International Journal of Cyber-Security and Digital Forensics (IJCSDF) 3(3): 170-182
The Society of Digital Information and Wireless Communications, 2014 (ISSN: 2305-0012)

Forensic Evidence of Copyright Infringement by Digital Audio Sampling Analysis - Identification – Marking

Stefan K. Braun

Faculty of Management. Comenius University
PhD Student, Comenius University
Bratislava, Slovak Republic
mail@stefanbraun.eu

ABSTRACT

In recent years, the number of attempts to use digital audio and video evidence in litigation in civil and criminal proceedings has increased. Technical progress makes editing and changing music, film and picture recordings much easier, faster and better. The methods of digital sampling differ from the conventional pirated copy in that using a sample involves extensive changes and editing of the original work. Different digital sampling methods make the technical analysis and the legal classification more difficult. Targeted analysis methods can clearly identify a case of sampling and belong to the main field of forensic analysis. If persuasive evidence of an unauthorized use of sampling cannot be produced, the proof is useless in the legal process. Labelling technologies that are applied correctly make an important contribution to the effective detection of unauthorised sound sampling. There are hardly any holistic approaches that integrate the problem of sound sampling into the fields of analysis, identification, and labelling. In combination with specific technical protective mechanisms against sampling, an unauthorised use of samples protected by copyrights can be prevented or reduced. Using and sampling somebody else's piece of music or video can be a copyright infringement. The copyright and the neighbouring rights of performing artists and the neighbouring rights of phonogram producers are affected by the consequences of illegal sampling. Part 1 of the article introduces the problems of digital audio sampling, Part 2 describes the typical manifestations of sampling, Part 3 illustrates various analytical procedures for the detection of audio sampling and Part 4 shows the identification by labelling strategies.

KEYWORDS

audio · authentication · bootlegging · digital techniques · single sound sampling · ENF · Electric Network Frequency · forensics · forensic audiology · real-time frequency analysis · cryptography · neighbouring rights · melody · mash-up · mix production · multi-sampling · phase inversion · remix · sample medley · sound sampling · sound separation · spectrogram · spectrometer measurement · sound collage · sound sequence sampling · copyright · watermarking

1 INTRODUCTION

1.1 The Problems and Classification of Digital Sound-Sampling

The word "sample" in this context stems from the piece of equipment known as a "sampler". The sampler is supplied with sound information by integrating sound or microphone recordings. From the fed-in oscillation curves, samples are taken and stored. With the use of modern software, removed samples can be, for example, transposed in pitch and tempo, mutilated, transformed, tampered or mixed as desired [1, 2].

From the sample source voices, instruments, rhythms and parts of melody can be removed ("sampled out") and incorporated into a new production. The purpose of sampling is the simple and inexpensive way of adopting desired sounds, instruments, or voices without having to invest in studio production costs, time and effort. Furthermore, the sound characteristics of performers can be imitated and used as inspiration without their knowledge or consent.

Users of samplers not only utilize notes but also sound from a specific production. The arrangement of individual sounds and timbres can be created, on the one hand, in the studio and, on the other, directly on the digital recording computer [1]. "Sound", "timbre" and "tone" are used more or less synonymously in literature. The limiting factor is that from a physical point of view, timbre is only one of the many components of sound [3].

Sounds and melodies can be generally adopted from both existing music productions and recordings. In contrast to this, there are sound databases that can be downloaded from the internet and also physical data carriers such as sound libraries.

In addition to shorter sound excerpts of a few bars or seconds, smaller melody parts, the so-called "licks" and smaller sequences are sampled. A specific sampled music sample therefore includes also the generated sound [4]. If there are, in addition to a certain sound, enough of these samples available to the user, he can put these together like a "mosaic" to create a "new" work. A very common form of sampling is taking foreign compositions from actual recordings into new music and film productions. Often pitches and characteristics are changed to differing degrees when adopting single tones or tone sequences in the sampling process.

Processing. The processing of a musical work is always associated with a transformation. When composing, the melodic, harmonic and rhythmic form is changed. When this is text, it is reworked, modified, supplemented, replaced completely or translated into another language, for example. The result of such a major rearrangement is a newly created work. The cover version shows the necessary individuality in the form of intellectual and approval-requiring creation [5]. The prerequisite is that the transformation in turn has the appropriate quality of "work". It should be determined which musical design elements cause the creative peculiarity of the work. To be considered in this context in particular are the tonal system, the duration of the tone, timbre, volume, rhythm and melody.

Processing eligible for protection. Processed work which is eligible for protection requires a recognizable creative performance of the editor, so that resulting from the compositional change or expansion of the musical substance of the original, a new, independent work is created. In contrast to such works which are eligible for protection are those which use an original work and take the musical substance of the original essentially unchanged and transfer the musical text of the original faithfully (e.g. editorial services) [6]. Works that have been created using other works or foreign melodies must be marked with the appropriate copyright information. For free works no permission for processing has to be sought from the originator. Protected works require this permission. Processing is the key feature when considering whether the original is eligible for protection [5]. It is crucial that the new work distinguishes itself from the old one and not only repeats an already existing one; the aesthetic overall impression of the new piece must not be present in the original work [7].

