

This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (<https://dspace.lboro.ac.uk/>) under the following Creative Commons Licence conditions.



For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

DX 24332



Pilkington Library

Author/Filing Title Temple

Vol. No. Class Mark T

**Please note that fines are charged on ALL
overdue items.**

Loan only

0402291956



**REAL-TIME FPGA IMPLEMENTATION
OF A
NEUROMORPHIC PITCH DETECTION SYSTEM**

by

Arthur Robert Temple B.Eng.(Hons)


A Doctoral Thesis

Submitted in partial fulfilment of the requirements
for the award of

Doctor of Philosophy
of
Loughborough University.

August 1999

© Arthur Robert Temple

 Loughborough University Engineering Library	
Date	Sept 12
Class	
Acc No.	040229195

M0002320CB

ABSTRACT

This thesis explores the real-time implementation of a biologically inspired pitch detection system in digital electronics. Pitch detection is well understood and has been shown to occur in the initial stages of the auditory brainstem. By building such a system in digital hardware we can prove the feasibility of implementing neuromorphic systems using digital technology.

This research not only aims to prove that such an implementation is possible but to investigate ways of achieving efficient and effective designs. We aim to achieve this complexity reduction while maintaining the fine granularity of the signal processing inherent in neural systems. By producing an efficient design we present the possibility of implementing the system within the available resources, thus producing a demonstrable system. This thesis presents a review of computational models of all the components within the pitch detection system. The review also identifies key issues relating to the efficient implementation and development of the pitch detection system. Four investigations are presented to address these issues for optimal neuromorphic designs of neuromorphic systems.

The first investigation aims to produce the first-ever digital hardware implementation of the inner hair cell. The second investigation develops simplified models of the auditory nerve and the coincidence cell. The third investigation aims to reduce the most complex stage of the system, the stellate chopper cell array. Finally, we investigate implementing a large portion of the pitch detection system in hardware.

The results contained in this thesis enable us to understand the feasibility of implementing such systems in real-time digital hardware. This knowledge may help researchers to make design decisions within the field of digital neuromorphic systems.

ACKNOWLEDGEMENTS

I sincerely wish to thank my supervisor, Professor Simon Jones, for his support, guidance and encouragement during the duration of this research. I must also thank him for giving me the opportunity to undertake this research and to attend international conferences during the last three years.

For my financial support I wish to thank the Department of Electronic and Electrical Engineering and the Faculty of Engineering at Loughborough University, without whom this work would not have been possible. I would also like to thank the Royal Academy of Engineering for contributing to the necessary funding I required to attend international conferences.

I must also thank all members of the Electronic Systems Design Group at Loughborough University, both past and present, who have always provided support and encouragement. In particular I must send a special thank you to Dr Ryan Lim. His guidance and friendship have been deeply appreciated. I would also like to thank Dr Julian Yeandel for help in devising a design flow for my work.

Finally, I must send a special thank you to my fiancée, Angie, for making my stay in Loughborough a happy one. Her relentless support and encouragement has proved vital in the completion of this work.

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

1.1	NEUROMORPHIC SYSTEMS.....	1
1.2	SENSORY SYSTEMS.....	2
1.3	AUDITORY SUBSYSTEMS.....	2
1.3.1	Sound Source Localisation.....	3
1.3.2	Pitch Detection.....	3
1.3.3	Source Streaming.....	3
1.3.4	Summary.....	4
1.4	MOTIVATION.....	5
1.4.1	System Modelling.....	5
1.4.2	Technology.....	6
1.4.3	Applications.....	7
1.5	AIMS OF RESEARCH.....	7
1.6	STRUCTURE OF THESIS.....	8

CHAPTER TWO

REVIEW

2.1	OBJECTIVES OF REVIEW.....	10
2.2	AUDITORY SYSTEM.....	10
2.2.1	Auditory Periphery.....	11
2.2.2	Auditory Brainstem.....	12
2.3	PITCH DETECTION.....	14
2.3.1	The Goldstein Model.....	14
2.3.2	The Terhardt Model.....	16
2.3.3	The Licklider Model.....	18
2.3.4	The Hewitt and Meddis Model.....	21
2.3.5	Summary.....	22
2.4	AUDITORY MODELS.....	23

2.4.1	The Cochlear Filter.....	24
2.4.2	The Hair Cell.....	27
2.4.3	The Auditory Nerve.....	29
2.4.4	The Dendrite Filter.....	31
2.4.5	The Stellate Cell Soma.....	31
2.4.6	The Coincidence Cell.....	33
2.5	ARITHMETIC METHODS.....	33
2.5.1	Standard Arithmetic Techniques.....	33
2.5.4	Summary.....	34
2.6	TECHNOLOGY.....	34
2.6.1	DSP Systems.....	34
2.6.2	Programmable Logic.....	35
2.6.3	Non-programmable Logic.....	35
2.6.4	Summary.....	36
2.7	SUMMARY.....	36

CHAPTER THREE

OUTLINE OF INVESTIGATIONS

3.1	OBJECTIVES OF CHAPTER.....	37
3.2	IDENTIFICATION OF RESEARCH TOPICS.....	37
3.3	STATEMENT OF RESEARCH OBJECTIVES.....	38
3.4	EXPERIMENTAL VEHICLES.....	39
3.4.1	Software Simulations.....	39
3.4.2	Hardware Simulations.....	40
3.4.3	Hardware Verification.....	40
3.5	EXPERIMENTAL ASSUMPTIONS.....	41
3.6	INTRODUCTION TO INVESTIGATIONS.....	43

CHAPTER FOUR

IMPLEMENTATION OF HAIR CELL

4.1	OBJECTIVES OF CHAPTER.....	45
4.2	INTRODUCTION.....	45
4.2.1	Chemical Processes Within the Model.....	46
4.2.2	Model Behaviour.....	47
4.3	SIGNAL BIT WIDTH.....	49
4.3.1	Measurement of Error.....	50
4.3.2	Error Results.....	50
4.3.3	Simplified Parameters.....	52
4.4	HARDWARE IMPLEMENTATIONS.....	53
4.4.1	Bit Parallel Implementation.....	53
4.4.2	Bit Serial Implementation.....	55
4.5	SATURATION FUNCTION.....	56
4.5.1	Look-up Table.....	57
4.5.2	Piecewise Linear Approximation.....	58
4.5.3	Saturation Function Conclusions.....	59
4.6	PIPELINING THE HAIR CELL.....	59
4.7	CONCLUSION.....	60

CHAPTER FIVE

AUDITORY NERVE AND COINCIDENCE CELL MODELS

5.1	OBJECTIVES OF CHAPTER.....	61
5.2	INTRODUCTION.....	61
5.2.1	Auditory Nerve Models.....	61
5.2.2	Coincidence Cell Models.....	63
5.3	BEHAVIOURAL EFFECT OF VARYING AN NOISE LEVEL.....	65
5.3.1	Sensitivity of Nerve Fibres.....	65
5.3.2	Noise Level of Simplified Model.....	66
5.3.3	NSR of AN Models.....	67

5.4	COMPARISON OF AUDITORY NERVE MODELS.....	68
5.4.1	Comparison of Behaviour for Auditory Nerve Fibre Models...	68
5.4.2	Comparison of Complexity for AN Models.....	70
5.5	VARIATIONS OF THE COINCIDENCE CELL MODEL.....	71
5.5.1	Original Model.....	71
5.5.2	Simplified Model.....	72
5.5.3	Variation to Simplified Model.....	72
5.6	COINCIDENCE CELL MODEL COMPLEXITIES.....	74
5.7	CONCLUSION.....	75

CHAPTER SIX

DEVELOPMENT OF STELLATE CELL ARRAY

6.1	OBJECTIVES OF CHAPTER.....	77
6.2	INTRODUCTION.....	77
6.3	STELLATE CELL DENDRITE.....	80
6.3.1	First-order or Second-order Filter.....	81
6.3.2	Noise Before or After.....	83
6.3.3	Summary of Hardware Impact.....	85
6.4	STELLATE CELL SAMPLE RATE.....	86
6.4.1	Behavioural Effect of Reducing Sample Rate.....	86
6.4.2	Effect on Complexity of Reduced Sample Rate.....	88
6.5	STELLATE CELL BIT WIDTH.....	89
6.5.1	Behaviour Effect of Reducing Signal Bit-width.....	89
6.5.2	Effect on Complexity of Reduced Signal Bit-width.....	90
6.6	SUMMARY.....	91
6.7	CONCLUSION.....	92

CHAPTER SEVEN

IMPLEMENTATION OF A SINGLE CHANNEL

7.1	OBJECTIVES OF CHAPTER.....	93
7.2	INTRODUCTION.....	93
7.3	ARCHITECTURE OF IMPLEMENTED SYSTEM.....	95
7.3.1	Entire System.....	95
7.3.2	Single Channel.....	97
7.3.3	Single Channel With Only One Stellate Bank.....	98
7.3.4	Single Channel With Only One Stellate Cell.....	98
7.3.5	Summary.....	99
7.4	HARDWARE ARCHITECTURE.....	100
7.5	TEST-BENCH SOFTWARE.....	100
7.5.1	Software Graphical User Interface.....	101
7.5.2	Software/Hardware Communication.....	102
7.6	BEHAVIOUR CHARACTERISATION.....	102
7.6.1	Pure Tone.....	103
7.6.2	White Noise.....	105
7.6.3	Amplitude Modulated Tone.....	106
7.7	CONCLUSIONS.....	108

CHAPTER EIGHT

CONCLUSIONS

8.1	OBJECTIVES OF CHAPTER.....	109
8.2	REVIEW OF OBJECTIVES.....	109
8.3	CONCLUSIONS.....	110
8.4	MEASUREMENT OF SUCCESS.....	113
8.5	LIMITATIONS OF WORK.....	114
8.6	SUGGESTIONS FOR FUTURE WORK.....	115
8.6.1	Extensions of Current Investigations.....	115

8.6.1 Further Investigations.....	115
8.7 SUMMARY.....	116
REFERENCES.....	117
PUBLICATIONS.....	122
 LIST OF TABLES	
Table 4-1 Parameter values of hair cell model.....	47
Table 4-2 Values of hair cell parameters.....	49
Table 4-3 Complexity and speed of bit parallel hair cell implementation.....	54
Table 4-4 Complexity and speed of bit serial hair cell implementation.....	56
Table 4-5 Hardware details of simplified hair cell.....	59
Table 5-1 Auditory nerve model using rejection method.....	62
Table 5-2 Hardware complexity and speed of AN models.....	70
Table 5-3 Detection window widths for each coincidence cell.....	73
Table 5-4 Complexity and speed of coincidence cell implementations.....	75
Table 6-1 Complexity of first and second order dendrites.....	85
Table 6-2 Sample rates and coefficient values for stellate banks.....	88
Table 6-3 Complexity of stellate chopper cell soma.....	88
Table 6-4 Complexity and speed of stellate soma for 12-bit signals.....	90
Table 7-1 Available resources for Gatefield chips.....	95
Table 7-2 Hardware complexity of entire system implementation.....	96
Table 7-3 Hardware complexity of single channel implementation.....	97
Table 7-4 Hardware complexity of simplified channel implementation.....	98
Table 7-5 Hardware complexity of one-of-each implementation.....	99
Table 7-6 Port map of top level blocks.....	100

LIST OF FIGURES

Figure 1-1	Tree structure of where pitch detection fits.....	4
Figure 2-1	Structure of the Auditory Periphery.....	11
Figure 2-2	Cross-section of cochlea duct.....	12
Figure 2-3	Simplified organisation of the auditory brainstem.....	13
Figure 2-4	Neural cell structure.....	13
Figure 2-5	Diagrammatic representation of Goldstein model.....	15
Figure 2-6	Structure of Terhardt pitch extraction model.....	17
Figure 2-7	Licklider model.....	19
Figure 2-8	Delay chain.....	19
Figure 2-9	Structure of Hewitt four stage pitch detection system.....	21
Figure 2-10	The roex auditory filter shape.....	25
Figure 2-11	Transmission line model of cochlea filtering.....	26
Figure 2-12	Response of inner hair cell to tones of 10dB and 40dB.....	27
Figure 2-13	Generalised representation of inner hair cell models.....	28
Figure 4-1	Block diagram of inner hair cell design.....	46
Figure 4-2	Saturation function.....	48
Figure 4-3	Response of inner hair cell to amplitude modulated tone.....	48
Figure 4-4	Graph of percentage error for bit widths 13 to 18.....	52
Figure 4-5	Error using simplified multiplication.....	52
Figure 4-6	Block diagram of design including arithmetic units.....	53
Figure 4-7	Block diagram of bit parallel division unit.....	54
Figure 4-8	Structure of the bit serial divider design.....	55
Figure 4-9	Accuracy of look-up table implementation.....	57
Figure 4-10	Accuracy of linear piece-wise implementation.....	58
Figure 5-1	Diagram of the simplified auditory nerve model.....	63
Figure 5-2	Graph of coincidence cell response.....	64
Figure 5-3	Diagram of simplified coincidence cell architecture.....	64
Figure 5-4	Graph of coincidence performance against h.dt value.....	65
Figure 5-5	Graph of coincidence cell performance against noise level.....	66
Figure 5-6	Graph of NSR against input intensity.....	68
Figure 5-7	Coincidence activity for a number of modulation frequencies.	69

Figure 5-8	Original model response as a function of input gain.....	71
Figure 5-9	New model response as a function of threshold.....	72
Figure 5-10	Coincidence cell response.....	74
Figure 6-1	Response of dendrite filter.....	78
Figure 6-2	Response of stellate soma.....	79
Figure 6-3	Parameters of stellate somas.....	79
Figure 6-4	Comparison of filter responses.....	82
Figure 6-5	Original and new system architecture.....	84
Figure 6-6	Comparison of noise before and after dendritic filtering.....	85
Figure 6-7	Response of system for varying sample rates.....	87
Figure 6-8	Response of coincidence cells for varying bit-width.....	90
Figure 7-1	Entire system implementation.....	96
Figure 7-2	Single channel implementation.....	97
Figure 7-3	Single channel implementation but only one stellate bank.....	98
Figure 7-4	Single channel implementation but only one stellate cell.....	99
Figure 7-5	Snap shot of the software's GUI.....	101
Figure 7-6	System response to pure tone stimuli.....	104
Figure 7-7	System response to white noise stimuli.....	106
Figure 7-8	System response to amplitude modulated tones at 30dB.....	108

CHAPTER ONE

INTRODUCTION

1.1 NEUROMORPHIC SYSTEMS

Neuromorphic systems are electronic implementations (analogue or digital) of mathematical models of the central nervous systems of living creatures. Naturally, the human central nervous system holds most interest for researchers. A high proportion of functional features in the human brain is common to all mammals. Research on live preparations of other mammals has resulted in a greater understanding of the human brain.

It is self-evident that the mammalian brain is an outstanding signal processing device, substantially beyond the capability of algorithmic techniques. Despite this evidence, very few electronic designs exist which exploit the capabilities of neuromorphic signal processing. The size and complexity (number of cells) of neural circuits discourages researchers from attempting to implement them in hardware. The rapid development of semiconductor technology, however, has meant that the complexity of the systems that can be implemented on a single integrated circuit (IC) continues to increase. Consequently, it is time to consider using this technology to implement the signal processing found in living brains.

The brain can be considered as two parts: the cortex (the central region) and the sensory brainstem (the peripheral region). Attempting to implement the cortex is problematical since a clear understanding of the operation and purpose of the cells within the cortex as yet eludes us. Modelling the response of the brain to sensory stimuli, however, is easier as we can directly control the input to the sensory systems.

Through extensive physiological modelling we have a clear knowledge of the structure, function and purpose of the cells in the early part of the sensory systems that are situated in the brainstem.

1.2 SENSORY SYSTEMS

The 5 senses are touch, taste, smell, sight and hearing. Taste and smell are senses that respond to the presence and quantity of certain chemicals in our food and in the air we breathe. Touch generally acts as a pressure sensor although it does have the ability to determine textures. The senses of sight and hearing, however, exhibit more interesting and complex signal processing properties than the other senses. It is for this reason that vision and audition have been the main focus of physiological research over recent years. This research has led to the discovery of directly equivalent processes that occur in both the auditory domain and the visual domain.

Our area of interest is within the field of audition. The main reason for this is that in audition we are dealing with a 1 dimensional input (pressure intensity on the eardrum). This compares with the vision system, which has a 3 dimensional input (light intensity on a point of the retina, and the x and y co-ordinates of the retinal area). This suggests that the electronic implementation of an auditory subtask is considerably less complex than the implementation its visual counterpart.

1.3 AUDITORY SUBTASKS

There are 3 main sub-tasks that exist within the auditory brainstem. These subtasks are described below:

- **Sound source localisation** tells us where somebody's voice, for example, is coming from.

- **Pitch detection** is a process whereby we can detect the perceived fundamental frequency of a complex sound signal.
- **Streaming** focuses our attention on a single sound source in the presence of multiple sound sources.

1.3.1 Sound Source Localisation

The auditory brainstem can perceive whether a sound is coming from the left or the right by calculating the inter-aural time difference (ITD) and the inter-aural intensity difference (IID) [Nandy96a, Nandy96b, Neti92]. The ITD is the time delay between a sound arriving at one ear and the other ear. The IID is the difference between the sound signal strength arriving at each of the ears. Additional information provided by the subtle filtering characteristics of the external ear (pinna), help in the perception of the angle of elevation of a sound source. These filtering characteristics are dependent on the pinna's physical shape and the angle a sound approaches it.

1.3.2 Pitch Detection

The perceived pitch of a natural complex sound signal can be a function of either the fundamental harmonic or the amplitude modulation frequency of a signal [Meddis97]. There exist a number of models that set out to represent the portion of the auditory system that performs pitch detection. These models can be categorised as either spectral (calculations in the frequency domain) or temporal (time domain).

1.3.3 Source Streaming

This is a process where the auditory system can manage to discriminate between a number of sound sources and concentrate on only one of them [McCabe97,

Beauvois94]. A number of models exist where a single sound source is associated with an active frequency band and the perceived locality of that band.

1.3.4 Summary

We must now consider the usefulness of implementing each of these subtasks in electronic hardware. Localisation has applications in guidance systems. Streaming has applications, for example, in signal enhancement by eliminating unrelated sound sources (noise). Pitch detection has applications in music and speech processing. Speech processing, in particular automatic speech recognition, is subject to a great deal of research. A speech signal can be considered as a series of syllables each with its own pitch. This pitch provides information about intonation and emphasis. The combination of speech (or syllable) recognition and pitch detection provides a potential form of speech signal compression, for example. For this reason we identify pitch detection as the most suitable system for electronic implementation.

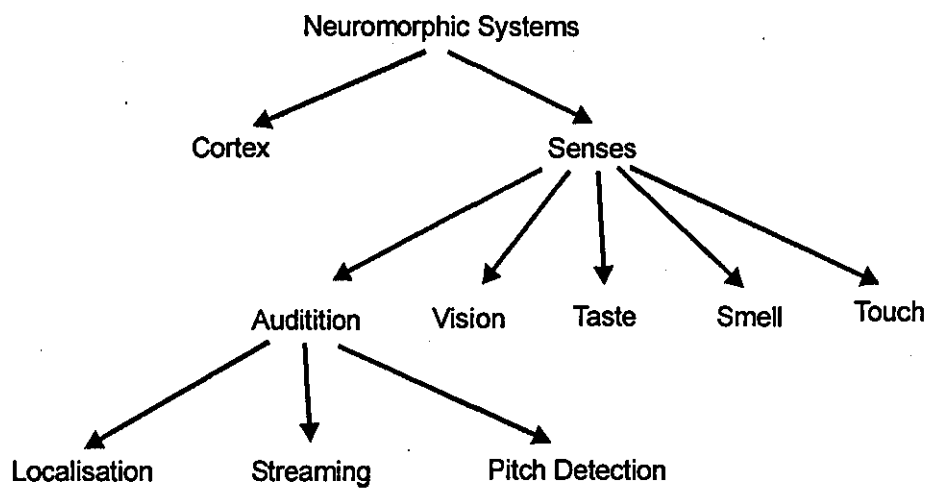


Figure 1-1: Tree Structure Of Where The Pitch System Fits.

In short we can describe our route to selecting the pitch detection system as the neuromorphic system we wish to implement with the tree structure seen in Figure 1-1.

1.4 MOTIVATION

1.4.1 System modelling

In the last decade there have been significant advances in the understanding of the auditory system. These advances include the development of models of individual cells as well as systems of cells. The problem faced by physiologists is that they are only able to measure the response of a few cells at a time in live samples. Although it is possible to capture the behaviour of a small number of cells in this way it is impossible to do the same for a complex system.

System models have to be constructed by proposing a collection of various component cells and how they are connected. Generally, these computational models are implemented in software as a means of evaluation. The advantages of using software simulations are:

- High performance workstations allow implementation of extremely complex models.
- Software flexibility allows response of all the components within the system to be recorded simultaneously.
- Availability of numerical software tools allows detailed analysis of results.

The drawback of software simulations is:

- Complex system simulations run extremely slowly.

We propose a direct electronic implementation of such models as a means of overcoming this drawback. The number of components that can be placed on a single chip continues to grow at an exponential rate while the cost of implementing such large circuits continues to drop. We believe that the advances in auditory modelling and semiconductor technology can be combined to make a significant contribution to the field of neuromorphic systems.

1.4.2 Technology

Electronic implementations of neuromorphic circuits can utilise either digital or analogue technology. A number of analogue electronic implementations of neural cells [Lazzaro92, Lyon88, Schaik96a, Schaik96b] and systems [Lazzaro89a, Lazzaro89b, Schaik97] have been developed in recent years. The main advantages of such implementations are:

- Analogue implementations run in real-time.
- Single transistors can be used to represent components as diverse as multipliers (in the linear region of operation) and filters (with only a few extra resistors and capacitors). This allows for models to be represented by very simple circuits.

There are, however, a number of disadvantages for analogue design which include:

- Analogue circuits cannot easily be programmed. The utilisation of external devices, such as variable resistors, to bias circuits is the only way to achieve any amount of programmability.
- Analogue implementations are constrained to using a one-to-one representation of components because multiplexing is not possible. Multiplexing would eliminate the continuous-time nature of analogue models.

We propose the use of digital technology to overcome these drawbacks. The advantages of using digital technology are:

- Digital electronic circuits can be multiplexed to allow a single processing unit to represent a number of components.
- Digital circuits can be programmed through the use of RAM or register blocks.

The disadvantages of digital technology are:

- A large number of gates are required to perform arithmetic operations. For example several thousand transistors are required to perform a multiplication in digital technology compared with a single gate in analogue technology.

1.4.3 Applications

There are a number of applications for hardware implementations of auditory models. Three examples of such applications are:

- As a front end to a speech recognition system [Jackowoski95, Ghitza88].
- The development of medical prosthetics. To date these implants use mainly analogue technology [Ay97, Kuraishi84, McDermott91].
- The development of training tools for physiologists. It is possible to have an electronic implementation of an auditory sub-system running in real-time and interfaced to a desktop computer. The computer could sample and display the response of each cell in the sub-system graphically.

1.5 AIMS OF RESEARCH

The aim of this work is to establish how a real-time digital implementation of a neuromorphic system can be realized such that the most productive use of limited resources can be made. Our resources comprise a 3 year 1 man time period. The system must exhibit some useful functionality that can be compared to known biological measurements. The knowledge gained in pursuing this goal will allow us to predict whether similar neuromorphic systems can be realised in digital circuitry. To fulfill our aim we set out three objectives.

Our first objective is to explore the architecture of a well-understood neuromorphic system. Such a system consists of richly connected components within a fine-grain architecture. Our goal is to understand how the connectivity and granularity influence

the behaviour of this system. From this knowledge we can develop a new system architecture which incorporates a minimum level of granularity and connectivity while maintaining the characteristic system behaviour. This simplified architecture will result in a more efficient (less complex) design.

Our second objective is to identify computationally efficient and representative models for the biological components that comprise the system. Biological components and cells can be described by a number of mathematical and algorithmic models. These models are assessed in terms of computational efficiency, behavioural accuracy and biological correspondence. Biological correspondence is a measure of how much a model is based on known biological structures and chemical processes. Our goal is to capture the behaviour of biological components and cells in the most computationally efficient way, while representing the biological structures and chemical processes.

Our third objective is to implement a significant portion of the system in real-time digital hardware. In accordance with our overall aim, this sub-system must exhibit some useful functionality. Furthermore, we set out to design software that will allow the user to stimulate the circuit with a range of representative stimuli types and intensities. This software allows the implementation of a demonstrable system. The successful completion of this system fulfills our overall aim. Specifically, we will have established that a real-time digital implementation of a neuromorphic system can be realized within set resource limits.

1.6 STRUCTURE OF THESIS

Chapter 1 identifies the pitch detection system within the auditory brainstem as the neuromorphic system most suitable for electronic implementation. It also identifies digital technology as the most appropriate technology for implementing such a system.

A review of related work is provided in **Chapter 2**. Initially we identify an appropriate pitch detection model for digital electronic implementation. Models of the cells that comprise the system are then assessed. This is followed by a review of computational methods that might be utilised to implement the pitch detection model. Finally, we identify the digital technology most appropriate for our implementation.

Chapter 3 introduces the experimental investigations contained in this thesis and describes more specific research objectives. Development of simplified models of the neural processes and the implementation and utilisation of a portion of the system in digital hardware are each identified as main topics of investigation.

The first-ever digital electronic implementation of the mammalian inner hair cell is presented in **Chapter 4**. Investigations are carried out to find the optimal signal bit-width as well as developing a simplification of the saturation function.

Chapter 5 investigates alternative models of the auditory nerve and the coincidence cell. These new models aim to provide a more efficient digital electronic implementation while capturing the essential behaviour of these biological components.

In **Chapter 6** we focus on simplifying the architecture of the stellate cell array. There is much to be gained from simplifying its architecture, as this is the most complex stage of the system.

Chapter 7 investigates implementing a large portion of the system in digital hardware. Also, it investigates designing software to utilise the hardware implementation and measure its response to a range of stimuli.

Chapter 8 concludes this thesis by reviewing the initial objectives and setting out the main conclusions and contributions to knowledge of this work. The limitations of this research are also considered and the direction for future work is discussed.

CHAPTER TWO

REVIEW

2.1 OBJECTIVES OF REVIEW

In this chapter we review the current state of auditory modelling and discuss suitable technologies for digital neuromorphic hardware. More specifically, the objectives of this chapter are to:

- Assess neuromorphic pitch detection models. These models are assessed in terms of their behavioural accuracy and their correspondence to known biological structures. We will select the most appropriate model following this assessment.
- Assess the models of each of the biological components that comprise the selected pitch detection system on the same basis. Again we select the most appropriate models in terms of behavioural accuracy and biological correspondence.
- Select the most appropriate computational technique. We will look at a number of deterministic arithmetic techniques.
- Select the most appropriate technology. We will assess programmable logic together with non-programmable technologies.

2.2 AUDITORY SYSTEM

Before we begin to describe specifically the cells and systems within the auditory system it is necessary to give a general description of its structure. The auditory system comprises two parts: the auditory periphery and the auditory brainstem [Pickles88].

2.2.1 Auditory periphery

A diagram of the auditory periphery is shown in Figure 2-1.

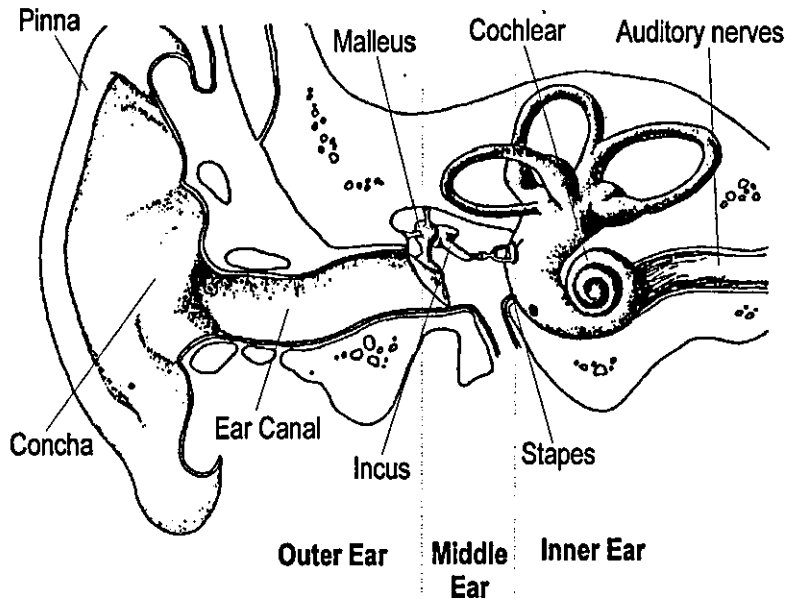


Figure 2-1: Structure of the auditory periphery.

We can see that the pinna, concha, and ear canal comprise the outer ear. These components are responsible for directing sound waves towards the middle ear.

The middle ear, comprising the tympanic membrane, malleus, incus, and stapes, acts as an impedance matching mechanism to allow sound waves in the ear canal to be transmitted efficiently into the cochlea, which is a fluid filled structure. Without such a mechanism much of the sound would be reflected. The function of impedance matching is achieved by the fact that the forces collected over the much larger area of the tympanic membrane are concentrated onto the smaller area of the stapes.

The cochlea is the spiral shape seen in Figure 2-1. The wide end of the coil is called the base and the narrow point in the centre of the coil is called the apex. The basilar membrane (see Figure 2-2) runs along the length of this coil and resonates at certain points when stimulated by a sound. The narrowing shape of the cochlea as it reaches the apex implies that the resonant frequency of the basilar membrane increases along its length. In fact, the opposite is true because the basilar membrane actually

widens as it approaches the apex. When the basilar membrane resonates outer hair cells, which are distributed along its length, also vibrate at the point of resonance. The outer hair cells have the ability to stiffen [Dallos99]. This stiffness is related to the intensity of the signal being applied. This means that that this mechanism acts as a kind of gain control. The outer hair cell response is transmitted to the inner hair cells. The inner hair cells produce electrical impulses. These impulses are transmitted along the auditory nerve. The resonant frequency at a certain point along the length of the basilar membrane is a function of that points position. This means that the basilar membrane acts as a cascade of band-pass filters. Each pass band is known as a cochlear channel.

