

$$\mathrm{cov}(x, y) = (\sum (x_i - \bar{x})(y_i - \bar{y}))/n$$

$$\mathrm{var}(x) = (\sum (x_i - \bar{x})^2)/n$$

$$\mathrm{var}(y) = (\sum (y_i - \bar{y})^2)/n$$

where n is the number of samples taken for variable X and Y and

$\bar{x}$ and $\bar{y}$ are mean of variable samples x and y.

The value of r lies between -1 and 1. Statistically independent variables will yield zero for the covariance value, so r is zero. The absolute covariance value of a pair of linearly related variables can be decomposed into the product of variance of the variables. Therefore, a closer value of r to +1 indicates a stronger positive linear relationship between variable X and Y; the closer is r to -1, the stronger is the negative linear relationship between X and Y.

3.4.2 *t* Test for Correlation Coefficients

The t test is a hypothesis test for a small number of samples. From elementary statistics (e.g. Williams, 1996), the t test for testing correlation coefficients of a two sets of variable samples is given by:

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}}$$

where n is the number of paired observations and r is the correlation coefficient.

From the value of t thus computed, a level of significance can be found using a table which is included in most statistics textbooks (Williams, 1996). For $n=8$ and if a 95% level of significance two-tailed test is desired, r must be larger than 0.632 for the null hypothesis to be rejected (i.e. there is significant correlation between the parameter and the listening test).

3.5 Summary

In this chapter the problem of measuring sound quality of audio system quantitatively was presented. Standard techniques used for characterizing headphones were also presented along with a review of the process of statistical correlation.

Chapter 4

Algorithms

4.1 Introduction

This chapter presents an introduction to the concepts of principal component analysis and some chosen clustering methods. The theory of principal axis transform, which is the core mechanism for PCA, is first introduced followed by an outline of how they can be applied for dimensionality reduction. Two categories of clustering techniques, namely the traditional clustering methods (including online template matching and K-means) and unsupervised neural learning (including simple competitive learning and ART2) are then presented.

4.2 Principal Component Analysis

Principal Component Analysis (PCA) is the optimal linear transform in a mean-squared error sense for dimensionality reduction. PCA preprocessing is often used in signal processing to extract the most important features from the input data while at the same time reducing the dimensionality of the input vector to be processed (see e.g. Jolliffe, 1986).

This section begins with discussion of dimensionality reduction in classification problem, followed by explanation of principal component analysis (PCA) transformation and implementation details of PCA algorithm.

4.2.1 Dimensionality Reduction

In many practical problems, pre-processing can have a profound effect on the performance of a classification problem is that adding new input feature variables beyond certain point can

actually lead to a reduction in the performance of the final system (Jurgen S., 1996). First of all, it is often the case that the input variables are generally correlated in some way, so that the data points do not fill out the entire input space but tend to be restricted to a sub-space of lower dimensionality. This is the fundamental reason that dimensionality reduction is feasible. Secondly, the value of the input variables will not change arbitrarily from one region of input space to another, but will typically vary smoothly as a function of the input variables. Thus it is possible to infer the values of the input variables at intermediate points by a process similar to interpolation when there is missing data.

Reducing the number of input variables can normally lead to improved performance for a given data set even though information is being discarded. The dimensionality reduction process can help many practical implementations of most of the signal processing system. Moreover, most of the computational process can be done in a much lower dimensional space, greatly saving computational resources.

4.2.2 PCA Transformation

The geometrical space of raw input data in most signal processing systems is initially determined by the basis vector system established incidentally during the measurement process. Presumably, this coordinate system may not be the optimally suitable one for subsequent processing. The measured data usually need to go through a transformation process in order to switch to a new basis vector system. The principal-axis transform fulfills the following optimization criterion: among all possible linear transformations, this transform has the property of making the individual variances of the transformed feature vector elements maximally non-uniform. This means most of information are stored in a few number of transformed vector element which can be considered as a form of information compression method.

The detail of the PCA transform is explained below based on the statistical approach described by Jurgen (Jurgen S., 1996):

The source data $\{\boldsymbol{v}\}$ is a distribution of data point in N-dimensional space centered at the data mean $\boldsymbol{\mu}$ and with orientations and widths given by the common covariance matrix $\boldsymbol{K}$ (computed by cov($\boldsymbol{v}$,$\boldsymbol{v}$), refer to Section 3.4 for explanation of cov). This covariance matrix $\boldsymbol{K}$ determines an orthogonal system of basis vectors (the principal axes), which can be shown to be in the optimum coordinate system for representing the stochastic process $\{\boldsymbol{v}\}$.

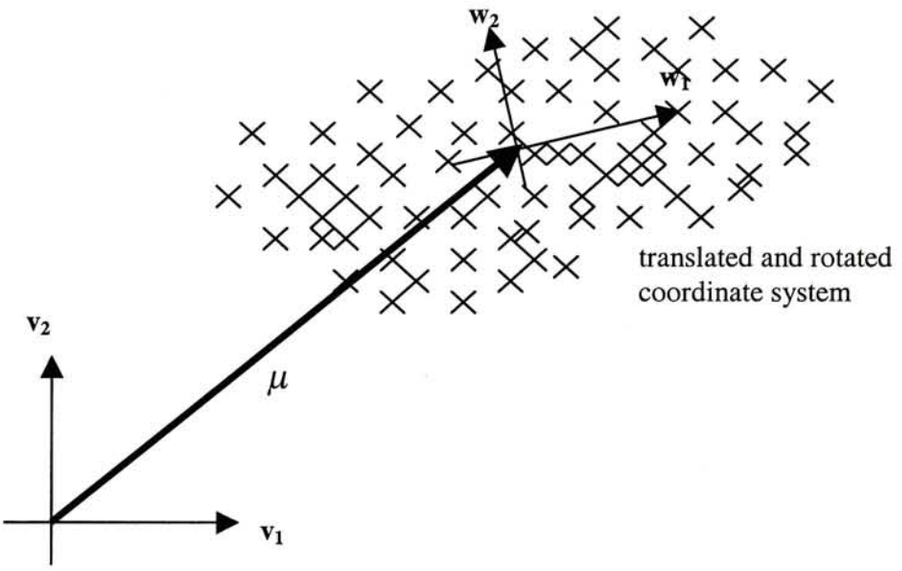

Figure 4-1: Principal axis transform. *v* is measured vector, *w* is transformed vector.

The covariance matrix $\boldsymbol{K}$ can be decomposed into the product $\boldsymbol{K}=\boldsymbol{B}\boldsymbol{D}\boldsymbol{B}^T$ of matrices $\boldsymbol{B}$ and $\boldsymbol{D}$. $\boldsymbol{B}$ is the orthogonal matrix of the eigenvectors of $\boldsymbol{K}$; that is each column vector in $\boldsymbol{B}$ is eigenvector of the corresponding row vector in $\boldsymbol{K}$. $\boldsymbol{D}$ is the diagonal matrix of the of eigenvalues of $\boldsymbol{K}$; that is each element in diagonal of $\boldsymbol{D}$ is the eigenvalue of the corresponding row vector in $\boldsymbol{K}$.

$$\boldsymbol{B}=(\boldsymbol{b}_1\ \boldsymbol{b}_2 \ldots \boldsymbol{b}_N)$$

$$D=\begin{bmatrix} \lambda_1 & 0 & . & 0 \\ 0 & \lambda_2 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & \lambda_N \end{bmatrix}$$

The N eigenvalue-eigenvector pairs $[\lambda_n,\boldsymbol{b}_n]$, $n=1,\ldots,N$, are ordered by decreasing eigenvalues.

In the given measurement geometrical space, a new coordinate system is introduced. The data are translated from the N-space origin into new center $\boldsymbol{\mu}$ and rotated according to the eigenvectors of $\boldsymbol{K}$ as represented in the matrix $\boldsymbol{B}$ of eigenvectors.

The feature vector $\boldsymbol{w}$ is computed from the given measurement vector $\boldsymbol{v}$ by

$$\boldsymbol{w}=\boldsymbol{B}^T(\boldsymbol{v}-\boldsymbol{\mu}).$$

The new data set $\{\boldsymbol{w}\}$ obtained by the above-stated transformation have zero mean value and uncorrelated:

$$\boldsymbol{\mu}=0 \text{ and } \operatorname{cov}\{\boldsymbol{w},\boldsymbol{w}\}=\boldsymbol{D}.$$

The optimal dimensionality reduction is by truncation of matrix $\boldsymbol{B}$ to $\boldsymbol{B}_M$ which consists of the M dominant eigenvectors in matrix $\boldsymbol{B}$.

$$\boldsymbol{B}_M = (\boldsymbol{b}_1\, \boldsymbol{b}_2 \ldots \boldsymbol{b}_M), \text{ where } M \leq N.$$

Performing the principal-axis transform with the truncated matrix $\boldsymbol{B}_M$,

$$\boldsymbol{w}_M=\boldsymbol{B}_M^T(\boldsymbol{v}-\boldsymbol{\mu}),$$

leads to the dimensionality reduction of the transformed data set $\boldsymbol{w}$ from N *to* M

The original measurement vector $\boldsymbol{v}$ can be approximately reconstructed from the feature vector $\boldsymbol{w}_M$ by:

$$\hat{\boldsymbol{v}} = \boldsymbol{B}_M \boldsymbol{w}_M + \boldsymbol{\mu}$$

The resulting reconstruction error is the expected value of difference between the original measurement vector and the reconstructed vector:

$$R^2 = E\{|\hat{\boldsymbol{v}}-\boldsymbol{v}|^2\}$$

The principal-axis transform is optimum with respect to the reconstruction error R^2 among all possible linear transformation of the coordinate system for a required dimensionality reduction from N to M. The mean-square reconstruction error R^2 can be measured by accumulating those eigenvalues λ_n belonging to those discarded eigenvectors $\boldsymbol{b}_n$, where $n=M+1,\ldots,N$. The reconstruction error R^2 is minimum if all the column vectors in $\boldsymbol{B}$ with largest eigenvalues are reserved.

4.2.3 PCA Implementation

The steps required to apply PCA for a set of input N×1 (N is dimensionality of input) vectors $\boldsymbol{x}^p$ *(p=1 .. P)* are as follows :

STEP 1: Calculate the mean vector

$$\bar{x} = \frac{1}{P}\sum_{i=1}^{P} x^i$$

STEP 2: Compute the covariance matrix

where ^{T} is the matrix transpose function and × represents matrix multiplication.

$$K = \frac{1}{(P-1)}\sum_{i=1}^{P}\left(x^i - \bar{x}\right)\times\left(x^i - \bar{x}\right)^T$$

STEP 3: Compute the eigenvalues and eigenvectors of $\boldsymbol{K}$. We will call the eigenvector corresponding to the *i'th* largest eigenvalue of $\boldsymbol{K}$ *as* $\boldsymbol{b}_i$, the *i'th* principal component. The principal components are orthogonal vectors in that space accounting for the maximum amount of variance in the data.

STEP 4 (Dimensionality reduction):

The principal components set is used to transform the given data set in orientation corresponding to the maximum variation of data. By reserving those components with biggest variation, reduction in dimensionality can be accomplished.

With a given input vector $\boldsymbol{x}$, we can express it as a weighting of the first L principal components $(L \leq P)$ as $\boldsymbol{w}$:

$$w_i = b_i^T\left(x - \bar{x}\right) \qquad i = 1,2,\ldots L$$

The $\boldsymbol{w}$ vector of length L is a reduced dimensionality representation of $\boldsymbol{x}$ that can be approximately reconstructed by:

$$x \approx \bar{x} + \sum_{i=1}^{L} b_i w_i$$

In the experiments detailed later, if PCA preprocessing is specified, $\boldsymbol{w}$ is used instead of $\boldsymbol{x}$ as the input to the clustering routine.

4.3 Traditional Clustering Methods

This section provides the information of two commonly used clustering methods: online template matching and K-means. Implementation detail of these methods is also included.

In the following descriptions, an Euclidean distance measure is the distance between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ that is defined by:

$$\|x - y\| = \sqrt{\sum_{i}^{1...N} (x_i - y_i)^2}$$

4.3.1 Online Template Matching (TM)

The template-matching algorithm used in this thesis was an online modification of an algorithm proposed by Millecchia and McIntyre (1978). This template-matching algorithm classifies on the basis of the similarity of the spike waveform with a set of previously determined template waveforms.

An overview of the template-matching clustering algorithm is given below: an automatic online template matching procedures in which no prior assumptions are made about the distribution of waveform distribution within a single cluster is outlined as follows (see Figure 4.2 for flow diagram). If an incoming waveform is less than a user-defined distance D (an Euclidean distance squared measure is used) from one or more templates, that spike is classified as belonging to that class with that closest matching template. The template is then adapted using a weighted average between the template and the new spike. A new template is generated when a spike occurs which is greater than D from all of the other templates.

The simple clustering process is controlled by one user-set parameter: the threshold number for promoting new template spike. The clustering process is sensitive to the setting of this threshold value. Only in cases when there is dramatic change in signal to noise ratio among recordings, the threshold value remained unchanged after being tuned to a specific value suitable for subsequent clustering. The threshold value can be selected by comparing two known spikes form same neuron using Euclidean measure (or SUMDIF method in Rene F.,

1991). The minimum difference that one can expect to find between two identical wave forms can be estimated by calculating standard deviation of the noise and multiply by the number of data points on a wave form (16 in our case). This minimum difference value is selected as the threshold and it serves to provide the noise margin for our classification.

The template is adapted by averaging the template vector with the new coming spike with a adaptation factor k, that is*: template vector = template vector* × *(1-k) + new spike vector* × k. This adaptation reduces the chance of multiple representation of what is actually a single class of spikes. Normally, the adaptation factor k is a number close to zero but in the interval between 0 and 1, allowing for gradual modification of spike template without the risk of corrupting templates.

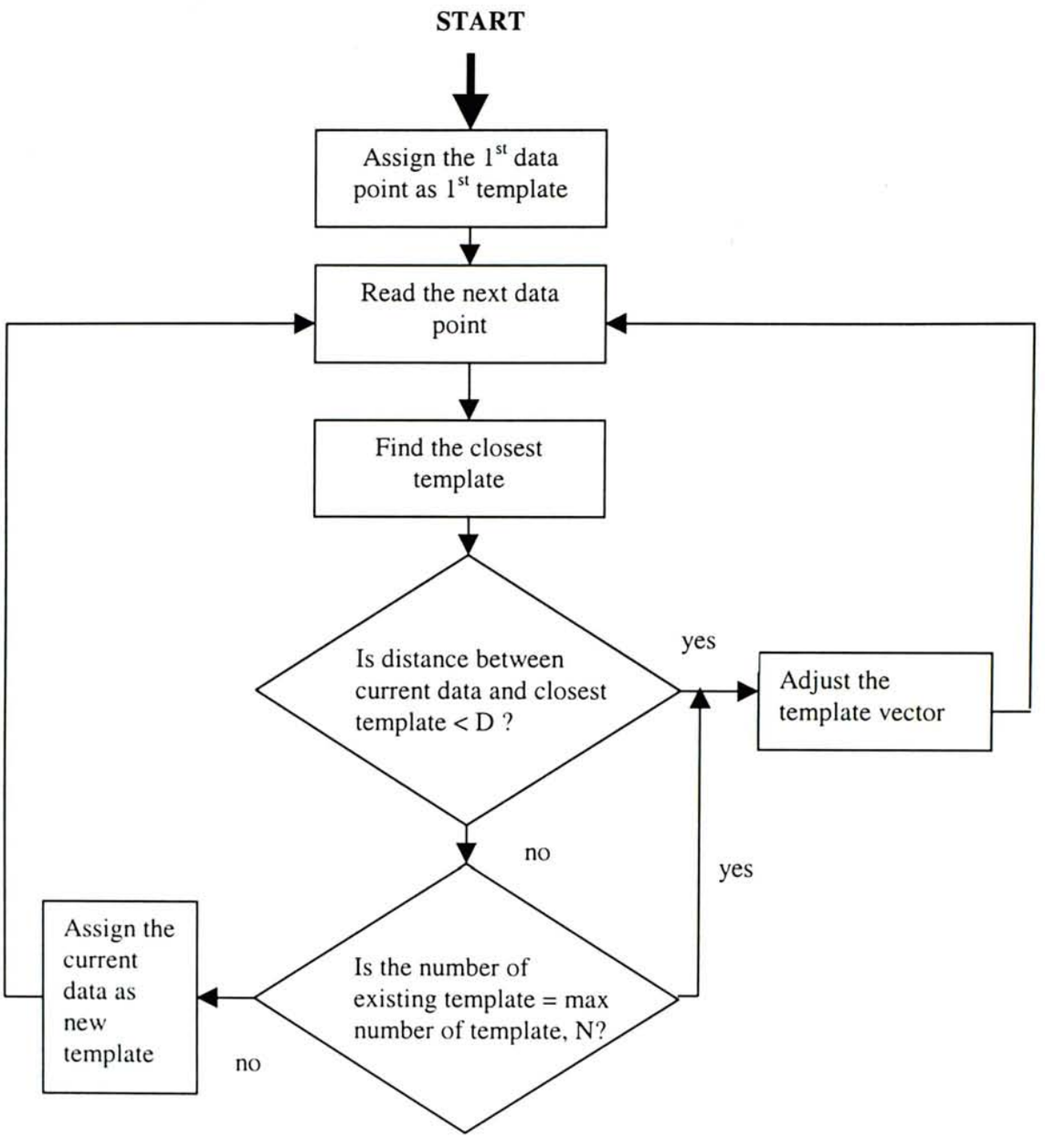

Figure 4-2: Flow diagram of online template matching

4.3.2 Online Template Matching Implementation

The p'th input waveform pattern is represented by $\boldsymbol{x}^p$, a column vector of length N and $1 \leq p \leq P$ where P is the total number of patterns in the training set.

The class the input $\boldsymbol{x}^p$ belong to is represented by C_p.

The online template-matching algorithm can be expressed as follows:

Step 1 (Initialization):

J: number of templates = 1;

$\boldsymbol{T}_1$: 1st template = $\boldsymbol{x}^1$;

C_1: cluster number of the1st input vector $\boldsymbol{x}^1$ = 1;

D: user-selectable threshold number for promoting new template;

M: the maximum number of templates allowed.

In our experiment, D = 1.5.

STEP 2 (Classification):

For each of the J templates $\boldsymbol{T}_j$, compute the squared distance d_j to the input vector $\boldsymbol{x}^p$ by:

$$d_j = \left\| x^p - T_j \right\|^2$$

Find the template $\boldsymbol{T}_k$ closest to the vector $\boldsymbol{x}^p$, namely the one satisfying the inequality:

$$d_k \leq d_j \quad \forall j$$

In the event of more than one template obeying the above inequality, a random template is chosen those having shortest distance with input pattern.

STEP 3 (Threshold test):

If $d_k > D$ and $J \leq M$, then go to Step 4;

Otherwise, $C_p = k$ and return to Step 2.

STEP 4 (New template):

The incoming input pattern is promoted as new template by:

$J = J + 1$;

$\boldsymbol{T_J = x}$;

and $C_p = J$.

4.3.3 K-Means Clustering

The K-means clustering algorithm (Hartigan, 1975) minimizes the sum of squared errors within K clusters (user-defined). This clustering method is a standard benchmark for comparison of various clustering techniques. It is Euclidean measure based and is intuitive in understanding its algorithms even for newcomer to engineering field. K-means is applied in both online spike discrimination problem and headphone sound quality measurement.

The algorithms proceed as follows: the algorithm starts by performing a rough initial clustering based on the sum of the data points in the waveform. For the case of spike sorting problem, the waveform is neuron spike. For the case of headphone sound quality measurement, the waveform is frequency magnitude response. This sum is used to group the training waveforms into one of the K initial clusters, where K is a user-selected number of cluster to be divided among the data set. A test is performed on each waveform by moving it from its current cluster to every other cluster and calculating the resulting total error. The cluster is moved (if necessary) to the cluster which gives the smallest total error. This local optimization procedure is repeated for each waveform until the algorithm iterates over the entire training set without changing the cluster assigned to any waveform.

The K-means algorithm proceeds as follows: Assume the initial clusters are C_1, C_2, ... C_k arbitrarily. The cluster mean for each cluster are assumed as $\boldsymbol{y_1}$, $\boldsymbol{y_2}$, ... $\boldsymbol{y_k}$ respectively.

The total error of the partitioning are then computed by summing up the Euclidean distance of each waveform from its corresponding cluster mean, that is by:

For every cluster C_i,

$$E_i = \sum_{x \in C_i} \|x - y_i\|^2$$

Total partition error: $E = \Sigma E_i$

Assume the number of waveforms in cluster C_i is N_i.

For every waveform in the data set, the change caused by transferring waveform from current

$$L_j = \frac{N_j \|x - y_j\|^2}{N_j + 1} - \frac{N_i \|x - y_i\|^2}{N_i - 1}$$

cluster to every other cluster is measured, that is by:

L_j is the error change in transferring the waveform number from present cluster C_i to cluster C_j.

