ESTABLISHING A STANDARD FOR DIGITAL AUDIO AUTHENTICITY:

A CRITICAL ANALYSIS OF TOOLS, METHODOLOGIES,

AND CHALLENGES

by

Daniel Lawn Rappaport

B.A., University of Virginia, 2000

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science, Recording Arts, Emphasis in Media Forensics

2012

This thesis for the Master of Science degree by

Daniel Lawn Rappaport

has been approved for the

Master of Science, Recording Arts, Emphasis in Media Forensics

by

Catalin Grigoras, Chair

Jeffrey M. Smith, Advisor

Lorne F. Bregitzer

Date April 27, 2012

Rappaport, Daniel, Lawn (M.S., Recording Arts, Emphasis in Media Forensics)

Establishing a Standard for Digital Audio Authenticity: A Critical Analysis of Tools, Methodologies, and Challenges

Thesis directed by Catalin Grigoras.

ABSTRACT

Audio recordings of acoustic events that include conversations, interviews, interrogations, wiretaps, surveillance situations, etc., are increasingly stored in a digital format. While digital recording has been possible for several decades, the technology is now ubiquitous in the form of recordable optical discs (CD-R, CD-RW, DVD, etc.), handheld digital voice recorders, digital voicemail services, mobile smart phones with voice notes/memo features, digital audio workstations (DAWs) that record to computer hard drives (hard disk or solid-state) , and other types of media that store binary audio information. Digital media offers many advantages over analogue, including better signal-to-noise ratio, fewer mechanical noises, longer battery life, and more, but it also makes proving authenticity much more challenging. This is particularly important when conducting forensic examinations involving criminal or civil cases. The fact that perfect clones are possible, combined with the relative ease of manipulating or editing digital recordings that sometimes defy aural detection, makes authentication a difficult process.

At present, no national or international standard exists for proving digital audio authenticity. There is neither a standard for defining “authentic” digital recordings (previous standards are for analogue), nor is there a standard means for presenting the findings of an expert. This lack of agreement is probably due to the difficulty in drawing generalizations in a discipline where every case is unique (to some degree), the technology is rapidly changing, and there is no examiner certification process or licensing board (aside from some advisory and technical committees with no regulatory authority). With some notable exceptions, relatively few articles and best practices guidelines have been published on digital audio authenticity. This paper examines some current and emerging techniques, to attempt to come closer to a reliable standard for digital audio authentication. Beyond the techniques themselves, which are sure to change, the paper critically analyzes the thought process that occurs during authenticity analysis: what the examiner learns (or does not learn) from each technique, the role that interpretation plays in guiding him to his ultimate conclusion, and the means of presenting findings in a clear, unbiased manner that do not overstate the scientific certainty of conclusions.

The form and content of this abstract are approved. I recommend its publication.

Approved: Catalin Grigoras

DEDICATION

I dedicate this work to President Richard Nixon whose infamous eighteen and one-half minute gap unwittingly inspired the field of audio forensics. I also dedicate this to my wife, Katie, for her incredible support and patience. I will start doing dishes again, I promise.

ACKNOWLEDGMENTS

I would like to thank my advisor Catalin Grigoras for his help and support with my thesis as well as Jeff Smith and the rest of my committee for their valuable comments. Also, I would like to thank my mother Connie and my stepfather Charles for their emotional and financial support. In addition, I am very grateful to Chris Buehler at the MITRE Corporation for his helpful insights and suggestions.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

AES	Audio Engineering Society
ASCII	American Standard Code for Information Interchange
DAW	Digital Audio Workstation
ENF	Electric Network Frequency
IOCE	International Organization on Computer Evidence
LTAS	Long Term Average Spectrum
MAC	Modified-Accessed-Created
NAS	National Academy of Sciences
NCMF	National Center for Media Forensics
ROI	Region(s) of Interest
SNR	Signal-to-noise Ratio
SWGDE	Scientific Working Group on Digital Evidence
VAR	Voice Activated Recording
VOR	Voice Operated Recording

CHAPTER I

INTRODUCTION

Background

From time to time, a sound recording has its authenticity brought into question, usually by one of the parties involved. With analogue recordings becoming increasingly obsolete, recordings in question are often on a digital medium. This can include portable digital audio recorders, voice mail machines or servers, CDs or DVDs, DATS or MiniDiscs, hard drives, flash memory, mobile phones or PDAs with voice memo or recording features, surveillance equipment, digital videos containing an audio track, and many other forms of digital media. Unlike analogue recordings where physical splices or cuts could be seen, felt, and often heard on the original tape, manipulations on digital media can be difficult to detect and sometimes impossible to hear (depending on the skill of the editor and recording conditions). They can also be fairly easy to make without requiring access to a lot of specialized tools or training.

Digital devices and computers are ubiquitous and inexpensive, so the barrier to entry for the unscrupulous person wishing to manipulate a recording is very low. Furthermore, digital audio workstation software (DAWs) that are designed for editing dialog or music are also plentiful, inexpensive, powerful, and somewhat easy to use. Training manuals, instructional videos, and courses in editing audio are accessible to anyone with an Internet connection seeking what was once considered specialized knowledge. One does not need a license or admission into a trade guild to learn these “trade secrets,” if they can even be called that. Articles on forensic audio techniques are even openly available to the public, thus perhaps giving inspiration to those wishing to carry out “anti-forensic” acts by exploiting known weaknesses. In short, there is little stopping someone from manipulating a digital audio recording and trying to pass it off as genuine, especially if the chance of detection is less than certain. Because of these factors, eliminating segments of the population as improbable or incapable of performing digital edits in not feasible and should not factor into an authenticity examiner’s conclusions (except possibly in circumstances when proprietary or non-standard equipment, unavailable to the consumer, was allegedly used).

Reasons for manipulation are many, but motives could include: political attacks, information warfare or propaganda, blackmail, defamation, disputes over inheritance (in the form of oral wills), liability, fraud, insurance claims, hiding police or professional misconduct (or framing it), documenting confession or innocence, proof/denial of domestic abuse, infidelity, harrassment, etc. Our entire legal system is based upon stories and the need to get the story straight in order to fairly dispense justice. [1] Possessing the tools and knowledge to manipulate stories by altering the perception of events, or actions, is a powerful force not to be overlooked. It will always be coveted and used by those members of society lacking moral fiber, and its increasing use and influence is regrettably part of living in the digital age. Seeing (or hearing) is no longer believing.

Throughout human history, forms of deception involving forged signatures and documents, copies of valuable art or artifacts, fool’s gold and counterfeit currency, doctored photographs, etc, were fairly common and created the need for authentication experts to authoritatively determine what was real and what was bogus. However, it could be argued that we are at a point now in time unlike any other, in that manipulated digital media appeals directly to our senses with more power and persuasion than ever before possible. Simulations are reality. [2] Because digital forgeries manipulate data instead of atoms, finding the "tell" (as in poker) by inspecting ones and zeros instead of physical evidence makes reliable authentication more difficult than ever. However, if we accept Locard's forensic exchange principal (in this case the "cyber exchange principal" [3]) that "every contact leaves a trace" to be true even in the digital age, then we need only to develop a sharper lens and keener methods in order to detect the ruse. As history can be seen as an arms race between the lock pickers and the lock makers, we on the side of science, honesty, law, and integrity must outsmart the criminals and deceivers. This is not a trivial challenge, for a lot is at stake that could affect what we hear and believe in our daily lives, not only through personal interactions but also in mass media. The ramifications of getting it wrong are enormous, to say the least.

Scope and Intent of Thesis

Although there is not a lot of literature on the subject, a few excellent articles have already been published on digital audio authenticity. The most notable example is the Audio Engineering Society (AES) article by Bruce Koenig and Douglas Lacey (Bek-Tek, LLC.) entitled "Forensic Authentication of Digital Audio Recordings." [4] I consider this paper to be a great achievement of scholarly work in this discipline and hold it in high regard, having read and studied it multiple times. This thesis is not meant to supplant that document, which should continue to be used as a reference in actual cases, and serves as one here. However, I wish to fill in the gaps, so to speak, and to attempt (perhaps foolishly) the not so insignificant challenge of explaining the interrelationships between disparate steps that are "too numerous and complex to be completely set forth." [4]

Empirical evidence is the foundation of science. While graphs, charts, and visualizations are useful, they often do not tell the whole story ("the expert's ultimate opining is likely to depend in some measure on experiential factors that transcend precise measurement and quantification.") [1] In addition to computational forensic tools, I wish to concentrate on the *thinking* behind the methodology and the interrelationships of the stages involved. I am less interested in examining individual analysis techniques, which have been previously published and are sure to change in the near future as new methods and tools are developed, although I will review many of them. I am more interested in what the forensic analyst *learns* from each analysis and how it informs his ultimate decision regarding the authenticity of the questioned recording.

Many of the scientific articles published on audio authenticity explain new tools that have been discovered, refined, or tested by the authors. The practical usage for these tools is left somewhat open ended because the authors know they can be employed in various ways and that many cases present different scenarios, requiring different approaches. A hammer can be used to drive in a nail or remove one, for example. The tool maker does not have much control over how his invention will be used, and may not even envision many ways in which it will be. This paper will attempt to show how an examiner can practically use some of the available tools, through a scientific method of deduction, to arrive at a decision regarding the authenticity of digital audio recordings.

CHAPTER II

PERSPECTIVE

Science and Uncertainty

Many fields employ some form of authentication process. Antiques, works of art, currency, precious metals, documents, photos, sound recordings, etc., need to be authenticated if their genuineness is called into question. There are perhaps cases when an object or thing can be empirically evaluated as undoubtedly authentic (by verifying the presence of a maker's mark or hidden seal, or through chemical testing, for instance), but more commonly an authentication ultimately comes down to the *opinion* of the expert conducting the examination.

It has been said that in science it is impossible to prove a hypothesis. One can only disprove alternative hypotheses and eventually accept the original hypothesis by default if no other explanation can be found. This is true in science as well as in the process of authentication. While it is often possible to prove 100% that an object (or thing) is a forgery, and is therefore "inauthentic," proving the opposite is much less certain. Albert Einstein once said: "No amount of experimentation can ever prove me right; a single experiment can prove me wrong." This concept is based on the notion of falsification, which was developed by philosopher Karl Popper. For statements or claims to be scientific, they must be falsifiable. [5] This does not mean they are necessarily false, it simply means that they can be refuted if a single unexplainable contradiction is found (not an anomaly, but an observation or experiment that yields a repeatable contradictory outcome).

Contradictions can be separated into explainable and unexplainable differences. For example, a falsifiable claim would be the following: “All polar bears are white.” If a scientist investigated the world’s polar bear population and found thousands of examples of white polar bears, but not a single non-white one, then he could say that the original claim is true (to a degree of scientific certainty). However, if a single black polar bear were discovered, the claim would be false or would have to be modified. Granted, if the black polar bear is examined more closely and is seen covered in oil, then that would be an explainable difference. The original claim does not need to take into account explainable differences and it would still be true. However, if unexplainable differences are discovered (such as a natural population of purple polar bears) then the claim is false or is said to have been refuted. A non-falsifiable claim would be something where the conclusion is not based on a categorical proposition. Examples would be a belief or singular observation. “The sound recording is inauthentic because I believe I hear an edit,” for example, is not falsifiable because it cannot be proved or disproved (the examiner’s hearing is subjective). Conversely, the assertion: “An edited recording will show signs of (re)compression, ENF discontinuities, or metadata inconsistencies” is a falsifiable claim, because contrary examples (that are not anomalies) may be found that refute it.

Popper proposed the idea of falsification as a means of demarcating science from non-science. If a theory is scientific, it is falsifiable. If a theory is non-scientific, it is not falsifiable. (It should be noted that non-falsifiable claims can still be meaningful.) Not everyone agrees with this definition of science, but is has been used in U.S. courts in concluding that “Creation Science” was not science (McLean v. Arkansas) and thus could not be taught in public schools. Therefore, falsification could be used as a criterion for admissibility in future court rulings and forensic practitioners should take note. As forensic audio experts, we must be careful not to put too much emphasis into analyses that yield results which are non-falsifiable. Results of analysis that produce figures and visualizations that require interpretation may not be entirely empirical, even though they are technical in nature. Forensic audio experts should strive to use falsifiable techniques whenever possible, to independently test the theories and assumptions behind those techniques to see if they are valid, and establish error rates. This is not as easy task in a field with such a brief history, with relatively few practitioners who often work in small isolated labs, and where there is little research funding or agreed upon standards.

While there are a number of peer-reviewed techniques and tools used in digital audio authentication, there is not yet an established methodology that is generally accepted in the field. Other forensic sciences such as fingerprint examination and DNA analysis developed their methodologies over time with the influence of new research, public comment, and refinements. However, they eventually arrived at a standard with the expectation that two different examiners in separate labs in different parts world can analyze the same evidence and obtain similar results and conclusions. This type of standard currently does not exist in digital audio authentication. We cannot expect one to be adopted over night, but hopefully the process will begin moving in the right direction, perhaps with the aid of this paper.

A forensic digital audio authentication case could include all of the analyses stages explored in this thesis, or only a portion of them. The number of stages employed, and the order, depends largely on the particulars of the case at hand and the preference of the examiner. However, I will argue that there are certain key decision points that should be followed in order to arrive at a sound conclusion. In addition, many stages employ a feedback loop in which results from one analysis cause the examiner to return to a previous one.

When devising a methodology, the examiner must always be aware of what forensic question needs to be answered [6]. This will determine which techniques should be used, possibly in what order, and how accurately the results can be determined and interpreted. While contextual bias must be avoided, at the same time the context of the case and the claims of the recording operator will in some part determine the methods used by the examiner and may affect the ultimate conclusion. However, it cannot be assumed that the methods of recording claimed by the operator are true. They must be verified by testing, and other potential explanations must be either accepted as viable alternatives or rejected. [7]

Analogies

The following sections will introduce some analogies that compare audio forensics to other disciplines or sciences. While there is a danger in straying away from one’s own field of expertise and perhaps comparing apples to oranges, it is also true that most people learn through the device of analogy by relating the known to the unknown. Even sound waves themselves (if reproduced through a speaker) are analogous to variations in electrical energy which in turn are analogous to the initial acoustic vibrations that the recording device captured: hence the term analogue. As forensic audio experts, we are required to explain our findings to the court consisting of lay jury members who likely have no background in our field and whose only exposure to it is through misrepresentations in the media (the so-called “CSI Effect”). Analogies are useful in this role of facilitating communication. But analogy alone is not perfect because two concepts or things, when compared closely, will always exhibit some fundamental differences that make the analogy fall apart. A more useful technique is to combine analogy and contrast. Physicist Richard Feynman used this device to describe quantum mechanics, which cannot be understood through direct experience because it is unlike anything else available to our senses. [8] If we only use analogy, we will fail. So we start with things we are familiar with and then use contrast to show how this new phenomena differs from what we understand. For example, a digital photographic image file has been compared to a can of soup. [9] The analogy is that both things have a container, contents, and an ingredients label (or metadata) with information about the product. However, an image file is not like a traditional soup can, but rather one that has nothing on the label besides a file name, extension, bar code, and thumbnail image. Unlike the exterior of a real soup can, there is little meaningful language describing the contents that can be read from the label of the “digital soup.” This information must be decoded by a hexadecimal viewer. Furthermore, while a physical label is fairly stable, the metadata is fragile and some of it may change (such as MAC times) depending on the operating system used to view it. There are countless other differences as well as similarities. Therefore, a digital image file is both like a can a soup and not like one. This technique of combining analogy and contrast is often used in science and is occasionally applied in this paper.

Container and Contents Analysis Framework

We can separate the task of digital audio authenticity into two main categories: *container analysis* and *content analysis*. The container relates to the file name, file structure, metadata (either stored in the file itself or generated by a software or operating system), etc. Content analysis relates to the bits and bytes that make up the audio portion of the file and the acoustic events that it reproduces. Renaming of the container may not necessarily affect the integrity of the contents, but may alter and/or damage the media support or wrapper. This would raise doubts about authenticity that require explanation and may make some types of analyses inconclusive. In conceptual terms, a container and its contents are often thought of as coming together in a pair, but in fact they are separate entities. In terms of logic and set theory, there are several possible combinations to consider: 1) authentic container/authentic contents; 2) authentic container/inauthentic contents; 3) inauthentic container/authentic contents; or 4) inauthentic container/inauthentic contents. The third option can be split further into mislabeled container, intentional or unintentional modification, etc. Just opening and saving a file may change the metadata and cause values to differ from previous checksums. Therefore, many recordings with authentic contents will have a compromised container. In authentication examinations, we must be clear if our ultimate conclusion is stated in terms of the *whole* media package (contents and container) or just the contents, which is really the fundamental issue being examined. In media forensics, it is understood that a container and its contents are separate entities that do not always stay attached to one another, even though they begin that way when first stored into memory.

Digital Audio Concepts

The differences between analogue and digital recording have been written about countless number of times. There is no point in repeating it here. This thesis deals only with digital audio authentication. Readers seeking information on authentication of analogue recordings should consult other references. Even while staying in the digital domain, however, there are a few important distinctions that should be made clear. First of all, an analogue recording can only have one master (except in somewhat rare circumstances when there are multiple masters created simultaneously on separate devices). Because creating copies from an analogue master requires transduction (i.e. converting magnetic flux to electrical energy and back again to magnetic flux) there will be a loss of quality in all subsequent generations. Therefore, it is vital to perform authentication analysis on the analogue master to ensure that transient events, low level signals (close to the noise floor), and characteristics of the recording system (such as wow and flutter) are preserved as accurately as possible without the influence of a secondary recording device. In addition, if there are any physical edits involving cuts and splices, or tampering with the reels or cassette housing, they will only be visible on the master tape. The same applies for mechanical impressions made by the erase and record heads, which are not transferred to copies. Electric Network Frequency (ENF) analyses are severely limited when conducted on an analogue tape, as opposed to digital recordings, because the inherent wow and flutter obscures the 60 Hz cycle (50 Hz in Europe). There is also no metadata on a traditional analogue tape, so that type of analysis only applies to digital audio authentication.

Because digital recordings are made up of discrete binary information, they are repeatable. Instead of transferring energy from master to copy as in the analogue domain, one transfers information in the digital domain. This information is perfectly repeatable so it produces copies that can be indistinguishable from the master (if done correctly). In the digital world we can substitute the word “clone” for “copy.” The existence of multiple digital masters or clones is a paradigm shift away from analogue recordings and requires a different approach to forensic authentication. To help illustrate this point, imagine a simple scenario involving a digital audio file and two identical make and model Secure Digital (SD) cards. The cards are indistinguishable and are the same in terms of appearance, storage capacity, format, etc. One card contains an original file that was recorded to it directly, which we will call the master. An exact bit-for-bit clone of this file is then placed onto the second card. A neutral party takes the two cards and plays a “shell game” so that even the person who made the recording and the clone has no idea which one is which. Since there are no distinguishing features on the cards themselves to tell them apart, and the files are exactly the same down to the smallest bit, there is no person or machine in the world that can now say which file is original and which is the clone. Therefore, for all intents and purposes, we can accept either one as the original and use it for forensic authentication purposes. It must be verified however, as part of evidence handling, that the two files are indeed forensic clones with matching checksums.

CHAPTER III

DEFINING AUTHENTICITY AND LACK OF STANDARDS

Historical

The Oxford English Dictionary defines “authentication” as the act of “establish[ing] the claims of (anything) to a particular character or authorship; to establish the genuineness of; to certify the authorship of.” Also: “To give legal validity to; to render valid, establish the validity of.” [10] The term “forensic digital audio authentication” therefore combines the classic definition of authentication (in a broad sense) with “forensic” and “digital audio.” To the layperson, it is worth explaining that “digital audio” refers to reproduced sound that is heard through a digitized representation of discrete samples: analogous to the way that fluid motion is represented as separate still frames in motion pictures. It should also be noted that the word “forensics” comes from the Latin word *forensis* whose root means “forum.” [11] In ancient Rome, the rhetorical skill of arguing was practiced in the forum and at times objects were given the chance to address the gathering. “Because they do not speak for themselves, there is a need for something like translation or interpretation. A person or a technology must mediate between the object and the forum, to present it and tell its story.” This concept can be applied not only to objects or evidence, but also to the findings (or work product) of a forensic analyst. The output of some forensic analyses can be presented to a jury or fact finder as raw data, such as face comparisons or magnified fingerprints, without much need for translation. It can be expected that a layperson will understand the material, to some extent, and make a fair decision based upon on it. In audio forensics, playing voice samples of a known and unknown recording for the jury to decide if it is the same person would be an example of where the evidence “speaks for itself” to some extent. However, other types of analyses used in audio forensics, and specifically in authentication, are more complex and would not normally be understood to a person without specialized experience. These types of analyses require more interpretation and presentation on the part of the forensic audio practitioner to act as the mediator between the evidence and the forum (i.e. the judge, jury, or fact finder).

Multiple Goals and Definitions

The first goal of digital audio authentication is to define what it means. Without a solid definition, each examiner will have a different idea of what they are looking for and a different threshold for reliably when designating a recording “authentic.” I will try to make sense of the various definitions of forensic digital audio authentication (when they are found), and the different methodologies that are being practiced in the field. The aim is to arrive closer to a standard methodology that is more clear, reliable, and adheres to the principles outlined in the 2009 National Academy of Sciences (NAS) report: “Strengthening Forensic Science in the United States: A Path Forward.” [12] Of course in actual casework, some variations to the method will arise due to the unique nature of particular cases or the legal system that the analyst is working under (in the U.S. we apply the Daubert and Frye standards).

The goal of digital audio authentication is to establish if the questioned recording is original, continuous, and unaltered. [4] The process tries to determine whether or not an unknown file is consistent with exemplar recordings made with the equipment and methods claimed by the party who submitted it. If there are any discrepancies, these must be explainable by means other than intentional alterations or else the recording is inauthentic. Eddy Brixen states that forensic audio authentication is generally understood as “every kind of documentation that can verify the origin and the contents of the audio recording and the recording media in question.” [13]

The goals of scientific truth and legal truth in the courtroom are not always the same. Durand Begault said that there is a difference between legal and technical authentication:

> *Legal authentication involves the need for proof that the material is original and is what it claims to be, whereas technical authentication involves claims of likelihood that the characteristics of a recording are consistent with those of an original recording. There is almost never 100 percent certainty that recorded material is original and unaltered. It is necessary to distinguish between intentional and unintentional modification, and it is sometimes difficult to distinguish between the two.* [14]

This paper will look at digital audio authentication only from a scientific view as taught at the National Center for Media Forensics (NCMF). Different legal systems can interpret these views according to their own standards, but by limiting the discussion to the science and logic of authentication, the methodology will have application beyond that of a single legal system. Digital audio authentication is a complex process of establishing the provenance of a questioned recording to determine whether it is consistent with an original one or if there is evidence of tampering. Since there is no “black box” that will reliably determine if a file is authentic or inauthentic, the examiner himself is the instrument of authentication. He must therefore apply scientific methodology, computational tools, experience, and logical interpretation to render an expert opinion regarding authenticity. In 1898 Thayer wrote: “In a sense all testimony as to matters of fact is opinion evidence, i.e. it is a conclusion from phenomena and mental impression.” [15] He went on to say: “As most language embodies inferences of some kind, it is not possible wholly to dissociate statement of opinion from statement of fact. All statements are in some measure inferences from experience." In order to arrive at that opinion, the general methodologies used in audio authentication arise from learning all the capabilities of the device on which the recording was made and using tools based on verification or reverse engineering the file format, in order to explain how the file could have arrived at its current state. Statistical tools that compare time windows or other content within an evidence file to itself, or to exemplars or database recordings, rely on the known characteristics of the file format: how it handles, organizes, and stores data. Although acoustic information is almost infinitely variable, the processes of digitization and binary storage impart a logical structure onto the file that is altered when edits are performed. Attempts at obfuscating these changes often result in even greater changes to the file. Although not always audible, the metadata or other parts of the file will exhibit traces of tampering, transcoding, or re-saving. Many of these changes will be permanently embedded in the file, even if later upsampled to a higher resolution. This will not necessarily be shown in the file format (such as a submitted WAV PCM file that was originally an MP3), but it can be revealed in the analysis of the content.

Need for Clarity and Standards

This paper will attempt to answer a simple question: how does an expert perform an examination for digital audio authenticity? It is a seemingly easy question with a not so easy answer. In fact, very little has been written on the subject, even though courts throughout the world admit experts to testify on this very matter. There is currently no standard procedure for establishing digital audio authenticity or even a clear definition of what is meant by an "authentic digital recording." This is a serious problem that needs to be clarified as soon as possible, but will require a meeting of the minds between practicing audio experts, professional societies, the Scientific Working Group on Digital Evidence (SWGDE), and the courts (both domestic and international). Without a standard, every audio expert or lab is basically operating as a "lone wolf" and will have different criteria for defining authenticity, different methods of examination and interpretation, and possibly different ways of presenting their findings. The AES definition for analogue recordings is the best we have at the present, so let us start there. The standards document, "AES Recommended Practice for Forensic Purposes- Managing Recorded Audio Materials Intended for Examination" (AES27-1996 r2007) provides the following definitions:

Authentic Recording- "…a recording made simultaneously with the acoustic events it purports to have recorded, and in a manner fully and completely consistent with the method of recording claimed by the party who produced the recording; a recording free from unexplained artifacts, alterations, deletions, or edits."

Accurate Recording- "…an original recording (or a copy of an original recording) of acoustic events that is, consistent with the limitations of the recording method, fully, exactly, and completely faithful to the original events with respect to time-varying amplitude, sequence, and completeness of program material."

Authenticity Analysis- "An examination, usually for forensic purposes, that seeks to determine whether a given recording was made of the acoustic events asserted by the parties who produced the recording, and in the manner claimed by the parties who produced the recording, and whether it is an original or copy." [16]

When applied to digital audio, the analogue definitions above are problematic in many ways. First of all, it may be implied that an “accurate” recording is part of the criteria for an “authentic” recording, since the relationship of the adjacent terms is not specified in the AES document. Other definitions of audio authenticity clearly do put the two terms together, such as the following: “Commonly, a review of audio authenticity is requested to determine whether the material presented *accurately* (emphasis added) represents the events recorded...” [17] These definitions require clarification. In theory, a recording that was made with the microphone unplugged or the gain turned all the way down would produce a continuous, unaltered, original recording. However, it would not be an accurate representation of events, and therefore it being designated as authentic depends on the inclusion or exclusion of the word “accurate.” A recording that suffers from excessive noise and unintelligibility faces the same question. If we accept that an authentic recording must be an accurate representation of the events that it purports to portray, then we must examine the meaning of the phrase: “completely faithful to the original events with respect to...completeness of program material.” It could be argued that a recording containing *any* stops/starts of any duration would not be a complete and accurate record of the original acoustic events. A documentation of an acoustic event containing starts and stops made for any reason is basically equivalent to a video containing dropped frames, or a still image filled with “swiss cheese” holes in the scene content. When the device stops it creates a gap in the time continuum of the events and is no longer complete. We do not know what we do not know. The listener must interpolate the missing information based on what came before and after. The gap could represent two seconds, two minutes, two hours, or two days. The concatenated segments may be contiguous, but they are not continuous. Even if a reliable independent time stamp, such as ENF, can verify that the missing gap is very short, there is no way to guarantee that the missing information is insignificant to the documentation and understanding of the events.

For example, in the case of INCARNATI v. SAVAGE, a jury's verdict was reversed on appeal and a new trial granted because of a stenographic transcription error in a physician's deposition. [18] The error was the typing of the word "inconsistent" when the doctor actually said "consistent." If this were an audio recording, the difference between adding or removing the syllable "in" is only a few milliseconds. It takes only that long to invert the meaning of a statement. Therefore, a recording containing any gaps for any reason or duration may have its accuracy and authenticity called into question, depending on which definition is accepted. It may be preferable to separate the word "accurate" from the definition of authenticity so that there can be two categories of authentic recordings: A) authentic and accurate; or B) authentic and inaccurate. The second classification would be a perfect fit for unaltered recordings that contain discontinuities introduced in real-time without human involvement, such as VOR pauses, dropouts, packet losses, RF interference, etc. Unaltered recordings containing mechanical stop/starts (but not edits) or hoaxes (actors reciting dialogue) that are authentic, but not complete or accurate portrayals of events, may also fit this designation.

Furthermore, the inclusion of the word "original" in definitions of authentic digital audio recordings should be clarified. It is generally accepted that authenticity examinations on analogue recordings have to be conducted on the original. However, because digital media is so easily replicated, there is the potential for confusion and experimental error if the wrong questions are asked or an improper file (i.e. non-forensic copy/clone) is submitted for analysis. While the original recording should be submitted, it should *not* be used directly for digital audio authentication analysis, since it should be preserved (which is a departure from analogue practices). Only a forensically verified clone should be used. The SWGDE document entitled, “Best Practices for Forensic Audio,” uses the word “original” but leaves many questions unanswered:

> *An audio authentication examination seeks to determine if a recording is original, unaltered or continuous, and/or consistent with the manner in which it is alleged to have been produced. Consequently, there is no catch-all means of declaring a recording “authentic” without having a clear understanding of what claims the creator holds true about its nature and what specific allegations are being levied against the recording. SWGDE recognizes that audio authentication is a complex examination that requires specific training. Best practices for this type of examination are reserved for a future document.”* [19]

In summary, digital audio authentication requires its own definition that is separate from previous definitions pertaining to analogue audio. It is hoped that such a definition can be drafted, posted for public comment, agreed upon and accepted by the forensic community.

CHAPTER IV

PYRAMID OF PHYSICAL EVIDENCE: OCCURRENCE, RECOVERY, ANALYSIS, INTERPRETATION AND PRESENTATION

There are five basic stages of forensic science (see figure 4.1): 1) occurrence of crime (or acoustical event); 2) seizing of evidence; 3) analysis; 4) interpretation; and 5) presentation. [15]

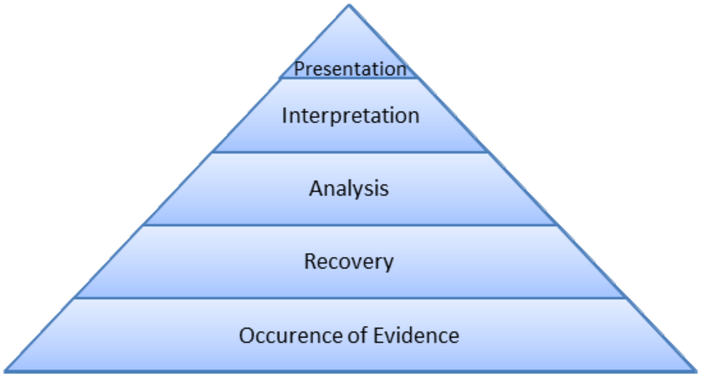


Figure 4.1 The Stages of Physical Evidence Process. [15]

This paper will focus mostly on the analysis, interpretation, and presentation stages, and more specifically on modern techniques and the methodology of applying them to a variety of cases. While evidence handling and general housekeeping procedures are vitally important to an authenticity examination, these topics have been well covered in previous papers [4] and will only be discussed here superficially. The forensic examiner has no control over the first stage. Typically, the second stage is performed by a first responder, law enforcement agent, or technician. The examiner does have control over the third, fourth, and fifth stages and should check to ensure that the evidence was collected properly, with nothing being added, lost, damaged, or contaminated in stage two. Evidence should be safely and securely packaged as soon as possible (anti-static bags are recommended for electronics), as well as photographed and logged. Access should be controlled and recorded so that chain-of-custody is maintained. While not always applicable to digital evidence, crime scenes and health and safety measures must be taken when collecting and transporting evidence. If a backup or working copy is made, which could occur in the second or third stage (recovery or analysis), it is vital to use a write blocker (with updated firmware) to ensure that metadata is not changed and that files are not accidentally written from the workstation computer to the evidence. Copies should be made only via a forensic disk image, which creates a bit-for-bit clone of the evidence. In addition, a hashing algorithm should be used, which creates a unique character string signature based on the data input of the file, folder, disk, or drive. Since altering a single bit of data will change the outcome of a hashing function, hash values of the evidence and working copies should match, and it should be verified that the image is mountable and not encrypted or damaged. Since some hash functions have been compromised, at least two robust algorithms should be used and the results stored in a hash history. Hashes should be calculated each time the evidence is copied, exported, or bit stream imaged to ensure that the evidence is exactly the same between the present and the time of the last hash. In addition to keeping the evidence secure, this process also allows independent examiners to repeat or reproduce the same analyses, which is only possible if the raw evidence is unchanged. Unfortunately, matching hash values does not guarantee that the evidence was possibly changed before it was seized.