Melody. The melody is, in occidental music, the most important parameter and main information carrier. Together with the harmony it is the most important forming structure in music. The term melody includes three elements: Harmony (harmonizing of tones), rhythmos [sic] (temporal structure) and logos (text). Melodies are differentiated in their function and their classification as a vocal melody (range, phrase length) or instrumental melody [8]. The melody forms a self-contained tone system (characteristic). It retains its own character even when accompaniment (rhythm) is eliminated or the sounds replaced (transposed). In music for easy listening and pop music, the vocal parts of the melody are considered to be the characteristic that can be assigned to the relevant song.

1.2 Services not Eligible for Protection

Typical techniques and thus ineligible for protection include mere conversions of sentences or sentence parts of a multi-part musical work, slight changes of melody, harmony and rhythm, or individual noise elements if the basic character of the original work remains the same [9]. Certain, recurring basic repeats or patterns, such as chord sequences, classic song structures or common elements of music are not eligible for protection [10]. Insignificant tonal variations, slight shortening or extensions taking into account the compositional or textual original work are permitted in this context [10]. Exceptions are to be seen under certain circumstances with regards to fingering in music course books when this characteristic forms the tone. The transposition of the pitch of the original is also one of the criteria ineligible for protection and does not change the melody.

Criteria for Activities not Eligible for Protection.

- Lack of originality.
- Insignificant, minor changes.
- Use of an original work, borrowing of partial works.
- Transposition to a different key or pitch for technical artistic reasons.
- Instrumentation and timbre of individual instruments, merely replacing an instrument.
- Adaptation of the melody to the vocal abilities of the singer.
- Making changes to the rhythm, replacement with another standard rhythm.
- Note-for-note transcription of existing voices to another instrument.
- Supplementing of performance indication, elaboration, fingering, applying punctuation.
- Addition, change of phrasing.
- Tempo and volume adjustments.
- Doubling of voices.
- Addition of accompanying voices in parallel motion (e.g. in the third or sixth).
- Reduction of existing parts in the score of a piano movement.
- Editorial services (publication of a pre-existing musical work).
- Digitization or compression into an MP3 file, for example.

2 TYPICAL MANIFESTATIONS OF SAMPLING

Cooper [11] divides audio editing into three levels: 1) Editing / tampering on a basic level, directly in the original material, during or after the recording; 2) Editing / tampering on an intermediate level, containing several fields copied from one or more original sources for a new recording; 3) Editing / tampering at a high level by means of appropriate editing and sound processing software. The edited version will then function as a “new original”.

According to their type of use, the sampling techniques can be divided into single-tone sampling and melody sampling. Single-tone sampling distinguishes again between the actual sampling of a single-tone and a variant called “Multi-Sampling”, one of the economically most important and technically difficult to detect sampling forms. It is referred to colloquially as “sound sampling”. The parties involved in each sampling are always the originator or author, the performing artist and, in the case of indirect sampling, the record producer. If a digital sample is used, there is inevitably always a reproduction of works or parts of works.

2.1 Origin of the Sound Material

Sampling of the Artists' Own Sound Material. Sound material can be recorded by the artists themselves or recorded and then sampled. This is usually done where there are certain fragments repeated in a musical work. Sampling is also carried out when certain figures of a piece have a repetitive character and do not differ in dynamics, articulation and rhythm. With this approach, difficult figures and phrases have to be recorded only once [12].

Sampling of Foreign Sound Material. Much more sampling material comes from external sources [12] such as sound recordings or individual tracks from multi-track tapes. Furthermore, so-called “factory sounds” and sound archives exist, for example, on CD or as downloads from internet archives.

Natural Sounds. These are divided into signals produced by oneself and others as well as natural sounds, meaning sounds not shaped by humans including animal sounds, machinery and everyday sounds [13] and meteorological noises [12].

2.2 Single-Tone Sampling

Direct Single-Tone Sampling. Under direct single tone sampling, sampling of individual instrumental sounds is understood. Here, a certain characteristic sound, for example, an instrument, a voice or a sound is taken in isolation, digitized, fragmented, and then imported into the sampling computer [12]. Using the keys of keyboards, the sound can be allotted to a button and then played. If there are sounds in different pitches, volumes and articulations, music can be played and modelled with specific musical characteristics. This process provides unrestricted access to the original sound of a music production.

Indirect Single-Tone Sampling. Single-tone indirect sampling is the term used to refer to the acquisition of sampled sounds from existing recordings, mostly audio recordings. A single tone can thus be isolated and the obtained sound then processed. The acquisition of single tones from a ready-mixed multi-track production by frequency superpositions of the single-tones and instrumental tracks later mixed together is not quite so simple. A single tone from single tracks of a recording, however, is very easy to remove and include and of high quality [1].

Multi-Sampling. The term multi-sampling is used when several individual notes with different pitch intervals and volumes are distributed on a sampler keyboard. The distribution usually takes place according to the original pitch. Often tones of mixed productions are extracted which have superimposed frequencies of other instruments. If only one sound as in the single-tone sampling is extracted, this would have to be transposed to a different pitch, which would lead to frequency distortions in any existing secondary frequencies. Therefore, different sounds according to their pitch ranges are extracted from different points of a piece in order to avoid this negative effect. An additional optimization is achieved by the blending (positional crossfading) of the samples with each other [1].