If the minimum value of Lj is negative for all $j \neq i$, the waveform is transferred from cluster Ci to Cj. The cluster mean $\mathbf{y}_i$ and $\mathbf{y}_j$ are adjusted according if changes occur.

If no movement of waveform from one cluster to another occur any more, the clustering process is completed. Otherwise, the comparison process repeats itself until a minimum total partition error E is found.

The correct number of cluster for waveform set can be determined visually. A better systematic way is done by the indication of the F_{ratio}.

$$F_{ratio} = \frac{E_{K+1clusters}}{E_{KClusters}}$$

This is a ratio of total partition error of K+1 clusters over that of K clusters. If the ratio is above 10, then K+1 or more clusters should be necessary and the calculation continues. For F-ratio below 10, the K-clusters are adequate for waveform set. Waveform template is then the cluster mean of each clusters.

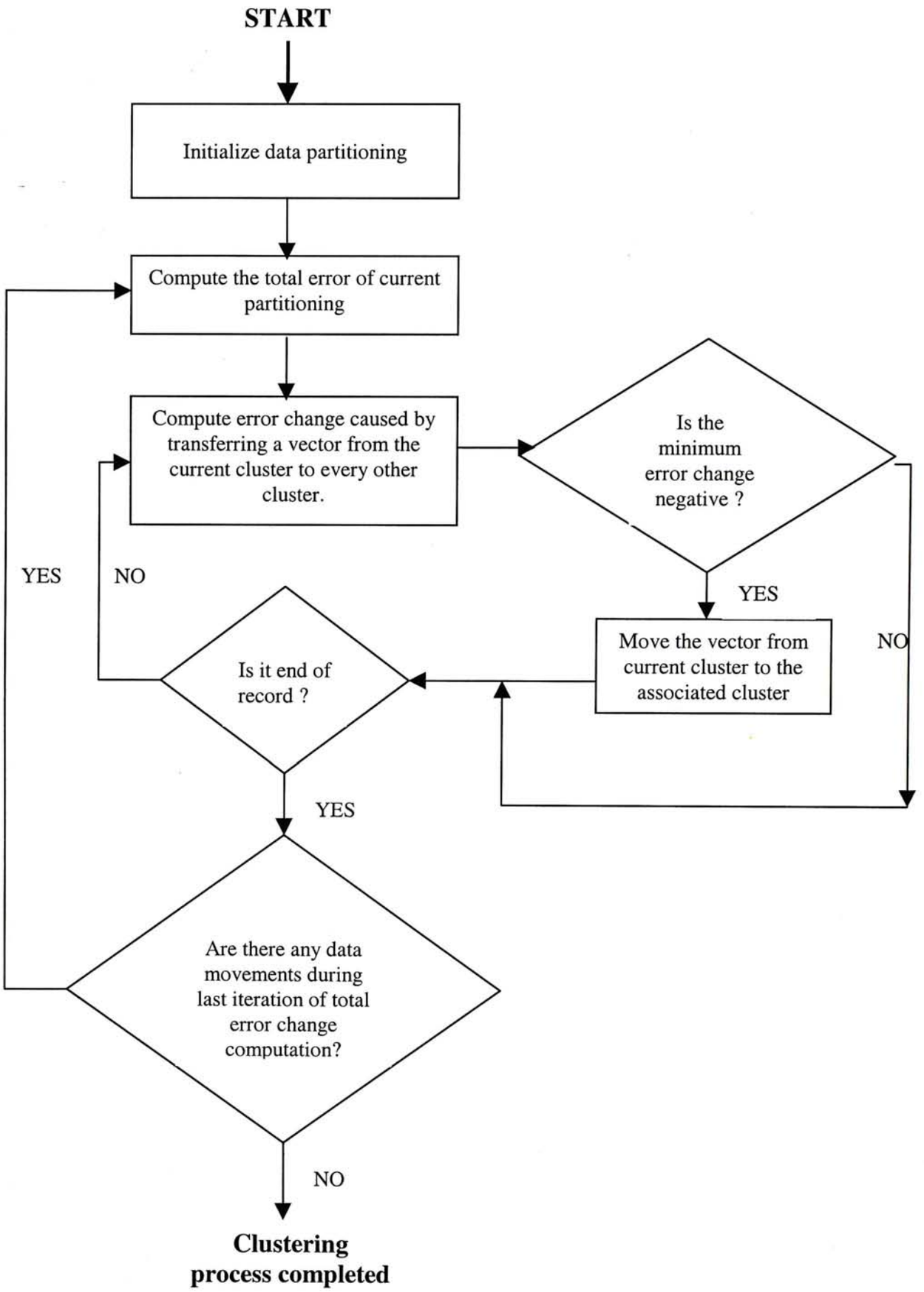

Figure 4-3: Flow diagram of K-means clustering

4.3.4 K-Means Clustering Implementation

The p'th input waveform pattern is represented by $\boldsymbol{x}^p$, a column vector of length N and $1 \leq p \leq P$ where P is the total number of patterns in the training set.

The class the input x^p belong to is represented by C_p.

The K-means algorithm can be described as follows:

STEP 1 (Initialization):

Each input pattern $\boldsymbol{x}^p$ is initialized to an initial cluster in the following way.

For every input $\boldsymbol{x}^p$ compute the sum of its scalar components $\boldsymbol{s}^p$ by:

$$s^p = \sum_i \left(x_i^p\right)$$

Also compute the maximum and minimum values of $\boldsymbol{s}^p$:

$$\boldsymbol{M} = \max(\boldsymbol{s}^p),$$

$$\boldsymbol{m} = \min(\boldsymbol{s}^p).$$

The initial clustering of each $\boldsymbol{x}^p$ is given by:

$$C^p = 1 + \text{floor}(K(s^p - m)/(M - m))$$

where floor(a) is the truncated integer part of a, and K is the user-selectable number of clusters.

STEP 2: Let $\boldsymbol{U}_k$ be the set of all input vectors assigned to cluster k, $\boldsymbol{M}_k$ be the mean vector of $\boldsymbol{U}_k$ and $\boldsymbol{N}_k$ the number of vectors in $\boldsymbol{U}_k$.

Compute the error associated with each input $\boldsymbol{x}^p$ by:

$$E^p = \left\| x^p - M_{C^p} \right\|^2$$

STEP 3: Compute the total error of the present partitioning by:

$$E = \sum_p E^p$$

STEP 4: For each input $\boldsymbol{x}^p$, compute the error $L^p{}_k$ which would result if it were transferred from its present cluster ($\boldsymbol{C}_p$) to another cluster k.

$$L_k^p = \frac{N_k \left\| x^p - M_k \right\|}{N_k + 1} - \frac{N_{C^p} \left\| x^p - M_{C^p} \right\|}{N_{C^p} - 1}$$

If the minimum of L_k^p is negative then transfer x^p from cluster C^p to cluster k and update M_{Cp}, M_k, C^p and E accordingly.

STEP 5: If there is a pass going through all p's without changing the clustering of any input, then stop; otherwise return to step 2.

4.4 Unsupervised Neural Learning

This section begins with basic concepts of neural computational learning and artificial neural network (ANN) model. The theory and implementation detail of simple competitive learning (SCL) and adaptive resonance theory (ART) are also presented.

4.4.1 Neural Network Basics

As first demonstrated by McCulloch and Pitts in 1943, the interactions among neurons in a network can be described by a logical calculus (James A. F. and David M. S., 1991). Networks of interconnected model neurons can therefore be designed and utilized to perform complicated calculations. Although the electrical properties of the artificial neurons used in these neural networks are usually quite different from real biological networks, the computing power of such networks can be highly impressive.

A network of artificial neurons built based on the above description can learn to associate a *given* set of different inputs with a given output. This kind of learning is called *supervised* learning. An offline training procedure with a given training data set is needed. After the network has learned a set of associations, it can use this newly acquired 'knowledge' to evaluate and classify new sets of inputs.

Another kind of learning is called *unsupervised* learning. During the training, no desired outputs are provided. The network learns automatically the associations among all the input

elements. This process is also referred as self-organized training or clustering, which will be described in more detail in later sections.

4.4.2 Artificial Neural Network Model

The individual computational elements that make up most artificial neural system model are often called nodes or processing elements (PEs) (James A. F. and David M. S., 1991). Figure 4.4 shows a general PE model. Each PE is numbered as shown in the figure. Each PE can have multiple input but only one single output. Each output terminal can fan out to multiple PEs in the network. The input stimulus the PE1 receives from the PE2 is indicated as x_j, which is also the output of the PE2. Each connection to the PE1 has associated with it a multiplication factor called a weight. The weight on the connection from the PE2 to the PE1 is denoted w_{ij}. The inputs to PE are classified into various types. An input connection may be excitatory or inhibitory. In general, excitatory connections have positive weights and inhibitory connections have negative weights.

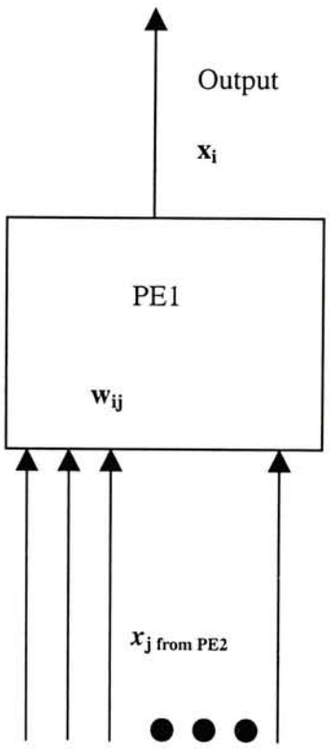

Figure 4-4: A single PE in an artificial neural network

Each PE determines its input value based on all input stimuli from its input terminals. In the absence of special connections, the net input value is given by summing all the input stimuli, each of which multiplied by their corresponding weighting in the synapse.

The net input n_1 to the PE1 unit can be written as:

$$n_1 = \sum_j x_j w_{ij}$$

where the index j runs over all input connection terminals to the PE. Excitation and inhibition are accounted for automatically by the sign of the weightings.

This sum-of-products calculation accounts for the core part of computation of network. Because there are often huge number of interconnects in a network, the speed at which this computation can be performed determines the network simulation performance.

Once the net input of the PE is calculated, the output value can be determined by:

$$x_i = f_i(n_i),$$

where f_i is called the output function and is generally a nonlinear function.

4.4.3 Simple Competitive Learning (SCL)

The neural structure model discussed in this section is known as competitive network. The network structure is simply composed of three layers: an input layer, a competitive layer and an output layer. This simple network architecture using unsupervised learning is referred to as simple competitive learning (SCL). An overview of network structure is given as follows (refer Figure 4.5): the input layer in our network is simply a set of pass-through units responsible only for distributing input data to other processing elements. The hidden middle layer is a set of processing elements that share the general properties of PEs discussed in previous sessions. This layer of processing units serve to classify all the input vectors because the position of the unit giving the strongest response for any given input identifies the region of space in which the input vector lies. In this layer, rather than examining the response of each unit to determine which is the largest, the task is simpler since the unit with the largest response is the only units to have a nonzero output. This effect can be accomplished by

having each unit competing with one another for the privilege of turning on. We call this a winner-take-all policy. No external judge exists to decide which unit has the largest output value. The unit must decide among themselves who is the winner. This decision process requires communication among all the units on the layer. The last layer is the output layer. This output layer is simply for the task of transferring the output vector to the outside. The output of L3 node represents a specific category.

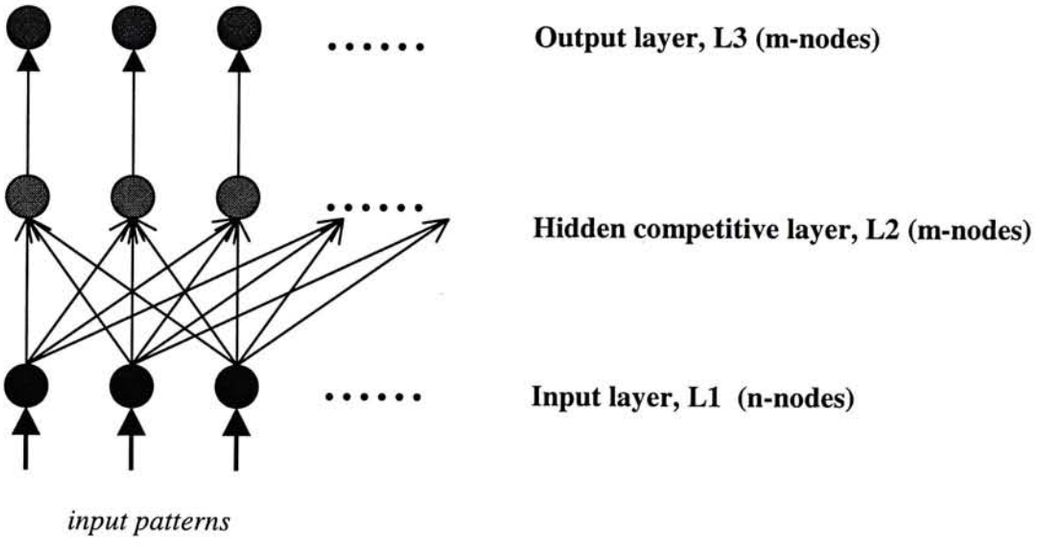

Figure 4-5: Unsupervised competitive layer structure

There are n nodes in input layer L1 that is also the dimensionality of input data. Full interconnection is made between L1 and L2. Each connection from L1 node to L2 node is unidirectional and associated with it a weighting number $\boldsymbol{W}$. L2 and L3 contain the same number of nodes m where m is a user-selected number of clusters to be divided among the input pattern set. Each L2 node is connected to the L3 node above it. The weighting associated with this connection is 1 and is unchanged, thus no learning is allowed in L3. When each input pattern is presented, the network learns the associations in the input pattern set by adapting the weightings between L2 and L1. The winner-take-all strategy allows only the winning node in the competitive layer L2 to adapt its weightings. Nodes in L2 compete with each other by comparing the magnitude of their output activation value. Only the largest

response-giving L2 node send output signal to L3. If there is a tie, we randomly select one node from those giving the same largest value of response.

The learning rule applied to adapt the weightings $\boldsymbol{W}$ is:

$$\Delta W(i, j) = \eta(x^p(j) - W(i, j))$$

where η is a user defined constant called the "learning rate".

η must be chosen in the interval [0,1]. The larger the learning rate, the more sensitive the network to new patterns and the faster the convergence. As the learning rate approaches 0, the stability of old pattern associations increase, but new-coming associations are less likely to be discovered.

If the weight changes $\boldsymbol{\Delta W}$ described above are accumulated over the entire set of input patterns before they are applied (i.e. a batch update is performed), the SCL algorithm becomes identical to the K-means algorithm (see e.g. Hertz et al., 1991). Thus this version of SCL can be regarded as an online version of K-means.

4.4.4 SCL Implementation

This SCL algorithm uses a single layer competitive neural network trained using the standard competitive learning rule (see e.g. Hertz et al., 1991). The network model in mathematical terms is shown in Figure 4.6.

The p'th input waveform pattern is represented by $\boldsymbol{x^p}$, a column vector of length N and $1 \leq p \leq P$ where P is the total number of patterns in the training set.

The class the input x^p belong to is represented by C_p.

The algorithm proceeds as follow: the inputs pattern $\boldsymbol{x^p}$ received by input layer L1 and is then propagated to L2 for multiplication with a weight matrix $\boldsymbol{W}$. The product is then summed with a set of bias values $\boldsymbol{B}$ to give vector $\boldsymbol{O}$. The competitive layer L2 of the network is a winner-take-all layer, is used to compute an output classification vector $\boldsymbol{a}$ from vector $\boldsymbol{O}$. $\boldsymbol{a}$ is of length N with exactly one output being set to 1 and all others zero and is finally passed to output layer L3. The cluster of the input pattern $\boldsymbol{x^p}$ is assigned as class C_p if a_p is set.

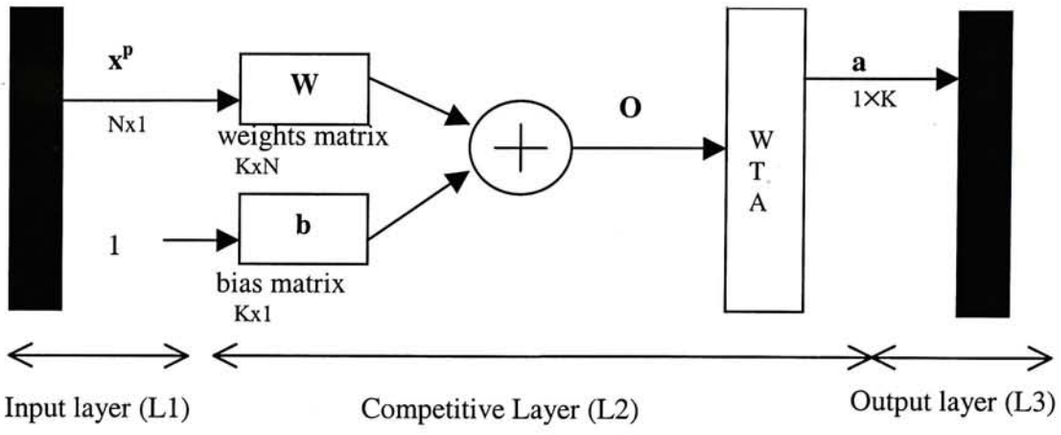

Figure 4-6: SCL network architecture

Mathematically, $\boldsymbol{W}$ is a randomly initialized $N \times K$ matrix, where N is size of input vector and K is the number of output nodes of the network. $\boldsymbol{b}$ is a vector of length K.

The output vector $\boldsymbol{O}$ is computed by:

$$O = x^p \times W + b$$

Vector $\boldsymbol{a}$ is computed by passing $\boldsymbol{O}$ through a "winner take all" function which satisfies

$$a_k = \begin{cases} 1 & if\ \forall j \quad a_k \geq a_j \\ 0 & otherwise \end{cases}$$

In the event of a tie, a random number is chosen among those satisfying the above function, so that there is only one nonzero value in $\boldsymbol{a}$.

After each presentation of an input pattern, the following learning rule is used to update $\boldsymbol{W}$

$$\Delta W(i,j) = \eta(x^p(j) - W(i,j)) \quad if\ a(i) = 1$$

where η is a user defined constant called the "learning rate". η lies in the interval [0,1].

In our experiments, η was chosen to be 0.3.

4.4.5 Adaptive Resonance Theory Network (ART)

Adaptive resonance theory (ART) (Carpenter and Grossberg, 1987) is a neural computational algorithm designed for real time self-organizing stable pattern recognition codes in response

to arbitrary sequences of analog input patterns. The self-organizing ART networks were developed based on studies of brain mechanism.

There is a series of ART networks. ART1 is the simplest and is purely binary in operation, while ART2 processes analog input data. ART networks contain a number of adjustable parameters that allow the categorization to be adjusted over a range from fine to coarse and enhance the features in the input patterns. This can be a disadvantage since the parameter tuning procedure is complicated.

The ART2 network is shown in Figure 4.7. There are two fields of nodes *F1* and *F2* that are interconnected by top-down and bottom-up weighted links. The *F1* and *F2* fields, the bottom-up and top-down adaptive filters are contained within the attentional subsystem. ART2 adopts a competitive learning model in which a stream of input patterns to a network can adapt the Long-Term-Memory (LTM) values that multiply the signals in the pathways between *F1* and *F2*. Layer *F2* is a competitive network capable of choosing the node that receives the largest total input (this corresponds to the closest matching template stored in the LTM). Each *F2* node represents a specific category so a top-down signal indicates a particular category is activated. The competitive learning method also appears in other artificial neural network. In ART, the second top-down adaptive filter serves to self-stabilize the association learning process in response to an arbitrary input environment. There is also an auxiliary orienting subsystem that will be activated when the top-down expectation from *F2* cannot match with input pattern.

The ART2 algorithm proceeds as follows: in an ART2 network, an input pattern is applied to the *F1* layer. The criterion for an adequate match between an input pattern and a chosen template is adjustable by the vigilance parameter ρ. An auxiliary orienting subsystem becomes active when a bottom-up input to *F1* fails to match the learned top-down expectation read-out by the active template representation at *F2*. The orientating subsystem is activated and causes a reset of the active template. Alternative templates are tested in the hope of

finding an adequate match. When the right template is found, the learned top-down expectation resonates with the input pattern. The search time constant is short in comparison with the learning time so significant changes in LTM are possible only after the searching stops. The bottom-up forward and top-down feedback structure in the *F1* layer encodes the input in the amplitude independent resonant state via interaction with the *F2* layer. If there is no successful match, the cycle will end by selecting an uncommitted node of *F2*, then the bottom-up and top-down adaptive filters linked to this node learn the *F1* activation pattern generated directly by the input. If the full capacity of the system is used, the system will not be able to accommodate new input pattern.

