The Shazam System for Music Identification

Sylvain Marchand

after the presentation of
Shazam Entertainment (2003)
by its founder and chief scientist :

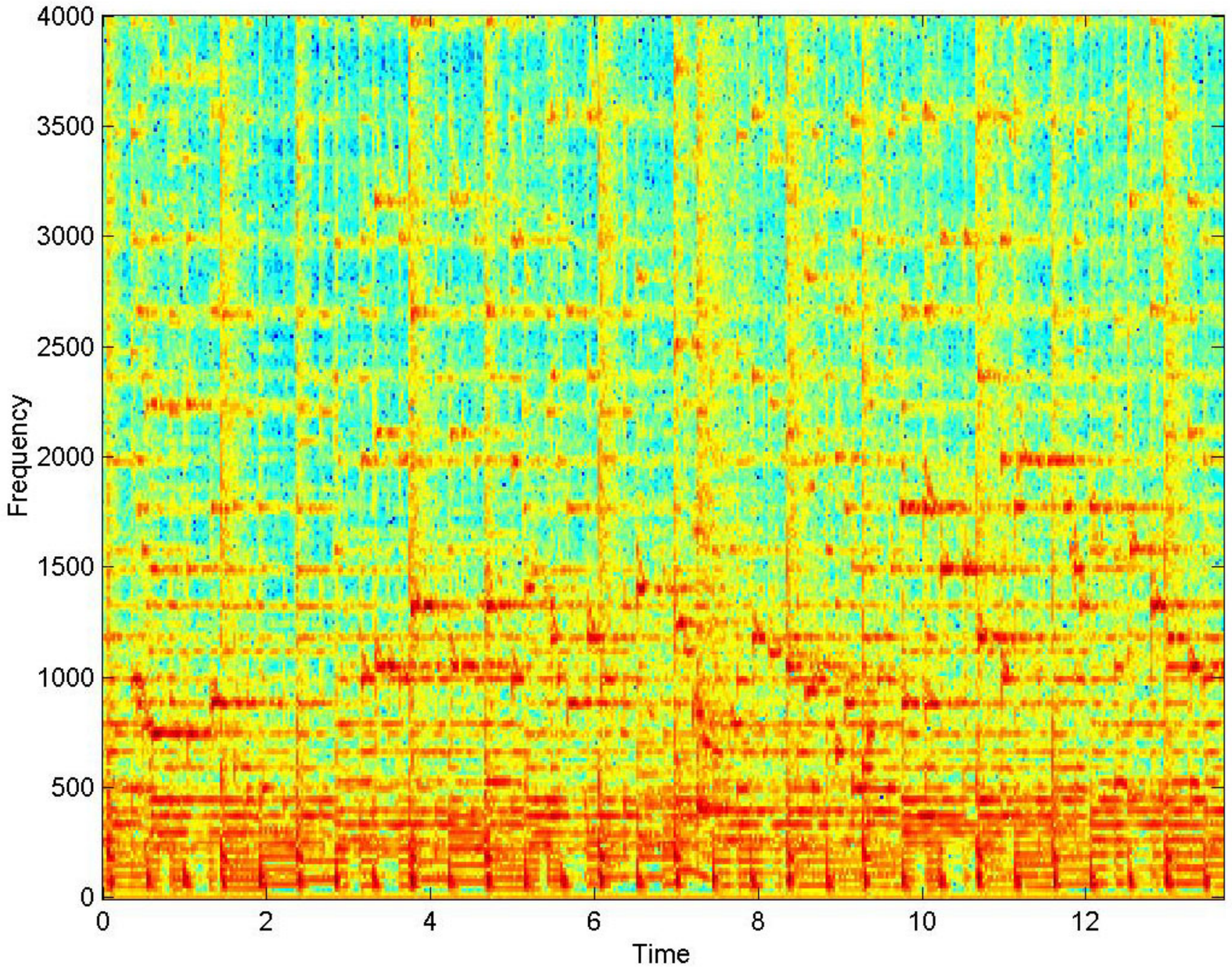
Avery Wang

What is Shazam?

How does it work?

