Content Comparison and Analysis (COCOA) of Contemporaneously Recorded Audio Material

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

The proliferation of handheld audio and video recorders and cheap data storage media in recent years has resulted in a large amount of audio and video evidence that is collected in both forensic and investigative tasks. It is also not unusual for an investigation to have several independent sources of audio, pictures or video, as in the case of the Boston marathon bombing in April 2013, where there was an appeal to the public for their recordings of the events leading up to the blasts. Searching, comparing, and extracting the relevant parts of the recordings as evidence is a time-consuming task, and can be helped significantly by automatic analysis techniques.

In Alexander et al (2012), we have presented a method for cancelling out music or television interference from forensic audio recordings using the so called, 'audio fingerprinting' method. The ability to 'fingerprint' a section of audio, based on the acoustic content present in it and to accurately time-align and 'subtract' the source material, allowed for significant improvement in the intelligibility of the target speech present in the audio. In this work, we extend this approach to all types of audio recording containing non-music speech or other sounds. We propose a novel method of comparing audio files using the acoustic content recorded in the files. If two or more different recordings contain the same acoustic events, it is possible to search for and identify the audio that is overlapping. This method will allow us to compare one audio file with many audio files in a directory, and provides a likelihood of match of a part or the whole of the files. A match provides the exact time of alignment between two recordings of the same event.

The proposed 'COCOA' method uses time correlations of the energy variation in the frequency spectrum to identify match points. The following are the three main applications of this method to forensic and investigative analyses:

- Content-based audio search: In certain forensic tasks, although a large quantity of audio or video data is analyzed, only a small section of audio is presented in evidence. However, it is sometimes necessary to provide the source recordings of various clips provided in court. This approach allows the forensic expert to search through audio recordings possibly from multiple cameras or recorders covering the same event.
- Intelligibility enhancement: Using time synchronization of a set of independent recorders, it is then possible either to use reference cancellation to reduce the effect of interfering noises or to mix devices for a better output. In Figure 1 and 2, we illustrate the exact time alignment of three recordings made using mobile telephones in a pub.
- Audio data de-duplication: During audio enhancement or speaker recognition work, many slightly modified copies of the audio or sections of the audio can be created on the expert's workstation. By analyzing the content of the audio files, it is possible to identify and group files that contain overlapping or similar audio content.

This approach successfully extends the scope of audio content comparison beyond recordings with distinct frequency patterns like music and television, to more general recordings. The initial results obtained using the approach show good performance even when comparing relatively clean recordings with severely degraded ones (e.g. from a damaged recorder). This audio content-based comparison approach can be applied to a variety of forensic audio and video related problems.

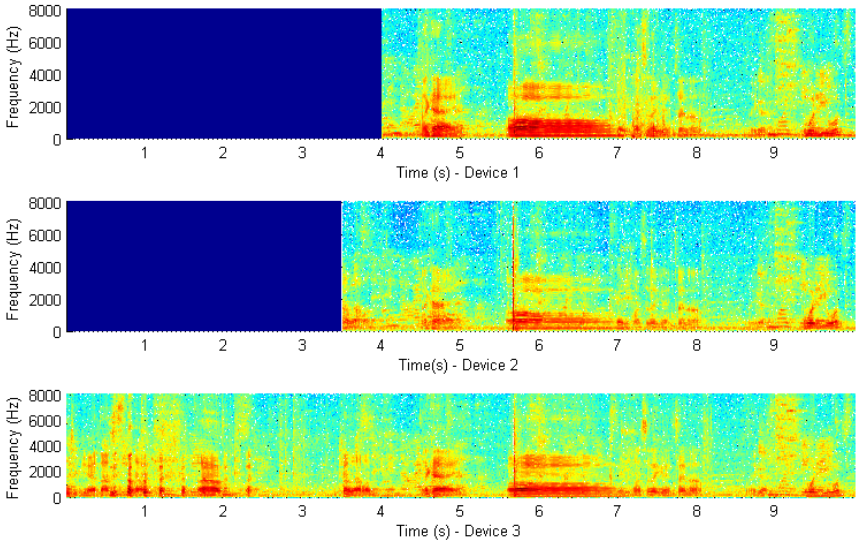

Figure 1 Time-aligned spectrograms of recordings from three independent mobile phones made in a pub environment using the COCOA method. All recorders were started at different times with device 3 started before both 1 and 2.

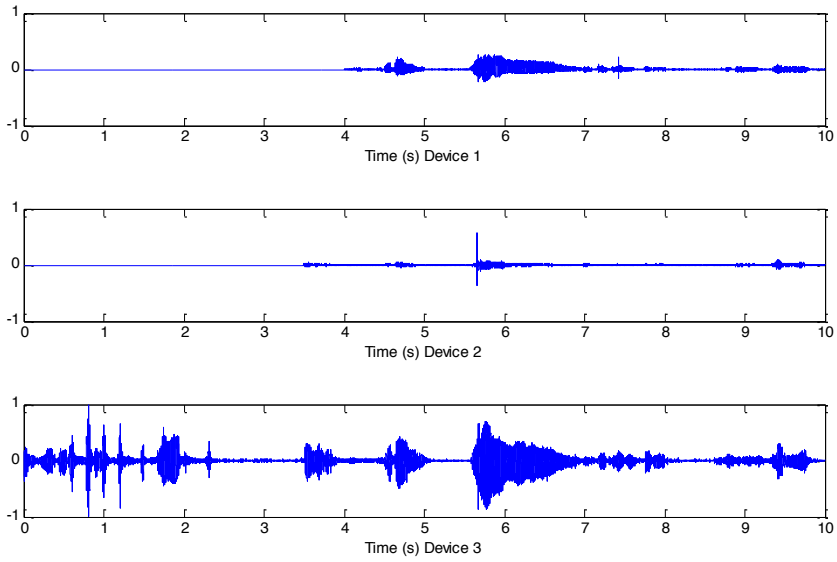

Figure 2 Time-aligned waveforms of the recordings from three independent mobile phones made in a pub environment using the COCOA method (same recordings as in Figure 1).

References

Alexander, A., Forth, O., and Tunstall, D., "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