Chapter 8

Conclusions

8.1 Summary and Conclusions

The issues concerning forensic audio analysis are broad and include scientific, technical, procedural and legal matters.

Forensic audio analysis techniques are well established for analogue recording technology. When these techniques are coupled with idiosyncrasies of individual recorders or recording systems, the integrity of a recording may be established. Advances in recording technology require a change in forensic techniques, the most significant advance has been the development of digital recording systems, where many of the analogue forensic analysis techniques are inappropriate.

The research aim was to provide a specific demonstration that a forensic digital recording is an original recording and not some form of copy. The requirement was to produce a technique that results in a simple indicator such as a single parameter value that is quick to calculate, statistically robust and provable to other experts in the forensic audio field. In achieving the latter the signal process should be founded on routinely applied and standard techniques, where the complexity of any step must not obscure the validity of the process.

As a result of potential interfacing problems confronting the forger, including copy protection, the analogue interface was considered likely to be the most common choice for getting signals in and out of an editing and copying system and this is where the research has been concentrated. However, finding chinks in the armour of a recording that has been copied by such interfacing, and that could be exploited by the forensic engineer in a reliable way has proved difficult.

Recording systems are designed to have a flat frequency response and there is little to distinguish an original recording process from a copied recording process in the digital field, except that imposed by the requirement for anti-aliasing filtering.

Copying via the analogue interface forms the starting point in chapter 3 for deriving a model of a recording system suitable for describing the process of generating original and copied recordings. The model is developed with the forensic audio recording environment as a reference and certain assumptions about the acoustic and electronic signal properties of such recordings have been made, resulting in a simplification of the final model proposed.

The hypothesis is that at high frequencies towards the Nyquist limit, the dominant signal is from electronic signal conditioning noise produced by the microphone pre-amplification circuits of the recorder. As the psd of this noise can be treated as white, the low-pass response of the original recorder over the high frequency region can be estimated. If the recording is subsequently copied, this high frequency response is modified and can be used as a basis for original or copied recording discrimination.

A simple novel transform (RDT) that produces a normalisation of the spectral response through zero dB magnitude at the centre frequency of the ROI, allows a visual comparison of recordings that have been made using completely different recording levels. The RDT also transforms the non-linear spectral envelope over the ROI into an approximately linear one and after the application of regression techniques, allows the envelope shape to be described by a single parameter which indicates original or copied recording status. Overall, the result of estimating the RDT slope is totally dependent on the spectral response of the recording system and independent of the acoustic signal and recording levels.

The basic signal processing procedure, defined and analysed in detail in chapter 5, consists of estimation of the averaged power spectrum, log transformation and application of the RDT. A number of signal-processing side effects arise from the analytical procedures. The chosen solution to reduce or eliminate these unwanted side effects produced from the signal processes is summarised in fig 8.1. Two important signal-processing requirements need to be met in order to reduce the effects of signal truncation on the estimate. The first is that a high-pass filter is to be applied prior to the FFT to stop the generation of spectral leakage from frequency components below the spectral ROI. The second is that the Fourier transform size must be high enough to reduce smearing caused by the width of the central lobe of the applied time domain window. Both leakage and smearing can result in a bias of the spectral envelope estimate. These solutions also allow the truncated time-domain data to be left un-weighted prior to the application of the power spectral estimation procedures. The unweighted or rectangular windowing approach is shown to produce the minimum MSE in the spectral estimate compared to other weighting windows. The MSE proves to be a function of main lobe bandwidth and this bandwidth is minimised using the rectangular window.