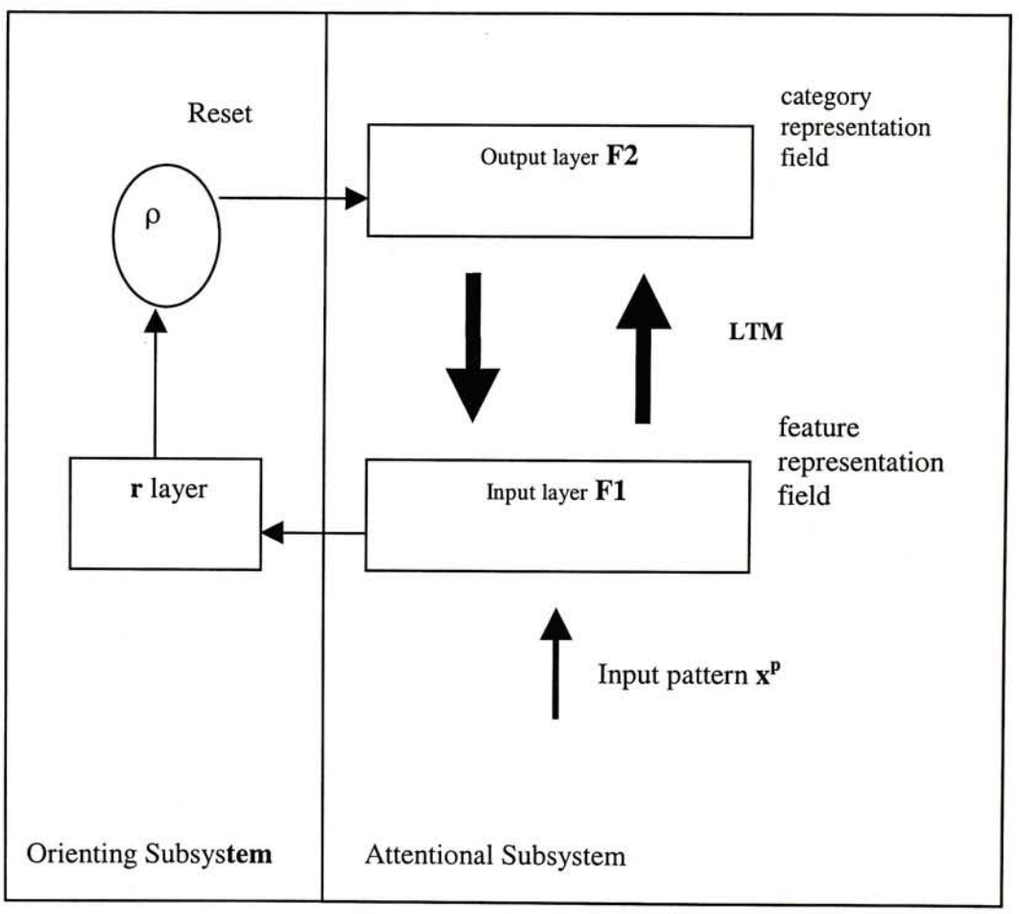

Figure 4-7: Block diagram of the ART2 network

4.4.6 ART2 Implementation

The complete ART2 algorithm is expressed mathematically below.

Parameter selection constraints:

The constraints applying to the constants used in the ART2 algorithms are defined below:

$$a, b > 0$$

$$0 \le d \le 1$$

$$cd/(1 - d) \le 1$$

$$0 \le \theta \le 1$$

$$0 \le \rho \le 1$$

In our experiments, the above constants are set as follow:

$a = 20$, $b = 20$, $c = 0.1$, $d = 0.9$, $\theta = 0.38$ and $\rho = 0.9$.

STEP 1 (Initialization):

All the top-down weights are initialized to zero:

$$z_{ij}^{0} = 0$$

Bottom-up weights are initialized according to:

$$z_{ji}^{0} = 0.5/(1-d)\sqrt{M}$$

All the following equations describe the steady-state values of a corresponding system of differential equations (Carpenter and Grossberg, 1987). The interconnection between sub-layers inside *F1* is shown in Figure 4.8. The reader should refer to Figure 4.8 while reading the description below.

Preprocessing on feature representation field *F1* (steps 2- 7)

STEP 2: At the lower layer of F1, vector $\boldsymbol{w}$ is the sum of an input vector $\boldsymbol{x}^p$ and the internal feedback signal vector $a\boldsymbol{u}$ where a is a constant, so that

$$\boldsymbol{w} = \boldsymbol{x}^p + a\boldsymbol{u}$$

STEP 3: $\boldsymbol{w}$ is normalized to yield

$$\boldsymbol{x} = N(\boldsymbol{w})$$

where the operator N,

$$N(\boldsymbol{w}) = \boldsymbol{w}/||\boldsymbol{w}||,$$

carries out an Euclidean normalization.

STEP 4: $\boldsymbol{x}$ is transformed to $\boldsymbol{v}$ via a nonlinear signal function defined by

$$\boldsymbol{v} = F(\boldsymbol{x})+bF(\boldsymbol{q}),$$

where $\boldsymbol{q}$ is defined in step 7 and function F is defined as:

$$F(\boldsymbol{x}_i) = \boldsymbol{x}_i \text{ if } \boldsymbol{x}_i > \theta, \text{ or 0 otherwise;}$$

STEP 5: $\boldsymbol{v}$ is normalized

$$\boldsymbol{u} = N(\boldsymbol{v})$$

Vector $\boldsymbol{u}$ is the output vector from *F1* to the orienting subsystem.

STEP 6: The top *F1* layer $\boldsymbol{p}$ sums both the internal *F1* signal $\boldsymbol{u}$ and all the *F2* to *F1* filtered signals. That is:

$$\boldsymbol{p} = \boldsymbol{u} + z_{ij} \times \boldsymbol{y};$$

where $\boldsymbol{y}$ is the output signal from the *F2* node computed in step 9 and z_{ij} is the LTM trace computed in steps 11 and 12.

STEP 7: $\boldsymbol{p}$ is normalized to give vector $\boldsymbol{q}$

$$\boldsymbol{q} = N(\boldsymbol{p})$$

Category Representation Field *F2:* (steps 8-9)

STEP 8: Output of the $\boldsymbol{p}$ sublayer is propagated to the *F2* layer. Net inputs $\boldsymbol{T}$ of *F2* is determined by

$$\boldsymbol{T} = z_{ji} \times \boldsymbol{p}$$

STEP 9: $\boldsymbol{T}$ undergoes the winner-take-all selection and produces the *F2* output vector $\boldsymbol{y}$.

$$\boldsymbol{y} = g(\boldsymbol{T});$$

where $\boldsymbol{g(T)}$ is function to select and amplify the maximum element in $\boldsymbol{T}$ and is defined as:

$$g(T_j) = dT_j = max\{T_k\}; \text{ or 0 otherwise.}$$

Any nodes marked as ineligible by previous reset signal from orienting subsystem do not participate in the competition.

Orienting Subsystem: (steps 10-11)

STEP 10: Vector ***r*** supervises the degree of match between the *F1* bottom-up input and *F2* top-down input.

Output of the ***r*** layer is:

$$r_i = \frac{u_i + cp_i}{\|u\| + c\|p\|}$$

where ***u*** and ***p*** are sub-layers in *F1* layer and c is constant.

STEP 11: System reset if

$$||\boldsymbol{r}|| < \rho;$$

where ρ is the vigilance parameter. If the system resets, the current active *F2* node will be marked ineligible for future competition and then return to step 1. In this way, the system reset will make ART2 activate the node with the next highest input in next iteration.

STEP 12: Adjust bottom-up weights on the winning *F2* unit by:

$$z_{Ji} = u_i/(1-d)$$

STEP 13: Adjust top-down weights coming from the winning *F2* unit by:

$$z_{iJ} = u_i/(1-d)$$

STEP 14: Restore all inactive *F2* units and go back to step 2 with new input pattern $\boldsymbol{x}^{p+1}$.

4.6 Summary

In this chapter, an introduction to the relevant aspects of principal component analysis, traditional clustering methods and clustering by unsupervised neural learning was presented. An overview of implementation details of PCA and the four chosen clustering techniques (namely online template matching, K-means, simple competitive learning and ART2) was also presented.

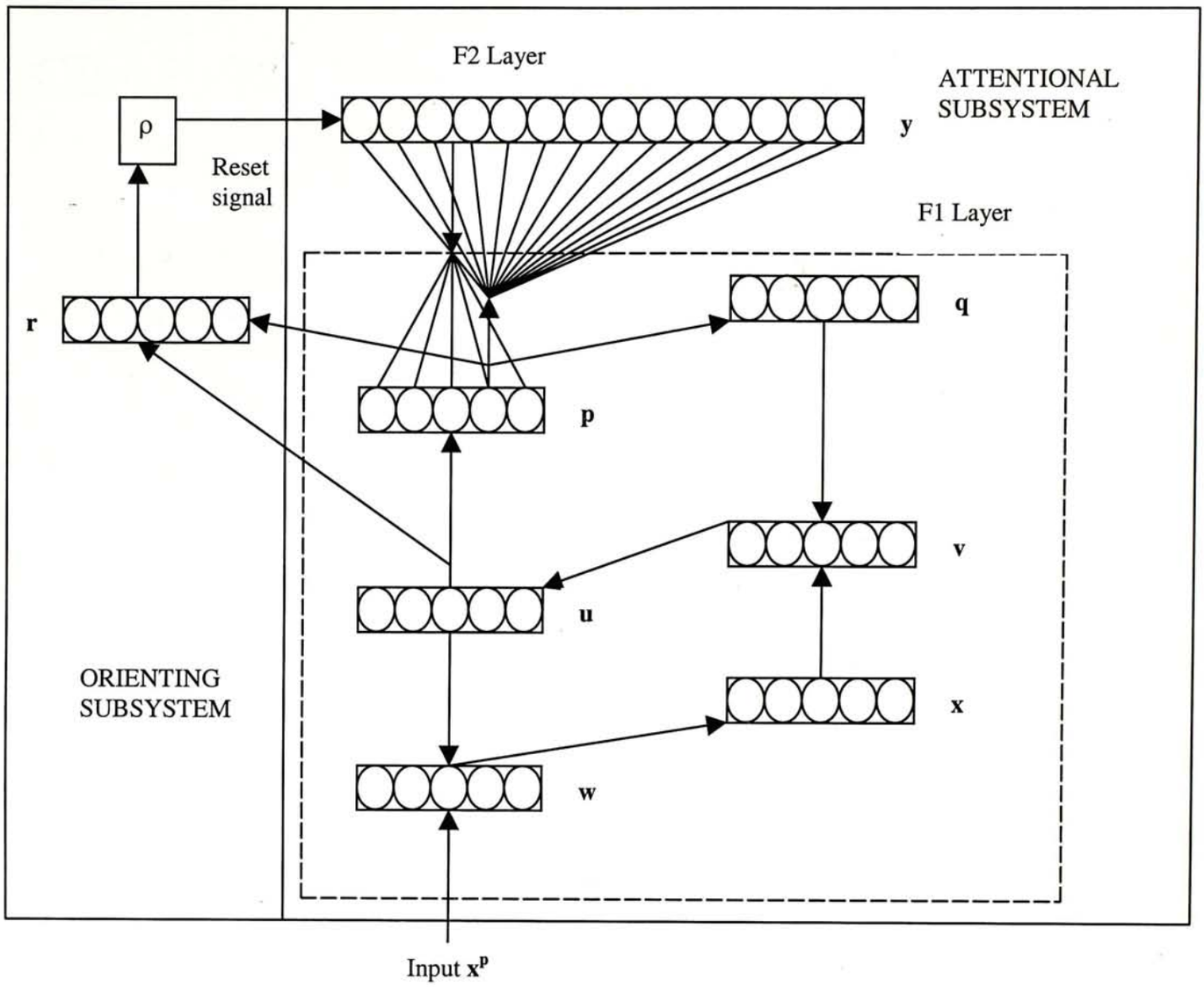

Figure 4-8: Layers, sublayers and interconnections inside ART2

CHAPTER 5

Experimental Design

5.1 Introduction

The objective of this chapter is to describe the experimental setups for the two problems introduced in chapters 2 and 3. The first problem is to demultiplex the electronically recorded neural spike signals into separate clusters where each cluster represents signals coming from a particular neuron. The second part of this chapter presents an approach to test mass produced headphones in a hope to find the crucial parameters affecting sound quality of a large batch of headphones in production line.

5.2 Electrophysiological Spike Discrimination

5.2.1 Experimental Design

The experimental design for the automatic online clustering of extracellular multi-neuron recordings from the nervous system is described in this section. The spike train are first processed by a Schmitt trigger threshold detector and optionally, principal component analysis (PCA). A clustering was applied to reconstruct separate spike trains from 21 different multi-neuron recordings. Quantitative measurements of the efficiency and utility of the algorithms were performed. Using these measurements, the performances of the four algorithms with and without PCA preprocessing were compared in terms of accuracy and efficiency.

5.2.2 Extracellular Recordings

The analysis of the spike waveform discrimination algorithms was performed on real physiological data obtained from the lateral geniculate nucleus[1] of the common marmoset, *Callithrix jacchus*, a small South American monkey. *Callithrix jacchus* has a visual system similar to that of humans and other primates. The signals from the extracellular electrodes were amplified by a factor of 50000. After passing through an anti-aliasing filter, they are digitized with a 12 bit National Instruments NI-MIO-16H data acquisition card at 22.5 kHz and the resultant data were saved to disk. All experiments described were performed off-line from recorded data. All algorithms were implemented using the Mathworks Inc. MATLAB version 5.0, running on an Ultra-Sparc 5 workstation. MATLAB provides a fast prototyping language for expressing computations, good data plotting facilities and an interactive environment for testing different algorithms.

Figure 5.1 illustrates the software implemented Schmitt trigger used to identify spikes. Two user-defined threshold values are used namely the spike recording trigger threshold $\delta 1$ and spike recording enable threshold $\delta 2$. A spike is triggered by an excursion of the input signal above $\delta 1$. The first 6 data points taken before the triggering point and 19 points at and after the trigger point are collected to form the spike vector. Before a new trigger can occur, the signal must pass below the second threshold $\delta 2$, implementing a hysteresis effect to reduce false triggering in the presence of noise.

[1] Part of the visual pathway between the retina and the visual cortex.

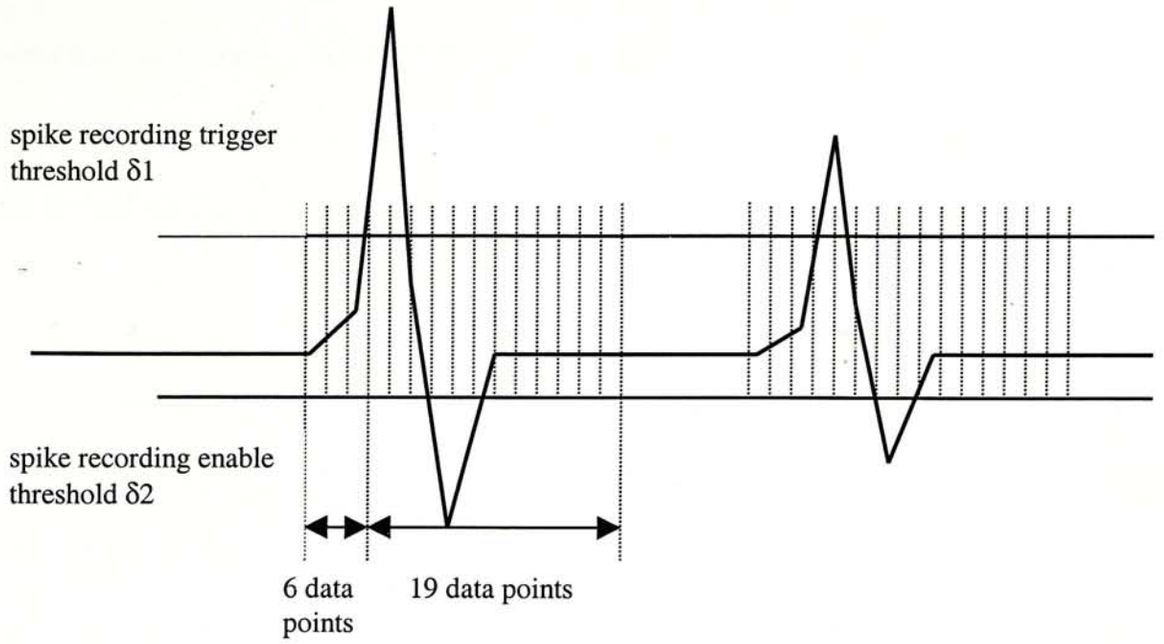

Figure 5-1: Schmitt trigger to record spikes

5.2.3 PCA Feature Extraction

All of the experiments were performed with and without a PCA preprocessing step and the effect of PCA on the computational requirements and accuracy of the results were also measured. PCA preprocessing is applied optionally to reduce the dimensionality of spike data from 25 to 2.

Refer to chapter 4 for detailed description of how PCA is implemented. The PCA preprocessing was performed on the entire data set if there are less than 1000 data. Otherwise, the first 1000 spikes were selected to compute the PCA parameters.

5.2.4 Clustering Analysis

Four clustering algorithms (template matching (TM), K-means, simple competitive learning (SCL) and ART2) were compared for accuracy and computational requirements. K-means is the standard clustering technique for multineuron spike sorting (Schmidt, 1984b; Salganicoff et. al., 1988; Sarna et. al., 1988) and used as a benchmark for comparison. Template matching is well known for its simplistic algorithmic design and computational efficiency. SCL and

ART2 were chosen as being representative of the relatively new nonlinear neural computation models that are becoming increasingly prevalent in contemporary signal processing systems. The correct number of cluster for waveform set was determined by a systematic way using the indication of the F_{ratio}.

$$F_{ratio} = \frac{E_{K+1clusters}}{E_{KClusters}}$$

This is a ratio of total partition error of K+1 clusters over that of K clusters. If the ratio is above10, then K+1 or more clusters should be necessary and the calculation continues. For F-ratio below 10, the K-clusters are adequate for waveform set.

In order to tune the clustering routines to give the desired number of cluster, the parameters of all clustering techniques (threshold value D for TM, learning rate parameter η for SCL and vigilance parameter ρ for ART2) were selected using a bisection search. By this bisection search method, experiments with labeled and noise-corrupted data were performed with parameter values adjusted linearly until the best results were obtained. There were no tunable parameters for K-means except the number of clusters K so no parameter tuning was necessary.

All the input patterns are normalized to ease the task of tuning the parameter values. Training was performed in an online manner for all clustering methods except K-means. In the case of K-means, the training step was performed over the entire data set. K-means is normally an offline clustering method. K-means is normally applied to spike sorting (Salganicoff et al., 1988) by clustering an initial representative data set to compute cluster means which are then frozen in value, real-time clustering being performed by assigning the spike to the nearest cluster mean.

5.3 Correlation of Headphone Sound Quality with physical Parameters

5.3.1 Experimental Design

This section presents an approach to test mass produced headphones in a hope to find the crucial parameters affecting sound quality of a large batch of headphones in production line. This proposed approach can be summarized as follow.

1. In order to reduce the number of headphones to undergo subjective listening tests, a large number of headphones are clustered based on frequency response into a small number of clusters. A few samples of headphones are selected from each cluster to form a reduced headphone test set.
2. Listening tests are conducted with each headphone being tested several times and compared against several different headphones by different people.
3. In order to simplify the task of the listener, each human sample is asked to grade 4 randomly selected headphones by making three simple binary comparisons.
4. All the measurable parameters of the headphones tested are measured to ensure comprehensive analysis is performed.
5. Correlation coefficient between measured parameters and the listening test scores is computed to provide a single figure measure of the ability of parameters to predict the listening test result.
6. The preliminary correlation result is applied to select new headphone samples for further confirmation tests.

This approach can be used to identify those parameters directly affecting the sound quality of the headphones (if any exist) in a production line. This makes it possible to automate the quality control process and allows design and manufacturing changes, which can lead to an overall improvement in the sound quality of the product.

5.3.2 Frequency Response Clustering

52 mass produced headphones of the type typically supplied with portable tape recorders and CD players were studied. In order to minimize the manufacturing variations, all of the headphones were of the same model produced on the same production line on the same day. The frequency response for each headphone was measured using Audiomatica Srl's CLIO system (Audiomatica, 1998). A plot of the frequency response of all headphones is shown in Figure 5-2.

It can be seen from the frequency response measurements of Figure 5-2 that the low frequency gain varies dramatically between headphones. We believe this was due to the mechanical device used to hold the headphones and the low frequency measurements should not be regarded as being accurate.

A clustering technique was applied to group headphones with similar frequency response into clusters. The procedure proceeds as follows:

1. Principal component analysis (PCA) was applied to the magnitude of the frequency response to extract the salient features from the data and achieve dimensionality reduction (Jolliffe, 1989) by choosing the first two principal components from the magnitude frequency response of each headphone.
2. This reduced dimensionality data set was then clustered using the K-means algorithm (Hartigan, 1975) to group them into clusters.

The above clustering technique was applied to cluster the frequency responses of the 52 different headphones into 4 clusters.

The result of this clustering process is shown in Figure 5-3 and Figure 5-4. Figure 5-3 shows the clustered plot of the first two PCA components of the clustered data set and Figure 5-4 showed the clustered frequency response plot of headphones. The frequency distribution of the clustering results is shown in Table 5.1.