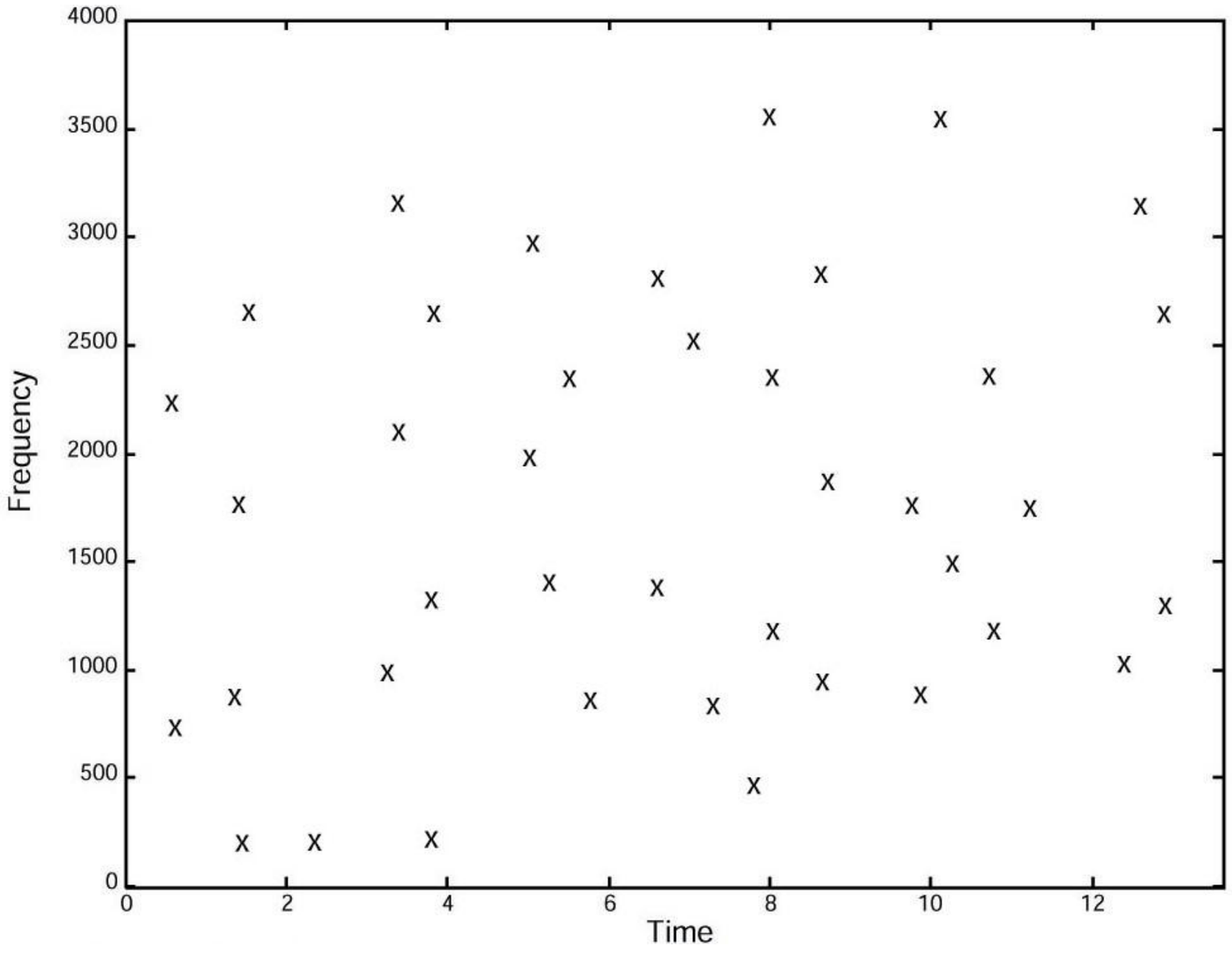
spectrogram

- magnitude of the short-term Fourier transform



spectral peaks

- local maxima in the magnitude spectrum

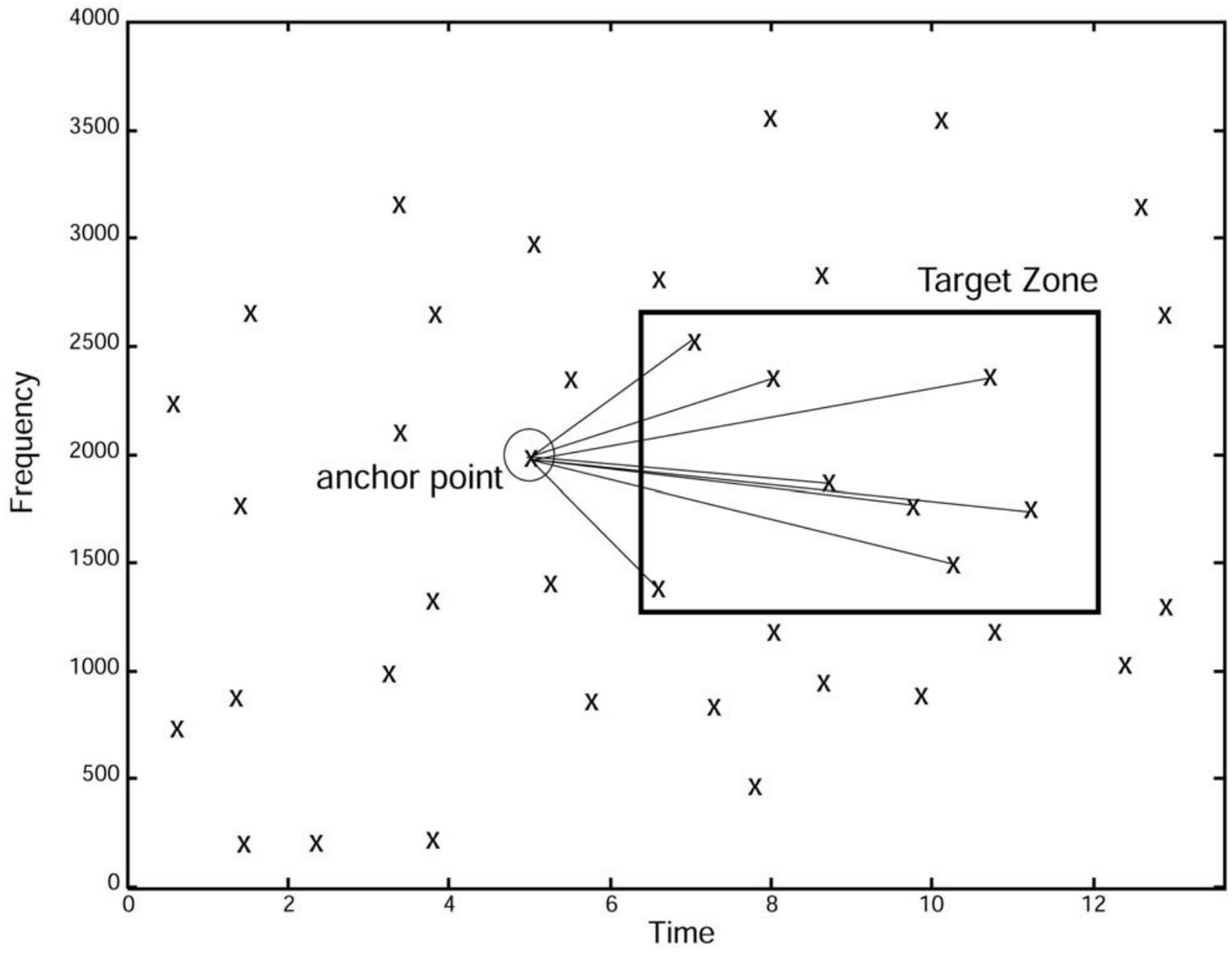


spectral peaks

- the most energetic components
- important to human perception
- tend to **survive in noisy conditions**
- but individual peaks have low entropy…

