Chapter 1

INTRODUCTION

1.1 PROBLEM DEFINITION

The study of elephant vocalizations is an important part of elephant research (Garstang, 2004; Langbauer Jr, 2000). One type of elephant vocalization, the infrasonic rumble, is especially important because of the abundance of information that it contains. Recordings of elephant rumbles can be used by researchers to determine the size and composition of the herd, the sexual state of an elephant as well as the emotional condition of an elephant (Clemins, Johnson, Leong and Savage, 2005; Garstang, 2004; McComb, Reby, Baker, Moss and Sayialel, 2003; Langbauer Jr, 2000).

Some of the practical issues that researchers are faced with include the difficulty of acquiring high volumes of continuous high quality acoustic data in adverse environments as well as the tedious and time consuming task of identifying elephant rumbles within recordings (Clemins *et al.*, 2005; Langbauer Jr, 2000). The identification of rumbles in a recording is done manually by experts who inspect spectrograms and listen to sped-up versions of the recordings (Poole 1999). If, however, high volumes of data were to become available, research time spent on manual detection of rumbles would be hard to justify.

This dissertation investigates methods to assist elephant researchers by applying engineering principles to develop research tools to solve these problems. The aim of

this study is to provide solutions to both the problems that have been identified. Firstly, an instrument that can record high volumes of continuous acoustic data in unfavourable circumstances needed to be developed. This included the design of a robust but powerful electronic system that has very low power consumption and high volumes of permanent memory resources as well as a sound mechanical design that ensures that the instrument will survive for the required period of time. Secondly, an algorithm to automatically detect infrasonic rumbles from recordings needed to be developed. Speech processing techniques that can automatically detect the presence of human speech in a recording were modified to achieve the automatic detection of these elephant rumbles.

1.2 APPROACH

Obtaining high quality, continuous recordings of elephant vocalizations in the wild is a difficult but necessary task. One way in which recordings have been made in previous work is via a RF (Radio Frequency) transmitting collar which contains a microphone and a RF transmitter (Clemins *et al.*, 2005; Clemins and Johnson, 2003). The sound received by the microphone is then modulated onto the RF signal and transmitted to a receiver station. This method works well for recording vocalizations from captive elephants. In the wild, where elephants can move hundreds of kilometres in a few days, this method is unreliable. If the elephant moves too far away from the receiver the RF signal will be lost. Also, even if the elephant that wears the collar does not move very far from the receiver station, the topography of the area may result in degradation or loss of the RF signal. This can typically occur if the elephant walks into a deep valley and could result in sporadic loss of data as the elephants move through their territory.

An alternative would be to design a recording collar that stores the data on the collar itself. The concept of a recording collar that stores data onboard is attractive because the unreliable RF link is not used. Except for the problems that could occur when using an RF signal as was discussed above, another potential problem that may occur with a collar is clogging of the microphone with dirt. Elephants continuously cover their bodies with mud and dust. If the microphone gets obstructed it could cause loss of sound input to the system. The immense size and strength of an elephant, as well as their fondness of pushing trees out of the ground (Ogada, Gadd, Ostfeld,

Young and Keesing, 2008; Sankaran, Ratnam and Hanan, 2008), mean that the collar needs to be immune to intense physical abuse. Elephants also swim, so the collar needs to be completely watertight. It should also be remembered that wild elephants need to be tranquilized in order to mount a collar. This is an expensive procedure which needs approval from an ethics committee and is a stressful experience for the elephants. The collar should be able to store data for periods of time long enough to justify tranquilising of elephants.

Another vocalization recording method that has been used in the past is a hand-held recorder (McComb *et al.*, 2003; Wood, McCowan, Langbauer Jr, Viljoen and Hart, 2005). A specific herd of elephants was followed and recorded from a safe distance. The person who did the recording was notified of the herd's position each morning via a GPS (Global Positioning System) collar worn by the matriarch of the herd (Wood *et al.*, 2005). He then drove to the site indicated by the GPS coordinates after which he tracked the exact position of the elephants using an RF locator (also worn in a collar by an elephant in the herd).

