Exploring Music Collections by Browsing Different Views

Elias Pampalk, Simon Dixon, Gerhard Widmer
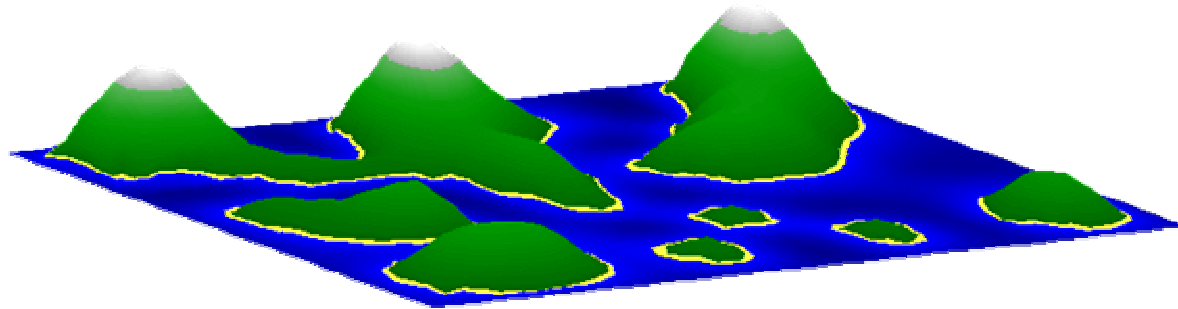
Austrian Research Institute for
Artificial Intelligence (OeFAI)
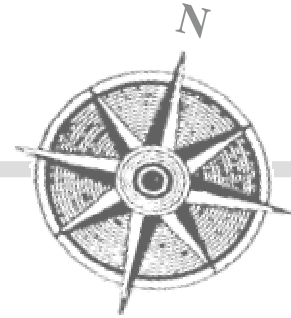Vienna, Austria

ÖFAI

Basic Idea

- ## Organize & Visualize Music Collections
  - Active (similarity-based) exploration
  - Previous work: Islands of Music with static definition of similarity

- ## Different Aspects of Similarity
  - Timbre, Rhythm, Melody, Lyrics, …

    Relevance depends on context

    → Allow user to interactively (and gradually) change focus

Outline

Self-Organizing Map (SOM)

- Map
  - Map with bins (aka units)
  - Each bin is associated with a prototype piece (aka weight vector)
  - Neighboring bins have similar prototypes
  - A song is placed in the bin with most similar prototype (BMB)

- Training (online version)
  - Random initialization of prototypes
  - Loop
    - Randomly select piece X
    - Find BMB
    - Adapt prototype of BMB to better represent X
    - Slightly adapt neighborhood of BMB to better fit X

- Characteristics
  - Clustering & Visualization
  - Non-linear projection
  - Focus on local distances ("Magnification")

Smoothed Data Histogram (SDH)

- Visualize distribution of pieces on the map

  Many pieces placed in a bin → cluster center

- Smoothing: robust ranking function

  Pieces "vote" for best-matching bins

  Best-matching gets N points, 2nd best gets N-1 points, …
  N-th best gets 1 point all others get 0 points.