Cluster	*Frequency*
1	23
2	19
3	8
4	2

Table 5.1: Frequency distribution of headphone clusters

If frequency response were a good measure of headphone quality, one might expect that a single cluster would contain a disproportionally high number of similar sounding headphones. Since listening tests based on 52 different headphones would be too large an undertaking, a subset of these headphones based on the clustering results was selected.

Random samples were taken from each cluster to represent the cluster with the number taken dependent on the number of headphones in the cluster. Since cluster 1 had the most headphones, three samples were taken from it. Two headphones were taken from clusters 2 and 3 and one sample from cluster 4. This subset of 8 headphones was chosen as a representative set of the different characteristics of the original 52 headphones to be used in the listening tests. Figure 5-5 shows the frequency response plot of the 8 selected headphones. A sample size of 8 headphones was selected since it made the number of listening tests small enough to be feasibly conducted in our laboratory using a simple randomly drawn listening test. A much larger sample size would increase our confidence in the results but would require an enormous number of listening tests. It is also possible to perhaps use a smaller number of listeners and ask them to assign a numerical score to each headphone. This procedure would enable a much larger number of headphones to be tested, however, listener fatigue and personal preference may lead to unreliable results. It must be remembered that the goal of the headphone industry is to produce headphones that sound good to the majority of listeners rather than satisfy a small number of critical headphone reviewers.

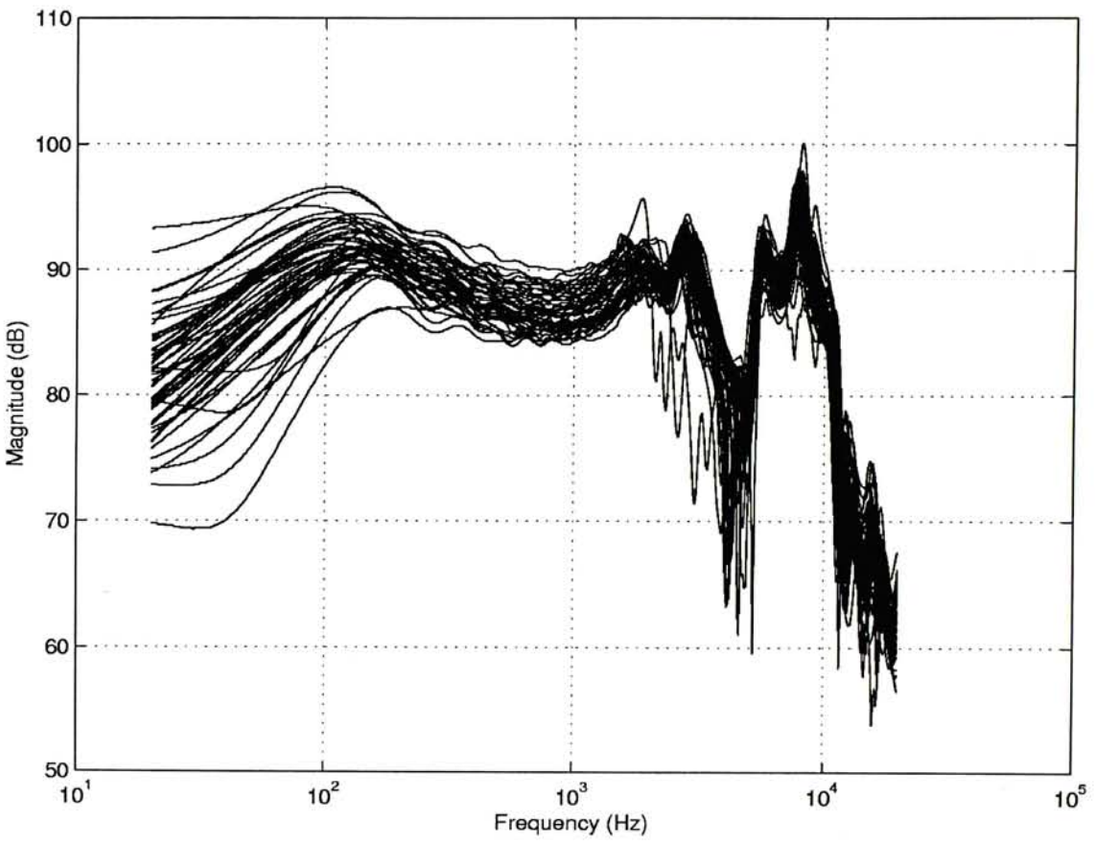

Figure 5-2: Frequency response of all headphones

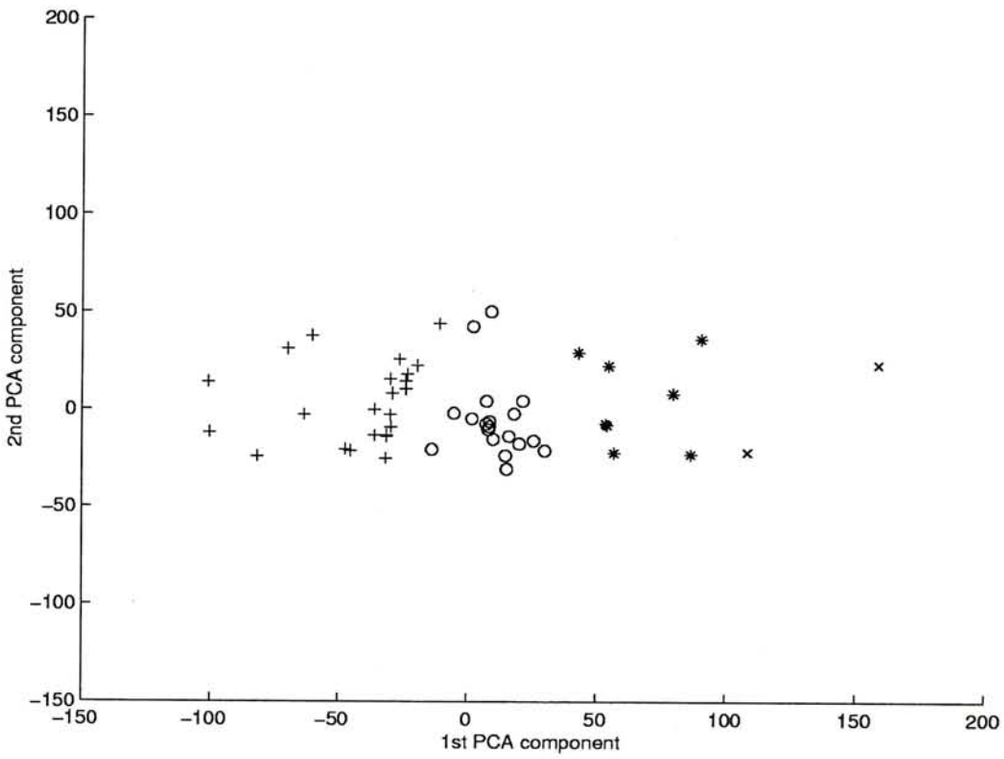

Figure 5-3: PCA components plot of spike data

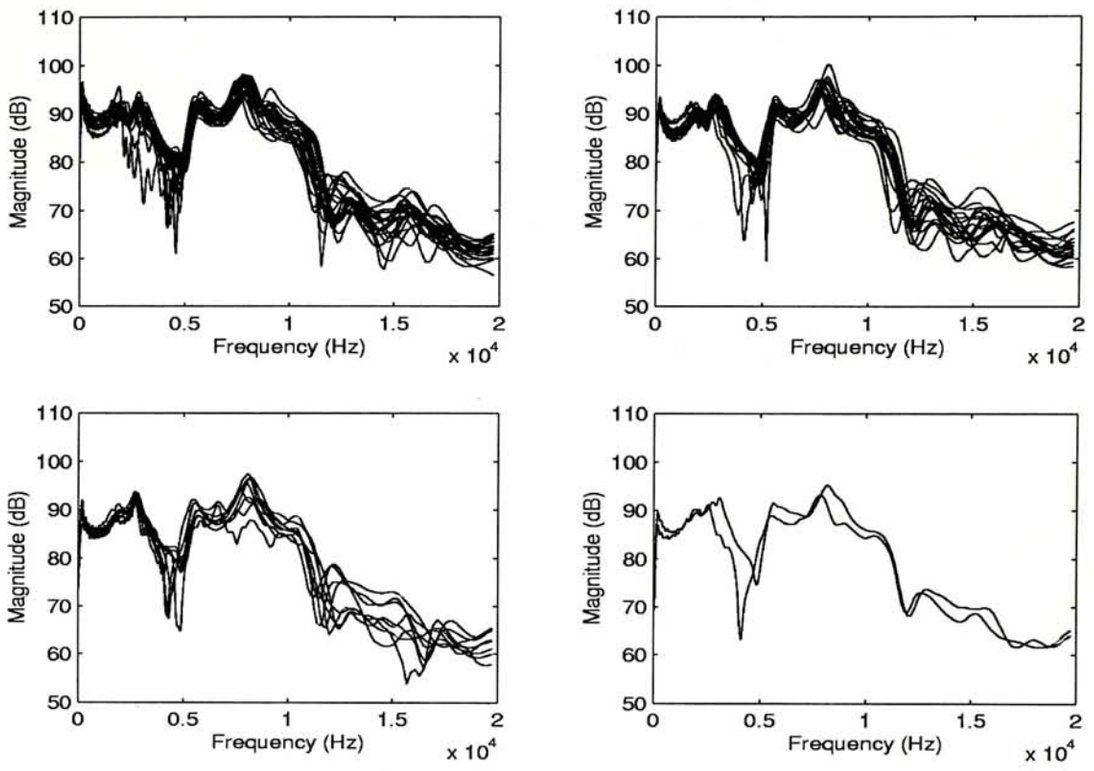

Figure 5-4: Frequency response plot of clustered headphones

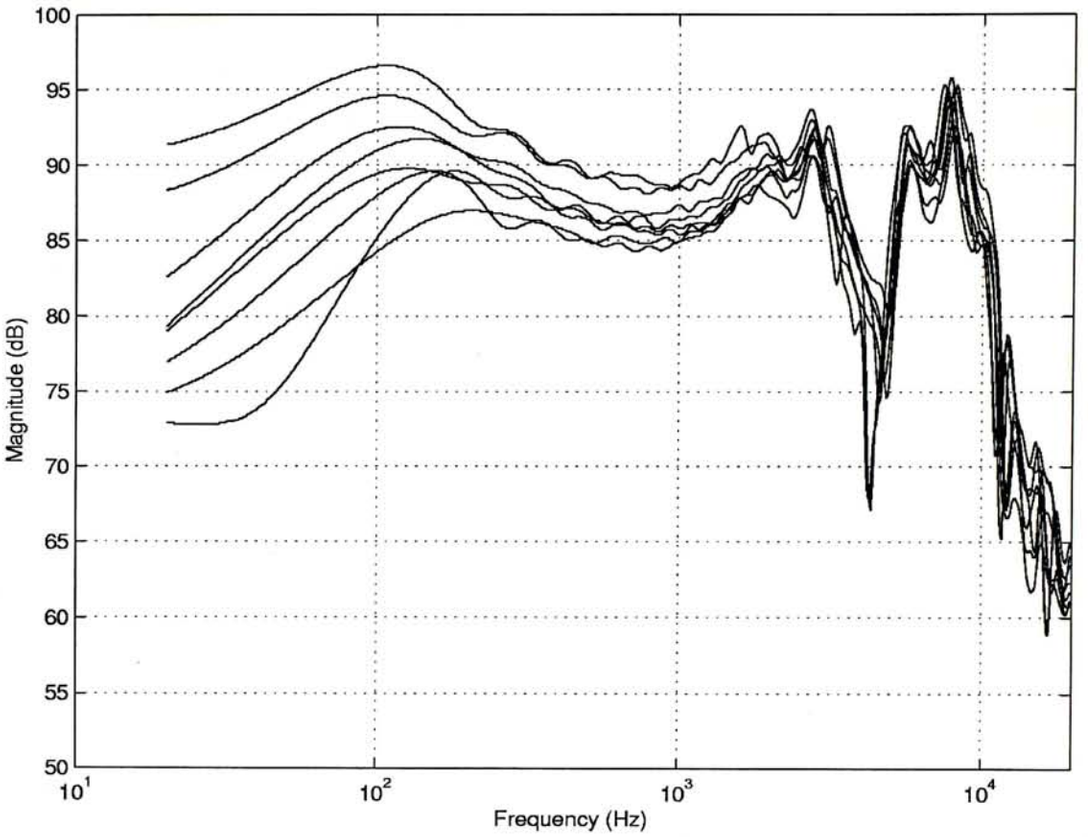

Figure5-5: Frequency response of the selected headphones

5.3.3 Additional Parameters Measurement

Additional parameters of the 8 selected headphones were also measured using Audiomatica Srl's CLIO system. The first set of parameters, A_{46Hz}, A_{1kHz}, $A_{11547Hz}$ were taken from the frequency response and correspond to the nominal low, midrange and high frequency responses of the headphones. The second set of measurements were D_2 and D_3 which correspond to the distortion levels of the headphones to a 1kHz sinusoidal input signal. Finally, the voice-coil driver parameters (R_E, F_s, Q_{es} , Q_{ms} , Q_{ts} , L_{1k} and L_{10k}) were measured which relate directly to the voice coil. The symbols of each parameter are listed below with brief note of their physical meanings:

Frequency Response Parameters:

A_{46Hz}	Gain at 46 Hz
A_{1kHz}	Gain at 1 kHz
$A_{11547Hz}$	Gain at 11547 Hz

Harmonic Distortion Parameters:

D_2	Second harmonic distortion (1~kHz input signal)
D_3	Third harmonic distortion (1~kHz input signal)

Voice-coil Driver Parameters:

R_E	DC resistance of the voice coil
F_S	Resonant frequency of driver
Q_{es}	Total Q of driver at F_s considering only electrical resistance
Q_{ms}	Total Q of driver at F_s considering only non-electrical resistance
Q_{ts}	Total Q of driver at F_s considering all system resistance
L_{1k}	Inductance of voice coil at 1~kHz
L_{10k}	Inductance of voice coil at 10~kHz

5.3.4 Listening Tests

The system used to evaluate the headphones was typical of the type of system with which such headphones would be used. It comprised of a Sony D-465 Discman CD player connected

to a distribution amplifier. The amplifier, constructed from Burr-Brown OPA604 opamps with a gain of 2, was used so that the 4 headphones could be evaluated simultaneously without disconnecting headphones. The song "Desafinado" from the CD "Jazz Samba" (Verve records 810 061-2) was used for all of the evaluation tests. This piece was selected for its transient nature, being a mix of drums, tenor sax, bass and guitar.

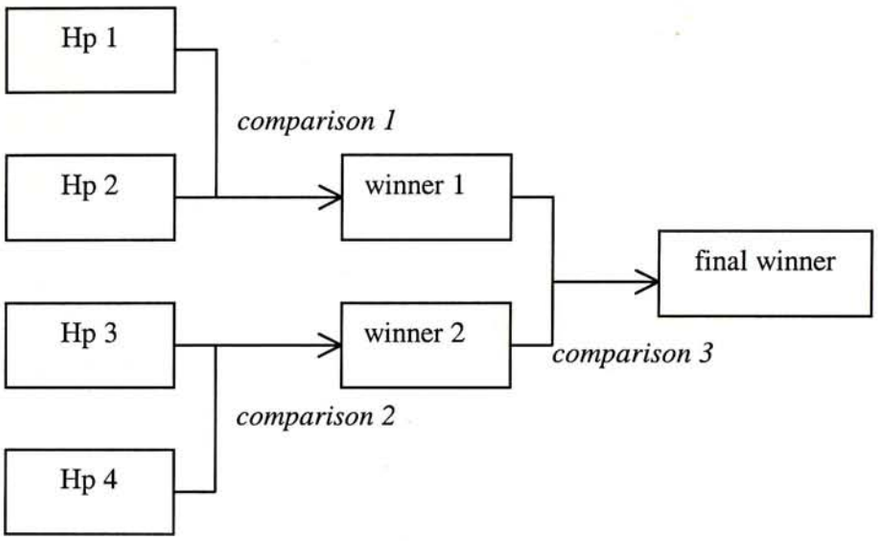

Figure 5-6: Listening test based on 3 binary comparisons

Using the clustered set of 8 headphones, a simple informal listening test (see Figure 5-6) was conducted in the following manner: a computer program was used to randomly draw four headphones in a random order. Each subject was asked to choose their preference between the first and second headphones and then the third and fourth headphones on the list. Then the subject was asked to select between the two winning headphones which determined the headphone they most preferred. In this manner, a winning headphone was selected from 4 different headphones based on three simple binary comparisons.

5.3.5 Confirmation Test

The correlation coefficients between test scores of headphones and measured parameters are found using the *t* test, which is a standard statistical test for correlation measurement. Those parameters with the highest correlation are used as criteria for further selection of headphones in confirmation test. A new test batch is formed by randomly choosing 4 from the original test batch and selecting 4 headphones from production line that pass the selection criteria

concluded from last listening test. The same listening test is performed on this new test set. This selection process itself is another form of binary clustering procedure since the headphones are divided into groups, one group selected according to their parameters similarity with headphones scoring high marks in previous listening tests and one group randomly selected with prior requirements to their parameters values. The correlation measurement is done again to confirm the validity of the previously concluded criteria for headphone selection.

5.4 Summary

In this chapter, the experimental setup for the two signal processing problems (namely, electrophysiological spike discrimination and headphone sound quality measurement) was described in detail. Clustering analysis methods had been applied to help solving the problems. The mixed spike signals were separated into separate spike-trains using four different clustering methods. Clustering was applied to help selecting the representative headphone sample for listening tests.

CHAPTER 6

Results

6.1 Introduction

In this chapter, the experimental results from both problems: (electrophysiological spike discrimination and headphone sound quality measurement) are presented. For spike discrimination, a comparison of four clustering techniques is given based on their robustness to noise and ability to achieve good convergence when applied to real spike data. The proposed experimental procedures to control headphone sound quality with high efficiency are also verified.

6.2 Electrophysiological Spike Discrimination: A Comparison of Methods

Quantitative measurements of the four clustering techniques (TM, K-means, SCL and ART2) for the automatic online clustering of extracellular multi-neuron recordings from the nervous system were made in terms of efficiency and accuracy. Rigorous comparison was performed based on their robustness to gaussian noise and convergence ability on real spike data.

The main theme of this thesis was to provide a quantitative analysis of various clustering techniques to facilitate future spike discrimination experiments in physiological laboratories. Readers should always bear in mind that real problem of electrophysilogical spike discrimination demands a real time classification process. Supervised learning methods like MLP with backpropagation learning (Jansen R.F., 1990) which require offline human

intervention are not an adequate method to fit in real spike discrimination situation so it is not chosen for comparison in this thesis.

6.2.1 Clustering Labeled Spike Data

Data set preparation

The first set of data was a set of labeled spike data. It was aimed to test the accuracy of clustering algorithms and was generated by adding different levels of white noise to human-clustered results. Four clearly differently shaped spikes were taken from a clustering of recorded electrophysiological data. Different amounts of Gaussian noise were added to each waveform to form a test set. For each of the 4 spike shapes, 30 corrupted versions were generated to form a test set of 120 spikes. This test set was generated at noise levels corresponding to 1%, 20%, 40%, 60% and 80% of the maximum amplitude (see Figure 6-1) to test robustness of the four clustering algorithms. The originality of each spike in this data set was well defined and this made a quantitative comparison of various clustering algorithms feasible.

Robustness to noise

The performances of each algorithm for the labeled test set described (with and without PCA preprocessing) were shown in Table 6-1. From Table 6-1, it could be seen that for a noise level less than 20%, all algorithms were able to cluster the data with no errors. This showed an overall robustness to noise, mainly due to the spikes being very different in shape. It also demonstrated the correctness of our clustering algorithm implementations. At 40% noise and above, K-means and SCL algorithms remained accurate whereas ART2 showed a somewhat degraded performance. Worst of all in performance was the TM algorithm without PCA preprocessing. TM without PCA collapsed at 40% noise level and its problem was mainly due to its simplistic method of creating new templates. In every case, PCA preprocessing has improved the accuracy of the clustering algorithms.

Efficiency comparison of clustering algorithms

The number of floating point operations (FLOPS) and the actual runtime (as measured using the built-in measurement facility in MATLAB) required to execute the algorithm running on an isolated Sun Ultra-Sparc workstation (CPU speed: 140MHz; operating system: Solaris 2.5.1; 64MB memory) were recorded for the 1% noise level case and the result were shown in Table 6-2. PCA computation was performed only once per data set and did not need to be computed for subsequent inputs. All 120 spikes were used to compute the PCA parameters for clustering. Thus comparison of the algorithms with PCA preprocessing was best made on the "Relative FLOPS / Runtime(sec) (CLUSTERING)" column in Table 6-2.