The use of hand held microphones have several clear disadvantages. Firstly, hand-held recordings are not continuous because the person that operates the recorder needs to stop the recording every time that the herd moves along and then start recording from a new position. Secondly, these recordings are prone to noise since the microphone is located relatively far from the source of the elephant noises. In addition, this method requires the constant presence of a researcher in the field who conducts the recordings, and a GPS collar and RF collar need to be mounted on two of the elephants in the herd for them to be located.

The benefits of using an elephant recording collar therefore include the possibility of continuous recordings, a shorter distance between the source of the vocalizations and the microphone, as well as a reduction in unwanted sounds, as reported with far field hand-held recordings. The main choices that have to be made regarding the design of the collar are the kind of data storage device to use, how to ensure that the device is mechanically robust and watertight and how to mount the microphone so as to prevent damage or blockage without compromising sensitivity. The details of the design and motivations for design choices will be discussed in Chapter 3.

It will be shown in Chapter 2 that sound production of elephants resemble those

of humans and that this justifies the use of speech processing techniques used in human voice detection on elephant vocalizations. The problem of detecting elephant rumbles in recordings is similar to that of detecting human speech in recordings. Thus, identification of a suitable VAD (Voice Activity Detector) from this literature may assist in solving this problem.

VAD techniques that are usually used in a telecommunication context are not very robust (Wu, Wang and Brown, 2003). These techniques mostly use the energy level in a signal to decide whether a certain segment of sound was speech or non-speech since speech is normally the only source of energy in the signal. VAD of speech in noisy circumstances are more complex and some of the characteristics of speech need to be exploited in order to distinguish it from other high energy components present in the signal (Wu *et al.*, 2003). The harmonic nature of speech is one such characteristic that may be exploited by using a pitch detection algorithm. There are three basic types of pitch detection algorithms in use, namely frequency domain, time domain and time-frequency domain algorithms.

Just like human speech, infrasonic elephant rumbles also have a harmonic nature but in a lower frequency range (Clemins *et al.*, 2005; Clemins and Johnson, 2003; Langbauer Jr, Payne, Charif and Thomas, 1989; Langbauer Jr, 2000; McComb *et al.*, 2003). The elephant vocalizations itself will not be the only energetic components of a recorded signal. Various noises that occur in the wild will also be recorded, including birds, footsteps, breaking of branches and wind. Therefore, the simple energy threshold VAD technique will probably not suffice. The approach that has been taken for the automatic detection of the elephant rumbles is based on a robust time-frequency domain pitch determination algorithm which has been used specifically for VAD in personal audio recordings which also contains various kinds of noise. The details of the development of the automatic rumble detector will be discussed in Chapter 3.

1.3 HYPOTHESIS AND RESEARCH QUESTIONS

The hypothesis of this study is that speech processing techniques used for human voice can be adapted effectively to detect infrasonic elephant rumbles from noisy recordings.

All mammals produce sounds in a similar way. Specifically, the fact that elephants, just

like humans, use vocal cords to produce vocalizations suggests that techniques used to detect human speech in sound data can be modified to detect elephant rumbles from recordings. A suitable technique needs to be identified from the wide range of VAD algorithms available in the literature. One should be able to modify the chosen technique so that it can detect elephant rumbles.

To test this hypothesis, the development of a tool that can record substantial amounts of data is needed. It becomes part of the problem addressed in this study to develop such a tool. With the swift advances in flash card technology it should be possible to develop a physically robust recording collar that can store vast amounts of high quality acoustic data safely onboard the device. The equally impressive advances in low power electronics should ensure that the device can be powered for a long enough period of time from a relatively small and lightweight battery.

Specifically, the research questions investigated in this dissertation are:

1. What are the components needed for developing a recording tool that can withstand the harsh conditions on an elephant for months at a time?
2. What is the optimal way of mounting the microphone so that sensitivity is maximized, an even pass band is maintained and the chances of microphone blockage or damage is minimized while ensuring that the memory cards can be retrieved after a recording session?
3. What is a suitable speech processing technique that can be used as a basis for automatically detecting infrasonic elephant rumbles?
4. What modifications should be made to the chosen speech processing technique so that it may be applicable to elephant rumbles?
5. What is the accuracy with which the algorithm can detect rumbles and under which circumstances will it fail?