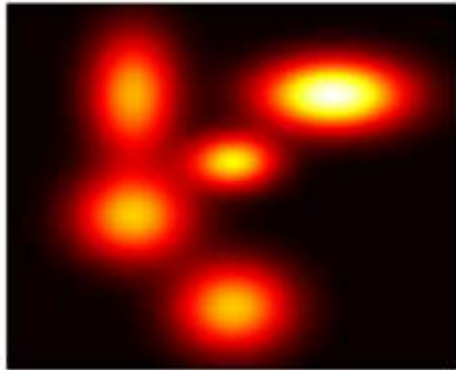 The distribution of votes is visualized

Matlab implementations (Toolboxes)
http://www.cis.hut.fi/projects/somtoolbox/ (Google: "SOM Toolbox")
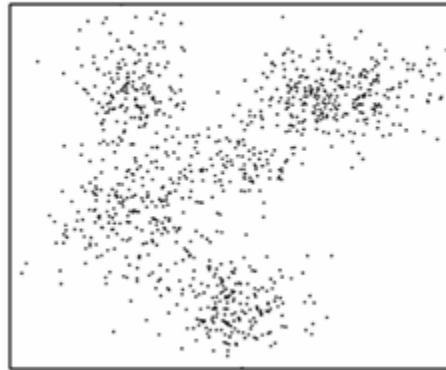http://www.ai.univie.ac.at/~elias/sdh/ (Google: "SDH Toolbox")
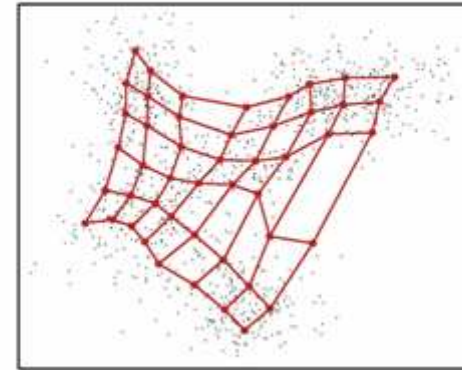
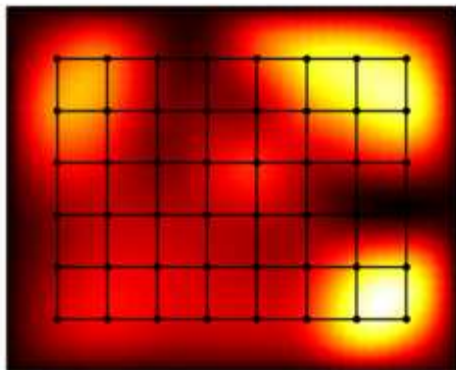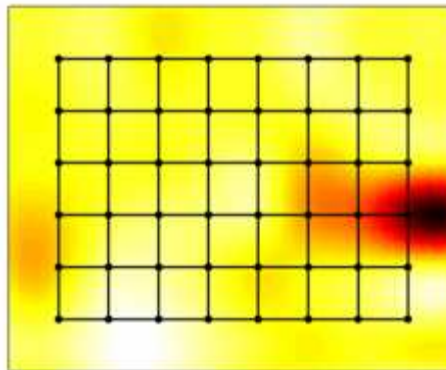# SOM & SDH Illustration

Data Space
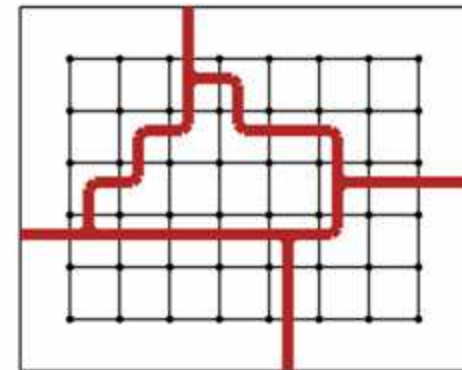
PDF

Sample (N = 1000)

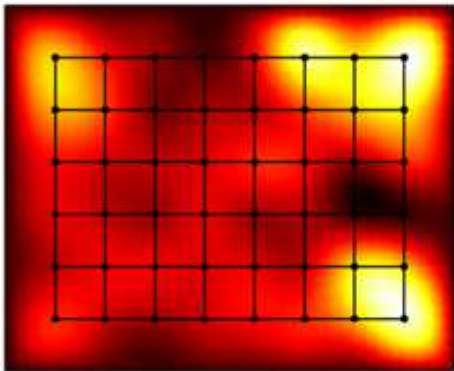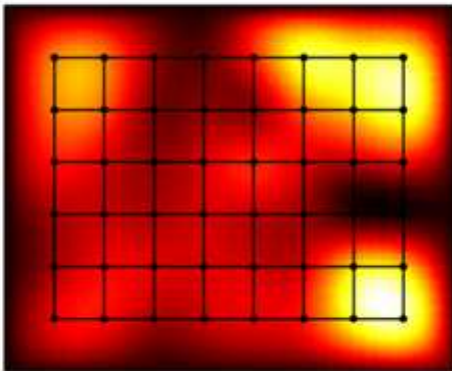SOM (8×6)