2.3 Melody Sampling

Contrary to the sound use of the single tone sampling, tone sequences sampling is about the (partial) adoption of melodies, harmonies and rhythms and the subsequent collage-like composition of new musical works. In general, a sequence of sampled parts from well-known music productions is used to maintain the recognition effect [12]. A variety of procedures can be distinguished.

Mixed Productions (Sample Medley). In mixed production consecutive characteristic music parts of a few seconds or bars are sampled and successively linked together in a newly created mixed production. Here, the new mixed production either contains parts of samples [1] or, in extreme cases, consists entirely of such. By using adjustment of the tempo the individual samples must be adapted, where necessary, before the mixing takes place. The purpose of this approach is the recognition effect of the sampled work parts. The more clearly the recognition of parts of the originator's work, the more successful the goal of the mixed production was implemented. Such mixed productions are created in the pop and dance genres by disc jockeys. Such productions were used before digital sampling technology existed, carried out by hand and the much more complicated and time-consuming tape cutting.

Sound Collages. Unlike mixed productions, sound collages disguise their origin [14]. Instead of stringing together sound samples, into sound collages these are layered over each other ("batch processing"). It is not unusual for several layers of samples to be superimposed. For example, a melody sequence can be taken as a sample from work 1, a rhythm from work 2 and a guitar sequence from work 3. In general, the individual samples must then be adjusted with regards to volume, tempo, pitch and timbre, so that they fit together in a new production, often cut as a "loop". As with mixed productions, the sound collages may consist either in part or entirely of samples.

Cover Versions. The sampling technique with cover versions and remixes is understood as "hit-recycling". Either the whole work or parts thereof, for example, the refrain, are taken from the original and backed with new rhythms and sounds. The purpose is the audible sound adaptation to new listening habits. Cover versions (interpretations of an earlier original) can be made without using the sampling technique. The sampling technique is still used consciously and for economic reasons, however, to maintain the successful part of the original. As with the mixed productions, sampled parts should be recognized [1].
If the artist leaves the limited scope for interpretation set for cover versions and moves towards a processing with independent creative input into the piece, this change is subject to approval.

Remixes. The remix follows the same rules as processing. Successful hits are frequently re-released as a remix. Individual tracks of a multi-track tape are often completely "broken down into pieces" and recomposed and remixed along with new recordings. There are also mixed sound effects, new recordings of instruments and a far-reaching change in the sound of the material. The remix, however, can take place with the extraction of a sample [12].

Mash-Up. Mash-ups (also known as bootlegging, bastard pop or collage) have been enjoying increasing popularity for years. At the beginning of the 1990s, it was usually only 2 different pop songs whose vocal and instrument tracks were mixed with each other to form a remix [12]; today there are multi-mash-ups with several dozen mixed and sampled songs, artists, video sequences and effects. It is a challenge to mix this combination of different styles to new danceable tracks.
The mash-up is a mix of sound collage and mixed productions. Usually known sequences of two or more (multi-mash-up) existing works are mixed to create a "new" work. The samples used are layered over each other (sound collage), as well as in series (mixed production). The incorporation of large parts of the original in the mash-up is the rule. In sampling, however, it is rather the exception [12].

3 ANALYSIS METHODS FOR DETERMINING THE USE OF A SAMPLE

Evidence of sampled parts in a musical work can be achieved by means of different methods of analysis.

3.1 Musical Aspects

Under certain circumstances, a simple listening test is sufficient. As a rule, a direct comparison of the musical notation is carried out. Since most samples were changed in speed and pitch, it can be helpful, to adapt these in terms of pitch and tempo to the original before starting the analysis.
Pitch changes and temporal extension have qualitative limits if a realistic overall impression should remain. Deviations of about 15-20% produce audible noise and alienate the original. This can be desirable for creative reasons. Often sampled parts are superimposed with other instrument and vocal tracks. A simple separation is then no longer possible.

3.2 Physicals Aspects

Analysing and measurement methods provide evidence of the use of sampling.

Electric Network Frequency Analysis (ENF). With regard to the validation of digital audio and video recording a common method recognized by expert forensics is the Electric Network Frequency Analysis (ENF). Each mains power supply leaves a characteristic frequency, a so-called "mains hum". This may not be audible but the oscillations can be detected in an audio file [15]. If digital recording devices are used such as cameras or audio recorders to record voice, music or film recordings these, in addition to the actual content, store the network frequencies of 50 or 60 Hz. This happens with battery powered devices in the same way [16]. The frequency of the carrier never has exactly the same value. The random fluctuations in the power supply are the result of the differences between produced and consumed current. The actual ENF signal can be extracted by using band-pass filters that filter out, for example, in a 50 Hz power supply the range 49-51 Hz. Interruptions or irregularities in the phase response can be an indication of tampering. The network frequency behaves in effect as a temporary digital watermark. The effectiveness of this method, however, is dependent on such a network signal existing at all [15]. In most situations, a visual comparison of the spectrogram with the frequencies which are stored in an ENF database is sufficient. More detailed studies require measurement and analysis in certain short time slots, which are compared to each other. With this method it is even possible to determine the exact location and the exact time of production of a recording. Corresponding reference samples prepared by continuous recording of the network frequencies in power networks (such as the German or European electricity grids) are a prerequisite [16].