pairs of peaks

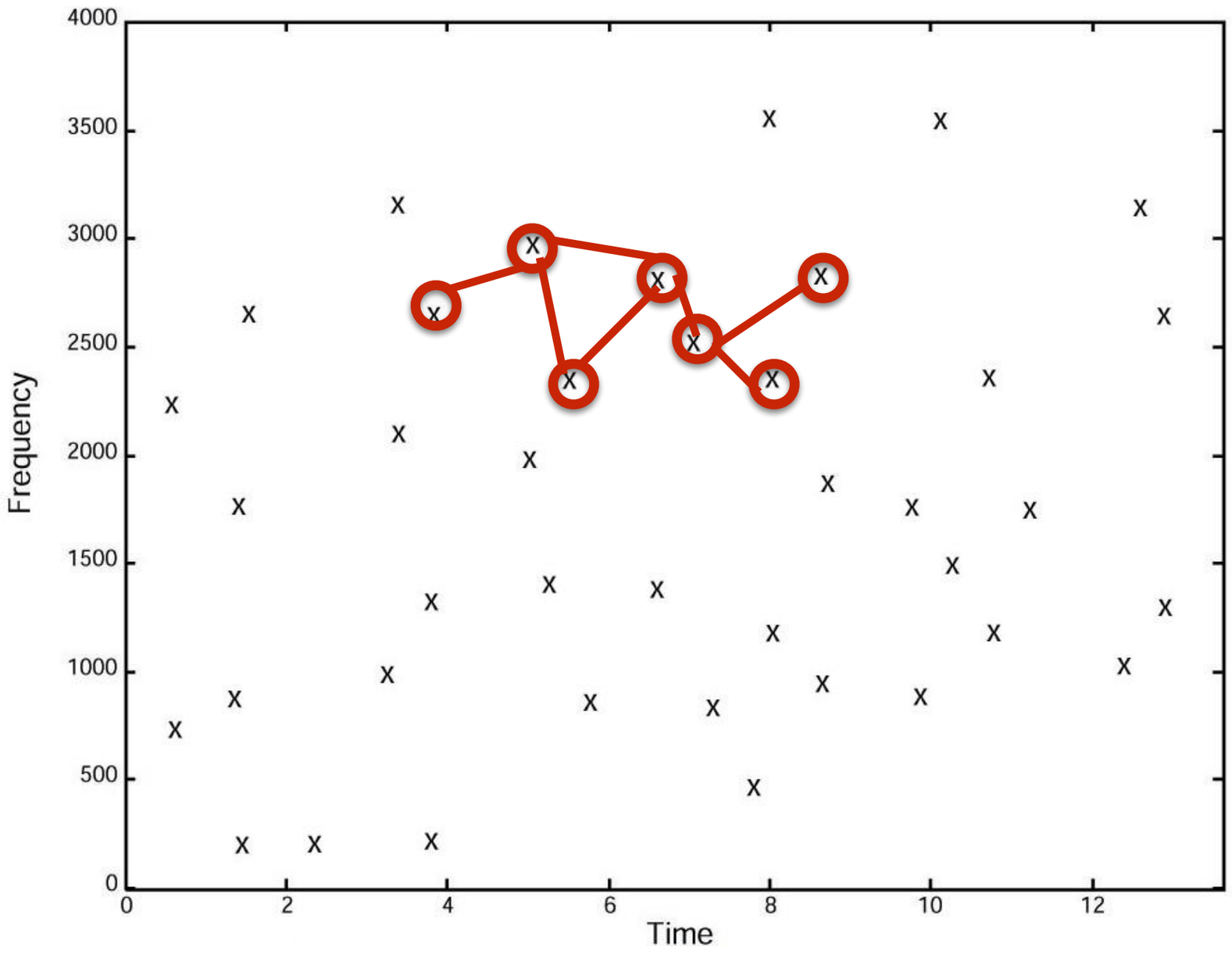
- each point is taken as an anchor point
- each anchor point has a target zone



pairs of peaks

- define the edges of a graph…
- this graph is a primary structure…
- suitable for the fingerprint of the sound!

fingerprint graph



graphs…

- How to compare (approximate) graphs?

- Well, don’t… focus on edges (peak pairs)!

combinatorial hashing

- store the peak pairs ((t1, f1), (t2, f2)) in a hash table
- the hash key is (f1, f2, t2-t1) to **resist time translation**
- the hash value is the time index t1 + song number
- remember that storing / searching to / from a hash table is O(1) complexity (thus **optimal**)…

search…

- in case of a match, there is a peak in the histogram of t1-t (returned database time - current time of the soundfile)