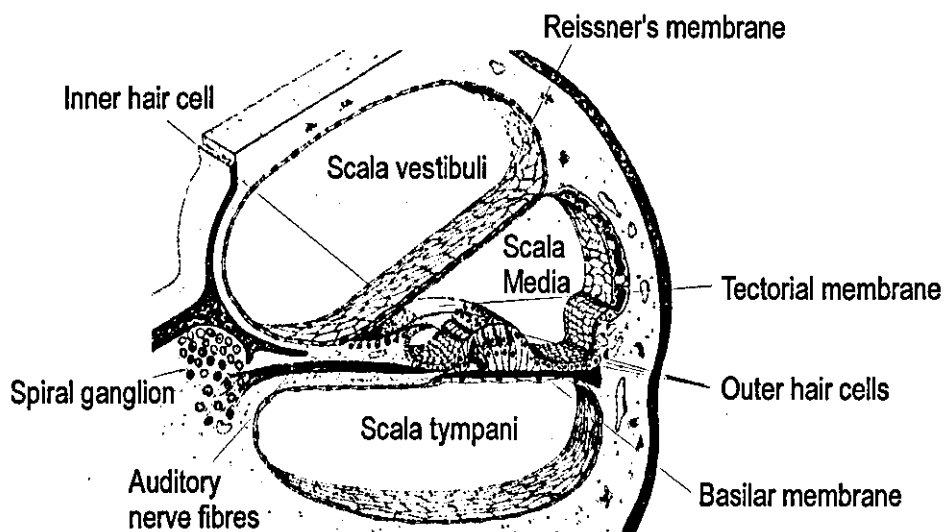


Figure 2-2: Cross-section of the cochlea duct. [Pickles88 ,Fig.3.1B]

2.2.2 Auditory brainstem

A diagram of the auditory brainstem can be seen in Figure 2-3. It can be seen that the auditory brainstem comprises a number of nuclei. Each nucleus is a region that contains certain types of neural cell. Furthermore, the cells within a nucleus tend to be arranged tonotopically. This means that their physical position within the nucleus is related to the frequency of stimulus to which they respond.

The structure of a typical neural cell can be seen in Figure 2-4. The dendrites are input connectors and the axon is output connector to the cell. The chemical processes of the soma (the cell's nucleus) define the behaviour of the cell and are characterised by the flow of many types of ions through the membrane that surrounds the soma. This membrane contains ion gates that are activated and deactivated at certain input potentials and membrane potentials. These gates allow (or prohibit) the flow of ions in and out of the cell soma.

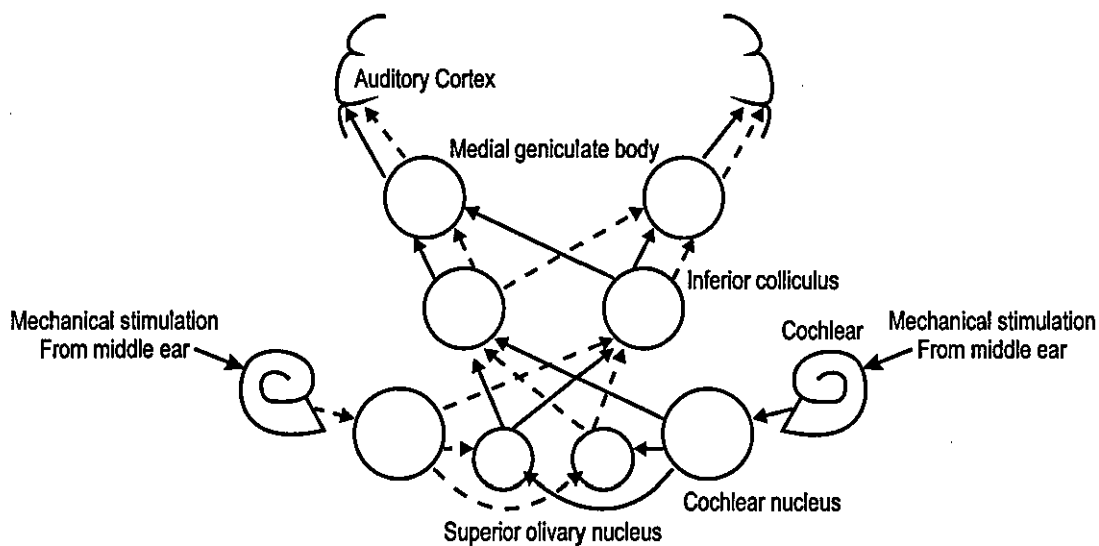


Figure 2-3: Simplified organisation of the auditory brainstem.

Such neural cells communicate with each other by transmitting action potentials, also called spikes. These spikes take the form of brief and sudden voltage changes that are transmitted down a cell's axon to other cells' dendrites.

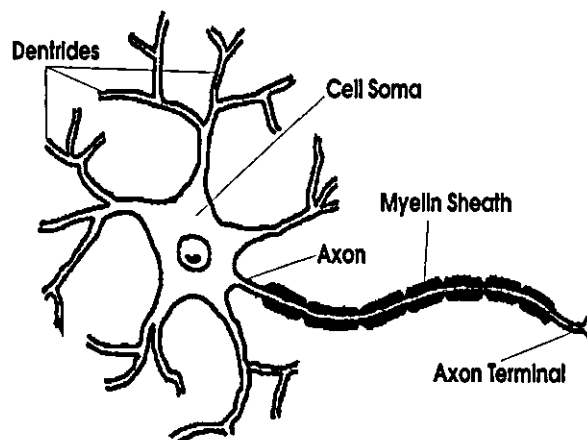


Figure 2-4: Neural cell structure.

2.2.3 Auditory cortex

The auditory cortex is responsible for the most complex tasks of the auditory system. Most importantly, for human, it performs speech processing allowing us to communicate with each other. In the case of other animals, it allows them to recognise, for example, the sound of predators.

2.3 PITCH DETECTION

In this section we will examine a number of computational models for pitch detection. These models fall into one of two categories: spectral theories and temporal theories. Spectral theories resolve their input sound signals into frequency components before calculating the pitch. Temporal theories perform calculations in the time domain to determine the pitch of a sound signal. Because of their computational nature, all of the models described below have the potential to be implemented in digital hardware.

2.3.1 The Goldstein Model

Early attempts to explain pitch detection were based on spectral models. Initially we will look at the Goldstein model [Goldstein73]. The structure of this model is shown in Figure 2-5 and consists of the following three stages:

- Spectrum analysis (band-pass filtering)
- Independent noisy channels
- Central processor

Stage 1 represents the band-pass filtering within the auditory periphery. The information passed on from this stage is an indication of whether a resolved component exists within this channel. A resolved component is where only a single frequency component exists within a cochlear channel. In this model no phase or amplitude information is conveyed.

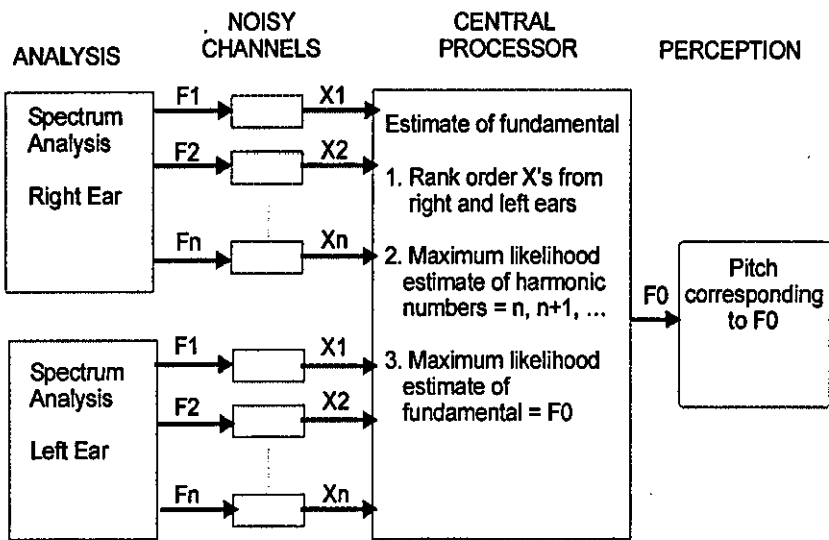


Figure 2-5: Diagrammatic representation of Goldstein model.

The second phase comprises a number of independent noisy channels, one for each frequency channel. The output from each of these channels is the resolved simple tone together with some random error. A Gaussian distribution, $G[f_k, \sigma_k]$, centred on the centre frequency, f_k , of the frequency channel represents this noisy signal and maps the resolved harmonic f_n onto the stochastic signal X_n . The Gaussian distribution, G , approximately represents the filter shape in the frequency domain of the cochlear filter.

The central processor that comprises the third stage of the system makes an optimum estimate of the fundamental frequency on the basis of the noisy representation of the stimulus. This input to the processor is constrained to consisting of successive harmonics. Furthermore, each of the harmonics must be resolvable by the first stage of the system.

To demonstrate the function of the processor we will use the example of 2 consecutive and resolvable harmonics f_1 and f_2 . The processor estimates the fundamental frequency (f_0) and the harmonic numbers ($n, n+1$). Each estimated pair of values of f_0 and n has an associated probability of being correct. The Gaussian functions are described as follows:

$$f_1 \rightarrow X_1, \text{pdf}[x_1] = G[f_1, \sigma_1],$$

$$f_2 \rightarrow X_2, \text{pdf}[x_2] = G[f_2, \sigma_2],$$

where

$$G[f_k, \sigma_k] \equiv (2\pi\sigma_k^2)^{-1/2} \exp[-(x_k - f_k)^2/2\sigma_k^2]$$

The overall probability of a pair of f_0 and n values being correct is:

$$L_{\max} = \{(2\pi\sigma_1^2)^{-1/2} \exp[-(x_1 - f_1)^2/2\sigma_1^2]\} \times \{(2\pi\sigma_2^2)^{-1/2} \exp[-(x_2 - f_2)^2/2\sigma_2^2]\}$$

The value of σ_k (the width of the Gaussian function) is proportional to f_k . This is representative of the fact that the bandwidth of each filter is proportional to its centre frequency.

The values of f_0 and n that give the maximum value of L_{\max} are found. A fundamental limitation of this model is that it cannot cope with unresolved harmonics. It is known that pitch can be detected in the case of unresolved harmonics. Also, there is no representation of the components within the auditory brainstem that might be responsible for performing the calculations.

2.3.2 The Terhardt Model

The Terhardt model [Terhardt82] is also a spectral model whose structure can be seen in Figure 2-6. This model consists of the following five stages:

- Spectrum analysis
- Extraction of tonal components
- Evaluation of masking effects
- Weighting of components
- Extraction of possible pitches

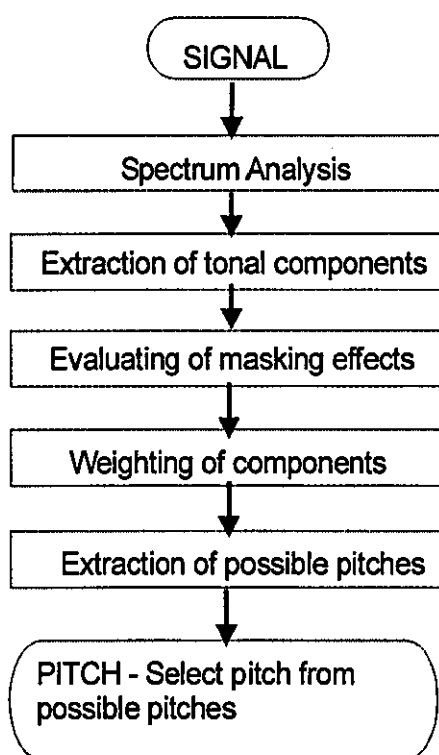


Figure 2-6: Structure of Terhardt pitch extraction model

The first stage performs a frequency analysis of the input. From this we gain a graph of the spectral components. The intensity of the spectral components is measured as sound pressure level (SPL) in dB.

The second stage extracts the tonal components from the frequency spectrum. These tones are found by detecting the local maxima in the SPL against frequency graph. The result of this process is a number of discrete tones and their intensities.

The third stage extracts the relevant components from those initially detected. The SPL excess is calculated for each tone and is measured as the SPL of a tone minus the effective masking SPL of the other tones. Masking is where a tone becomes difficult to hear because of the presence and intensities of other tones. This process makes sure that the tones extracted above are actually audible and therefore relevant to the pitch detection process.

Weighting of components is performed in stage four. This process takes into account the phenomenon of spectral dominance, which represents the way in which humans are more sensitive certain frequencies. This stage assigns weights (WS) to the components depending on their frequency and their SPL excess.

In stage five the pitch is extracted. A number of sub-harmonic tones for each component are evaluated. Each sub-harmonic represents a possible virtual pitch. A weight is calculated for each of these sub-harmonics. There are four criteria that influence the calculation of this weight:

- 1) The number of spectral components which provide the same, or nearly the same, pitch; the weight increases with the number of components.
- 2) The spectral pitch-weight, WS, of the relevant components; pitch weight increases with spectral dominance of involved components.
- 3) The sub-harmonic numbers involved; the weight decreases with increasing sub-harmonic number.
- 4) The accuracy of the near coincidences mentioned in (1); the weight increases with accuracy, attaining a maximum with perfect coincidence.

This model is an improvement on the Goldstein model only because it is not constrained by unresolved harmonics. The frequency analysis stage, however, makes no attempt to model the band pass channels created by the cochlea. Furthermore, the model fails to identify which components within the auditory system might perform the above tasks.

2.3.3 The Licklider Model

The first temporal model appeared nearly fifty years ago [Licklider51, Licklider59]. Here we will describe this model in some detail (see Figure 2-7). The model assumes that the acoustic patterns delivered to the two ears are subjected to mechanical frequency analysis in the cochleas and that the products of that analysis are then

subjected to correlation analysis in unspecified centres of the nervous system. The system performs four discrete functions in sequence on the sound signals entering the ears. These functions are labeled F, G, H, and J. The input to the system is labelled $\theta_1(x,t)$. The x refers to the cochlear channel the stimulus is presented to.

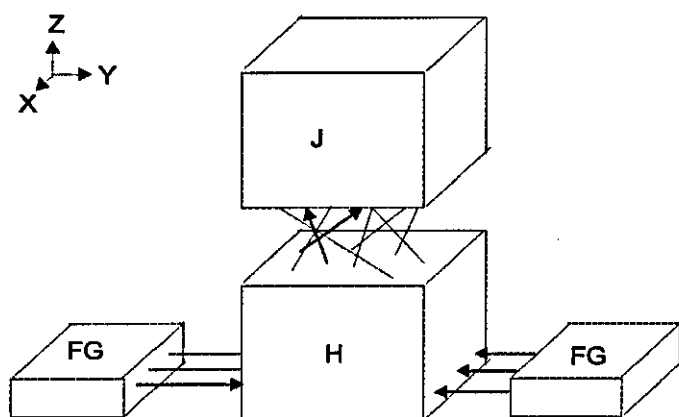


Figure 2-7: Licklider model.

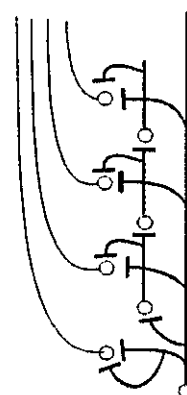


Figure 2-8: Delay chain.

Functions F and G are performed in the cochlea. Function F performs cochlear filtering where a complex sound is separated into a number of frequency channels by a cascade of band-pass filters. The G function is performed on each of the frequency channels. This function initially half-wave rectifies the signal before low-pass filtering it. This low-pass filtered signal is labelled $\theta_2(x,t)$. This signal is presented to a number of spiking neurons. The spiking rate of these neurons is approximately proportional to the intensity of their input. The sum of these spiking waveforms, $I_3(x,t)$, is representative of the signal $\theta_2(x,t)$. The inherent low-pass filtering in functions F and G means that signal $I_3(x,t)$ represents the actual input waveform, $\theta_1(x,t)$, for low frequency signals or the envelope of the signal for high frequency signals.

The H function is performed by an array of spiking neural cells. These cells are arranged in a three-dimensional block. The cochlear filters in the F function are cascaded in the x-dimension therefore activity in a certain x plane represents activity in the associated cochlear channel. Neurons of the same x value receive inputs from both the left and the right inputs. The y-dimension represents inter-aural time difference. If the activity within the cell array is concentrated around a y-value close

to the left side of the array then the signal arrived at that side later than the right side. Activity of the neurons on the base of the H block is described by $I_4(x,y,t)$.

The z-dimension represents temporal coincidence or correlation. Auto-correlation is performed in the z-dimension using delay lines in conjunction with 'straight-through' lines (no delay). These delay and straight-through line can be seen in Figure 2-8. The positions of greatest activity in the z-dimension represent the points of temporal coincidence and have a certain delay time τ associated with them. The time delay associated with a particular area where there is a high level of activity represents a possible harmonic of the pitch value. The auto-correlation function is described mathematically by:

$$I_5(x,y,\tau,t) = k \int_{T=-\infty}^t I_4(x,y,T) \bullet I_4(x,y,T-\tau) dT$$

The J function is described as a self-organising neural network. Initially, all the outputs of H are connected to all the inputs of J neurons. The strength of a connection (weight) between a neuron A, in H and a neuron B, in J is strengthened whenever A participates in the spiking of B. The connection weight is weakened when A fires and B does not. This behaviour is closely related to Hebbian learning in neural networks. Over time the neural network evolves such that a transformation from patterns in H to points in J occurs. This description implies that through exposure to a number sounds (like speech and music) the system 'learns' to detect pitch.

It is clear that functions F and G represent the function of the cochlea and the auditory nerve. Furthermore, a neural structure is proposed to perform the H function. The relation between H and J, however, implies that pitch detection is learned by the system. This contradicts the widely held belief that pitch perception is inherent and is not something that has to be learned.

2.3.4 The Hewitt and Meddis Model

More recently temporal models have continued to be the subject of investigations. Here we describe the most recent model in the temporal field [Hewitt94]. The structure of the model can be seen in Figure 2-9.

The model is made up of four stages:

- Peripheral (mechanical) band-pass filtering,
- Half-wave rectification and low-pass filtering,
- Within-channel periodicity extraction,
- Across-channel aggregation of periodicity estimates.

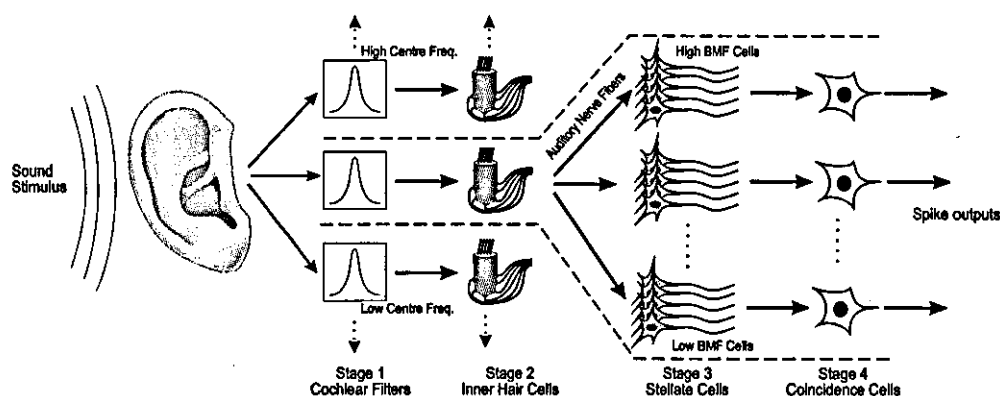


Figure 2-9: Structure of Hewitt four stage pitch detection system.

Stage 1 represents the filtering carried out in the cochlea, which resides in the inner ear. This cochlea acts as a cascade of band-pass filters, which splits the stimulus into a number of frequency channels. The filter centre frequencies are logarithmically spaced between 100 and 8000 Hz.

Half-wave rectification is achieved by hair cells (stage 2) within the cochlea. This cell also performs automatic gain control as well as being responsible for converting the sound (mechanical) signal into a neural (electrical) signal. This neural signal is represented as a series of spikes which are transmitted down the auditory nerve (AN).

The sum of spikes across the AN fibres is representative of the input signal to the hair cell with noise added.

The within-channel periodicity extraction in stage 3 of the system is achieved by a correlation function. This correlation is performed by stellate chopper cells, which reside in the cochlear nucleus region of the auditory brainstem. These are spiking cells and are arranged in banks. Cells within the same bank have the same natural spiking frequency, which is known as the best modulation frequency (BMF). Although these local cells have the same spiking frequency the exact timing of their individual spiking patterns are not necessarily identical. This is due to the fact that the noise added by the AN is uncorrelated for each stellate cell input. However, when the frequency of the cell's input signal matches the spiking frequency of the stellate cell then phase locking occurs. In this case, the stellate cell's output spike is produced at the same time as the positive peak of the stimulus arrives at the input of the cell.

Coincidence cells detect this phase locking in stage 4 of the system, and are believed to reside in the inferior colliculus region of the auditory brainstem. The coincidence cell is also a spiking cell. Its input is the sum of spikes from a single bank of stellate chopper cells. When phase locking occurs in the stellate chopper cell bank, the input to the coincidence cell comprises periodic high intensity stimulus levels. These high intensity bursts cause the coincidence cell to produce a spike at its output. When phase locking does not occur, the input to the coincidence cell is an irregular low intensity signal, which is not intense enough at any time to cause the coincidence cell to spike.

2.3.5 Summary

From the above descriptions we can identify which model has the behaviour which corresponds most to psychological measurements and whose elements have most correspondence to known biological structures and cells. All the models represented the cochlea as a cascade of band-pass filters except the Terhardt model. This discrete

cascade-of-filters representation corresponds to current knowledge of the cochlea. From then on the models diverge in the techniques they employ. Goldstein's model applies noise to the resolved components before a central processor performs a number of calculations upon the noisy signals. This process is algorithmic in nature and cannot be associated with any known biological mechanism. Terhardt's model takes into account a number of known masking and spectral preference characteristics before applying another algorithmic function to predict the pitch. Again, this estimation technique does not correspond in function to any neural structure. This lack of biological correspondence eliminates these two spectral models from our investigations.

The first of our temporal models is that of Licklider. The G function in this model is representative of that of the inner hair cells within the cochlea. The H function is based on spiking neurons and delays lines. Although the existence of this cell array has not been proved the structure is based on neuromorphic principles. The J function within the Licklider model is based on a self-organising neural network. It is believed, however, that pitch perception is inherent to the auditory system and not a result of training.

The Meddis model is similar to that of Licklider except that the left and right signals are processed separately. This removes any inter-aural time difference information. Furthermore, the Meddis model describes known neural cells to perform the correlation-based pitch detection. From this we can say that the Meddis model shows the most accurate behavioural characteristics and biological correspondence. As a result, we will use this model in the investigations carried out in this thesis.

2.4 AUDITORY MODELS

In this section we review a number of computational models for the various biological components that make up the pitch detection system. Although there are four stages to

the pitch detection system there are six different types of computational block that make it up. The six components are:

- The cochlear filter,
- The inner hair cell
- The auditory nerve,
- The stellate cell dendrite,
- The stellate cell soma, and
- The coincidence cell.

We will now review mathematical models for each of these biological components individually.

2.4.1 The Cochlear Filter

The filtering characteristics of the cochlea are such that it acts as a bank of band-pass filters. The models that exist for this filtering vary widely in complexity. It has been shown from biological measurements that the filter shape of these band-pass filters varies, or adapts, with the intensity of the input. The more complex cochlear models exhibit these adaptive characteristics, whereas simpler models have filter shapes that remain constant. Two models that characterise these adaptive effects are the rounded-exponential (roex) and the gamma-tone filters.

The Roex Filter: The first adaptive filter developed was the rounded-exponential (roex) filter [Patterson82, Patterson86]. The roex term refers to the shape of the filter in the frequency domain. This shape is described by the function:

$$W(g) = (1-r)(1 + pg)e^{-pg} + r$$

The g term refers to the frequency of the input signal relative to the centre frequency of the filter. The p term describes the pass-band width and the rate of fall of the skirts

that bound the pass-band. Finally, the r term describes the dynamic range limit beyond which no more attenuation occurs. The $(1 + pg)$ term rounds off the exponential terms at the centre frequency and makes the filter shape continuous. By allowing the parameters to be different for each side of the centre frequency (p_u and r_u for above the centre frequency and p_l and r_l for below it) the filter shape becomes asymmetric. This asymmetry has been observed in experiments on the cochlea. Furthermore, by allowing the parameters p_l and r_l to be a function of sound pressure level P_s , a level dependant filter shape is achieved. This level dependence can be seen in Figure 2-10 and is described as follows:

$$p_l = c_1 + c_2 P_s \quad \text{where } c_1 \text{ and } c_2 \text{ are constants.}$$

$$r_l = c_3 + c_4 P_s \quad \text{where } c_3 \text{ and } c_4 \text{ are constants.}$$

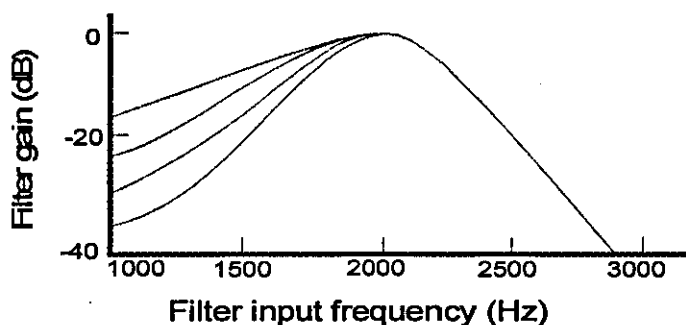


Figure 2-10: The roex auditory filter shape.

The transmission line model: The non-adaptive models of the cochlea filtering effects use transmission line theory [Giguere94, Zweig76]. The cochlea, in these theories, is uncoiled into a straight line. The structure is simplified further by describing this uncoiled structure as a fluid-filled rigid-walled compartment with an elastic partition running along its length. The elastic partition represents the basilar membrane. The behaviour of this structure is described by the classic representation of a transmission line as seen in Figure 2-11. L_{sn} represents the acoustic mass of the scala vestibulae and therefore the inertia associated with that mass. L_n represents the acoustic mass (inertia) of the basilar membrane at point n . C_n is a measure of

compliance of the basilar membrane at that point. R_n represents the resistance of the BM at that point. The resonant frequency and bandwidth of each stage in this representation is describes as:

$$f_n = \frac{1}{2\pi\sqrt{C_n L_n}}$$

$$Q_n = \frac{\sqrt{L_n}}{R_n\sqrt{C_n}}$$

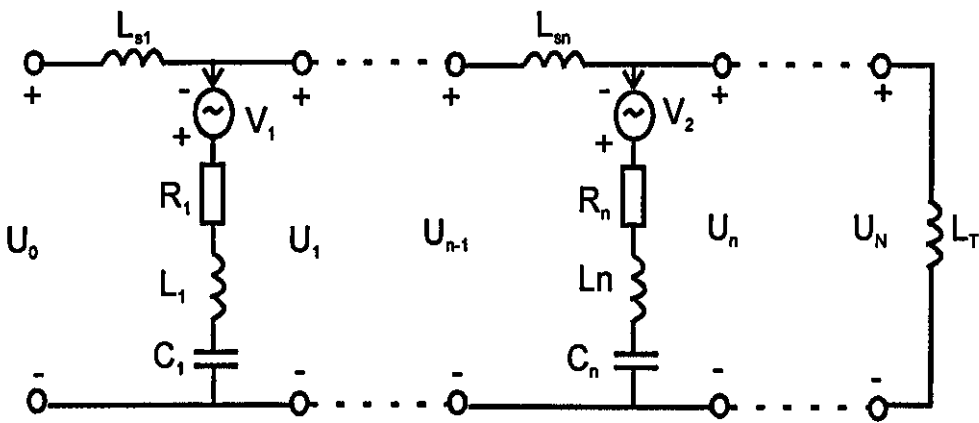


Figure 2-11: Transmission line model of cochlea filtering.

The requirements of our system are such that a non-active cochlear model will meet our needs. The subtle characteristics of the active cochlear filtering models do not influence the behaviour of the pitch detection system. The transmission line model described above maps easily into analogue technology but not digital circuitry. For this reason we use a second order infinite impulse response (IIR) filter. This model takes the form:

$$y(t) = x(t)A_1 + x(t-1)A_2 + x(t-2)A_3 + y(t-1)B_1 + y(t-2)B_2$$

This model, where A_1 , A_2 , A_3 , B_1 and B_2 are constants, provides the simplest implementation of the cochlear filtering in digital technology.

2.4.2 The Inner Hair Cell

In this section we assess a number of inner hair cell models in terms of how accurately they represent the behaviour of the biological cell. We will also assess the models in terms of computationally efficient and how appropriate they are for implementation in digital hardware.

Before we can assess the models we must understand the operation of the inner hair cell [Hewitt91]. The main properties of the inner hair cell, as discovered through physiological studies are:

- Onset and steady-state response to stimuli
- Two-component adaptation
- Recovery of spontaneous activity after stimulus offset

These properties are best observed by looking at the response of a real inner hair cell to the application of 2 tones of different intensities.

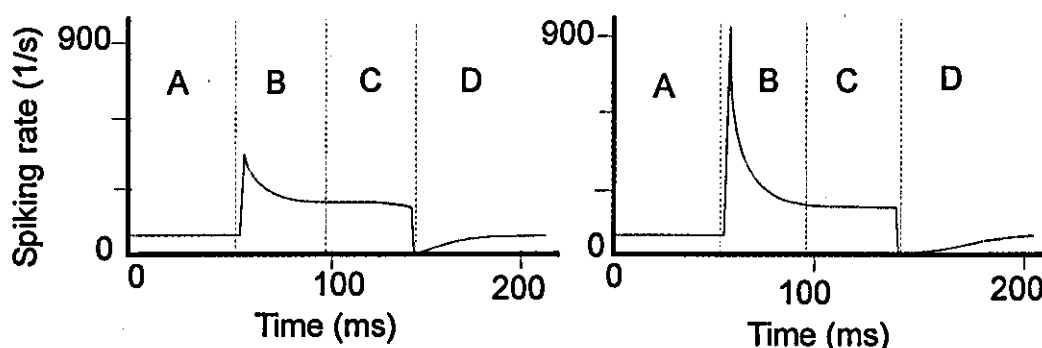


Figure 2-12: Response of inner hair cell to tones of 10 dB and 40 dB.

The above characteristics can be seen in Figure 2-12. In this figure the spike rate produced in the auditory nerve by the hair cell is seen. In this case we apply tones of intensities 10 dB and 40 dB. The general model structure, whose mechanism is responsible for this behaviour, can be seen in Figure 2-13. This general model comprises a number of reservoirs that contain a transmitter substance. The greater the amount of transmitter in the cleft the greater the spiking activity is caused in the

auditory nerve fibres. The permeability of the cell membrane is related to the intensity of hair cell's input. The tones are applied in regions B and C but no stimulus is applied in regions A and D. Regions A of the responses represent the spontaneous activity of the cell when no stimulus is applied. In region B we can see the onset response of the cell and the two-component adaptation. We can also see that the magnitude of the onset response is related to the intensity of the stimulus. This onset is caused by a sudden increase in the membrane permeability combined with full reservoirs. The two-component adaptation consists of a rapid component, causing the steep decline from the onset response, and a slower component, which contributes to the gentle final stage of adaptation to the steady state response. Regions C show that the steady state response is similar for both stimulus intensities. This demonstrates the automatic gain control mechanism inherent within the hair cell. This automatic gain control is caused by a balance between the depleted reservoirs (tending to decrease the level of activity) and the increase in the membrane permeability (tending to increase the level of activity). Regions D show the recovery of the hair cell back to its spontaneous level of activity. We can see that after the application of a high intensity stimulus the hair cell takes longer to recover. This is caused by heavily depleted reservoirs taking time to refill. There is another characteristic of the hair cell, phase locking, which occurs when low frequency tones are applied ($<2\text{kHz}$ in humans). This effect results in the hair cell stimulating spiking in the auditory nerve fibres only during the positive cycle of the stimulus tone. The overall effect is that the temporal information in the signal is preserved within the auditory nerve.