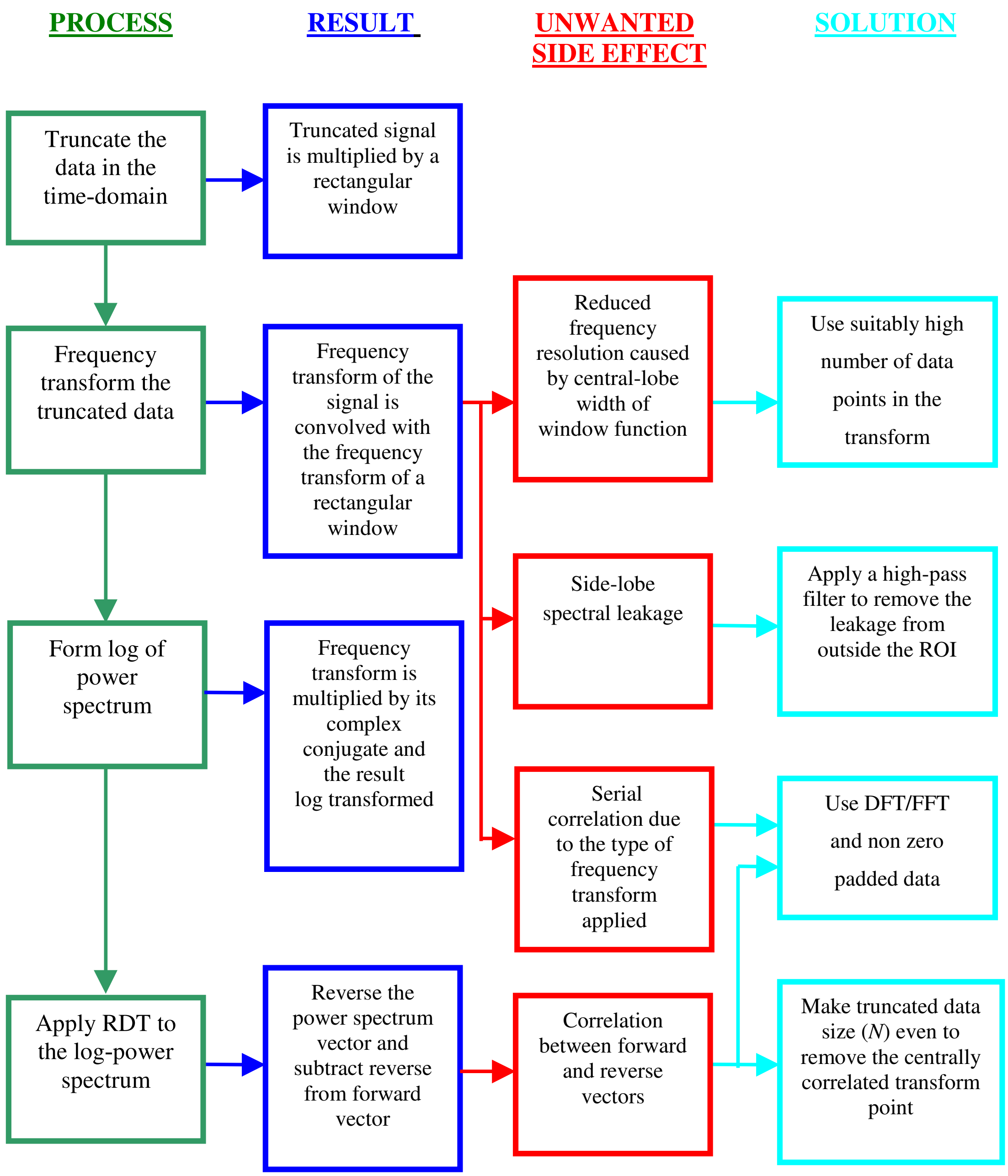

Fig 8.1: Problems and applied solutions to the basic signal processing procedure.

The overall model becomes inaccurate when the ROI is contaminated by high frequency spectral components of the acoustic signal, but additional signal processing can reduce the effects of both long and short time duration spectral contaminants. For shorter time spectral contamination, the contaminant is considered in terms of its probability of occurrence. By applying non-linear averaging techniques, specifically the log of the geometric mean, the spectral estimate can be weighted towards the underlying spectral response produced by the signal conditioning noise of the system, greatly reducing the effects of acoustic signal contamination. Bias of the envelope shape of the ROI caused by long term acoustic contamination has been shown from an experiment conducted on a large number of evidential recordings, to be effectively removed by an extrapolation and cancellation process. Other factors discussed in chapter 6, are robustness of the estimation process to relatively high additive noise levels introduced in a copying process, intra and inter recorder variability and the effects on the slope estimate from aliasing. It may be concluded that little can be done to stop the influence of these last three factors but it is important to understand the mechanisms by which they occur and the likely magnitude of their effects.