Microphones also leave a particular frequency spectrum in the audio material. Should several different spectra show up in a recording, this can also be an indication of tampering [15]. The evaluation of the digital audio recording by detecting the exact measurements, the comparison, and a mapping of the individual frequencies in the reference database is, therefore, of great importance.
Suitable methods are the spectrogram representation, “re-sampling”, real-time-frequency analysis (spectrometer measurement) and the phase inversion.
Besides ENF analysis the spectrogram representation, “re-sampling”, real-time frequency analysis (spectrometer measurement) as well as phase inversion are appropriate forensic methods.

Spectrogram Representation. In a spectrogram display the spectral density of a signal is displayed over time. Figure 1 shows in A the recording A (an original music recording from 1990) and B the recording B (unauthorized editing by sampling removed from the original A and mingling with new instrumental tracks in 2007). The recordings were first equalized in tempo and pitch before the examination by means of a spectrogram representation and then directly compared (see “A” (left channel) as recording A and “B” (right channel) recording as B). A time section of 6 seconds is depicted. This corresponds to about 4 synchronous parallel bars from the two recordings. The Y-axis shows the frequency spectrum of 0 Hz (Hertz) to about 20 kHz and depicts approximately the hearing range of the human ear[1]. In the lower area, the lower frequencies are shown and the higher frequencies at the top. The horizontal X-axis is the time axis. Frequency is the number of periods that are run through in a second. The unit of frequency is the Hertz (Hz). An oscillation is composed of a positive and a negative half-wave, i.e. the to-and-fro swing of the electrons is called an oscillation, wave or period [17].

With this representation, the audio material can be visualized. The representation in the frequency spectrum is used to gain direct access both to specific frequency ranges as well as certain time ranges in contrast to standard waveform processing (see Figure 1, Area C) which is always performed for the entire frequency domain. These frequency ranges can be shown in colour by means of analysis software. High and low frequencies are represented by different colours. The intensity and the level of the frequencies are displayed in a colour spectrum that extends from blue and white (the highest intensity) to purple and black (the lowest intensity). In simple terms, a bell sound in a piece of music can, for example, be reduced, replaced or removed by using the “Copy & Paste” software function to copy a part without a bell and insert it over the desired place. In spectral processing, there are diverse modes that can be used. For example, it is possible to reduce levels by means of band, low and high pass filters (“damping”) - the peak level is blurred by mixing the frequencies and thus they “disappear” or are covered up. Furthermore, it is possible to transform the dynamics without changing the actual frequency content (“dispersion”).

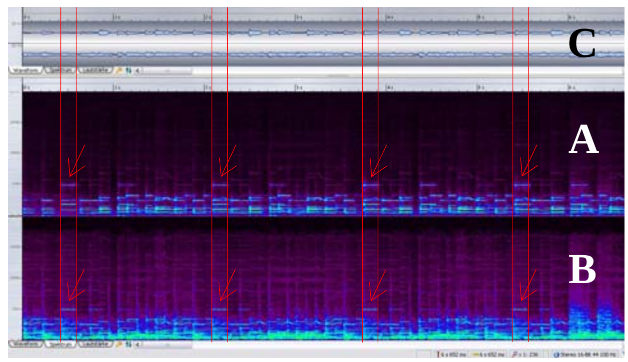

Figure 1. Comparison of the recordings A and B in the spectrogram representation (Source: Stefan K. Braun).

In Figure 1 a bell can be clearly seen in the sampled recordings A and B (red arrow). It lies in the frequency range of approximately 5400 Hz. In the AB-comparison it is very clear to see that all the different frequencies correspond, the patterns of both images are identical. The frequency spectrum of recording B is much richer. This is due to the mixed instrumental tracks added to the recording. The temporal distribution of the frequency phase and the significant characteristic features such as the existing bell in the recording have not been changed by the sampling. The area C represents the waveform processing. Visual procedures such as the spectrogram representation are important methods for aiding detection of manipulation.

“Re-Sampling”. Under certain circumstances, a sampling procedure can be carried out via a so-called “re-sampling”. Here, in simple terms, the numerical values of the digital samples are compared with those of the original. This pre-

[1] The hearing range (auditory sensation area) of human ear is from about 16-21 Hz to 16-19 KHz.

supposes, however, that there are identical comparative pieces. Usually the samplings used do not exist in isolation, but in the final product mixed together inseparably with other audio and instrument tracks, distorted with effects and changes in tempo and pitch. A direct comparison is no longer possible.

Spectrometer Measuring. In sampling, a digital copying process cannot always be compared purely by listening. With spectrometer measurement a coherent frequency diagram can be displayed and a very accurate and detailed real-time frequency analysis performed. In this case, the frequency spectrum is represented as a linear graph. Peak levels are depicted as short horizontal lines showing the last reached maximum values (see Figure 2). Spectrometer measurements are also used in forensic analyses, e.g. vocal comparisons in the field of criminology.