The three columns correspond to the number of floating point operations required and runtime to perform PCA, clustering and PCA + clustering. Figure 6-2 compared the efficiency of PCA + clustering in term of FLOPS consumed and Figure 6-3 compared efficiency of PCA + clustering in term of actual runtime. The comparison results based on FLOPS and runtime were similar because the targeted algorithms were all highly computational with critical routines performing mostly floating point operations. FLOPS was suggested as a better measure of computational efficiency than actual runtime measurement since it was machine independent and could be used as criteria for selecting computer platforms feasible for real time spikes classification. The actual runtime were included to give reader a concrete idea of how fast those clustering algorithms could perform in nowadays high-speed computer platforms.

The algorithms in order of computational efficiency in terms of FLOPS (most efficient first) were SCL, TM, SCL + PCA, TM + PCA, K-means + PCA, ART2 + PCA, K-means and ART2. The decreasing order of efficiency of cases using PCA was the same as the non-PCA case. In the case of no PCA preprocessing, SCL and TM had approximately the same computational requirements whereas K-means and ART2 requires 6 and 12 times more FLOPS and runtime than SCL respectively. PCA preprocessing reduced dimensionality of data set from 25 to 2, thus reducing the number of FLOPS required for clustering by factors of 3, 4, 3 and 7 over the non-PCA case for TM, K-means, SCL and ART2 respectively.

For the case of running SCL and TM with PCA on selected machine platform, the classification rate is 17 spikes/sec. The computing facility is capable of handling real-time spike classification as the normal spike firing frequency is below 10 Hz. For both K-means and ART2, the spike classification rate is about 4 spikes/sec so the algorithms may fail for cases with spikes firing rate higher than 4 Hz.

	Number of Errors (1% noise)	Number of Errors (20% noise)	Number of Errors (40% noise)	Number of Errors (60% noise)	Number of Errors (80% noise)
TM	0	0	90	90	90
TM + PCA	0	0	8	10	17
K-means	0	0	1	7	12
K-means + PCA	0	0	8	7	10
SCL	0	0	10	17	21
SCL + PCA	0	0	5	14	19
ART2	0	1	21	29	31
ART2 + PCA	0	0	9	24	26

Table 6-1: Results for selected data corrupted with noise. The entries correspond to the number of incorrectly classified data in a 120 point test set.

Algorithm	Relative FLOPS / Runtime (sec) (PCA)	Relative FLOPS / Runtime (sec) (CLUSTERING)	Relative FLOPS / Runtime (sec) (TOTAL)
TM	0/0	3.6/20	3.6/20
TM + PCA	8.5/60	1.2/7	9.8/67
K-means	0/0	22/180	22/180
K-means + PCA	8.5/60	4.9/41	13/101
SCL	0/0	3.4/24	3.4/24
SCL + PCA	8.5/60	1 {FLOPS=13,168} /7	9.5/67
ART2	0/0	41/211	41/211
ART2 + PCA	8.5/60	6/34	15/94

Table 6-2: Comparison of the efficiency of the algorithms for the 1% noise level case of Table 6-1. All FLOPS results are expressed relative to the most efficient algorithm, SCL+PCA, for which the absolute number of flops is given (i.e. RELATIVE FLOPS(X) = FLOPS(X)/FLOPS(SCL+PCA)). The results in the first column show the relative number of FLOPS required to perform PCA, the second column are the relative number of flops required to perform the clustering and the final column is the total relative FLOPS for PCA plus clustering.

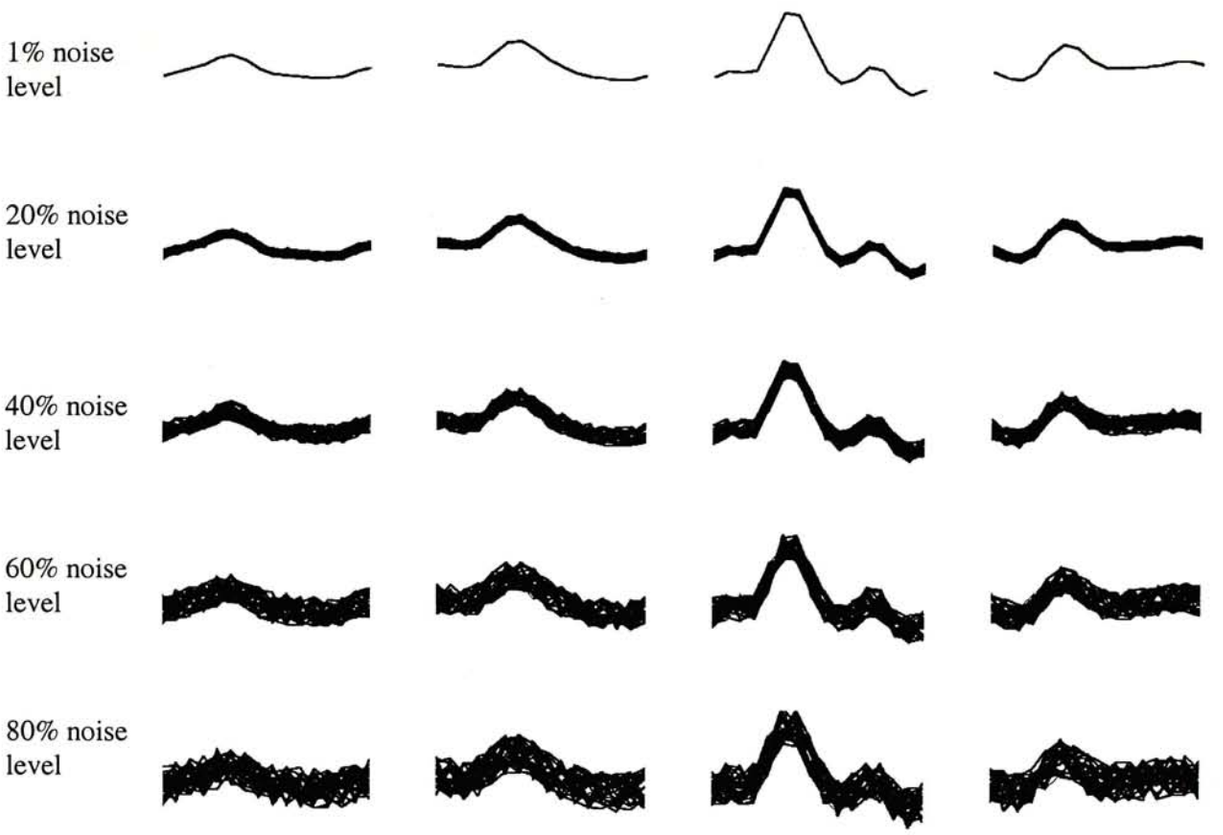

Figure 6-1: Selected data sets corrupted with 1%, 20%, 40%, 60% and 80% noise.

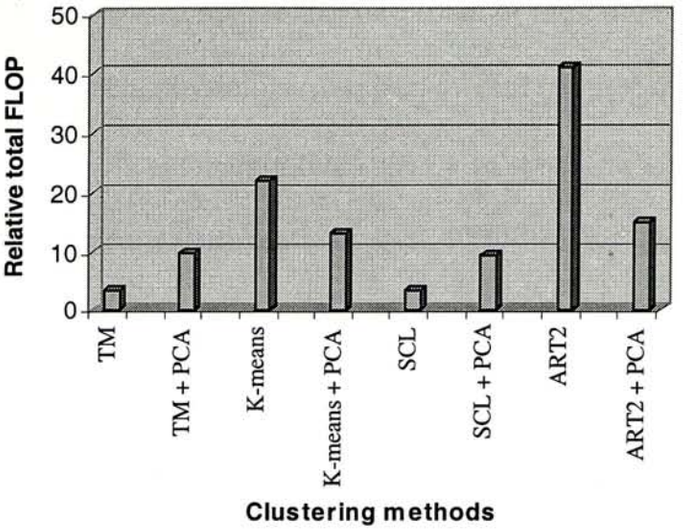

Figure 6-2: Comparison of efficiency of clustering algorithm on artificial generated data in term of relative flops

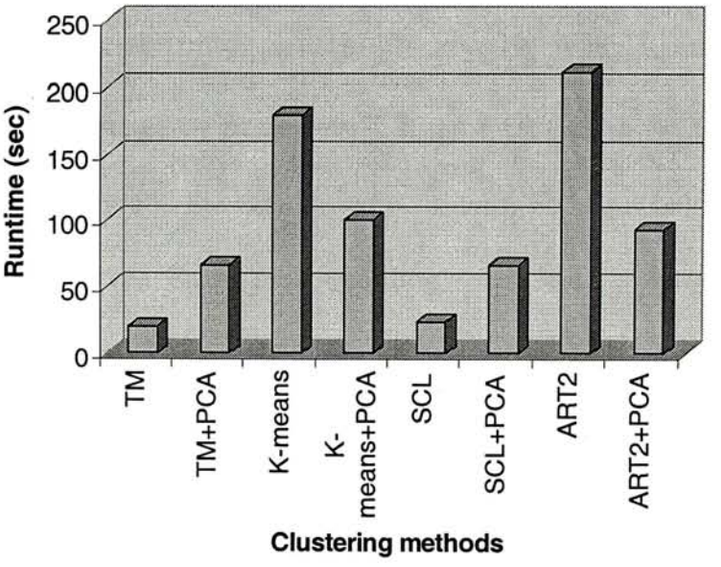

Figure 6-3: Comparison of efficiency of clustering algorithm on artificial generated data in term of runtime measurement

6.2.2 Clustering of Unlabeled Data

The second data set consisting of unlabeled spikes taken from various multi-neuron recordings were used to make quantitative measurements of the efficiency and utility of the algorithms. For detail of electrophysiological spikes recordings preparation, please refer to section 5.2.2. The four classification algorithms were also tested on a set of real recordings taken from electrophysiological experiments. In total, 25150 spikes were classified from 21 different recordings.

Number of cluster definition and parameter tuning for individual recordings

The correct number of cluster for each spike recording was determined by a systematic way using the indication of the F_{ratio} accompanied with K-means clustering.

$$F_{ratio} = \frac{E_{K+1clusters}}{E_{KClusters}}$$

This is a ratio of total partition error of K+1 clusters over that of K clusters using the K-means clustering method. There were no tunable parameters for K-means except the number of clusters K so no parameter tuning was necessary. By using K-means, the number of cluster can be easily determined by parameter K. If the F ratio is above 10, then K+1 or more clusters should be necessary and the calculation continues. For F_{ratio} below 10, the K-clusters are adequate for waveform set. Each spike recording were clustered into a desired number of cluster using this F_{ratio} value. The numbers of cluster formed for each recording were shown in the last column of TableA-1.

All the input patterns were normalized to ease the task of tuning the parameter values. In order to tune the clustering routines to give the predetermined number of cluster, the parameters of all clustering techniques (threshold value D for TM, learning rate parameter η for SCL and vigilance parameter ρ for ART2) were selected using a bisection search. By this bisection search method, experiments with labeled and noise-corrupted data were performed with parameter values adjusted linearly until the data set were divided into the predetermined number of clusters.

Quantitative comparison of clustering techniques

The *TDIST* value for each algorithm is the total Euclidean distance between the template for each cluster (defined automatically by the clustering algorithms) and each spike in the data set. The *TDIST* value is a measure of its convergence properties, a smaller value being more desirable.

Figure 6-4 compared the average convergence result (*TDIST*) of clustering methods over the 21 recordings with and without PCA preprocessing. Figure 6-5 compares the average efficiency (*FLOPS*) of clustering methods with and without PCA. Figure 6-6 compares the average efficiency (*Runtime*) of clustering methods with and without PCA. The runtime were taken from computer platform Sun Sparc Ultra (CPU speed: 140MHz; Memory: 64MB; Operating system: Solaris 2.5.1). Figures 6-7 and 6-8 showed the result of the dividing the data set T21 into 4 clusters with and without PCA respectively. (Table A-1 shows the results of the experiments without PCA preprocessing and Table A-2 shows the results with PCA preprocessing.)

Referring to Figure 6-4, the order of clustering methods in increasing *TDIST* value (decreasing convergence) is K-means, SCL, TM and ART2 for both cases with and without PCA preprocessing. The experimental result using PCA preprocessing showed a high degree of similarity with those without using PCA. PCA preprocessing had proved its significance to improve the convergence ability of all the four clustering algorithms.

Again, it can be shown from Figure 6-5 and Figure 6-6 that the comparison of computational efficiency using actual runtime measurement and FLOPS is very similar. It is clear from Figure 6-5 and Figure 6-6 that the TM and SCL algorithms are the most efficient, followed by K-means and then ART2 for both cases with and without using PCA. Using PCA analysis prior to clustering reduced the amount of computation required for performing clustering alone since the reduction in data meant less computation was needed. TM, K-means, SCL and ART2 enjoyed a 7, 17, 4 and 11 times speedup respectively (see Table A-2 for detail).

K-means algorithm produced the smallest value of *TDIST* for both case of with and without PCA preprocessing. This reflects the fact that K-means could optimize its distance measure using all of the input patterns whereas the other, online algorithms could only update based on the current pattern (to put it another way, K-means was applied in an off-line fashion whereas the others were online algorithms). A direct result of this global optimization was that the K-means without PCA was the slowest among all of the other algorithms (see Figure 6-5 and Figure 6-6). The computational loading of K-means without PCA was disproportionally greater than other algorithms (about 63 time of SCL without PCA). PCA preprocessing highly improved the efficiency of K-means, but K-means + PCA was still the second slowest clustering methods among all.

The SCL algorithm managed to achieve the next best *TDIST* measure, an average 2% higher than K-means (refer to Tables A-1 and A-2 in Appendix A.1) while SCL without PCA was the fastest algorithms. The TM algorithm requires 30% more FLOPS than SCL but achieving a poorer *TDIST* measure than SCL. The ART2 algorithm achieved a *TDIST* measure similar to TM but requires about 5 times more FLOPS than SCL.

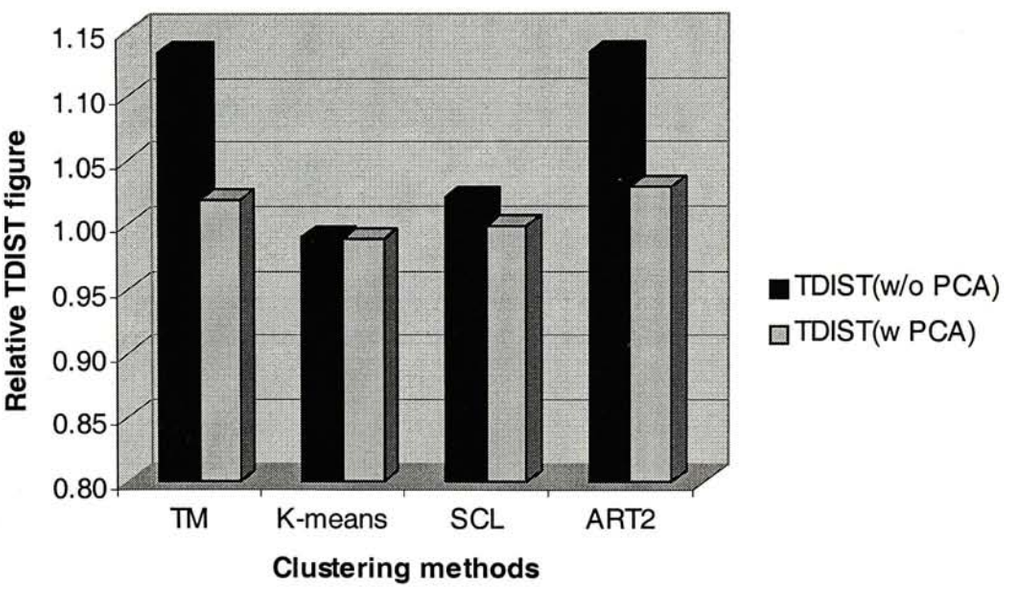

Figure 6-4: Comparison of performance measure (*TDIST*) of clustering algorithm with and without PCA preprocessing on real spike data

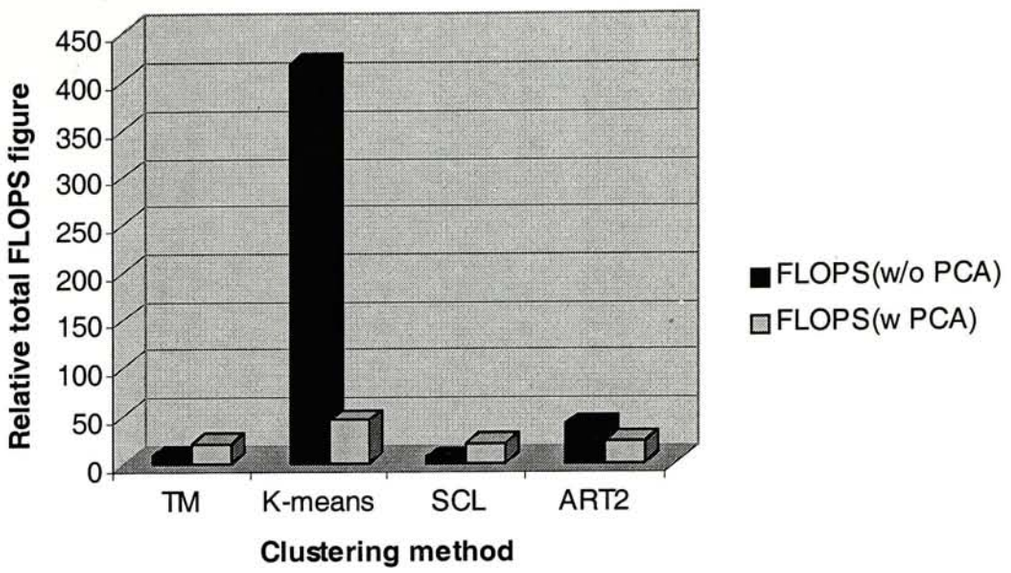

Figure 6-5: Comparison of efficiency (*FLOPS*) of clustering algorithm with and without PCA preprocessing on real spike data (operation of PCA included in *FLOPS* if needed)

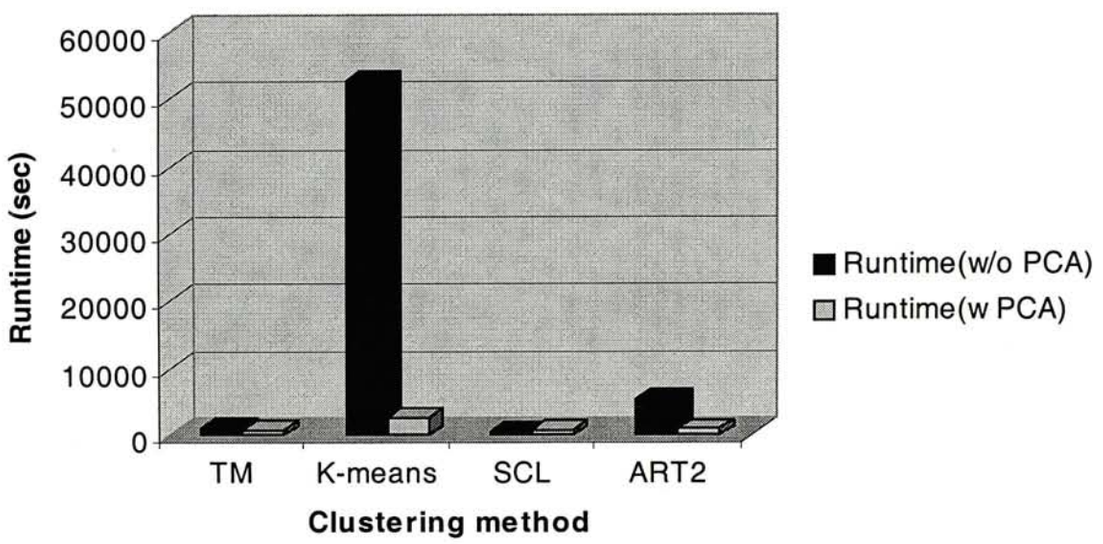

Figure 6-6: Comparison of efficiency (*Runtime*) of clustering algorithm with and without PCA preprocessing on real spike data (operation of PCA included in *FLOPS* if needed)

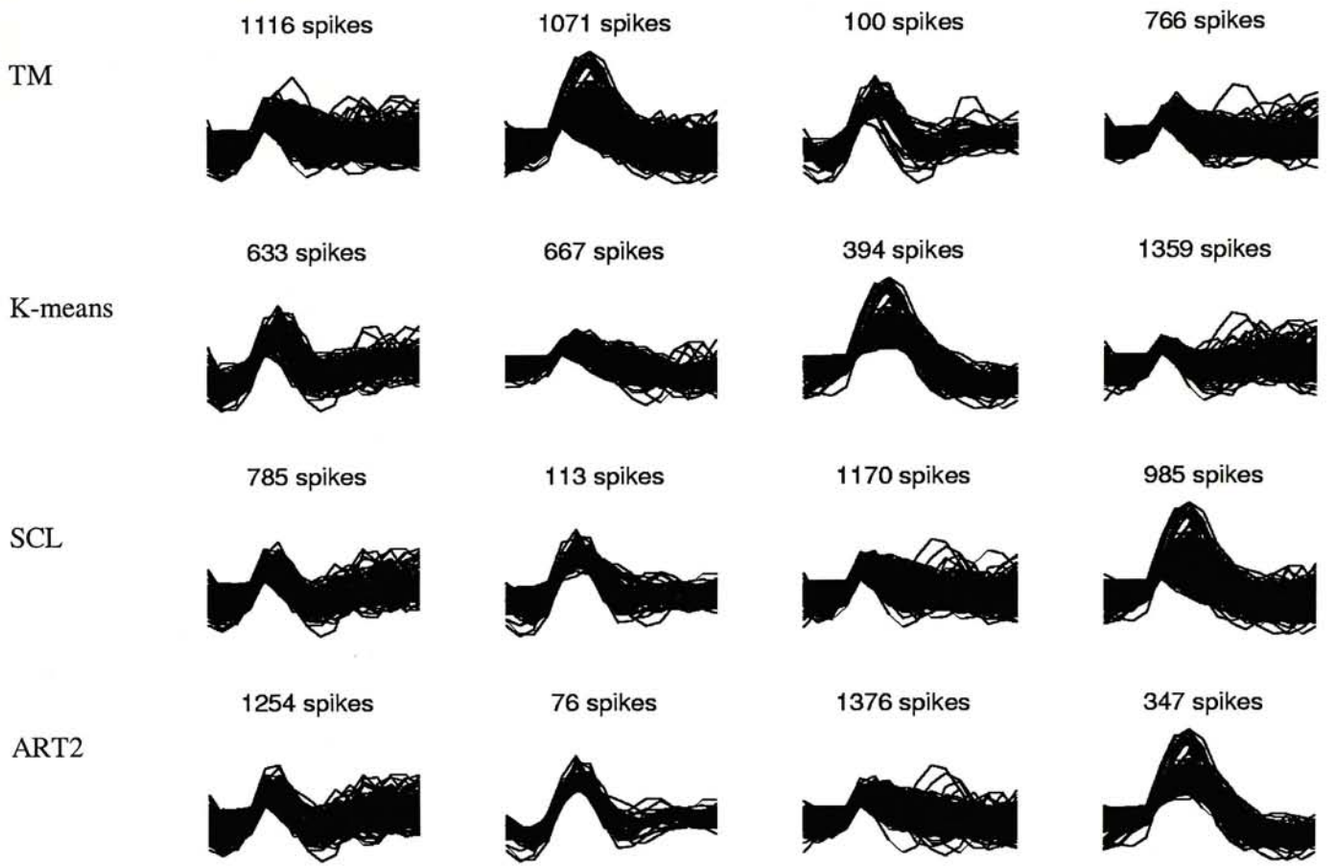

Figure 6-7: Comparison of clustering performance without PCA preprocessing on recorded electrophysiological data set T21.

