Audio fingerprinting to detect illegal content in digital media files

Anil Alexander, Oscar Forth, and Alankar Atreya
Oxford Wave Research Ltd, Oxford, United Kingdom
{anil|oscar|alankar}@oxfordwaveresearch.com

Police seizures of computers or mobile devices as part of an investigation can often lead to the collection of a large amount of digital evidence which includes documents, images, audio recordings and videos. In the case of terrorism or sexual abuse related investigations, it may be required to check for prohibited content in images, audio or video files on seized devices. If present, this material may further require identification and classification according to severity and importance to the investigation. The assignment and classification of such content is normally performed manually by trained police officers. This is a time-consuming task that can affect those who do it. According to the Association of Chief Police Officers (ACPO) Good Practice Guide for Digital Evidence (2011), watching disturbing material can have a significant psychological effect on the officers who view it. Additionally, copies of media files present on one device, for instance a mobile phone, may be present on other devices such as computers and tablets as well. Further, the same or similar files or even sub-clips may be found as part of other investigations. The possibility of 'fingerprinting' files with illegal content would provide an efficient and less human-intensive means of finding and identifying such material in the future. The term fingerprinting is appropriate in this context because it is based on the premise that the series of acoustic events happening in a recording do indeed have a unique signature. It is important to distinguish the well-established term 'audio-fingerprinting' (Cano et al, 2005), which relies on the assumption of uniqueness, from 'voice-printing' of spectrograms where such uniqueness has not been observed.

If the media file is an unaltered copy, it is relatively straightforward to produce some hash value that is related to the bitwise content of the files (e.g. MD5 cryptographic checksums). However, both the video and audio can be changed by intentional and unintentional format conversions, and editing, requiring analysis of the visual and acoustic content of the files. Video comparison techniques that use visual content need to be robust to rotation, cropping, and colour conversions. In addition, they should be able to detect if a small sub-clip of a longer file is present within another file. In cases where audio is present in the recordings, the audio signal can be used to create a robust signature of what is being recorded. Audio fingerprinting can also be applied to video files as audio stands a better chance of not being affected by changes normally made in the visual space. In this work, we have developed a compact acoustic signature for audio and video that it is robust to changes in formats, levels and also the addition of extraneous noise. Figure 1 shows an example of such a signature.

As the original audio or video data cannot be reconstructed from the fingerprints, privacy and evidential concerns need not limit the use of fingerprints across cases, and indeed across police forces and jurisdictions. This allows just the fingerprints to be transferred without necessitating the actual data to be provided. Using a light-weight fingerprint format for cross-comparisons is computationally more efficient. If a file has been flagged up as containing objectionable content in one search, its signatures can be stored and comparisons can be run against this signature.

Audio fingerprinting thus provides a robust and efficient method of detecting illegal material within video and audio files and thus reduces the effort and impact of performing this task manually.

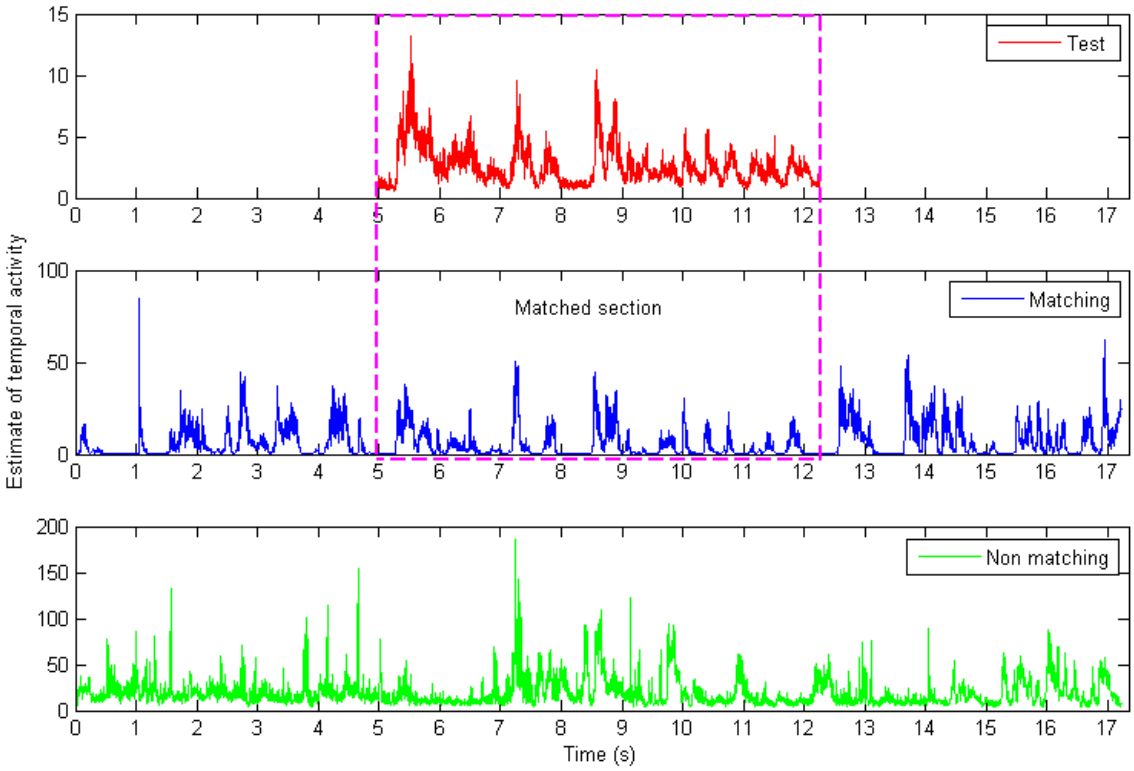

Figure 1 Example of a (simplified) audio fingerprint being compared against a matching and a non-matching file. In case of a match this provides the exact time of match. The purple-dashed matched section shows matching temporal activity.

References

ACPO (Association of Chief Police Officers) Good Practice Guide for Digital Evidence, Version Five (October 2011)

Anil Alexander, Oscar Forth, and Donald Tunstall, "Music and noise fingerprinting and reference cancellation applied to forensic audio enhancement," Audio Eng. Soc. 46th Int. Conf.: Audio Forensics, Denver, CO, pp. 29-35, June 2012

Pedro Cano, Eloi Batlle, Ton Kalker and Jaap Haitsma 2005 'A review of audio fingerprinting', Journal of VLSI Signal Processing,. 41 271–84

Oscar Forth and Anil Alexander, 'Content Comparison and Analysis (COCOA) of Contemporaneously Recorded Audio Material' International Association of Forensic Phonetics and Acoustics (IAFPA) conference 2014, Zurich, Switzerland