CHAPTER V

BASIC METHODOLOGY: GLOBAL AND LOCAL ANALYSIS

We have established that digital audio authentication involves both container and content analyses. Furthermore, content analysis is divided into two main branches: *global* and *local* analyses. Some of the analysis tools can have a global component when analyzing characteristics of the whole file, or a local component when applied to regions of interest (ROI). In the proposed methodology, the examiner works from the global level to the local level. This is somewhat analogous to the way a doctor examines a patient. If he is looking at a corpse, it only takes a quick global analysis to tell that the patient is dead. Perhaps one or two local analyses, such as checking for a pulse or heart beat, are performed that confirm the suspicion. If he wishes to go farther, he can try to assess the cause of death. This could involve a full autopsy and many detailed analyses in order to come up with a definitive answer. Similarly in audio authentication, the examiner begins with global analyses tools that perform computations on the entire file. He may be able to conclude from these analyses alone that the file is not consistent with an authentic recording. If he cannot make such a conclusion, the file may be authentic, or it may contain better disguised manipulations that could be revealed through more intensive and time consuming local analyses.

A common approach in forensic authentication is for the examiner to begin with a neutral stance. The idea is to have no opinion regarding whether the evidence is genuine or tampered with, but to merely perform the analysis to see if each step produces a positive, negative, or inconclusive indicator. [4] While this is a good starting point, it could potentially be influenced by bias (even subconsciously) if the results are not black and white and require subjective interpretation. There is danger of inferring too much from results that cannot be necessarily deduced through the experimental procedure. Therefore, each stage of analysis (whenever possible) should be framed in a manner that includes the scientific method, which always begins with a hypothesis. In science, we try to prove a hypothesis through experimentation. If we fail, we either modify the experiment and retest the same hypothesis, or accept that it has been refuted and move onto a new inquiry. The point being that we do not begin with a completely neutral stance, but with a classical hypothesis test. There are only two possible hypotheses for authentication: A) the evidence is authentic; or B) the evidence is not authentic. Since we never definitively prove A, we conduct our analyses to refute it by proving B. If none of our tests are able to prove B, then we must accept A by default. This is known as the null hypothesis. [20] We can presume authenticity of questioned evidence unless proven otherwise. Therefore, the null hypothesis that the recording is authentic will persist until evidence is found to disprove it. By conducting our analysis in order to find inconsistencies, we now have a clear goal that will produce meaningful results. If we succeed in proving B, and thus refuting A, we conclude that the evidence is not authentic. We can stop there or choose to go further to show what manipulations were done (if any) if one wishes to try to establish different degrees of inauthenticity: innocuous copy (non-clone), unintentional alteration, intentional alteration, forgery, etc. These are all technically “inauthentic” recordings, but some may have different implications pertaining to the ultimate question in the case.

In summary, it is more scientific to approach each step of authenticity not with a completely neutral stance, which lacks a hypothesis and is vulnerable to multiple interpretations, but instead with conjecture based on inductive reasoning that attempts to answer a specific question.

CHAPTER VI

ANALYSIS TOOLS

Introduction

The following sections will briefly describe some of the more technical processes that are currently being used in forensic digital audio authentication. Not every technique will be covered. To avoid redundancy, I will not go into as much detail as other papers that are already published. Readers are encouraged to consult all the primary documents in the field. However, I will explain logically how these tools are intended to be used for authentication, how they can practically be used, and what I perceive some of the weaknesses to be. New methods are being devised to fill in the current gaps, but more research is needed. The examiner should approach each analysis technique with a general scientific hypothesis, or premise that allows for necessary deductions. For scientific theories to work, they need to be true for both general and specific cases. “What goes up, must come down” is a general claim that is true in every instance where the force of gravity is present. One cannot find an example, or construct an experiment, that falsifies it. However, if we begin with the premise: “An MP3 file that is allegedly tampered with will show signs of either external software, second generation lossy re-compression, or DC offset inconsistencies,” we cannot say that statement is true in all cases. There are very plausible explanations that contradict all parts of the premise (some perhaps more likely than others). Although Occum’s Razor states that the simplest explanation is usually the most plausible one, we cannot know the capabilities of the people who had access to the evidence before it was seized, especially when the tools and knowledge necessary for tampering are available to almost anyone. Forensic audio experts should analyze only the physical evidence and data, rather than guess about the actions or capabilities of persons who made the recording or had access to it. Therefore, the scientific method is used to try to determine what physical, mechanical, electronic, or data specific processes necessarily had to occur for the recorded evidence to arrive in its current form. If alternate explanations can be found, every effort must be made to eliminate them in order to have strong conclusions. For a scientific premise to be true, it must be both valid and sound. If we begin with an invalid premise, our methods may be sound (and even highly technical, quantitative, and repeatable), but the only thing they proved was a false assumption. We must therefore begin analyses with a falsifiable hypothesis and conduct our analyses to disprove it.

A digital audio file can be expressed as a two column vector representing time and quantization level. Several types of analyses compute the similarities and differences between vectors of questioned and known recordings. One mathematical equation used for this comparison is Correlation Coefficient with Mean Subtraction (*CC*), represented as:

$$CC = \frac{\sum_{k=1}^{N}(x_k - xm)(y_k - ym)}{\sqrt{\sum_{k=1}^{N}(x_k - xm)^2}\sqrt{\sum_{k=1}^{N}(y_k - ym)^2}}$$

Another important equation is Mean Quadratic Difference (*MQD*):

$$MQD = \log\left(\frac{1}{N}\sum_{k=0}^{k=N}(x_k - y_k)^2\right)$$

Evidence Handling

When handling digital evidence, the International Organization on Computer Evidence (IOCE) principles should be respected, as well as those by SWGDE, AES, and other appropriate professional organizations. The basic principle of any forensic evidence handling is not to change anything. Since digital evidence is often volatile in nature, this principle must be respected even more closely. Generally in forensics, most physical evidence brought into a lab for analysis is essentially like a corpse. It does not significantly change (if preserved) unless a person handles it improperly. However, a lot of digital evidence can be changed, often inadvertently, simply by turning on the device or by plugging it into a computer without a write blocker. While the audio itself will likely not change (unless the device initializes auto-record mode or connects to a network), the metadata, MAC times, and possibly the internal clock used for time stamps could easily. Furthermore, if the examiner needs to make exemplar recordings on the device, doing so could overwrite important deleted information in the slack space, change the contents of the memory (ensuring that new hash values will not match previous disk images), or accidentally erase evidence if he is unfamiliar with the operation of the device. In the situation just described, the best approach is to make test recordings onto separate non-evidentiary removable media (if the device uses SD cards, etc.) or to get permission in writing to make recordings on the actual device. Test recordings on the actual device should only be made after the entire memory has been forensically imaged, verified, and hash values compared and notated. It is also wise to read the owner's manual before turning on the device, although manuals often contain errors or omissions. When a device is turned on for the first time, the examiner should immediately inspect the settings of the internal clock and note the offset compared to the current time (using an accurate reference). While this may aid the examiner somewhat in determining the relevance of times stamps in the evidence metadata, the fact is that the internal clock setting could have been changed at any time between the moment the questioned recording was made and the moment the device was seized. A time stamp always depends on the setting of the clock used to generate it. [21] We may not know if the clock was set correctly during the time of the recording, if it was intentionally changed between the time of the recording and collection of evidence, or if battery ran out and the clock drifted. In law enforcement recordings, some of these factors may be known and documented. We can note the time offset when we power on the device, but this may or may not be relevant. We can at least verify how the unit records times stamps in the metadata (or if it does) by making test recordings. If there is something unexpected about the time stamps, such as the start and end times always being one second smaller than the total time recorded, then we know not to go searching for a "missing" second in the questioned recording.

Other housekeeping procedures such as taking photographs, noting serial numbers, physical inspection, setting to write protect mode (if available), assigning evidence number, etc., have been covered in other publications. [4] [16] [19] It should be mentioned that the examiner should be careful to prevent electrostatic discharges and ensure there is no biological evidence (fingerprints, blood, hair, etc.) on the submitted media that must be preserved.

Marking evidence is a topic of debate. Some lab technicians and forensic examiners do it as standard practice. However, a pen mark on an optical disc in the wrong location, with too much pressure applied, or with the wrong type of tip could damage the media and make it unreadable. If we mark a label or envelope instead, the evidence is not damaged, but it could potentially become separated from its evidence number. On the other hand, if we mark the evidence, we change it in some way. An examiner would not write an evidence number onto a bullet casing or blood splatter, for example. In most types of physical evidence, we simply mark the label or envelope. Including a description and photographs in the log along with the evidence helps verify that the label has not been switched (accidentally or purposefully). If the evidence itself has to be marked, it must be done with the proper type of permanent pen that is known to be harmless. The use of invisible ink may also be considered.

File Structure Analysis

Format Analysis

File structure analysis relates to the *container* of the evidence and is made up of three parts: file format, header, and hexadecimal analysis. File format analysis consists of previewing the file to look at the sampling rate, quantization (bit depth), bit rate (samples per second and bits per sample), number of channels, etc. The goal is to ensure these are consistent with the file type, extension, codec, and capabilities of the device claimed to have made the original recording. For example, if the file is in a 64 kbps MP3 but the alleged device does not support this particular bit rate, it is inconsistent with an authentic, original file. Analysis of the file format is also used to ensure proper playback for critical listening so that the recording is heard at the native format, speed, and resolution. The examiner also looks to see if the format is what it purports to be (i.e. consistent with the media wrapper) so he knows what types of analyses will be relevant. For example, the file extension can indicate WAV but there are several compression formats that store audio in WAV files such as: Microsoft ADPCM, DVI/IMA ADPCM, A/μ-Law. The examiner can also search for steganography traces, because the presence of a hidden file will impact results of subsequent analyses.

MAC times (which stands for Modified-Accessed-Created) should also be viewed to determine if they are consistent with original and unaltered test recordings. However, these may be unreliable because they can easily be changed by altering the computer's clock settings before transferring to computer, or with a MAC editor. In a situation where not just the original device, but the original computer used to make the recording or extract it is supplied, a record of changes to clock settings could be obtained (depending on the operating system) which may help in an authenticity analysis. Not all versions of operating systems retain this data, however. [22] Therefore, MAC times should be taken with a grain of salt and other means of external timestamp verification should be obtained whenever possible.

Header Analysis

A computer file requires metadata in order to be properly stored and read by the operating system. Usually, this information resides in the form of allocated bytes of data at the beginning of the file, called the file header. It could also reside at the end of the file or in a separate metadata file which points to the audio file. Header analysis is done in a hexadecimal viewer or editor. While digital image files have a fairly standard EXIF, there is no real equivalent with digital audio files and the header information contained within them can vary widely. Using a hexadecimal reader, the examiner should look at the header information to see if the file format matches the file name extension (such as RIFF or hex 52 49 46 46 indicating WAV, hex 49 44 33 indicating MP3, hex 30 26 B2 indicating WMA, etc.). Depending on the device and brand, there may be information about the model, serial number, firmware version, time, date, and length of the recording (as determined by the internal clock settings). Figure 6.1 on page 32 shows an example of a digital recording (WAV file) viewed in hex editor.

It is useful to note the time stamps and compare them to the date and time claimed by the recording operator as to when the file was made. However, if the user can change the internal clock on the device, this data may be debatable or even unreliable. However, it could provide a useful starting place to search an ENF database for external time verification, if the signal is present and a database available. If the original recording device is available, or the same make and model, the examiner should make exemplar recordings and compare the resulting header information to ensure consistency with the questioned recording. Even a header that is consistent with an authentic file cannot assure that the recording is free from alterations, as the data can often be surreptitiously changed with a hex editor. Also, some devices allow for internal editing without the interaction of external software, so recordings altered in this way may still reflect an authentic header. Although header information relates to the container of the file and not the audio content, it can provide valuable information for authentication purposes that is specific to the recording device. Since it can be easily changed (even accidentally by opening and re-saving) it should be mandatory to provide the original file/clone to the examiner. [23] Copies of evidence made innocently but in a non-forensic manner, which have had their header information altered, should be avoided.

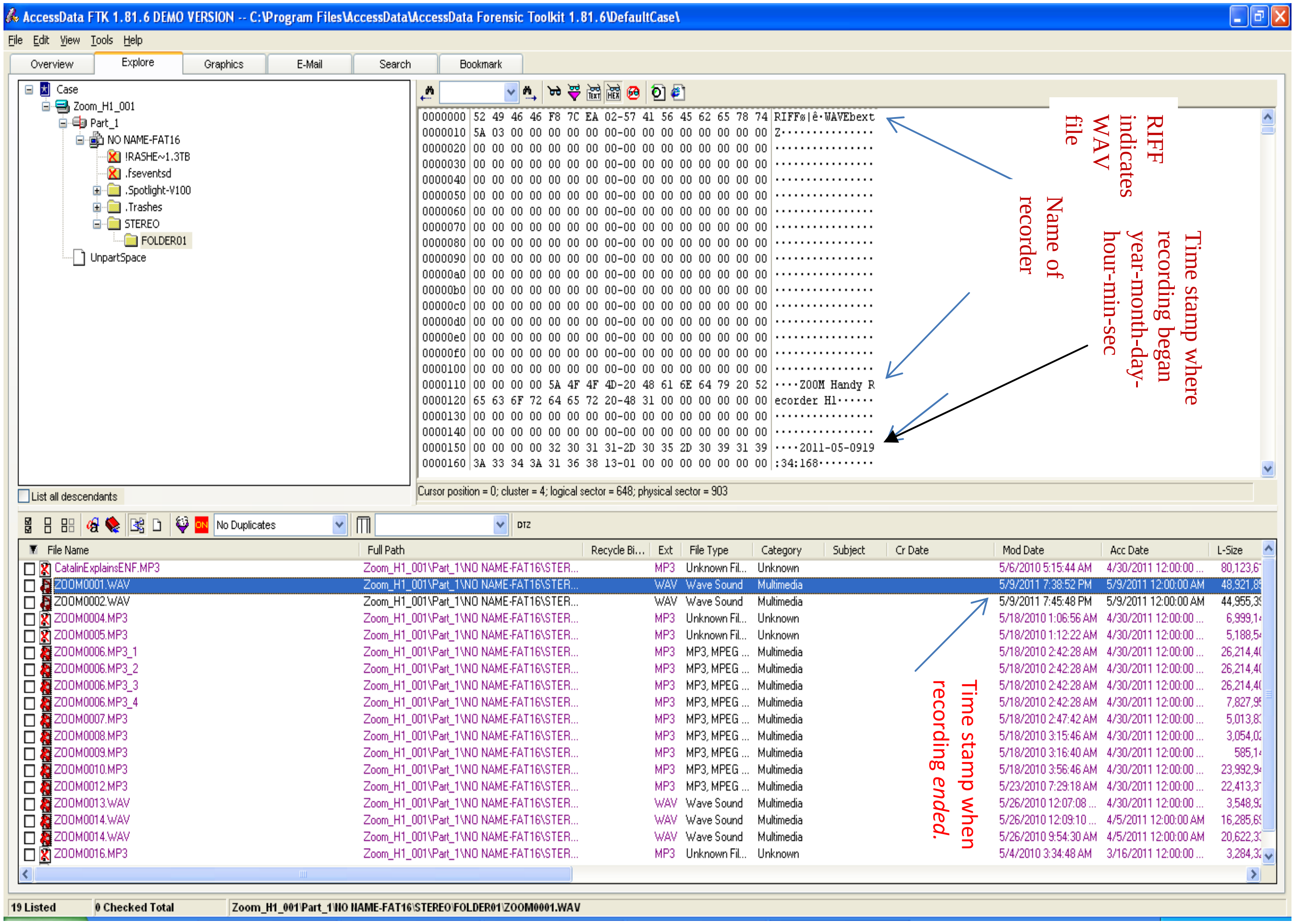
Hexadecimal Analysis

While not as straightforward as header information, the raw digital data of the file may contain useful information that can be examined in a hexadecimal reader with an ASCII (American Standard Code for Information Interchange) character viewer. See Figure 6.1. Block addresses of audio information, titles of external software (if present), post-processing operations, and other useful information may be displayed. Keyword searches can be performed and automated scripts written to scan the data for periodic structure or even missing blocks. [4] Interaction with external software will often leave changes in the hex data that are more difficult to obfuscate than header information. Attempts at covering up changes to hex data may make the file unplayable.

Critical Listening

Critical listening is one of the earliest stages, after evidence handling and data analysis, which an analyst uses for authentication. It is done to assure playback optimization, note intelligibility, observe acoustic events (location of recording, number of speakers, etc.), study background noises, detect voice discontinuities or sudden interruptions, breath flow, abrupt transitions, changes in noise floor or frequency response, and much more.

Techniques for proper selection, setup, and operation of equipment and lab space have been covered in other publications and should be followed to provide the examiner with an optimum listening environment. Quality, full-range headphones should be used, as opposed to speakers, and external over internal sound cards are favored to reduce noise and interference from the CPU motherboard and power supply. [4] This stage of analysis is highly subjective and varies considerably depending on the case.



```
0000000 52 49 46 46 F8 7C EA 02-57 41 56 45 62 65 78 74 RIFFø|ê·WAVEbext
0000010 5A 03 00 00 00 00 00 00-00 00 00 00 00 00 00 00 Z···············
0000020 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000030 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000040 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000050 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000060 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000070 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000080 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000090 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
00000a0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
00000b0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
00000c0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
00000d0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
00000e0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
00000f0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000100 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000110 00 00 00 00 5A 4F 4F 4D-20 48 61 6E 64 79 20 52 ····ZOOM Handy R
0000120 65 63 6F 72 64 65 72 20-48 31 00 00 00 00 00 00 ecorder H1······
0000130 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000140 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 ················
0000150 00 00 00 00 32 30 31 31-2D 30 35 2D 30 39 31 39 ····2011-05-0919
0000160 3A 33 34 3A 31 36 38 13-01 00 00 00 00 00 00 00 :34:168·········
```

Cursor position = 0; cluster = 4; logical sector = 648; physical sector = 903

	File Name	Full Path	Recycle Bi...	Ext	File Type	Category	Subject	Cr Date	Mod Date	Acc Date	L-Size
□	CatalinExplainsENF.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	Unknown Fil...	Unknown			5/6/2010 5:15:44 AM	4/30/2011 12:00:00 ...	80,123,6
□	ZOOM0001.WAV	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		WAV	Wave Sound	Multimedia			5/9/2011 7:38:52 PM	5/9/2011 12:00:00 AM	48,921,8
□	ZOOM0002.WAV	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		WAV	Wave Sound	Multimedia			5/9/2011 7:45:48 PM	5/9/2011 12:00:00 AM	44,955,3
□	ZOOM0004.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	Unknown Fil...	Unknown			5/18/2010 1:06:56 AM	4/30/2011 12:00:00 ...	6,999,1
□	ZOOM0005.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	Unknown Fil...	Unknown			5/18/2010 1:12:22 AM	4/30/2011 12:00:00 ...	5,188,5
□	ZOOM0006.MP3_1	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 2:42:28 AM	4/30/2011 12:00:00 ...	26,214,4
□	ZOOM0006.MP3_2	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 2:42:28 AM	4/30/2011 12:00:00 ...	26,214,4
□	ZOOM0006.MP3_3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 2:42:28 AM	4/30/2011 12:00:00 ...	26,214,4
□	ZOOM0006.MP3_4	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 2:42:28 AM	4/30/2011 12:00:00 ...	7,827,9
□	ZOOM0007.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 2:47:42 AM	4/30/2011 12:00:00 ...	5,013,8
□	ZOOM0008.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 3:15:46 AM	4/30/2011 12:00:00 ...	3,054,0
□	ZOOM0009.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 3:16:40 AM	4/30/2011 12:00:00 ...	585,1
□	ZOOM0010.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/18/2010 3:56:46 AM	4/30/2011 12:00:00 ...	23,992,9
□	ZOOM0012.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	MP3, MPEG ...	Multimedia			5/23/2010 7:29:18 AM	4/30/2011 12:00:00 ...	22,413,3
□	ZOOM0013.WAV	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		WAV	Wave Sound	Multimedia			5/26/2010 12:07:08 ...	4/30/2011 12:00:00 ...	3,548,9
□	ZOOM0014.WAV	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		WAV	Wave Sound	Multimedia			5/26/2010 12:09:10 ...	4/5/2011 12:00:00 AM	16,285,6
□	ZOOM0014.WAV	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		WAV	Wave Sound	Multimedia			5/26/2010 9:54:30 AM	4/5/2011 12:00:00 AM	20,622,3
□	ZOOM0016.MP3	Zoom_H1_001\Part_1\NO NAME-FAT16\STER...		MP3	Unknown Fil...	Unknown			5/4/2010 3:34:48 AM	3/16/2011 12:00:00 ...	3,284,3

19 Listed | 0 Checked Total | Zoom_H1_001\Part_1\NO NAME-FAT16\STEREO\FOLDER01\ZOOM0001.WAV

RIFF indicates WAV file

Name of recorder

Time stamp where recording began year-month-day-hour-min-sec

Time stamp when recording *ended.*

Figure 6.1 Example of Digital Recording (WAV file) Viewed in Hex Editor (with additional notes). Recording began on May 9, 2011 at 7:34 pm and ended at 7:38pm.

An established protocol is to divide the task into four steps: 1) preliminary overview of the recording; 2) identifying possible stops/starts, pauses, and edits; 3) observing background sounds; 4) analyzing foreground voices and other sounds. [4] Some examiners use the presence of steady tones, uninterrupted noise, or music, etc., as indicators that a recording is unedited. It should be noted that any of these signals could have been added through post-processing after a file has been altered to convey the illusion of continuity. Even the presence of a continuous TV or radio broadcast in the background (specific to a time, date, and region) does not guarantee authenticity, as it could be an archived recording that was added during the editing process. While these types of forgeries will likely have artifacts of convolution (or lack thereof), re-recording, ENF inconsistencies, or external software, they might not be detected through critical listening alone. Although critical listening analysis is useful to some degree, results directly obtained from it are not always objective or quantifiable, may involve intuition as opposed to logical deduction, and should not be overly relied upon to reach scientific conclusions.

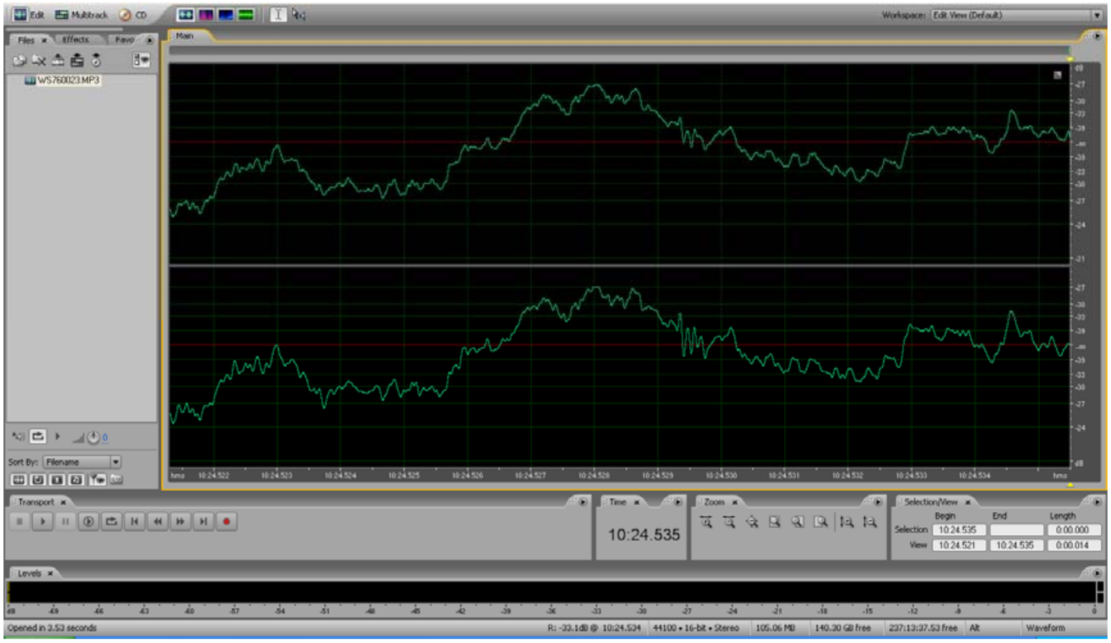
A common misapplication of critical listening for authentication purposes assumes that the forensic examiner has years of experience and highly trained “golden” ears that will detect information that is not picked up by the average listener (otherwise an expert would not be needed). While it is possible to develop highly trained listening abilities, there are still physical limits as well as the possibility of alterations that defy aural detection by even the most critical ear, and most importantly, subjective evaluations like critical listening are susceptible to bias which can affect or cloud the examiner’s judgment. While it may be possible to detect poorly made edits in this manner, the absence of such indicators does not guarantee the absence of alteration(s). Nor does the presence of regions of interest (ROI) necessarily imply edits. Therefore, an examiner who believes he can reliably detect alterations through critical listening is making a dangerous assumption, especially if he uses it as his main authentication tool. This logic is rooted in the early days of natural science when a scientist only made observations about the world around him that were available to his senses (such as practicing astronomy by simply gazing at the stars without a telescope). However, it is a very small section of natural phenomena that one understands through direct experience. While human hearing is an incredibly precise instrument, it does not measure everything, and there are many things that our ears hear which our mind does not comprehend. "It is only through refined measurements and careful experimentation that we can get a wider vision...and then we see unexpected things." [9] Therefore, we must use other tools beyond critical listening to perform a reliable authentication.

An analyst may claim to know that a recording is authentic or inauthentic based on training, experience and what they subjectively hear, but these claims are not falsifiable or scientific. Although we cannot rely on critical listening alone to reach an ultimate conclusion, we can use it as a valuable tool to help shape the order and methods of the examination, to identify areas of interest that need more investigation, and to form hypotheses that will be tested in later stages of analyses that yield empirical results. If the findings of the examiner corroborate visually, mathematically, and aurally, the conclusions will be much stronger. The results of some analyses may cause the examiner to return to critical listening (feedback loop) and form new hypotheses, or modify or discard old ones. Therefore, it is important not to infer too much from critical listening, especially early on (which can create expectation biases) until the results of all analyses are evaluated.

Waveform Analysis

This analysis technique is probably the most familiar to non-forensic audio professionals because it can be performed in any standard DAW or editor program. It involves zooming in closely on a monitor or printout to inspect areas of interest in the waveform. Only time and amplitude are displayed. The horizontal x-axis represents time in hours:minutes:seconds:milliseconds or samples, if preferred (bars and beats should not be used). The audio file always begins at 0:00:000 or sample 0 and the timeline is independent of any counter that the recording device may have embedded in the file, spoken slate indicators, or time reference from an associated video (if applicable). The timeline in the DAW becomes the master time clock for the purposes of the examination. The presence of clipped or distorted samples should be noted because they can affect the outcome of subsequent statistical analyses. The examiner should not trim any excess leader or normally make any alterations (unless for enhancement or some other purpose that is documented). The leader may contain system noise, DC samples, or non-speech information that is very useful for other analysis techniques. If enhancements are made, they shall be presented as demonstratives only and are not to be confused with the raw evidence file. The vertical y-axis represents amplitude and can be expressed in decibels relative to full scale (dBFS), percentage, or amplitude quantization levels. In authentication examinations, waveform analysis is primarily used to inspect possible edits and discontinuities. A waveform with a drastic change in amplitude that is not typical of sounds that occur in Nature may be an indication of an edit. However, peculiarities of the recording device must be eliminated as possible explanations. Viewing parts of the waveform at the sample level is a useful technique for identifying the presence of consecutive quantization levels. A 16 bit digital audio recording can represent any sample in a waveform as one of 65,536 discrete quantization steps. These steps correspond to voltage levels, which are analogous to fluctuations in sound pressure. When a naturally occurring acoustic sound wave is quantized, it is highly unlikely that consecutive samples will have the exact same quantization level. Typically, repeated consecutive levels only occur if there is clipping, zero amplitude, or zero order interpolation (which can be introduced by packet losses, read/write error correction, editing in certain software, etc.). [24] [25] Therefore, the examiner should inspect areas with repeated quantization levels to determine if they are consistent with the alleged method of recording, or are possibly indications of alteration. For example, zero amplitude levels may be an indication of deletions, unless automatically introduced by VOR or normal operation of the device. Some file formats, such as MP3 and WMA, introduce leading DC samples that precede the recorded information. [23] These will be explained more in the DC Analysis section.

Waveform analysis is one method of detecting butt-splice edits, but because a single edit can represent such a small portion of the entire sound file and may be inaudible, it is impractical for the examiner to study the entire signal at sufficiently high resolution (potentially examining millions of samples). Therefore, automatic methods of edit detection are recommended to identify possible edits that can then be examined more closely with waveform analysis. [26] If a discontinuity is identified, there are several ways it could have occurred including: signal interruption (disconnected cable, etc.), transmission dropouts, RF interference, disk write or buffer overrun errors, manual pauses, automatic VOR, automatic song split feature, internal editing on the device, post-processing editing in a DAW, and more. Many of these variables will only be applicable for certain recording methods, channels, or devices and may be eliminated as possible causes. However, if the only explanation for discontinuities is alteration or deletion, then the examiner may conclude the file is inauthentic. Therefore, it is important to know how the recording was allegedly produced. Based partly on this information, the examiner can make deductions about the origin of certain discontinuities, but he is also subject to contextual bias (i.e. different contextual information could lead to different conclusions). It is vital to have access to the original recording device and to make exemplars in order to establish a normal baseline of operation, which the examiner uses for comparison to the evidence file. A waveform signature, transient, or discontinuity that routinely appears on authentic test recordings made on one device may be a definite indication of tampering on another device. If the original device is not available, then the ability of the examiner to form hypotheses and make deductions regarding the cause of certain signatures will be severely limited. Since he will not have device-specific knowledge to apply to the questioned recording, he can only use his resources of general knowledge and experience regarding acoustics, electronics, digital audio, signal processing, and editing techniques, to make interpretations. Since very few presumptions about digital audio can be applied to all cases involving any device, the possibility of reaching false conclusions may be unacceptably high. Past experience may not necessarily be applicable to the current case, since it is the nature of forensics to deal with unique phenomena. [15] In order to be able to apply general observations about known editing signatures to specific cases and accurately conclude that editing did or did not take place (as determined by the shape of the waveform), an extensive amount of independent testing and verification must be performed.



10:24.535

	Begin	End	Length
Selection	10:24.535		0:00.000
View	10:24.521	10:24.535	0:00.014

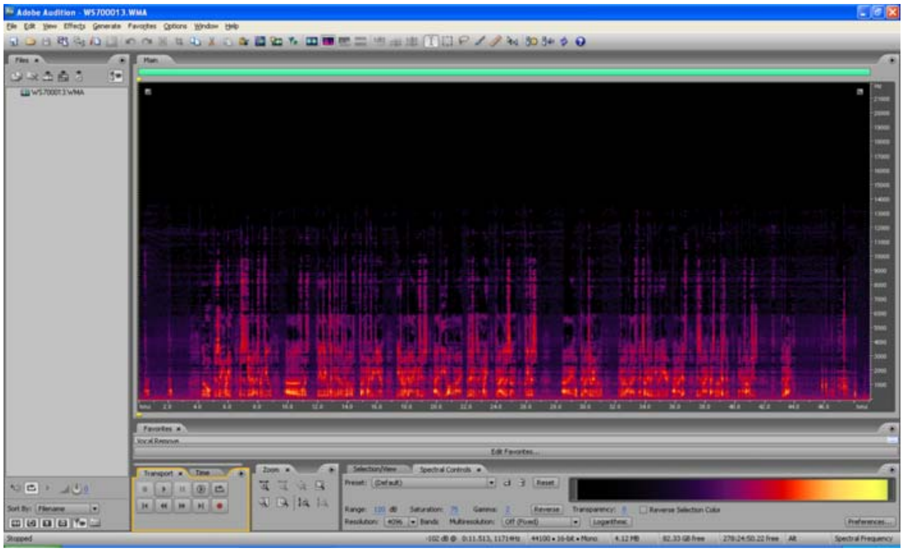
Opened in 3.53 seconds R: -33.1dB @ 10:24.534 44100 • 16-bit • Stereo 105.06 MB 140.30 GB free 237:13:37.53 free Alt Waveform

Figure 6.2 Example of Stop Signature Using High Resolution Waveform Analysis.