Single Reservoir Models: The first inner hair cell models developed consisted of a single reservoir that released transmitter to stimulate the post-synaptic membrane [Schroeder74]. The amount of transmitter released depended on the stimulus-related permeability of the cell membrane. Such a simple model is unable to represent the two-component adaptation known to exist in the inner hair cell.

Multiple Reservoir Models: More sophisticated models emerged such as that of Schwid and Geisler [Schwid82] whose model consisted of six independent immediate stores of transmitter. The reservoirs were ordered in increasing threshold such that transmitter be released from only 1 reservoir at low stimulus levels and by all 6 at high stimulus levels. These reservoirs were arranged like the immediate stores in Figure 2-13.

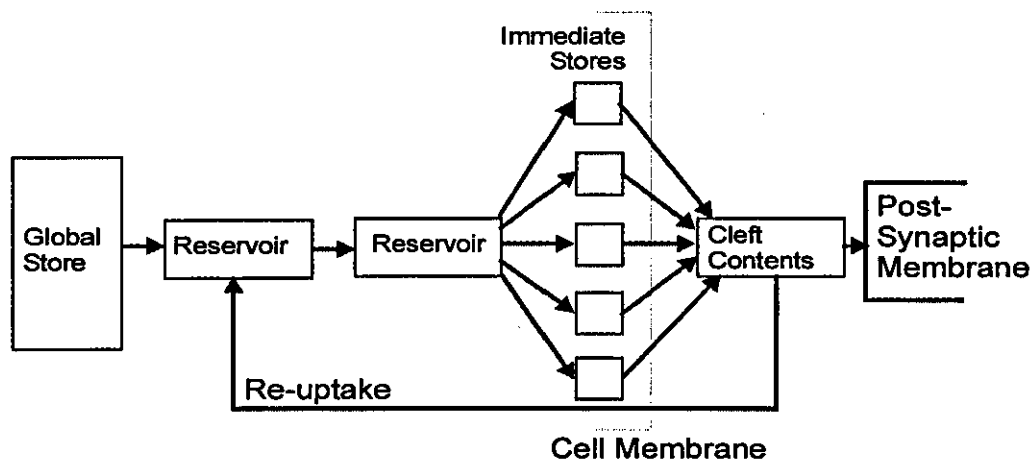


Figure 2-13: Generalised representation of inner hair-cell models.

A model by Ross employed 4 reservoirs in series [Ross82]. The permeability between the reservoirs is constant and so flow rate between them is determined by the concentration or level differences. The transmitter release from the final reservoir is governed by 2 permeabilities. The first permeability is fixed and the second is related to the stimulus intensity.

Feedback Models: A model proposed by Meddis [Meddis86, Meddis88] consisted of 3 reservoirs that incorporated a novel re-uptake and re-synthesis loop. The permeability of the membrane fluctuated as a function of the input intensity. This is the most computationally efficient model that incorporates the three inner hair cell characteristics listed above. For this reason we will use it in the investigations carried out in this thesis.

2.4.3 The Auditory Nerve

The auditory nerve is made up of a collection of nerve fibres. Each nerve fibre is capable transmitting electrical impulses or spikes (sudden and brief changes in voltage potential) from the hair cell cleft to various cells within the auditory brainstem. In the case of the pitch detection system, the AN connects to the input of the stellate cell. An implementation of the auditory nerve has to take the spiking probability output of the hair cell model and calculate how many AN fibres are spiking at that instant. This is a stochastic process and therefore requires random (or pseudo-random) numbers to generate uncorrelated spiking patterns in the auditory nerve.

In each channel of the pitch detection system the auditory nerve connects the hair cell to each stellate cell using 60 nerve fibres. Meddis uses binomial distribution statistics [Hastings75] together with random number generation to calculate which AN fibres are spiking. Pseudo-random numbers are most efficiently generated using linear-feedback shift-registers (LFSRs) as described in [Tsui87]. The binomial spike generation technique can be employed using 2 methods: the rejection method and the geometric distribution method.

The Geometric distribution method: This method uses binomial statistics together with a random number generator. This is applied to a Bernoulli function that calculates how many auditory nerve fibres are firing. In this method p is the probability value produced by the inner hair cell and R_i is a random number between 0 and 1 inclusive. The total number of auditory nerve fibres is n .

$$\sum_{i=1}^k G_i > n \quad \text{where} \quad G_i = \frac{\log(R_i)}{\log(1-p)}$$

$$s = k - 1$$

In this function we add the values of G_i until the sum exceeds n . The number of iterations, minus 1, is number of auditory nerve fibres spiking, s . When p is small this function can be completed in very few iterations. The disadvantage, however, is that it

employs logarithmic and division functions both of which are expensive to implement in digital technology.

The Rejection method: In this case n is a random number whose value is between 0 and 1 inclusive. The value p is the probability value produced by the hair cell.

In this case we take 60 values of n and the number of these that exceeds the value of p is the number of nerve fibres firing, s .

2.4.4 The Stellate Cell Dendrite

The dendrite is the input stage of the stellate cell. Its job is to collect the signals that have been transmitted along a number of auditory nerve fibres (typically 60 fibres). The dendrite in this case has passive electrical characteristics such as resistance and capacitance. The shape of the dendrite is elongated which gives it a low-pass characteristic. There are no other signal processing properties associated with the dendrite.

Meddis simply uses a first order IIR to represent the dendrite filtering. This design provides the simplest solution for our implementation.

2.4.5 The Stellate Cell Soma

The stellate cell that applies to this system is the stellate chopper cell. The 'chopper' in the title refers to the constantly spiking output of the cell when a dc level is applied to its input. The spiking rate of the cell is proportional (but not linearly) to the input dc level. In this section we present a number of models of the stellate chopper cell for evaluation.

The Hodgkin-Huxley Model: This model was one of the first to be published [Hodgkin52, Hodgkin52b] and represents all the chemical processes to be found in a

stellate cell soma. The model of the cell membrane can be represented by the formulae:

$$I_{Na} = g_{Na}(E - E_{Na}),$$

$$I_K = g_K(E - E_K),$$

$$I_l = g_l(E - E_l).$$

The ionic current in the cell is divided into three components: the sodium current (I_{Na}), the potassium current (I_K) and a small leakage current (I_l). The conductances g_{Na} , g_K and g_l vary with cell membrane potential, E , which represents the difference between the potential outside and inside the cell. The potentials E_{Na} , E_K and E_l are caused by the difference in the ion concentrations on each side of the membrane and by ion leakage. The drawback of this model is that it characterises a single spiking process and cannot model a continuous series of spikes.

The Fitzhugh-Nagumo Model: This model [Fitzhugh61, Linares91] is based on the Hodgkin-Huxley model but uses a non-linear function to represent the sodium current. This makes it computationally complex. The model can, however, represent the continuous spiking nature of the stellate chopper cell.

The MacGregor PTNRN10 Model: This model is a simplified version of the Hodgkin-Huxley model [MacGregor87]. It consists of 3 differential equations. These equations represent the cell membrane potential (E), the potassium conductance (G_K) and the potential threshold (Th). These equations can be seen below.

$$\frac{dE}{dt} = \frac{-E + V(t) + Gk(Ek - E)}{Tmem}$$

$$\frac{dGk}{dt} = \frac{-Gk + B(S)}{TGk}$$

$$\frac{dTh}{dt} = \frac{-(Th - Th(0)) + C(E)}{TTh}$$

$$S = \begin{cases} 0 & E < Th \\ 1 & E \geq Th \end{cases}$$

Initially the values of E and G_k are zero and the value of Th is Th_0 . The applied voltage, V , causes an increase in the value of E . When the value of E exceeds the value of Th , the potassium channel is opened. The introduction of potassium ions reduces the value of E until the potassium current decays. The decay of the potassium current allows the value of E to increase again and the cycle is repeated. When the value of E exceeds Th a spike is produced ($S=1$). This continuous cycle produces a constantly spiking output. This model is selected for our investigations because it is computationally efficient and corresponds to the internal chemistry of the real cell.

2.4.6 The Coincidence Cell

The coincidence cell has the job of detecting phase locking within a group of stellate chopper cells. The model used in the Meddis simulations is based on the MacGregor PTNRN10 Model (just like the stellate chopper cell). The parameters used to represent the coincidence cell are similar to those used for the stellate chopper cell.

2.5 ARITHMETIC METHODS

In this section we describe different numerical techniques that could be utilised to implement our neuromorphic system. We consider traditional deterministic methods such as bit-parallel and bit-serial techniques.

2.5.1 Standard Arithmetic Techniques

There are a number of techniques for multiplication, division, addition and subtraction depending on whether speed or the compactness of a design are the main requirements. These techniques also depend on whether data is flowing in bit-serial [Baldwin78, Denyer85] or bit-parallel format [].

The advantage of bit-serial arithmetic is that it provides built in pipelining with a compact design. The disadvantage is that bit-serial arithmetic units require a larger number of clock cycles to perform a calculation than a bit parallel design. For example, when multiplying two 16-bit numbers in bit-serial there is a 17 clock cycle latency before the result begins to emerge on top of the 32 clock cycles it takes for the result to be presented in full. In terms of complexity, however, this bit-serial design only requires 16 full adders and 60 flip-flops to store intermediate results (or partial products).

Bit-parallel implementations, however, can be pipelined to produce a result every clock cycle irrespective of the bit-width of the numbers it is multiplying. The disadvantage is that the complexity of such a multiplier is considerably more than that of an equivalent bit-serial multiplier. Such a multiplier requires 256 full adders and 512 flip-flops to store intermediate results (or partial products).

2.5.2 Summary

Within the conventional arithmetic field, we have specifically chosen bit-serial arithmetic because it provides us with designs that are considerably smaller than its bit-parallel counterpart. This makes the bit-serial technique more appropriate for the limited gate resources of our single FPGA implementation.

2.6 TECHNOLOGY

2.6.1 DSP Systems

DSPs are a type of processor specifically designed to tackle mathematically intensive problems. It has been shown that they can be applied to the field of auditory signal processing [Ahn90, Bogli90]. The pitch detection system is indeed mathematically intensive. In simple terms, DSPs are designed to carry out pairs of multiplication and

addition/subtraction calculations at a very high rate. This type of system is suitable for implementing models that solely comprise difference equations.

By using DSPs we can only judge the complexity of our implementation by the complexity of the micro-code used to program the DSP. We are more interested in the hardware complexity of our system. Furthermore, we are unable to represent the fine-grain architecture of neuromorphic systems by employing a DSP to implement the pitch detection system.

2.6.2 Programmable Logic

Programmable logic comes in the form of field programmable gate arrays (FPGAs) and complex programmable logic devices (CPLDs). The advantages of such programmable logic are:

- Rapid prototyping. This allows the implementation and verification of a design to occur in a short space of time.
- Reuse of resources. An FPGA or CPLD employing SRAM or Flash technology can be reprogrammed many times thus making more efficient use of resources.

The disadvantages of programmable logic are:

- Slower clock speeds. This is a result of limited available routing resources and the hardware overhead involved in implementing the programmability of each logic unit and routing resource.
- Less gate density. Again, this is a result of implementing the programmability within the device.

2.6.3 Non-Programmable Logic

Non-programmable logic (such as an ASIC) has the following advantages:

- High complexity designs can be realised on a single chip.
- High speeds (of the order of 100 MHz) can be realised.

The disadvantages of non-programmable technology are:

- Slow production turnaround time. This is especially true for low volume production.
- Design alterations and corrections mean a further long period of production turnaround.

2.6.4 Summary

Considering the fact that time and flexibility are critical, we will choose programmable logic for our hardware implementations. In particular we have chosen the FPGA technology offered by Gatefield Corporation.

2.7 SUMMARY

This chapter has reviewed the current state of auditory modelling together with technological issues that influence the hardware implementation of such auditory models. The conclusions drawn are as follows:

- The Hewitt and Meddis pitch detection model demonstrates the greatest correspondence to known biological structures as well exhibiting the most accurate behaviour with respect to biological measurements.
- The simplest and most accurate models were identified to represent the biological components within the pitch detection system.
- Bit-serial arithmetic was selected as the most appropriate to perform the calculations within each system component as it provides inherent pipelining and compact designs.
- Programmable technology was identified as the most appropriate because it provides rapid prototyping times as well as reusability.

CHAPTER THREE

OUTLINE OF INVESTIGATIONS

3.1 OBJECTIVES OF CHAPTER

The objectives of this chapter are to select and introduce the experimental investigations to be carried out within the context of this thesis. More specifically the objectives are to:

- From the knowledge gained in previous two chapters identify areas of neuromorphic research that remain unexplored.
- State the research topics by identifying which of these unexplored areas we wish to expand upon.
- Describe the experimental vehicles used to carry out these experiments.
- Clarify the experimental assumptions made before any investigations are undertaken.
- Introduce the investigations by giving a detailed description of the experiments to be carried out.

3.2 IDENTIFICATION OF RESEARCH TOPICS

Our research focuses on the electronic implementation of a neuromorphic system. From the knowledge gained from the previous two chapters we can identify a number fields of interest which remain unexplored. These fields of interest are:

- Auditory processing: we have identified this field as providing interesting signal processing problems, having a simple 1 dimensional input, and whose input is easy to control.
- Pitch detection: we have identified that this process as a useful within speech processing systems and have identified an appropriate model.
- Digital technology: we have identified that digital technology provides a flexible medium for implementing the system and that it has benefits over analogue technology.
- Biological cell models: we have identified that although suitable models exist, there is a need to develop them in order to suit them to digital technology.

3.3 STATEMENT OF RESEARCH TOPICS

This thesis aims to explore the digital implementation of a neuromorphic pitch detection system. More specifically, it aims to implement each of the processing units that make up the system in hardware. Also, it seeks to implement a large proportion of the pitch detection system in digital hardware. Initially, four investigations have been proposed to address these aims and they are discussed below.

The first experiment is an investigation into the hardware implementation of the inner hair cell. This is a key component in the operation of all processes performed by the auditory system. This experiment also represents the first attempt to implement this cell in digital hardware.

The second experiment investigates the digital hardware implementation of the auditory nerve and the coincidence cell. The models that currently exist for these components are highly complex and inappropriate for hardware implementation. It is for this reason that we see a need to develop simpler algorithms.

The most complex stage of the system, in terms of the number of cells, is the stellate cell array. The third experiment sets out to reduce the complexity of this stage while preserving the overall system behaviour.

The final experiment investigates the implementation, characterisation and utilisation of a large portion of the pitch detection system. Characterisation is achieved by measuring the response of the implementation to a number of stimuli. The design of software to interface with the hardware system allows utilisation the hardware implementation.

3.4 EXPERIMENTAL VEHICLES

This section outlines the environment in which the experiments are carried out. There are three aspects to the experiments. First of all there are the software simulation, then there is the hardware simulation and finally there is hardware verification.

3.4.1 Software Simulation

Our software simulations were carried out using the Labview graphical programming language. These simulations model the behaviour of the system but not the circuit timing. The advantages of using Labview for simulations are:

- It runs much faster than an equivalent VHDL simulation (by a factor of 100).
- Built-in facilities allow results to be displayed in a graphical format as simulations run.
- Once familiar with the graphical nature of the language it is much quicker to produce useful programs than in text-based high level languages.

The disadvantages are:

- Run-time speed slower than other high level languages such as C.

- Programs do not compile to stand-alone executables. This means that they are not portable to machines without Labview installed.

Taking into consideration these advantages and disadvantages Labview is still favourable to other languages because development speed and graphical capabilities are priorities over run-time speed.

3.4.2 Hardware Simulations

The hardware is designed using VHDL (VHSIC Hardware Description Language). We use Computer Aided Engineering (CAE) tools from Veribest to synthesise and verify our VHDL code. The behaviour of both the VHDL code and of the synthesised netlists are verified using test-bench software written in VHDL.

By using a VHDL based test-bench we can verify the behaviour of the both the VHDL code and the synthesised netlist using the same program and vector file. This allows consistency in the verification process.

3.4.3 Hardware Verification

Once we have verified the behaviour of the netlists we place, route, and download them onto an FPGA. To verify the behaviour of the design on the FPGA we use Labview again. We have written a test-bench that utilises a Data Acquisition (DAQ) card to apply test vectors to the hardware and measure its response. The hardware test-bench uses exactly the same stimulus file format for its test vectors as the VHDL test-bench. There are a number of advantages and disadvantages in using this test-bench to verify the hardware design. The advantages are:

- There is no need to generate a new test vector file for verifying the hardware.

- Using the same file ensures that each stage of the design flow is subject to the same vectors. This ensures that the behaviour of the design is preserved through the various design stages.

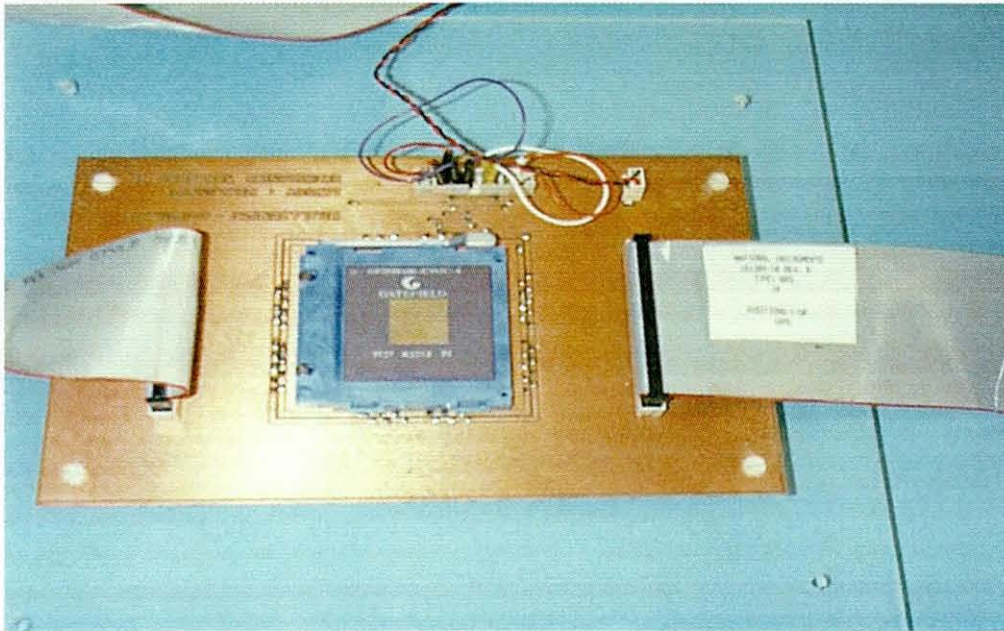


Figure 3-1: Test set up for Gatefield chip.

The disadvantage is:

- The test-bench runs too slowly to test real-time operation of the design.

The consistency of the test vectors throughout the design flow is of much more importance than the fact that we cannot verify the design in real time. The test set-up with connections to the DAQ card at the sides of the FPGA can be seen in Figure 3-1.

3.5 EXPERIMENTAL ASSUMPTIONS

In Chapter 2 we described the pitch detection system that we plan to implement. There is a number of simplifying assumptions that we make about the model (see Figure 3-1). First of all, we assume that the architecture comprises:

- 30 cochlear band-pass channels. The centre frequencies of these filters range from 80 Hz to 5 kHz in logarithmic steps.
- There is one hair cell for each cochlear filter.
- Each channel comprises 20 banks of stellate cells. The BMF of these banks range from 80 Hz to 300 Hz in logarithmic steps.
- Each stellate bank comprises 40 stellate cells.
- The auditory nerve in each channel produces 40 uncorrelated noisy signals instead of 800. This assumption is made because the noise applied to each stellate cell within a stellate bank needs to be uncorrelated. It is not necessary to have uncorrelated noise for each stellate cell within a channel. This is true because there is no cross talk between the stellate banks.
- Similarly, there are 40 dendrite filters instead of 800. The system requires 1 dendrite filter for each noisy AN signal.

The assumptions made about the number of cochlear channels, stellate banks, and stellate cell per bank are based on biological measurements. The assumptions made about the number of noisy AN signals and dendrite filters are intended to keep the connectivity of the system at a minimum.

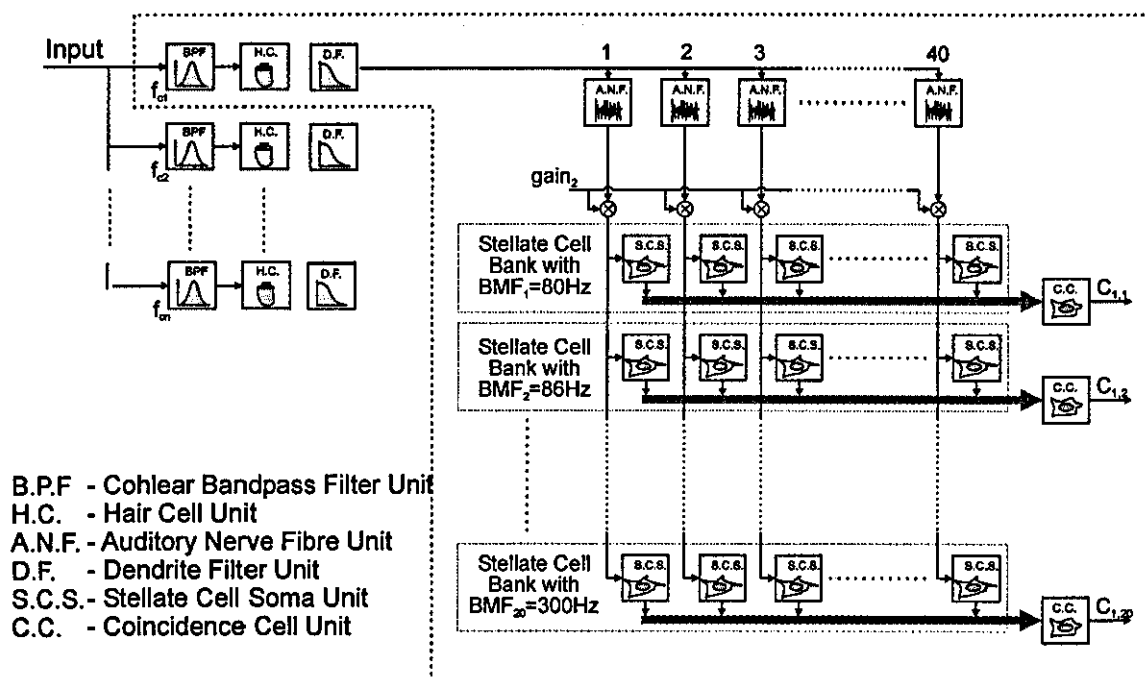


Figure 3-2: Assumed system architecture.

3.6 INTRODUCTION TO INVESTIGATIONS

The first experiment is designed to find out the most efficient way of implementing the inner hair cell in real-time digital hardware. There are three aspects of this hardware design that we address. Firstly, we look at the relation between the accuracy of the model and the bit-width of the internal signals. We are looking to find the point at which the expected behaviour of the cell brakes down. The second part of the experiment addresses the complexity of the hair cell's saturation function. The original function employs a divider that makes inefficient use of the hardware resources. Consequently, we explore alternative methods of implementing the saturation function. The final part explores pipelining the hair cell. Pipelining would allow a single hair cell implementation to represent all 30 hair cells in the system.

In our second experiment we investigate alternative models of the auditory nerve and coincidence cell. In the case of the auditory nerve, current models use probability

functions together with random number generation to determine the spiking patterns of individual auditory nerve fibres. We set out to model the overall effect of the auditory nerve, which is to add noise to output signal of the hair cell. In the case of the coincidence cell, current models use standard ion flow neural cell models. We believe that such a model is too complex for the role that it plays. As such, we compare the performance of this model against one that is more suited to its task and to a digital hardware implementation.

The third experiment is concerned with the stellate cell array. The stellate cell array is by far the most complex stage of the pitch detection system simply because it comprises 800 cells. It is situated in the cochlear nucleus within the auditory brainstem. The stellate array receives signals from the auditory nerve and produce a phase locked spiking response from some of the cells when an amplitude modulated signal is presented to it. For this reason we attempt to simplify our representation of this cell array. The main goal is to reduce the sample rate of the stellate cell somas thereby reducing the number of processing units required to represent them. Before doing so, we need to ensure that the dendrites feeding the somas attenuate the high frequency components of the input signal sufficiently such that these components do not interfere with the behaviour of the somas.

The final experiment investigates the implementation and utilisation of a portion of the pitch detection system. Firstly, we have to identify a portion of the system that exhibits some useful functionality while being compact enough to fit within an FPGA. Also, by designing software we can exploit this functionality we aim to develop a demonstrable system. Furthermore, this software can apply representative signals to the hardware such that its response to typical stimuli can be measured.

CHAPTER FOUR

HAIR CELL IMPLEMENTATION

4.1 OBJECTIVES OF CHAPTER

The object of this chapter is to develop an efficient digital hardware implementation of the mammalian inner hair cell. More specifically our aims are to:

- Investigate the relationship between the bit width of the internal signals and the accuracy of the Meddis hair cell model,
- Identify an efficient method of implementing the model's saturation function, and
- Explore how pipelining can be incorporated to allow a single processing unit to represent a number of inner hair cells.

4.2 INTRODUCTION

The inner hair cell plays a crucial role in the operation of the auditory system: it is responsible for converting the mechanical vibrations in the cochlea into electrical impulses in the auditory nerve. The output of the inner hair cell is a transmitter substance which synapses with the auditory nerve thereby producing spikes which are then transmitted down the auditory nerve fibres. It is the auditory nerve that drives neurons in the auditory brainstem. As a consequence of this importance this cell plays a vital role in the function of the pitch detection system. Stage 2 of the system comprises a bank of inner hair cells. The reasons for selecting the Meddis model to represent the inner hair cell are discussed in Chapter 2. In brief, the Meddis model encapsulates behavioural accuracy (in comparison to biological measurements) and

computational simplicity. Furthermore, it represents the chemical reservoirs and processes that are responsible for the inner hair cell's behaviour.

4.2.1 Chemical Processes Within the Model

The behavioural features are characteristics of the chemical processes that occur within the inner hair cell. The chemical involved in these processes is called transmitter substance and it exists in three reservoirs. These reservoirs are the *free transmitter pool*, the *cleft*, and the *reprocessing store*. The transmitter can flow between these reservoirs as described in Figure 4-1.

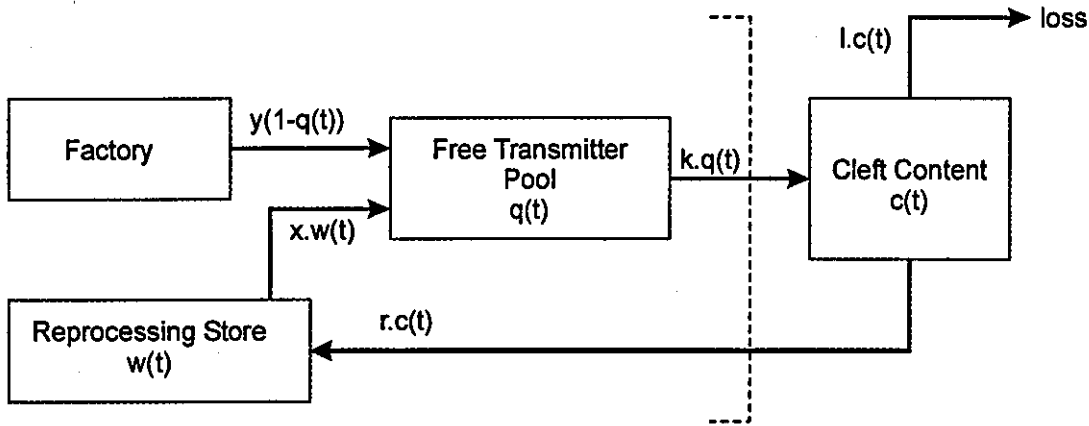


Figure 4-1: Block diagram of inner hair cell design

The variables q , c , and w describe the contents of the free transmitter pool, the cleft and the reprocessing store and respectively. The behaviour of the model is described by the following differential equations:

$$\begin{aligned}\frac{dq}{dt} &= y[m - q] - kq + xw \\ \frac{dc}{dt} &= kq - lc - rc \\ \frac{dw}{dt} &= rc - xw\end{aligned}$$

In a digital system it is useful to consider these in difference equation format. Assuming that dt is the length of a sample time and that t is the current sample number we get the following difference equations:

$$q(t+1) = q(t) + \{y[m - q(t)] - kq(t) + xw(t)\}dt$$

$$c(t+1) = c(t) + \{kq(t) - lc(t) - rc(c)\}dt$$

$$w(t+1) = w(t) + \{rc(t) - xw(t)\}dt$$

The values of q , c and w range between 0 and 1 where m (maximum value) is a constant of value 1. The values of the parameters in these equations are described in Table 4-1. There is another parameter, h , which we will discuss in Chapter 5.

Parameter	Description	Units	Value
dt	Sample time	ms	50.00
A	Permeability constant	s^{-1}	240.00
B	Permeability constant	s^{-1}	5000.00
x	Reprocessing rate	s^{-1}	170.00
y	Replenishment rate	s^{-1}	5.05
l	Loss rate	s^{-1}	1650.00
r	Recovery rate	s^{-1}	8500.00
g	Release rate	s^{-1}	500.00

Table 4-1: Parameter values of hair cell model.

All of the above parameters are constants. However, we can see from Table 4-1 that we have not defined the variable k . This variable is defined as follows:

$$k = \frac{A + stim}{A + B + stim} gdt, \quad A + stim \geq 0 \quad \dots(7)$$

$$k = 0, \quad A + stim < 0 \quad \dots(8)$$

These equations define how the input stimulus intensity affects the release of transmitter from the free transmitter pool into the cleft.

4.2.2 Model Behaviour

The inner hair cell has a number of behavioural features, which are important to its role in the auditory system. These features are half-wave rectification, saturation, and adaptation. The rectification process is self-explanatory.

The saturation function is a characteristic of the non-linear response of the membrane permeability between the free transmitter pool and the cleft. The overall effect is to limit the input signal dynamic range (input intensity range) before it is presented to the core function of the inner hair cell. Figure 4-2 shows the relationship between the input intensity and the value of k .

