Chapter 1

Introduction

The term *forensic audio* refers to audio material that may provide evidence in legal proceedings. The party introducing the recorded audio evidence must show that it has not been altered since the time of its production and this is achieved using standardised evidence-handling procedures and chain-of-custody records. When these safeguards fail, the reliability of the evidence may depend upon a technical analysis to prove the originality and integrity of the recording. Analysis techniques used for the authentication of analogue recordings have been developing over four decades, however, forensic audio authenticity examination is still a relatively unknown branch of forensic science.

Advances in recording technology have increased the opportunity to gather audio evidence by both the legal authorities and the public. Developments in digital signal processing hosted by powerful desktop personal computers have also increased the potential for economically and efficiently editing and manipulating recorded audio material. Audio evidence is therefore increasingly vulnerable to unauthorised modification. The digital audio recorder in its many guises has further compounded the consequent problem for the legal system by making the detection of falsified evidence produced by such equipment more difficult than its analogue counterpart.

The signal conditioning principles associated with digital audio recordings are completely different to those used for analogue recordings, requiring forensic audio engineers to develop new tools and techniques for assessing the integrity and authenticity of suspect recordings produced in the digital domain.

Nearly all the literature relating to techniques for the authentication of digital audio media is based on adding data to the audio signal or meta-data prior to recording. Two classes of method are common: watermarking, requiring authentication data to be embedded into the signal and fingerprinting, which allows a signature to be extracted from the signal relating to its content [1]. These techniques may be used separately or in combination [2] and are designed to produce tamperproof recordings. The research reported in the thesis is concerned with establishing the authenticity of an audio recording that has been produced on consumer based recording formats which are not considered tamperproof.

Forensic authenticity examination of digital audio media is a young field of forensic science and to date very little published research material has been available. An extensive literature search, including databases of the AES and IEEE produced only two papers concerning the forensic authenticity examination of digital audio recordings. One paper describes a new technique developed to determine the date and time of production of a digital audio recording, achieved by analysis of induced power line signals [3]. The second paper has been written by the author, and examines the significance of the Serial Copy Management System (SCMS) in the forensic analysis of digital audio recordings [4]. As there is little known previous work in the field, all of the experimentation reported in the thesis is original unless otherwise referenced.

With the breadth of available digital recording technologies, choices had to be made about the type of recording and recording systems that would be given priority for the research into authenticity investigation. Based on the author's experience of forensic audio laboratory work profiles, the decision was taken to concentrate the research on portable digital audio recording systems that store audio data that has not been perceptually encoded. The broad aims of the research were:

1. Establish reliable signatures for the recording process by examination of the recorded signal.
2. Using digital signal processing techniques, develop a robust method for the detection of digital copy recordings that have been produced by analogue interfacing.
3. Produce a straightforward method of analysis based on standard signal processing techniques that may be applied by forensic audio practitioners.
4. Having identified a method, establish the statistical attributes and inferences concerning the method.

The proposal is that a digital audio recording copied using analogue interfacing, can be detected using a single measurement parameter. The single parameter is derived from estimates of the magnitude transfer function of the recording system using only the audio data from the recording under investigation.

Under English law there are no particular legal requirements for a new forensic technique:

> "*Currently, a scientific technique does not have to pass any formal test in order for evidence derived from it to be allowed before a court. It is up to the judge at each trial to decide whether a particular piece of evidence can be admitted. This allows the courts to take advantage of the very latest developments in scientific knowledge*." [5]

However, experts presenting scientific evidence to the courts may be challenged on the basis of the techniques applied and the Crown has to be satisfied that the techniques in question are valid [6]. Similarly in the USA under rule 702:

> "*The testimony is the product of reliable principles and methods*" [7].

To assist the courts, simple techniques derived from the combination of long-established procedures that can be demonstrated in a straightforward manner are advantageous.

1.1 Synopsis

The thesis consists of 8 chapters, with chapters 3 to 7 containing the original contributions.