Spectral/Spectrogram Analysis

Several different displays use a Fourier transform to convert time domain signals into the frequency domain for analysis. The basic concept behind a Fourier transform is that any periodic signal (or wave) can be represented as an infinite sum of sine and cosine waves with various frequencies and amplitudes. [24] It is up to the examiner whether to use waterfall plots, narrow band frequency display, or a spectrogram depending on what is most appropriate. This analysis is useful in determining the overall frequency limits and response of the recording system, identifying precise pitches of particular sounds or speech, convolution effects, channel or transmission characteristics, presence or absence of ENF, etc. [4] Much is left to interpretation by the examiner. Tones that abruptly start and stop, which are visualized in the display, may be indications of editing unless they can be explained by acoustic events. The Nyquist limit determined by the sampling rate should be verified to ensure consistency with the claimed settings and device capability. The presence of anti-alias filtering, resampling, or perceptual coding may be determined by visual inspection.

A sound spectrogram is a visual representation of time on the x-axis, frequency on the y-axis, and amplitude indicated by color intensity or grayscale range. Dark colors indicate weak, or quiet signals, while bright indicate louder signals. By analyzing a spectrogram, one can visually distinguish background noise from speech, for instance, or identify many things about a recording. Spectrographic analysis can also be used to determine the presence or absence of ENF signal which can be compared to a database and used in authenticity examinations.



Preset: (Default)
Range: 120 dB Saturation: 75 Gamma: 2 Transparency: 0
Resolution: 4096 Bands Multiresolution: Off (Fixed)
Sort By: Filename
-102 dB @ 0:11.513, 11714Hz | 44100 • 16-bit • Mono | 4.12 MB | 82.33 GB free | 278:24:50.22 free | All | Spectral Frequency

Figure 6.3 Example of Spectrographic Analysis of WMA File. FFT resolution is 8192 bands; Blackmann-Harris windowing.

Phase Consistency

Audio phase is the pattern of peaks and troughs in a sinusoidal waveform, which can be viewed on an oscilloscope or high-resolution waveform display. The Bolt report on the Watergate tapes explained phase consistency: “The phase of a wave is a measure of the relative locations of the peaks and valleys in the wave. As long as the peaks and valleys follow each other by the exact same amount, the phase is said to be continuous. If at some point the wave pattern shifts abruptly one way or the other, then a phase discontinuity is said to occur at that point.” [8] The committee concluded that the “buzz” sections, which were not continuous but separated from one another by short pieces of transient events, had to have been caused by starting and stopping the tape machine and not by alternate explanations such as line interference, etc. If the tape machine had been running at the same speed without stopping, they inferred, the phase of all of the buzz sections would be continuous when compared to a grid of evenly spaced vertical divisions. The phase did not line up consistently, which therefore implied that each separate buzz section represented a time when the machine was recording, stopped, and started again.

This same technique can be applied to forensic digital recordings where there is a stable tone in the background, such as a ground hum or oscillation from a fan or motor (at a pitched frequency). If sections of the recording have been edited and concatenated together in order to appear continuous, it is possible that the editor did not line up the phase accurately between cuts. Therefore, the phase would appear continuous throughout each section, but not across the edit points. If the editor is aware of phase continuity, he could carefully chose his edits to occur at zero crossings (points where the sinusoidal wave crosses the X axis at zero amplitude) so the phase would appear continuous, but he may not be able to if his edits choices are based on closely timed events or dialogue. However, it is common in post-production voice editing for music, film, or television to use time compression or expansion processing to move slices of waveforms to zero crossings so that the edit is audibly smooth. This same technique can be applied in anti-forensics by a knowledgeable editor.

Phase consistency can also be viewed by using a computer generated sine wave, aligning it to a sinusoidal waveform (tone) present in the questioned recording, and using visual or automatic methods to compare their differences. The accuracy of the recording device's word clock should be taken into account for long recordings. Phase analysis is used differently for mono or stereo signals.

ENF Comparison

Electric Network Frequency (ENF) is a characteristic of alternating current (AC) that is supplied to customers by power companies all over the world. It operates close to 60 Hz in North America and 50 Hz in Europe and elsewhere. In theory, the power is a stable sinusoidal frequency, but in reality there is a constant fluctuation based on changes in production and consumption across the grid. If ENF is captured in a digital audio recording, the signal can be extracted and the fluctuations compared to an ENF database as part of an authentication analysis. Many publications have outlined the Electric Network Frequency (ENF) criterion. [7] [27] [28] The fundamentals that make the ENF-criterion possible is the variation of the frequency value at any moment in time is relatively the same across all points on a grid; and the frequency variations are unique and not repeatable over a period of time. [14] Therefore, a segment of recorded audio exhibiting ENF forms a vector shape that, if continuous, unaltered, and from the date/time/location claimed by the operator, should prove a positive match with an equal length signal from an ENF database. If such a match is obtained, there is a high likelihood that the questioned recording is original and authentic.

ENF analysis is a promising technique with great potential for aiding digital audio authenticity. While it has been used successfully in European courts, its use in U.S. courts has thus far been limited. The future use of ENF will depend largely on the maintenance of accurate databases (at least one per grid) and further research and testing. If smart grids become a reality, this could mean significant changes to the criterion for using ENF analysis. There will certainly be more data to collect in a smart grid than in the current conventional ones which may even improve the accuracy of the technique. However, if the frequency is not consistent across the entire grid, then the correlation between a questioned recording and a database becomes impossible.

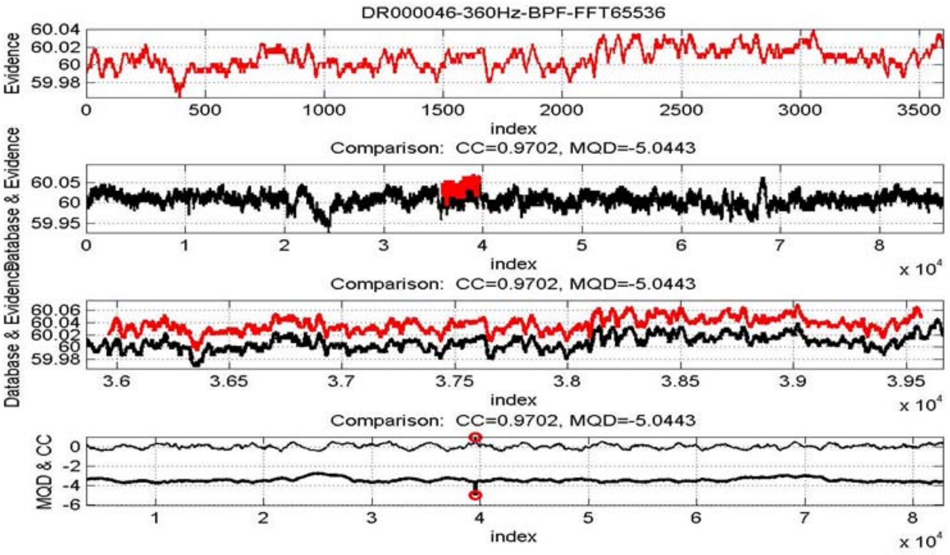


Figure 6.4 Example of ENF Automatic Comparison Method with Successful Match.

The two main methods of comparison are visual (spectrographic) and semi-automatic (FFT). The former relies on finding visual consistencies between spectrograms of the ENF signal in both the questioned recording and database. This method can be very time consuming for a blind search, but is effective if the suspected date/time/location of the recording falls within a reasonable time frame contained in the database. It is also facilitated if the questioned recording is of longer length. In some cases, it may be the only method possible if the resolution of the ENF signal is very low, which can be improved with downsampling, bandpass filtering, and normalization. A very low signal-to-noise (SNR) ratio could make the technique inapplicable for some recordings. The spectrographic method should be used first to determine if ENF signal is present, whether it is of sufficient resolution for comparison, and if the overall shape is consistent with mains power fluctuation (since ENF from a UPS, inverter, or generator can only be used to check continuity but not for database comparison). [27] If the shape is consistent with A/C mains power, the examiner can prepare the file and begin checking against ENF databases, starting with the date/time/grid most likely to contain a match. He may prefer to use the semi-automatic method, which relies on splitting signals into 1 or 2 second time windows, finding the peak FFT in each frame to reduce the amount of data, and automatically comparing vectors of the same length against a database. The segments with the best "score" in terms of highest CC and lowest MQD are evaluated by the examiner. They are said to match if there are no unexplainable differences between them. It should be fairly obvious if the score is relevant (by both visual and numerical verification) or if more comparisons are needed on a different time/date/database. If there is a good match between the questioned recording and database, this is a strong verification that the file is continuous and unaltered. As noted by Brixen, however, there still remains a prospect of cheating the criterion. “This would normally be in cases where the ENF signal (including the harmonics!) is removed from the audio signal and is replaced by another ENF-recording.” [14] To do so, the ENF signal would have to be acquired from a private database and artificially added to a tampered recording to make it appear genuine. While this type of anti-forensics would require skill, knowledge, access, and possibly forethought on the part of the forger, the ubiquity of the ENF signal at all locations on the grid makes it a remote possibility. [7] A file altered in this manner would likely have traces of external editing software or analogue re-recording. It could also contain harmonics of the original ENF signal (if captured), which are more difficult to filter out than the fundamental. If the questioned recording and the database have a highly correlated ENF signal, but there are momentary amplitude spikes present in one file and not the other, these are likely the result of local voltage fluctuations or power surges rather than tampering, as the 50/60 Hz frequency itself has not varied. Disk write errors are also possible [29] which will create short variations or missing samples in one file but not the other. However, if there are partial accordances and discordances in frequency and shape between the two files, and not merely transient variations, more interpretation is required. If the two signals begin at the same point, but slowly diverge over time, it could be attributed to an unstable word clock and not tampering. If there is ENF signal present in some sections on the questioned recording, but not others (i.e. attenuated below the noise floor), the power source may have changed from mains to battery (which would likely be corroborated with a spike and/or changes in DC offset), or sources or electro-magnetic interference may have moved or switched on or off. [Note: Devices such as

Tascam DR-100mkII allow switching from three different power sources while recording without interruption. Laptops can also switch from mains power to battery, or vice versa.] These explanations not consistent with tampering would normally be supported through other analyses. However, if the ENF signals have definite segments of correlation and non-correlation that destroys the time continuum, it is a pretty sure sign of editing, especially if they are out of order (which indicates cut/paste), repeated, or come from different dates and times entirely (as verified through more ENF comparisons). The examiner should be aware that voice operated recording (VOR) or stop/starts made by the recording operator will create skips and concatenate segments of ENF into a smaller time frame when compared to the database, but will not change the basic chronology or introduce ENF from another date/time unless re-recording or editing was performed. If two or more ENF signals are present in the questioned recording, this is normally a sign or re-recording unless a radio transmission or other broadcast was originally captured. The absence of multiple ENF signals does not necessarily indicate the absence of re-recording, however, since battery powered devices could have been used away from electro-magnetic fields. If the zero-crossing method is used for ENF comparison (which is not discussed here) it is necessary to remove DC offset, which changes the spacing of semi-periods and can lead to false conclusions.

Compression Analysis

Compression analysis looks at the relationship between neighboring samples in an audio file. Recordings that are saved to lossy compression formats contain artifacts that become a permanent signature even if the file is later up-sampled or resaved in a higher quality, uncompressed format. There are many tools for detecting the presence of these artifacts, which can help answer questions about originality and the method of recording. It is recommended to combine multiple techniques, or all of them, in order to corroborate findings. Since there is a wide variety of different codecs and resolutions available on the market, and the possibility of more than one combination used on a single recording, artifacts not detected by one analysis tool may be found by another.

While an original uncompressed or lossless signal has a low correlation between nearby samples, which are essentially random (depending on the content of the signal), "a lossy compression algorithm will introduce correlations of varying degrees between neighboring samples." [28] When the file is analyzed, these artifacts are expressed in the form of periodicity because the algorithms are basically storing the differences between nearby samples as opposed to capturing all of them at the highest resolution. This is somewhat analogous to the way that .JPEG compression introduces blocking artifacts in digital photos by interpolating values of neighboring pixels. Lossy compression permanently removes information in order to decrease file size or increase bit stream efficiency. Codecs such as MP3, .AAC, WMA, etc. use advanced perceptual coding algorithms to discard information that will be least objectionable to human listeners. Derivatives in calculus, which measure the rate of change of a function (such as instantaneous acceleration), may be used to indicate that the variation of overlapping frames is slower for interpolated samples than it is for naturally occurring acoustic signals. Variable Bit Rate (VBR) codecs are adaptive to the input signal and assign higher priority to complex portions and lower priority to simpler portions. Constant Bit Rate (CBR), on the other hand, maintains the same quality for the entire length of the file. At its most basic level, compression analysis can begin by a visual spectral analysis to see if there are characteristics of lossy compression, especially in the higher frequencies. For example, a music recording of the same material will look very different for an uncompressed WAV PCM file than it would for a perceptually coded MP3. There will be a steep drop off of upper harmonics, convolution, and low amplitude information. After visual inspection, the examiner can go further using the *CL* function. This produces numeric values, tables and figures that help distinguish between compressed and uncompressed signals. [28]

An important question to ask during authentication is what was the audio file’s native format when was it originally created? It cannot be assumed that the contributor of the evidence is telling the truth about the alleged device or methods used to create the recording, or that the present file format, as submitted, is the actual original format. However, the examiner can begin by looking at the file format, header metadata, and the possible formats that the device is capable of making (by consulting the manual and making test recordings) to see if they are consistent with the evidence. Previously, many portable digital audio recorders used proprietary formats that required specialized software to open and play back the files. Since the software usually did not allow re-saving in the native format, it was essentially its own write blocker, so creating successful forgeries or alterations was very difficult. While proprietary devices and formats are still sometimes used in law enforcement, devices are now commonly available that use standard formats. Recordings made on such devices often have a data connection (such as USB), can be opened with most software, re-saved, and moved from a computer back to the device. Compression analysis is one way of detecting evidence of the interaction of external software with the questioned file. May devices give the option of recording to both uncompressed and lossy formats (such as WAV and MP3), or even to both formats at the same time (such as the Roland R-05). If the questioned recording allegedly originated on a device that only records WAV PCM or lossless files, but contains perceptual coding artifacts, then it is suspected that the file has been converted and re-saved and is thus not original. On the other hand, if the file allegedly originated as a perceptually coded file (such as MP3) then it would be expected to see one generation of lossy compression artifacts. However, a WAV PCM file that is free from compression artifacts is not necessarily authentic (based simply on compression analysis) because it could have been edited and re-saved as a WAV without ever having compression introduced. Therefore, the presence or absence of compression artifacts in regards to authentication depends on many factors and the questions that one is trying to answer. There is little stopping the contributor from making a recording as a WAV PCM file, performing edits or alterations, and saving it as an MP3 to help obfuscate some of the traces of manipulation. If he erases the original file and claims that the MP3 is the true original, the examiner will not suspect alteration merely from looking at the compression level. In terms of compression artifacts, it may appear no different than an authentic file that is recorded natively to MP3 format, assuming the device is capable of recording to this format and resolution. (Note: the bit rate should be inspected to verify consistency with the device options; this can sometimes change with firmware updates.) If the device is capable of multiple formats and a perceptually coded file is presented as the supposed original, the process of deduction is much harder and not all alternate hypotheses can always be dismissed.

Compression analysis can also be used as an authentication tool to try to determine if a lossy compressed sound file is a first or second generation. In order to tamper with a compressed file, it normally must be opened in a DAW, edited, and saved again. This process necessitates transcoding the file to an intermediate file format and afterwards re-encoding back to the original format in order to make the file appear original. Alternatively, the edited file could be saved to an uncompressed or different compressed format without using the same algorithm, perhaps to save the file onto a CD or different device. If the audio passes through a lossy compression codec twice, it becomes a second generation fitting one of several descriptions: A) same codec, same setting; B) same codec, different setting; C) different codec, similar compression level; D) different codec, different compression level. Each possibility, quality setting, and combination will likely change the file in a different way. However, more research is necessary to detect these various signatures reliably so the examiner can scientifically accept or rule out different scenarios. The examiner can conduct his own tests during the authentication procedure, but his time and research may be limit exhausting all variables. Currently, the examiner operates under the assumption that a second generation file should exhibit higher compression artifacts than one that is first generation, since the effect of destructive compression codecs is cumulative. This is because the encoding process for a second generation file is interpolating information that has already been interpolated, and therefore should exhibit less random sample correlation and more periodicity. If an examiner can reliably identify second generation artifacts, he can conclude that a file containing them is unoriginal. However, if a file is identified as a second generation, the alteration may be innocuous or deceptive. There are legitimate reasons for saving to a lossy format, and some programs do it by default (such as Apple iTunes). The recording operator or contributor may have simply opened the file to listen to it, and re-saved or re-encoded it without realizing the changes it made to the container and/or contents. However, the examiner cannot determine the state-of-mind of the recording operator; he can only say that the file is not original. If the contributor has nothing to hide and claims the content was not changed, he should supply the original file for analysis. Providing only a second generation file for analysis without explanation is suspect, and will make reliable authentication difficult or impossible. Although a second generation file cannot be original, it may still be an accurate representation of the acoustic events. However, the examiner may never know the extent of the alteration(s), if any, without the original to compare it to.

Being able to accurately identify whether an audio file is uncompressed, perceptually coded, and if so, which generation is helpful in determining authenticity. The presence or absence of compression artifacts is merely a characteristic of the file that requires interpretation as to whether it supports the likelihood of either authenticity or alteration. If more than one format was combined to produce a forgery, the possible combinations of codecs, bit rates, and quality levels becomes very complex. It is important to use compression level analysis on both the entire file and well as specific regions of interest. If the examiner suspects that certain words or sections of a recording have been copy and pasted from other recordings (either identified through critical listening, discontinuities in the ENF signal, DC analysis, or other methods) an effective technique is to split up the file and perform compression analysis on these regions of interest, as well as neighboring areas, to see if they show the same compression artifacts.

If the compression levels differ, then the regions likely came from different recordings (although variations in acoustic signal complexity or gain could affect the encoding process). This type of forgery sometimes known as compositing, if indicated, should be corroborated by the results of other analyses to strengthen conclusions. If the compression level is the same for all sections, the file could be unaltered or possibly regions may have been deleted and concatenated from within the same recording. A third possibility is that regions were copy/pasted from a separate recording of the same format and compression level, which would have the same compression level for all regions (but possibly second generation artifacts across the entire file). Empirically, computational tools can only detect levels of compression which then must be interpreted by the examiner. Evidence of compression is not necessarily evidence of tampering. The presence of (re)compression artifacts only gives the examiner definitive proof of tampering in very specific circumstances. While there are some scenarios that yield definite inconsistencies that strongly indicate tampering, the evidence of authenticity, conversely, is less certain since alternate hypotheses may persist. It is thus a complex process of elimination that aids the examiner by increasing or decreasing probabilities towards one conclusion or the other.

It should be noted than when comparing the results of the *CL* function on different regions or files, the most accurate results will be obtained from using silent (non-speech) portions of audio. This is because the dynamics or complexity of the audio material, which varies with the acoustic events, can impact the adaptive encoding process of many algorithms. For example, comparing a noisy recording containing many voices (cocktail party effect) with a recording containing a single voice and little background noise may yield irrelevant results even if both files used the same format and settings. It is recommended to select silent regions from the questioned recording and known exemplars to provide a more stable baseline for comparison. Also, signals containing lossy compression will yield irrelevant results with butt-splice detection and other tools that look for fast transitions between samples, since these characteristics are often obscured by the compression algorithm(s).

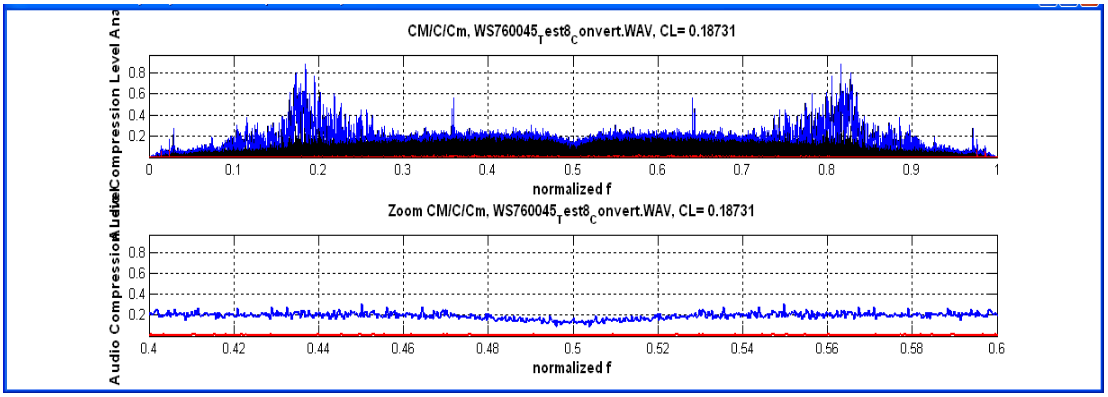


Figure 6.5 Compression Level Analysis of WAV PCM File.

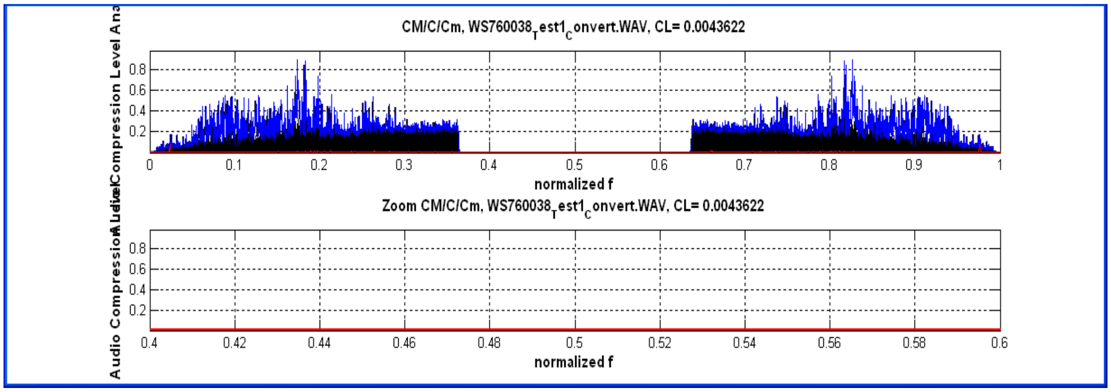
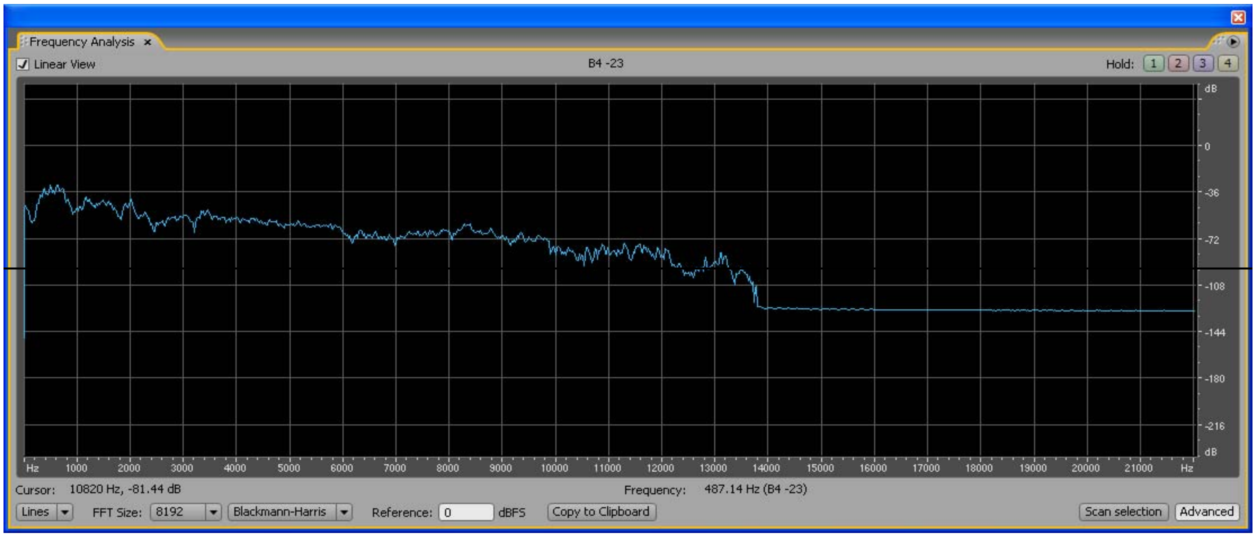


Figure 6.6 Compression Level Analysis of MP3 File at 128 kbps.

Long-Term Average Spectrum (LTAS)

A signal's long-term average spectrum (LTAS) can reveal information about the environment, the acoustic sources, characteristics of the recording equipment and transmission channel, other signals picked up by the transmission channel and/or the recording equipment, possible traces of filtering, down sampling, or lossy (re)compression. LTAS is a mathematical computation that can be used for many purposes including the statistical comparison of different signals. To obtain LTAS, the signal is divided into short time windows, then frame functions are applied, FFT and power spectral density are computed on each resulting frame, and the results for each frame are averaged. The result is a two-column vector consisting of frequency and amplitude that can be saved and used with various algorithms. Furthermore, a histogram can be derived from the LTAS and show the number of appearances of each energy level. The examiner can view a global spectrum of the signal and verify if the curve is consistent a typical recording obtained with the methods, equipment, and settings claimed by the recording operator.



☑ Linear View

B4 -23

Cursor: 10820 Hz, -81.44 dB

Frequency: 487.14 Hz (B4 -23)

Lines

FFT Size: 8192

Blackmann-Harris

Reference: 0 dBFS

Figure 6.7 Example of Long-term Average Spectrum (LTAS) for WMA File.

Sorted Spectrum

Discrete Fast Fourier Transform (DFFT) computes the minimum, maximum, and mean frequencies over all the frames in a recorded signal, known as M3. [28] A perceptually coded file will show a steep drop off in magnitude above a certain frequency in contrast to an uncompressed file, which is more consistent across the full range of frequencies. In a sorted spectrum, the x-axis represents an index of all the frequencies, and the y-axis represents the values from high to low. Sorted spectrum is often combined with DFFT. Derivatives can be applied in order to detect traces of previous signal lossy (re)compression. An uncompressed recording will produce a figure with a smooth curve, whereas a recording featuring lossy compression will have a distinctive bump. In some cases, the differentiated sorted spectrum may provide more useful and detailed results.

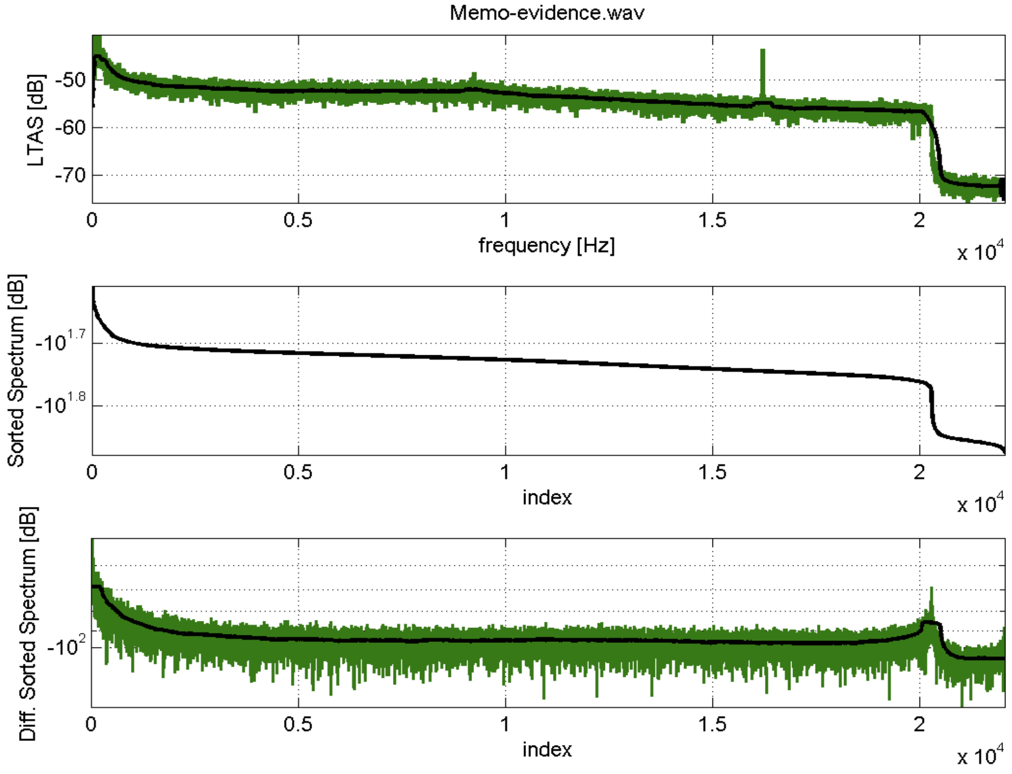


Figure 6.8 Example of LTAS, Sorted Spectrum, and Differentiated Sorted Spectrum for M4A File.

DC Offset

Along with the desired acoustic events, a digital audio recording captures information about the characteristics of the recording device and, more specifically, the linearity of the analogue-to-digital convertor. DC Offset analysis can indicate forgeries where regions were copied and pasted from separate recordings, because it is unlikely that the gain for both recordings was identical. It can also reveal changes in amplification within a single recording. All digital audio recording devices contain an analogue-to-digital (A/D) convertor or a sound card. When electrical energy from the microphone(s) or transducer(s) passes through the convertor, it becomes a bit stream of binary information that is recorded onto the device. Some direct current (DC) is introduced into the bit stream from the circuit so that the overall waveform is modulated, or shifted, above or below the zero line. In general, poor convertors will have more DC offset and less linearity than quality ones. [30] In an ideal recording with no DC offset, the average of all the samples would be zero because the positive and negative amplitude values would cancel out. In reality, an authentic recording that has not been processed will exhibit a slight positive or negative bias. In order for an original signal to have an ideal zero DC, it is necessary for the entire signal to be predicted. For such a signal to occur in near real time without post-processing, the device would have to store the entire recording in a buffer (up to an hour or more in length), and then remove the DC before saving the file to memory. No digital audio recorders have been reported to operate in this manner, so therefore questioned recordings that have a zero mean (no DC offset) have undergone post-processing and cannot be original.

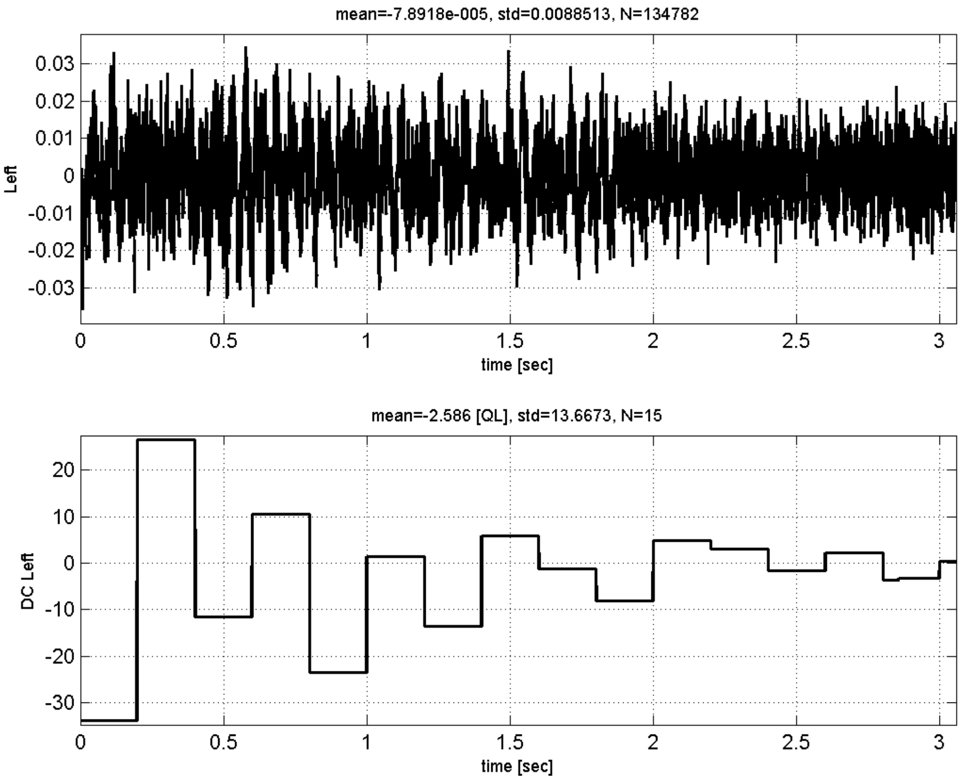


Figure 6.9 Example of audio signal (top) with DC analysis (bottom). Each step is an average of DC levels per time division. Mean and standard deviation of all samples is also indicated.