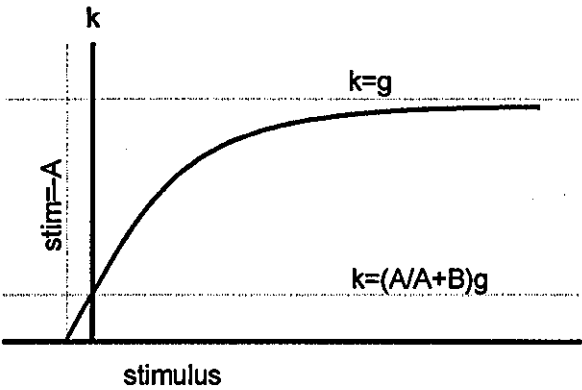


Figure 4-2: Saturation function.

Adaptation is a process similar to automatic gain control. When we increase the input stimulus level there is an initial increase in the amount of transmitter ejected into the synaptic cleft from the free transmitter pool. Soon afterwards, however, there is depletion in the amount of the transmitter substance left in the free transmitter pool. This results in there being insufficient transmitter in the free pool to maintain the ejection rate. The effect is that the amount of transmitter in the synaptic cleft decays down to the value it had before we increased the stimulus level. The response of the inner hair cell to an amplitude-modulated tone can be seen in Figure 4-3.

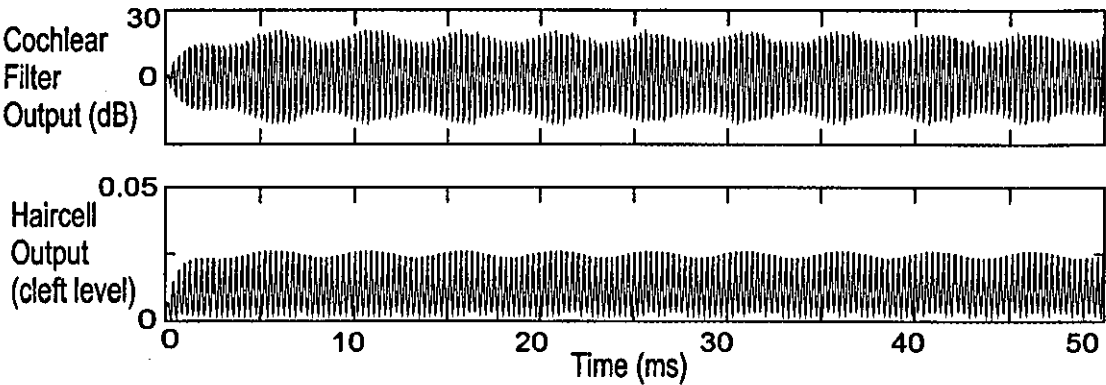


Figure 4-3: Response of inner hair cell to amplitude modulated tone.

In the following sections we will investigate the real-time digital electronic implementation of the Meddis inner hair cell model. This particular model has been selected because it possesses the above characteristics and is computationally efficient. We will also identify aspects of the model that can be altered in order to reduce the hardware complexity of this implementation. In short we must optimise the model for digital implementation while maintaining the cell behaviour.

4.3 SIGNAL BIT WIDTH

On analysis of the difference equations in the previous section, we can see that the flow rate parameters need to be multiplied by dt . Both dt and the flow parameters remain constant therefore we chose to provide the model with the parameters $x.dt$, $y.dt$, and so on. As a first approximation we implement the hair cell using 16 bits to represent the internal signals. The signals and their values can be seen in Table 4-2. We can see from the table that for values A and B are fixed point integers whereas the other parameters have a fixed point situated between bit 16 (the most significant bit) and bit 15.

Signal name	Number of fractional bits	Decimal value	16 bit binary value
A	1	240.0000	0000000011110000
B	1	5000.0000	0001001110001000
x.dt	15	0.0085	0000000100010111
y.dt	15	0.0002	0000000000001000
l.dt	15	0.0825	0000101010001111
r.dt	15	0.4250	0011011001100110
g.dt	15	0.0250	0000001100110011

Table 4-2: Values of hair cell parameters.

We can see that the $y.dt$ parameter has only one significant bit representing its value. This may or may not present accuracy problems. In case accuracy is a problem we will simulate the hair cell using wider bit widths than 16.

4.3.1 Measurement of Error

The bit-width used to represent the internal signals of the model affects both the accuracy and the complexity of the hardware implementation. In this section we shall look at the behavioural effect of varying the bit-width before looking at the effect it has on the implementation complexity. We will compare the accuracy of the bit-limited models with an extended-precision floating-point software implementation of the model. We calculate the error in terms of the cleft content. It is the cleft content that is the output of the hair cell. This error is calculated as a mean percentage difference over a 500 ms simulation time. We will also set the maximum allowable mean error to be 10% because it is believed that the noise added to the output of the inner hair cell by the auditory nerve is in excess of 10%. This means that there is no gain in making the model any more accurate.

4.3.2 Error Results

In this section we will apply an amplitude-modulated signal to the models. This signal is representative of a complex signal that has already passed through the cochlear filter. In this case the carrier frequency is 3350 Hz and the modulation frequency is 198 Hz. The modulation index is 25 %. The signal is also presented at a wide range of intensities ranging from 0 dB to 90 dB. We have chosen this wide intensity range to identify any possible underflow or overflow problems that might exist.

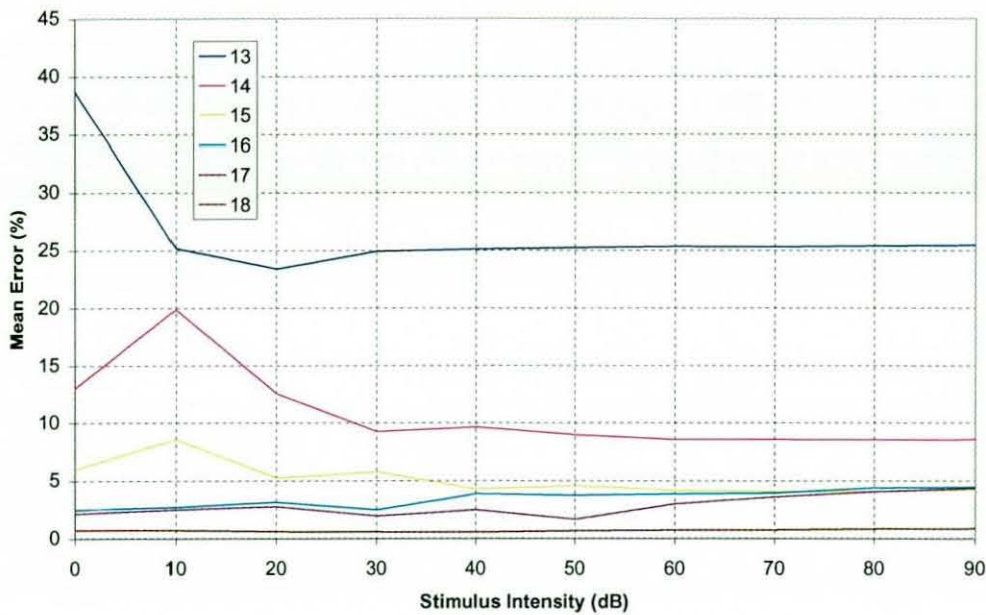


Figure 4-4: Graph of percentage error for bit widths 13 to 18.

Initial experiments show that below 13-bit representation we experience very large errors (over 50%) mainly caused by 100% quantisation error in the y_{dt} parameter. A way of avoiding this problem is for each parameter to have its decimal point in a different place. This allows for more accurate calculation of internal signals. Unfortunately, if these varying representations are used then problems occur when calculating of the new values of q , c , and w . Extra control logic would be required to align decimal points for subtraction and addition operations. For this reason a consistent decimal point position is maintained.

We can implement the inner hair cell model using 15 bits instead of the original 16-bit model and still maintain an accuracy of over 90 %. This reduction in bit-width results in a reduction in complexity of the hardware implementation. We do not believe, however, that the advantages gained from reducing the bit width from 16 to 15 are sufficient to warrant an alteration in the model.

4.3.3 Simplified Parameters

The exact values produced by the inner hair cell model are not as important as its general behaviour. This is mainly due to the fact that the output of the model is subjected to a relatively high amount of noise added by the auditory nerve. For this reason we hope to use simplifications of the constant parameters of the hair cell model in order to reduce the complexity of the design. The way we plan to do this is to replace each bit-serial multiplier with relatively few shift-and-add operations. The total complexity of one or two bit-serial adders and a shift registers will prove far less complex than a bit-serial multiplier. This hardware simplification is achieved automatically by an optimisation tool that recognises multiplication by a constant and removes the unnecessary parts of the multiplier hardware.

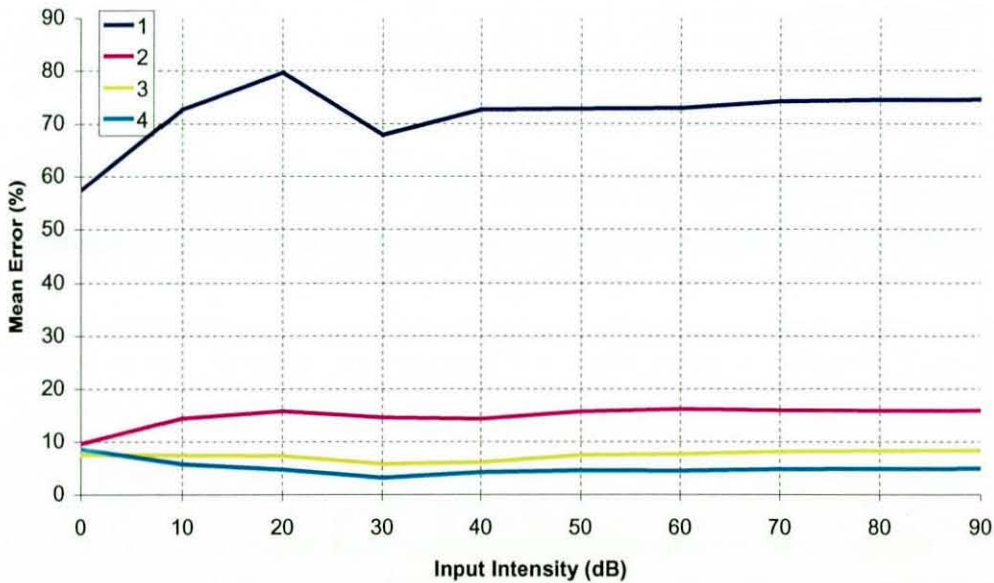


Figure 4-5: Error using simplified multiplication against number of significant bits used in multiplication.

We can see from Figure 4-5 that using the 3 most significant bits of the parameters, which only requires 2 shift-and-add operations, we can keep the mean error below 10%.

4.4 HARDWARE IMPLEMENTATIONS

We have developed two different digital electronic hardware implementations of the inner hair cell model. Before describing them we will discuss the architecture on which they will both be based. The top level of the architecture comprises four blocks. These are REG_SET, CALC_FLOW and CALC_RES. These blocks can be seen in Figure 4-6. The REG_SET block is the register block in which we store the previous values of q , c and w . The CALC_FLOW block calculates the amount of transmitter that flows between the reservoirs and the amount of transmitter lost. Finally, CALC_RES calculates the new reservoir levels. The arithmetic units that comprise the model can also be seen in Figure 4-6.

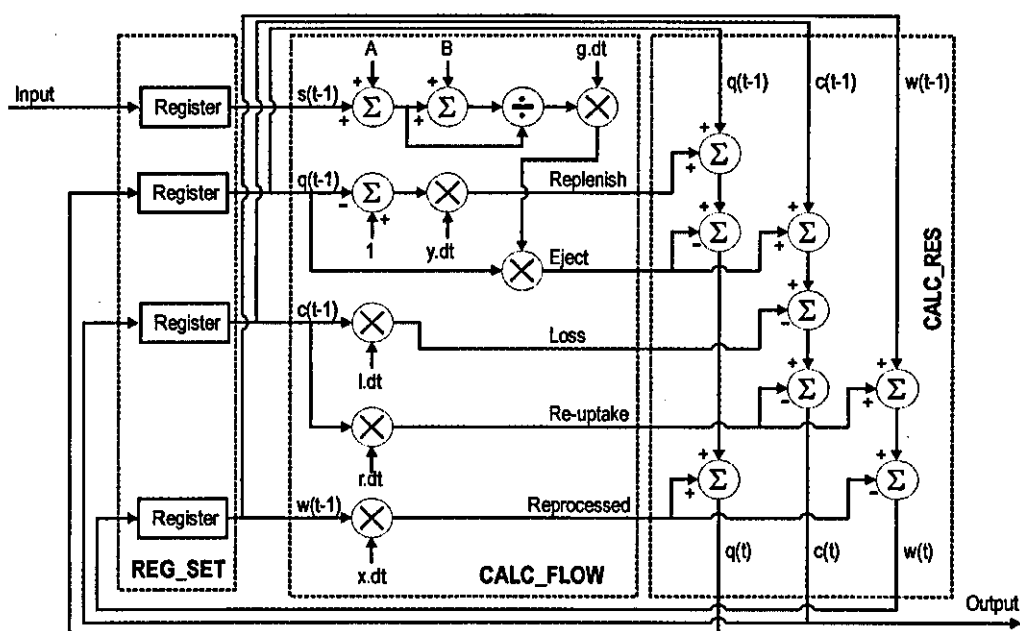


Figure 4-6: Block diagram of design including arithmetic units.

4.4.1 Bit-Parallel Implementation

The bit-parallel version is the first-ever implementation of the inner hair cell in digital electronics. For this reason we decided to initially use the simplest possible method of

implementation to allow a rapid design time. We chose to use bit-parallel arithmetic techniques specifically using ‘ripple-carry’ parallel adders, and ‘shift-and-add’ parallel multipliers. We also designed a ‘conditional shift-and-subtract’ divider circuit. The divider design process was simplified by the fact that the numerator and denominator are both positive and that the numerator is never greater than the denominator. This means that the output from the divider is always a positive fractional number. The architecture of the divider can be seen in Figure 4-7.

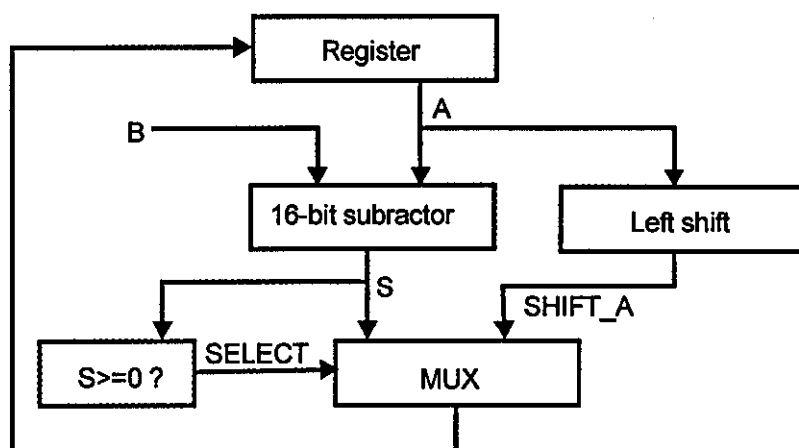


Figure 4-7: Block diagram of bit parallel division unit for $A \div B$.

Figure 4-6 shows the architecture of this design in more detail. This figure shows how the arithmetic components link together within the context of the design top-level blocks.

Block	Complexity (Gates)	Speed (MHz)	Cycles per epoch
REG_SET	758	55.37	2
CALC_FLOW	6167	7.52	35
CALC_RES	2275	12.85	5
HAIR CELL	9200	7.52	42

Table 4-3: Complexity and speed of bit parallel hair cell implementation

Table 4-3 shows the complexity and speed of the bit-parallel implementation of the hair cell. We can see that the implementation meets the speed requirement for the cell to run in real-time.

Epoch frequency = clock frequency / cycles per epoch

$7.52 \text{ MHz} / 42 = 179.05 \text{ kHz}$ (over 8 times faster than the required 20 kHz)

The critical path for this circuit is the parallel addition. If we had used more sophisticated addition techniques (such as carry-look-ahead) then the speed of the design could have been increased further. The emphasis at this stage, however, is to develop a simple implementation, therefore, such techniques are not pursued.

4.4.2 Bit-Serial Implementation

Bit-serial arithmetic provides two advantages over the previous parallel implementation. First of all, bit-serial techniques are inherently pipelined thus reducing the delay in the critical path. This presents the opportunity for us to increase the clock speed of the design. Furthermore, the potential increase in clock speed comes at no cost in terms of hardware complexity. The block diagram for this implementation is the same as for the bit-parallel implementation (Figure 4-4).

Apart from the divider, which we designed ourselves, all of the arithmetic units used in this implementation were designs from the literature. Normally division in bit-serial is too complex to consider but the special circumstances in this design, described above, ensure that we can simplify the design task. The structure of our bit-serial divider design can be seen in Figure 4-8.

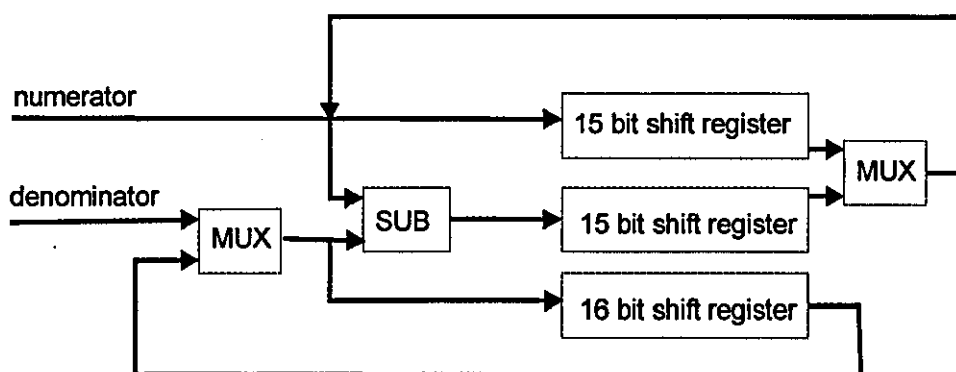


Figure 4-8: Structure of the bit-serial divider design.

As the numerator and the denominator enter the divider they pass through a bit serial subtractor. This subtracts the denominator from the numerator to produce a result. The three values (numerator, denominator, and subtraction result) are stored in constantly rotating shift registers. The length of the denominator shift register is 16 bits whereas the other two registers are 15 bits. If the value of the subtraction is negative then the denominator is subtracted from the numerator again. If the result of the subtraction is positive then the denominator is subtracted from the previous subtraction result. The difference in the length of the registers causes the numerator and subtraction result to be multiplied by 2 after each iteration. Every time a positive result is returned from the subtractor a single bit of division result is set. The divider calculates the result in the order MSB down to the LSB.

	Complexity (Gates)	Speed (MHz)	Cycles per epoch
REG_SET	941	41.70	(runs in parallel) 0
CALC_FLOW	5606	34.38	325
CALC_RES	1516	53.19	16
HAIR CELL	8063	34.38	341

Table 4-4: Complexity and speed of bit serial hair cell implementation.

Table 4-4 shows the complexity of the bit-serial implementation of the hair cell. We can see that for a similar complexity we have an implementation that runs with a much higher clock speed. The disadvantage is that the number of clock cycles required for a single iteration is very high. The only way to solve this problem is to implement a simplification of the saturation function, thereby removing the need for the time-intensive bit serial divider.

4.5 SATURATION FUNCTION

In the previous section we evolved the hardware implementation from bit-parallel arithmetic to bit-serial. The motivation of this was to increase the clock speed. Although we have successfully achieved this aim, we now have much longer latency (number of clock cycles) for each calculation. The large majority of this latency is due to the division operation within the saturation function. As a result we set out in this

section to assess alternative methods for implementing the saturation function. The two approximate methods that we now investigate are a look-up table and a piece-wise linear approximation.

4.5.1 Look-up Table

In this section we try using look-up tables with varying numbers of values to implement the saturation function. We endeavour to find the minimum number of values required for the cell's function to be maintained. First of all we need to find a way of mapping the input to the saturation function to the addresses of the output values. In this investigation we will map an approximate logarithmic function. For this reason we use the position of the MSB of the input value as the MSBs of the look-up table address. The remaining bits of the look-up table address correspond to the bits that follow the MSB of the input value.

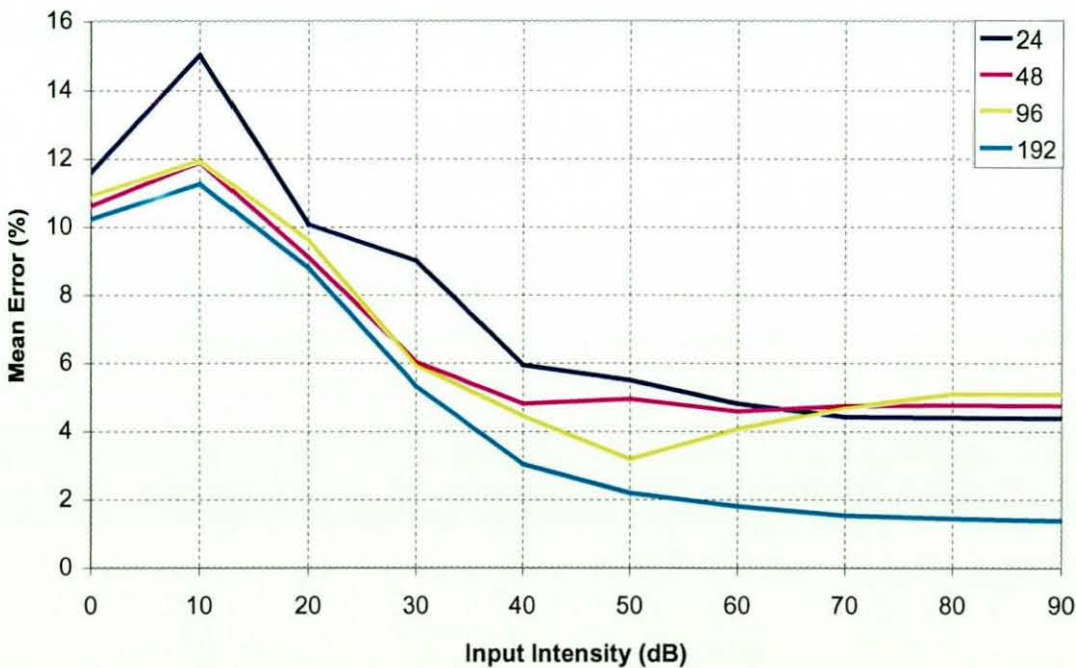


Figure 4-9: Accuracy of look-up table implementation.

We will be testing the behaviour of look-up tables with 24, 48, 96 and 192 entries. In this case there are 24 possible positions of the inputs MSB which gives us 24 look-up

values in the simplest case. This is doubled for each following input significant input bit that we use to address the look-up table.

We can see from Figure 4-9, that the behaviour of the hair cell is still outside the acceptable error limits when we use a look-up table with 198 entries.

4.5.2 Piece-wise Linear Approximation

The second method is a piece-wise linear approximation for the saturation function. Again we have varied the number of linear sections in the curve to find the optimal solution. Again, we have matched the joins between the linear sections at logarithmic intervals on the x-axis.

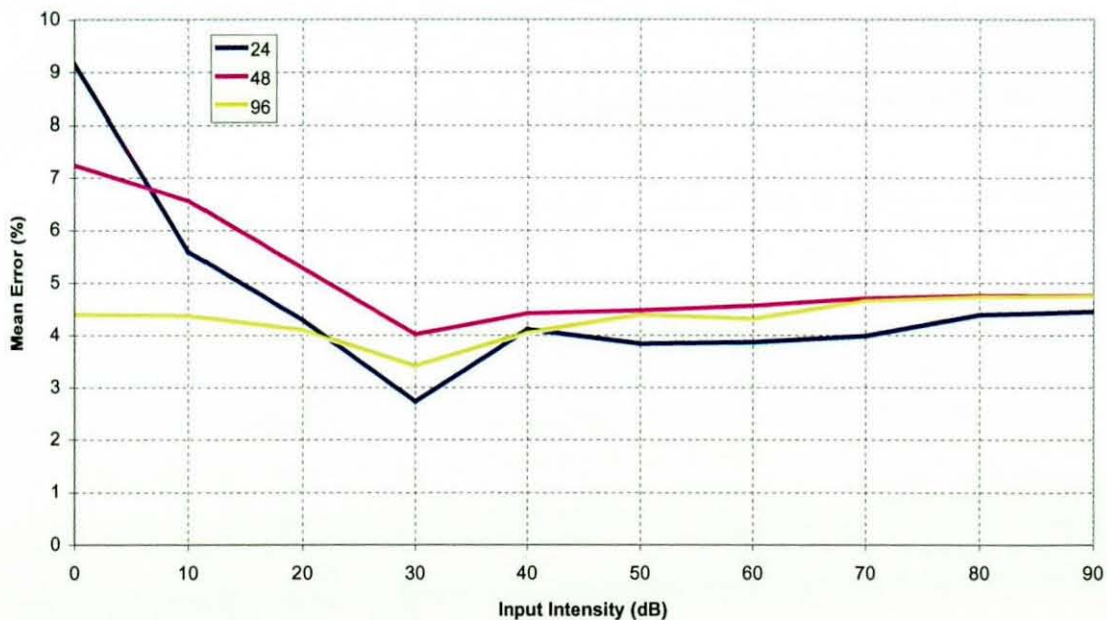


Figure 4-10: Accuracy of linear piece-wise implementation.

In Figure 4-10 we can see that with only 24 sections the approximation is more than adequate to maintain the behaviour of the inner hair cell model.

4.5.3 Saturation Function Conclusions

It is clear that the linear piecewise approximation of the saturation function is less mathematically complex than the original divider based algorithm while maintaining an acceptable level of accuracy in the overall inner hair cell design. This makes it suitable for hardware implementation.

Block	Complexity (Gates)	Speed (MHz)	Cycles per epoch
REG_SET	941	41.70	(runs in parallel) 0
CALC_KT	2006	34.38	69
CALC_RES	4563	53.19	16
HAIR_CELL	7020	34.38	85

Table 4-5: Hardware details of simplified hair cell.

Table 4-5 shows the reduced complexity and number of clock cycles required by the piecewise linear approximation implementation of the saturation function.

4.6 PIPELINING THE HAIR CELL

Our implementation of the pitch detection system consists of 30 cochlear channels. Each channel starts with a cochlear filter followed by an inner hair cell. In order to reduce the complexity of the system we plan to pipeline as many processing units as possible. In this section we will investigate using a single processing unit to represent all 30 inner hair cells. For this to be possible we need each of the top-level blocks in the cell design to take 48 clock cycles between the start of consecutive inputs entering the block. The 48 clock cycles are required because the saturation function that we employ require 24 clock cycles with no data between the 24 clock cycles with data in order to operate properly. This, together with a required sample frequency of 20 kHz, means that we require a clock speed of 28.8 MHz. With this pipelining in place and our system clock frequency of 34.8 MHz, one hair cell processing unit can represent the 30 inner hair cells.

4.7 CONCLUSION

The objective of this chapter was to develop an efficient digital hardware implementation of the mammalian inner hair cell. The first step was to investigate the relationship between the bit width of the internal signals and the model's accuracy. Initially, it was discovered that a bit width of 15 was the lowest allowed while maintaining a mean error of less than 10%. We decided, however, to use a 16 bit representation. We then determined that the number of significant bits required by the parameter values to keep the error within the 10% limit was 3.

After developing a bit serial implementation of the cell it became clear that the saturation function was the bottleneck in the design. As a result we investigated the use of a lookup table and a linear piecewise approximation of the saturation function. We determined that the piecewise linear approximation provided a simple solution while maintaining the accuracy of the model.

The final stage of development was structuring the top-level blocks such that pipelining could be implemented. This achievement gives us the possibility of implementing all the inner hair cells in the pitch detection system using a single processing unit.

CHAPTER FIVE

AUDITORY NERVE AND COINCIDENCE CELL MODELS

5.1 OBJECTIVES OF CHAPTER

The objective of this chapter is to investigate alternative ways of implementing the auditory nerve and coincidence cell. More specifically, the objectives of this chapter are to:

- Develop an auditory nerve model that produces a more efficient digital hardware implementation than current models.
- Develop a coincidence cell model that produces a more efficient digital hardware implementation than current models.
- Measure the performance of the system using these alternative models against the system using the original models.

5.2 INTRODUCTION

5.2.1 Auditory Nerve Models

The hair cells in the cochlea drive the auditory nerve (AN). The AN is made up of a collection of nerve fibres. Each fibre transmits spikes (brief and sudden changes in voltage potential) to the neural centres of the auditory brainstem. Each AN fibre has a

maximum spike rate of approximately 1 kHz. In order to transmit signals of several kHz we require several AN fibres to transmit the signals to their destination without loss of information. It has been shown that approximately 60 AN fibres connect hair cell output signals to a single stellate cell.

The first model describes the spiking patterns within the auditory nerve by utilising a binomial probability function. In Table 5-1 we can see an example of this algorithm running. The input to the algorithm is the probability output from a hair cell and therefore has a range between 0 and 1. We also generate 60 random numbers in the range 0 to 1. The number of random numbers in this group less than the hair cell probability value is the number of auditory nerve fibres spiking at that instant. If this process were deterministic then the total spiking nerve fibres would be 36 out of 60 to represent the probability of 0.6.

Fibre Number	Random Value (R)	Probability (p)	R < p ? (1=Yes, 0=No)
1	0.636	0.6	1
2	0.243	0.6	0
3	0.175	0.6	0
4	0.956	0.6	1
5	0.732	0.6	1
...
60	0.346	0.6	0
Total spiking nerve fibres			32

Table 5-1: Auditory nerve using rejection method.

This process needs to be performed 40 times to provide the required 40 noisy signals for the stellate cell array. When the spiking values are transmitted to the stellate cell dendrite they have to be transduced back into a voltage value. Simply summing the number of spiking nerve fibres and multiplying by a suitable scaling constant achieves this.

The simplified model of the auditory nerve uses a far simpler algorithm. In this model there is no representation of actual nerve fibres. Instead there is a proportion of noise

added to the probability output from the hair cell. The flow diagram for this design can be seen in Figure 5-1.

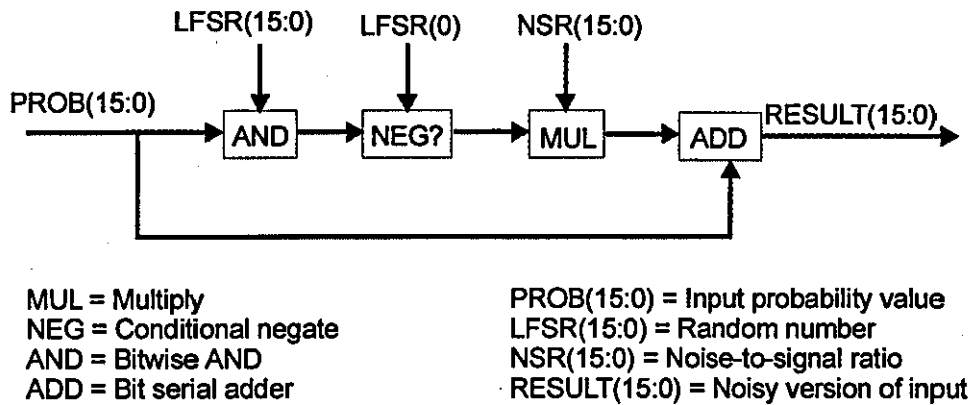


Figure 5-1: Diagram of the simplified auditory nerve model.