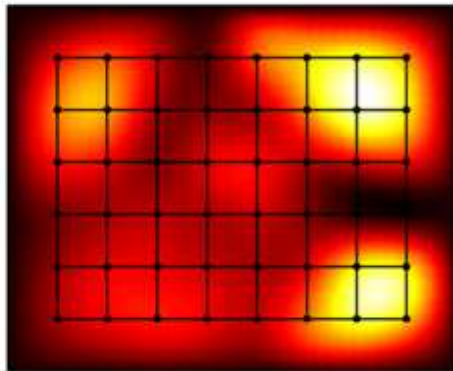**Visualization Space**

SDH (s = 3)

U–Matrix

K–Means

SDH Parameter
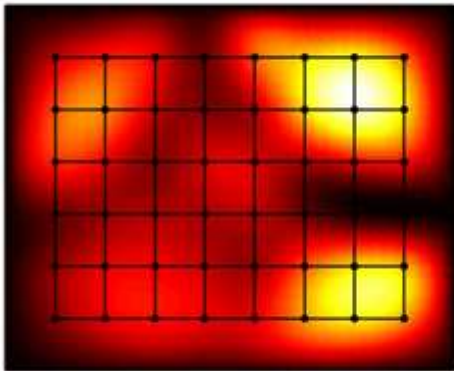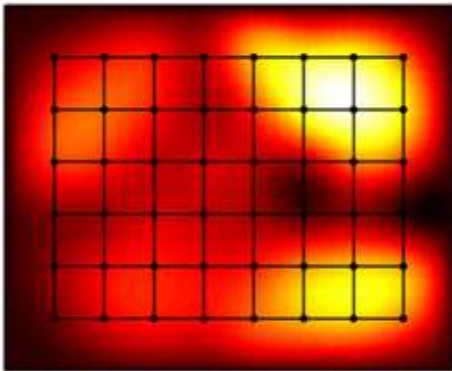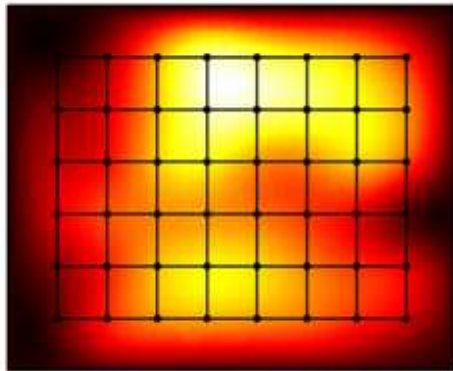
Aligned-SOMs Architecture

p_{min} p_{max} Parameter Values
(define different views of the data)