A new technique for device verification involving DC analysis has recently been developed. Although not yet validated, upcoming papers by Koenig et al. will outline the method. It involves several steps. First, DC analysis is conducted on a global level for the entire length of the file. If there are areas with a sudden shift in DC levels, this is an indication of gain changes and possibly editing. Once suspected areas of copy/paste are identified, they can be corroborated with other analyses by splitting them into separate files and doing LTAS and Compression Level comparisons (as well as miscellaneous techniques) to see if they originated from different formats or resolutions.

The next step is to compare the DC mean and standard deviation from of a questioned recording to exemplars made on the same device in silent conditions. Due to the variety of electronic components and circuits, different devices will have a different "normal" range of expected DC values. This varies depending on whether the built-in microphone, external mic, or line-input is used, so the range for the relevant recording technique should be considered (as claimed by the recording operator), as well as eliminating alternate methods. The power source, or battery charge, can also affect measurements. Test recordings must be made to establish a baseline for the device. Once the normal range or intra-variability is established, the questioned recording is compared to the known exemplar(s) to determine if the DC offset is consistent with authentic recordings made on the device.

Before the comparisons are made, it is important to remove all the speech from a working copy of the questioned recording and select at least several seconds with no tones or acoustic energy (only system noise). Otherwise, false conclusions may be reached. This technique may not be possible for all cases if a sufficient amount of silence cannot be extracted. In the future, it will be useful to quantify the minimum recording length and maximum SNR (in this case "noise-to-signal" since values with more noise than signal are desired) necessary to achieve accurate results. It should be noted that the beginning of a sound file will often have non-linear DC samples compared to the rest of the recording due to residual voltage in the start signature, or from the encoding process. Therefore, it is important to use the beginning of both files (evidence and exemplar), or from neither file, to make relevant comparisons. Care must be taken in the preparation of test files to ensure that a region containing a start signature is not compared to a silent region from another recording where the start signature was removed (do not mix and match). Otherwise, experimental errors may occur and results could be irrelevant.

Attempts at covering up DC variations resulting from copy/paste are difficult to achieve without detection. Even DAWs with a DC removal function only shift the mean up or down closer to zero. They do not change the standard deviation, or spread, of the DC levels, which is a signature of the recording system. It should be noted that some lossy compression algorithms introduce segments of zero DC in non-speech segments containing low background noise or acoustic information. The DC statistical analysis feature in some software, such as Adobe Audition or Sony Sound Forge, is not accurate for this type of analysis. Therefore, custom scripts should be used.

Butt-Splice Detection

Digital audio signals are mathematical functions of frequency and amplitude over time. Butt-splice detection functions [26], like the first derivative, can be run over the entire signal (if uncompressed) to show the rate of change of one value in relation to the previous. By calculating the differences between consecutive values in a vector, algorithms detect butt-splice edits where the shift to a high level transient is very rapid in relation to the quantization levels of surrounding samples. Since samples generated from naturally occurring acoustic events do not exhibit such a fast transition compared to butt-splices, inconsistencies can be located automatically even in long signals containing millions of samples. Low signal-to-noise ratio (SNR) or the presence of perceptual coding may negatively impact the results. This method is effective for showing butt-splice edits, but not smooth edits that were done at zero crossings, or with cross fades or interpolation. Therefore, it must be combined with other analysis tools to determine if the recording is likely free of edits.

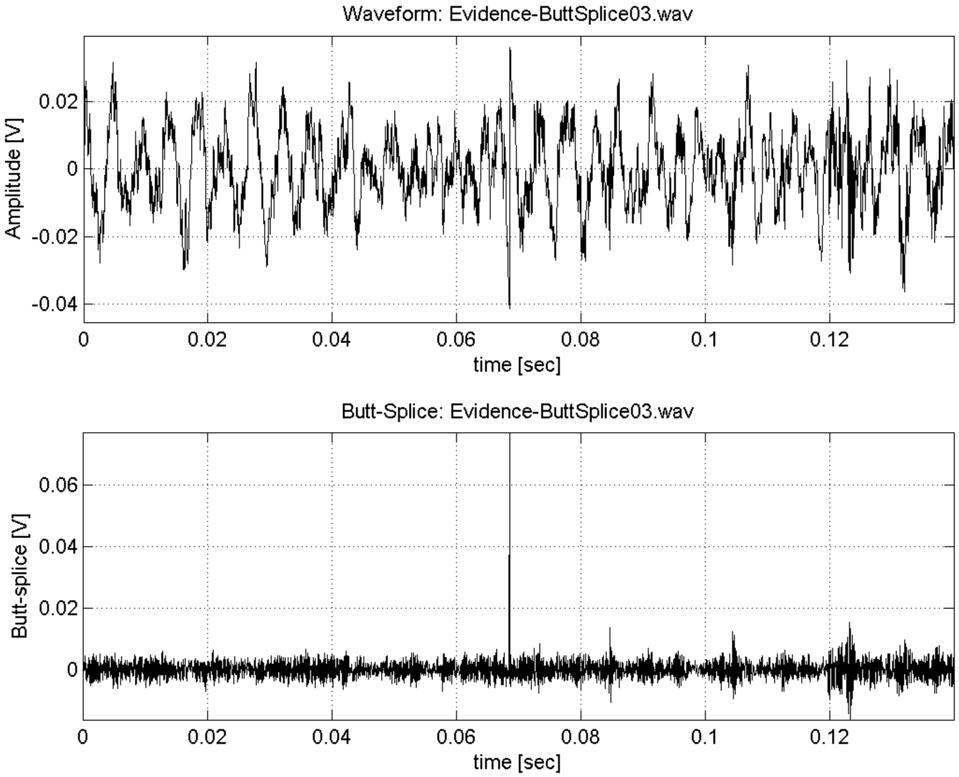


Figure 6.10 Example of Waveform and Automatic Butt-splice Detection.

Device Verification

Establishing the origin or provenance of evidence is part of the authentication procedure. Several of the techniques already mentioned can be used to verify that the questioned recording originated on the device from which it allegedly came. The original device, as opposed to a similar one, is needed to do these types of tests most effectively. Exemplars should be recorded on removable media that is separate from the original evidence. If the device only has internal memory, written permission should be obtained to make test recordings. These should only be made after the internal memory has been forensically imaged and verified, because creating new files may overwrite deleted data in the slack space and will change hash values. Plus, accidental erasure of evidence could occur especially if the device is unfamiliar. Once exemplars are made, a simple, but effective technique is to examine the file format, header, and hex data on exemplar recordings and compare them to the evidence. If the device creates time stamps or other meaningful metadata in the sound files, it can be quickly determined if the questioned recording is consistent or inconsistent with an original in terms of the container. Also as part of device verification, test recordings should be made with various different power sources (mains and battery) in various locations and with different settings to determine if the device is capable of capturing ENF signal. A questioned recording that contains ENF allegedly made on a device that is not known to capture it could be a strong indication of tampering.

Other techniques for device verification include DC analysis (previously mentioned) and LTAS comparison. Many forensic disciplines compare a questioned to a known sample, such as fingerprints or hair and fiber analysis. In any such comparison, including digital audio files, it is imperative to compare like objects or phenomena (apples to apples). Audio signals, however, represent unique acoustic events that are not replicable. Even in a speech recording, if the same speaker re-records the same dialogue, the resulting waveforms will never be an exact match due to many random variables. Therefore, the only way to make a relevant comparison between a known and unknown recording is to use recorded regions or entire files containing relative silence, which will provide correlated signals that reflect the characteristics of the recording system (i.e. microphone, electronics, DSP, etc). Silence in this case is defined as non-speech segments with no acoustic information or tones detected through critical listening or spectral viewing. In other words: pure system noise. Therefore, working copies of questioned recordings must be prepared by separating as much silence as possible from speech and other acoustic signals, and saving it to a separate WAV PCM file. Exemplar recordings are then made in the same acoustic environment, using the same equipment and settings. The orientation of the microphones must also be the same, and ambient noise must be similar. While this approach has not yet been validated, there is ongoing research and upcoming publications investigating it and preliminary tests at the NCMF have been successful. To be most effective, multiple exemplar recordings should be made and histograms representing intra- and inter-variability of devices formed. This reduces outliers and increases accuracy in the ability to include or exclude the evidence as having possibly originated from the claimed device. Since this is a time consuming operation and not every examiner will have the appropriate resources, it is recommended that shared databases become established that store raw data of many exemplar recordings from various devices. If the original device is available, it should be verified that it is in the same general condition as when the questioned recording was made. Otherwise, results may be irrelevant. If the device is missing, destroyed, or unobtainable, comparing a questioned recording to the same make and model number unit is useful for establishing class characteristics, but not individual characteristics of the evidence. Therefore, it might help in terms of excluding a possible device, but not in making a strong identification.

CHAPTER VII

OBJECTIVITY

Computer based processing and algorithms are a double-edged sword. On one hand, they greatly aid the examiner in terms of objectivity and performing computations that would otherwise be too time intensive or impossible. They yield results that our limited unaided senses would not otherwise be able to access. They are repeatable, reproducible, and largely free of bias. However, they can sometimes represent a proverbial “black box” and should not be used without validation and the examiner’s true understanding of the functions they are performing. If the software tools are proprietary, this poses a problem in terms of transparency, and makes it difficult for an examiner to explain precisely how he arrived at his conclusions. The examiner should not be sponsored by any one vendor, and should always use secondary or redundant tools by a competing manufacturer to see if he obtains similar results. Relying on the results of a particular tool only because a respected government agency uses it is not acceptable, and is an example of the appeal to authority fallacy. It does not necessarily make a tool or technique accurate and reliable just because an authority uses it, which history has proven in countless examples. Accuracy can only be established through independent testing and validation.

CHAPTER VIII

OTHER TECHNIQUES AND CHALLENGES

Re-Encoding and Transcoding Standard Audio Formats

When a file format is converted to another format, through a process of decoding and re-encoding, it is said to be transcoded. Often, it is not directly re-encoded but passes through an intermediate uncompressed format. Many lossy compressed audio file formats add a few milliseconds of samples, or preroll, to the beginning of files before the initial acoustic sound waves are digitized. [23] These should not be confused with part of the audio signal, since they are added by the algorithm buffer. These leading samples can be viewed in waveform analysis and may have zero amplitude or a DC offset that is an order of magnitude greater than the rest of the recording, which can throw off results. While an uncompressed WAV PCM file should not have leading samples, other formats, such as MP3, generate them as part of the encoding process. Just opening and saving a file to the native lossy compressed format may change the length of the audio file by a few milliseconds if the program re-encodes the file (as opposed to saving a clone). A second generation, compressed file that has been encoded twice could have a greater, or fewer, number of samples than a first generation file, depending on the software and codec used. [23] Samples that were not present in the first generation file may also be added to the end. As part of the authentication procedure, the examiner should look at the manner in which the recording was claimed to have been made, make exemplar recordings on the device using the same format and settings, and see if leading samples are introduced. If the evidence recording is not consistent with the exemplars, it may be an indication that the file is not original and was trimmed and resaved in an editor. Conversely, if an allegedly original recording is submitted as a WAV PCM or other uncompressed file, but contains leading DC samples, it may have been previously upsampled from a compressed format. This would be corroborated with the results of a Compression Level or Sorted Spectrum analysis.

Analogue Re-Recording

Nearly all digital audio devices contain a D/A convertor that passes signal through either a speaker, line-output, or headphone jack. Tampered recordings that are played back through analogue conversion and recorded to separate device (or looped back to the original device) are difficult to detect through current methods. Since the files they create are new, original, and probably consistent with authentic ones made on the claimed device, by some definitions these recordings may be considered “authentic,” even though the events they capture are essentially staged. While these types of forgeries may pass many of the current analyses undetected, there is hope that new tools will be developed to reliably discover them. There *are* intrinsic differences between original digital recordings and ones that have been re-recorded in this manner. The signal of the former will have passed through one stage of electronics, A/D conversion, and anti-alias filtering whereas the latter will have passed through two additional stages. One exception would be if both devices have digital inputs and outputs, in which case only one stage of conversion is still possible. Alan Cooper developed a method of detecting these types of forgeries that was effective in laboratory tests. [31] However, the algorithm can only be used for recordings free of lossy compression, that do not contain high frequencies near the Nyquist limit, and when the original device is available. Therefore, more tools should be developed to increase applications in other cases.

CHAPTER IX

FRAMEWORK PROPOSAL AND GRADING SYSTEM

After occurrence, collection, analysis, and interpretation, the final stage of forensic evidence is presentation. There is currently no standard for the manner in which examiners present their findings to the court or requesting party. However, it is a violation of ethical codes of conduct to overstate or understate the certainty of conclusions. Currently, audio experts present their findings in many ways. Since there is no consensus or concrete rules to adhere to, it is hard to find fault with any of them as different examiners have their own preferences. Most examiners believe their method is superior. Some express authenticity conclusions in the form of an opinion (with varying degrees of explanation), some use probability or statistics with numbers generated through computations or (regrettably) according to a “feeling,” some use likelihood ratios (Bayes’ Theorem) or verbal scales, and others use language implying 100% certainty. Probably the most neutral way to present evidence is to simply allow the evidence to speak for itself. For example, in voice identification or elimination a jury can listen to samples of the known and unknown voice and decide if they are the same person. However, this type of presentation is not realistic in audio authentication because the results of the analyses require specialized knowledge and translation in order to be understood. The examiner is the mediator between the science and the assessment of the ultimate truth made by the judge or jury.

This paper proposes an authentication framework developed by the NCMF to be used as both a means of presenting the findings of the examiner, and in helping him form an unbiased ultimate conclusion. Both the tools and the methodology will change and improve over time, as they are dependent upon one another. This thesis is meant to begin a discussion, rather than end the conversation. Like an experiment, we do not yet know the outcome and should not try to force the results. We should not begin with the conclusion that a rock solid framework for determining authenticity is possible with the currently available tools, and then look for evidence to support that presumption. An effective methodology must be shaped above all by the scientific method and by no external interests.

The method used at the NCMF is to say a questioned recording is: *inconsistent, inconclusive,* or *consistent* with an authentic recording. The following table is used to present the interpretation of the results obtained from all the analyses conducted on a questioned recording. Although the appearance is simple on the surface, all of the analyses should be supported with thorough work notes and whenever possible, by repeatable and reproducible experiments. Upon questioning, the examiner can clearly show how he arrived at his opinion and provide more detailed information.

Table 9.1 Authentication Framework Proposal.

File Structure	File Format
	Header
	Hex Data
Global Analysis	DC Offset
	Long Term Average Spectrum, Sorted Spectrum, Differentiated Sorted Spectrum

	Compression Level Analysis
	ENF
Local Analysis	Critical Listening
	Waveform Analysis
	Signal's Power
	Spectrum/Spectrogram Analysis
	DC Offset
	Butt-Splice Detection
	Traces of Interpolation Detection
	ENF Comparison
	Phase Continuity (mono)
	Phase Continuity (stereo)
Device Verification	Header (exemplars)
	Hex Data (exemplars)
	Long Term Average Spectrum, Sorted Spectrum, Differentiated Sorted Spectrum
	DC Offset
	ENF

The examiner evaluates the results of each analysis and assigns an appropriate grade, which helps form his ultimate conclusion. Results where no alteration is found are designated “consistent with an authentic recording” (*C*), whereas detected alterations or unexplained inconsistencies are labeled “inconsistent with an authentic recording” (*I*). Results of individual analyses that support claims of neither authenticity nor inauthenticity are labeled “inconclusive” (—). In some cases, the ultimate conclusion of an examination may also be inconclusive. [22] Nothing is labeled “authentic.” The differences between “authentic” and “consistent with an authentic recording” are not merely semantics, but are in fact different definitions. If the fact finder does not understand this distinction, then the framework may confuse, rather than clarify, the matter in question.

CHAPTER X

CONCLUSION

The work product of a digital audio authenticity examination is ultimately an opinion delivered by the forensic expert. However, the process of arriving at this opinion should be shaped, as much as possible, by scientific methods. It is never possible to eliminate bias completely and it is natural that circumstances surrounding the case will somewhat shape the investigation (to help the examiner decide where to look and what to test). It is even possible for the examiner to change his opinion if new facts are found, different questions are asked, or circumstances change. However, the scientist must remain objective and as unbiased as possible. Automatic methods and statistical analysis can never replace the need for a trained human examiner, but if he employs them as part of his work they will help mitigate bias. The idea that a tampered recording could fool one or more of the analyses, but not all of them, is a theory should been tested. While it is logical to assume that it is harder to fool multiple analyses as opposed to one, it cannot be assumed that a questioned recording that successfully passes all the analyses is 100% authentic. However, it can be claimed that the examiner did an exhaustive search for detecting alterations.

APPENDIX A

CASE STUDIES

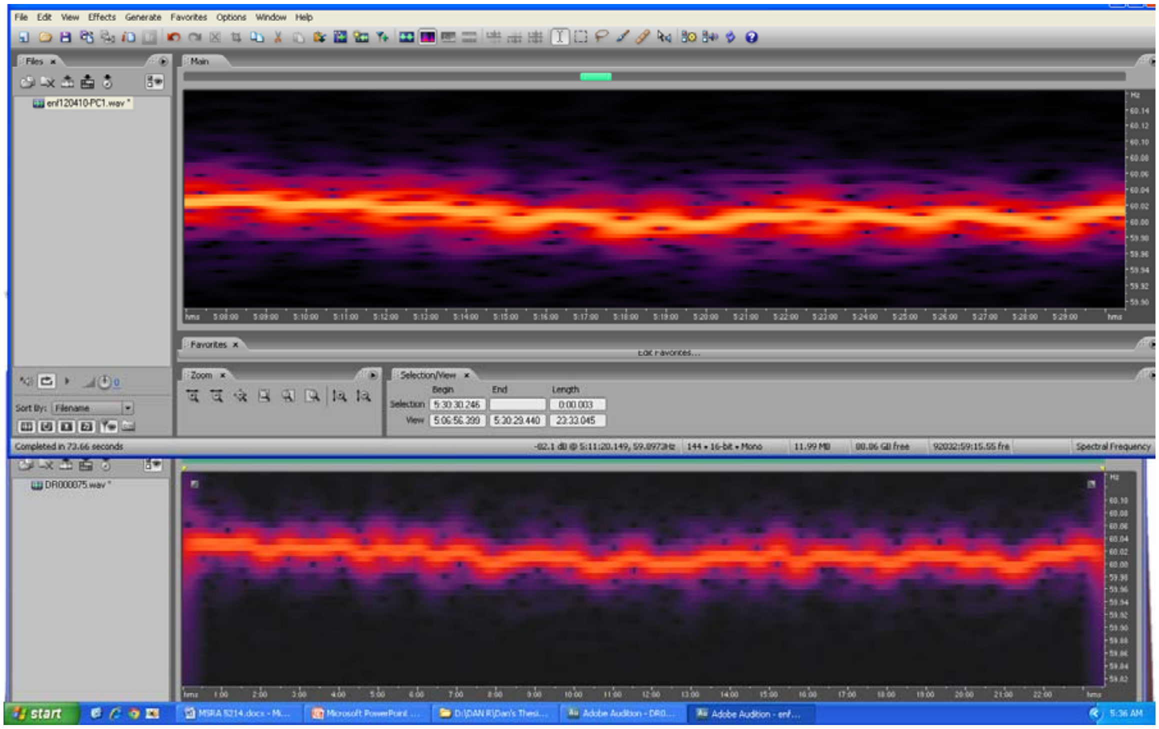
Preface

The following section will demonstrate case studies that are applied to the methodology and authentication framework proposal. These are not quite representative of actual cases because I created the sound files, performed the manipulations (in some cases), and did the analyses with the prior knowledge of the probable outcomes, which is a classic example of expectation bias. Still, I tried to remain as objective as possible. In real cases, the analyst would not know which files were authentic (if any), which were manipulated (if any), and if so, what the manipulations were. Regardless of these limitations, these case studies may be of some use to the reader.

Methods

For the test cases, a single acoustic event was recorded to a number of different recording devices and formats. Some of the resulting files were altered and some were not. Unlike image manipulations that can be demonstrated visually, audio manipulations must be heard aurally. It is unfortunate that the reader cannot experience the content differences between the original and tampered files from reading a paper.

Anticipating that some of the devices would capture an ENF signal in addition to the desired speech (which would need to be obfuscated for an effective manipulation) a private ENF database was collected using a mains powered probe constructed using the schematic outlined by Grigoras, Smith, and Jenkins. [32] The probe steps down the 120V signal from a standard electrical outlet to line level, which can be plugged directly into a digital recording device or sound card. This enables collecting pure ENF signal without added noise, frequency masking, or convolution effects. The recording device used was a Tascam DR-07 portable digital recorder that stores audio files to flash memory. It was set to record a mono 16 bit/44.1 kHz WAV PCM file with all settings flat (no high pass filter). The recording device and the probe were plugged directly into a wall outlet with no power strip, UPS, or surge protector, and the input gain was level 4, which ensured good SNR without clipping. The location of the device was a residence located approximately 1.2 miles from the ENF databases housed at the NCMF, both on the Western Grid. Below are spectrographic and automatic ENF comparisons of the two databases, which provided a match.


File Edit View Effects Generate Favorites Options Window Help
Files
enf120410-PC1.wav *
Main
Hz
60.14
60.12
60.10
60.08
60.06
60.04
60.02
60.00
59.98
59.96
59.94
59.92
59.90
hms 5:08:00 5:09:00 5:10:00 5:11:00 5:12:00 5:13:00 5:14:00 5:15:00 5:16:00 5:17:00 5:18:00 5:19:00 5:20:00 5:21:00 5:22:00 5:23:00 5:24:00 5:25:00 5:26:00 5:27:00 5:28:00 5:29:00 hms
Favorites
Edit Favorites...
Zoom
Sort By: Filename
DR000075.wav *
hms 1:00 2:00 3:00 4:00 5:00 6:00 7:00 8:00 9:00 10:00 11:00 12:00 13:00 14:00 15:00 16:00 17:00 18:00 19:00 20:00 21:00 22:00 hms
start
MSRA 5214.docx - M...
Microsoft PowerPoint ...
D:\DAN R\Dan's Thesi...
Adobe Audition - DR0...
Adobe Audition - enf...
5:36 AM

Selection/View

	Begin	End	Length
Selection	5:30:30.246		0:00.003
View	5:06:56.399	5:30:29.440	23:33.045

Completed in 73.66 seconds | -82.1 dB @ 5:11:20.149, 59.8973Hz | 144 • 16-bit • Mono | 11.99 MB | 80.86 GB free | 92032:59:15.55 free | Spectral Frequency

Figure A.1 Spectrographic Comparison of ENF from NCMF Database (top) to Private Database (bottom). Comparing selection 23:33 in length, both downsampled to 144 Hz. Different recording devices and sound cards were used, which may account for clock drift.

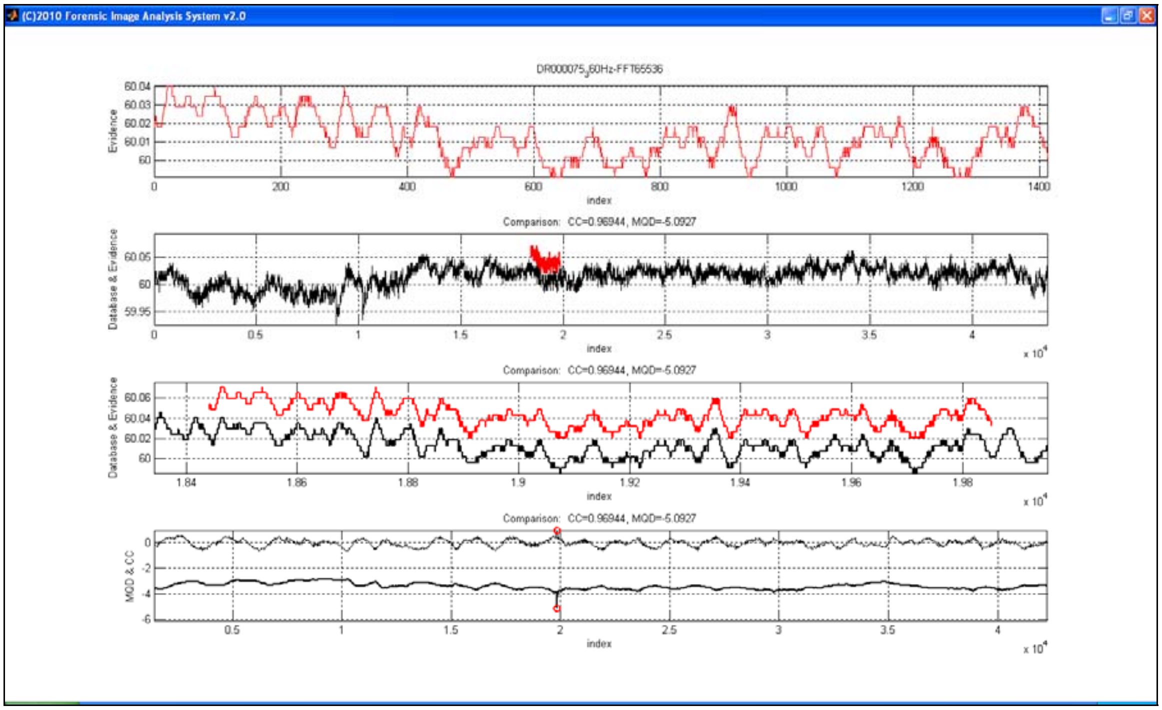


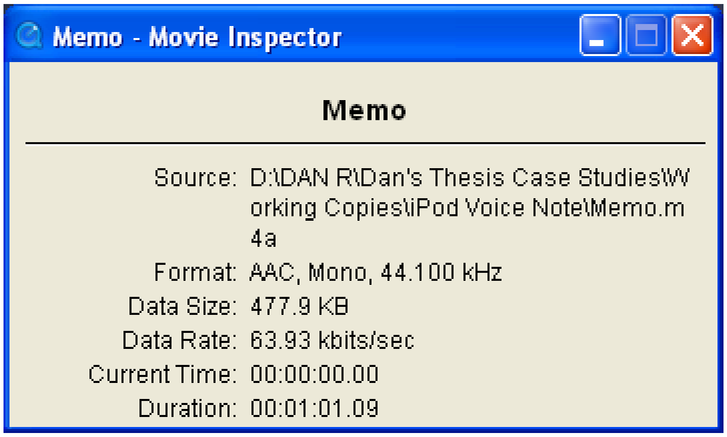
Figure A.2. Automatic Comparison of ENF from NCMF Database (black) to Private Database (red). Both files were downsampled to 360 Hz. Vectors containing peak FFT value per second were compared to find segment containing highest CC and lowest MQD and a relevant match was obtained.

Since all three test cases consisted of lossy compressed files, Butt-Splice Detection analysis and Traces of Interpolation Detection were not used since these are only relevant for non-compressed files. Also, while LTAS and DC were used for global and local analyses, they were not used for device verification. This is beyond the scope of this paper and requires more research and validation.

Test Case #1

An Apple iPod Touch (4th generation) was used to record an oral will and its authenticity has been questioned. The device was mains powered and the internal microphone was used with the built-in Voice Memo application, which has no user defined settings (including gain control). The file was extracted via email and downloaded to a forensic PC workstation. After evidence handling and computing hash values, the native M4A file (which is the audio portion of an MP4 file) was converted to WAV PCM using the QuickTime software application.

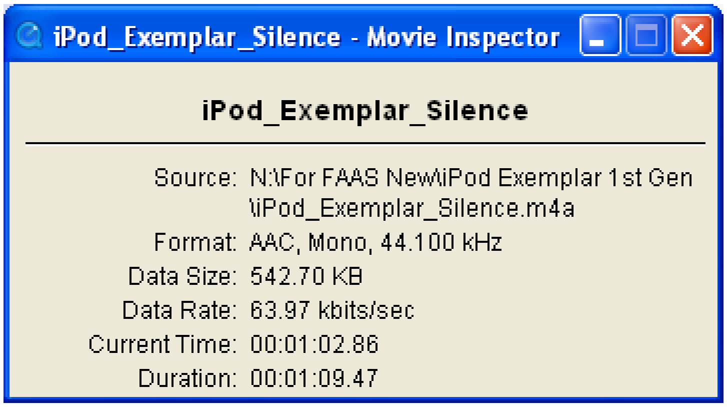
The file format analysis began with a procedural inspection of the sampling rate, bit rate, and channel configuration in QuickTime, which were consistent with exemplar files made on the device.



Memo - Movie Inspector

Memo

Source: D:\DAN R\Dan's Thesis Case Studies\Working Copies\iPod Voice Note\Memo.m4a
Format: AAC, Mono, 44.100 kHz
Data Size: 477.9 KB
Data Rate: 63.93 kbits/sec
Current Time: 00:00:00.00
Duration: 00:01:01.09



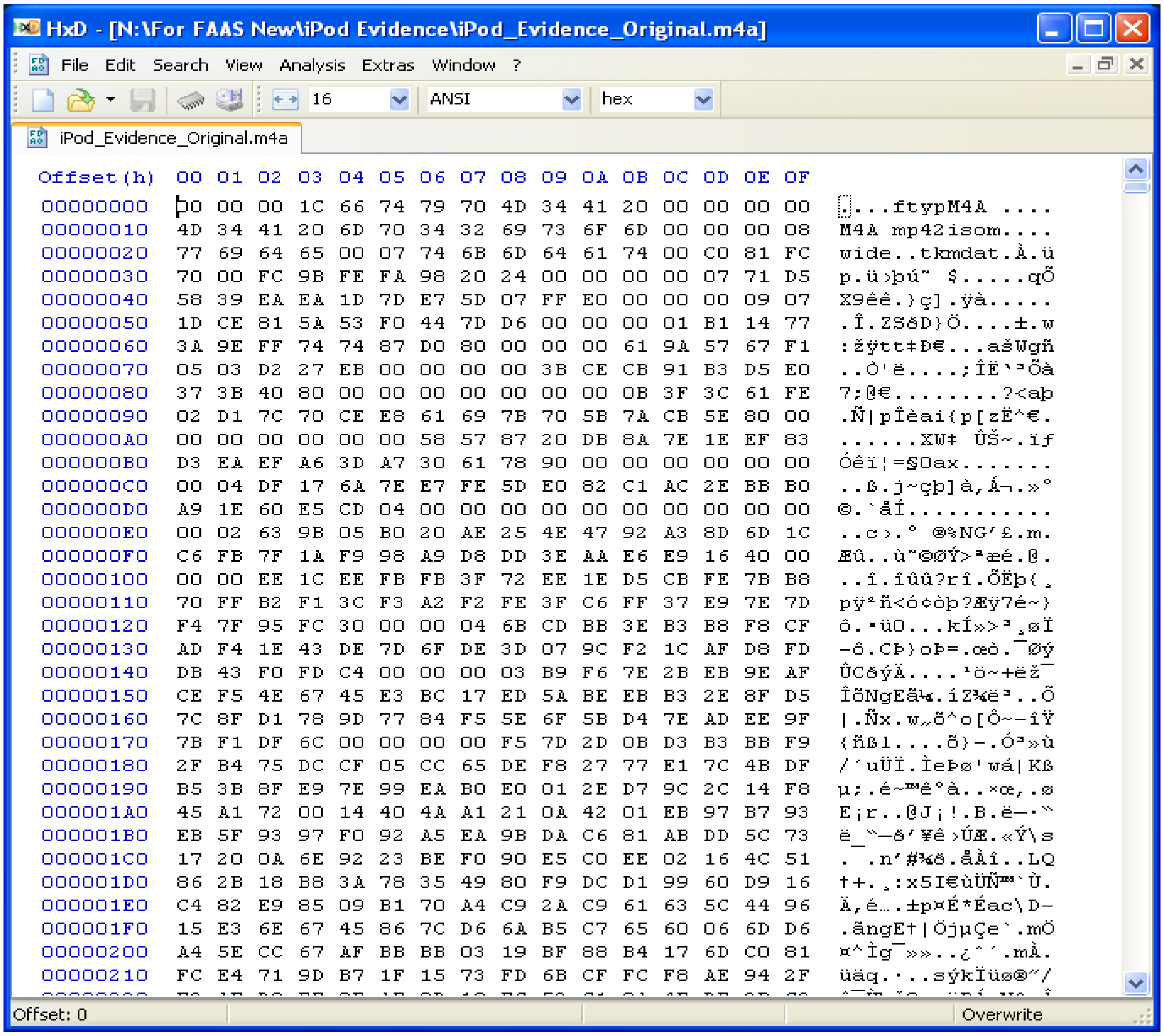
iPod_Exemplar_Silence - Movie Inspector

iPod_Exemplar_Silence

Source: N:\For FAAS New\iPod Exemplar 1st Gen\iPod_Exemplar_Silence.m4a
Format: AAC, Mono, 44.100 kHz
Data Size: 542.70 KB
Data Rate: 63.97 kbits/sec
Current Time: 00:01:02.86
Duration: 00:01:09.47

Figure A.3 File Format Analysis of Evidence (top) and Exemplar (bottom) M4A File Made on iPod Touch. Data rates, channel configuration, and sampling rate are consistent.