In this simplified model a pseudo-random number is ANDed with the input probability value. This process effectively scales the noise value to within the magnitude of the probability value. This positive noise value is then conditionally negated (depending arbitrarily on the least significant bit of the LFSR value) in order to give an equal likelihood of a positive or negative noise value. This scaled noise signal is then multiplied by a variable parameter in order to give desired overall ratio of noise intensity to signal strength.

5.2.2 Coincidence Cell Models

Coincidence cells are situated in the Inferior Colliculus and detect phase locking within the stellate chopper cell array. This cell has similar properties to a stellate cell soma. The output spikes from stellate group are summed and then scaled to give the equivalent to a voltage signal. This voltage signal is fed into the input of the coincidence cell. When phase locking occurs, the coincidence cell receives brief, high intensity input potentials.

This high intensity signal drives the cell's membrane potential to exceed a threshold value, which results in an output spike being produced.

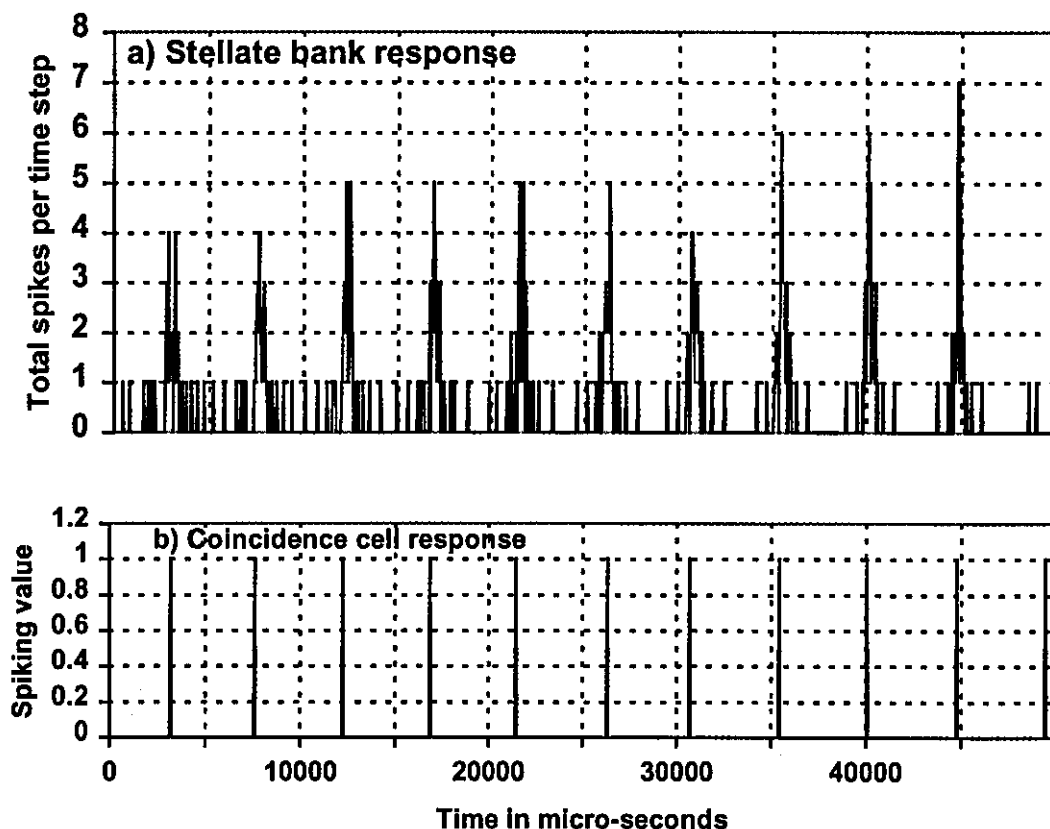


Figure 5-2: Graph of coincidence cell response.

Despite the complex electrochemical characteristics of this coincidence cell, it in effect performs a thresholding function. Such a function is simple to implement in digital hardware. For this reason we will investigate the behaviour of a simple thresholding algorithm in comparison to the existing biological model.

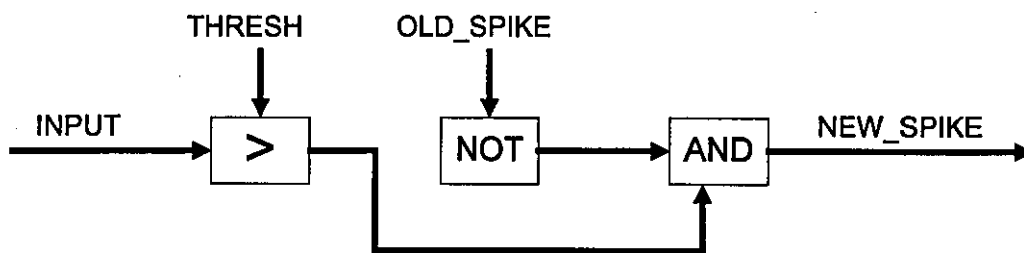


Figure 5-3: Diagram of simplified coincidence cell architecture.

We can see that there is a large difference in the behavioural complexity of the two coincidence cell designs. The real test of the simplified model is to see whether it produces results comparable to the original cell model.

5.3 BEHAVIOURAL EFFECT OF VARYING AN NOISE LEVEL

5.3.1 Sensitivity of Nerve Fibres

The variable *h.dt* (see Chapter 4) in the inner hair cell model converts the contents level of the cleft into a spiking probability value. This means that *h* effectively scales the function *c(t)* in the hair cell model. The higher this value, the higher the average spiking rate of the AN fibres. This value also affects the amount of noise added to the output from the hair cell by the AN. The variable *h.dt* has a value of 2.5 in the experiments carried out by Meddis. This value is derived from biological measurements and typically results in a maximum of 4 out of 60 nerve fibres spiking at any one time for steady state hair cell responses. It is easy to imagine that representing amplitude modulated signals using only 4 value levels results in an extremely inaccurate representation. We wish to identify which value of *h.dt* gives us the best results. Figure 5-5 shows a graph of value of *h.dt* against the coincidence cell results of the system.

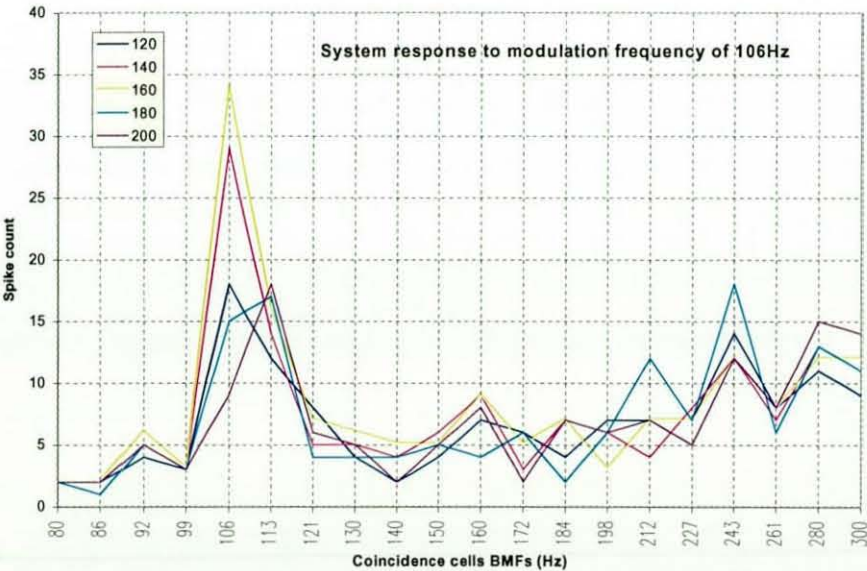


Figure 5-4: Graph of coincidence performance against *h.dt* value.

From the graph we can see that the h.dt value of 160 gives the best performance. We can explain this by the fact that for lower values of h.dt the amount of noise added to the hair cell response is much greater than for higher values of h.dt and thus the modulation information is destroyed. The ideal situation is for the maximum steady state response value of the hair cell to be represented by the full range of 60 firing nerve fibres. This occurs in Figure 5-4 when h.dt has a value of 160. When h.dt exceeds this value then there are not enough levels to represent the full range of steady state values and so the AN nerve fibres clip the hair cell response signal. This causes the behaviour of the system to break down.

5.3.2 Noise level of Simplified Model

We will now vary the noise level in our simplified auditory nerve model to investigate the effect this has on the performance of the system. It should be noted that the noise level represents the maximum noise-to-signal ratio at any one time and not the mean value.

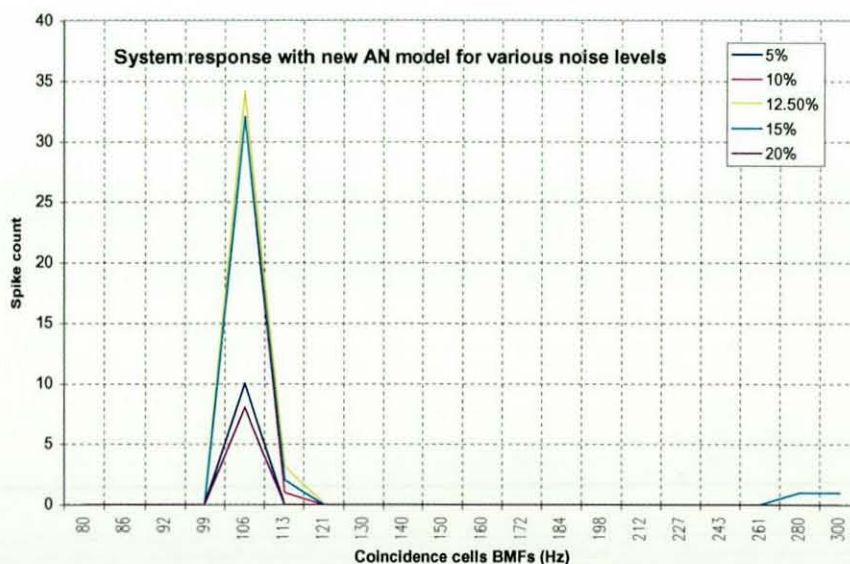
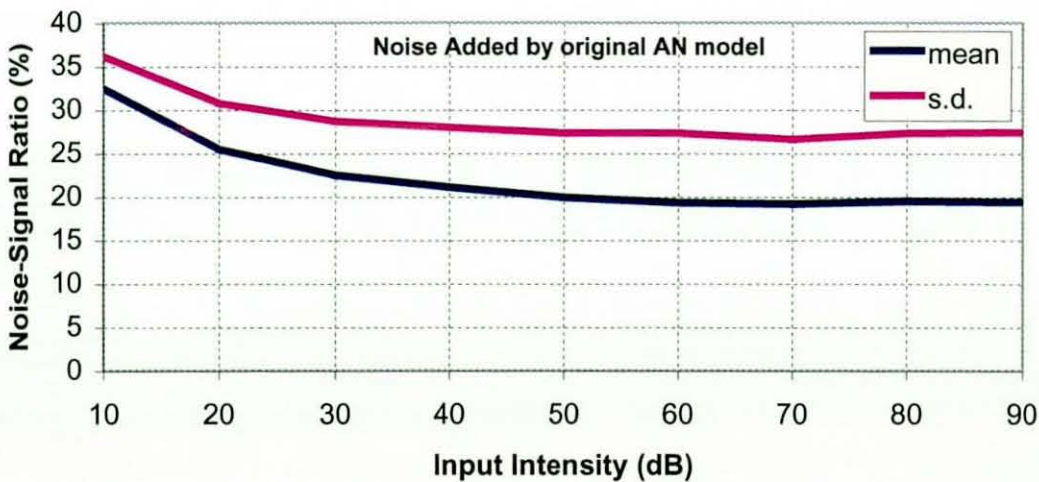


Figure 5-5: Graph of coincidence cell performance against noise level.

In Figure 5-5 we can see that the noise value of 12.5% gives the best performance of the system.

5.3.3 NSR of AN Models

We will now quantify the noise added to a typical signal by these auditory nerve models. This is achieved by calculating the mean noise-to-signal ratio (NSR) and the standard deviation (s.d.) over a simulation time of 500ms. This s.d. value is a measure of how widely spread the individual NSR values are. In Figure 5-6 the results are recorded for the original AN model using a value of h.dt of 160 and the new AN model using the noise level of 12.5%. These are the optimal noise levels for both models.



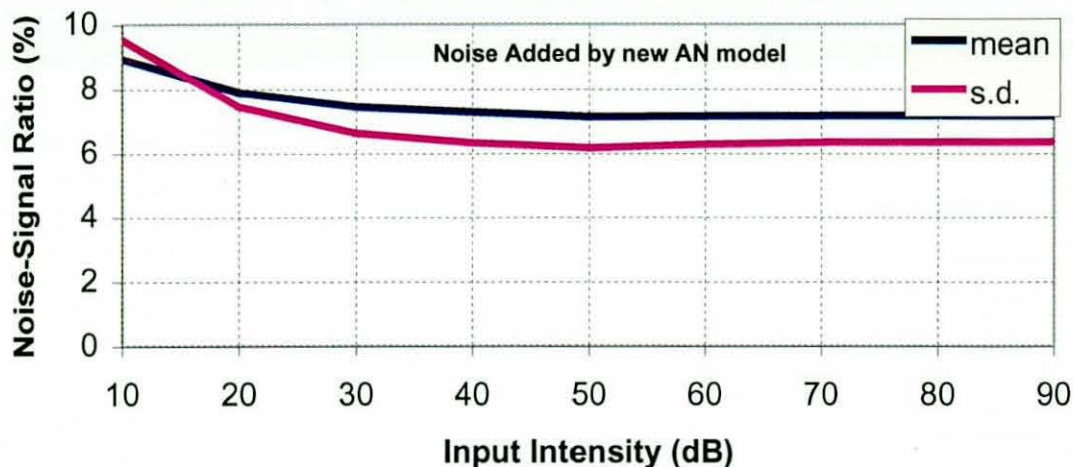


Figure 5-6: Graph of NSR against input intensity.

We can see that the amount of noise added by original model is greater than the noise added by our simplified model. All that remains is to be seen is which model gives a better system performance for a range of stimuli.

5.4 COMPARISON OF AUDITORY NERVE MODELS

5.4.1 Comparison of Behaviour for Auditory Nerve Fibre Models

When we identified the optimal noise level for the models we measured the performance of the system when applying a 3.3kHz tone amplitude modulated at 106Hz with an intensity of 30dB. We will now apply the same tone modulated at a range of modulation frequencies. This will provide a more thorough comparison of the two models.

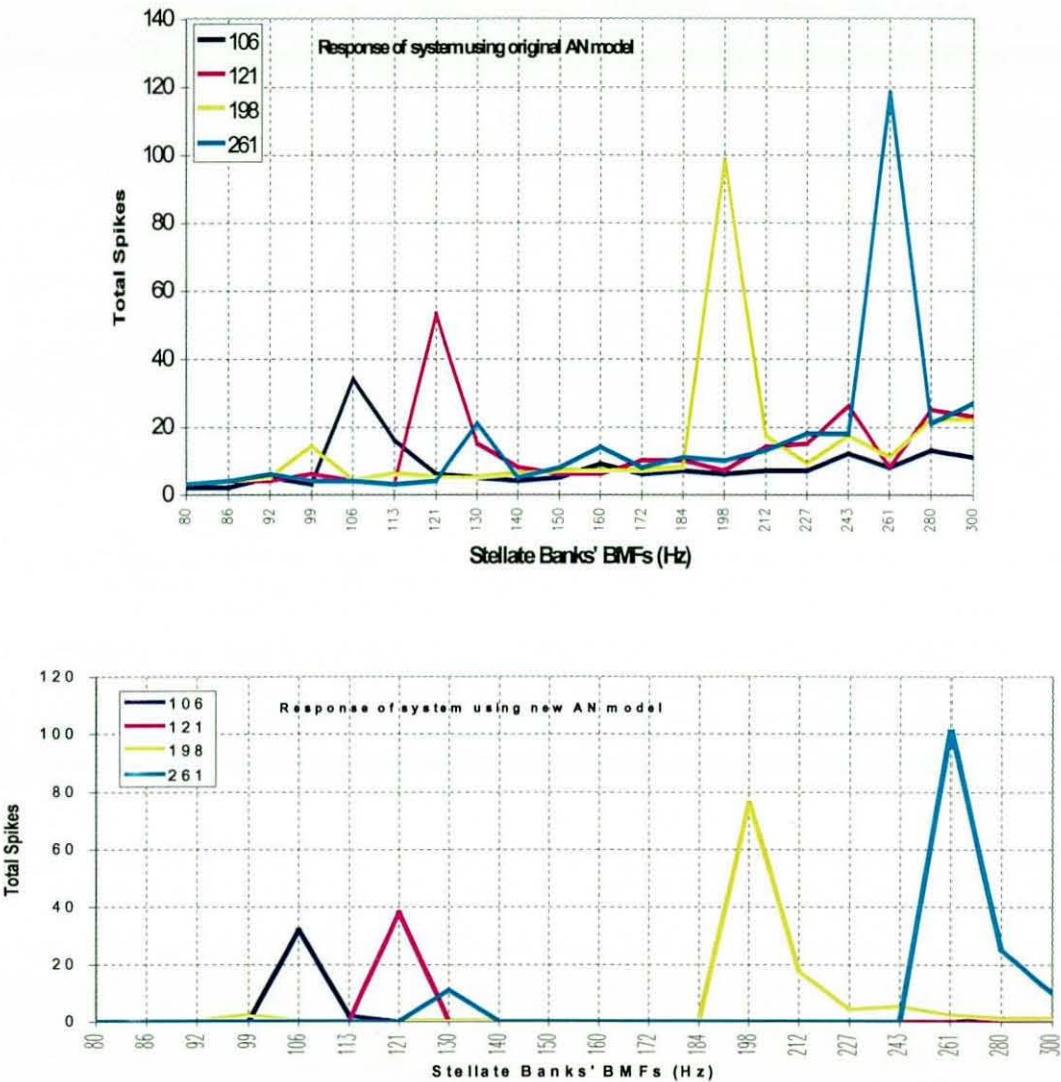


Figure 5-7: Coincidence activity for tone modulated at 106, 121, 198, and 261 Hz.

We can see the performance of both models in Figure 5-7. Both graphs clearly identify the modulation frequency of the applied tones although using the original AN model results in a slightly higher spike count for all coincidence cells for each stimuli. In general, though there is very little to chose between the behaviour of the two models.

5.4.2 Comparison of Complexity for AN Models

The complexity of each of the models can be seen in Table 5-2. We can see that the complexity of the simplified model is less than that of the original model. It should also be noted that we require 40 implementations of the original model every one of the simplified model.

Original	Complexity (Gates)	Speed (MHz)	Cycles per epoch
LFSR	191	53.45	40 x 15
COMPARE	1252	40.29	40 x 15
ACCUMULATE	417	46.15	40 x 2
SYNAPSE	1860	40.29	1280

Simplified	Complexity (Gates)	Speed (MHz)	Cycles per epoch
LFSR	191	53.45	40 x 15
SCALE_NOISE	1475	53.36	40 x 15
ADD_NOISE	1033	31.86	40 x 2
SYNAPSE	2699	31.86	1280

Table 5-2: Hardware complexity and speed of AN models.

The most complex part of each design is where the multiplier is used to scale the amount of noise added. In the case of the original model this is a multiplication of the response of the hair cell (cleft contents level) by h.dt. In the case of the new model the multiplier scales the noise value (LFSR value ANDed with probability value) by multiplying it with the noise level ratio.

5.5 VARIATIONS OF THE COINCIDENCE CELL MODEL

5.5.1 Original Model

When optimising the performance of each of the two models we have to ensure that we are modifying equivalent parameters. In the case of the coincidence cell, we vary the effective threshold values of the two models. In the original model we achieve this by varying the voltage associated with each input spike. This is effectively varying the input gain of the model.

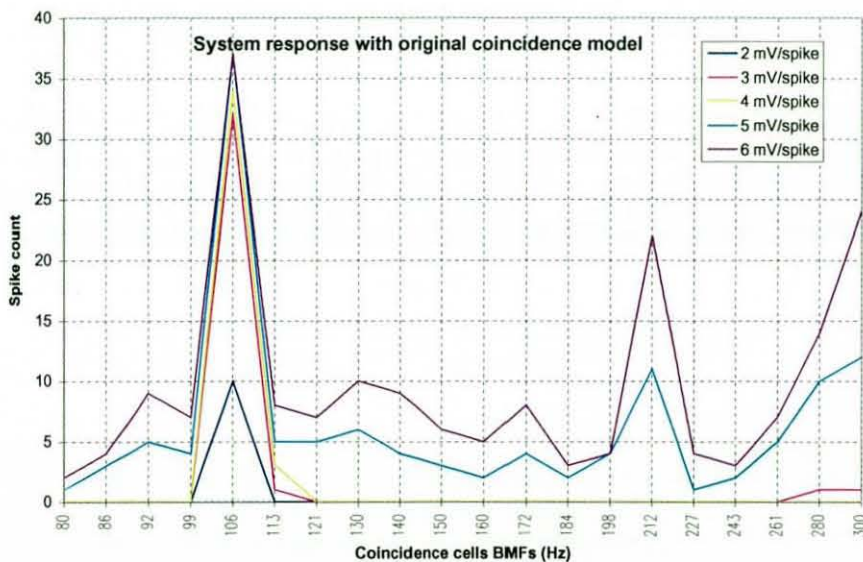


Figure 5-8: Original model response as a function of input gain.

In Figure 5-8 we can see the response of the coincidence cells as a function of input gain. In the above figure we used the same tone modulated at 106 Hz as in the previous section. We can see here, however, that we have some spiking activity at coincidence cells sensitive to 212 Hz. This can be explained by these cells responding to sub-harmonics of their BMFs. This gain is measured as volts per spike. We can see that the optimal input gain is 4mV per spike. Below this value the coincidence cell is not sensitive

enough to the phase locking that occurs in the stellate chopper cell banks. Above this value the coincidence cell is over-sensitive to stellate chopper cell banks where no phase locking is occurring.

5.5.2 Simplified Model

In order to alter the sensitivity of the new coincidence cell model we adjust its spiking threshold level. In order to find the optimum threshold level we carry out an experiment in which the behaviour of the cell is measured as a function threshold level.

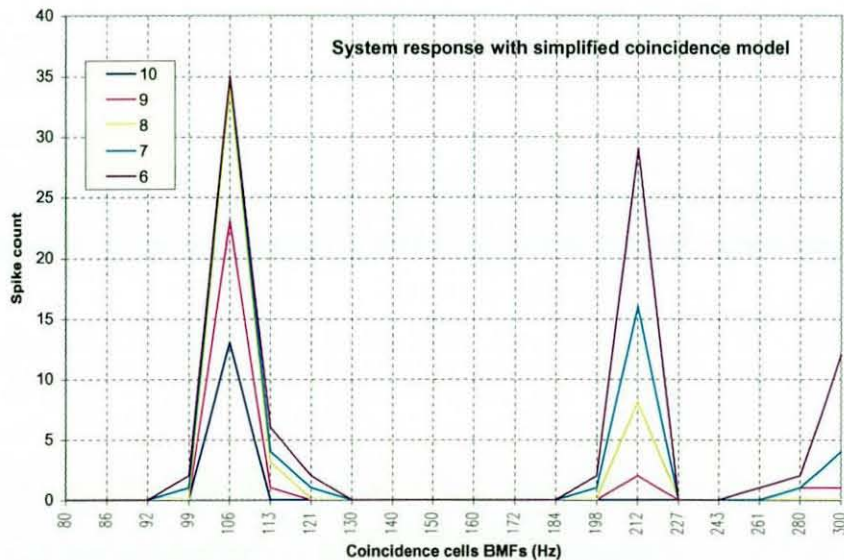


Figure 5-9: New model response as a function of threshold.

In Figure 5-9 we can see that the maximum modulation gain is provided a threshold of 9. However, we find that the optimum threshold is different for each stimulus. To accommodate this characteristic of the system we require different thresholds for each coincidence cell. Again, we can see some sub-harmonics effects at the coincidence cell sensitive to 212 Hz.

5.5.3 Variation to Simplified Model

It is clear that the response of the new coincidence model is a function of its instantaneous input value. This is in contrast to the original model, which is a function of its input values over a number of previous time steps. It is this 'leaky integration' process that allows the original model to detect phase locking in stellate chopper cell banks of different BMFs. For lower BMF stellate chopper cell banks, a phase locked group of spikes tends to be more loosely grouped than for higher BMFs. This means that checking for an individual peak in the input sum of spikes is not a reliable method of coincidence detection in the case of lower BMF stellate banks. In order to rectify this situation we propose an alteration to the simplified coincidence cell model. This alteration comprises summing the input values over a number of time steps before applying a thresholding function. Furthermore, the number of time steps over which the input is summed (which we will now call the detection window) depends on the BMF of the stellate bank. The detection window width should ideally be inversely proportional to the BMF of the stellate chopper bank to which the coincidence is attached. The closest we can come to this ideal situation is described in Table 5-3.

BMF (Hz)	Window size	BMF(Hz)	Window size
80	6	160	3
86	6	172	3
92	6	184	3
99	6	198	3
106	6	212	3
113	4	227	2
121	4	243	2
130	4	261	2
140	4	280	2
150	4	300	2

Table 5-3: Detection window widths for each coincidence cell.

The detection window widths in Table 5-3 are only approximately proportional to the BMF periods. We are obviously restricted by using as narrow a detection window as possible. Windows of greater than those described above allow lower levels of phase locking to be detected as readily as tightly packed phase locking. This is not desirable.

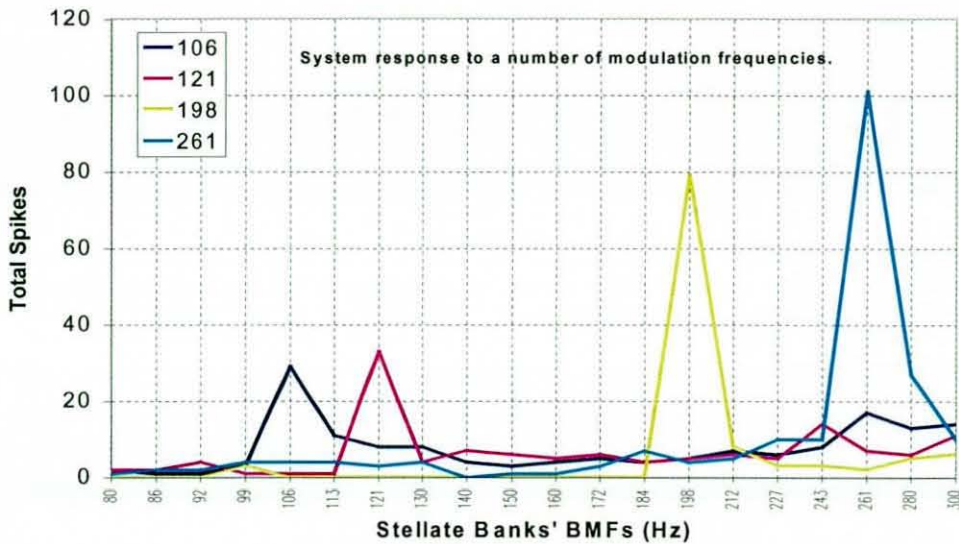


Figure 5-10: Coincidence cell response.

The response of the variable detection window coincidence cell model can be seen in Figure 5-10. We can see that at a threshold value of 10, the operation of the coincidence cells is maintained.

5.6 COINCIDENCE CELL MODEL COMPLEXITIES

In Table 5-4 we can a comparison between the complexities of the 2 coincidence cell designs. It is clear that the complexity of the simple threshold function is considerably lower than that of the MacGregor model implementation.

	Complexity (Gates)	Speed (MHz)	Cycles per epoch
REG_SET	206	57.34	(in parallel) 3
CALC_E	4441	62.15	77
CALC_GK	1886	60.98	77
CALC_S	306	70.13	77
COINC_CELL	6839	57.34	80

	Complexity (Gates)	Speed (MHz)	Cycles per epoch
ACCUMULATE	417	46.15	12
THRESHOLD	337	48.66	12
COINC_CELL	754	46.15	24

Table 5-4: Complexity and speed of coincidence cell implementations.

Considering the fact that the 2 models give similar behavioural responses it is clear that the simplified model is the most appropriate as it is the least complex implementation.

5.7 CONCLUSIONS

In this chapter we have discovered the NSR ratio of the original auditory nerve model and shown how it varies with the value of h.dt. We have applied the varying NSR values of the original AN model and of our simplified model to the system and measured the response. From this we have discovered that the models give similar results. Furthermore, we have measured the hardware complexity of the two models. From this we chose to use the simplified model because it is far less complex than the original model.

We have also compared the response of the original coincidence cell model and our simplified coincidence cell model. Again we have varied comparable parameters in both versions of the cell and measured the response. The comparison in the hardware complexity of the 2 models has also been made. Again, both models gave similar behavioural results but the simplified model resulted in a lower hardware complexity.

We have seen from the experiments that our simplified versions of both the auditory nerve model and the coincidence cell model give results that show a good correspondence to the original models while proving far simpler to implement in digital hardware.

CHAPTER SIX

DEVELOPMENT OF STELLATE CELL ARRAY

6.1 OBJECTIVES OF CHAPTER

In this chapter we attempt to reduce the hardware complexity of the pitch detection system implementation. We will focus our attention the most complex stage of the system, the stellate cell array. More specifically the objectives of this chapter are to:

- Investigate the effect of adding noise to the system after dendritic filtering. The aim is to reduce the number of dendrite filters required.
- Reduce the sample rate of the stellate cells. This will result in a reduction in the number of processing units required to implement the stellate cell array.
- Reduce the bit width of the stellate cell signals and further reduce the number of processing units required to implement the stellate cell array.

6.2 INTRODUCTION

The pitch detection system's input signal is processed by the first two stages of the system before reaching the stellate cell array. The stellate cell comes in two parts: the dendrite and the cell soma (nucleus). The dendrite acts as a low pass filter with a cut-off frequency of approximately 300 Hz. We have initially implemented the dendrite filter using a first order IIR filter. The input to the dendrite is the noisy output of the auditory nerve. Figure 6-1 shows the response of the dendrite filter.