1.4 OBJECTIVES

The primary objective of this study is fourfold. Firstly, a tool that can record high volumes of continuous high quality acoustic data had to be developed. Secondly,

this instrument needed to be tested on actual elephants to evaluate its performance. Subsequently, an algorithm needed to be developed that can be used to automatically detect elephant rumbles within recordings. The last goal of the primary objective was to evaluate the algorithms developed and identify circumstances under which they would fail.

The primary objective mentioned above was achieved by completing the following steps:

1. Do a concept design of the electronic system
2. Focus on low power usage and robustness
3. Test each sub unit individually
4. Do a final electronic design integrating all sub units
5. Implement the electronic design in hardware
6. Mould the electronics hardware unit into a collar
7. Ensure that weight remains under the maximum level
8. Run tests to find an acceptable way in which to mount the microphone
9. Test correct operation under a range of temperatures
10. Test the collar on an elephant for extended periods of time
11. Identify a suitable VAD technique for use in this application
12. Adapt the chosen technique to detect elephant rumbles instead of human voice
13. Test the level of success of the algorithm using actual recordings from an elephant collar
14. Identify the conditions under which the detector will not produce correct results

The blocks enclosed by the dotted line in Figure 1.1 shows where the work done in this study fits in. The long term objective of this study is to provide a way of recording high volumes of elephant vocalizations and finding a way to automatically detect elephant rumbles within these recordings to enable researchers to do further signal processing on these isolated rumbles.

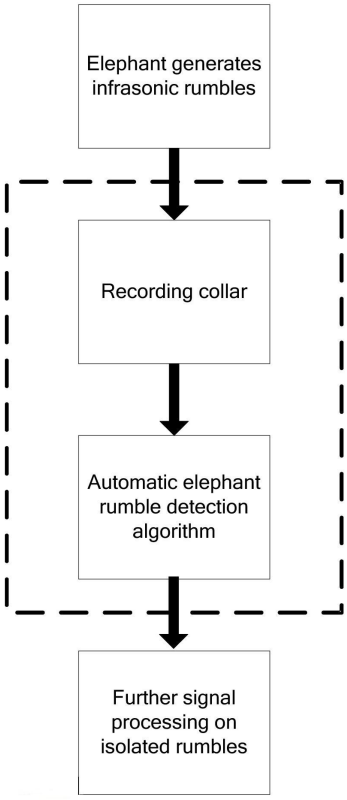

Figure 1.1: A diagram indicating where the current study fits into acoustic elephant research.

1.5 OVERVIEW OF THE STUDY

In the chapters that follow, the details of the development processes for the hardware tools as well as the development of the elephant rumble detection algorithm are discussed. An overview of the relevant literature is given in Chapter 2. This literature study motivates the importance of research on elephant rumbles and describes some of the methods currently used to obtain recordings of elephant vocalizations and the way in which these calls are isolated. The one known instance where speech processing has been applied to elephant vocalizations is also discussed. A short overview of different VAD techniques are given and the choice of technique for the current study is motivated from the literature.

Chapter 3 describes the methods used for the development of the research tools. The design and implementation aspect of the elephant recording collar are explained. The combination of electronic, mechanical, acoustic and ergonomical factors that need to be addressed are discussed in sequence. The electronic design is discussed and motivated after which the software that controls the instrument is explained on a functional level. The mechanical design of the collar is explained. The automatic detection of infrasonic elephant rumbles will then be considered. The details of every processing step of the detection technique are explained and the modifications that were necessary to apply the techniques to elephants are described.

The results from the elephant recording collar and the automatic elephant rumble detection (and pitch determination) algorithm are given in Chapter 4. The results of two field tests that were done with the elephant collar are discussed and the results of optimum microphone placement experiments are given. Further experiments that verify correct operation under certain environmental circumstances are presented. The performance of the automatic elephant rumble detection algorithm is evaluated and situations where the algorithm might not function correctly are identified.

The relevance of this study in context of the current literature is discussed in Chapter 5. The findings and implications of the different aspects in the study are reported and discussed. The contribution of this study to the current body of literature is pointed out. Chapter 6 states the conclusion that is drawn from the results that have been observed in this study. A summary of the work that has been done is given. Possible improvements to both the software and hardware that were developed in this

study are proposed and recommendations for further research are made.