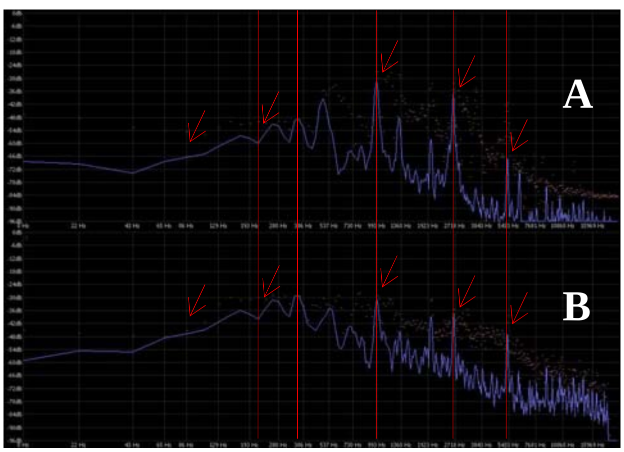

Figure 2. A frequency spectrum in a real-time frequency analysis with a linear graph at a randomly selected point in time of the investigated sample (Source: Stefan K. Braun).

In Figures 2 and 3, the amplitude of the wave of + / -0 dB (decibels) is represented by -96 dB on the vertical Y-axis, the horizontal X axis indicates the frequency band of 0 Hz to about 16 kHz. As described above, the recordings A and B are compared directly. In the authenticity analysis, the determination of the originality and continuity of the recordings and the detection of changes are of particular importance [18]. For real-time frequency analysis a random location of the samples to be examined was selected and fixed as a linear graph. The graphs show the result of a very similar, almost matching curve. Their frequency forms correspond in the typical manifestations in the characteristic points (e.g. shallow rise, steep climb, strong peaks between 500 Hz and 2700 Hz, falling from 7500 Hz). Within the investigated samples of 6 seconds duration all investigated linear graphs of the frequency spectrum show a relatively similar curve in terms of characteristics and patterns. Figures 4 and 5 show two or more overlapping linear graphs. The relatively similar curves from randomly selected positions in time on the same samples show clear similarities between the original A and sample processed B.

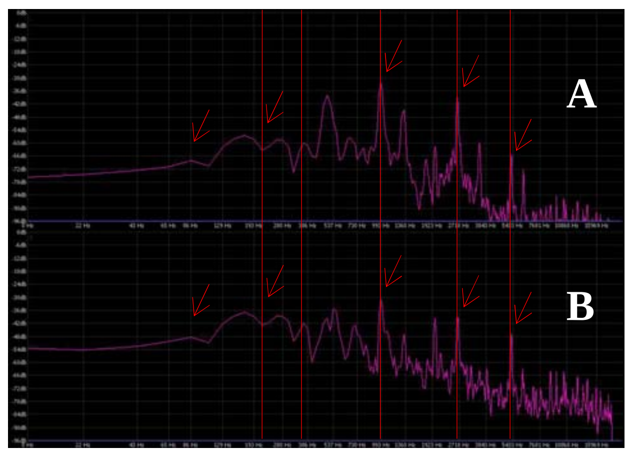

Figure 3. A frequency spectrum in a real-time frequency analysis with a linear graph at a different point in time of the same sample (Source: Stefan K. Braun).

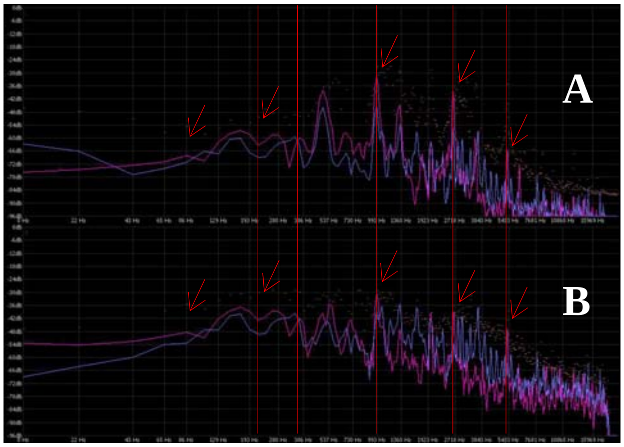

Figure 4. A frequency spectrum in a real-time frequency analysis with two overlapping linear graphs at a randomly selected point in time of the same sample (Source: Stefan K. Braun).

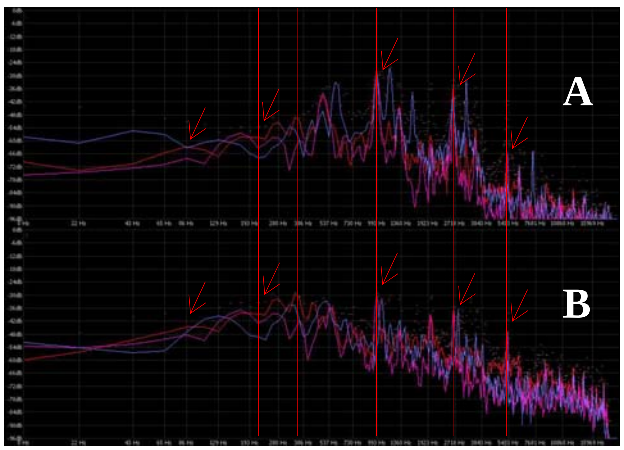

Figure 5. A frequency spectrum in a real-time frequency analysis with three overlapping linear graphs at a randomly selected point in time of the same sample (Source: Stefan K. Braun).