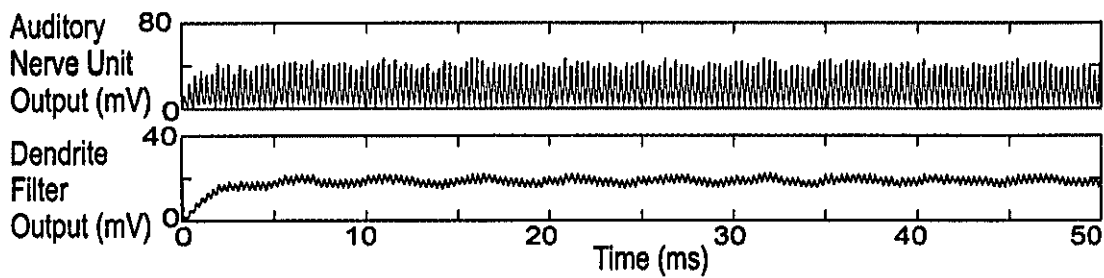


Figure 6-1: Response of dendrite filter.

On inspection we can see that the response of the dendrite output shows some remains of the carrier signal. As mentioned in the previous section we intend to decrease the sample rate of the output of the dendrite filter. It is possible that this remaining carrier signal could disrupt the output of the dendrite when the sample rate is decreased. As a result it may be necessary to employ a higher order filter in order to reduce the amount of carrier that remains after dendritic filtering.

Noise is added to the system by the auditory nerve. This noise is important specifically to the operation of the stellate cell somas. As a result, it is not necessary to add noise to the system before dendritic filtering. We will investigate the implications of moving the point of noise injection in terms of both system behaviour and hardware complexity.

The stellate cell soma is the spiking part of the cell. The input to the soma is the low-pass filtered response of the dendrite. The cut-off frequency of the filter is 300 Hz and the sample rate of the system is 20 kHz. This suggests that the stellate soma is currently subject to over-sampling. The response of the soma is shown in Figure 6-2. The value of the threshold, Th , varies so little that we have decided to make it a constant.

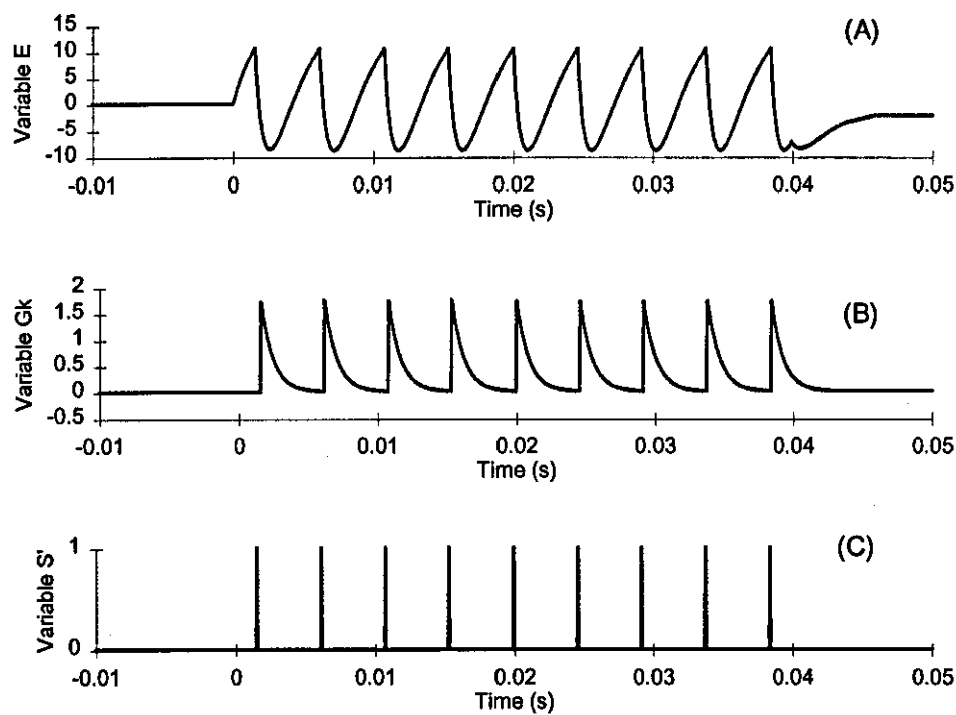


Figure 6-2: Response of stellate soma.

The spiking frequency of the stellate soma is determined by the parameters τ_{mem} and τ_{Gk} . The greater these values are the slower the spiking rate of the soma. We have discovered that when there is a ratio of 2:1 between t_{mem} and t_{Gk} then the maximum hyper-polarisation value of E is equal in magnitude to the threshold value, Th, as thus the maximum depolarisation value. This is only true when the inner hair cell has adapted to its steady state response.

BMF (Hz)	t_{mem} (s)	BMF(Hz)	t_{mem} (s)
80	5.08	160	2.58
86	4.74	172	2.40
92	4.43	184	2.25
99	4.12	198	2.09
106	3.86	212	1.95
113	3.62	227	1.83
121	3.39	243	1.71
130	3.16	261	1.59
140	2.94	280	1.49
150	2.74	300	1.39

Figure 6-3: Parameters of stellate somas.

The value of t_{mem} (and therefore double the value of t_{gl}) for each stellate bank BMF is shown in Figure 6-3. The number of fractional bits required to avoid underflow determines the number of bits required to represent the stellate soma signals. Underflow occurs when the change in a signal (for example E) is so small that it is rounded to zero. The size of the change in a signal increases and the sample rate decreases. If the sample rate of a soma is reduced then the minimum bit width required to avoid underflow is reduced. Once we have determined an appropriate sample rate, we will endeavour to find a suitable bit width for the stellate cell soma signals.

Our implementation uses a threshold function to implement the coincidence cells in Stage 4 of the system. The threshold function counts the number of stellate cells in a bank, which spike within a certain time frame. The length of this time frame depends on both the sample rate and the BMF of its input: the stellate cell. If we change the sample rate of the stellate cells then the sample rate of the coincidence cells has to change accordingly. This may result in us having to adjust further the detection window width of the coincidence cell.

6.3 STELLATE CELL DENDRITE

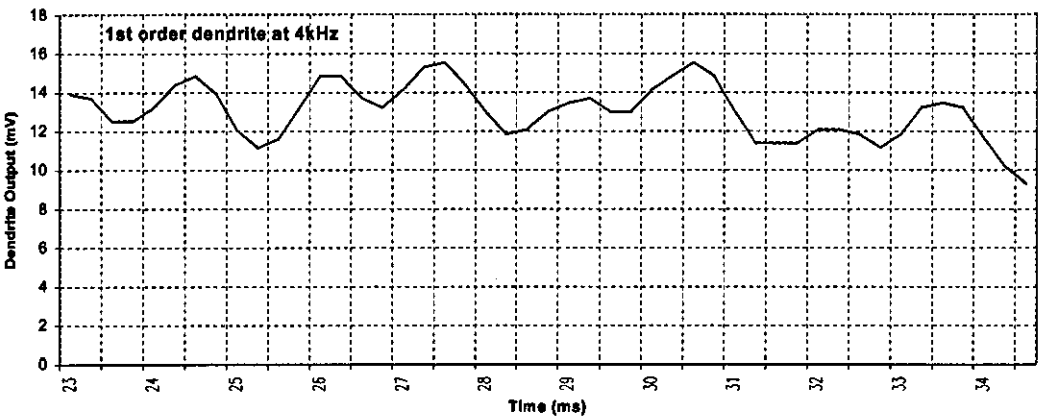
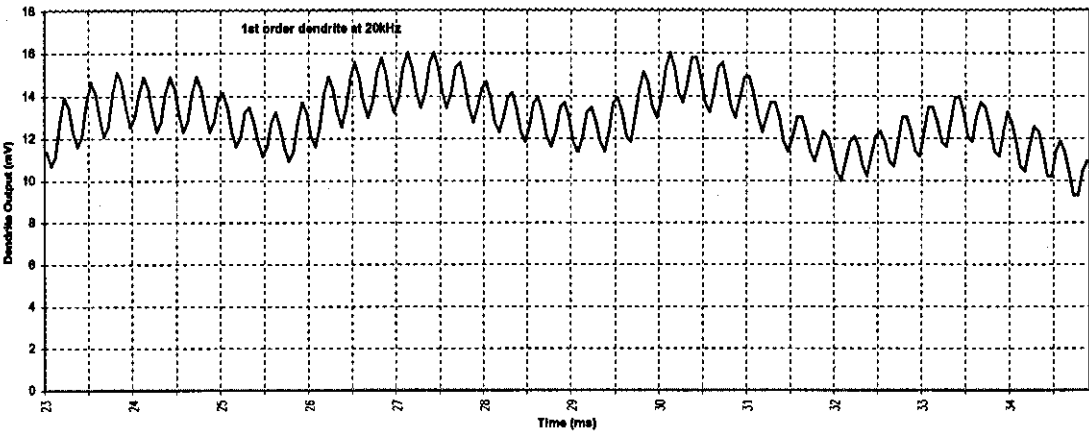
In this section we will assess the required filter cut-off gradient of the dendrite filter. In particular we assess which order of IIR is required to represent the dendrite filter. The initial implementation of the system uses a first order dendrite filter because that is what was used in the software simulations of Meddis *et al.* This implementation, however, does not eliminate the entire carrier frequency component. If we intend to reduce the sample frequency of the stellate cell soma then we need to be sure that any remaining carrier frequency component does not interfere with its operation.

We will also assess whether it is suitable to add noise after or before dendritic filtering. In the biological system and in our initial implementation noise is added before dendrite filtering. In our implementation this resulted in 40 noisy hair cell

signals being individually filtered by 40 dendrites. If we apply noise after dendritic filtering then we need only 1 dendrite for each channel of the system.

6.3.1 First Order or Second Order Filter

We will run 2 simulations in which the input to the filters is the output of a hair cell that has been stimulated by an amplitude-modulated signal. The initial sampling rate of the filter outputs is 20 kHz. Dividing the initial 20 kHz by integer values then reduces the sampling frequency giving us values of 10 kHz, 6.67 kHz, 5 kHz, 4 kHz, 3.33 kHz and so on.



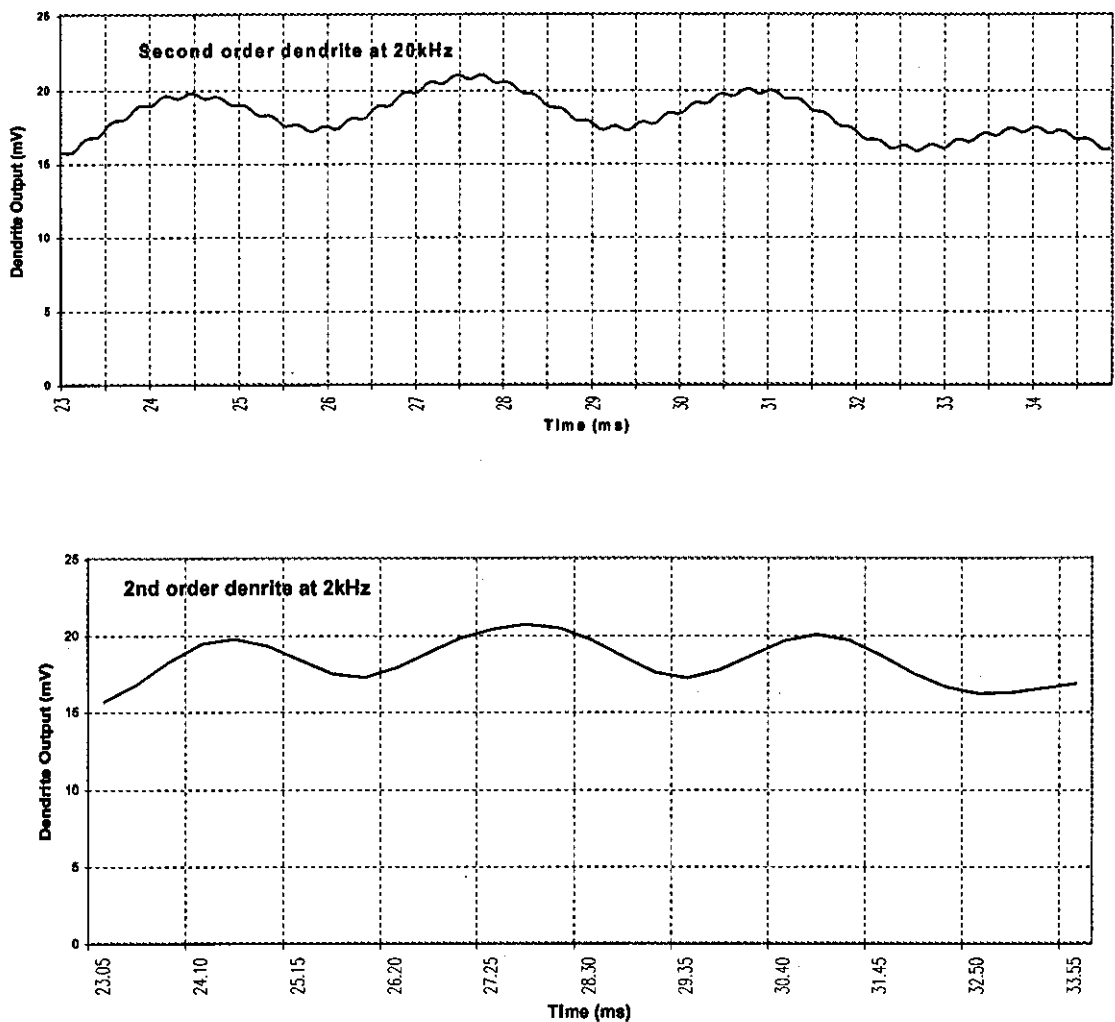


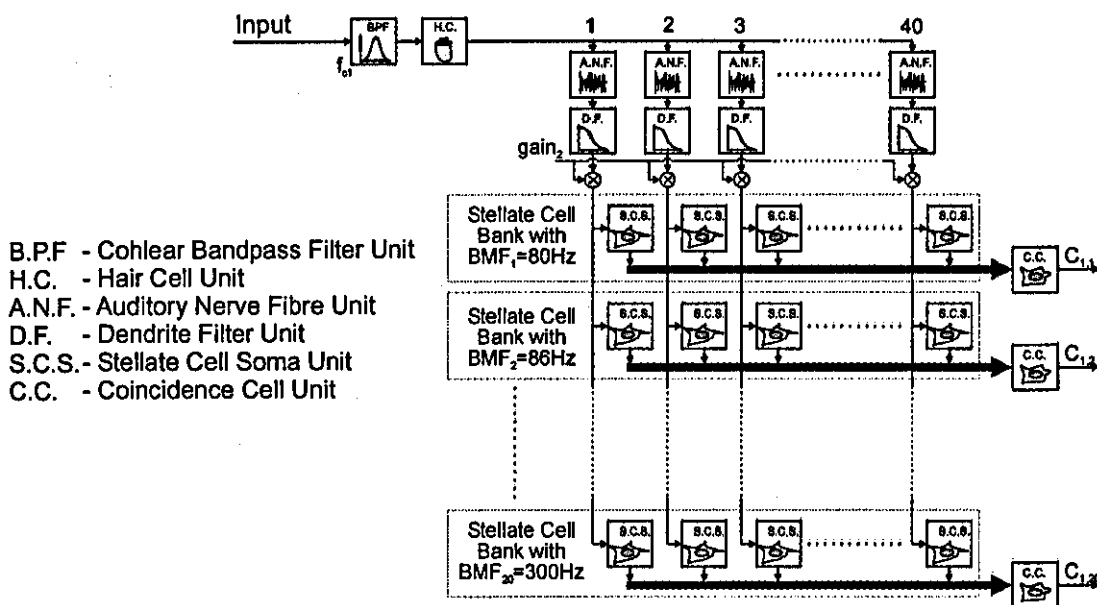
Figure 6-4: Comparison of filter responses.

As an experiment we apply a tone of 3.35kHz, which is modulated at 198Hz, to the system and observe the output of the dendrite filters at various sample frequencies. We can see from Figure 6-4 that the first-order filter corrupts the modulation information at a sample frequency of 4 kHz. The second order filter, however, maintains the modulation frequency information at the sample frequency drops below 2 kHz. This shows that the minimum sample rate of the second-order filter is lower by at least a factor of 2. The reason that we have not reduced the sampling rate further is that we have gone beyond the Nyquist limit of the 3.3 kHz carrier frequency. We have

demonstrated that we have eliminated enough of the carrier signal such that it no longer effects the output signal. This also means that do not need to evaluate any higher order filters than this second-order IIR filter.

6.3.2 Noise Before or After

In the auditory system noise is added to the signal as it travels down the auditory nerve before it reaches the stellate cells. Noise is added to the signal during this transmission. We have made a number of changes to each of the components models up until now, but none to the structure or architecture of the system itself. In this section we propose the first change to the system architecture. We believe that by adding noise to the system after dendrite filtering instead of before we can reduce the complexity of the architecture and thus reduce the hardware complexity of the system. We will now investigate whether we can add noise after the dendrite filters and still maintain the behaviour of the system.



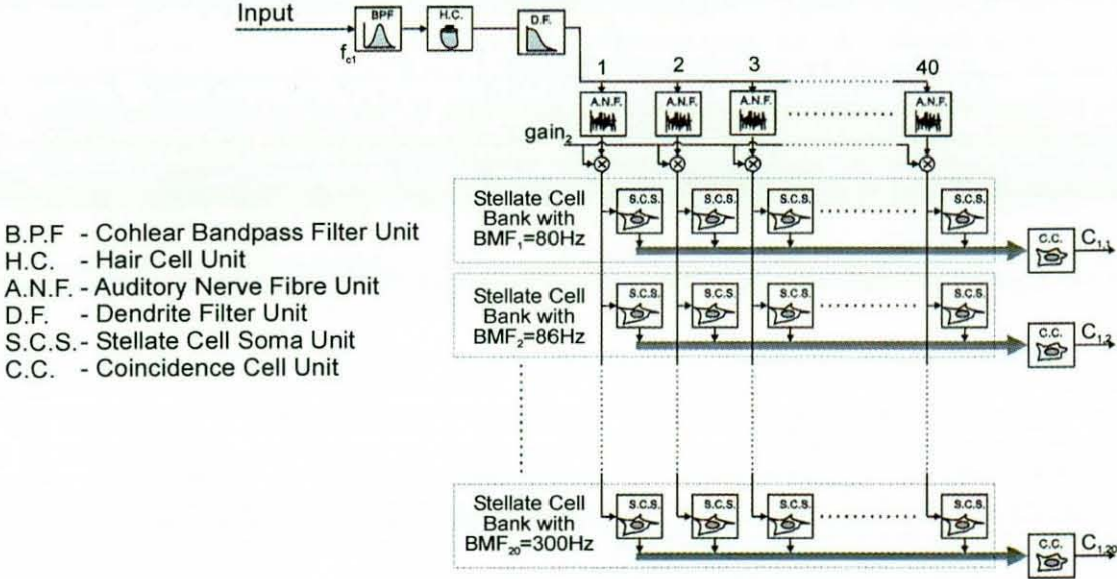
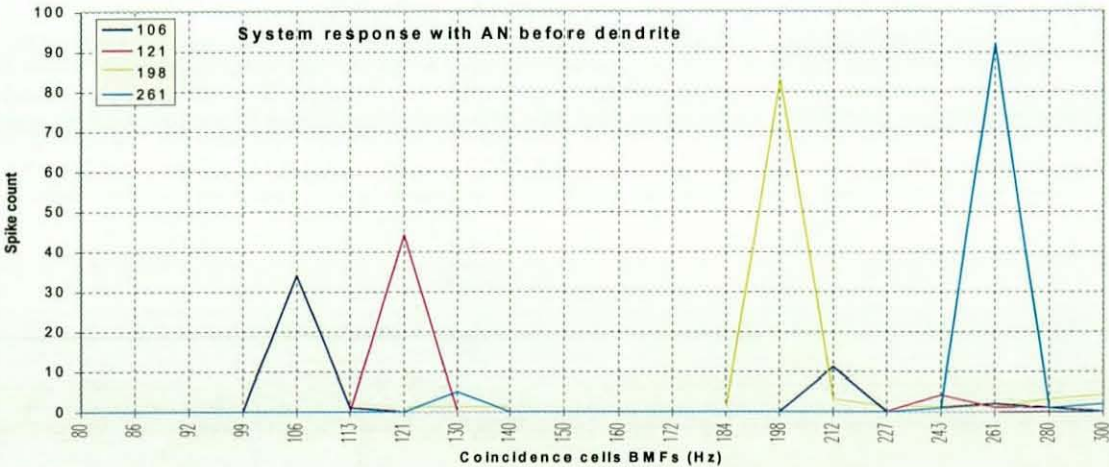


Figure 6-5: Original and new system architecture.

Our motivations for this proposed alteration to the system architecture are that the alteration will result in a reduction in the number dendrite filters that need to be represented in the system. We can see from Figure 6-5 that this architectural change means that we only require 1 dendrite filter per channel therefore providing a reduction by a factor of 40.



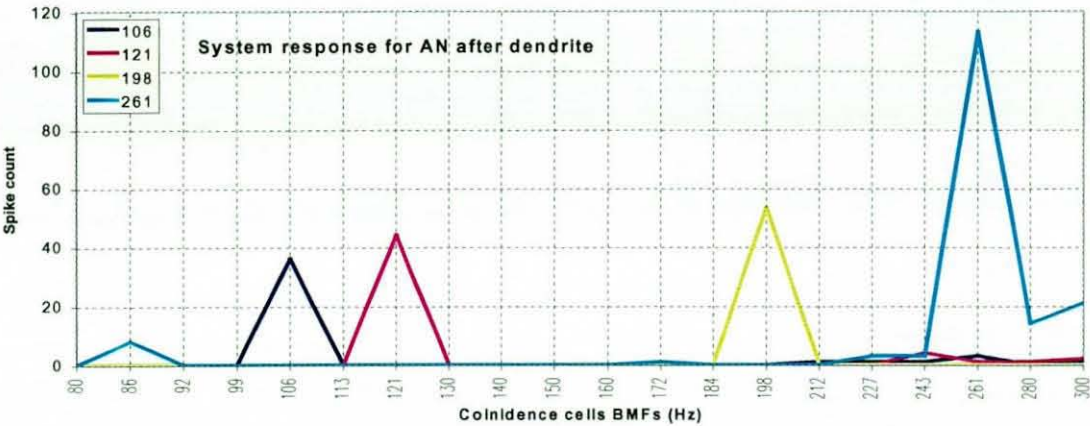


Figure 6-6: Comparison of noise before and after dendritic filtering.

We can see from Figure 6-6 that we can add noise after dendritic filtering the behaviour of the system is preserved.

6.3.3 Summary of Hardware Impact

By choosing a second-order filter and adding noise after filtering we have reduced the complexity of the system.

	Complexity (Gates)	Speed (MHz)	Cycles per epoch
DEND_S1	4033	56.98	34
REG_SET	633	32.22	(in parallel) 6
DEND_CELL	4666	32.22	40

	Complexity (Gates)	Speed (MHz)	Cycles per epoch
DEND_S1	6534	56.98	34
REG_SET	1121	42.09	(in parallel) 6
DEND_CELL	7655	42.09	40

Table 6-1: Complexity of first and second order dendrites.

We can see from Table 6-1 that we have actually increased the complexity of the individual dendrite processing unit. However, we can assume that the dendrite can be pipelined such that a single processing unit can represent up to 40 dendrite filters. This means that instead of one processing unit per cochlear channel we now only require one processing unit for the whole system.

6.4 STELLATE CELL SAMPLE RATE

The sample rate of the stellate cell somas is 20 kHz and the input to each soma is the output from a low-pass filter of cut-off frequency 300 Hz. The only way to discover if this is a case of over-sampling is to experiment with reducing the sampling rate in a systematic way until we find the point where the behaviour of the system breaks down. Our motivation for seeking this reduction in sample rate is to reduce the hardware complexity of this stage of the system.

6.4.1 Behavioural Effect of Reducing Sample Rate

We apply a tone of 3.3 kHz modulated at 106 Hz to the channel. We then measure the response of the coincidence cell of BMF 106 Hz. Figure 6-7a shows a graph of the activity of this coincidence cell as a function of sample rate. We can see that at sample rates below 3.3 kHz the activity of the coincidence cell drops significantly. We repeat the experiment for the tone modulated at 261 Hz and measure the response of the coincidence cell of BMF 261 Hz. Figure 6-7b shows a graph of the activity of this coincidence cell as a function of sample rate. In this case the activity of the coincidence drops significantly for sample rates below 10 kHz. These graphs show that the critical sample rate varies for each stellate bank. On repeating this experiment for each of the stellate banks we build up a graph of the critical sample rates of the stellate banks.

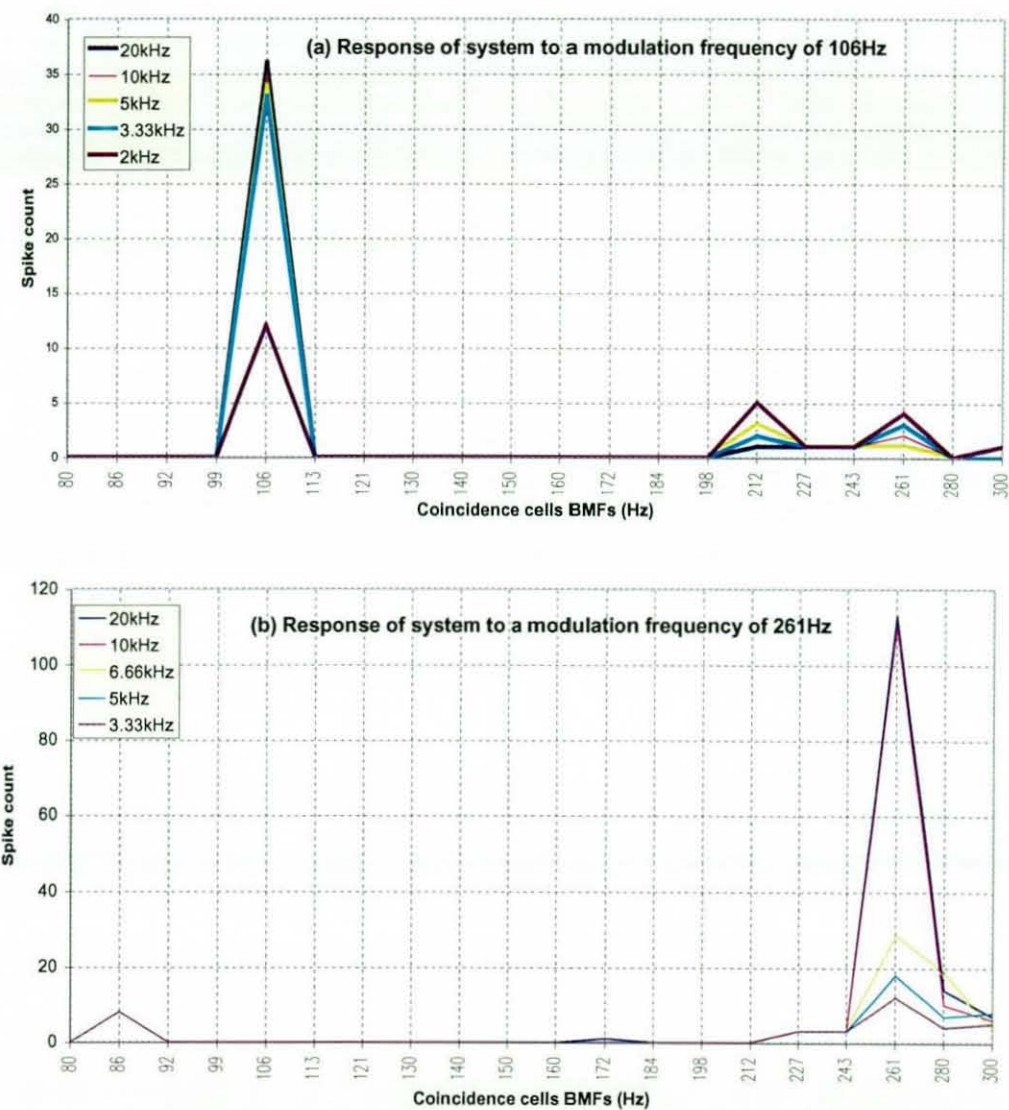


Figure 6-7: Response of system for varying sample rates.

When we try to detect lower modulation frequencies we find that we can reduce the sample rate further. This suggests that there should be different sample rates for different stellate banks. Table 6-2 shows the sample rates for each of the stellate cell soma banks. These values were calculated empirically to give lowest sample rate while maintaining system behaviour.

BMF (Hz)	Sample rate (kHz)	Tmem (s)	BMF(Hz)	Sample rate (kHz)	Tmem (s)
80	3.33	30.48	160	6.66	30.48
86	3.33	28.48	172	6.66	28.48
92	3.33	26.58	184	6.66	26.58
99	3.33	24.72	198	6.66	24.72
106	3.33	23.16	212	6.66	23.16
113	5.00	18.10	227	10.00	18.10
121	5.00	16.95	243	10.00	16.95
130	5.00	15.80	261	10.00	15.80
140	5.00	14.70	280	10.00	14.70
150	5.00	13.70	300	10.00	13.70

Table 6-2: Sample rates and coefficient values for stellate banks.

We can see from Table 6-2 that the coefficients values for banks 1 to 10 are identical to the coefficients for banks 11 to 20. We can account for this by looking at the example of banks 10 and 20. Bank 10 has a BMF half the value of bank 20’s BMF but also has half the sample rate.

6.4.2 Effect on Complexity of Reduced Sample Rate

The reduction in sample rate causes a significant reduction in the hardware complexity of the pitch detection system. The reason for this is that a processing unit that usually represents a single stellate bank at a sample rate of 20kHz can represent 2 stellate banks each sampling at 10kHz and so on. In Table 6-2 we see exact values involved.

Block	Complexity (Gates)	Speed (MHz)	Clock Cycles
MEM_CONTROL	206	57.34	(in parallel) 3
CALC_E	4441	62.15	77
CALC_GK	1886	60.98	77
CALC_S	306	70.13	77
TOTAL	6839	57.34	80

Table 6-3: Complexity of stellate chopper cell soma.

Table 6-3 shows that we can reduce the number of stellate soma processing units from 20 to 10. This is not as low as we would expect considering the large sample rate reductions in some of the stellate banks. The limited reduction in processing units can be explained in terms of the system architecture. Let us consider the 20 original stellate banks in 4 groups of 5. Each group has its sample rate reduced by certain amount. The highest BMF group (representing BMFs of 221Hz to 300Hz) has its sample rate reduced by a factor of 2. This means that in theory we now require 2.5 processing units which in practise are 3 processing units. The same applies to the other 3 groups, which now reduce to 2, 2, and 1 processing units. This is how we reach the total of 8 processing units.

6.5 STELLATE CELL BIT WIDTH

Now that we have reduced the sample rate we will now go on to find the lowest bit width possible to represent the stellate cell soma signals. Again, we will conduct a decremental progression in the bit-width. We will start at 16 bits and work down until the system behaviour breaks down.

6.5.1 Behavioural Effect of Reducing Signal Bit-width

Research has already been carried out [Lim97], which describes the relation between the sample-rate and the bit-width of the stellate cell soma signals and the accuracy of the representation. What has not been established is how much accuracy is required to maintain the system performance. In this section we aim to find out just how much accuracy is required.

