Determine Audio File Authenticity

Audio Forensics

Author:

David Conner

PGP:
9E01 3E94 5908 1C7D 2C10
DEB6 9FCE 9B30 64CE D760

Audio Forensics

Data Collected on May 14, 2023, 4:21 AM EST

Analysis

The spoofed sample is *sound_s.wav*.

- The two WAV files are encoded at separate sampling rates.

- `Sound B` is encoded in mono at `16,000 Hz` which is particularly low. The Nyquist frequency is half the sample rate.

$$f_N = \frac{f_S}{2} \tag{1}$$

- When a sound is discretized into WAV, each value is an integer sample representing the amplitude. The fourier transform cannot extract frequencies approaching and exceeding f_N.

- `Sound S` is encoded at `48,000 Hz`. This file's spectral domain has a sharp dropoff at frequencies above it's f_N. Since the recording is at a higher sample rate, its STFT can return values at higher frequencies. These can be seen above `8 kHz` as a noise signature of less than `-60dB`.

- However, in the spectrogram for `Sound S`, there is no clear signal above `8,000 Hz`. If the cutoff was gradual, this could indicate frequency limitations in the microphones, like in the lower end of the spectrogram for `Sound B`. The cutoff is sharp, which indicates a reencoded recording without interpolation.

Sound	Label	Channels	Nyquist	Sampling	Resolution	Byte Count	Frame Count
sound_b.wav	bonafide	1 (mono)	8,000 Hz	16,000 Hz	16 Bit	288000	144000
sound_s.wav	spoof	2 (stereo)	24,000 Hz	48,000 Hz	16 Bit	1812480	453120

- The byte count represents the WAV audio data and does not include the metadata.

- There are `44 Bytes` of metadata in each file.

Tools

- Language: Python

- Libraries: `numpy`, `matplotlib`, `wavinfo`, and `librosa`

Methods

- Plotted the time-domain wave form for each WAV file.

- Performed STFT on each channel in each WAV file.

- Transformed STFT output from Amplitude to dB relative the maximum dB for all channels in the WAV file.

- Plotted the resulting spectral domain for each channel.

Sound B

Sound	Label	Channels	Nyquist	Sampling	Resolution	Byte Count	Frame Count
sound_b.wav	bonafide	1 (mono)	8,000 Hz	16,000 Hz	16 Bit	288000	144000
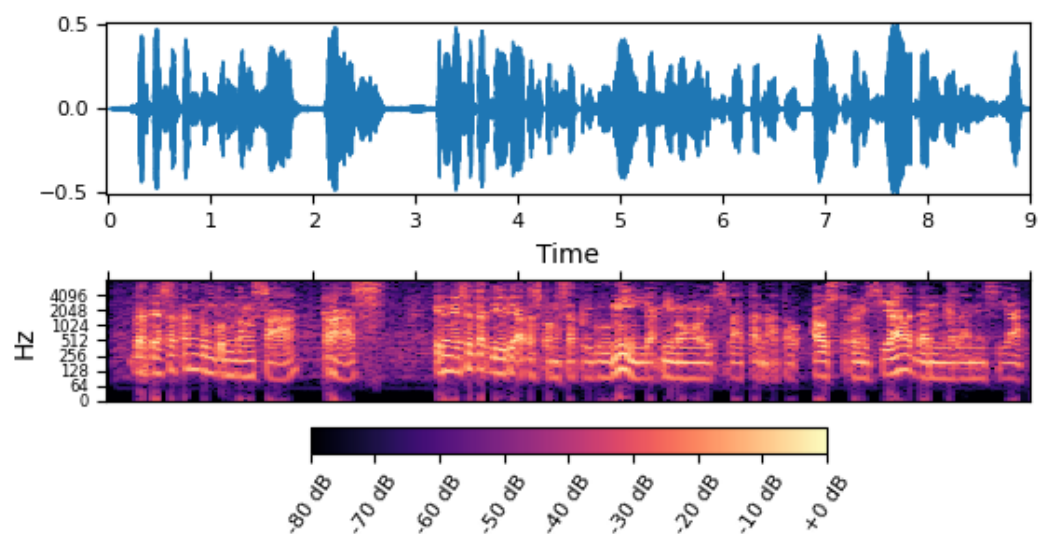
Sound S

Sound	Label	Channels	Nyquist	Sampling	Resolution	Byte Count	Frame Count
sound_s.wav	spoof	2 (stereo)	24,000 Hz	48,000 Hz	16 Bit	1812480	453120