The problem may be verification when a sampling was not created by copying, but by an extensive technical sound remake. Here there is a difference in the technical and legal view. While in terms of law, a remake "sample" can still be considered as such, it is technically a different object. If a sample is taken from an original, it can be determined relatively easily due to whether the frequency plot of the linear graphs is the same or different in the analysed sample. For example, physical characteristics of the same or different audio tracks of vocals can be represented by this method. Adopted or remade instrument passages can be revealed and checked for sameness with this method. Even non-audible differences of different blowing techniques for brass instruments or different striking techniques with keyboard instruments can be seen in the graph representation [13]. It is not possible to achieve congruent sound and frequency structures by imitating ways of playing and singing. If they are identical, everything points to a sampled adoption of the original. The limits of an identical representation of the linear graphs are reached when the samples in one object which are being compared are changed dramatically with respect to sound and are superimposed with other vocal and instrument tracks.

Phase Inversion. In recording studio technology, phase reversal (phase inversion) is often used to correct wrongly polarized audio signals in the phase. In order to achieve certain effects, phases with correct polarity can also be reversed deliberately. Using this, undesired and reverse-poled phases can be added/mixed with the phases of the original signal, so that they cancel each other out, in whole or in part. For example, in a piece of music with vocals, the vocals are "filtered out" by phase inversion in order to obtain an instrumental or karaoke version.

In the forensic evidence of phase reversal, a destructive interference is sought; the matching points (oscillations) of the samples cancel each other out. An oscillation is composed of a positive and a negative half-wave, and thus corresponds to a full circle of 360 degrees [17]. If two sine-phases in the fundamental frequency are shifted 180 degrees of the phase, they are opposed (mirrored or inverted) and so cancel each other out completely.

If two or more waves are added, their amplitudes are reinforced; this is referred to as constructive interference. If the waves cancel each other out, destructive (complete) interference is the term used.

Theoretically, both recordings must be completely identical in this experimental arrangement, i.e. tempo, pitch, volume and the course of the wave form match completely. If in a recording, a phase inversion is performed and this phase is mixed together with the other identical recording without phase inversion, it results in a complete cancellation of the part concerned.

In a study of the phase reversal, destructive interference was sought in order to mutually cancel the corresponding parts of the samples. Under practical conditions, the physical alignment of both recordings on exactly the same pitch is very difficult. The more accurate this process is, the greater the cancellation in the end. In the next step, recording A is inverted in phase and levelled with the pitch of recording B. Then both phases are superimposed. The result is shown in Figure 6. While the phases do not cancel themselves out completely, they clearly correlate with each other. This correlation is particularly evident in the direct comparison with the unprocessed recording B.
Comparison objects are seldom completely identical in practice. A phase cancellation is therefore mostly only partially possible. The affected sample part has partial cancellations. What can be heard after a partial phase inversion is a clear "flanging effect". This effect is caused by artificial zeroes which are the result of the cancellation of the audio signal in the frequency spectrum. At the same time, in the previous phase reversal, a phase shift will take place, which causes a shift duration ("delay"). Now both the (partially) erased places and also the shifting of the phases to each other are audible
"Flanging" altered audio signals produce a kind of "floating" effect. Often the effect is described like a jet ("jet effect") which moves through the music [19]. In simplified terms, the "flanging" effect is similar to that of a tape and tape recorder. If a spool is "braked" by hand, then it accelerates again when released. This creates the effect of "flanging".
Complete cancellation cannot be achieved even with perfect alignment of pitch as the examined images A and B are different in their characteristics. This is mainly due to, as mentioned above, the superposed instrumental and rhythm tracks in recording B.

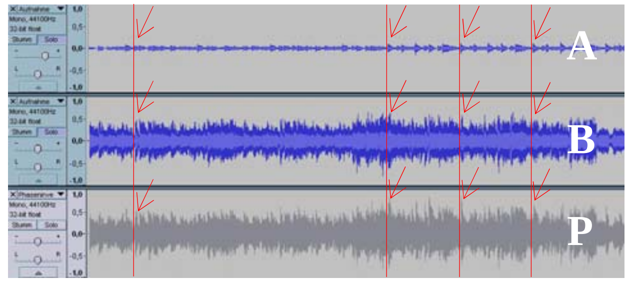

Figure 6. Phase inversion. A (input A, phase inverted), B (B recording, normal phase), P (mixed phases from the recordings A and B and audible "flanging"). Recording B is "shorter" at the parts with the included sampling at the end of the sample, i.e. it stops earlier than at the corresponding part in recording A. Due to this, the flanging effect stops at the end of this part. At this point the previous partial cancellation is particularly apparent (Source: Stefan K. Braun).