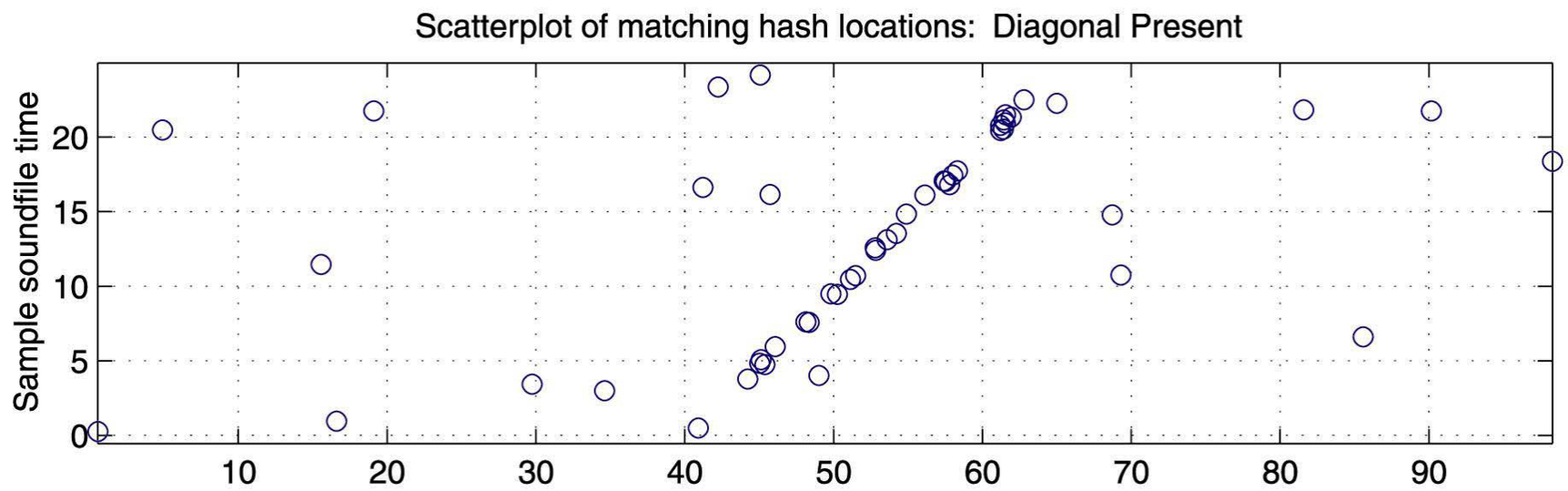


Fig. 3A

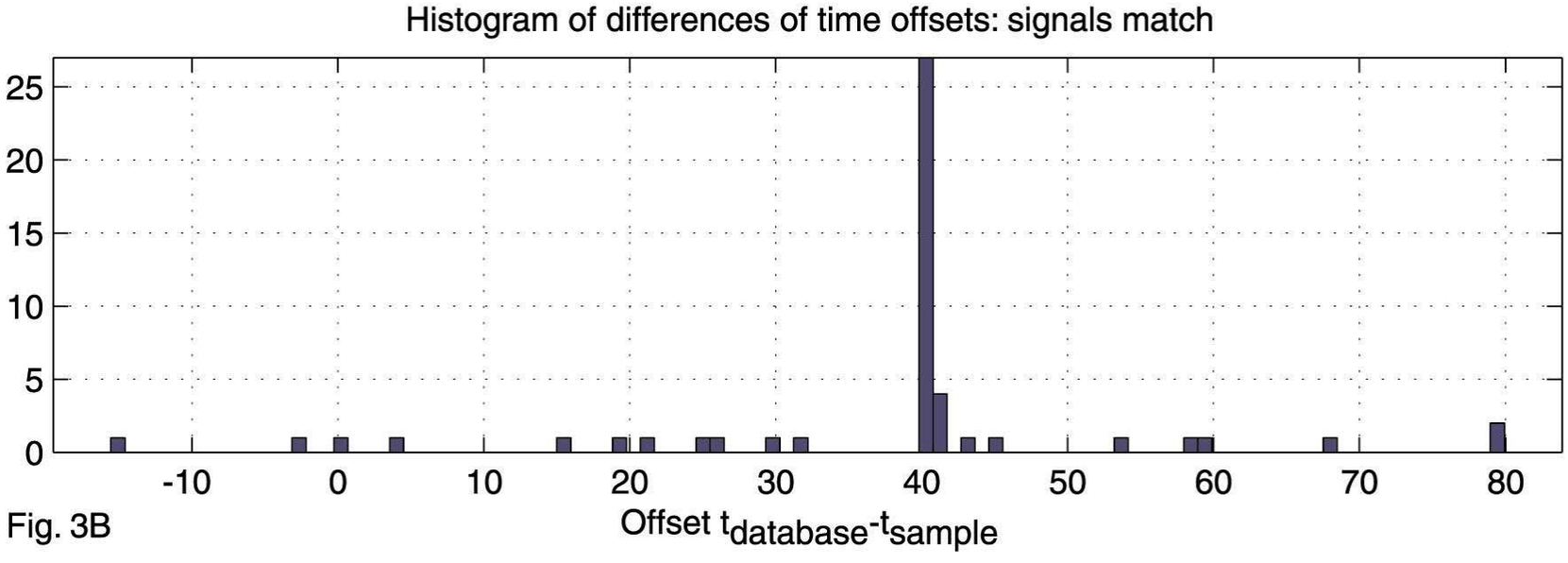


Fig. 3B

summary of the method

- extract spectral peaks
- form pairs of peaks
- use a hash table to store / query them
- and make some histogram of the queries