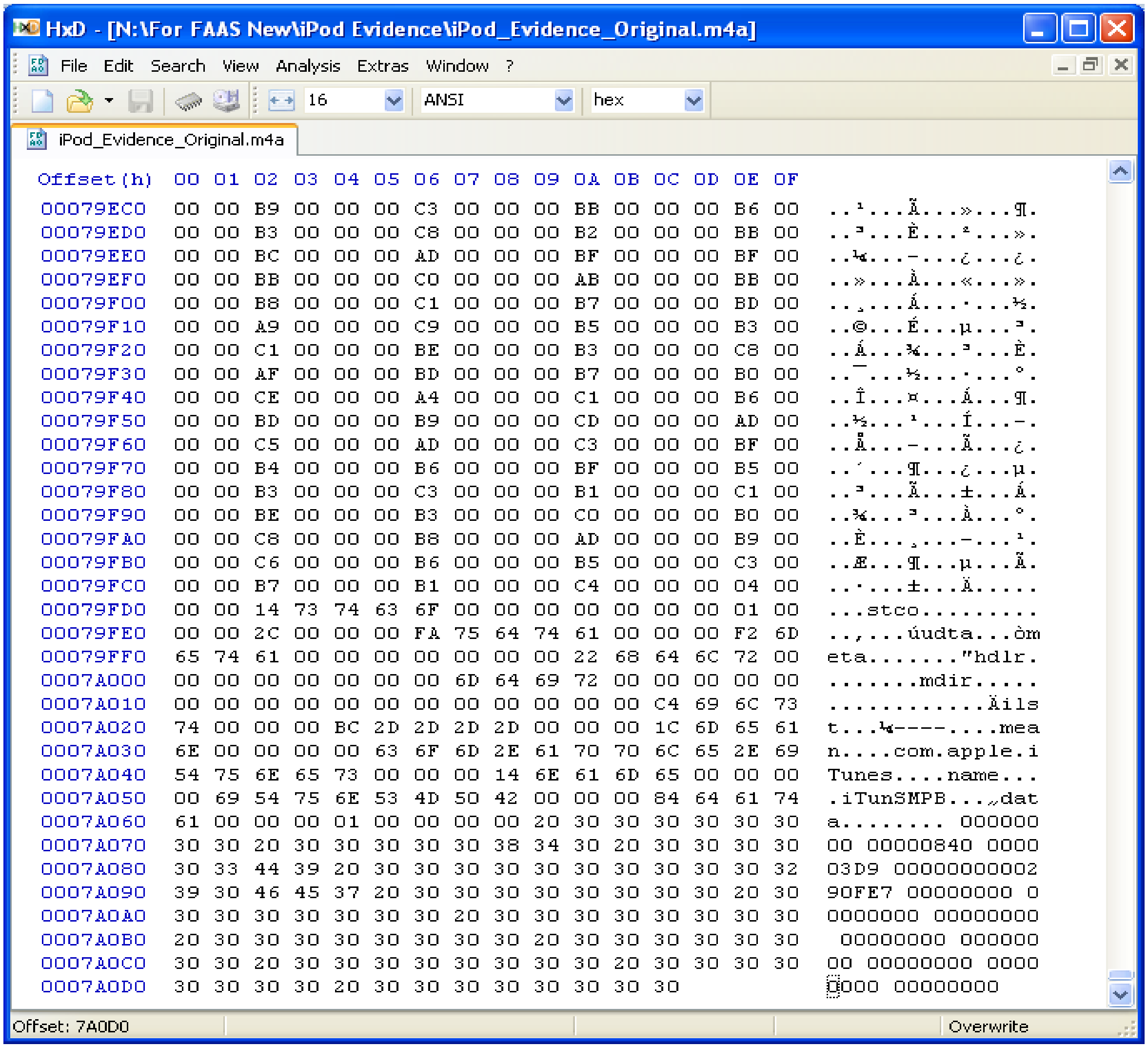
The MAC times were viewed which revealed identical Created and Modified times. Normally in an original file, the Modified time is when the file was first written to the original device and the Created time is when it first interacted with the operating system. However, since the file was extracted through email (no direct extraction from the device was possible), and exemplar files made on the original device produced similar results, this analysis proved consistent with an authentic file. The header and footer were also consistent with unedited exemplars, although they none indicated time stamps.


HxD - [N:\For FAAS New\iPod Evidence\iPod_Evidence_Original.m4a]
File
Edit
Search
View
Analysis
Extras
Window
?
16
ANSI
hex
iPod_Evidence_Original.m4a
Offset: 0
Overwrite

```
Offset(h)  00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00000000   00 00 00 1C 66 74 79 70 4D 34 41 20 00 00 00 00  ....ftypM4A ....
00000010   4D 34 41 20 6D 70 34 32 69 73 6F 6D 00 00 00 08  M4A mp42isom....
00000020   77 69 64 65 00 07 74 6B 6D 64 61 74 00 C0 81 FC  wide..tkmdat.À.ü
00000030   70 00 FC 9B FE FA 98 20 24 00 00 00 00 07 71 D5  p.ü›þú˜ $.....qÕ
00000040   58 39 EA EA 1D 7D E7 5D 07 FF E0 00 00 00 09 07  X9êê.}ç].ÿà.....
00000050   1D CE 81 5A 53 F0 44 7D D6 00 00 00 01 B1 14 77  .Î.ZSðD}Ö....±.w
00000060   3A 9E FF 74 74 87 D0 80 00 00 00 61 9A 57 67 F1  :žÿtt‡Ð€...ašWgñ
00000070   05 03 D2 27 EB 00 00 00 00 3B CE CB 91 B3 D5 E0  ..Ò'ë....;ÎË'³Õà
00000080   37 3B 40 80 00 00 00 00 00 00 00 0B 3F 3C 61 FE  7;@€........?<aþ
00000090   02 D1 7C 70 CE E8 61 69 7B 70 5B 7A CB 5E 80 00  .Ñ|pÎèai{p[zË^€.
000000A0   00 00 00 00 00 00 58 57 87 20 DB 8A 7E 1E EF 83  ......XW‡ ÛŠ~.ïƒ
000000B0   D3 EA EF A6 3D A7 30 61 78 90 00 00 00 00 00 00  Óêï¦=§0ax.......
000000C0   00 04 DF 17 6A 7E E7 FE 5D E0 82 C1 AC 2E BB B0  ..ß.j~çþ]à‚Á¬.»°
000000D0   A9 1E 60 E5 CD 04 00 00 00 00 00 00 00 00 00 00  ©.`åÍ...........
000000E0   00 02 63 9B 05 B0 20 AE 25 4E 47 92 A3 8D 6D 1C  ..c›.° ®%NG'£.m.
000000F0   C6 FB 7F 1A F9 98 A9 D8 DD 3E AA E6 E9 16 40 00  Æû..ù˜©ØÝ>ªæé.@.
00000100   00 00 EE 1C EE FB FB 3F 72 EE 1E D5 CB FE 7B B8  ..î.îûû?rî.ÕËþ{¸
00000110   70 FF B2 F1 3C F3 A2 F2 FE 3F C6 FF 37 E9 7E 7D  pÿ²ñ<ó¢òþ?Æÿ7é~}
00000120   F4 7F 95 FC 30 00 00 04 6B CD BB 3E B3 B8 F8 CF  ô.•ü0...kÍ»>³¸øÏ
00000130   AD F4 1E 43 DE 7D 6F DE 3D 07 9C F2 1C AF D8 FD  -ô.CÞ}oÞ=.œò.¯Øý
00000140   DB 43 F0 FD C4 00 00 00 03 B9 F6 7E 2B EB 9E AF  ÛCðýÄ....¹ö~+ëž¯
00000150   CE F5 4E 67 45 E3 BC 17 ED 5A BE EB B3 2E 8F D5  ÎõNgEã¼.íZ¾ë³..Õ
00000160   7C 8F D1 78 9D 77 84 F5 5E 6F 5B D4 7E AD EE 9F  |.Ñx.w„õ^o[Ô~-îŸ
00000170   7B F1 DF 6C 00 00 00 00 F5 7D 2D 0B D3 B3 BB F9  {ñßl....õ}-.Ó³»ù
00000180   2F B4 75 DC CF 05 CC 65 DE F8 27 77 E1 7C 4B DF  /´uÜÏ.ÌeÞø'wá|Kß
00000190   B5 3B 8F E9 7E 99 EA B0 E0 01 2E D7 9C 2C 14 F8  µ;.é~™ê°à..×œ,.ø
000001A0   45 A1 72 00 14 40 4A A1 21 0A 42 01 EB 97 B7 93  E¡r..@J¡!.B.ë—·"
000001B0   EB 5F 93 97 F0 92 A5 EA 9B DA C6 81 AB DD 5C 73  ë_"—ð'¥ê›ÚÆ.«Ý\s
000001C0   17 20 0A 6E 92 23 BE F0 90 E5 C0 EE 02 16 4C 51  . .n'#¾ð.åÀî..LQ
000001D0   86 2B 18 B8 3A 78 35 49 80 F9 DC D1 99 60 D9 16  †+.¸:x5I€ùÜÑ™`Ù.
000001E0   C4 82 E9 85 09 B1 70 A4 C9 2A C9 61 63 5C 44 96  Ä‚é….±p¤É*Éac\D–
000001F0   15 E3 6E 67 45 86 7C D6 6A B5 C7 65 60 06 6D D6  .ãngE†|ÖjµÇe`.mÖ
00000200   A4 5E CC 67 AF BB BB 03 19 BF 88 B4 17 6D C0 81  ¤^Ìg¯»»..¿ˆ´.mÀ.
00000210   FC E4 71 9D B7 1F 15 73 FD 6B CF FC F8 AE 94 2F  üäq.·..sýkÏüø®"/
```

Figure A.4 Header and Hex Data of M4A File Obtained from iPod Touch.



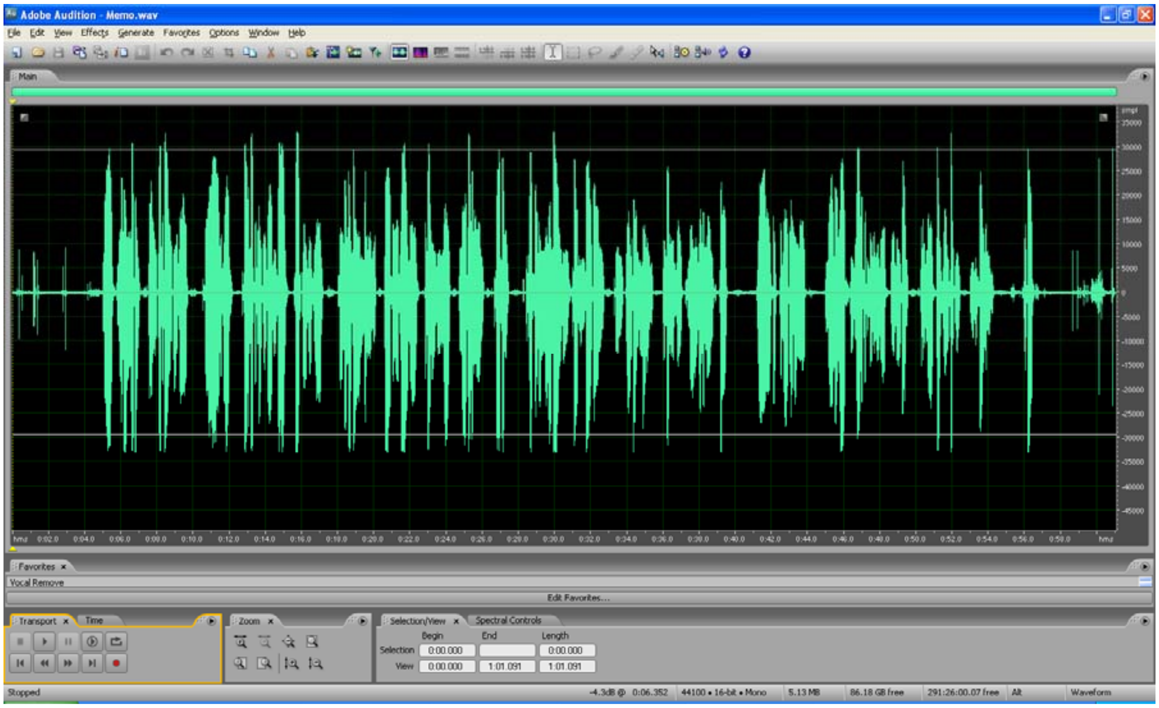
HxD - [N:\For FAAS New\iPod Evidence\iPod_Evidence_Original.m4a]
File Edit Search View Analysis Extras Window ?
16
ANSI
hex
iPod_Evidence_Original.m4a
Offset: 7A0D0
Overwrite

