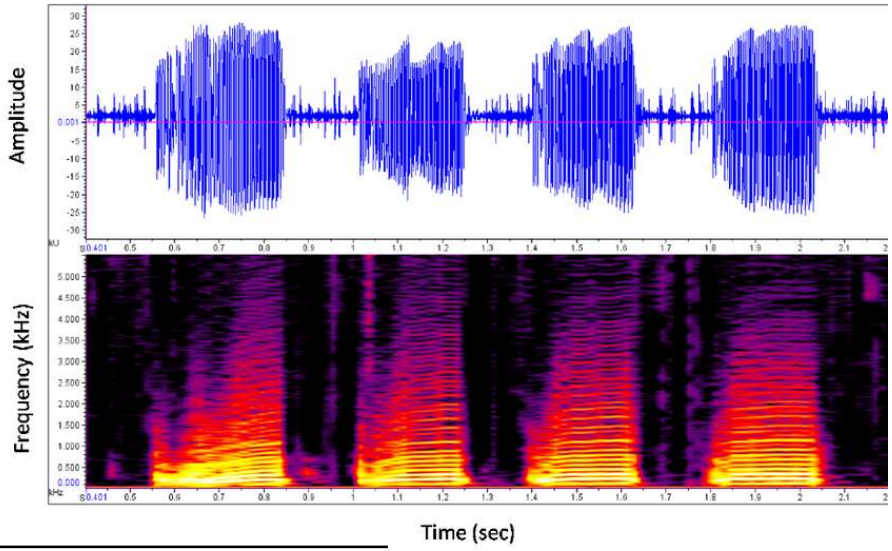
Multimedia fingerprinting

Audio fingerprinting

Shazam[2].

1. Compute the spectrogram of the audio piece.
   - A spectrogram is a 3D graph that relates time, frequency and amplitude of a signal, and is produced by computing an FFT/DCT on overlapping time windows of the original signal.



[2]Wang, An Industrial-Strength Audio Search Algorithm, 2005

Multimedia fingerprinting

Audio fingerprinting

2. Identify points with a higher amplitude than its neighbors. These points form a *robust constellation*. From that point on, amplitude is irrelevant.
3. Anchor points are chosen, each anchor point having a target zone associated with it. Each anchor point is sequentially paired with points within its target zone, each pair yielding two frequency components plus the time difference between the points $(f_1, f_2, \Delta t = t_2 - t_1)$.
4. Populate the DB with items $[H(f_1, f_2, \Delta t) : t_1 : Songname]$ for each pair point.