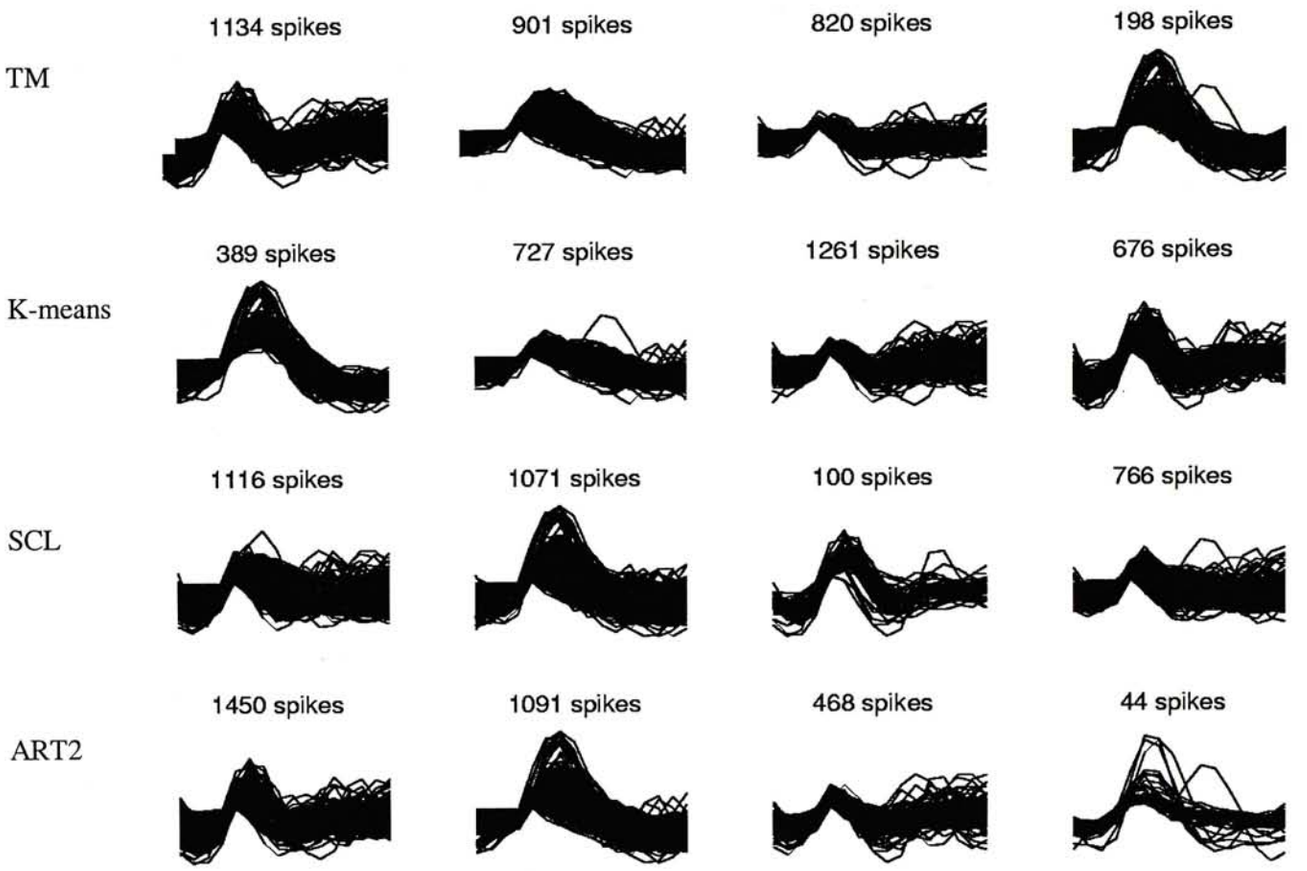

Figure 6-8: Comparison of clustering performance with PCA preprocessing on recorded electrophysiological data set T21

6.2.3 Remarks

Principal Component Analysis

It had been established that applying principal component analysis to the waveforms before clustering was beneficial in terms of convergence (see Figure 6-4) in the presence of noise. The computational efficiency of performing clustering algorithm alone (without including the cost of PCA) was also improved. The first two PCA components accounted for 75% of variance of the original spikes with dimensionality reduced from 25 to 2 (8% of original dimensionality). The selection of number of PCA components L used in our analysis was justified as reasonable since clustering with 2 PCA component had improved convergence ability of all the four selected clustering techniques accompanied with a favorable linear dimensionality reduction of 92%.

It was envisaged that the PCA be performed in an off-line manner before the real-time clustering of spikes. Though the convergence performance was improved, PCA was considerably computationally expensive as a feature extraction method for SCL and TM (computing the covariance coefficients for 1000 spikes for dimensionality reduction was more computationally expensive than performing SCL/TM clustering of original spikes). The whole clustering process for SCL and TM became significantly slower when PCA was included (see Figure 6-5). PCA was always recommended for K-means and ART2 since it improved convergence performance slightly with a high reduction of total computational cost. However, memory requirement of PCA was high. Although this was not a problem for a standard personal computer, it could be a problem for some memory constrained DSP systems. Another possible problem was that eigenvectors were computed from a data set which should be representative of the actual spike data's covariance. If the actual data being clustered changes significantly during the clustering process, the principal components will no longer point in the directions of greatest data variance and the efficiency of the compression would be reduced. In such cases it would be prudent to use a larger value of L in the dimension reduction step of the PCA algorithm.

K-means Clustering

K-means was found to be the most reliable clustering algorithm with the best convergence result. It main advantage was that it could update the templates based on the global spike population while other templates update was based on current incoming spike data for other clustering methods. K-means employed a global optimization policy in which all the spike waveforms participated in the decision process in order to make the best partitioning. However, the main problem of K-means was that it cannot work in an online manner and had computational requirements much greater than other clustering techniques. Adaptation process was slow and involved much computation. This explained why K-means performed well in clustering convergence but was inefficient in term of consumed computational resources. There was no tunable parameter in K-means except the user needed to decide the number of cluster. K-means was highly unsupervised to use with the minimum human intervention required among the four clustering methods. Therefore, K-means could be the best choice for off-line clustering if speed of execution was not important, but not in real situation where speed was a critical factor.

Due to its global optimization of partitioning policy, training of all the K-means clustering algorithms was performed in an off-line manner in contrary to other clustering methods. In K-means, the training step was performed over the entire data set. Therefore, K-means is normally applied to spike sorting (Salganicoff et al., 1988) by clustering an initial representative data set to compute cluster means which are then frozen in value, real-time clustering being performed by assigning the spike to the nearest cluster mean using Euclidean distance comparison.

Simple Competitive Learning (SCL)

The SCL algorithm was found to be faster, more accurate and had better convergence properties than the TM and ART2 algorithms. SCL was the fastest because of its simplicity in algorithmic design as a single layer of neural computational network. The convergence performance of SCL and TM was similar and was close to K-means which had been

recognized as the most accurate clustering algorithm. Since SCL was reasonably fast and accurate accompanied with its unsupervised and online nature, it was fulfilling most of the requirements that were specified before for an ideal spikes discrimination algorithm. The SCL clustering performance showed a successful application of simple unsupervised neural network learning in real time signal recognition problem.

Although SCL was found to be a fast and accurate algorithm for the clustering of these data, its neural learning algorithm was much less intuitive and its convergence ability was highly sensitive to selection of the learning rate parameter value η. This made it difficult for users to understand how a particular clustering was derived. In contrast, TM and K-means used simple Euclidean distance comparison and were more intuitive algorithms than SCL. This made it easier for physiologist to tune the input parameters in order to obtain a desired clustering result.

Online Template Matching

Template matching had similar convergence and efficiency performance with SCL from the analysis. It employed a simple computational algorithm by comparing the Euclidean distance of incoming spike with current template set. This implementation of online template matching was unsupervised in nature. The only tunable parameter D controlled the convergence ability in an intuitive way and users could tune to the desired result without much difficulty. The convergence of TM was much less sensitive than the case of SCL, which was a clearly desirable property as a solution to real time spike discrimination as users could control the clustering performance easily. For classifying the highly spontaneous spike firing activities, TM appeared as an another adequate clustering method beside SCL. TM required close to zero knowledge about the waveforms to be sorted and its implementation was easily manageable. This explained why it was one of the earliest applied classification techniques applied to the problem. The low memory requirement of TM enhanced its applicability since

we had to handle a large number of spikes with shape having multiple inflections which could drought memory resource of the system if memory was not utilized efficiently.

In some definitions, template matching is a supervised learning technique in which the templates are chosen prior to real time classification. However, the implemented online template matching technique used in this thesis does not require any prior knowledge of data set and all the templates are defined and modified during runtime of the experiments without human intervention. No prior training epoch is necessary and the incoming data are assigned to various clusters in an unsupervised manner without prior teaching. Therefore, this version of template matching can be reasonably justified as an unsupervised learning method.

ART2

ART2 had the worst convergence result among the four selected clustering algorithms and also the second most computational expensive algorithm. In aspect of ease of handling to users, the ART2 algorithm was the hardest of all. The main problem of ART2 was its parameter tuning difficulty had greatly limited its ability to achieve highly convergent result when the same parameter tuning effort was placed as applied to other clustering methods. ART2 had not only the vigilance parameter ρ which was adjusted in our experiments, but also a, b, c, d and θ which were kept as constants but could conceivably cause dramatic change the clustering results. All these factors made ART2 limited in its applicability to sort spike trains signal. Other experimenters may question on this unfavorable result of ART2 since ART2 had been successfully applied to other field of pattern recognition problems. These unfavorable remarks on ART2 were aimed at arousing attention that the clustering convergence control of ART2 was distributed among the whole parameter set. To tune ART2 to the best performance is a very complicated issue and its parameters tuning problem has limited ART2 to application of spike discrimination problem when compared with other clustering techniques using the same parameter tuning method. Based on spike discrimination algorithm design criteria stated in section 2.3.1, the most favorable clustering algorithm should be highly

unsupervised with minimum human intervention. Obviously, the parameter tuning procedure required much human intervention if satisfactory clustering result was to be found using ART2. Therefore ART2 was remarked as unfavorable clustering solution for real-time spikes discrimination problem in this analysis.

ART2 is more computational complex than SCL. ART2 incorporates more functionality to give solutions to some difficult tradeoff in pattern recognition problems like the stability-plasticity tradeoff. Orientating subsystem processing and the repeated signal back-propagation procedure from F2 layer to F1 layer in case when a spike from a new family was handled consume much computational resources. Based on other successful works of ART2 in other pattern recognition problem, it is still believable that this increased computational loading can be a good tradeoff in order to gain better clustering convergence performance. This would require the parameters set in ART2 to be feasibly finely tuned to give a desirable result using a reasonably unsupervised parameters-tuning method. Since lengthy analysis of parameters tuning in ART2 was out of scope of this thesis, the parameter tuning aspect of ART2 was left as further research work to other interested experimenters.

Applicability of clustering analysis to spike discrimination

Clustering analysis techniques were to be a successful solution to electrophysiological spike discrimination problem. A rigorous comparison all of the algorithms had been made in terms of their convergence performance and computational efficiency. K-means was found to be the most accurate but consuming much heavier computational loading than other methods. SCL was found to successfully trade off some accuracy but achieving much better efficiency. The overall performance of clustering analysis method for spike discrimination was promising. Computers continue to improve in speed at an exponential rate and real-time implementations of most of the algorithms have already been reported. It is believed that as both electrode and computer technology advances, electrophysiologists will increasingly make simultaneous extracellular recordings from multiple sites to gain a more global view of neuronal computation. In order to discriminate spikes from tens to hundreds of probes in real-time,

clustering systems would require the most computationally efficient algorithm that does not sacrifice accuracy.

6.3 Headphone Sound Quality Control

Experimental results in listening tests, statistical correlation measurement and confirmation tests of designed headphones sound quality control method are presented in this section.

6.3.1 Headphones Frequency Response Clustering

52 headphones of the same model were clustered into 4 groups using K-means algorithms based on frequency response parameters. The number of cluster was determined by the F_{ratio} indication method as described in section 4.3.3. Frequency response parameters were chosen to cluster the headphones because it had been a well-known criterion to compare sounding quality of different audio equipment or audio headsets from different models. The headphones in our analysis were all manufactured with the same model, coming from the same production line and produced on the same day. Classification of headphones with close inherent physical properties based on frequency parameters is still a method with unproven success and applicability. If frequency response were a good measure to classify headphone quality, one might expect that a single cluster would contain a disproportionally high number of similar sounding headphones. However from Table A-3, it was found that headphones with the highest listening score and lowest score came from the same cluster. It could be concluded that frequency response was not a good indicating measure to classify headphone sound quality for headphones with close inherent physical properties.

In order to reduce the sample size for making subsequent listening test and parameters measurement feasible, the previous clustering result of 52 headphones using K-means algorithms based on frequency response parameters were used for sample selection. Selection of a small number of samples from each cluster then followed to form a new test set with greatly reduced sample size. This selection procedure based on clustering result of frequency response parameter of headphones helped to ensure that the whole spectrum of headphones

frequency parameter property had been included in the selected sample set. A good analogy to explain this methodology is that it is a commonly well established statistical technique to perform population census based on samples taken from clustered population groups according to the age range of the whole population.

6.3.2 Listening Tests

Listening tests were then performed on the new test sets to determine subjective sound quality of the headphones. A total of 51 different people were used as test subjects. A single figure of merit was derived from the listening results by assigning a score of 3 to each occurrence of a first place, 1 for a second place and 0 for a third place result and then dividing by the total number of results to normalize the data. The resulting scores are shown in Figure 6-8 (the frequency distributions are shown in Table A-3.)

Looking at the scores associated with each headphone, they are grouped around three values. G was clearly superior to the other headphones. Headphones C, D, A and B followed although their differences in score were not large. The worst scoring (and presumably the worst sounding) headphones were F, H and E.

The four most favored headphones according to score (G, C, D and A) originated from four different clusters. This again indicated that a strong relationship between frequency response and listening test preference did not exist as predicted.

6.3.3 Correlation with Measured Parameters

Correlation test at 95% significance level was performed to determine correlation between all the measured parameters with listening test scores. Measurable parameters from headphones are often adjusted by engineers in the hope that optimization in these parameters can directly improve the sound quality of the headphones. Common design criteria include wide frequency response and low distortion. Such an assertion can only be convincible if the parameters are strongly correlated with a listener's perceived sound quality of that headphone. The measured parameters were all tested for their correlation with the score obtained from listening tests. The correlation coefficients of frequency parameters and distortion parameters were shown in

Figure 6.9 (or refer to Tables A-5 and A-6, in Appendix A.2) while those of voice-coil driver parameters were shown in Figure 6.9 (or refer to Tables A-7 and A-8, in Appendix A.2).

Frequency Parameters

Figure 6.9 showed the correlation of each of the 8 headphones gain at low, midrange and high frequencies with listening test scores. It had been observed that none of the correlation values, r, were greater than 0.632. Thus testing the correlation coefficients for significance at the 95% level, we cannot reject the null hypothesis and thus cannot find significant correlation between these parameters and the score. This confirmed the findings in Section 6.3.1 that frequency response was not a good measure for determining the sound quality of these headphones.

Distortion Parameters

Distortion was a commonly used measure of the non-linearity of a system. The measurements shown in Figure 6.9 showed measurements of the 2nd and 3rd harmonic distortions. For the 8 headphones measured. As with the frequency parameters, none of these measurements were correlated with the listening test results. (The total harmonic distortion (THD) was not obtained because our version of the CLIO software did not support this measurement.)

Voice-coil Driver Parameters

The values of driver resonance frequency and Q values measured in Figure 6.10 did not show good correlation with the listening test scores. However, all of the measurements of dc resistance and inductance (R_E, L_{1k} and L_{10k}) had correlation coefficient $r > 0.632$ and hence showed a statistically significant correlation (at the 95% level) with the listening test scores.

These driver parameters (R_E, L_{1k} and L_{10k}) were all derived as a function of the voice coil impedance. Strong correlation of sounding performance with voice coil impedance is reasonable in the sense that voice coil impedance is the basic quality-governing parameters in headphone production. Production of these headphones was done in a very manual fashion and variations in the coil wire length and inductance could result simply from the manner in which the components were placed in the coil-winding machine. That meant sound qualities

of headphones were highly sensitive to quality control in voice coil production stage. This can be a very valuable information to production engineers since they know how to allocate resource for quality improvement in the most sensible and efficient way. However, this correlation result is still a preliminary one. The following confirmation test serves to test validity of this result.

6.3.4 Confirmation Listening Test

More headphones from the same batch which had R_E, L_{1k} and L_{10k} parameters similar to those of headphone G were selected for further listening tests. This confirmation test was applied in an effort to determine whether for this particular model of headphone, the sound quality was indeed directly related to the parameters (R_E, L_{1k}, L_{10k}) found in previous test. The same listening test was performed on a new test set. The new test set was again of sample size 8. It was formed by selecting 4 new headphones with similar parameter values as the headphones scoring the highest marks in the previous test together with 4 headphones randomly from the last test set. The same 52 people were invited to act as subjects in the listening tests.

The newly selected headphones were named AA, BB, CC and DD. As shown in Figure 6.11 (or refer to Table A-4, in Appendix A.2), those new headphones showed disproportionally higher scores than the rest of the set. The correlation measurement results were shown in Figures 6-12 and 6-13 (or refer to Table A-9, A-10, A-11 and A-12, in Appendix A.2). Figure 6-13 showed that the correlation of the three selected parameters (R_E, L_{1k} and L_{10k}) all attained 95% significance level, which was highly consistent with previous correlation test results. All the eight headphones came from the same batch of production. The result indicated the preference of listener towards headphones set with specific R_E, L_{1k} and L_{10k} value. The confirmation test results reinforced the reasoning foundation of previous listening test. The listening test accompanied by statistical correlation measurement process made out the relationship between listener's preference and specific measurable parameters (R_E, L_{1k}, L_{10k}) of headphones in our batch.

6.4 Summary

In this chapter, the experimental results of two problems namely: electrophysiological spike discrimination and headphone sound quality evaluation, were presented. K-means was found to be the most accurate algorithm for sorting spike signal whereas SCL was the most computationally efficient while showing convergence ability close to K-means. A 95% significant correlation was found between physical parameters R_E, L_{1k}, L_{10k} and subjective perceived sound quality of headphones.