Empirical results have been reported in chapter 7 that validate the original and copied recording system models and theoretical arguments put forward. It may also be concluded from the empirical results, that the slope estimate for the RDT of the log of the power spectrum is not significantly affected when signal contamination occurs.

The empirical work resulted in recommendations concerning the parameters of the estimation procedure that influence the variance and bias in the final slope estimate and conclusions are valid only if these recommendations are followed.

8.2 Principal Conclusions

From the empirical work reported, a high discrimination exists between an original recording slope distribution and a copy recording slope distribution. This leads to a high degree of confidence in the final conclusion regarding the generation status of an unknown recording. However, relying on just the slope estimate alone to determine original or copied recording status is not enough to produce a correct assessment 100% of the time. Consideration has to be given to examination of the spectral response below, across and beyond the ROI in order to eliminate the possibility of false detection caused by unusual signal contaminants or relatively high additive noise levels, as indicated by the three negative results produced in the blind test described in chapter 7.

In order to make inferences concerning the original or copied status of an unknown recording, the make/model of the original recording machine or ideally the alleged original recording machine itself is required. This is in order to estimate an expected slope value for an original recording made on that equipment so that it can be used for comparison purposes. The thesis discusses the copying of recordings by using the analogue interface and makes the assumption that the copy recording would be produced at the same sampling rate as the original. This is a logical assumption based on the following:

a) The aim of the forger would be to keep the copy that has been tampered with as close to the original recording as possible.
b) Copying to a lower sampling rate reduces the bandwidth and therefore the potential quality of the copy recording compared to the original recording is reduced.
c) Other recording sampling rates may not be available for the format/recorder make model.

Two important conclusions can be made when the copy recording is produced at a different sampling rate. The first is that a copy produced at a higher sampling rate than the original is relatively straight forward to detect with simple spectral analysis, as the copy recording stop band of the low-pass response is now much wider, extending up to the new Nyquist limit. The second is that copying to a lower sampling rate would not be detectable, as the transition band of the original recording is effectively removed by the low-pass response of the copy recorder. A flowchart based on these conclusions, describing the decision process as to the status of a questioned recording is shown in fig 8.2.

Overall it may be concluded that a robust, efficient and relatively simple method of detecting copy recordings produced by analogue interfacing has been achieved. The possible sources of error on the RDT slope estimate in the form of increased variance and bias have been thoroughly investigated and their effects comprehensively analysed and modelled. A significant contribution has been made to the science of forensic audio, relating to the detection of copied and therefore potentially edited recordings.

8.3 Future Work

The motivation for further research into copy detection of forensic digital audio recordings will be provided by the increasing use of such recordings by both law enforcement agencies and the public. Challenges to the authenticity of evidence produced in this way will require forensic science to supply the criminal justice system with confidence that a recording has not been tampered with, and can be safely used as evidence.