4 IDENTIFICATION BY LABELLING STRATEGIES

There is almost no effective protection that prevents unauthorized copying. In the last 20 years or so, the affected industries have developed and used the most diverse digital copy protection and labelling systems. Known systems include Digital Rights Management (DRM), the Content Scrambling System (CSS), different types of holograms, signatures such as RIFD, Serial Copy Management System (SCMS) or, for example, digital watermarks. For novice users there might be restrictions in use as not all the playing devices are able to deal with the copy protection mechanisms such as the DRM restrictions. The technically versed professional is, regardless of the legal regulations, capable of getting round these precautions more or less easily. Although overall markings such as holograms, bar codes [20] or ISRC codes (International Standard Recording Code) [21] identify the product (recorded music, digital file) in terms of its originality, they do not protect or prevent a possible further illegal use. Of importance is a modular approach between the requirements of sound sampling, in conjunction with a proper identification method: Protection and recognition of very small clippings that are superimposed with other signals in foreign productions reappear. All procedures which can be used have a main problem in common: the more they cost, the less value these are in a practical use. Fundamentally it must be distinguished between “data hiding” and “watermarking”. While data hiding conceals information in the medium or in a channel, the watermarking binds the information into the medium. Data hiding is used interchangeably with "information hiding", although the latter is more likely to be used for the cryptic method [22].

The following procedures seem appropriate for marking, identification and authentication of sound samples for further use:

4.1 Cryptographic Processes

Cryptographic processes can be divided into asymmetric, symmetric and hybrid[2], as well as strong and weak methods.
According to Lynch / Lundquist a cryptic-secure data exchange is confronted with the following system requirements: Identification, authentication, verification, non-repudiation and privacy. If all five demands are met, this is referred to as a secure data exchange [23].

Asymmetric, cryptic processes are characterized by the fact that digital signatures have a private and a public code. With the use of the private codes it is ensured that only the owner of the product rights can assign an individual signature [20]. A test of the encryption is provided by the public code. Signatures, e.g. in the form of identification numbers (“Identification Keys”) in connection with a verification database allow the tracking of marked objects (“Tracking & Tracing”).

[2] https://www.datenschutz.rlp.de/downloads/oh/ak_oh_kryptographie_version1.pdf

4.2 “Watermarking”

The watermarking technology is a promising technology for the protection and prosecution of copyright infringement. The basic technique and main focus of research in digital watermarking consists of an integral, invisible [22] “interweaving” of identification (copyright information, names, logos, etc.) with the main channel without interfering with or impairing this. Audio signals (music and speech), images, movies, software, e-books and texts can be provided with individual markings in this way [24].
There are two important main groups of watermarking use: 1 Piracy resistant use, which prevents an attack on the watermark. Applications are copy-protection measures, “fingerprint” techniques and other preventive measures (e.g. hash functions). 2 A use that is weak in terms of being piracy resistant, the watermark is dissolved or minimally changed in the case of a piracy attack. When the watermark has been changed or is absent, copies of the originals are no longer recognized as originals [25].
There are important requirements for the labelling: 1 The easy readability of the watermark in retrospect, 2 Resistance to destruction, 3 The receipt of the signal in the case of the use of very small excerpts of the original file [25] and 4 The additional information must not be perceptible to the human ear [22].

There are several, often conflicting, properties that are the focus of watermarking: The amount of hidden or inserted information, the robustness and security of the data, the invisibility and the reading of the introduced data [22].
Labelling and identification systems which are based on an authentication and so distinguish the copy from the original can be used independently or with a database [20]. A check on the authenticity of the watermark and the control of the authentication is done, for example, using database systems. For audio files, for example, a watermark can be set as an “inaudible” frequency over the actual audio frequency

band. To read the information it needs the same algorithm, a "Watermark Key", which was necessary for the earlier interweaving of the information. The recognition of copyright infringement takes place via a verification comparison on the database server. Disadvantages of such systems are a not quite closed security chain, as markers are not created directly at the premises of the copyright owner, but in the sales shop. If only digital files using the watermark process are detected, a direct use of recorded music media and trade on exchange platforms cannot be prevented. Piracy resistance has limits with the use of watermarking technology. A frequent copying and transforming creates a "fuzzy", unreadable watermark. A significant advantage is in the aforementioned limitations of the preservation of the watermark even with format changes, compression, filtering, re-sampling, re-quantization, as well as recognizing the violation of even the smallest excerpts, as they occur with the sound sampling [24].

5 Conclusion

In principle, only the adoption of free or lawfully licensed works is allowed for processing as a sample. If it is unclear whether sampling should be carried out, a sample-clearing with their respective rights holders and collecting societies can help.

With regard to the validation of digital audio and video recording a common method recognized by expert forensics is the Electric Network Frequency Analysis (ENF). With this method it is even possible to determine the exact location and the exact time of the production of a recording. Visual procedures such as the spectrogram representation are important methods for aiding detection of manipulation. In the authenticity analysis, the determination of the originality and continuity of the recordings and the detection of changes are of particular importance.

Sound Sampling will continue to win in importance and new extraction methods (sound separation) which can extract the whole melody will exacerbate the problem of piracy. On the other hand, the improved analysis and marking processes such as the watermarking technology offer more possibilities for the protection and detection as well as prosecution of copyright violations.

References

[1] M. Häuser, *Sound und sampling: Der Schutz der Urheber, ausübenden Künstler und Tonträgerhersteller gegen digitales Soundsampling nach deutschem und US-amerikanischem Recht*. Dissertation. München: Beck, 2002.