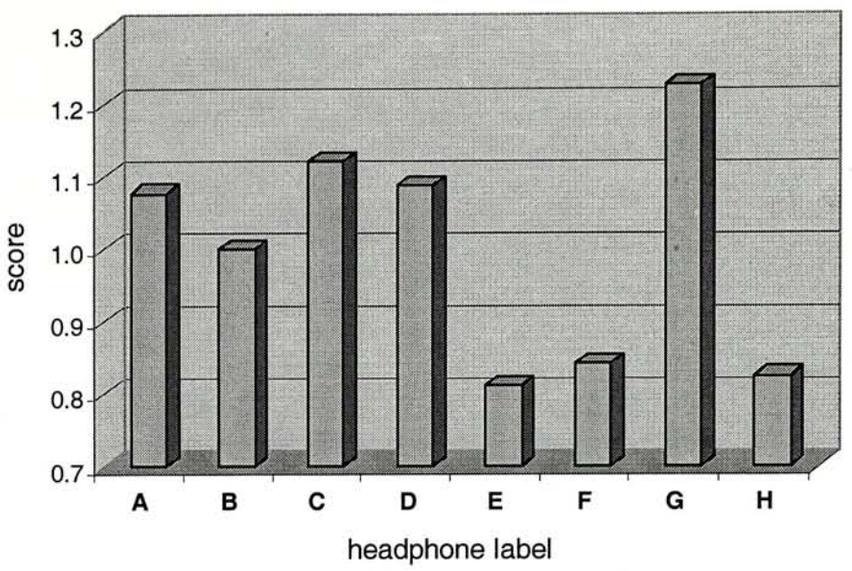

Figure 6-9: Listening test scores of headphone test set

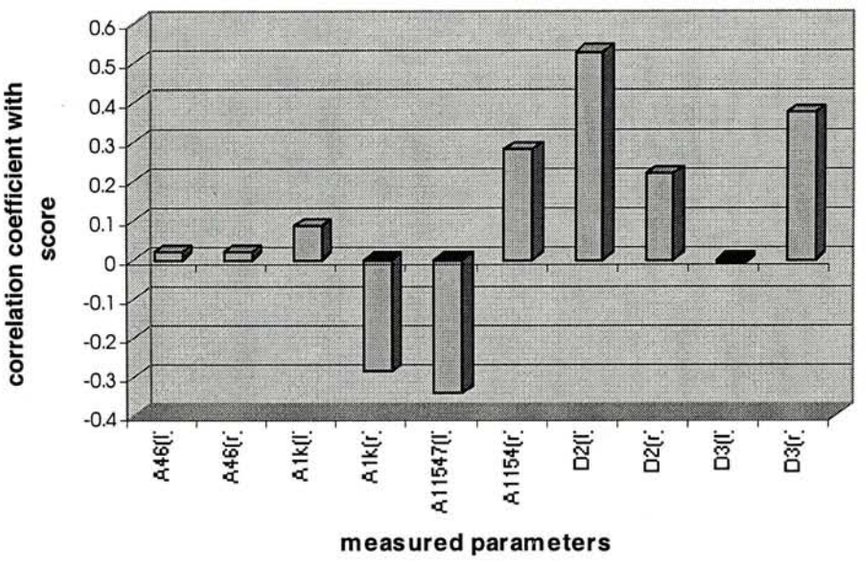

Figure 6-10: Correlation of frequency and distortion parameters with listening test scores

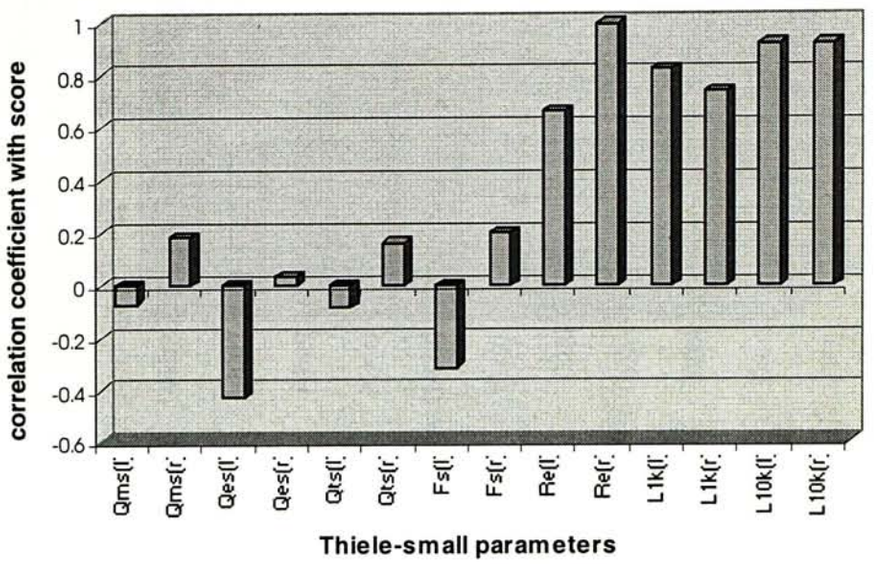

Figure 6-11: Correlation of voice-coil driver parameters with listening test scores

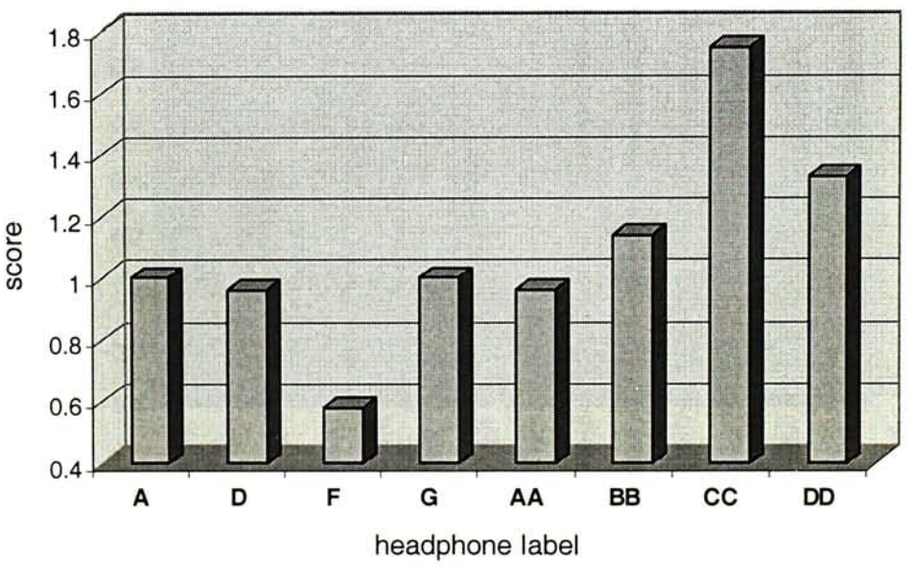

Figure 6-12: Confirmation listening test scores of new headphone test set

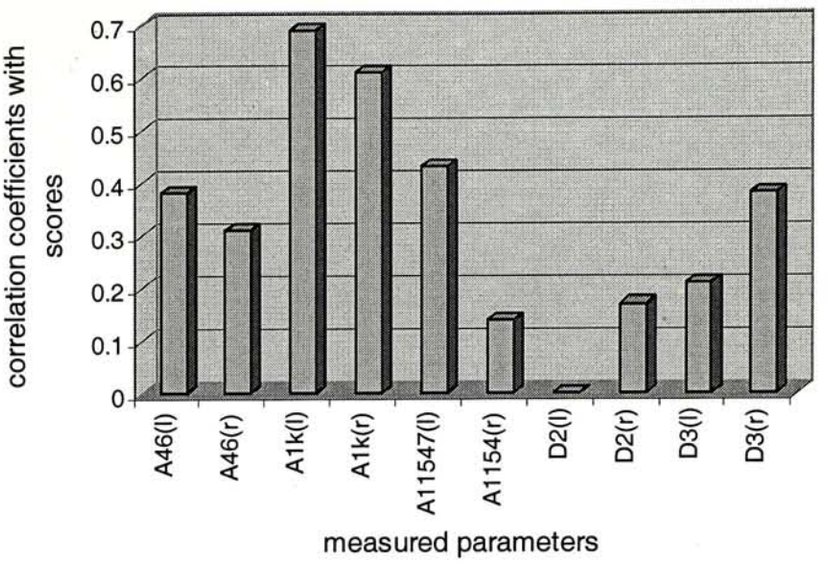

Figure 6-13: Correlation of frequency and distortion parameters with confirmation listening test scores of new headphone test set

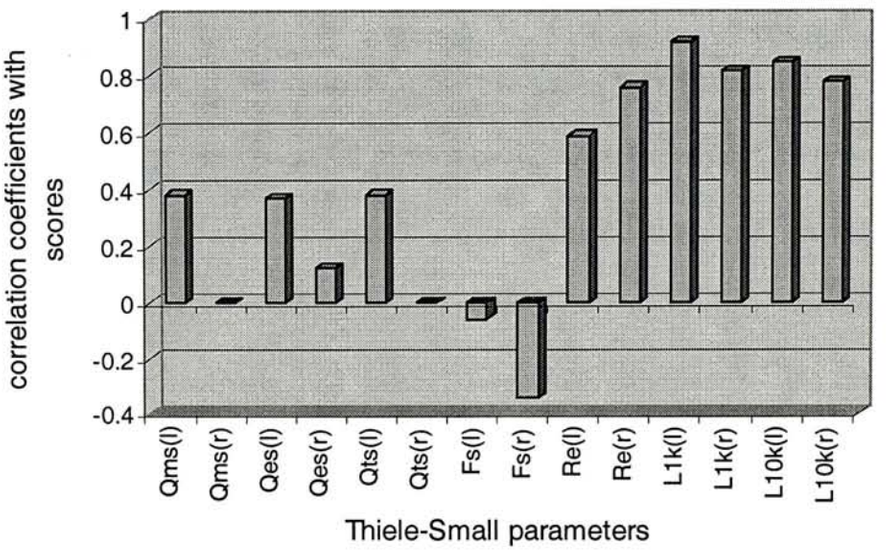

Figure 6-14: Correlation of voice-coil driver parameters with confirmation listening test scores of new headphone test set

CHAPTER 7

Conclusions

The aim of this thesis was to study clustering analysis on practical signal processing problems. The first application, electrophysiological spike discrimination is a natural application for clustering analysis. For this problem, a rigorous comparison of the standard algorithms was made. In the headphone sound quality problem, a novel application of clustering for choosing headphones was developed to find the relationship between sound quality and measured result.

Techniques for the automatic online clustering of extracellular multi-neuron recordings from the nervous system were compared in terms of their efficiency and accuracy. The experimental setup to record electrophysiological spike wave-train signal and extract individual spike waveforms was described in Section 5.2. After identical preprocessing using a Schmitt trigger threshold detector and optionally, principal component analysis (PCA), the template matching, K-means, competitive learning and ART2 algorithms were applied to the data sets. All clustering methods had been successfully applied to the problem with satisfactory overall convergence. Among all the methods, SCL was found to be the most successful clustering method considering the tradeoff between computational efficiency and accuracy. Preprocessing of spike data with principal component analysis can also help in speeding up the clustering process and improving overall accuracy. Although many different algorithms had been proposed in the past (Millecchia et. al., 1988; Gozani and Miller, 1994; Marks, 1965; Wyss, U.R. and Handwerker, 1971; Oghalai et. al., 1994; Gardicke and Albus,

1995), this study was the first to quantitatively compare clustering algorithms for electrophysiological spike sorting, and the first to quantify the benefits of PCA preprocessing.

A method was presented in Section 5.3 for identifying correlation between subjective listening tests and physical parameters. Clustering analysis had been applied in such a way that a large batch of headphones produced from the same production line were clustered so that a manageable number of headphones can be selected for subjective listening tests. A score related to the subjective sound quality of the headphone was determined from these tests and its correlation with measurable parameters was determined. Confirmation tests were performed to validate any statistical findings. The presented approach had successfully identified parameters directly affecting the subjective sound quality of the headphones, showing the reliability of clustering for sample selection. Using the t test for testing correlation coefficients, correlation at the 95% significance level with parameters related to the dc resistance and impedance of the voice-coil driver were identified. The approach developed, to the best of my knowledge, is the only one reported which can directly identify the measurable parameters in headphone design which directly affect the sound quality, and the only study to successfully predict the subject sound quality from measurements in headphones. This problem has applications not only in quality control but can also provide feedback to the headphone designer.

7.1 Future Work

7.1.1 Clustering Analysis

The presented works on spike discrimination and headphone sound quality measurement, of course are not the final words on these problems. Future works can surely further improve the present spike classifier and measurement of headphone sound quality.

Simple competitive learning (SCL) requires the least computational resource while accomplishing the second best convergence result next to K-means. The tradeoff of accuracy when applying SCL in real spike discrimination system can be minimized by fine tuning the

parameters in the ANN model including the learning rate, number of neuronal units and layers. More sophisticated ANN models also have the potential to achieve better clustering results.

Clustering analysis helped in choosing headphone samples in listening tests but it was later shown that frequency response was not strongly correlated with subjective sound quality. It is also feasible to classify the headphone set according to other physical parameters like sound field characteristics (Tannaka Y. et. al., 1989), sound level, cutoff frequency, Thiele-Small parameters, intermodulation distortion and harmonic distortion etc. Different optimization criterions and parameter weightings may also help improving the automatic headphone sound quality clustering performance. Regarding the physical parameters measurements, there are potential problems in the measurement set-up in this analysis that can affect the accuracy of correlation test result. The measurement problems can be readily addressed by repeating the tests using dummy heads which are sophisticated model resembling human heads. Also, multiple repeated listening tests can be done in order to account for random errors likely involved in subjective listening tests.

It should be noted that the presented method for classifying headphones with good sounding quality is not yet a generalized method. It is still arguable that this method can only classify headphones manufactured successfully in the same way as those samples used in the test. In order to generalize this proposed method to other headphone models, further research work is necessary.

7.1.2 Potential Applications of Clustering Analysis

Besides those methodologies presented in this thesis, there are many clustering analysis techniques like exchange algorithm, seriation and graph partitioning (Mirkin B., 1996) which are potentially applicable in many other disciplines of science, such as biomedical engineeering, automatic analysis of images and signals and other industrial problems.

One potential application of clustering is structure and substructure searching in chemical information systems. The two dimensional chemical structure diagram is the prime means of

communication between chemist and is the basis for computerized systems that deal with storage and retrieval of chemical information. One way to improve the searching mechanism is to rank the compounds according to the degree of similarity to the targeted compound (Willett P., 1987), that is to cluster the output from a structure search. Clustering methods can also be used when compounds are being chosen for random testing against primary biological screening by clustering the entire collection of compounds so as to identify the main classes in the data (Willett P., 1987). Compounds may then be selected systematically from the clusters so that all types of structure can be tested for activity in the screening.

Another potential application of clustering analysis is image segmentation. Image segmentation is a basic problem in scene analysis and machine vision. Essentially, image segmentation is to divide image pixels into different categories so that pixels of each category are relatively identical and similar in certain properties whereas two adjacent pixels of different categories shows evidently different properties. It is known that image segmentation is still a difficult problem and that performances of various algorithms are data dependent. Clustering analysis is one of the potential solutions to give a stable segmentation methodology and is widely applied in computer vision.

7.2 Closing Remarks

The aim of this thesis was to verify the applicability of clustering analysis methods in new domains of signal processing problems. Techniques for the automatic online clustering of extracellular multi-neuron recordings from the nervous system were studied. The presented headphone evaluation approach had succeeded in identifying parameters directly affecting the subjective sound quality of the headphones with the help of clustering analysis. The clustering approach was shown to be an elegant method of classifying a pool of neural spike data and headphone frequency response data into groups of high associations. This kind clustering approach should also be suitable for many other signal processing applications.

Appendix

A.1 Tables of Experimental Results: (Spike Discrimination)

Table A-1 and A-2 shows the experimental result of different clustering methods on real spike recordings. The rows of tables correspond to the different recordings and the average and standard deviation are given in the bottom two rows of the table. The average value gives a single figure of merit, summarizing the performance of each algorithm over all of the 21 different recordings. The entries of the SCL case are absolute measured figures whereas the other table entries show results that were normalized to the SCL results and expressed as percentages according to the formulae given in the figure caption.

The results in Table A-2 show the performance of each algorithm with PCA preprocessing. In the first column, the computational requirements of PCA as a percentage of the FLOPS of the SCL clustering algorithm on the reduced data set is shown. PCA computation is performed only once per data set and does not need to be computed for subsequent inputs. Thus its relatively large computation (10 times the cost of computing SCL), mostly in computing the eigenvectors of the covariance matrix, is not included in the FLOPS measures for each of the algorithms.

	TM		K-means		SCL (all entries $\times 10^5$)		ART2		No of Clusters
	FLOPS(%)/ Runtime(sec)	TDIST(%)	FLOPS(%)/ Runtime(sec)	TDIST(%)	FLOPS/ Runtime(sec)	TDIST	FLOPS(%)/ Runtime(sec)	TDIST(%)	
T1	13%/ 925	18	20,433%/ 16345	-5	{17.1}/ 1005	{11.9}	451%/ 4523	15	4
T2	41%/ 978	10	7,993%/ 7565	-1	{4.86}/ 356	{6.89}	567%/ 5623	11	3
T3	46%/ 1125	12	1,182%/ 9865	-1	{1.11}/ 78	{1.39}	572%/ 5895	7	3
T4	-8%/ 810	21	16,235%/ 15655	-2	{14.5}/ 856	{11.1}	450%/ 5623	21	4
T5	84%/ 1520	3	4,848%/ 4256	-2	{4.51}/ 356	{7.22}	569%/ 8516	1	3
T6	31%/ 1100	3	45%/ 1026	-3	{0.667}/ 45	{0.907}	577%/ 5645	-3	3
T7	31%/ 1000	22	651%/ 5600	-1	{0.440}/ 34	{0.572}	583%/ 5462	22	3
T8	19%/ 956	25	13,488%/ 10256	-2	{11.9}/ 756	{9.64}	566%/ 5623	28	3
T9	43%/ 1200	11	4,431%/ 15215	-1	{4.71}/ 325	{6.61}	567%/ 5487	10	3
T10	1%/ 800	2	1,207%/ 12567	-1	{1.45}/ 85	{1.96}	571%/ 5873	12	3
T11	38%/ 1000	1	7,195%/ 18956	-1	{8.04}/ 586	{9.76}	452%/ 5563	0	4
T12	49%/ 1200	-5	4,738%/ 40563	-10	{6.73}/ 456	{8.02}	453%/ 4456	-6	4
T13	47%/ 1000	13	7,388%/ 6567	0	{9.22}/ 562	{5.55}	567%/ 5623	14	3
T14	46%/ 957	3	1,886%/ 18652	-4	{2.01}/ 125	{2.71}	569%/ 5687	3	3
T15	39%/ 1056	0	497%/ 4562	0	{0.504}/ 35	{0.829}	746%/ 5468	0	2
T16	13%/ 896	7	7,334%/ 56321	-3	{5.92}/ 451	{6.50}	453%/ 5556	7	4
T17	36%/ 1056	11	10,129%/ 8523	-5	{13.1}/ 754	{8.39}	451%/ 5462	8	4
T18	6%/ 856	15	7,063%/ 56235	-5	{8.91}/ 562	{4.82}	452%/ 5632	16	4
T19	39%/ 1265	36	8,368%/ 84561	-7	{10.2}/ 599	{6.06}	452%/ 5687	42	4
T20	21%/ 988	16	1,012%/ 8956	-3	{0.796}/ 46	{1.14}	575%/ 4452	15	3
T21	12%/ 863	4	5,808%/ 45614	-7	{10.7}/ 600	{8.03}	452%/ 5867	5	4
Average	31%/ 526	11	6,282%/ 16012	-3	{6.54}/ 412	{5.71}	528%/ 2523	11	
Std	20%/ 165	10	5,292%/ 41175	3	{4.98}/ 292	{3.51}	76%/ 963	11	

Table A-1: Comparison of efficiency and performance measure of algorithms without PCA for clustering spike data. SCL is used as a reference and its actual values of FLOPS and TDIST are enclosed in parentheses. A percentage comparison for the other algorithms computed by $TDIST_X(\%)=100(TDIST_X-TDIST_{SCL})/TDIST_{SCL}$ and $FLOPS_X(\%)=100(FLOPS_X-FLOPS_{SCL})/FLOPS_{SCL}$ appears in the other columns. The bottom rows are the average and standard deviation over the 21 test cases.