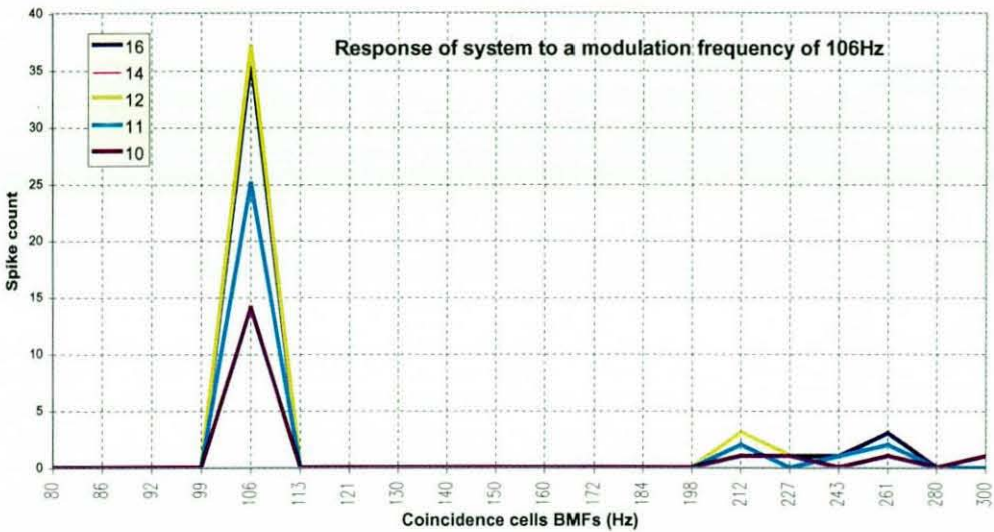


Figure 6-8: Response of coincidence cells to a varying bit-width.

In Figure 6-8 we can see the performance of the system as we vary the sample rate of the stellate cell soma. It shows that below the signal bit width of 12 the behaviour of the system begins to breaks down.

6.5.2 Effect on Complexity of Reduced Signal Bit-width

Reducing the signal bit width has 2 effects on the complexity of the system. Firstly, the size of the multipliers is reduced, and secondly, the number of clock cycles required to complete a multiplication is reduced. These effects are reflected in the complexity results shown in Table 6-4.

Block	Complexity (Gates)	Speed (MHz)	Clock Cycles
MEM_CONTROL	347	57.34	9
CALC_E	3864	62.15	51
CALC_GK	1850	60.98	51
CALC_S	306	70.13	51
TOTAL	6367	57.34	60

Table 6-4: Complexity and speed of stellate soma for 12-bit signals.

By considering the fact that the number of clock cycles per iteration is reduced by 25% we now only require only 8 processing units.

6.6 SUMMARY

In this chapter we reduced the complexity of the system considerably. This has been achieved in main ways:

- We have added noise after dendritic filtering thus reducing the number of dendrite filter processing units from 30 to 1.
- We have reduced sample rate of each of the stellate chopper cell soma. The reduction in sample rate varies depending on the BMF of the stellate bank.
- We have reduced the bit-width from 16 to 12 bits. This has reduced the complexity once more.

These developments have reduced the complexity of the system in a number of ways. We will now summarise this complexity reduction. The initial implementation of the system must be assumed to comprise:

- A clock frequency of 28.8MHz.
- 1 cochlear filter processing unit (= 1×10968 gates).
- 30 bit-serial inner hair cell processing units (= 30×8063 gates).
- 40 original auditory nerve processing units per channel (= $30 \times 40 \times 1860$ gates).
- 30 1st order dendrite filter processing units (= 30×4666 gates)
- 20 pipelined stellate cell processing units per channel (= $30 \times 20 \times 6839$ gates).
- 1 coincidence cell processing units per channel (= 30×6839) gates.

This gives us an overall system complexity of 6,933,408 gates. By the end of this chapter the system comprises the following:

- A clock frequency of 28.8MHz.
- 1 cochlear filter processing unit ($= 1 \times 10968$ gates).
- 1 pipelined bit-serial inner hair cell processing unit ($= 1 \times 7020$ gates).
- 1 pipelined dendrite filter processing unit ($= 1 \times 7655$ gates)
- 30 simplified auditory nerve processing units ($= 30 \times 2699$ gates).
- 8 pipelined stellate cell processing units per channel ($30 \times 8 \times 6367$ gates).
- 1 coincidence cell processing units per channel (30×754) gates.

This gives us a new total of 1,657,313 gates. This is clearly a significant reduction in complexity (23.90% of the initial gate count).

6.7 CONCLUSIONS

In this chapter we have reduced the complexity of the system through both architectural and timing alterations. Initially we moved the noise injection from before to after dendritic filtering. The result is that only a single processing unit is required to implement the systems dendrite filters.

We have also reduced the sampling rate of the stellate cell somas. This has resulted in the sample rate of the stellate somas being closer to cut-off frequency of the dendrite filters. Furthermore, this has caused a reduction in the hardware complexity of the system.

Finally, we have reduced the bit width of the stellate cells. When employing bit serial arithmetic, this reduces the number of clock cycles required to perform a calculation. Again, this causes a reduction in the complexity of the stellate soma processing units and a reduction in the number of stellate soma processing units required.

CHAPTER SEVEN

IMPLEMENTATION OF A SINGLE CHANNEL

7.1 OBJECTIVES OF CHAPTER

This chapter details the implementation and characterisation of a large portion of the pitch detection system. Specifically, our objectives are to:

- Implement a major portion of the pitch detection system. This sub-system should be able to represent any part of the entire system.
- Design and implement software to utilise the hardware. Our aim is that this software should provide the user with a selection of input waveform types and intensities with which to stimulate the hardware implementation.
- Apply a range of input types and intensities to the hardware using the software tool and measure the response. We aim to compare the response of the hardware to biological measurements.

7.2 INTRODUCTION

In Chapter 2 we described the architecture of the system. We now wish to identify a portion of the system to implement in hardware. By considering the available resources we aim to identify which parts of the system we can implement. As described in the thesis objectives we wish this implementation to be able to represent any part of the pitch detection system simply by adjusting its parameters. We will

initially look at the entire system and progressively break it down into smaller sub-systems until one of these sub-systems fit into our limited resources.

The software that communicates with the hardware must meet certain criteria. These criteria are:

- Ease of use
- Hardware transparency
- Control over the type of stimuli applied to hardware
- Clear representation of results

Ease of use comes about through presenting the user with clear and finite options. By definite options we mean providing a list of stimulus intensity options instead of typing a value into a box.

Hardware transparency is where the user requires no knowledge of how the PC interacts with the hardware design. The user should only have to understand what stimulus they wish to apply to the pitch detection system. This means that the software should handle all the required format changes from the user options to the hardware's input signals and deal with converting the hardware response into an intelligible format for non-engineers.

The stimuli that we can apply using this software should provide some useful results. In other words the response of the system must be comparable to biological measurements. Although amplitude modulated signals are the most common type of stimulus in biological experiments, the response of the system to pure tones and noise may also provide interesting results.

Clear representation of results is easily achieved using the built in waveform functions in the Labview language. This allows the display of the hardware response as it is produced and not after the hardware has finished being stimulated.

7.3 ARCHITECTURE OF IMPLEMENTED SYSTEM

In this section we identify the possible implementations that can be made. We will also calculate the resources required by each and compare them with the resources available to us, described in Table 7-1.

Part Type	Logic Gates	Embedded RAM bits
GF250F100	50,000	N/A
GF260F180	42,000	23,000

Table 7-1: Available resources for Gatefield chips.

We have identified four progressively simpler implementations of the system. These implementations are as follows:

- The entire system
- A single channel
- A single channel with only one stellate bank
- A single channel with only one stellate cell

Each of these sub-systems have a certain usefulness and complexity. We will now describe these sub-systems and their attributes.

7.3.1 Entire System

The implementation of the entire system is clearly the best result we can hope for. Such an implementation would clearly exhibit some usefulness.

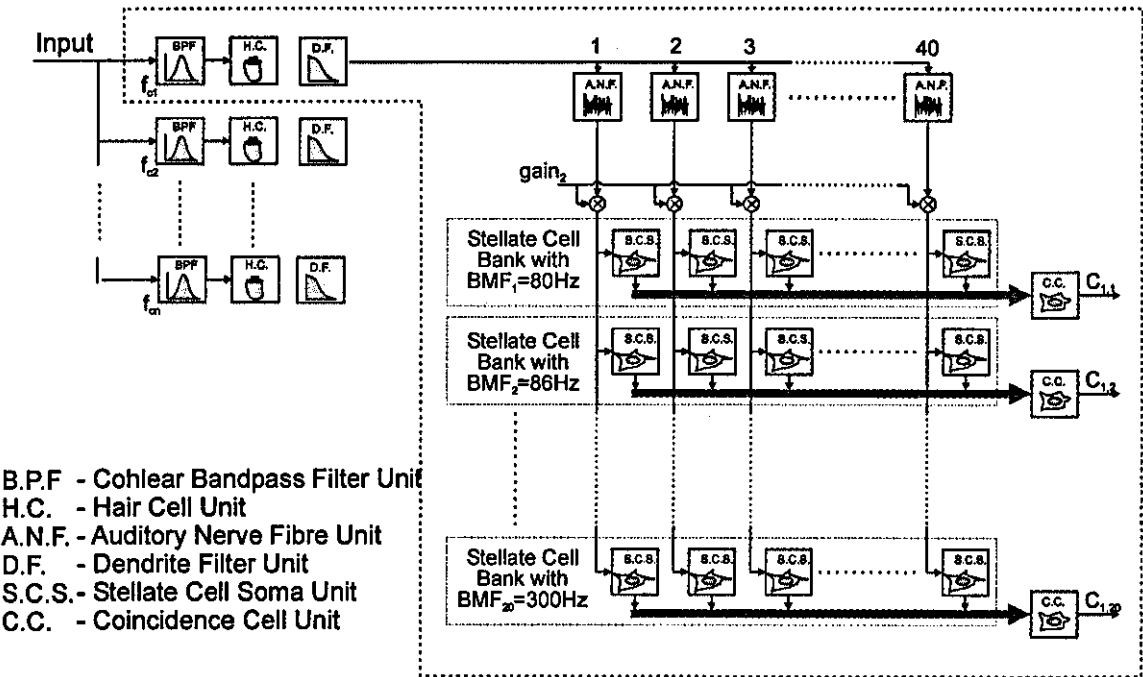


Figure 7-1: Entire system implementation.

The complexity of this full implementation is described in Table 7-2.

Component	Gates per unit	Number of units	Gates per unit
Cochlear Filter	10968	1	10968
Inner Hair Cell	7020	1	7020
Dendritic Filter	7655	1	7655
AN Noise Injection	2699	30	80970
Stellate Bank	6367	240	1528080
Coincidence Cell	754	30	22620
Total			1656313

Table 7-2: Hardware complexity of entire system implementation.

We can see that the hardware complexity of the entire system is too great to be implemented within our available devices.

7.3.2 Single Channel

The single channel implementation can represent any channel in the system by changing the cochlear filter parameters.

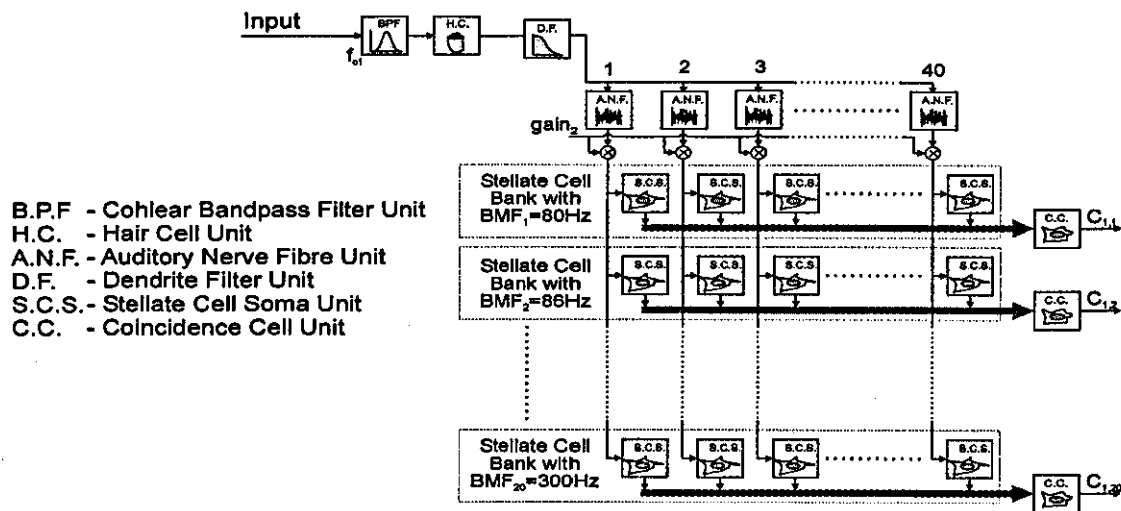


Figure 7-2: Single channel implementation.

The complexity of this single channel implementation is described in Table 7-3.

Component	Gates per unit	Number of units	Gates per unit
Cochlear Filter	10968	1	10968
Inner Hair Cell	7020	1	7020
Dendritic Filter	7655	1	7655
AN Noise Injection	2699	1	2699
Stellate Bank	6367	8	50936
Coincidence Cell	754	1	754
Total			80032

Table 7-3: Hardware complexity of single channel implementation.

We can see that the hardware complexity of the entire system is too great to be implemented within our available devices.

7.3.3 Single Channel With Only One Stellate Bank

This implementation can represent any stellate bank in the system by altering the cochlear filter parameters and the stellate bank parameters. Furthermore, by programming its parameters we can alter the BMF of the stellate bank.

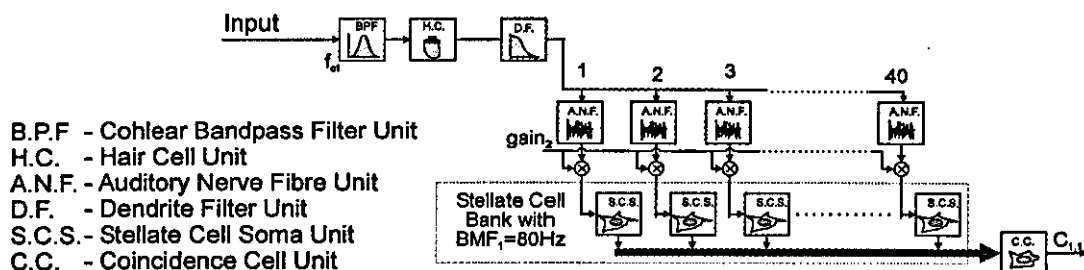


Figure 7-3: Single channel implementation but only one stellate bank.

The complexity of this simplified channel implementation is described in Table 7-4.

Component	Gates per unit	Number of units	Total gates
Cochlear Filter	10968	1	10968
Inner Hair Cell	7020	1	7020
Dendritic Filter	7655	1	7655
AN Noise Injection	2699	1	2699
Stellate Bank	6367	1	6367
Coincidence Cell	754	1	754
Total			36163

Table 7-4: Hardware complexity of simplified channel implementation.

We can see that the hardware complexity of this implementation is such that it fits a single GF260F device.

7.3.4 Single Channel With Only One Stellate Cell

This implementation can represent any data route in the system up to and including the stellate cells. The coincidence cell, however, is now redundant because we can no longer represent a bank of stellate cells.

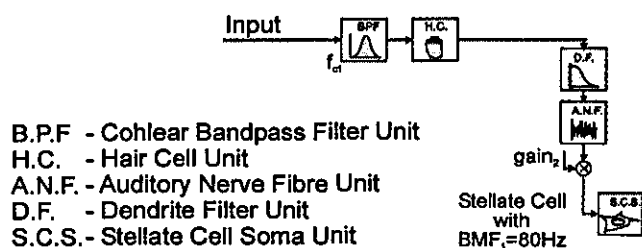


Figure 7-4: Single channel implementation but only one stellate cell.

The complexity of this one-of-each-cell implementation is described in Table 7-5.

Component	Gates per unit	Number of units	Total gates
Cochlear Filter	10968	1	10968
Inner Hair Cell	7020	1	7020
Dendritic Filter	7655	1	7655
AN Noise Injection	2699	1	2699
Stellate Bank	6367	1	6367
Coincidence Cell	754	1	754
Total			36163

Table 7-5: Hardware complexity of one-of-each implementation.

We can see that the hardware complexity of this arrangement can be realised in either a single GF250F device or a GF260F device (no memory required).

7.4.5 Summary

We have chosen to implement the channel with only a single stellate bank. This implementation allows us to represent any part of the system. Furthermore, it allows to implement a bank of stellate cells which gives us some usefulness from the hardware.

7.4 HARDWARE ARCHITECTURE

The simplified implementation consists of 6 top-level blocks. The blocks represent:

- 1) The cochlear filter
- 2) The hair cell
- 3) The dendrite filter
- 4) The synaptic noise injection
- 5) The stellate cell bank
- 6) The coincidence cell

These blocks have a standard input/output. This interface is described in Table 7-2. We can see that the bit-serial nature of our implementations allows a low pin count for each block even when considering the parameter loading.

Port Name	Port Type	Signal Description
CLK	Input	System clock
RST	Input	Active-low synchronous reset
SYNIN	Input	Active-high start signal
VI	Input	Input data stream
LOAD	Input	Active-high parameter load enable
PARAM	Input	Parameter data stream
VO	Output	Result data bit stream
SYNOUT	Output	Active high result start signal

Table 7-6: Port map of top-level blocks.

7.5 TEST-BENCH SOFTWARE

The first two criteria for the test-bench software are ease of use for non-engineers, and the intuitive display of results. These criteria are addressed in the next section, which describes the software's graphical user interface. The final criterion is the application

of representative stimuli. The response of the system to these stimuli is described in following section.

7.5.1 Software Graphical User Interface

One of the purposes of the software is to allow the hardware behaviour to be evaluated. This evaluation is likely to be carried out by physiologists and psychologists. As such the software must be easy to use and must make the interfacing with the hardware transparent to the user.



Figure 7-5: Snap shot of the software's GUI.

The response of the system must be displayed in a meaningful format. This is best achieved by displaying the results graphically. The software, therefore, must convert the binary values produced by the hardware and convert them into decimal format and also perform some scaling before producing a graphical result.

7.5.2 Software/Hardware Communication

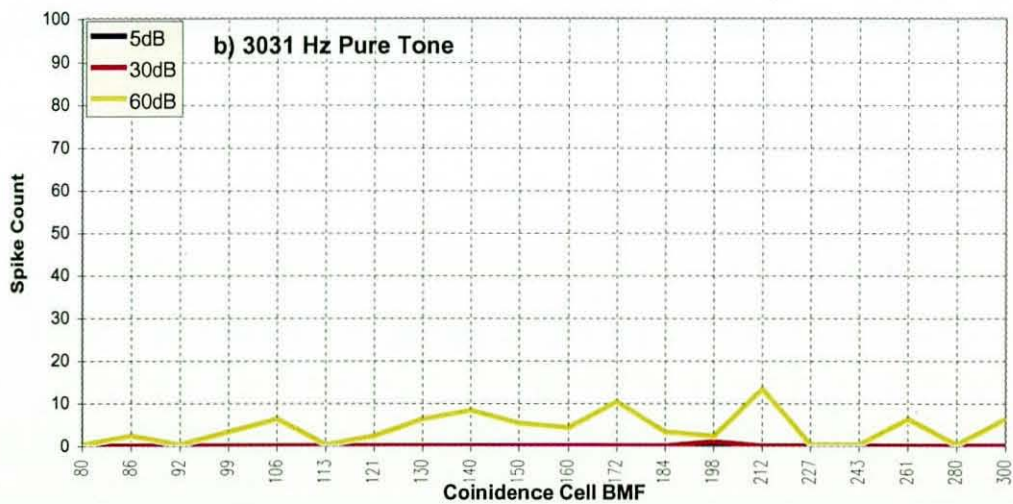
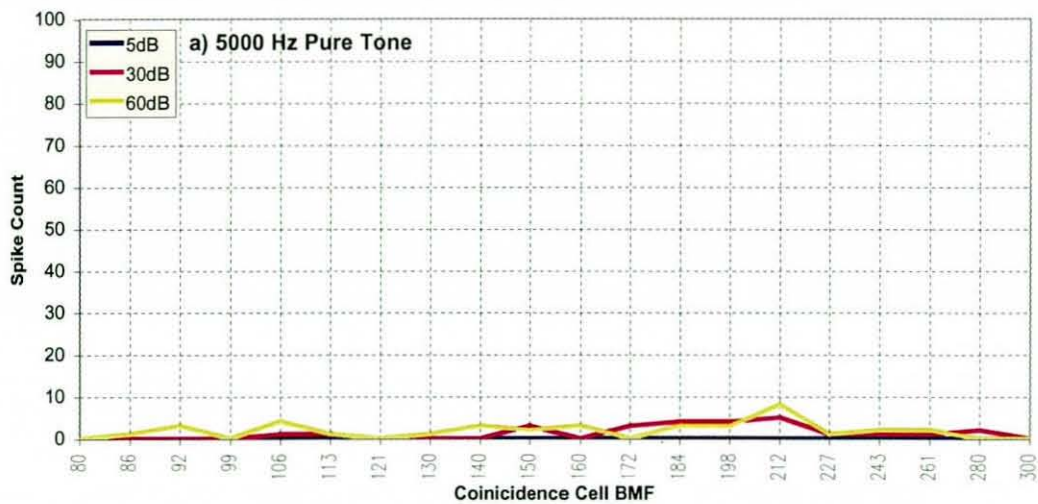
The software communicates with the hardware at bit level. Because we have implemented a synchronous design we need to apply new data to the input of the hardware at every half clock cycle. The list of input and outputs for each block can be seen in Table 7-6. This brings the total number of input pins to 26 if we assume that only require one CLK pin and one RST pin. We now have to consider the size of data blocks we apply to and read from the hardware at a time. The total number of clock cycles required to apply an input value to the cochlear filter and receive a result from the coincidence cell is 1569. This means that we need to present 3138 consecutive inputs to the system per iteration. In total this brings our data requirements to an array of 3138 by 26 bits per iteration. Such an array is generated in full before being applied to the hardware. A similar output array is generated as the input array is applied to the hardware and its response measured. Again, this output array is not analysed until the iteration is complete.

7.6 BEHAVIOUR CHARACTERISATION

Not only must we compare the response of the system against biological measurements but also against the original implementation of the system that was characterised in [Lim98]. For this reason we must allow the application of pure tones, white noise, and amplitude modulated tones to the system. These waveforms must also be applied at a number of intensities.

7.6.1 Pure Tone

As mentioned in the previous section there are a number of waveforms that we must apply to the system in order to characterise it. Initially, we will deal with the application of a pure tone.



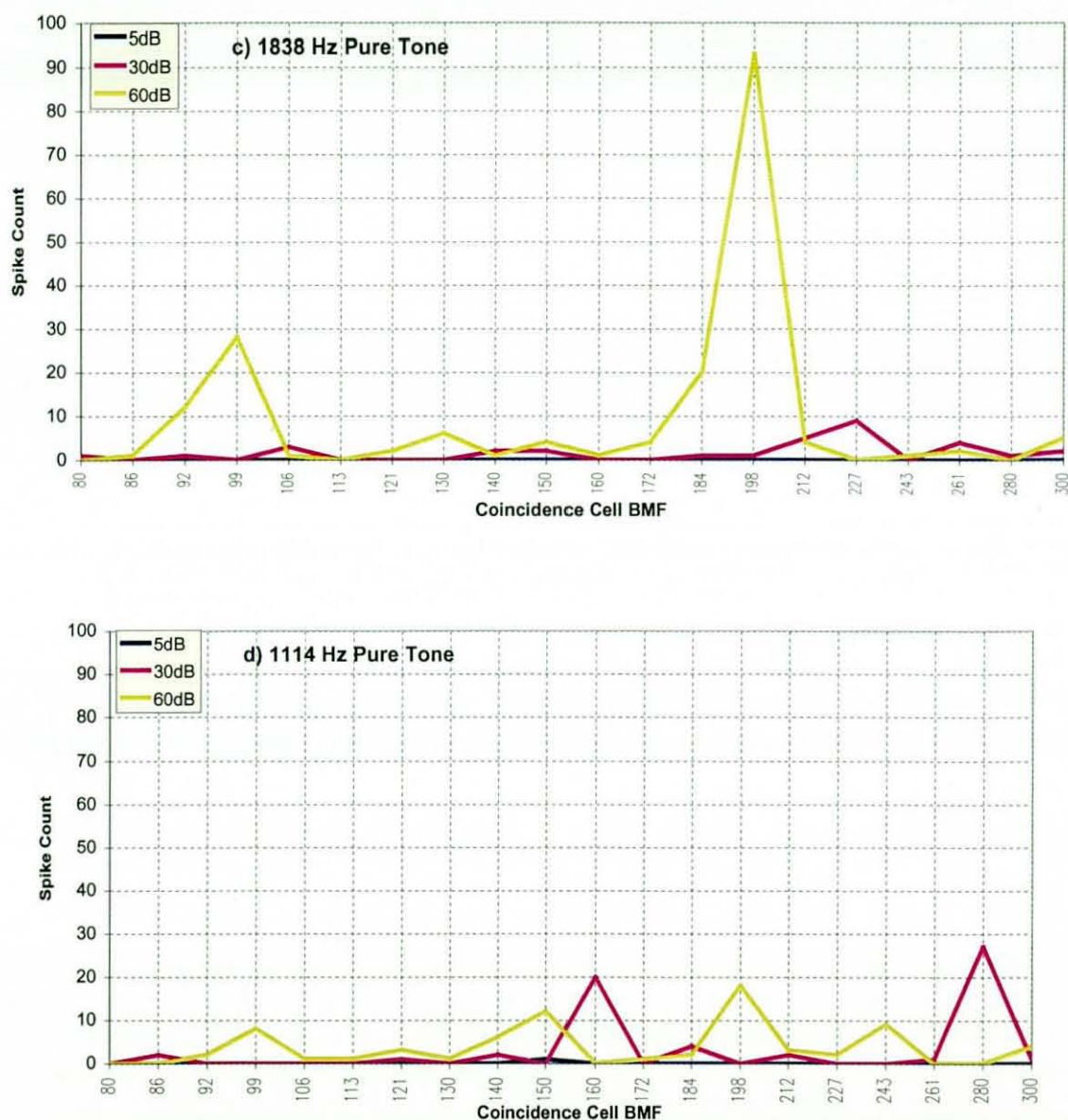
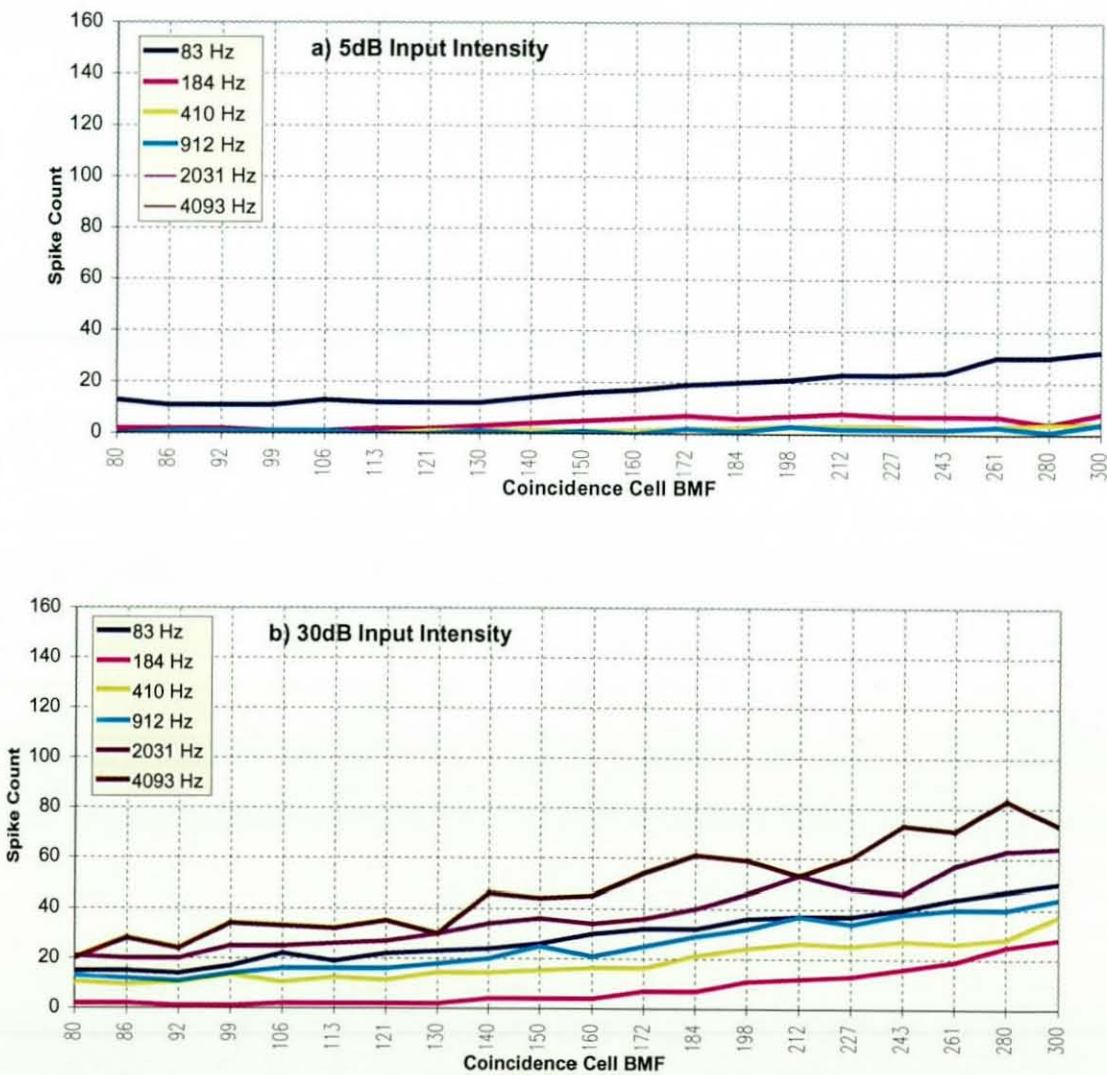


Figure 7-6: Response of the system to pure tone stimuli.

In Figure 7-7 we can see the response of the system to a number of tones. We expect the system to show little activity when presented with these tones because the dendrite filter has a low-pass cut off frequency of 300Hz. This is true in all cases except graph (c). In the graph (c) we can see a high level of activity in the coincidence cell with BMF 198Hz. This is unexpected since 198 is not an integer factor of 1838.

7.6.2 White Noise

The response of the system to the application of white noise can be seen in Figure 7-8. We can see that the systems coincidence cells give a uniform response. The highest activity, however, occurs in the lower frequency cochlear channels. This can be explained by the fact that the frequency components of the noise signal below 300 Hz can pass through these lower frequency cochlear filters and the dendrite filters. Higher frequency components, however, are cut-off when they reach the dendrites.