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00079EC0  00 00 B9 00 00 00 C3 00 00 00 BB 00 00 00 B6 00  ..¹...Ã...»...¶.
00079ED0  00 00 B3 00 00 00 C8 00 00 00 B2 00 00 00 BB 00  ..³...È...²...».
00079EE0  00 00 BC 00 00 00 AD 00 00 00 BF 00 00 00 BF 00  ..¼...-...¿...¿.
00079EF0  00 00 BB 00 00 00 C0 00 00 00 AB 00 00 00 BB 00  ..»...À...«...».
00079F00  00 00 B8 00 00 00 C1 00 00 00 B7 00 00 00 BD 00  ..¸...Á...·...½.
00079F10  00 00 A9 00 00 00 C9 00 00 00 B5 00 00 00 B3 00  ..©...É...µ...³.
00079F20  00 00 C1 00 00 00 BE 00 00 00 B3 00 00 00 C8 00  ..Á...¾...³...È.
00079F30  00 00 AF 00 00 00 BD 00 00 00 B7 00 00 00 B0 00  ..¯...½...·...°.
00079F40  00 00 CE 00 00 00 A4 00 00 00 C1 00 00 00 B6 00  ..Î...¤...Á...¶.
00079F50  00 00 BD 00 00 00 B9 00 00 00 CD 00 00 00 AD 00  ..½...¹...Í...-.
00079F60  00 00 C5 00 00 00 AD 00 00 00 C3 00 00 00 BF 00  ..Å...-...Ã...¿.
00079F70  00 00 B4 00 00 00 B6 00 00 00 BF 00 00 00 B5 00  ..´...¶...¿...µ.
00079F80  00 00 B3 00 00 00 C3 00 00 00 B1 00 00 00 C1 00  ..³...Ã...±...Á.
00079F90  00 00 BE 00 00 00 B3 00 00 00 C0 00 00 00 B0 00  ..¾...³...À...°.
00079FA0  00 00 C8 00 00 00 B8 00 00 00 AD 00 00 00 B9 00  ..È...¸...-...¹.
00079FB0  00 00 C6 00 00 00 B6 00 00 00 B5 00 00 00 C3 00  ..Æ...¶...µ...Ã.
00079FC0  00 00 B7 00 00 00 B1 00 00 00 C4 00 00 00 04 00  ..·...±...Ä.....
00079FD0  00 00 14 73 74 63 6F 00 00 00 00 00 00 00 01 00  ...stco.........
00079FE0  00 00 2C 00 00 00 FA 75 64 74 61 00 00 00 F2 6D  ..,...úudta...òm
00079FF0  65 74 61 00 00 00 00 00 00 00 22 68 64 6C 72 00  eta......."hdlr.
0007A000  00 00 00 00 00 00 00 6D 64 69 72 00 00 00 00 00  .......mdir.....
0007A010  00 00 00 00 00 00 00 00 00 00 00 00 C4 69 6C 73  ............Äils
0007A020  74 00 00 00 BC 2D 2D 2D 2D 00 00 00 1C 6D 65 61  t...¼----....mea
0007A030  6E 00 00 00 00 63 6F 6D 2E 61 70 70 6C 65 2E 69  n....com.apple.i
0007A040  54 75 6E 65 73 00 00 00 14 6E 61 6D 65 00 00 00  Tunes....name...
0007A050  00 69 54 75 6E 53 4D 50 42 00 00 00 84 64 61 74  .iTunSMPB...„dat
0007A060  61 00 00 00 01 00 00 00 00 20 30 30 30 30 30 30  a........ 000000
0007A070  30 30 20 30 30 30 30 30 38 34 30 20 30 30 30 30  00 00000840 0000
0007A080  30 33 44 39 20 30 30 30 30 30 30 30 30 30 30 32  03D9 0000000002
0007A090  39 30 46 45 37 20 30 30 30 30 30 30 30 30 20 30  90FE7 00000000 0
0007A0A0  30 30 30 30 30 30 30 20 30 30 30 30 30 30 30 30  0000000 00000000
0007A0B0  20 30 30 30 30 30 30 30 30 20 30 30 30 30 30 30   00000000 000000
0007A0C0  30 30 20 30 30 30 30 30 30 30 30 20 30 30 30 30  00 00000000 0000
0007A0D0  30 30 30 30 20 30 30 30 30 30 30 30 30           0000 00000000
```

Figure A.5 Footer and Hex Data of M4A File Obtained from iPod Touch.

Next, critical listening and waveform analysis were performed using Adobe Audition 3.0. Playback was optimized according to the standards set forth by Koenig and Lacey. [4] The length of the recording was determined to be 1:01:091 (in minutes:seconds:milliseconds). A questionable sound, possibly handling noise, was noted and visually inspected at approximately 0:44 but no signs of manipulation were found in subsequent analyses. High resolution analysis showed 67 milliseconds (ms) of zero amplitude samples at the end of the file following the termination of acoustic information. This is attributed to the M4A encoding algorithm. Similar samples were not found at the beginning of the file. An exemplar recording had 58 ms of samples at the end, and a re-encoded second generation file had 48 ms.

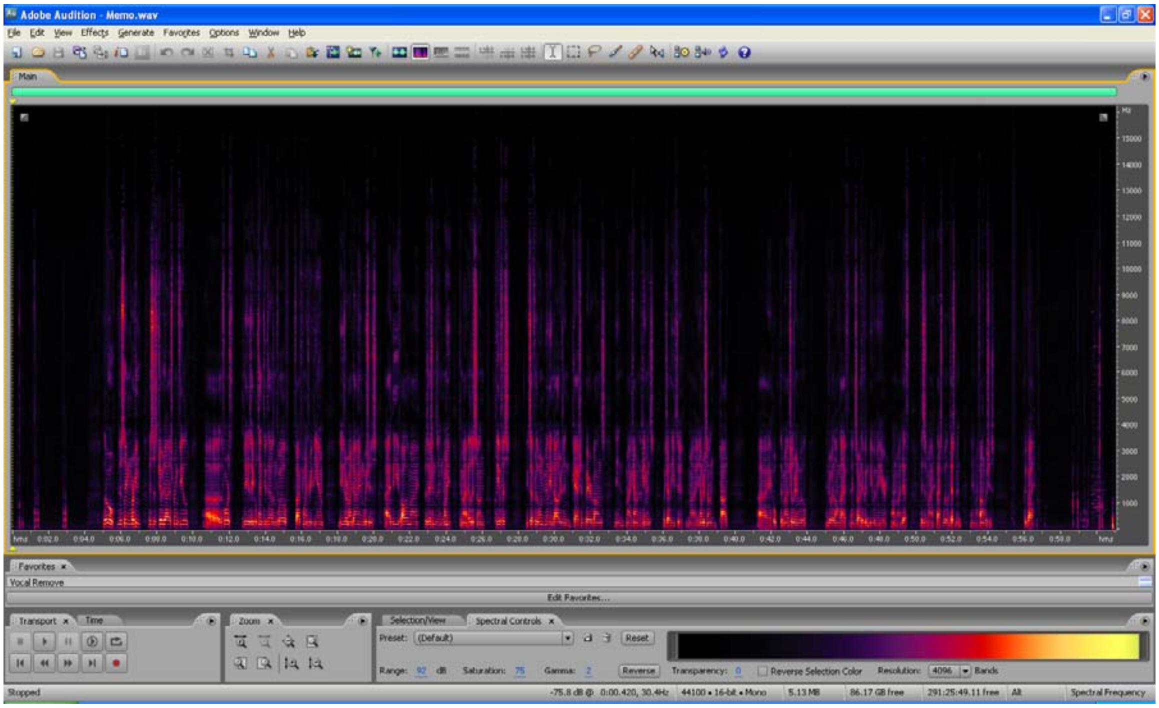


	Begin	End	Length
Selection	0:00.000		0:00.000
View	0:00.000	1:01.091	1:01.091

Stopped -4.3dB @ 0:06.352 44100 • 16-bit • Mono 5.13 MB 86.18 GB free 291:26:00.07 free Alt Waveform

Figure A.6 Waveform Analysis of M4A File.

Spectrum/spectrographic analysis were also consistent with an authentic recording. No steady tones for phase analysis or sufficient ENF signal were found even though the device was mains powered, so those techniques were not available for this case.


Adobe Audition - Memo.wav
File Edit View Effects Generate Favorites Options Window Help
Main
Hz
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
hms 0:02.0 0:04.0 0:06.0 0:08.0 0:10.0 0:12.0 0:14.0 0:16.0 0:18.0 0:20.0 0:22.0 0:24.0 0:26.0 0:28.0 0:30.0 0:32.0 0:34.0 0:36.0 0:38.0 0:40.0 0:42.0 0:44.0 0:46.0 0:48.0 0:50.0 0:52.0 0:54.0 0:56.0 0:58.0 hms
Favorites
Vocal Remove
Edit Favorites...
Transport
Time
Zoom
Selection/View
Spectral Controls
Reset
Reverse

Preset: (Default)

Range: 92 dB Saturation: 75 Gamma: 2 Transparency: 0 □ Reverse Selection Color Resolution: 4096 Bands

Stopped -75.8 dB @ 0:00.420, 30.4Hz 44100 • 16-bit • Mono 5.13 MB 86.17 GB free 291:25:49.11 free All Spectral Frequency

Figure A.7 Spectrographic Analysis of M4A Evidence File. FFT resolution is 4096 bands; Blackmann-Harris windowing. No ENF found. No inconsistencies or steady tones.

For DC analysis, the entire file was analyzed and no indications of gain changes or copy/paste were detected. Then, a WAV file was prepared with speech removed from the questioned recording to obtain only noise. This was compared to a noise-only exemplar recording made on the same device in the same location, with the built-in mic and Voice Memo application. The mean and standard deviation were consistent for both files, as seen in the figures below.

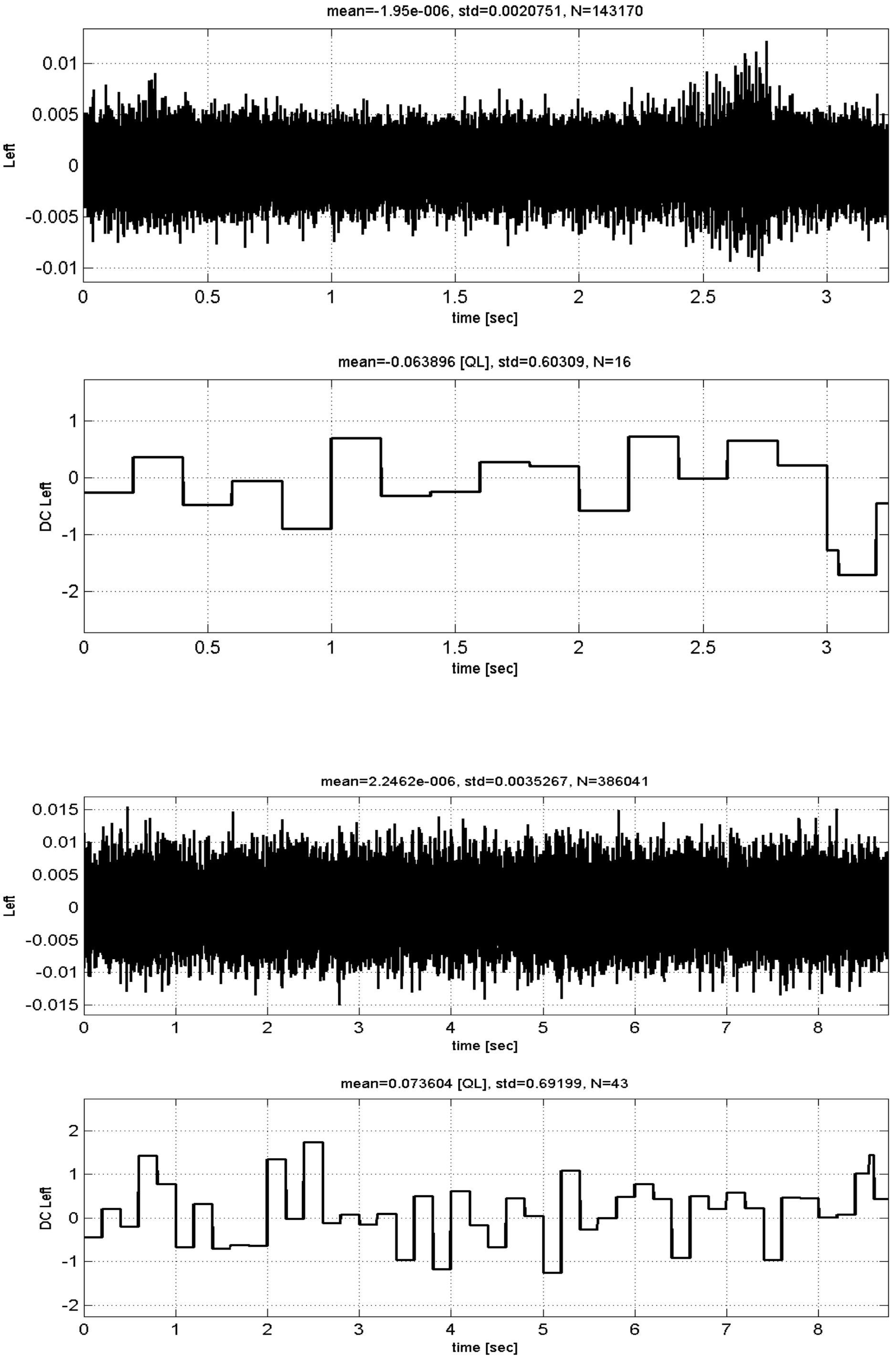
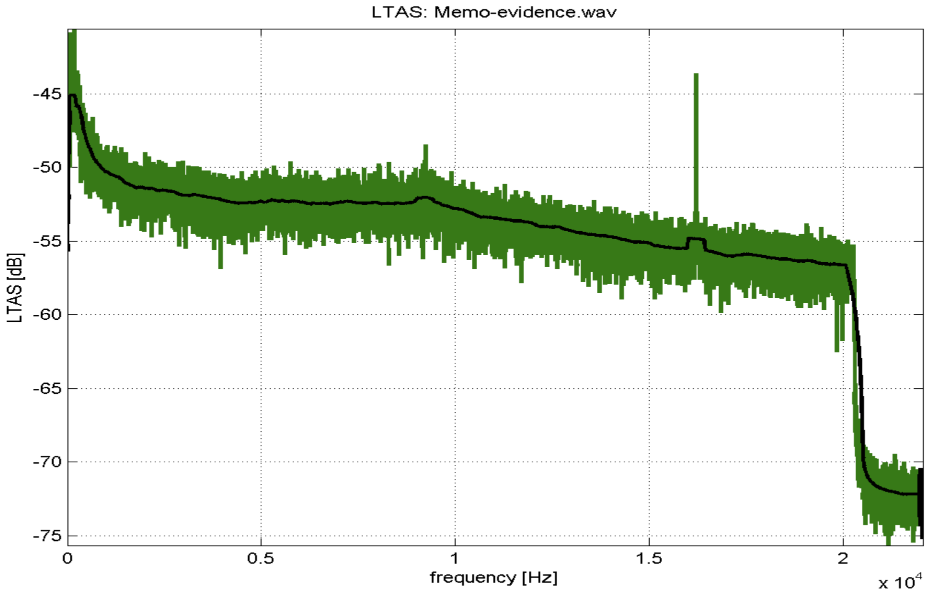


Figure A.8 DC Analysis of Noise from Questioned M4A Recording (top) and Exemplar (bottom).

The LTAS of the questioned file as well as an exemplar recorded on the same device both had a steep cut off at approximately 18 to 21 kHz and were consistent.



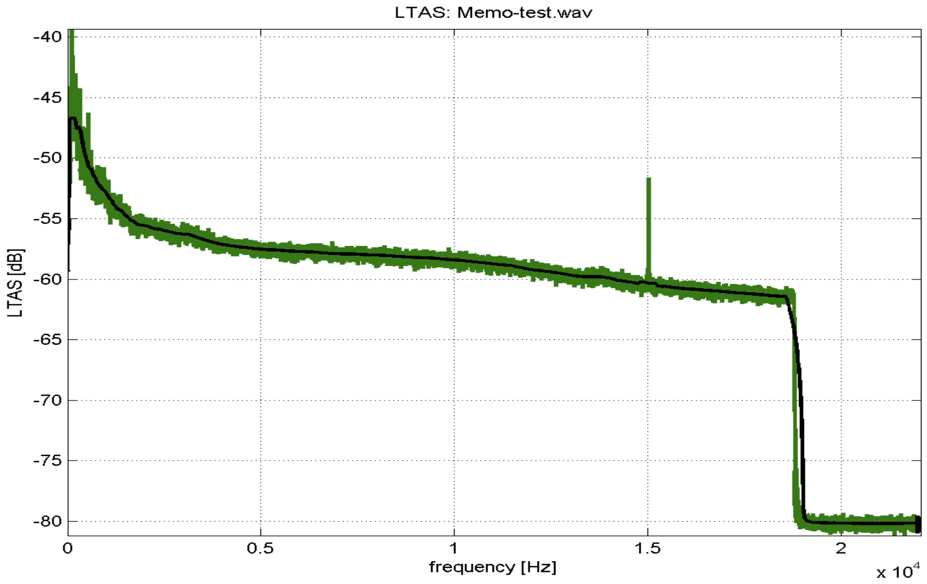


Figure A.9 LTAS of Evidence (top) and Exemplar (bottom) M4A Files.

Sorted Spectrum analysis was consistent with the characteristics of the M4A file format in test recordings.

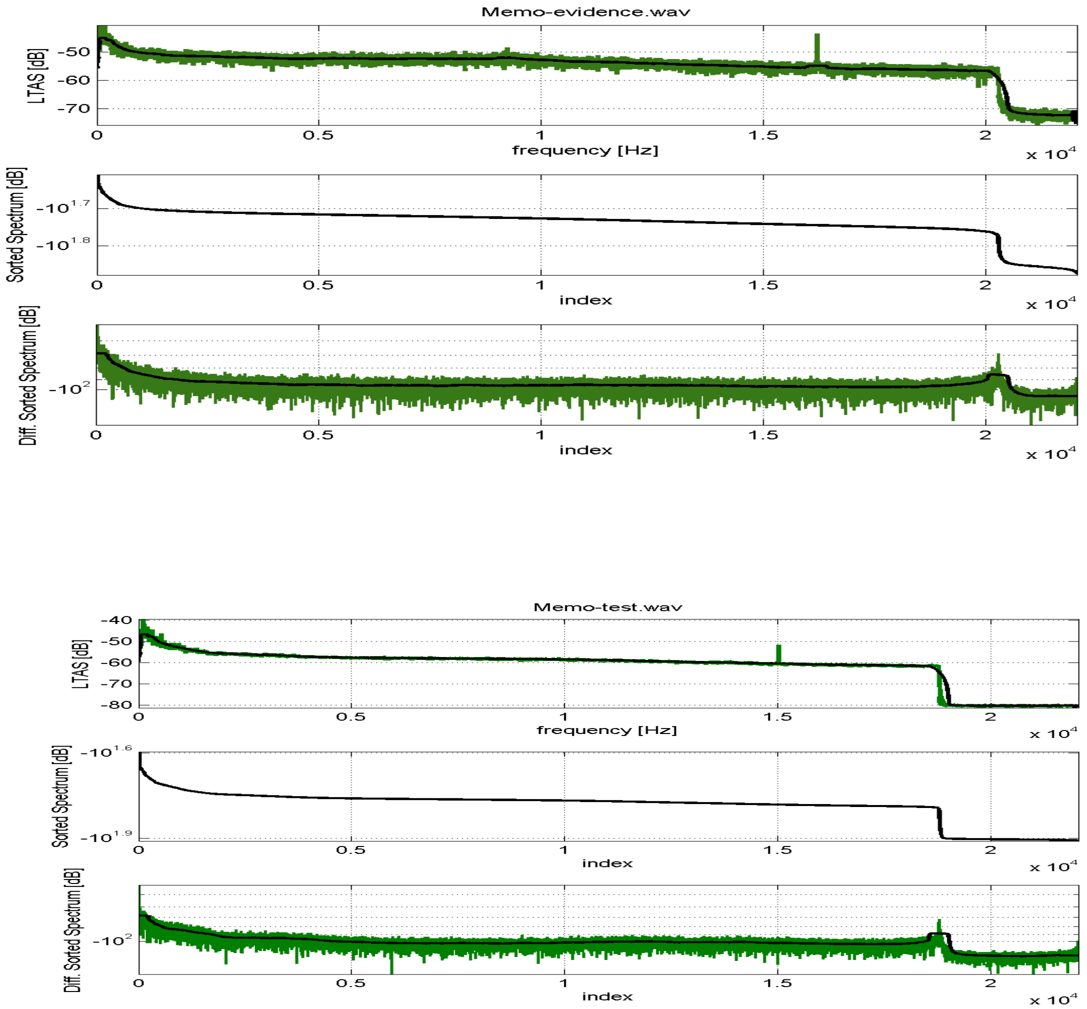


Figure A.10 Sorted Spectrum of Questioned Recording (top) and Exemplar (bottom).

Finally, Compression Analysis (*CL*) of non-speech was consistent with a first generation exemplar, whereas a second generation exemplar showed a much wider gap. It is likely the questioned recording is a first generation M4A file as claimed by the contributor and verified by *CL* analysis.

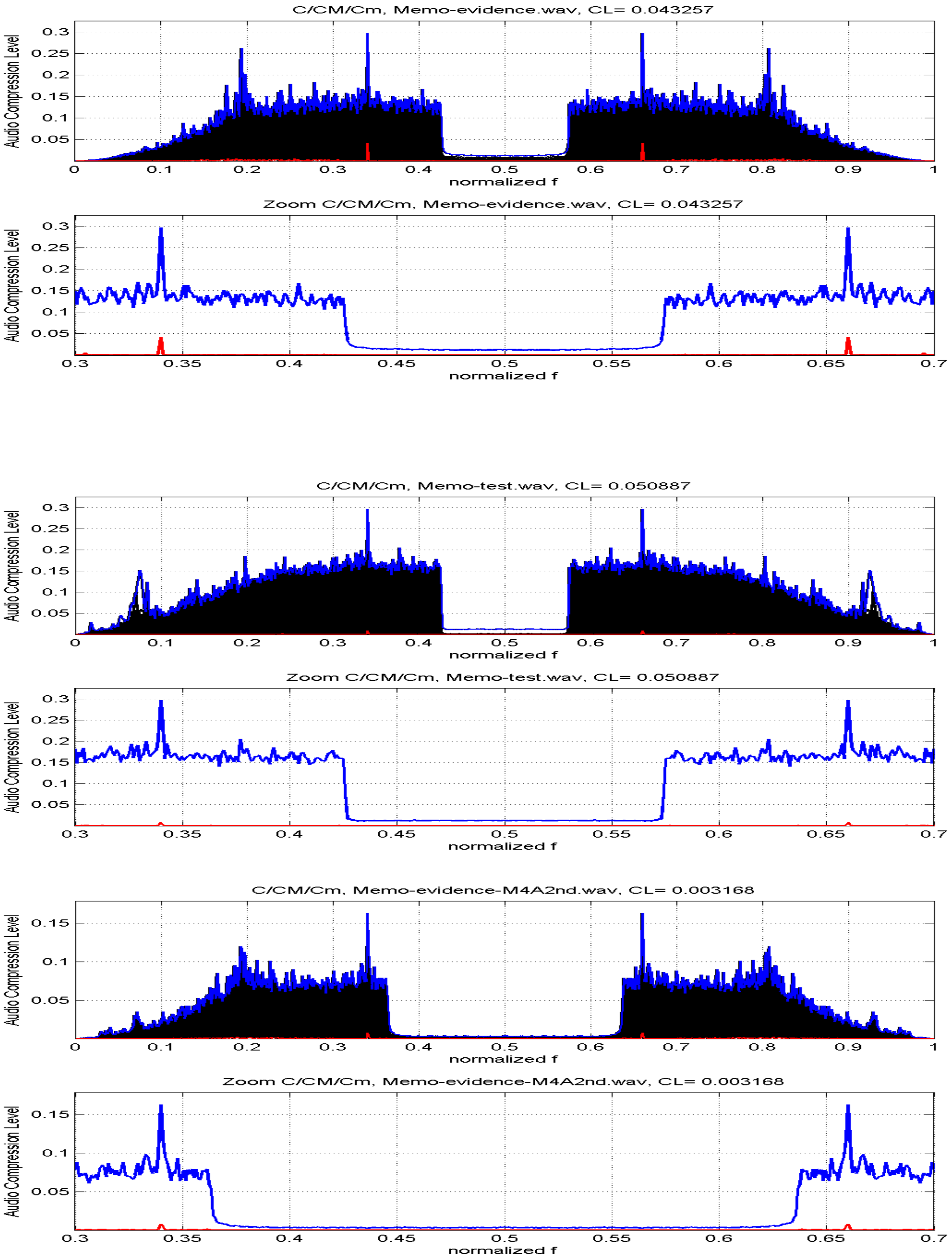


Figure A.11 Compression Analysis for iPod Voice Memo M4A Evidence File (top), Exemplar (middle), and Second Generation Exemplar (bottom).

The interpretations of all the analyses are presented in the table below. The examiner's ultimate conclusion is that the questioned file is *consistent* with an authentic recording.

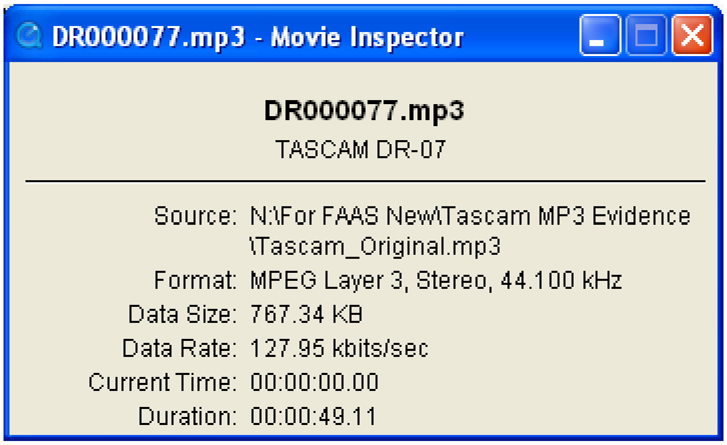
Table A.1 Authentication Framework for Case #1.

File Structure	File Format	*C*
	Header	*C*
	Hex Data	*C*
Global Analysis	DC Offset	*C*
	LTAS, Sorted Spec., Differentiated Sorted Spec.	*C*
	Compression Level Analysis	*C*
Local Analysis	Critical Listening	*C*
	Waveform Analysis	*C*
	Spectrum/Spectrogram Analysis	*C*
	DC Offset	*C*
Device Verification	Header (exemplars)	*C*
	Hex Data (exemplars)	*C*

Test Case #2

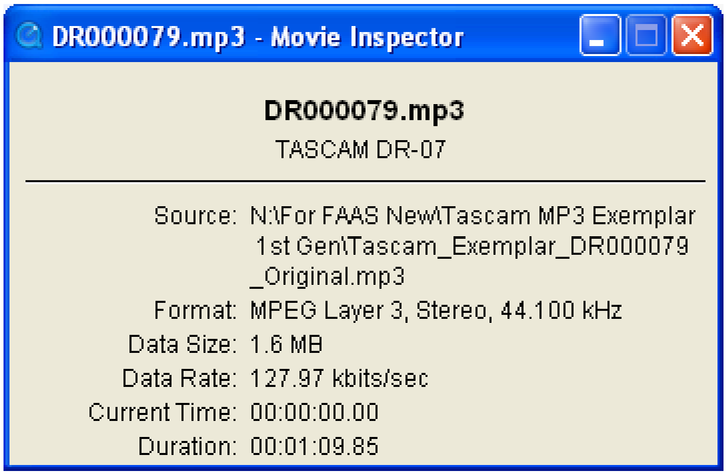
The same acoustic event as in the previous test case, an oral will, was simultaneously captured to a Tascam DR-07 portable digital recorder. Using the internal microphones and mains power source (directly connected to outlet with no power strip), a stereo MP3 file was created at 44.1 kHz and 128 kbps. The file was edited in Adobe Audition, saved as a second generation MP3 and presented as the original evidence.

File Format analysis was consistent with an exemplar recording.


DR000077.mp3 - Movie Inspector

DR000077.mp3
TASCAM DR-07

Source: N:\For FAAS New\Tascam MP3 Evidence \Tascam_Original.mp3
Format: MPEG Layer 3, Stereo, 44.100 kHz
Data Size: 767.34 KB
Data Rate: 127.95 kbits/sec
Current Time: 00:00:00.00
Duration: 00:00:49.11

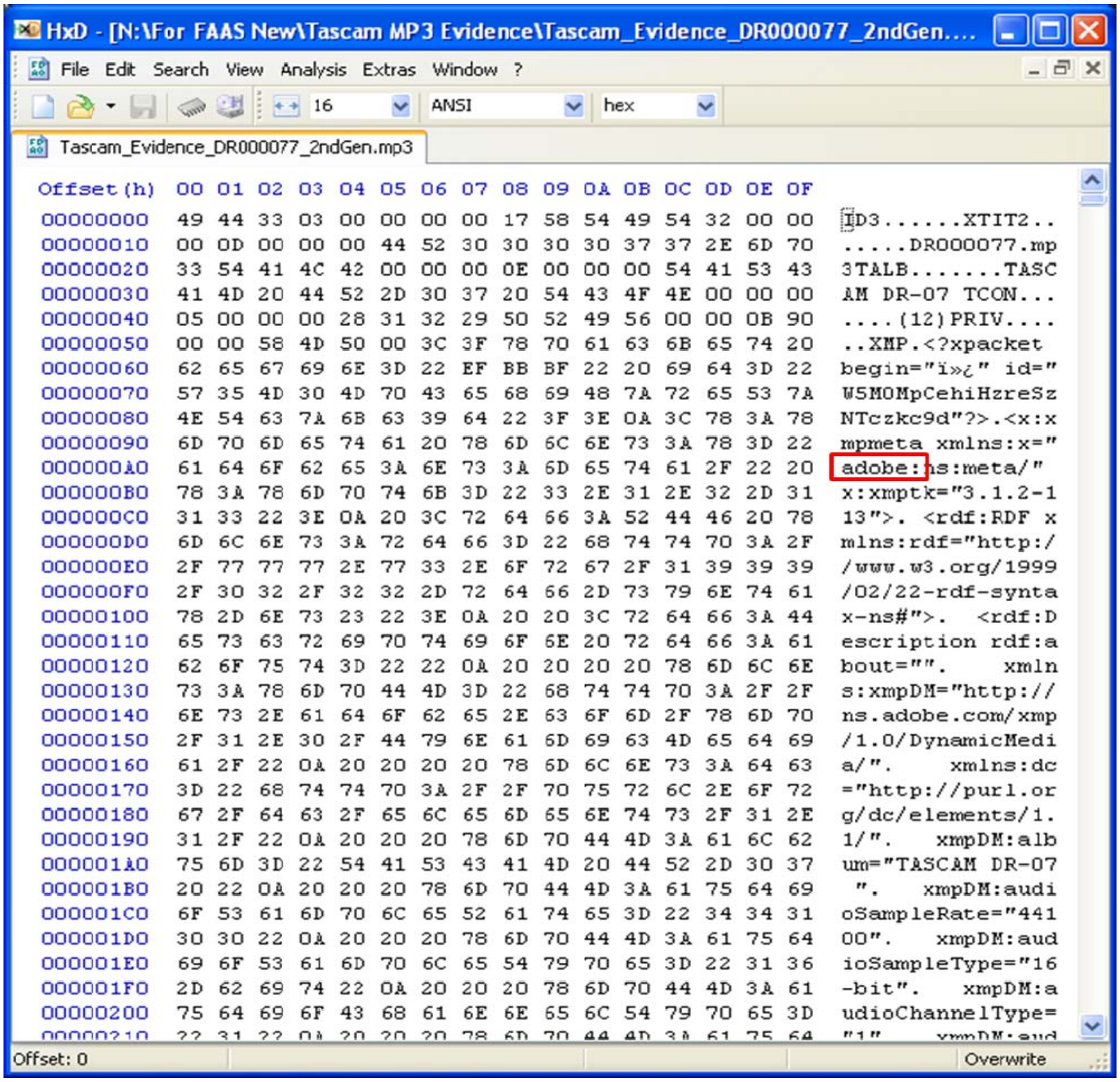

DR000079.mp3 - Movie Inspector

DR000079.mp3
TASCAM DR-07

Source: N:\For FAAS New\Tascam MP3 Exemplar 1st Gen\Tascam_Exemplar_DR000079 _Original.mp3
Format: MPEG Layer 3, Stereo, 44.100 kHz
Data Size: 1.6 MB
Data Rate: 127.97 kbits/sec
Current Time: 00:00:00.00
Duration: 00:01:09.85

Figure A.12 File Format Analysis of Evidence (top) and Exemplar (bottom).

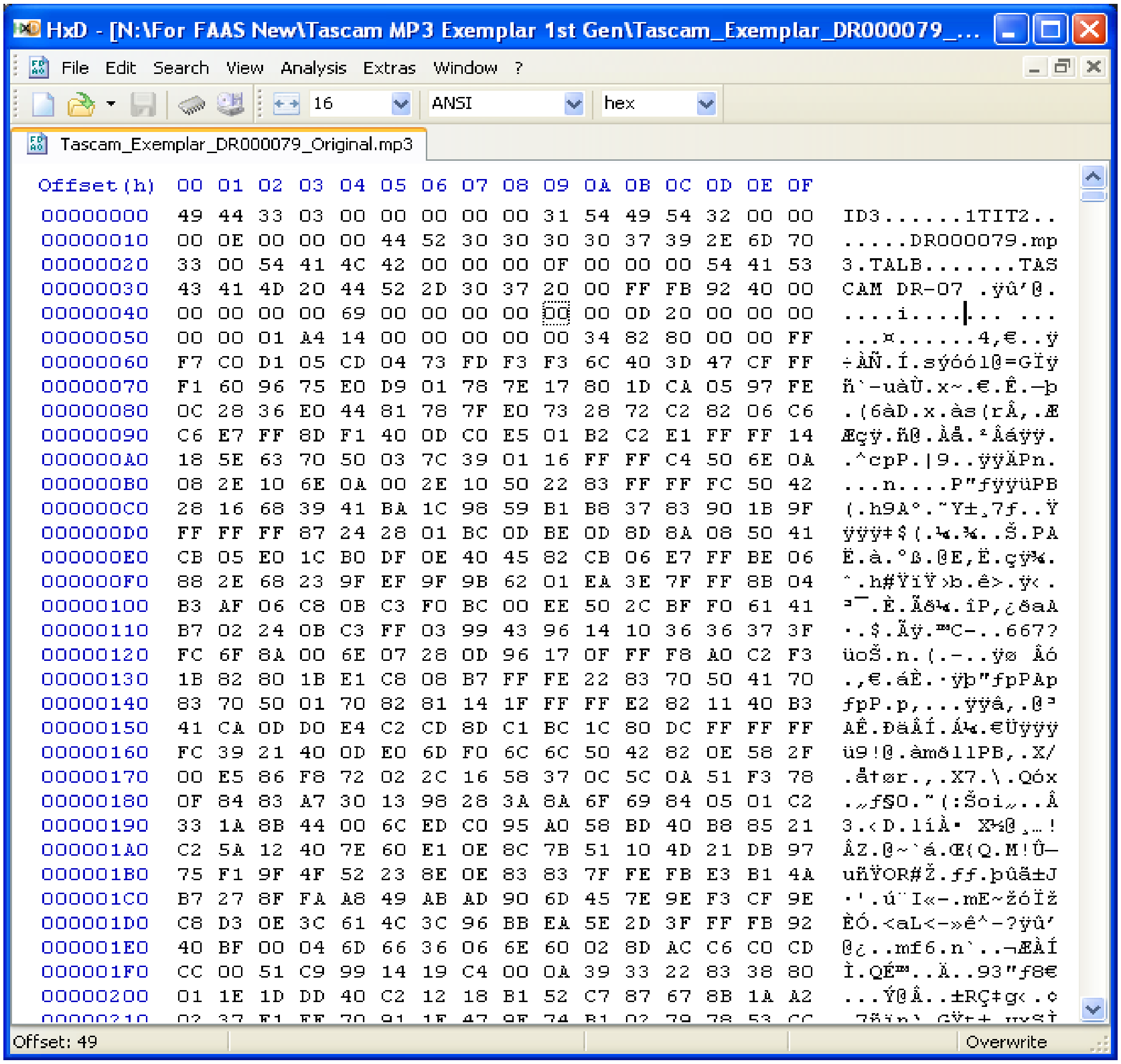
Header analysis showed evidence of Adobe software in the questioned recording but not in the exemplar. Hex data showed differences a well that were inconsistent with an authentic file.



HxD - [N:\For FAAS New\Tascam MP3 Evidence\Tascam_Evidence_DR000077_2ndGen....
File
Edit
Search
View
Analysis
Extras
Window
?
16
ANSI
hex
Tascam_Evidence_DR000077_2ndGen.mp3
Offset: 0
Overwrite

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00000000  49 44 33 03 00 00 00 00 17 58 54 49 54 32 00 00  ID3......XTIT2..
00000010  00 0D 00 00 00 44 52 30 30 30 30 37 37 2E 6D 70  .....DR000077.mp
00000020  33 54 41 4C 42 00 00 00 0E 00 00 00 54 41 53 43  3TALB.......TASC
00000030  41 4D 20 44 52 2D 30 37 20 54 43 4F 4E 00 00 00  AM DR-07 TCON...
00000040  05 00 00 00 28 31 32 29 50 52 49 56 00 00 0B 90  ....(12)PRIV....
00000050  00 00 58 4D 50 00 3C 3F 78 70 61 63 6B 65 74 20  ..XMP.<?xpacket 
00000060  62 65 67 69 6E 3D 22 EF BB BF 22 20 69 64 3D 22  begin="ï»¿" id="
00000070  57 35 4D 30 4D 70 43 65 68 69 48 7A 72 65 53 7A  W5M0MpCehiHzreSz
00000080  4E 54 63 7A 6B 63 39 64 22 3F 3E 0A 3C 78 3A 78  NTczkc9d"?>.<x:x
00000090  6D 70 6D 65 74 61 20 78 6D 6C 6E 73 3A 78 3D 22  mpmeta xmlns:x="
000000A0  61 64 6F 62 65 3A 6E 73 3A 6D 65 74 61 2F 22 20  adobe:ns:meta/" 
000000B0  78 3A 78 6D 70 74 6B 3D 22 33 2E 31 2E 32 2D 31  x:xmptk="3.1.2-1
000000C0  31 33 22 3E 0A 20 3C 72 64 66 3A 52 44 46 20 78  13">. <rdf:RDF x
000000D0  6D 6C 6E 73 3A 72 64 66 3D 22 68 74 74 70 3A 2F  mlns:rdf="http:/
000000E0  2F 77 77 77 2E 77 33 2E 6F 72 67 2F 31 39 39 39  /www.w3.org/1999
000000F0  2F 30 32 2F 32 32 2D 72 64 66 2D 73 79 6E 74 61  /02/22-rdf-synta
00000100  78 2D 6E 73 23 22 3E 0A 20 20 3C 72 64 66 3A 44  x-ns#">.  <rdf:D
00000110  65 73 63 72 69 70 74 69 6F 6E 20 72 64 66 3A 61  escription rdf:a
00000120  62 6F 75 74 3D 22 22 0A 20 20 20 20 78 6D 6C 6E  bout="".    xmln
00000130  73 3A 78 6D 70 44 4D 3D 22 68 74 74 70 3A 2F 2F  s:xmpDM="http://
00000140  6E 73 2E 61 64 6F 62 65 2E 63 6F 6D 2F 78 6D 70  ns.adobe.com/xmp
00000150  2F 31 2E 30 2F 44 79 6E 61 6D 69 63 4D 65 64 69  /1.0/DynamicMedi
00000160  61 2F 22 0A 20 20 20 20 78 6D 6C 6E 73 3A 64 63  a/".    xmlns:dc
00000170  3D 22 68 74 74 70 3A 2F 2F 70 75 72 6C 2E 6F 72  ="http://purl.or
00000180  67 2F 64 63 2F 65 6C 65 6D 65 6E 74 73 2F 31 2E  g/dc/elements/1.
00000190  31 2F 22 0A 20 20 20 78 6D 70 44 4D 3A 61 6C 62  1/".   xmpDM:alb
000001A0  75 6D 3D 22 54 41 53 43 41 4D 20 44 52 2D 30 37  um="TASCAM DR-07
000001B0  20 22 0A 20 20 20 78 6D 70 44 4D 3A 61 75 64 69   ".   xmpDM:audi
000001C0  6F 53 61 6D 70 6C 65 52 61 74 65 3D 22 34 34 31  oSampleRate="441
000001D0  30 30 22 0A 20 20 20 78 6D 70 44 4D 3A 61 75 64  00".   xmpDM:aud
000001E0  69 6F 53 61 6D 70 6C 65 54 79 70 65 3D 22 31 36  ioSampleType="16
000001F0  2D 62 69 74 22 0A 20 20 20 78 6D 70 44 4D 3A 61  -bit".   xmpDM:a
00000200  75 64 69 6F 43 68 61 6E 6E 65 6C 54 79 70 65 3D  udioChannelType=
00000210  22 31 22 0A 20 20 20 78 6D 70 44 4D 3A 61 75 64  "1".   xmpDM:aud
```


HxD - [N:\For FAAS New\Tascam MP3 Exemplar 1st Gen\Tascam_Exemplar_DR000079_...
File
Edit
Search
View
Analysis
Extras
Window
?
16
ANSI
hex
Tascam_Exemplar_DR000079_Original.mp3
Offset: 49
Overwrite