	PCA	TM		K-means		SCL (all entries $\times 10^5$)		ART2	
	FLOPS(%)/ Runtime(sec)	FLOPS(%)/ Runtime(sec)	TDIST(%)	FLOPS(%)/ Runtime(sec)	TDIST(%)	FLOPS/ Runtime(sec)	TDIST	FLOPS(%)/ Runtime(sec)	TDIST
T1	335%/ 987	19%/ 265	6	4,669%/ 12563	-1	{4.27}/ 256	{11.4}	173%/ 600	6
T2	1,069%/ 845	18%/ 89	1	1,690%/ 1456	0	{1.37}/ 78	{6.86}	191%/ 165	1
T3	1,44%/ 265	17%/ 25	-2	236%/ 74	-3	{0.315}/ 18	{1.42}	190%/ 45	0
T4	399%/ 1024	21%/ 245	4	2,916%/ 6458	-1	{3.61}/ 245	{10.9}	173%/ 512	5
T5	1,081%/ 897	18%/ 74	1	820%/ 712	-1	{1.28}/ 78	{7.09}	191%/ 187	2
T6	1,779%/ 210	16%/ 14	4	23%/ 15	0	{0.189}/ 12	{0.907}	190%/ 26	4
T7	2,199%/ 165	14%/ 8	1	367%/ 41	-1	{0.1.25}/ 8	{0.572}	189%/ 17	4
T8	432%/ 1035	18%/ 256	3	1,837%/ 4123	0	{3.38}/ 195	{9.51}	191%/ 513	3
T9	1,072%/ 876	18%/ 78	0	1,314%/ 1056	-1	{1.33}/ 87	{6.61}	191%/ 197	0
T10	1,321%/ 356	10%/ 26	3	340%/ 104	0	{0.409}/ 25	{1.94}	191%/ 56	0
T11	736%/ 912	21%/ 138	0	1,873%/ 2187	-1	{2.00}/ 112	{9.75}	173%/ 256	0
T12	886%/ 875	22%/ 108	-2	1,006%/ 912	-2	{1.68}/ 93	{7.47}	173%/ 226	0
T13	562%/ 945	18%/ 160	12	489%/ 801	0	{2.61}/ 145	{5.57}	191%/ 354	15
T14	1227%/ 421	16%/ 40	6	417%/ 175	0	{0.569}/ 35	{2.59}	191%/ 82	5
T15	1,931%/ 245	13%/ 14	0	69%/ 17	0	{0.168}/ 9	{0.829}	211%/ 25	0
T16	1,002%/ 865	11%/ 90	2	2,166%/ 1750	0	{1.47}/ 83	{6.34}	173%/ 187	1
T17	444%/ 1025	21%/ 216	2	2,392%/ 4875	-1	{3.26}/ 213	{8.08}	173%/ 489	3
T18	657%/ 1025	22%/ 158	1	2,080%/ 2785	-2	{2.22}/ 135	{4.70}	173%/ 310	3
T19	570%/ 923	21%/ 178	2	2,384%/ 3564	-3	{2.55}/ 147	{5.81}	173%/ 356	6
T20	1,617%/ 205	17%/ 20	-1	344%/ 74	-2	{0.225}/ 13	{1.12}	190%/ 28	1
T21	544%/ 915	21%/ 212	-1	3,070%/ 5754	-6	{2.66}/ 167	{7.88}	173%/ 345	2
Average	1,014%/ 680	18%/ 107	2	1,453%/ 2088	-1	{1.70}/ 90	{5.59}	184%/ 252	3
Std	529%/ 348	3%/ 83	3	1,192%/ 2623	1	{1.25}/ 69	{3.42}	10%/ 191	3

Table A-2: Comparison of efficiency and performance measure of algorithms using PCA for clustering spike data. SCL is used as a reference and its actual values of FLOPS and TDIST are enclosed in parentheses. A percentage comparison for the other algorithms computed by $TDIST_X(\%)=100(TDIST_X-TDIST_{SCL})/TDIST_{SCL}$ and $FLOPS_X(\%)=100(FLOPS_X-FLOPS_{SCL})/FLOPS_{SCL}$ appears in the other columns. Also, efficiency of PCA component extraction is compared in the same way with SCL execution time as a reference by $FLOPS_{PCA}(\%)=100(FLOPS_{PCA}-FLOPS_{SCL})/FLOPS_{SCL}$. The bottom rows are the average and standard deviation over the 21 test cases.

A.2 Tables of Experimental Results: (Headphones Measurement)

#	*Cluster*	*1st place*	*2nd place*	*3rd place*	*Score*
A	1	8	4	14	1.08
B	4	5	9	10	1.00
C	3	7	7	11	1.12
D	2	7	4	12	1.09
E	4	7	5	20	0.81
F	2	3	7	9	0.84
G	1	10	8	13	1.23
H	1	4	7	12	0.83

Table A-3: Frequency distribution of listening test results. The first column is the headphone identification letter, the second column is the cluster number from which it was selected, the 3rd-5th columns represent the frequency of a 1st-3rd ranking respectively and the last column represents the normalized score given to that headphone.

#	*1st place*	*2nd place*	*3rd place*	*Score*
A	9	7	18	1
D	5	3	11	0.96
F	3	6	17	0.58
G	6	10	12	1
AA	6	5	13	0.96
BB	6	6	9	1.14
CC	12	6	6	1.75
DD	5	9	4	1.33

Table A-4: Frequency distribution of confirmation listening test results. The first column is the headphone identification letter, the 2nd-4th columns represent the frequency of a 1st-3rd ranking respectively and the last column represents the normalized score given to that headphone.

#	Score	A_{46Hz} (left, dB)	A_{46kHz} (left, dB)	A_{1kHz} (right, dB)	A_{1kHz} (left, dB)	A_{1157Hz} (left, dB)	A_{1157Hz} (right, dB)
G	1.23	80.23	77.67	85.71	85.71	68.14	76.09
C	1.12	85.1	84.92	88.02	88.42	79.63	75.57
D	1.09	78.41	81.28	85.49	86.18	69.83	75.2
A	1.08	84.63	84.63	87.51	87.51	80.02	80.02
B	1.00	84.84	86.48	88.72	88.42	81.25	83.61
F	0.84	82.42	84.15	84.77	87.69	77.93	79.6
H	0.83	84.27	82.21	87.69	87.15	76.37	70.82
E	0.81	77.29	82	86.68	85.32	75.27	71.98
r		0.0195	0.0212	0.0843	-0.286	-0.344	0.280

Table A-5: Measured frequency response parameters from the 8 headphones (sorted by score). The bottom row is the correlation (r) of the parameter with the listening test score.

#	Score	D_2 (left, dB)	D_2 (right, dB)	D_3 (left, dB)	D_3 (right, dB)
G	1.23	49.68	33.24	29.81	20.72
C	1.12	41.45	41.47	25.36	27.05
D	1.09	30.19	31.08	24.76	31.54
A	1.08	37.25	37.25	28.52	28.52
B	1.00	39.76	42.17	26.12	22.97
F	0.84	38.87	40.31	25.77	10.13
H	0.83	39.21	40.46	33.64	23.6
E	0.81	28.87	15.1	23.95	25.31
r		0.528	0.219	-0.008	0.376

Table A-6: Measured distortion parameters from the 8 headphones (sorted by score). The bottom row is the correlation (r) of the parameter with the listening test score.

#	Score	Q_{ms} (left)	Q_{ms} (right)	Q_{es} (left)	Q_{es} (right)	Q_{ts} (left)	Q_{ts} (right)	F_S (left, Hz)	F_S (right, Hz)
G	1.23	0.54	0.48	9.04	8.09	0.51	0.45	128.72	140.33
C	1.12	0.65	0.58	8.59	7.79	0.6	0.54	144.14	133.65
D	1.09	0.57	0.64	8.81	7.7	0.53	0.59	139.12	132.48
A	1.08	0.6	0.53	8.2	7.31	0.56	0.49	136.34	129.85
B	1.00	1.27	0.67	9.36	7.31	1.12	0.61	139.95	140.33
F	0.84	0.67	0.53	11.06	6.98	0.63	0.5	141.56	127.3
H	0.83	0.53	0.58	9.08	8.51	0.5	0.54	134.85	141.56
E	0.81	0.66	0.42	8.72	7.83	0.61	0.4	139.12	131.34
r		-0.0739	0.185	-0.430	0.0325	-0.0834	0.164	-0.315	0.201

Table A-7: Measured voice-coil driver parameters (I) from the 8 headphones (sorted by score). The bottom row is the correlation (r) of the parameter with the listening test score.

#	Score	R_E (left, Ω)	R_E (right, Ω)	L_{1k} (left, mH)	L_{1k} (right, mH)	L_{10k} (left, mH)	L_{10k} (right, mH)
G	1.23	34.7	35.4	1.09	1.17	0.16	0.17
C	1.12	34.8	34.3	1.08	1.14	0.16	0.16
D	1.09	34.9	34.3	1.09	1.11	0.16	0.16
A	1.08	35	34.3	1.1	1.13	0.16	0.16
B	1.00	34.3	34.2	1.05	1.17	0.16	0.16
F	0.84	33.7	33	0.99	1.1	0.15	0.15
H	0.83	34.5	32.5	0.92	1.01	0.15	0.14
E	0.81	34.4	33.2	1.03	1.09	0.15	0.15
r		0.660	0.995	0.822	0.735	0.916	0.918

Table A-8: Measured voice-coil driver parameters (II) from the 8 headphones (sorted by score). The bottom row is the correlation (r) of the parameter with the listening test score.

#	Score	A_{46Hz} (left, dB)	A_{46kHz} (left, dB)	A_{1kHz} (right, dB)	A_{1kHz} (left, dB)	A_{1157Hz} (left, dB)	A_{1157Hz} (right, dB)
A	1	84.63	84.63	87.51	87.51	80.02	80.02
D	0.96	78.41	81.28	85.49	86.18	69.83	75.2
F	0.58	82.42	84.15	84.77	87.69	77.93	79.6
G	1	80.23	77.67	85.71	85.71	68.14	76.09
AA	0.96	95.29	95.29	96.55	96.55	77.35	77.35
BB	1.14	84.64	84.64	93.76	93.76	79.74	79.74
CC	1.75	89.13	89.13	96.45	96.19	79.14	79.14
DD	1.33	89.06	89.06	94.00	93.88	78.49	78.49
r		0.38	0.31	0.69	0.61	0.43	0.14

Table A-9: Measured frequency response parameters from the new headphones test set. The bottom row is the correlation (r) of the parameter with the listening test score.

#	Score	D_2 (left, dB)	D_2 (right, dB)	D_3 (left, dB)	D_3 (right, dB)
A	1	37.25	37.25	28.52	28.52
D	0.96	30.19	31.08	24.76	31.54
F	0.58	38.87	40.31	25.77	10.13
G	1	49.68	33.24	29.81	20.72
AA	0.96	48.24	52.29	42.15	51.20
BB	1.14	47.26	45.22	37.16	24.71
CC	1.75	38.57	40.12	29.24	32.46
DD	1.33	41.64	50.58	37.89	34.87
r		0.00	0.17	0.21	0.38

Table A-10: Measured distortion parameters from the new headphones test set. The bottom row is the correlation (r) of the parameter with the listening test score.

#	Score	Q_{ms} (left)	Q_{ms} (right)	Q_{es} (left)	Q_{es} (right)	Q_{ts} (left)	Q_{ts} (right)	F_S (left, Hz)	F_S (right, Hz)
A	1	0.6	0.53	8.2	7.31	0.56	0.49	136.34	129.85
D	0.96	0.57	0.64	8.81	7.7	0.53	0.59	139.12	132.48
F	0.58	0.67	0.53	11.06	6.98	0.63	0.5	141.56	127.3
G	1	0.54	0.48	9.04	8.09	0.51	0.45	128.72	140.33
AA	0.96	1.38	0.68	29.54	10.79	1.31	0.64	132.58	127.17
BB	1.14	1.38	0.98	29.54	14.78	1.31	0.91	132.58	126.94
CC	1.75	1	0.25	18.88	3.03	0.95	0.23	137.68	125.16
DD	1.33	1.28	1.33	29.73	28.59	1.23	1.27	138.58	124.78
r		0.38	0.00	0.37	0.12	0.38	0.00	-0.06	-0.34

Table A-11: Measured voice-coil driver parameters (I) from the new headphones test set. The bottom row is the correlation (r) of the parameter with the listening test score.

#	Score	R_E (left, Ω)	R_E (right, Ω)	L_{1k} (left, mH)	L_{1k} (right, mH)	L_{10k} (left, mH)	L_{10k} (right, mH)
A	1	35	34.3	1.1	1.13	0.16	0.16
D	0.96	34.9	34.3	1.09	1.11	0.16	0.16
F	0.58	33.7	33	0.99	1.1	0.15	0.15
G	1	34.7	35.4	1.09	1.17	0.16	0.17
AA	0.96	36	35	1.09	1.13	0.16	0.13
BB	1.14	36	35.5	1.09	1.02	0.16	0.12
CC	1.75	35.5	35.5	1.5	1.52	0.28	0.3
DD	1.33	35.5	35.5	1.15	1.26	0.17	0.18
r		0.59	0.76	0.92	0.82	0.85	0.78

Table A-12: Measured voice-coil driver parameters (II) from the new headphones test set. The bottom row is the correlation (r) of the parameter with the listening test score.

Bibliography

Bessou, P. and Perl, E.R. (1969), Response of cutaneous sensory units with unmyelinated fibres to noxious stimuli, J. Neurophysiol., Vol. 32, pp. 1025-1043.

Birkett A. N. and Goubran R. A. (1996), Nonlinear loudspeaker compensation for hands free acoustic echo cancellation, Electronic Letters, Vol. 32, Issue 12, pp. 1063-4.

Carpenter G. A. and Grossberg S. (1987) ART2: Self-organizing of stable category recognition codes for analog input patterns, Appl. Optics, Vol. 26, pp. 4919-4930.

Ding Y. S. (1985), A tone-burst method for measuring loudspeaker harmonic distortion at high power levels, J. Audio Engineering Society, Vol. 33, Issue 3, pp. 145-7.

Everitt B. (1974), Cluster Analysis, An H.E.B. Paperback.

Fee M.S., Mitra P.P. and Kleinfeld D. (1996) Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability, J. Neurosci. Methods, Vol. 69, pp. 175-188.

Freeman J A., Skapura D.M. (1992) Neural Networks: Algorithms, applications and programming techniques, Addison Wesley.

Gadicke R. and Albus K. (1995) Real-time separation of multineuron recordings with a DSP32C signal processor, J. Neurosci. Methods, Vol. 57, pp. 187-193.

Garbrielsson A. and Lindstrom B. (1985), Perceived sound quality of high-fidelity loudspeakers, J. Audio Engineering Society, Vol. 33, No. 1/2, pp. 33-52.

Garbrielsson A., Hagerman B., Bech-Kristensen T. and Lundberg G. (1990), Perceived sound quality of reproduction with different frequency responses and sound levels, J. Acoustical Society of America, Vol. 88(3), pp. 1359-66.

Goto T., Kido K. and Yamada A. (1975), Feedback amplifier distortion-cancelling circuit, Matsushita Electric Ind. Co. Ltd., (Patent).

Gozani S. N. and Miller J. P. (1994) Optimal Discrimination and Classification of Neuronal Action Potential Waveforms from Multiunit, Multichannel Recordings Using Software-Based Linear Filters, IEEE Trans. Bio-Med. Eng., Vol. 41, pp. 358-372.

Harrison J. (1996), An integral limitation upon loudspeaker frequency response for a given closure volume, J. Audio Engineering Society, Vol. 44, No. 12, pp. 1097-1103.

Hartigan, J. A. (1975), Clustering Algorithms, Wiley, New York.

Harwood H. D. (1976), Audibility of phase effects in loudspeakers, Wireless World, Vol. 82, Issue 1481, pp. 30-32.

Hertz, J., Krogh, A. and Palmer, R.G. (1991), Introduction to the theory of neural computation, Addison Wesley.

Jansen R. F. (1990), The reconstruction of individual spike trains from extracellular multineuron recording using a neural network emulation program, J. Neurosci. Methods, Vol. 35, pp. 203-213.

Jennett S. (1989), Human Physiology, Churchill Livingstone.

Jolliffe I.T. (1986), Principal Component Analysis, Springer-Verlag, New York.

Jurgen Schurmann (1996), Pattern Classification, Wiley Interscience.

Kates J. M. (1984), A perceptual criterion for loudspeaker evaluation, J. Audio Engineering Society, Vol. 32, No. 12, pp. 938-944.

Leach W.M. (1979), On the specification of moving-coil drivers for low-frequency horn-loaded loudspeakers, J. Audio Engineering Society, Vol. 27, No. 12, pp. 950-959.

Leach W.M., Schafer R.W. and Barnwell T.P. (1979), Time-domain measurement of loudspeaker driver parameters, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-27, Issue 6, pp. 734-9.

Millecchia R. and McIntyre T. (1978), Automatic nerve impulse identification and separation, Comput. and Biomed. Res., Vol. 11, pp. 459-468.

Mirkin B. (1996), Mathematical Classification and clustering, Kluwer Academic Publishers.

Moller H., Jensen C. B., Hammershoi D., and Sorensen M. F. (1995), Design criteria for headphones, J. Audio Engineering Society, Vol. 43, No. 4, pp. 218-232.

Nakayama T. (1988), Practice of sound quality evaluation techniques I, J. Acoustical Society of Japan, Vol. 44, Iss. 4, pp. 325-331.

Oghalai J.S., Street W.N. and Rhode W.S. (1994), A neural network-based spike discriminator, J. Neurosci. Methods, Vol. 54, pp. 9-22.

Öhberg F., Johansson H., Bergenheim M., Pedersen J. and Djupsjöbacka M. (1996), A neural network approach to real-time spike discrimination during simultaneous recording from several multi-unit nerve filaments, J. Neurosci. Methods, Vol. 64, pp. 181-187.

Penfold R. A. (1976), Total harmonic distortion filter, Practical Electronics, Vol. 12, Issue 3, pp. 216-220.

Russotti J.S., Santoro T.P.,and Haskell G. B., Proposed technique for earphone calibration, J. Audio Eng. Soc., Vol. 36, No.9, 1998 September.

S.R.L. A. (1998), CLIO User's manual.

Salganicoff M., Sarna M., Sax L. and Gerstein G.L. (1988) Unsupervised waveform classification for multineuron recording: a real-time, software-based system. I: Algorithms and implementation, J. Neurosci. Methods, Vol. 25, pp. 181-187.

Sarna M. F., Gochin P., Kaltenbach J., Salganicoff M. and Gerstein G.L. (1988), Unsupervised waveform classification for multineuron recording: a real-time, software-based system. II. Performance comparison to other sorters, J. Neurosci. Methods, Vol. 25, pp. 189-196.

Schmidt, E.M. (1984a), Instruments for sorting neuroelectric data: a review, J. Neurosci. Methods, Vol. 12, pp. 1-24.

Schmidt, E.M. (1984b), Computer separation of multi-unit neuroelectric data: a review, J. Neurosci. Methods, Vol. 12, pp. 95-111.

Schurer H., Slump C. H. and Herrmann O. E. (1995), Second order Volterra inverses for compensation of loudspeaker nonlinearity, 1995 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics, 284 preprint.

Small R. (1972), Closed-Box loudspeaker systems Part I: Analysis, J. Audio Eng. Soc., Vol. 20, pp. 798-808.

Small R. (1973), Vented-box loudspeaker systems part ii: Large-signal analysis, J. Audio Eng. Soc., Vol. 21, pp. 438-444.

Tannaka Y. and Koshikawa T. (1989), Correlation between sound field characteristics and subjective ratings on reproduced music sound quality, J. Acoustical Society of America, Vol. 86(2), pp. 603-620.

Tannaka Y., Muramori. K., Kohashi M. and Koshikawa T. (1990), Correlations between harmonic distortion, sound field characteristics and reproduced sound quality change in listening tests for loudspeakers, J. Acoustical Society of Japan, Vol. 11, Issue 1, pp. 29-42.

Thurmond B and von Reckinghausen D. R. (1992), Measurement and perception of quality in sound systems, Proceedings of the AES 11th International Conference. AES Test and Measurement Conference, 359 preprint.

Tsujimoto K., Subjective evaluation of audio equipment (1996), J. Acoustical Society of Japan, Vol 42, Issue. 10, pp. 816-21.

Vander A. J., Sherman J. H. and Luciano D. S. (1970), Human physiology: The mechanisms of body function, McGraw-Hill.

Williams J. (1996), Statistical Methods, University Press of America.

Wyss, U. R. and Handwerker (1971), Stap-12: A library system for online assimilation and offline analysis of event/time data, Computer Programs in Biomedicine, Vol.1, pp. 209-218.

Yajima M. (1997), Loudspeaker design-for preferable sound quality, J. Society of Instrument and Control Engineer, Vol. 36, Issue 12, pp. 864-867.

Publications

Journal Paper

- Philip Leong, Alex Sim and Craig Jin, "Online Electrophysiological Spike Discrimination: A Comparison of Methods", submitted to the Journal of Neuroscience Methods.

Conference Papers

- A.W.K.Sim, C.T.Jin, L.W.Chan and P.H.W.Leong, "A Comparison of Methods for Clustering Electrophysiological Multineuron Recordings", Proceedings of 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1381-1384.
- P.H.W.Leong, Y.S.Moon, D.W.P.Lam and A.W.K.Sim, "Sound Quality Measurements in Headphones", 106th Convention 1999 of Audio Engineering Society, preprint 4874(B6).
- A.W.K.Sim,"ART2 for Clustering Electrophysiological Multineuron Recordings", ACM 1998 Postgraduate Research Day, pp. 123-126.