Chapter 2 discusses the background to the science and technology of forensic audio, concentrating mainly on the authenticity examination of recordings. The problems of detecting recordings that have been tampered with are reviewed and the motivation for the research is established by examining the significance of copy detection presented against a background of audio editing techniques. Finally, the chapter considers the use of analogue or digital interfacing in a copying process, with the conclusion that the analogue interface may be the more likely route chosen by a forger carrying out an editing/tampering process.

Chapter 3 develops a model for an uncompressed recording system that produces an original digital audio recording in the time domain. The model is then extended to a system capable of producing a copy recording, made by analogue interfacing two recording systems. Making a number of assumptions concerning the various signal and noise sources, the model is then reduced to elements consisting of acoustic signals, wide-band signal conditioning noise, gain stages and a series of low-pass filters.

The time domain model is transformed into a frequency domain model in chapter 4. Over the transition region of the low-pass response of the recording system, the assumption is made that energy in acoustic signals under most recording conditions are negligible, leaving signal conditioning noise only. The rate of change of attenuation over the transition region of the overall low-pass response of the recording system increases for an increase in the number of cascaded low-pass filters. In order to estimate these attenuation characteristics, the logarithm of an averaged power spectrum is proposed. To improve the performance of the estimator and to make the technique independent of the recording levels, a novel simple transform was devised and given the name "Reverse Difference Transform" (RDT). The transform normalises the data in order to remove differences between recordings due to different recording levels. This produces a result that is totally dependent on the spectral response of the recording system and independent of the acoustic signal. Further, the transform linearises the envelope of the normally non-linear response of the log-power spectrum and is of major benefit for estimation and comparison of original and copied recordings.

Having formed the basic requirements for the estimation technique, the signal processes are defined and analysed in detail in chapter 5. As the requirement was to establish the frequency response over a limited frequency region, the chirp transform has been investigated and compared to the fast Fourier transform. A power spectral estimation is made based on the well established method known as the periodogram. Logarithmic transformation of the power spectrum is then discussed and is shown to produce statistical advantages for the subsequent processes. The RDT is analysed in depth, both in its effects on the overall log-power spectral envelope and on individual RDT coefficient distributions. The RDT is shown to remove all even order and constant terms from the spectral envelope shape and produces a probability distribution for individual frequency coefficients having a symmetrical bell shape, similar to that of a Gaussian distribution.

In order to produce a single parameter to describe the transformed spectral response that indicates the provenance of the recording, linear and higher order regression techniques are investigated. The spectral signal conditioning achieved by the RDT is shown to satisfy the requirements when using linear regression techniques. Both methods of least squares and maximum likelihood are discussed as estimators for this single 'slope' parameter. The final estimation for the slope value is shown to be unbiased having a minimum variance dependent on both time domain data size and bandwidth of the spectral region of interest.

Chapter 6, identifies, analyses and offers potential solutions to problems that result in errors in the estimates. The sources of the errors can loosely be classified as: spectral leakage due to time domain segmentation of the signal, both short and long term acoustic signal contamination over the spectral region of interest, low recording levels, intra/inter recorder variability and aliasing.

The effects of spectral leakage are shown to be reduced to negligible levels by application of a high-pass filter to the data. Reduction of signal contamination over the spectral region of interest is achieved by non-linear spectral averaging and extrapolation techniques. Low recording levels may produce bias in the slope estimate and the robustness of the estimate due to a change in recording level is examined, along with a method of detecting if such bias is present for any particular recording. Finally, the effect of aliasing on the estimate is investigated, where it is found that aliasing reduces the estimated slope value used in assessing the provenance of the recording.

Chapter 7 provides empirical analysis in support of the assumptions and theoretical arguments put forward for the models, signals, systems and signal processes, presented in the previous chapters. The results of large scale testing on original and copied evidence is reported, and a case study presented for a particular recording format, providing data to estimate intra-recorder and inter-recorder slope variability. In the penultimate section of the chapter, the results of a large-scale blind test designed to establish if the recordings are original or some form of copy are presented.