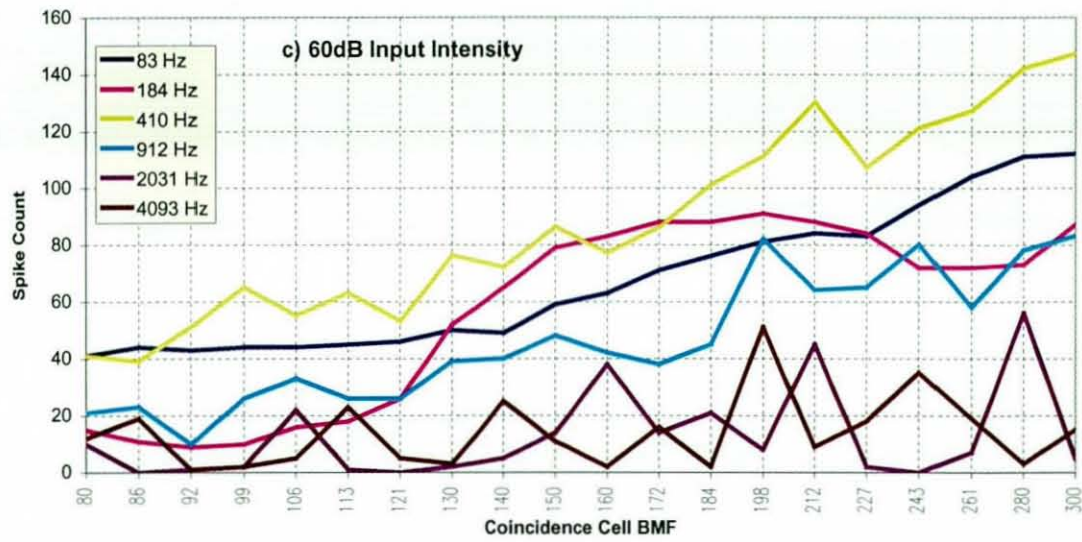
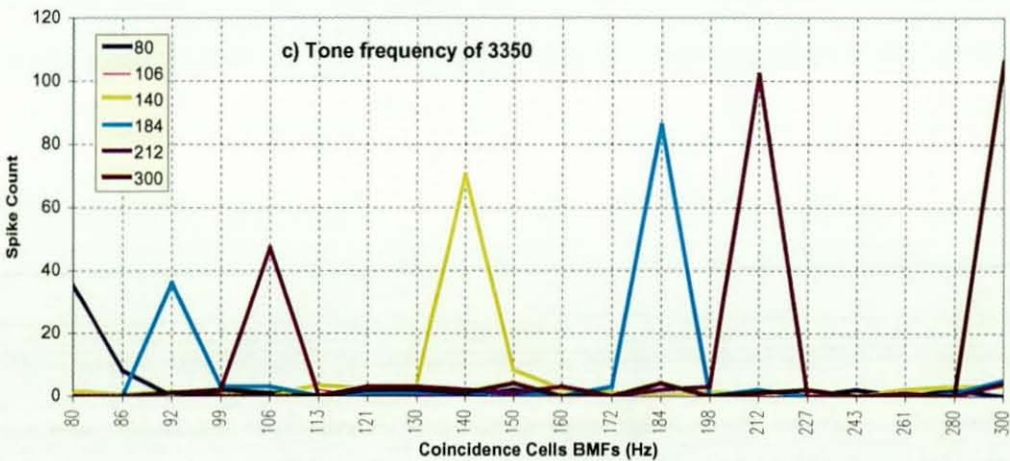
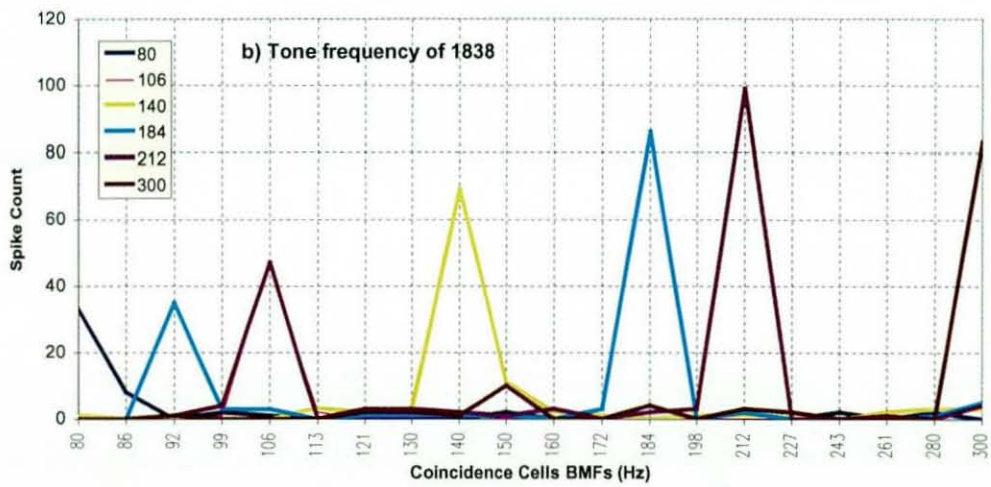
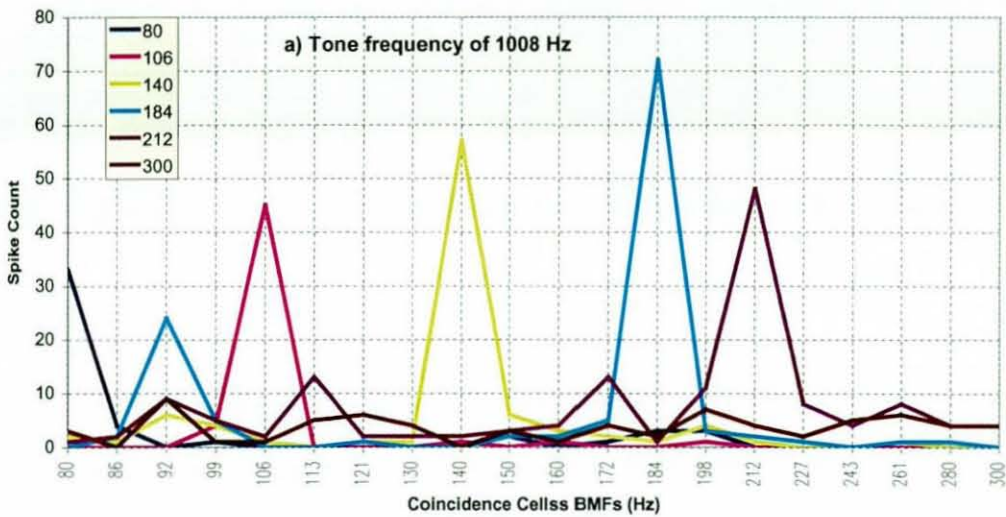


Figure 7-7: Response of the system to white noise stimuli.

7.6.1 Amplitude Modulated Tone

When we apply a modulated tone the appropriate coincidence cell clearly shows a higher level of activity than the other coincidence cells (see Figure 7-9). These results are similar to those shown by Lim. The difference in the results shown below is that the modulation frequency component is detected more clearly than in Lim’s results at lower carrier frequencies. The higher order dendrite filter causes this improvement by removing the carrier component of the modulated signal more efficiently.



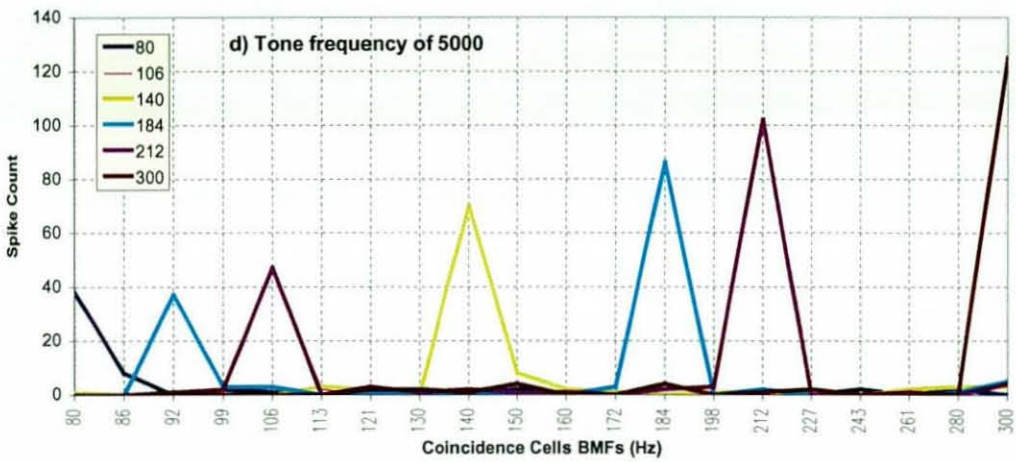


Figure 7-8: Response of the system to amplitude modulated at 30dB.

7.7 CONCLUSIONS

In this chapter we identified a suitable portion of the pitch detection system for implementation in digital hardware within our available resources. We identified a single channel with only one stellate bank instead of 20. The parameters of this implementation can be altered such that it can represent any portion of the system.

We then implemented this portion of the system on an FPGA. Once the implementation had been verified we designed software to provide an interface to the hardware. This software facilitates ease of use for non-engineers, together with the ability to apply a variety of waveforms at different intensities to the system. Furthermore, the software takes the low-level response of the circuit (binary format) and produces a graphical display.

We then applied a number stimuli which are representative of natural sounds that would exist in the biological system. The system managed to identify the modulation frequency of amplitude modulated tones. It also largely managed to avoid being active when presented with pure tones.

CHAPTER EIGHT

CONCLUSIONS

8.1 OBJECTIVES OF CHAPTER

This chapter concludes the thesis. It reviews the objectives of the thesis and summarises the conclusions from each of the investigations. From this the main conclusions and contributions of this thesis are set out. This chapter also considers the limitations of this work and suggests ways in which this research can be extended.

8.2 REVIEW OF OBJECTIVES

Our objectives are to develop and implement a neuromorphic pitch detection system in digital electronic hardware. The review of neuromorphic systems and pitch detection identified three key issues:

- Architecture of neuromorphic systems.
- Models of biological components.
- Implementation of biological components and systems in digital hardware.

Neither the pitch detection system nor its components had been implemented in digital hardware before. For this reason it was necessary to gain a better understanding of the issues surrounding such digital designs. This would lead to a better understanding of the feasibility of implementing neuromorphic systems in general.

Leading on from this, Chapter 3 presented an analysis of these issues which lead on to a detailed choice of four investigations and a statement of the research objectives:

- The investigation into the digital electronic implementation of the inner hair cell.
- The development of simple models and digital hardware implementations of the auditory nerve and the coincidence cells.
- The reduction in hardware complexity of stage three of the system: the stellate array. This is the most complex stage of the system.
- The investigation into the implementation characterisation and utilisation of a large portion of the pitch detection system.

In Chapters 4 to 7 investigations were carried out in response to these objectives. Chapter 4 investigated the digital hardware implementation of the inner hair cell and Chapter 5 developed models of the auditory nerve and coincidence cell appropriate for digital hardware. Chapter 6 set out to reduce the hardware complexity of stage 3 of the design. Finally, Chapter 7 investigated implementing a large portion of the system in digital hardware.

8.3 CONCLUSIONS

The investigations in this thesis have addressed a number of key questions. This section presents the main conclusions drawn from these investigations.

The investigation into the digital electronic implementation of the inner hair cell looked at themes such as the bit-width of internal signals and simplified algorithms for internal functions. From this investigation the following conclusions can be drawn:

- By enforcing a mean error limit of 10% we discovered that 15-bit signal representation was the lowest resolution possible for the Meddis inner hair cell model.

- It is possible to implement the Meddis hair cell model in digital electronics and run it in real-time. Also the speed of the design is such that a single processing unit can be multiplexed to represent a number of inner hair cells.
- It is possible to simplify the saturation function using a piece-wise linear approximation. This not only reduces the complexity of the design but also reduces the number of clock cycles required to carry out each calculation.
- By reducing the number of clock cycles required by each block in the design to perform a calculation it is possible to pipeline the design. This allows the possibility of using a single processing unit to represent all the inner hair cells within the pitch detection system.

The models that exist for the auditory nerve and the coincidence cell are complex. This is because they try to represent exact biological operations and not general behaviour. In terms of a digital hardware implementation, however, it is not possible to implement such models efficiently. For this reason we investigated alternative models for these entities that provided less complex digital hardware implementations. The main conclusions that can be drawn from these investigations are:

- A new model for the auditory nerve was developed. By applying noise, in proportion to the signal value, directly to the output signal of the hair cell the overall effect of the auditory nerve is replicated.
- The optimal noise-to-signal ratio, in terms of system behaviour, was measured for each model while applying a typical stimulus. A comparison was then made between the two models to measure their effect on the system behaviour for a number of stimuli. The new model was observed to maintain the system behaviour and provide a less complex design.
- A new model for the coincidence model was developed. This model applies a simple threshold function to the sum of spikes, which is its input, instead of using the MacGregor neuron model.
- The optimal input gain and threshold level for each model was measured while applying a typical stimulus. A comparison was then made between the two models

to measure their effect on the system behaviour for a number of stimuli. The new model is shown to maintain the system behaviour and results in a less complex design.

The next step is to reduce the hardware complexity of the system. The most complex stage in the system is Stage 3, the stellate cell array. In particular the stellate cell somas take up the majority of the design gate count. The main developments from this investigation are:

- Changing the order of noise injection and dendritic filtering results in a reduction in the number of dendrite filters required. This is because the noise injection within the auditory nerve takes 1 input and produces 40 outputs.
- Reducing the sample rate of the stellate cell soma from 20 kHz to 4kHz in some cases does not affect the behaviour of the system. The minimum sample rate of each stellate bank is found to be proportional to its natural frequency.
- Reducing the bit-width of the stellate cell soma model internal signals to less than 12 bits causes the behaviour of the system to break down. The reduction to 12, however, allows the each operation to be completed in less clock cycles.

Finally, we implemented a significant proportion the system onto an FPGA. First of all we had to select which portion of the system made best use of the available gate complexity.

- A single channel of the pitch detection system incorporating only 1 bank of stellate cells was implemented. This fully utilised the available gates while providing a useful system.
- Software was developed that stimulates the hardware and displays the systems response. The ease of use inherent in the software design and graphical nature of the response display allow non-engineers, such as physiologists and psychologists, to assess the behaviour of the hardware.

- A number of types and intensities of stimuli were applied. The response to these stimuli proved that the system functions as expected, and in correspondence to biological measurements.

This final investigation proves that we can implement useful neuromorphic systems in digital hardware and in real-time. Furthermore, we have shown that such an implementation can be used as a demonstrator to encourage growth in this emerging field.

8.4 MEASUREMENT OF SUCCESS

The broad aim of this thesis was to implement a significant portion of the pitch detection system in real-time digital hardware. The review of pitch detection systems and neuromorphic modelling identified three key issues concerning the implementation of such a system. The objective of this thesis was then to investigate these issues. Four investigations were proposed to address this objective:

- *The inner hair cell study aimed to produce an efficient digital hardware implementation of the Meddis model.* This was investigated in Chapter 4. We developed a design that not only runs in real-time but also can be pipelined to represent all the cells in the system using only one processing unit.
- *The development of models of the auditory nerve and coincidence cell that were appropriate for digital electronic implementation.* This was addressed in Chapter 5. Models were developed that are much more efficient in terms of hardware complexity than the original models.
- *The investigation into the reduction of the system complexity.* This investigation was carried in Chapter 6. Firstly, the performance of dendrite filtering was repositioned to before noise injection. Also, the sample rate and signal bit-width of the stellate soma was reduced. These changes resulted in a significant reduction in the hardware complexity of the system.

- *Implementing a significant portion of the system.* This investigation was carried out in Chapter 7. This implementation comprised a single cochlear channel with only 1 bank of stellate cells instead of all 20. This implementation can represent any part of the complete system and has shown the expected behavioural characteristics.

This thesis has, therefore, proposed and undertaken investigations that have addressed the main objective of the thesis. The main strength and contribution to knowledge of this thesis is the fact that it begins to address the many challenges in the largely unknown field of digital neuromorphic systems. As such the results contained in this thesis may be used in guiding future research efforts in this emerging field.

8.5 LIMITATIONS OF WORK

The following limitations apply to this work:

The complexity of the system was reduced in the stellate cell array investigation. The result of this investigation was that to implement the entire system we required only a single processing unit each for the dendrite filter. We also required five processing units to implement the stellate soma array for each channel. This investigation was limited, however, by the fact that we restricted our interest to bit serial arithmetic. This meant that we could not reduce the complexity of the stellate soma array to that of a single processing unit through the use of, for example, a parallel multiplier.

Furthermore, the generation of the noisy signals through the auditory nerve model requires 1 processing unit per channel. Again, if parallel arithmetic techniques were used, it would be possible to develop an auditory nerve model where 1 processing unit is required for the whole system.

The limitations of the system implementation investigation are that the design was not tested in real-time. Furthermore, we did not test the response of the system to a real

speech signal. The reason for not applying this type of stimulus is that it has not been applied in biological experiments.

8.6 SUGGESTIONS FOR FUTURE WORK

The limitations of the work that were highlighted in the previous section suggest that there is scope for a number of extensions to this work. This section will detail these and will suggest alternative systems that can be based on the designs developed in this thesis.

8.6.1 Extensions of Current Investigations

In terms of system architecture, a stellate soma processing unit should be developed to further reduce the complexity of stage 3 of the system. Such a development would perhaps mean that a single processing unit would be required to represent all 800 components in a channel.

Furthermore, the auditory nerve model should be developed such that a single processing unit can represent all the required noise injection stages.

Finally, a channel of the system should be implemented on array of FPGAs and allowed to run in real-time. This is the only way to prove the system speed capabilities. This would involve placing a number of programmed FPGAs on a PCB together with a clock generator.

8.6.2 Further Investigations

Further investigations that could be undertaken may involve finding ways of implementing the first 2 stages of the system to facilitate a digital cochlear implant.

This implant would be able to represent the spikes in the auditory nerve. Implementing the spike generation algorithm, described in Chapter 5, in parallel arithmetic can do this. Obviously, the use of such a circuit will require the connection (or splicing) of electrical connections with nerves. It is unclear whether such an operation is possible. What is clear is that this type of operation will become possible at sometime in the future. In order for this design to become a full system there must be analogue-to-digital circuitry at the input (together with a microphone) and digital-to-analogue circuitry at the output.

8.7 SUMMARY

This chapter has presented the main conclusions that were reached as a result of the investigations described in this thesis. This serves the function of relating the intermediate conclusions to form an overall impression of what has been achieved.

In addition to combining the results from the investigations, this chapter has also assessed the degree to which the investigations have addressed the aims and objectives set out in Chapter 1. The limitations imposed by time and facilities were set out thus enabling the scope of the conclusions to be determined.

Finally, this chapter has presented a summary of the contributions made by this research concluding with suggestions for work in this area that would benefit from future research investment.

REFERENCES

- [Ahn90] Ahn, S.G., Westerkamp, J.J. (1990). 'Cochlear modelling using a general purpose digital signal processor,' IEEE Proceedings of the National Aerospace and Electronics Conference, Vol 1, pp 57-63.
- [Ay97] Ay, S.U., Zeng, F.G., Sheu, B.J. (1997). 'Hearing with bionic ears,' IEEE Circuits and Devices Magazine, Vol. 13, No. 3, pp. 18-23.
- [Baldwin78] Baldwin, G.L., Morris, B.L., Fraser, D.B, Tretola, A.R. (1978). 'A modular high-speed serial pipeline multiplier for digital signal processing,' IEEE Journal of Solid-State Circuits, 13(3).
- [Beauvois94] Beauvois, M, Meddis, R. (1994). 'Computational modelling of auditory streaming phenomena,' Journal de Physique IV, Vol. 4, No. C5 Pt. 1, pp.371-374.
- [Bogli90] Bogli, H., Dillier, N., (1990). 'Digital speech processor for the nucleus 22-channel cochlear implant,' Proceedings of the Annual Conference on Engineering in Medicine and Biology, Vol.13, No.4., pp 1901-1902.
- [Denyer85] Denyer, P., Renshaw, D., (1985). *VLSI Signal Processing: A Bit-Serial Approach*, Addison-Wesley.
- [Fitzhugh91] Fitzhugh, R. (1961). 'Impulses and physiological states in theoretical models of nerve membrane,' Biophysical Journal, Vol.1, pp 445-466.
- [Ghitza88] Ghitza, O., (1988). 'Temporal non-place information in the auditory nerve firing patterns as a front-end for speech recognition in a noisy environment,' Journal of Phonetics, 16(1), pp 109-123.
- [Goldstein73] Goldstein, J.L. (1973). 'An optimum processor for the central formation of pitch of complex tones,' Journal of the Acoustical Society of America, 54(6), pp 1496--1516.
- [Giguere94] Guigere, C., Woodland, P.C. (1994). 'A computational model of the auditory periphery for speech and hearing research. I. Ascending path,' Journal of the Acoustical Society of America, 95(1), pp 331-342.

- [Hastings75] Hastings, N.A.J., Peacock, J.B. (1975). *Statistical distributions*, Butterworths.
- [He99] He, D.Z.Z., Dallos, P. (1999). 'Somatic stiffness of cochlear outer hair cells is voltage dependent.' *Proc. Natl. Acad. Sci. USA* 96 pp 8223-8228.
- [Hewitt91] Hewitt, M.J., Meddis, R. (1991). 'An evaluation of eight computer models of mammalian inner hair cell function,' *Journal of the Acoustical Society of America*, 90(2), pp 904-917.
- [Hewitt92] Hewitt, M.J., Meddis, R. and Shackleton, T.M. (1992). 'A computer model of a cochlear-nucleus stellate cell: Responses to amplitude modulated and pure tone stimuli,' *Journal of the Acoustical Society of America*, 91, pp 2096-2109.
- [Hewitt93] Hewitt, M.J., Meddis, R. (1993). 'Regularity of cochlear nucleus stellate cells: A computational modelling study,' *Journal of the Acoustical Society of America*, 93, pp 3390-3399.
- [Hewitt94] Hewitt, M.J., Meddis, R. (1994). 'A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus,' *J. Acoust. Soc. Am.* 95, pp 2145-2159.
- [Hodgkin52a] Hodgkin, A.L., Huxley, A.F. (1952). 'Currents carried by sodium and potassium ions through the membrane of the giant axon of loligo,' *Journal of Physiology*, 116, pp 449-472.
- [Hodgkin52b] Hodgkin, A.L., Huxley, A.F. (1952). 'A quantitative description of membrane current and its application to conduction and excitation in nerve,' *Journal of Physiology*, 117, pp 500-544.
- [Jackowoski95] Jackowoski, C.R., Vo, H.D.H., Lippmann, R.P. (1995). 'A comparison of signal processing front ends for automatic word recognition,' *IEEE Transactions on Speech and Audio Processing*, 3(4), pp 286-293.
- [Kaufman91] Kaufman, A., Gupta, M.M. (1991). *Introduction to fuzzy arithmetic: Theory and applications*, Van Nostrand Reinhold, New York.
- [Kuraishi84] Kuraishi, Y., Nakayama, K., Miyadera, K., Okamura, T. (1984). 'A single-chip 20-channel speech spectrum analyser using a multiplexed switched-capacitor filter bank,' *IEEE Journal on Solid State Circuits*, Vol. SC-19, No. 6, pp. 964-970.

- [Lazzaro89a] Lazzaro, J., Mead, C.M., (1989). 'A silicon model of auditory localisation,' *Neural Computation*, 1, pp 47-57.
- [Lazzaro89b] Lazzaro, J., Mead, C.M., (1989). 'Silicon modelling of pitch perception,' *Proceedings of the National Academy of Science*, 89, pp 9597-9601.
- [Lazzaro92] Lazzaro, J. (1992) 'Low-power silicon spiking neurons and axons,' *IEEE International Symposium on Circuits and Systems*, pp 2220-2223.
- [Licklider51] Licklider, J.C.R. (1951). 'A duplex theory of pitch perception,' *Experientia*, 7, pp 128-133.
- [Licklider59] Licklider, J.C.R. (1959). 'Three Auditory Theories,' in *Psychology: A Study of Science*, McGraw-Hill Book Company.
- [Linares91] Linares-Barranco, B., Sanchez-Sinercio, E., Rodriguez-Vasquez, A., Huertas, J.L. (1991). 'A CMOS Implementation of Fitzhugh-Nagumo neuron model,' *IEEE Journal of Solid State Circuits*, 26(7), pp 956-965.
- [Lyon88] Lyon, R.F., Mead, C.A. (1988). 'An analog electronic cochlea,' *IEEE Transactions on Acoustics Speech and Signal Processing*, 36, pp 1119-1134.
- [McCabe97] McCabe, S.L., Denham, M.J. (1997) 'A model of auditory streaming,' *Journal of the Acoustical Society of America*, Vol.101, No.3, pp.1611-1621.
- [McDermott91] McDermott, H. (1991). 'A custom-designed receiver-stimulator chip for an advanced multiple channel hearing prosthesis,' *IEEE Journal of Solid State Circuits*, Vol. 26, pp. 1161-1164.
- [MacGregor87] MacGregor, R.J. (1987). *Neural and Brain Modelling*, Academic Press Inc.
- [Mars81] Mars, P., Poppelbaum, W.J. (1981). 'Stochastic and deterministic averaging processors,' in *IEE Digital Electronics and Computing Series 1*, Institution of Electrical Engineers, London.
- [Meddis86] Meddis, R. (1986). 'Simulation of the auditory-neural transduction in the auditory receptor,' *Journal of the Acoustical Society of America*, 79, pp 702-711.

- [Meddis88] Meddis, R. (1988). 'Simulation of auditory-neural transduction: Further studies,' *Journal of the Acoustical Society of America*, 83, pp 1056-1063.
- [Morgan95] Morgan, N., Bourland, H., Greenberg, S., Hermansky, H., Wu, S.L., (1995). 'Stochastic perceptual models of speech,' *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol.1, pp.397-400.
- [Nandy96a] Nandy, D., Benarie, J. (1996) 'Auditory localization model-based on high frequency spectral cues,' *Annals of biomedical engineering*, Vol. 24, No. 6, pp 621-638.
- [Nandy96b] Nandy, D., Benarie, J. (1996) 'Estimating the azimuth from the binaural spectral amplitude,' *IEEE Transactions on Speech and Audio Processing*, Vol.4, No.1, pp.45-55.
- [Neti92] Neti, C., Young, E.D., Schneider, M.H. 'Neural network models of sound localisation based on directional filtering by the pinna,' *Journal of the Acoustical Society of America*, Vol. 92, No. 6, pp.3140-3156.
- [Oertel83] Oertel, D. (1983). 'Synaptic responses and electrical properties of cells in brain slices of the mouse anteroventral cochlear nucleus,' *Journal of Neuroscience*, 3(10), pp 2043-2053.
- [Oertel85] Oertel, D. (1985). 'Use of brain slices in the study of the auditory system: Spatial and temporal summation of synaptic inputs in the cells in the anteroventral cochlear nucleus of the mouse,' *Journal of the Acoustical Society of America*, 78(1), pp 328-333.
- [Patterson82] Patterson, R.D., Nimmo-Smith, I., Weber, D.L. and Milroy, R. (1982). 'The Deterioration of Hearing With Age: Frequency Selectivity, the Critical Ratio, the Audiogram and Speech Threshold,' *Journal of the Acoustical Society of America* 72, pp 1788-1803.
- [Patterson86] Patterson, R.D., Moore, B.C.J. (1986). 'Auditory filters and excitation patterns as representations of frequency resolution,' in *Frequency Selectivity in Hearing*, Academic Press, London, pp 123-177.
- [Pickles88] Pickles, J.O. (1988). *An Introduction to the Physiology of Hearing*, 2nd Edition, Academic Press.
- [Ross82] Ross, S. (1982). 'A model of the hair cell-primary fibre complex,' *Journal of the Acoustical Society of America*, 71, pp 926-941.

- [Schaik96a] van Schaik, A., Fragniere, E., Vittoz, E., (1996). 'Improved silicon cochlea using compatible lateral bipolar transistors,' in *Advances in Neural Information processing Systems 8*, MIT Press, pp. 671-677.
- [Schaik96b] van Schaik, A., Fragniere, E., Vittoz, E., (1996). 'An analogue electronic model of ventral cochlear nucleus neurons,' *Proceedings of MicroNeuro '96*, pp. 52-59.
- [Schaik97] van Schaik, A., Fragniere, E., Vittoz, E., (1997). 'A silicon model of amplitude modulation detection in the auditory brainstem,' in *Advances in Neural Information processing Systems 9*, MIT Press, pp. 741-747.
- [Schroeder74] Schroeder, M.R., and Hall, J.L. (1974). 'Model for mechanical to neural transduction in the auditory receptor,' *Journal of the Acoustical Society of America*, 55, pp 1055-1060.
- [Schwid82] Schwid, H.A., and Geisler, C.D. (1982). 'Multiple reservoir model of neurotransmitter release by cochlear inner hair cell,' *Journal of the Acoustical Society of America*, 72, pp 1435-1440.
- [Terhardt82] Terhardt, E., Stoll, G. and Seewann, M. (1982). 'Algorithm for extraction of pitch and pitch salience from complex tonal signals,' *Journal of the Acoustical Society of America*, 71(3), pp 679--688.
- [Tsui87] Tsui, F.F. (1987) *LSI/VLSI testability design*, McGraw-Hill Book company.
- [Zweig76] Zweig, G., Lipes, R., Pierce, J.R. (1976). 'The cochlear compromise,' *Journal of the Acoustical Society of America* 59(4), pp 975-982.

PUBLICATIONS

A.R. Temple, S.C. Lim, S. Jones and R. Meddis, '**Digital Realisation of the Mammalian Hair Cell**', MicroNeuro'97, Proceedings 6th International Conference on Microelectronics for Neural Networks, Evolutionary and Fuzzy Systems, University of Dresden Press, ISBN 3-86005-190-3, pp. 312-317, September 1997.

S.C. Lim, A.R. Temple, S. Jones and R. Meddis, '**VHDL-Based Design of Biologically-Inspired Pitch Detection System**', Proceedings of the IEEE International Conference on Neural Networks '97, IEEE Computer Society Press, pp 922-927, June 1997.

A.R. Temple, S.C. Lim, S. Jones, '**Digital Implementation of a Stellate Cell Using the MacGregor Model**', Proceedings of the Workshop on Design Methodologies for Signal Processing, Zakopane Poland, August 1996, pp 111-117, ISBN 83-904743-5-2.

Lim, S.C., Temple, A.R., Jones, S.R. (1997) '**Digital Hardware Implementation of a Neuromorphic Pitch Extraction System.**' Proceedings of the 1st European Workshop on Neuromorphic Systems, Stirling, Scotland, 29-31 August 1997.