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00000000  49 44 33 03 00 00 00 00 00 31 54 49 54 32 00 00  ID3......1TIT2..
00000010  00 0E 00 00 00 44 52 30 30 30 30 37 39 2E 6D 70  .....DR000079.mp
00000020  33 00 54 41 4C 42 00 00 00 0F 00 00 00 54 41 53  3.TALB.......TAS
00000030  43 41 4D 20 44 52 2D 30 37 20 00 FF FB 92 40 00  CAM DR-07 .ÿû'@.
00000040  00 00 00 00 69 00 00 00 00 00 00 0D 20 00 00 00  ....i....... ...
00000050  00 00 01 A4 14 00 00 00 00 00 34 82 80 00 00 FF  ...¤......4,€..ÿ
00000060  F7 C0 D1 05 CD 04 73 FD F3 F3 6C 40 3D 47 CF FF  ÷ÀÑ.Í.sýóól@=GÏÿ
00000070  F1 60 96 75 E0 D9 01 78 7E 17 80 1D CA 05 97 FE  ñ`–uàÙ.x~.€.Ê.—þ
00000080  0C 28 36 E0 44 81 78 7F E0 73 28 72 C2 82 06 C6  .(6àD.x.às(rÂ,.Æ
00000090  C6 E7 FF 8D F1 40 0D C0 E5 01 B2 C2 E1 FF FF 14  Æçÿ.ñ@.Àå.²Âáÿÿ.
000000A0  18 5E 63 70 50 03 7C 39 01 16 FF FF C4 50 6E 0A  .^cpP.|9..ÿÿÄPn.
000000B0  08 2E 10 6E 0A 00 2E 10 50 22 83 FF FF FC 50 42  ...n....P"ƒÿÿüPB
000000C0  28 16 68 39 41 BA 1C 98 59 B1 B8 37 83 90 1B 9F  (.h9Aº.˜Y±¸7ƒ..Ÿ
000000D0  FF FF FF 87 24 28 01 BC 0D BE 0D 8D 8A 08 50 41  ÿÿÿ‡$(.¼.¾..Š.PA
000000E0  CB 05 E0 1C B0 DF 0E 40 45 82 CB 06 E7 FF BE 06  Ë.à.°ß.@E,Ë.çÿ¾.
000000F0  88 2E 68 23 9F EF 9F 9B 62 01 EA 3E 7F FF 8B 04  ˆ.h#ŸïŸ›b.ê>.ÿ‹.
00000100  B3 AF 06 C8 0B C3 F0 BC 00 EE 50 2C BF F0 61 41  ³¯.È.Ãð¼.îP,¿ðaA
00000110  B7 02 24 0B C3 FF 03 99 43 96 14 10 36 36 37 3F  ·.$.Ãÿ.™C–..667?
00000120  FC 6F 8A 00 6E 07 28 0D 96 17 0F FF F8 A0 C2 F3  üoŠ.n.(.–..ÿø Âó
00000130  1B 82 80 1B E1 C8 08 B7 FF FE 22 83 70 50 41 70  .,€.áÈ.·ÿþ"ƒpPAp
00000140  83 70 50 01 70 82 81 14 1F FF FF E2 82 11 40 B3  ƒpP.p,...ÿÿâ,.@³
00000150  41 CA 0D D0 E4 C2 CD 8D C1 BC 1C 80 DC FF FF FF  AÊ.ÐäÂÍ.Á¼.€Üÿÿÿ
00000160  FC 39 21 40 0D E0 6D F0 6C 6C 50 42 82 0E 58 2F  ü9!@.àmðllPB,.X/
00000170  00 E5 86 F8 72 02 2C 16 58 37 0C 5C 0A 51 F3 78  .å†ør.,.X7.\.Qóx
00000180  0F 84 83 A7 30 13 98 28 3A 8A 6F 69 84 05 01 C2  .„ƒ§0.˜(:Šoi„..Â
00000190  33 1A 8B 44 00 6C ED C0 95 A0 58 BD 40 B8 85 21  3.‹D.líÀ• X½@¸…!
000001A0  C2 5A 12 40 7E 60 E1 0E 8C 7B 51 10 4D 21 DB 97  ÂZ.@~`á.Œ{Q.M!Û—
000001B0  75 F1 9F 4F 52 23 8E 0E 83 83 7F FE FB E3 B1 4A  uñŸOR#Ž.ƒƒ.þûã±J
000001C0  B7 27 8F FA A8 49 AB AD 90 6D 45 7E 9E F3 CF 9E  ·'.ú¨I«.mE~žóÏž
000001D0  C8 D3 0E 3C 61 4C 3C 96 BB EA 5E 2D 3F FF FB 92  ÈÓ.<aL<–»ê^-?ÿû'
000001E0  40 BF 00 04 6D 66 36 06 6E 60 02 8D AC C6 C0 CD  @¿..mf6.n`..¬ÆÀÍ
000001F0  CC 00 51 C9 99 14 19 C4 00 0A 39 33 22 83 38 80  Ì.QÉ™..Ä..93"ƒ8€
00000200  01 1E 1D DD 40 C2 12 18 B1 52 C7 87 67 8B 1A A2  ...Ý@Â..±RÇ‡g‹.¢
```

Figure A.13 Header Analysis of Evidence (top) and Exemplar (bottom).

File structure analysis has already determined that the container is not authentic, but it is possible that the external software was used innocuously and the contents are still unaltered. This will be determined in global and local analyses.

DC analysis on the entire evidence file showed no indications of gain changes or copy/paste. A non-speech version of the evidence file was compared to a silent exemplar that was recorded with the internal microphones and was consistent, having close mean and standard deviations. Device verification using DC offset was not performed.

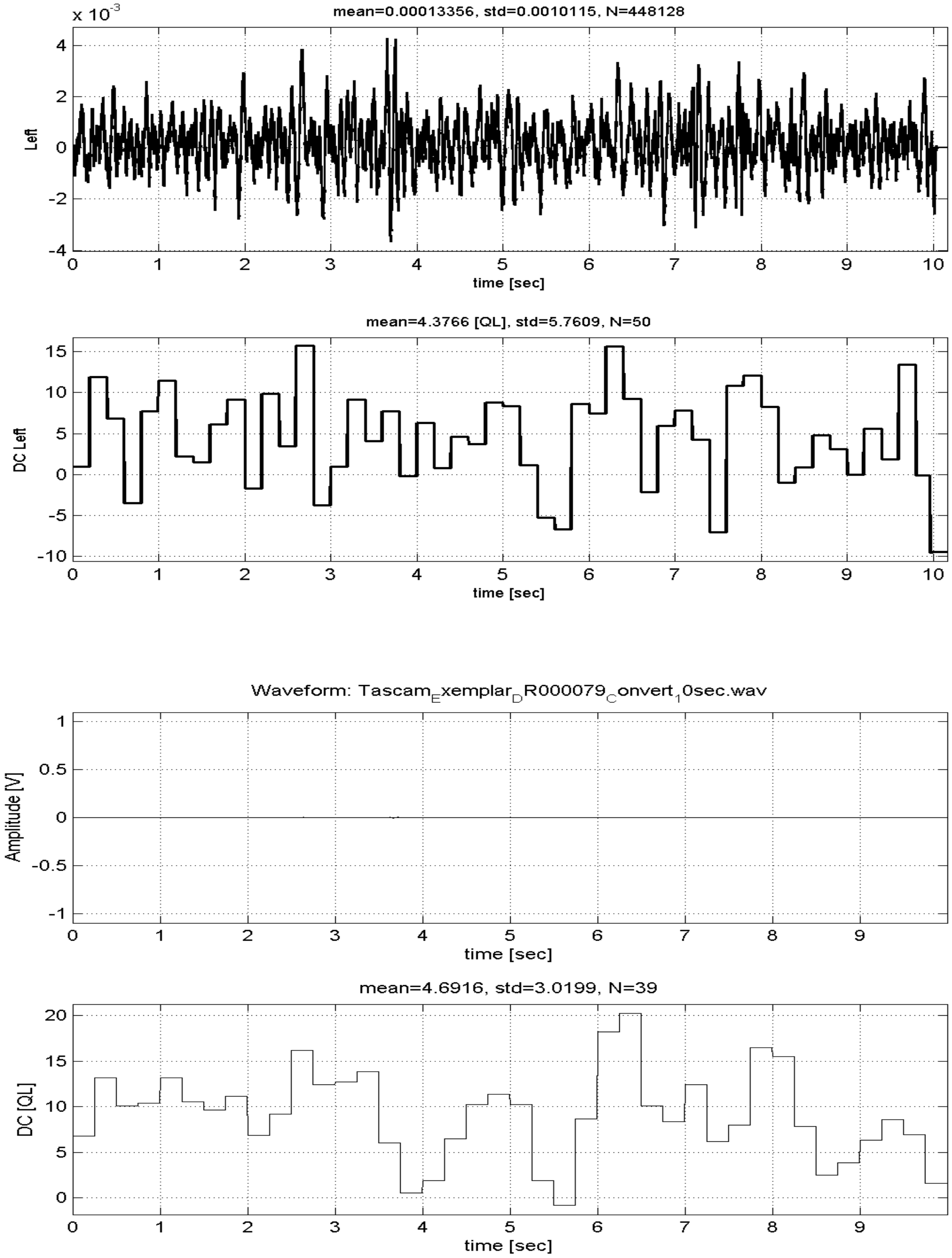
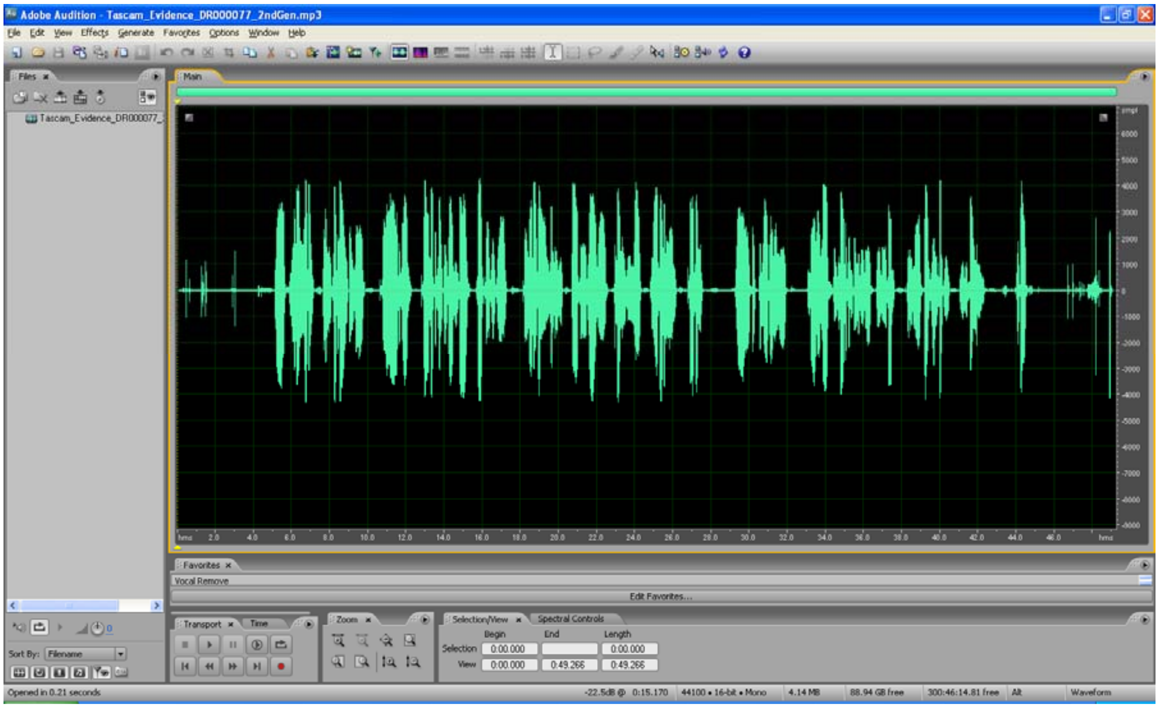


Figure A.14 DC Offset Analysis of Evidence (top) and Exemplar (bottom).

Critical listening was consistent with an authentic recording. Waveform analysis did not reveal any visible edits or ROI. However, there were 83 ms of zero amplitude samples at the beginning of the evidence file and 129 ms at the end. While some preceding zero samples would be expected in an MP3 recording, it was inconsistent with a first generation exemplar, which had 50 ms at the beginning and none at the end. Interestingly, a second generation exemplar had approximately 20 ms at the beginning and 120 ms at the end. The evidence is inconsistent with an authentic, original recording in the claimed format.



Adobe Audition - Tascam_Evidence_DR000077_2ndGen.mp3

Tascam_Evidence_DR000077_

Selection: Begin 0:00.000, End (blank), Length 0:00.000

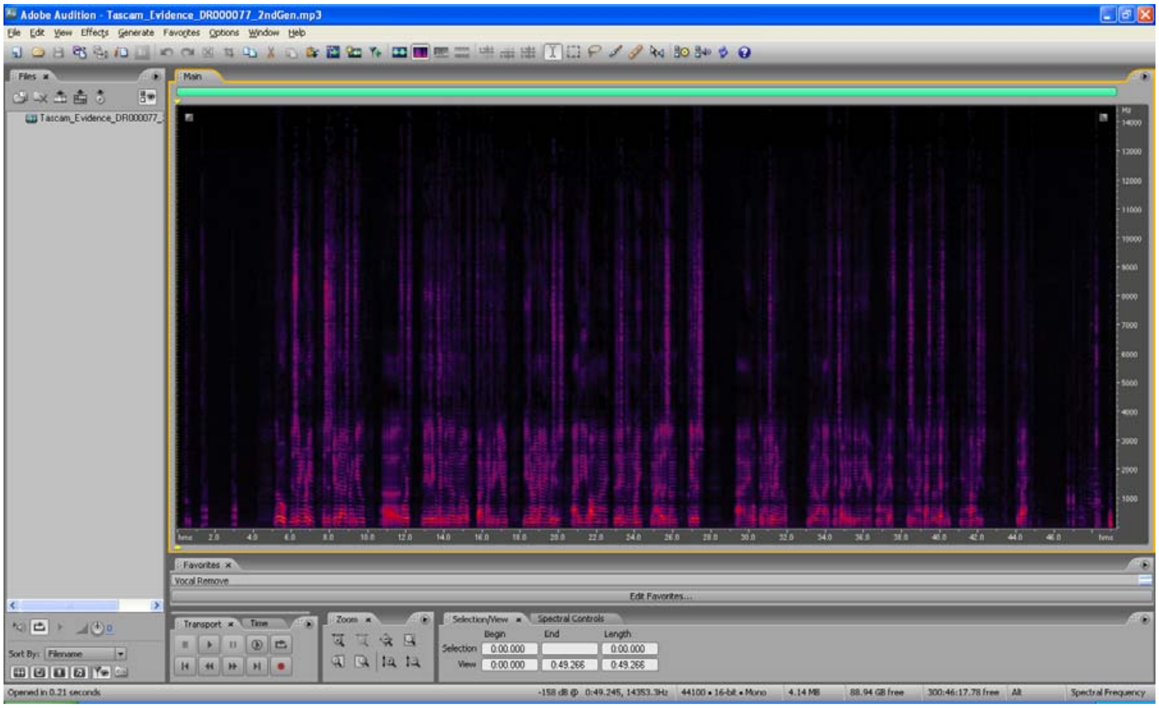
View: Begin 0:00.000, End 0:49.266, Length 0:49.266

Opened in 0.21 seconds

-22.5dB @ 0:15.170 | 44100 • 16-bit • Mono | 4.14 MB | 88.94 GB free | 300:46:14.81 free | Alt | Waveform

Figure A.15 Waveform Analysis of Tascam MP3 Evidence.

Spectrographic analysis did not find any inconsistencies, but there was not sufficient ENF signal present to do a database comparison or phase analysis.



Adobe Audition - Tascam_Evidence_DR000077_2ndGen.mp3

Tascam_Evidence_DR000077_

Selection/View

	Begin	End	Length
Selection	0:00.000		0:00.000
View	0:00.000	0:49.266	0:49.266

Sort By: Filename

Opened in 0.21 seconds | -158 dB @ 0:49.245, 14353.3Hz | 44100 • 16-bit • Mono | 4.14 MB | 88.94 GB free | 300:46:17.78 free | Alt | Spectral Frequency

Figure A.16 Spectrographic Analysis of Tascam MP3 Evidence. FFT Resolution is 8192 bands; Blackmann-Harris Windowing.

LTAS comparisons between evidence and exemplars were consistent.

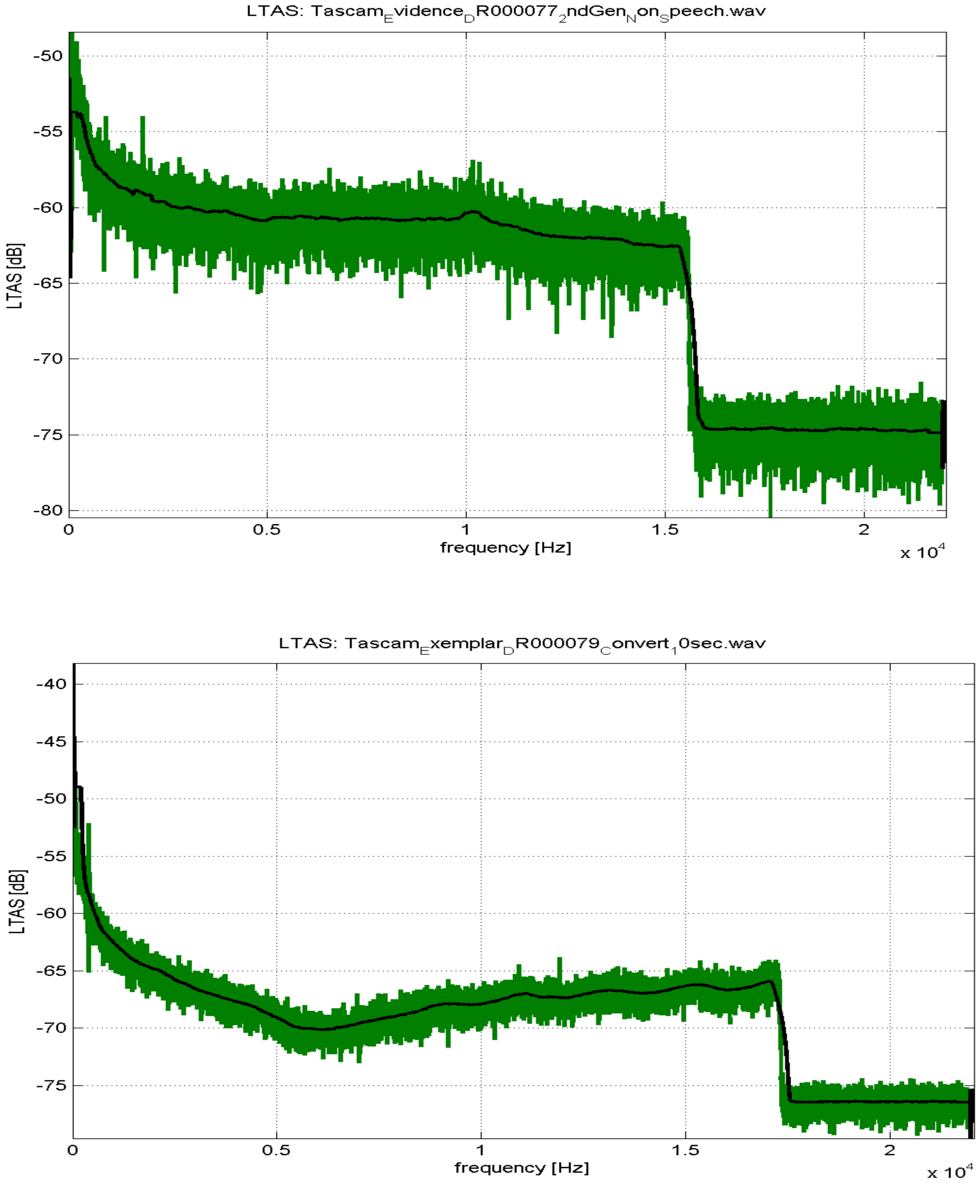


Figure A.17 LTAS of Tascam MP3 Evidence (top) and Exemplar (bottom).

Sorted spectrum analysis showed indications of lossy compression, as indicated by a bump in the curve that represents a steep drop off in amplitude beyond a cutoff frequency. Conversely, uncompressed files exhibit a smooth curve. The analysis of the evidence was consistent with an authentic MP3 exemplar.

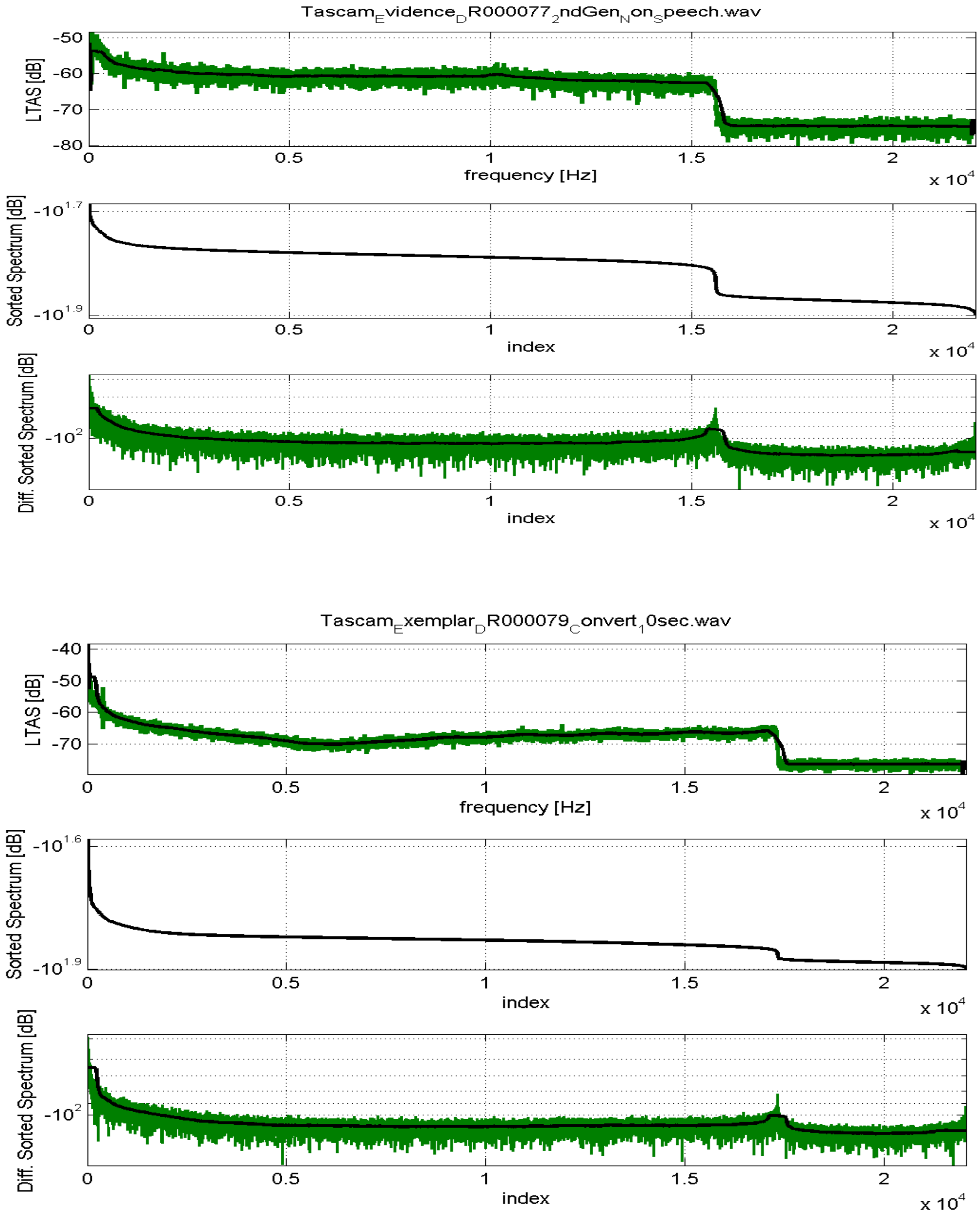


Figure A.18 Sorted Spectrum of Evidence (top) and Exemplar (bottom).

Compression Level Analysis showed a higher correlation between the evidence recording and a second generation exemplar, as opposed to a first generation. This is a very strong indication that the questioned recording has been re-encoded to MP3 and is not original, which corroborates with the results of waveform analysis.

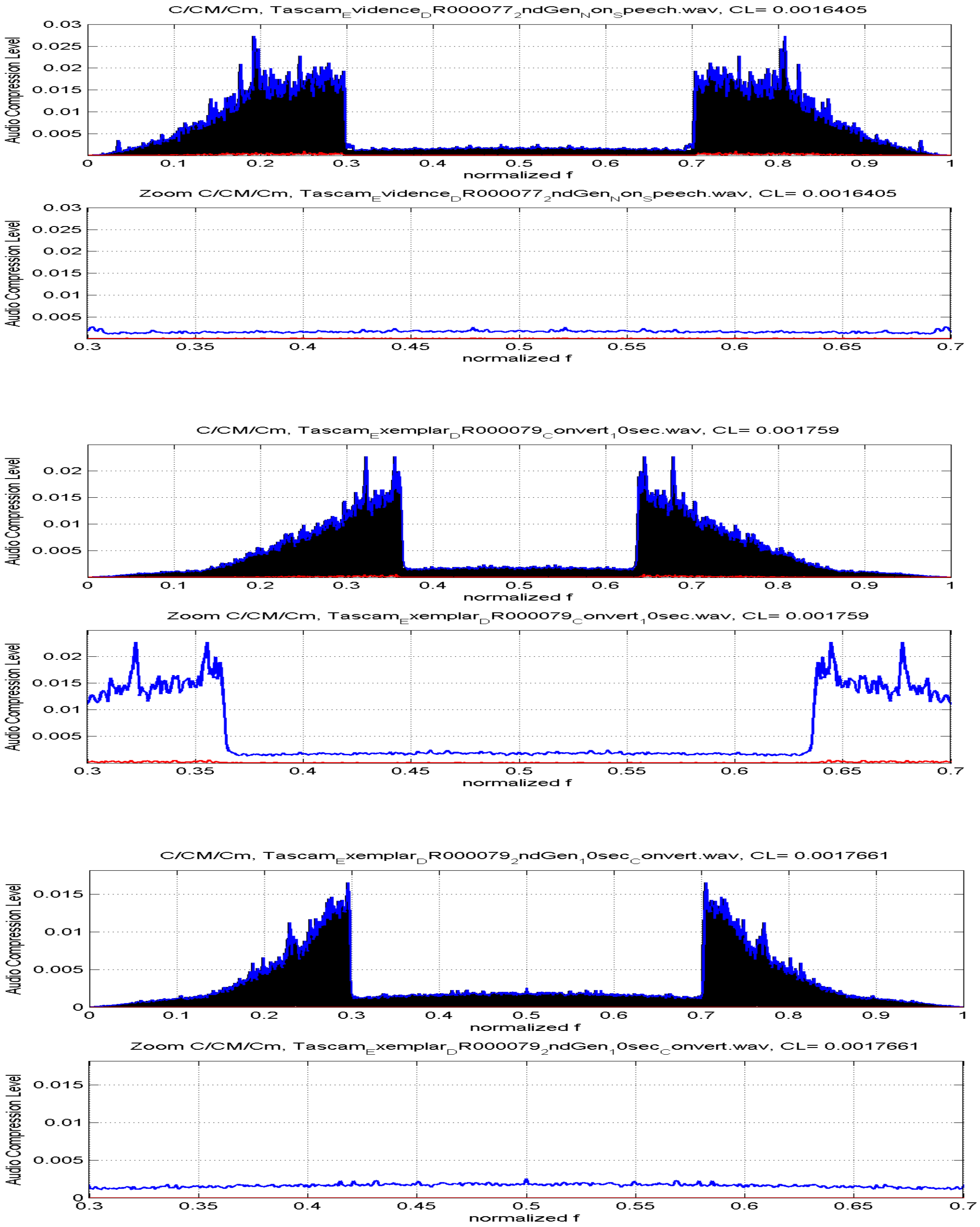


Figure A.19 Compression Level Analysis of Questioned MP3 (top), First Generation Exemplar (middle), and Second Generation Exemplar (bottom).

The examiner’s ultimate conclusion, as presented in the table below, is that the questioned recording is *inconsistent* with an authentic file made with the device and methods claimed by the contributor. Although no malicious alterations were found and the content may still be accurate, the file is not original and is very likely a second generation MP3 that interacted with digital editing software.

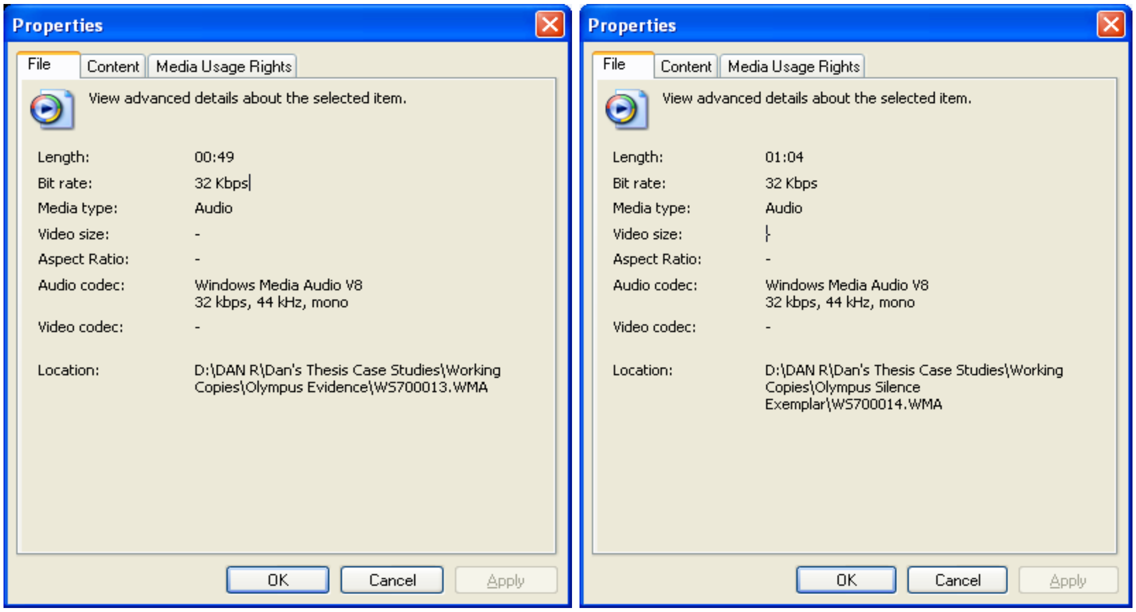
Table A.2 Authentication Framework for Case #2.

File Structure	File Format	*C*
	Header	*I*
	Hex Data	*I*
Global Analysis	DC Offset	*C*
	LTAS, Sorted Spec., Differentiated Sorted Spec.	*C*
	Compression Level Analysis	*I*
Local Analysis	Critical Listening	*C*
	Waveform Analysis	*I*
	Spectrum/Spectrogram Analysis	*C*
	DC Offset	*C*
Device Verification	Header (exemplars)	*I*
	Hex Data (exemplars)	*I*

Test Case #3

This case also involved a manipulated recording that was presented as original. The original iPod Touch M4A voice memo file from case #1 was converted to WAV PCM and editing using Pro Tools software on a laptop computer. A few sentences were deleted from the oral will, which changed the meaning of the acoustic events (and who received the inheritance). The split regions were concatenated with a crossfade instead of butt-splice or interpolation. The edit sounded smooth and there was little background noise to create an audible discontinuity. The entire file was processed with a high pass filter at 65 Hz to remove any possible traces of ENF fundamental. The mains power for the laptop was unplugged, and all surrounding electrical equipment was shut off, including fluorescent lights. The edited recording was played back through the laptop speakers and captured onto an Olympus WS-700M portable recorder, using the internal microphone and battery power (external power supply could not be obtained). The internal date and time were previously set to April 10, 2012 at 5:20pm, which was consistent with the original acoustic event. A new WMA file was created at HQ quality (40 Hz to 13 kHz frequency response; 44.1 kHz sampling rate claimed in user manual) with all settings flat, and presented as an unaltered original. Microphone sensitivity was medium and battery charge was full. The file was recorded to a removable MicroSD card and the original device was available for analysis.

The file format was viewed in Windows Media Player. The bit rate, sampling rate, and channel configuration were consistent between the evidence and an authentic exemplar.



View advanced details about the selected item.

Length: 00:49
Bit rate: 32 Kbps
Media type: Audio
Video size: -
Aspect Ratio: -
Audio codec: Windows Media Audio V8
32 kbps, 44 kHz, mono
Video codec: -
Location: D:\DAN R\Dan's Thesis Case Studies\Working Copies\Olympus Evidence\WS700013.WMA

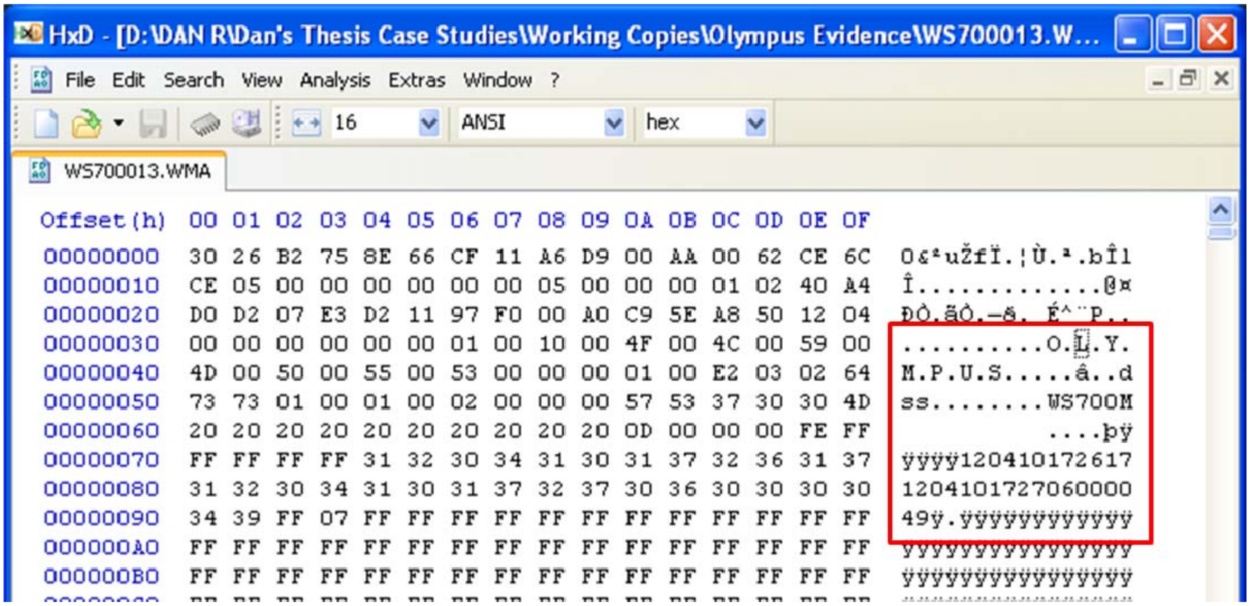
View advanced details about the selected item.

Length: 01:04
Bit rate: 32 Kbps
Media type: Audio
Video size: -
Aspect Ratio: -
Audio codec: Windows Media Audio V8
32 kbps, 44 kHz, mono
Video codec: -
Location: D:\DAN R\Dan's Thesis Case Studies\Working Copies\Olympus Silence Exemplar\WS700014.WMA

Figure A.20 File Format Analysis of Evidence (left) and Exemplar (right).

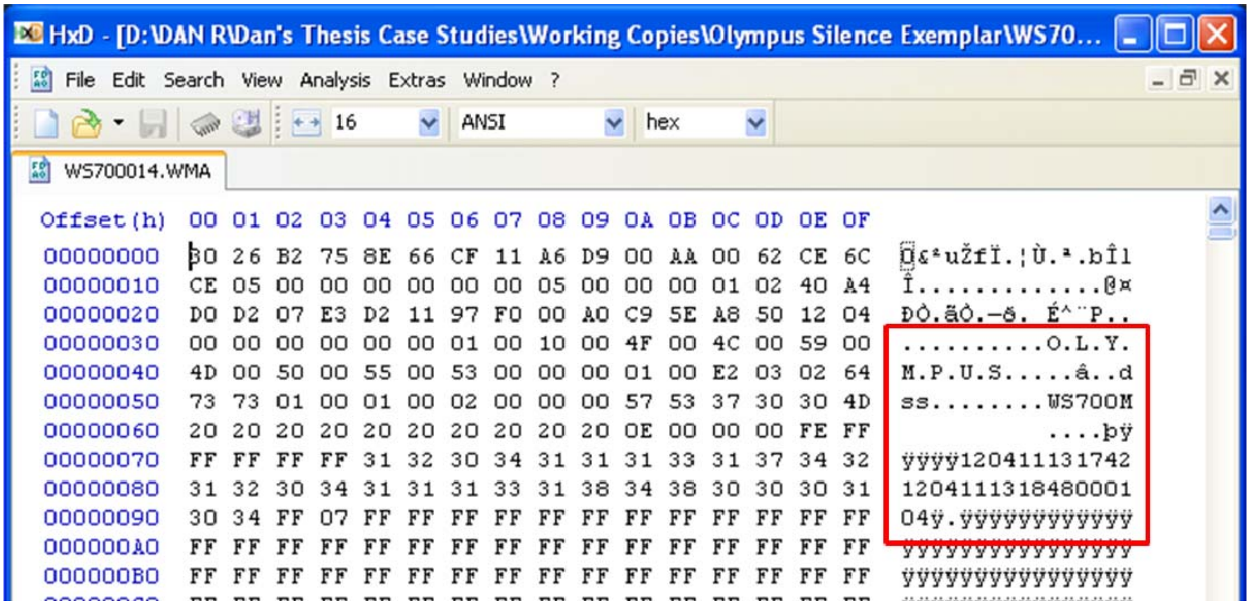
The MAC times also appeared consistent with test recordings made on the device: the Modified time reflected the moment the recording was first manifested on the device (as determined by the internal clock) which differed from the Created time.

The header showed the brand name “Olympus,” model number “WS700M” and detailed information about the recording including the date and time it started and ended, as well as the total length in minutes and seconds. By making test recordings on the same device, it was determined that the numbers 120410 represented the date, April 10, 2012 in the order year:month:day. This was followed by 172617, which translates to 5:26pm and 17 seconds in 24 hour time. 120410 was the date the recording ended (same as the start date) and 172706, or 5:27pm and 6 seconds, was the end time. Finally, the number 49 is the total length of the file in seconds. As seen below, the header was consistent with an authentic exemplar. It was confirmed, however, that the time stamp could be altered by manually changing the internal clock on the device before recordings were made.


HxD - [D:\DAN R\Dan's Thesis Case Studies\Working Copies\Olympus Evidence\WS700013.W...
File Edit Search View Analysis Extras Window ?
16
ANSI
hex
WS700013.WMA

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00000000  30 26 B2 75 8E 66 CF 11 A6 D9 00 AA 00 62 CE 6C  0&²uŽfÏ.¦Ù.ª.bÎl
00000010  CE 05 00 00 00 00 00 00 05 00 00 00 01 02 40 A4  Î.............@¤
00000020  D0 D2 07 E3 D2 11 97 F0 00 A0 C9 5E A8 50 12 04  ÐÒ.ãÒ.—ð. É^¨P..
00000030  00 00 00 00 00 00 01 00 10 00 4F 00 4C 00 59 00  ..........O.L.Y.
00000040  4D 00 50 00 55 00 53 00 00 00 01 00 E2 03 02 64  M.P.U.S.....â..d
00000050  73 73 01 00 01 00 02 00 00 00 57 53 37 30 30 4D  ss........WS700M
00000060  20 20 20 20 20 20 20 20 20 20 0D 00 00 00 FE FF            ....þÿ
00000070  FF FF FF FF 31 32 30 34 31 30 31 37 32 36 31 37  ÿÿÿÿ120410172617
00000080  31 32 30 34 31 30 31 37 32 37 30 36 30 30 30 30  1204101727060000
00000090  34 39 FF 07 FF FF FF FF FF FF FF FF FF FF FF FF  49ÿ.ÿÿÿÿÿÿÿÿÿÿÿÿ
000000A0  FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
000000B0  FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
```