[2] E. Adeney, "The sampling and remix dilemma: What is the role of moral rights in the encouragement and regulation of derivative creativity," (English), *Deakin Law Review*, vol. 17, no. 2, pp. 335–348, 2012.

[3] T. M. Jörger, *Das Plagiat in der Popularmusik*. Dissertation, 1st ed. Baden-Baden: Nomos Verlagsgesellschaft, 1992.

[4] B. Wessling, *Der zivilrechtliche Schutz gegen digitales Sound-Sampling: Zum Schutz gegen Übernahme kleinster musikalischer Einheiten nach Urheber-, Leistungsschutz-, Wettbewerbs- und allgemeinem Persönlichkeitsrecht*. Dissertation. Baden-Baden: Nomos Verl.- Ges, 1995.

[5] M. Pendzich, *Von der Coverversion zum Hit-Recycling: Historische, ökonomische und rechtliche Aspekte eines zentralen Phänomens der Pop- und Rockmusik*. Dissertation. Münster: LIT, 2004.

[6] Ohne Verfasser, *GEMA: Schutzfähige Bearbeitungen freier Werke*. Journal Article

[7] U. Loewenheim and B. von Becker, *Handbuch des Urheberrechts*. Textbook, 2nd ed. München: Beck, 2010.

[8] R. Amon, *Lexikon der Harmonielehre: Nachschlagewerk zur durmolltonalen Harmonik mit Analysechiffren für Funktionen, Stufen und Jazz-Akkorde*. Textbook. Wien, Stuttgart: Doblinger; Metzler, 2005.

[9] R. Moser, *Handbuch der Musikwirtschaft*. Textbook, 6th ed. Starnberg u.a: Keller, 2003.

[10] G. Berndorff, B. Berndorff, and K. Eigler, *Musikrecht: Die häufigsten Fragen des Musikgeschäfts ; die Antworten*. Textbook, 6th ed. Bergkirchen: PPV-Medien, 2010.

[11] A. J. Cooper, *Detection of copies of digital audio recordings for forensic purposes*. Milton Keynes: Open University, 2006.

[12] P. Wegener, *Sound Sampling: Der Schutz von Werk- und Darbietungsteilen der Musik nach*

schweizerischem Urheberrechtsgesetz. Dissertation. Basel: Helbing Lichtenhahn, 2007.

[13] R. Münker, *Urheberrechtliche Zustimmungserfordernisse beim Digital Sampling*. Dissertation. Frankfurt am Main, New York: P. Lang, 1995.

[14] T. Meschede, *Der Schutz digitaler Musik- und Filmwerke vor privater Vervielfältigung nach den zwei Gesetzen zur Regelung des Urheberrechts in der Informationsgesellschaft*. Dissertation. Frankfurt/Main, New York: P. Lang, 2007.

[15] R. Korycki, "Time and spectral analysis methods with machine learning for the authentication of digital audio recordings," (Englisch), *Forensic Science International*, vol. 230, no. 1-3, pp. 117–126, 2013.

[16] C. Grigoras, "Applications of ENF criterion in forensic audio, video, computer and telecommunication analysis," (Englisch), *Forensic Science International*, vol. 167, no. 2-3, pp. 136–145, 2007.

[17] H. Meister, *Elektronik: Mit Versuchsanleitungen und Rechenbeispielen,* 8th ed. Würzburg: Vogel-Buchverlag, 1986.

[18] B. E. Koenig and D. S. Lacey, "An Inconclusive Digital Audio Authenticity Examination: A Unique Case," (Englisch), *Journal of Forensic Sciences*, vol. 57, no. 1, pp. 239–245, 2012.

[19] J. Webers, *Tonstudiotechnik: Handbuch der Schallaufnahme und -wiedergabe bei Rundfunk, Fernsehen, Film und Schallplatte,* 4th ed. München: Franzis, 1985.

[20] M. Abramovici, *Kennzeichnungstechnologien zum wirksamen Schutz gegen Produktpiraterie: Mit Ergebnissen aus Projekten MobilAuthent, O-Pur, EZ-Pharm*. Research Project. Frankfurt am Main: VDMA, 2010.

[21] M. Schäfer and T. Hansen, *ISCR - International Standard Recording Code: Das ISRC-Handbuch*. Textbook. Available: http://www.musikindustrie.de/fileadmin/news/publikationen/vb_isrc_handbuch.pdf (2014, Feb. 17).

[22] M. Faundez-Zanuy, J. J. Lucena-Molina, and M. Hagmüller, "Speech Watermarking: An Approach for the Forensic Analysis of Digital Telephonic Recordings*," (Englisch), *Journal of Forensic Sciences*, vol. 55, no. 4, pp. 1080–1087, 2010.

[23] D. C. Lynch and L. Lundquist, *Digital money: The new era of Internet commerce*. Journal Article. New York: Wiley, 1996.

[24] S. V. Dhavale, "Lossless Audio Watermarking Based on the Alpha Statistic Modulation," (Englisch), *IJMA*, vol. 4, no. 4, pp. 109–119, 2012.

[25] M. A. Nematollahi and S. A. R. Al-Haddad, "An overview of digital speech watermarking," (Englisch), *Int J Speech Technol*, vol. 16, no. 4, pp. 471–488, 2013.