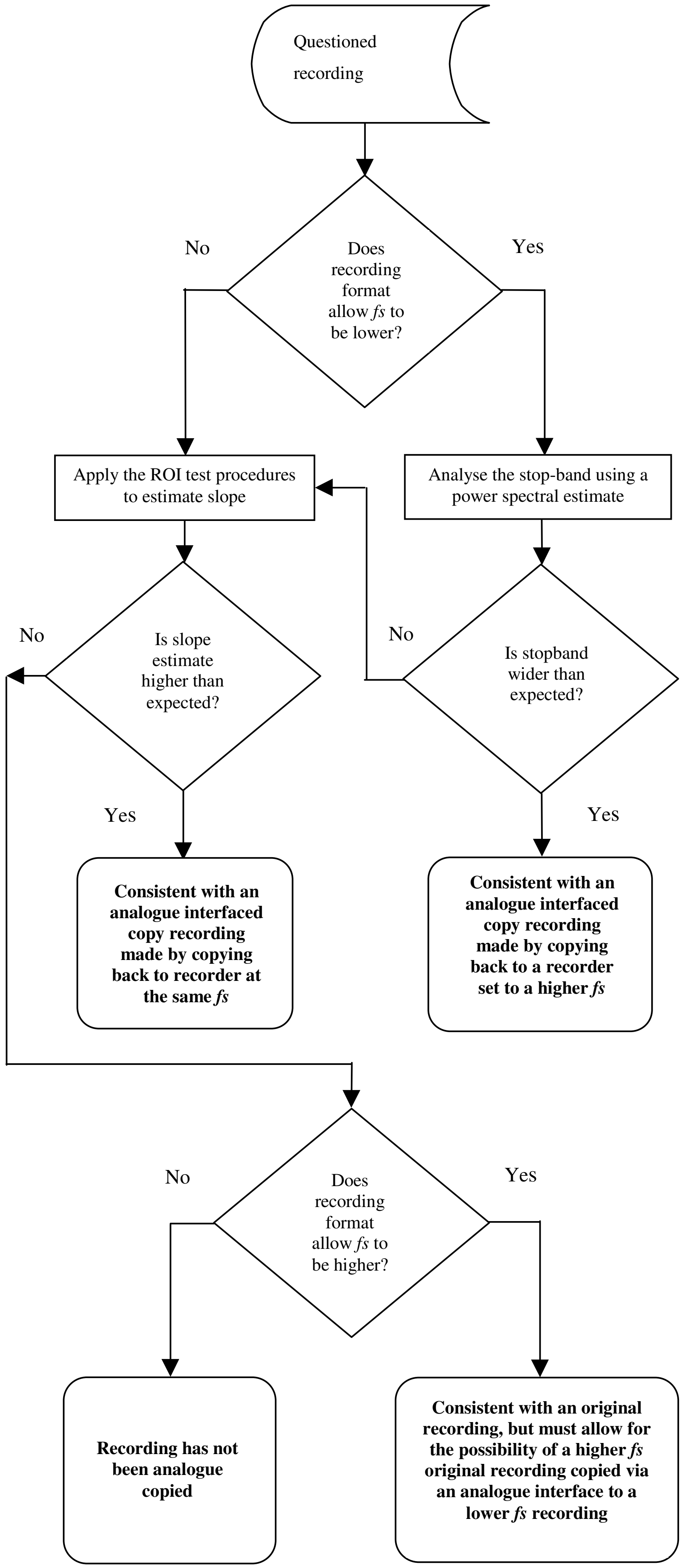

Fig 8.2: Decision process leading to the status of a questioned recording when copied sampling rates may be different to the original.

In the particular work reported in this thesis, low sampling rates below circa 24 kHz present particular problems due to masking of the spectral ROI of the low-pass response as lower frequency acoustic components can mask the spectral ROI. This acoustic contamination cannot usually be eliminated by the non-linear averaging or extrapolation techniques proposed for the higher sampling rates. An area of research that may provide a solution to this problem is based on examining short sections of a recording where little acoustic activity is taking place and for this limited region, conduct a spectral analysis based on a parametric approach, where it is known that superior results to the periodogram may be obtained when the data sample size is very small [126].

Clearly, an important area of research is in the detection of copy recordings produced entirely in the digital domain. This may seem an improbably achievable goal, as the result should produce copies that are bit for bit the same as the original data apart from any deleted, added and moved material that may result from an editing process. However, when a recording is copied to a computer based editing system and edited in a tampering process, often other signal processes are applied to the original signal. A forger may try to conceal tampering processes by using gain manipulation, amplitude compression/limiting and dc offset removal. These kinds of processes may be detectable by conducting statistical analysis on the pdf of the quantisation level distribution. Some research work related to this area, undertaken on commercial music produced onto Compact Disc has been reported by Knee and Hawksford [173] and Weida [174].

Further, research into the detection of copied material, produced by either analogue or digital interfacing needs to be considered for recordings produced using perceptual coding techniques such as MPEG audio formats [175] and the Adaptive Transform Acoustic Coding (ATRAC) system used for the Sony mini-Disc recorder [176]. Editing and copying of such data usually requires two stages of compression, once for the original and once again when the data is re-compressed when the edited copy is produced. Such a process is termed “tandem coding” [177]. The effects of tandem coding may be detectable.

In closure, forensic investigations into the integrity and authenticity of audio recordings are often centred on establishing that the recording is the original. Establishing the originality of a recording may rely on analysis techniques that exploit the idiosyncrasies of the underlying recording technology. As recorder technology evolves, the known idiosyncrasies disappear, leaving the task of the forensic engineer to find techniques to exploit a new set of idiosyncrasies that are associated with the new recording technology.