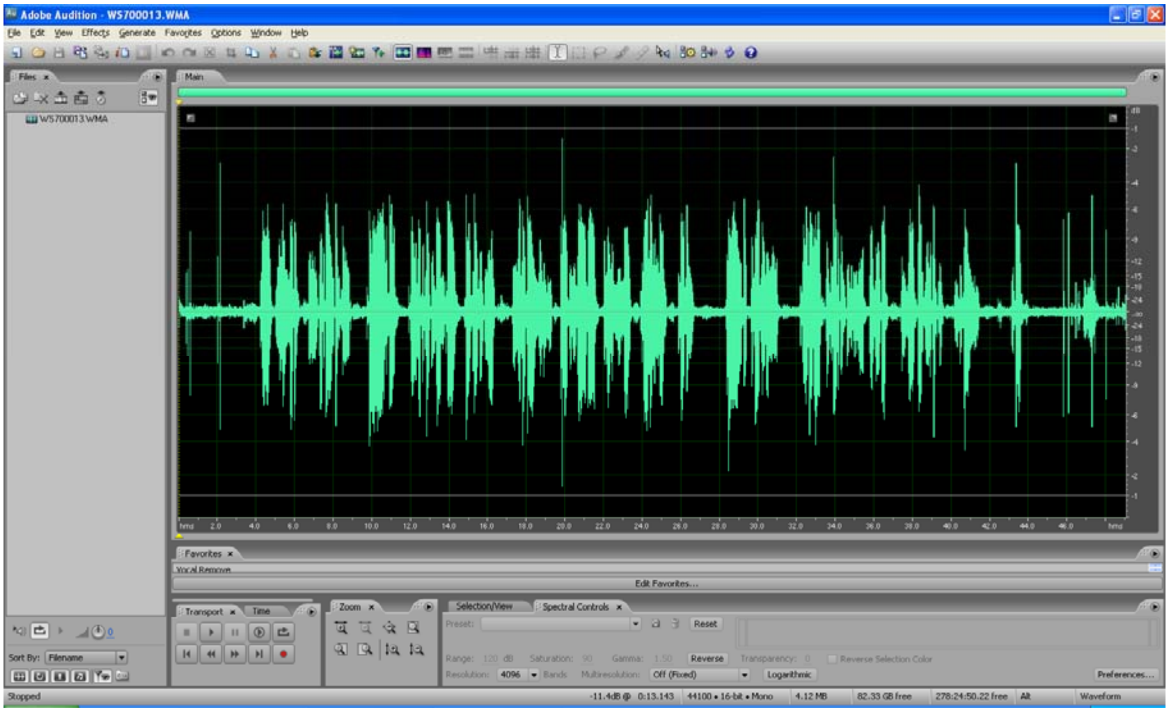

HxD - [D:\DAN R\Dan's Thesis Case Studies\Working Copies\Olympus Silence Exemplar\WS70...
File Edit Search View Analysis Extras Window ?
16
ANSI
hex
WS700014.WMA

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00000000  30 26 B2 75 8E 66 CF 11 A6 D9 00 AA 00 62 CE 6C  0&²uŽfÏ.¦Ù.ª.bÎl
00000010  CE 05 00 00 00 00 00 00 05 00 00 00 01 02 40 A4  Î.............@¤
00000020  D0 D2 07 E3 D2 11 97 F0 00 A0 C9 5E A8 50 12 04  ÐÒ.ãÒ.—ð. É^¨P..
00000030  00 00 00 00 00 00 01 00 10 00 4F 00 4C 00 59 00  ..........O.L.Y.
00000040  4D 00 50 00 55 00 53 00 00 00 01 00 E2 03 02 64  M.P.U.S.....â..d
00000050  73 73 01 00 01 00 02 00 00 00 57 53 37 30 30 4D  ss........WS700M
00000060  20 20 20 20 20 20 20 20 20 20 0E 00 00 00 FE FF            ....þÿ
00000070  FF FF FF FF 31 32 30 34 31 31 31 33 31 37 34 32  ÿÿÿÿ120411131742
00000080  31 32 30 34 31 31 31 33 31 38 34 38 30 30 30 31  1204111318480001
00000090  30 34 FF 07 FF FF FF FF FF FF FF FF FF FF FF FF  04ÿ.ÿÿÿÿÿÿÿÿÿÿÿÿ
000000A0  FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
000000B0  FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
```

Figure A.21 Header of Olympus WMA Evidence file (top) and Exemplar (bottom).

Critical listening was inconclusive. Although no discontinuities or areas or interest were identified, the recording sounded slightly tinny and off axis, but was highly intelligible. The possibility of the evidence being a first generation digital recording made with poor microphone technique, bad acoustics, or an article of clothing blocking the microphone could not be eliminated. While the method of tampering used introduces some amount of convolution to the signal, the examiner might mistake that for being part of the acoustic ambience in the original recording without a “dry” signal to compare it to (or the amount of convolution may be insignificant). No evidence of the extra stages of D/A and A/D conversion could be heard by the examiner. Techniques presently available for detecting analogue re-recording [31] could not be used, since the file was a lossy compressed format.


Adobe Audition - WS700013.WMA
File Edit View Effects Generate Favorites Options Window Help
Files
WS700013.WMA
Main
hms 2.0 4.0 6.0 8.0 10.0 12.0 14.0 16.0 18.0 20.0 22.0 24.0 26.0 28.0 30.0 32.0 34.0 36.0 38.0 40.0 42.0 44.0 46.0 hms
dB
Favorites
Vocal Remove
Edit Favorites...
Transport
Time
Zoom
Selection/View
Spectral Controls
Reset
Reverse
Reverse Selection Color
Logarithmic
Preferences...
Sort By:

Sort By: Filename

Preset:

Range: 120 dB Saturation: 90 Gamma: 1.50 Transparency: 0

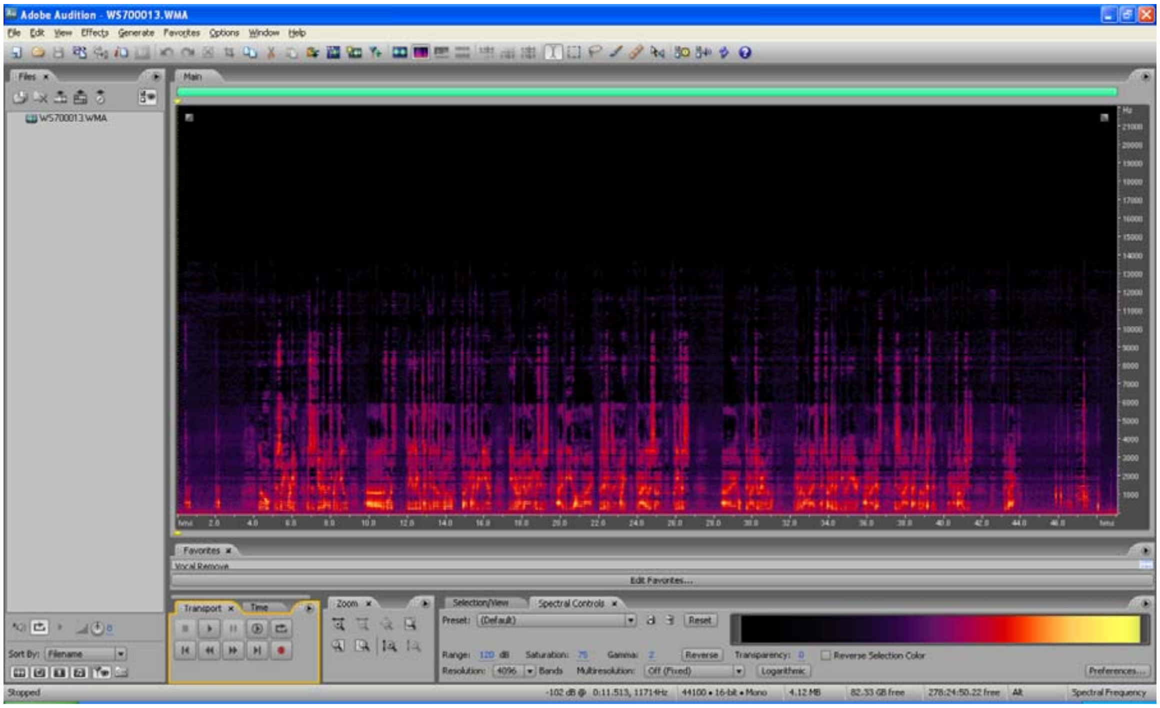
Resolution: 4096 Bands Multiresolution: Off (Fixed)

Stopped -11.4dB @ 0:13.143 44100 • 16-bit • Mono 4.12 MB 82.33 GB free 278:24:50.22 free Alt Waveform

Figure A.22 Waveform Analysis of Olympus WMA Evidence.

Waveform analysis showed no discontinuities or ROI and was consistent with an authentic recording. No clipping, normalization, or zero amplitude samples were found. Neither first nor second generation WMA test recordings produced zero amplitude samples at the beginning or end of the file. However, a second generation test file added 24 ms of noise to the length of the original.

Spectrographic analysis showed no inconsistencies, no ENF signal or steady tones.



Sort By: Filename

Preset: (Default)

Range: 120 dB Saturation: 75 Gamma: 2 Transparency: 0 □ Reverse Selection Color

Resolution: 4096 Bands Multiresolution: Off (Fixed)

Stopped -102 dB @ 0:11.513, 11714Hz 44100 • 16-bit • Mono 4.12 MB 82.33 GB free 278:24:50.22 free Alt Spectral Frequency

Figure A.23 Spectrogram of Olympus WMA Evidence File. FFT resolution is 8192 bands; Blackmann-Harris windowing.

LTAS analysis of the evidence with speech removed was inconsistent with authentic exemplars made in silent conditions with the same device and settings. The upper frequency limit was similar, but the overall shape of the exemplars was smoother.

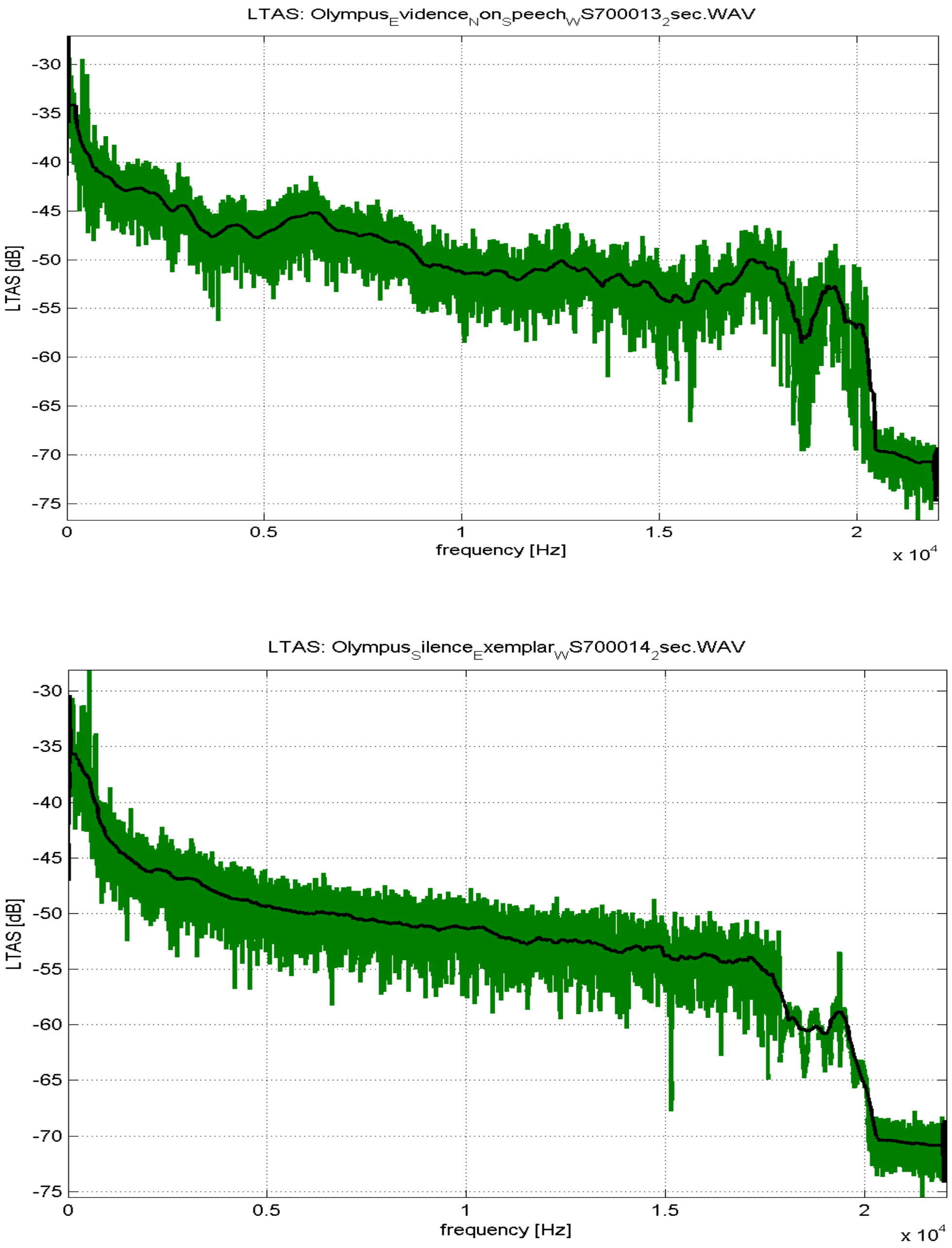


Figure A.24 LTAS of Olympus WMA Evidence with Speech Removed (top) and Exemplar (bottom) made in Silent Conditions.

DC analysis on the overall questioned file did not show indications of gain changes or copy/paste. DC comparisons were also made on the evidence file with speech removed to exemplar made with the internal microphones. The mean and standard deviation were consistent.

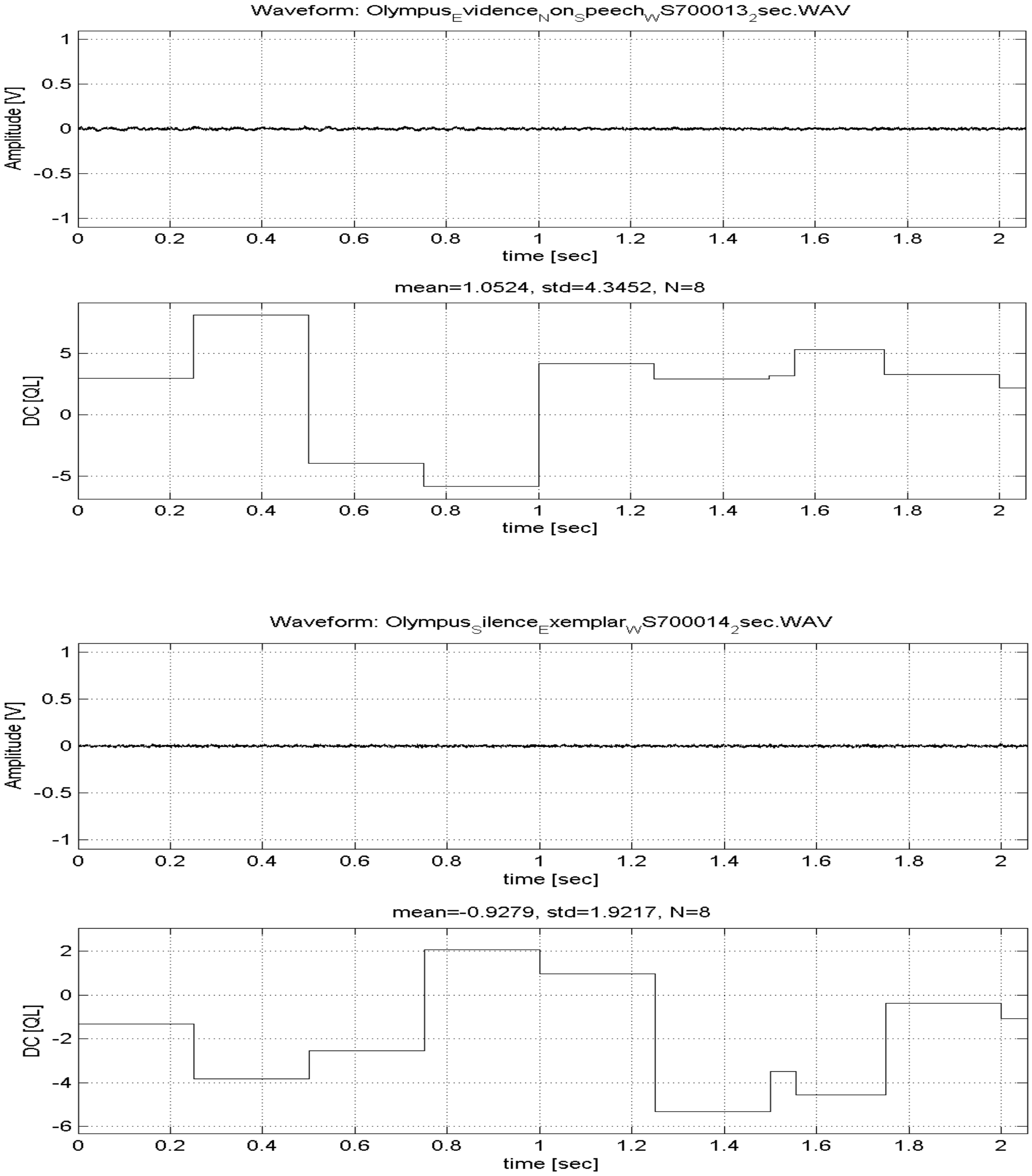


Figure A.25 DC Analysis of Evidence with Speech Removed (top) and Silent Exemplar (bottom).

The sorted spectrum of the questioned recording showed indications of a lossy compressed file, which was consistent with a WMA exemplar.

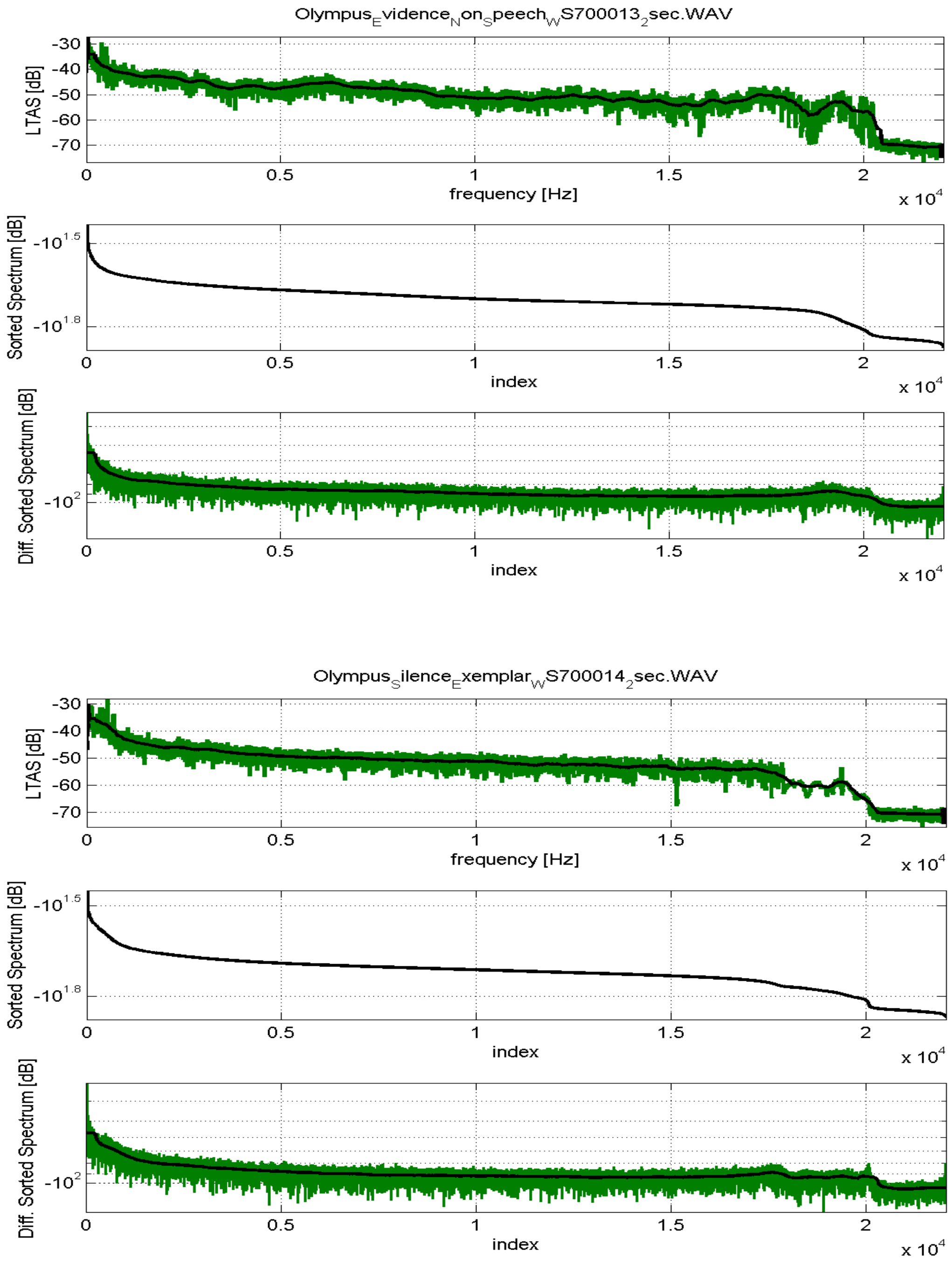


Figure A.26 Sorted Spectrum Evidence File (top) and Exemplar (bottom).

Compression Level analysis was consistent with a first generation WMA file, exhibiting the same sized gap, while a second generation exemplar had a wider gap.

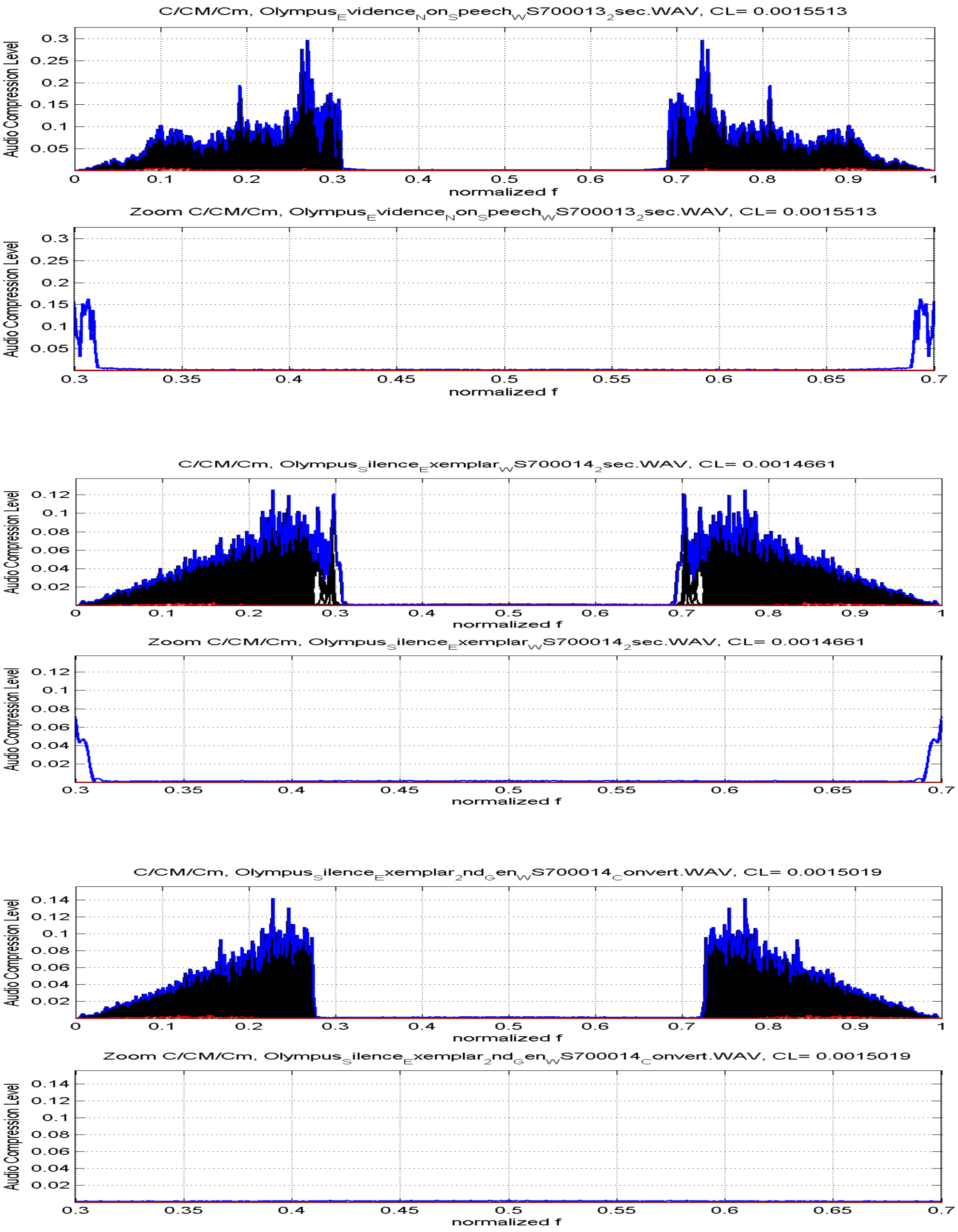


Figure A.27 Compression Level Analysis of WMA Evidence with Speech Removed (top), First Generation Exemplar (middle), and Second Generation Exemplar (bottom).

The examiner’s ultimate conclusion is that the questioned recording is *inconsistent* with an authentic recording using the methods claimed by the recording operator.

Table A.3 Authentication Framework for Case #3.

File Structure	File Format	*C*
	Header	*C*
	Hex Data	*C*
Global Analysis	DC Offset	*C*
	LTAS, Sorted Spec, Differentiated Sorted Spec.	*I*
	Compression Level Analysis	*C*
Local Analysis	Critical Listening	–
	Waveform Analysis	*C*
	Spectrum/Spectrogram Analysis	*C*
	DC Offset	*C*
Device Verification	Header (exemplars)	*C*
	Hex Data (exemplars)	*C*

REFERENCES

1. Smith, Fred Chris, and Bace, Rebecca Gurley. *A Guide to Forensic Testimony: The Art and Practice of Presenting Testimony as an Expert Technical Witness*. Boston: Addison-Wesley, 2003.

2. Benjamin, Walter. “The Work of Art in the Age of Mechanical Reproduction.” 1936. Accessed February 10, 2012. http://www.marxists.org/reference/subject/philosophy/works/ge/benjamin.htm

3. Zatyko, Ken, and Bay, John, “The Digital Cyber Exchange Principle.” *Forensic Magazine* Vol. 8. No. 6 (2011-2012): 13-15.

4. Keonig, Bruce, and Lacey, Douglas. “Forensic Authentication of Digital Audio Recordings.” *Journal of the Audio Engineering Society* Vol. 57, No. 9 (2009): 662-695.

5. Gorham, Geoffrey. *Philosophy of Science: A Beginner’s Guide*. England: Oneworld Publications, 2009.

6. Scientific Working Group on Digital Evidence (SWGDE). “Electric Network Frequency Discussion Paper, Version 1.2” May 20, 2010.

7. Advisory Panel of White House Tapes. “The EOB Tape of June 20, 1972.” 1974.

8. Feynman, Richard. “Probability and Uncertainty: The Quantum Mechanical View of Nature.” BBC Television. November 18, 1964. The Messenger Lectures: The Character of Physical Law. Accessed February 29, 2012. http://www.youtube.com/watch?v=w_vT1TaT3Mg.

9. Anderson, Scott D. “Digital Image Analysis: Analytical Framework for Authenticating Digital Images.” Master’s Thesis, University of Colorado Denver, 2011.

10. *Oxford English Dictionary*. Oxford: Oxford University Press. Accessed January 2, 2012. http://www.oed.com/view/Entry/13320.

11. Keenan, Thomas, and Weizman, Eyal. “Mengele’s Skull: The Advent of Forensic Aesthetics.” *Forensics* Issue 43 (2011).

12. Committee on Identifying the Needs of the Forensic Sciences Community, et.al. *Strengthening Forensic Science in the United States: A Path Forward.* Washington, D.C: The National Academies Press, 2009. Accessed February 15, 2012. http://www.nap.edu/catalog/12589.html.

13. Brixen, Eddy. “Techniques for the Authentication of Digital Audio Recordings.”

Presented at the 122nd convention of the Audio Engineering Society, Vienna, Austria May 5-8, 2007.

14. Rumsey, Francis. “Forensic Audio Analysis.” *Journal of the Audio Engineering Society* Vol. 56, No. 3 (2008): 211-217.

15. Gaudette, B.D. “Basic Principles of Forensic Science.” *Encyclopedia of Forensic Sciences*. Elsevier, (2000): 297-302, 716.

16. Audio Engineering Society. AES-27 1996 (r2007). “AES Recommended Practice for Forensic Purposes- Managing Recorded Audio Materials Intended for Examination.” Audio Engineering Society, Inc. 2007.

17. Grigoras, Catalin, and Smith, Jeff M. “Audio Enhancement and Authentication.” *Encyclopedia of Forensics Sciences*. Elsevier, accepted for publication 2012 (in press).

18. Bronstein, Daniel A. *Law for the Expert Witness: Third Edition*. Boca Raton: CRC Press, 2007.

19. Scientific Working Group on Digital Evidence (SWGDE). “SWGDE Best Practices for Forensic Audio, Version 1.0” January, 2008.

20. Hooke, Robert. *Introduction to Scientific Inference.* United States: Holden-Day, Inc. 1963.

21. Cosic, Jasmin, and Baca, Miroslav. “(Im)Proving Chain of Custody and Digital Evidence Recovery with Time Stamp.” Presented at MIPRO 2011 Conference, Opitija, Croatia, May 24-28, 2011.

22. Koenig, Bruce, and Lacey, Douglas. “An Inconclusive Digital Audio Authenticity Examination: A Unique Case.” *Journal of Forensic Sciences* Vol. 57, No. 1 (2012): 239-245.

23. Keonig, Bruce, and Lacey, Douglas, “Forensic Authenticity Analysis of the Header Data in Re-Encoded WMA Files from Small Olympus Audio Recorders.” Peer reviewed and accepted for publication in the *Journal of the Audio Engineering Society* with a presently unknown issue date.

24. Pohlmann, Ken C. *Principles of Digital Audio: Fifth Edition*. New York: McGraw-Hill, 2005.

25. Balasubramaniyan, Vijay A., et al. “PinDr0p: Using Single-Ended Audio Features to Determine Call Provenance.” Converging Infrastructure Security (CISEC) Laboratory Georgia Tech. Information Security Center (GTISC) Georgia Institute of Technology, Atlanta, GA.

26. Cooper, Alan J. “Detecting Butt-Spliced Edits in Forensic Digital Audio

Recordings." Presented at the 39th AES International Conference Audio Forensics- Practices and Challenges, Hillerød, Denmark, June 17-19, 2010.

27. Grigoras, Catalin. "Application of ENF Analysis Method in Forensic Authentication of Digital Audio and Video Recordings." Presented at the 123rd convention of the Audio Engineering Society, New York, New York, October 5-8, 2008.

28. Grigoras, Catalin. "Statistical Tools for Multimedia Forensics." Presented at the 39th AES International Conference Audio Forensics- Practices and Challenges, Hillerød, Denmark, June 17-19, 2010.

29. Jenkins, Christopher W. "An Investigative Approach to Configuring Forensic Electric Network Frequency Databases." Master's Thesis, University of Colorado Denver, 2011.

30. Brixen, Eddy. *Audio Metering Measurements, Standards and Practices: Second Edition*. United States: Elsevier, 2011.

31. Cooper, A.J. "Detection of Copies of Digital Audio Recordings Produced Using Analogue Interfacing." *International Journal of Speech, Language, and the Law* Vol. 15, No.1 (2008): 67-95.

32. Grigoras, Catalin, et al. "Advances in ENF Database Configuration for Forensic Authentication of Digital Media." Presented at 131st convention of the Audio Engineering Society, New York, New York, October 20-23, 2011.