Chapter 8 gives a summary and final conclusions regarding the principal results, identifies further work and proposes new areas of related research. An overview of the structure of the thesis is provided in fig 1.1.

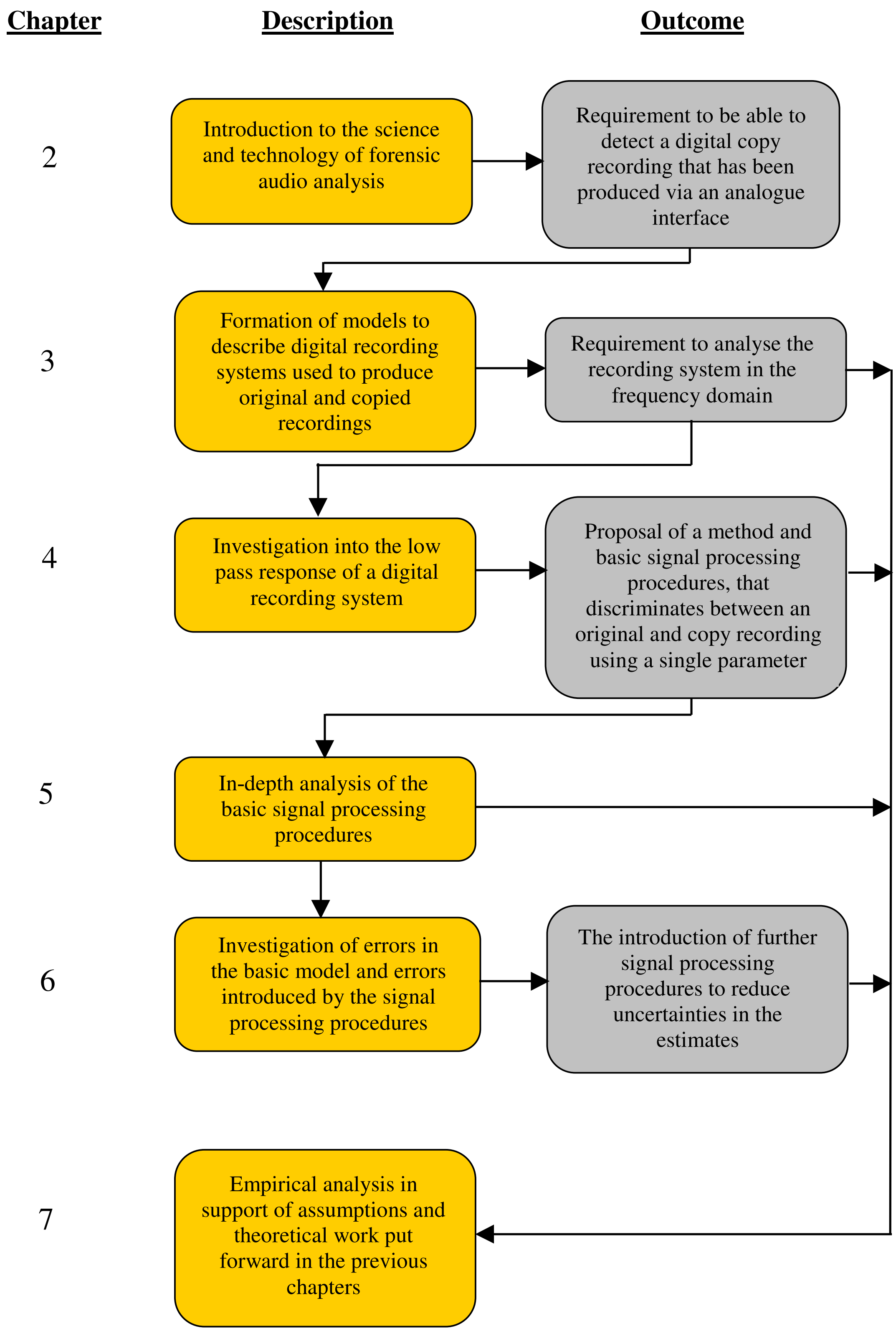

Figure 1.1: Overview of thesis structure.