...   Stack of SOMs
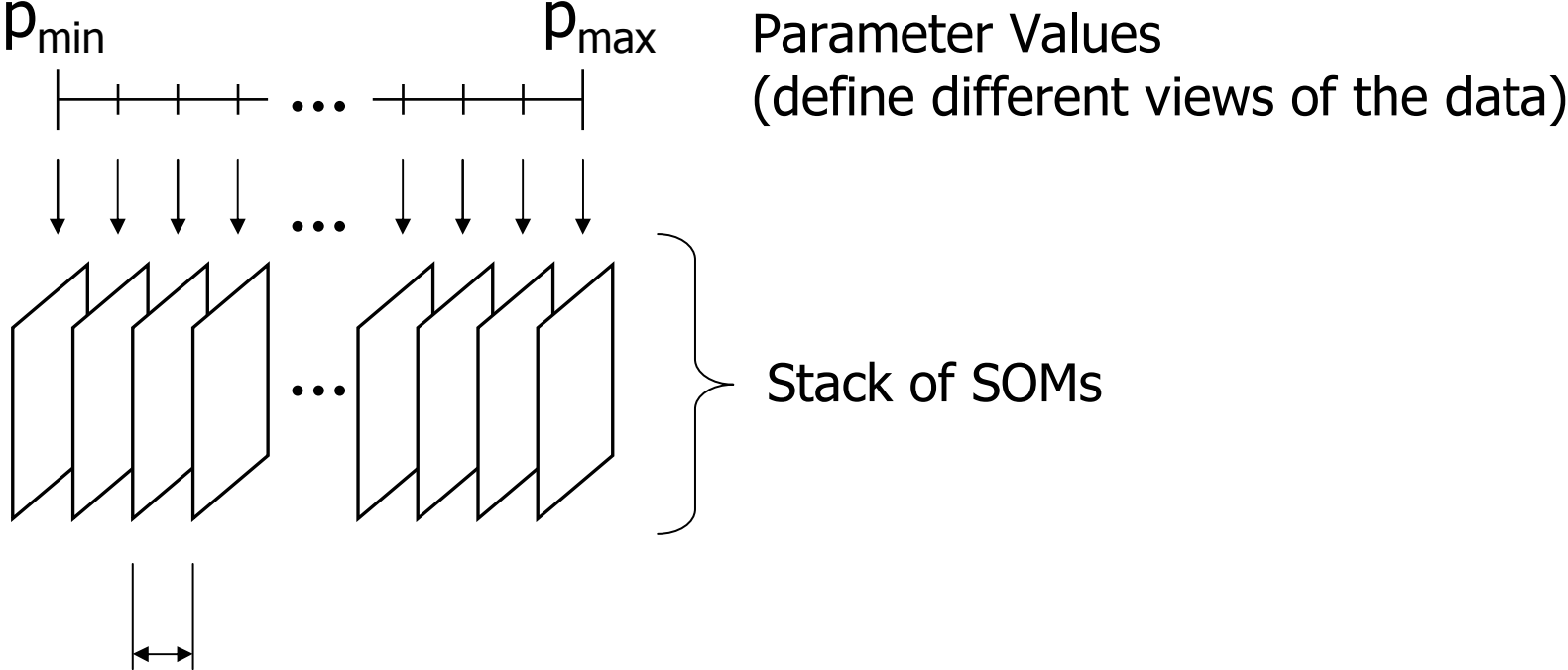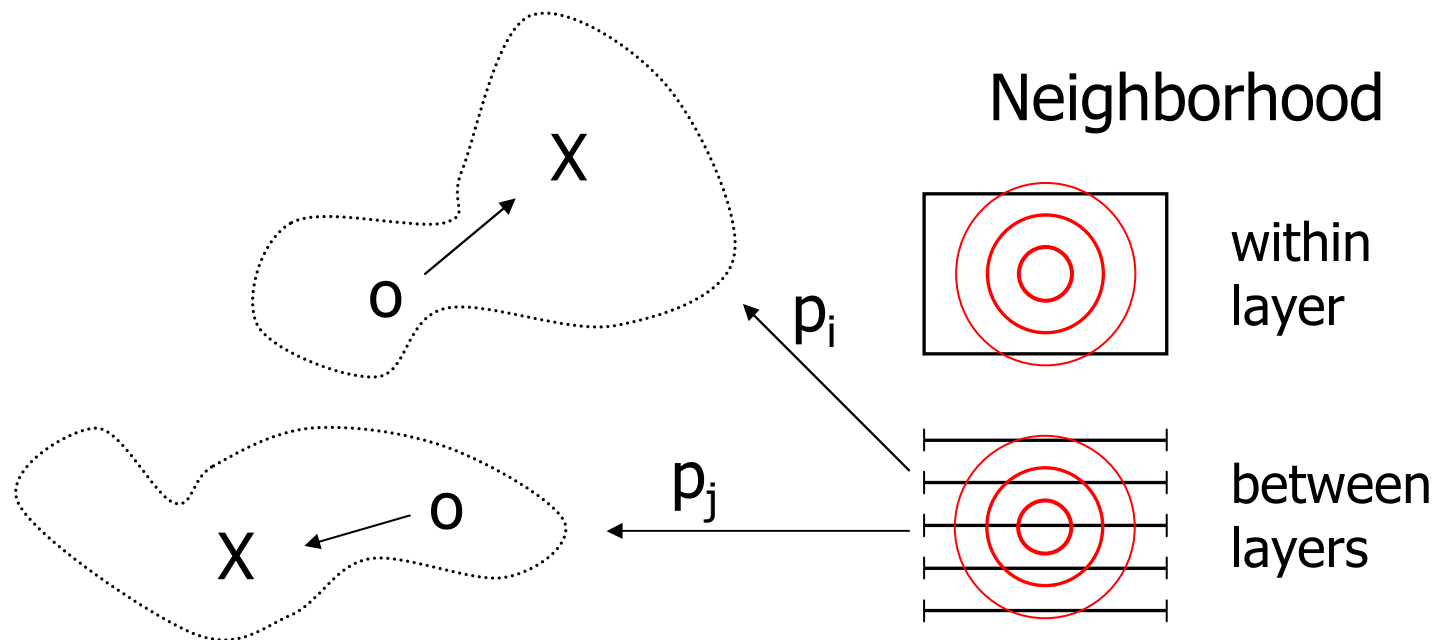
Distance between layers
(relative to distance between
units in same layer)
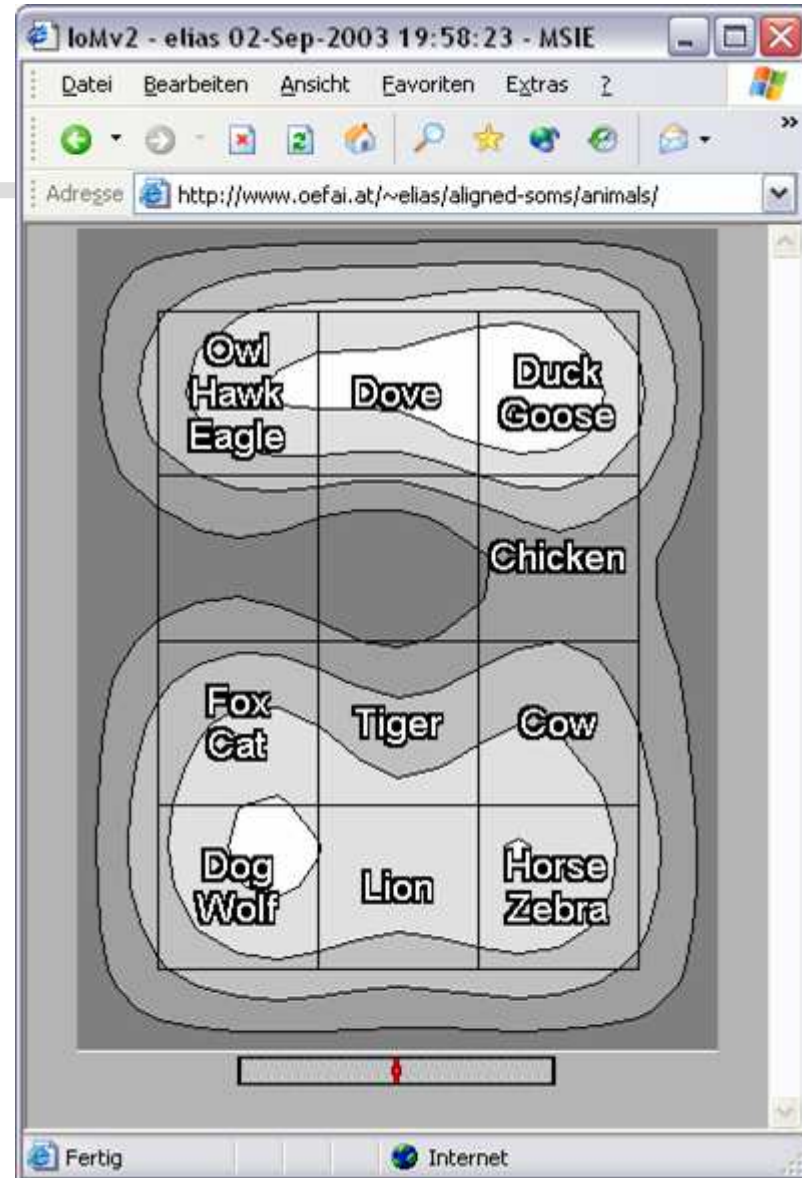
Aligned-SOMs Training (online version)

- Initialize all layers
- Loop
  - Randomly select piece X and layer L
  - Find BMB for X in L
  - Adapt neighborhood of BMB

Neighborhood

within layer

between layers

Aligned-SOM Illustration

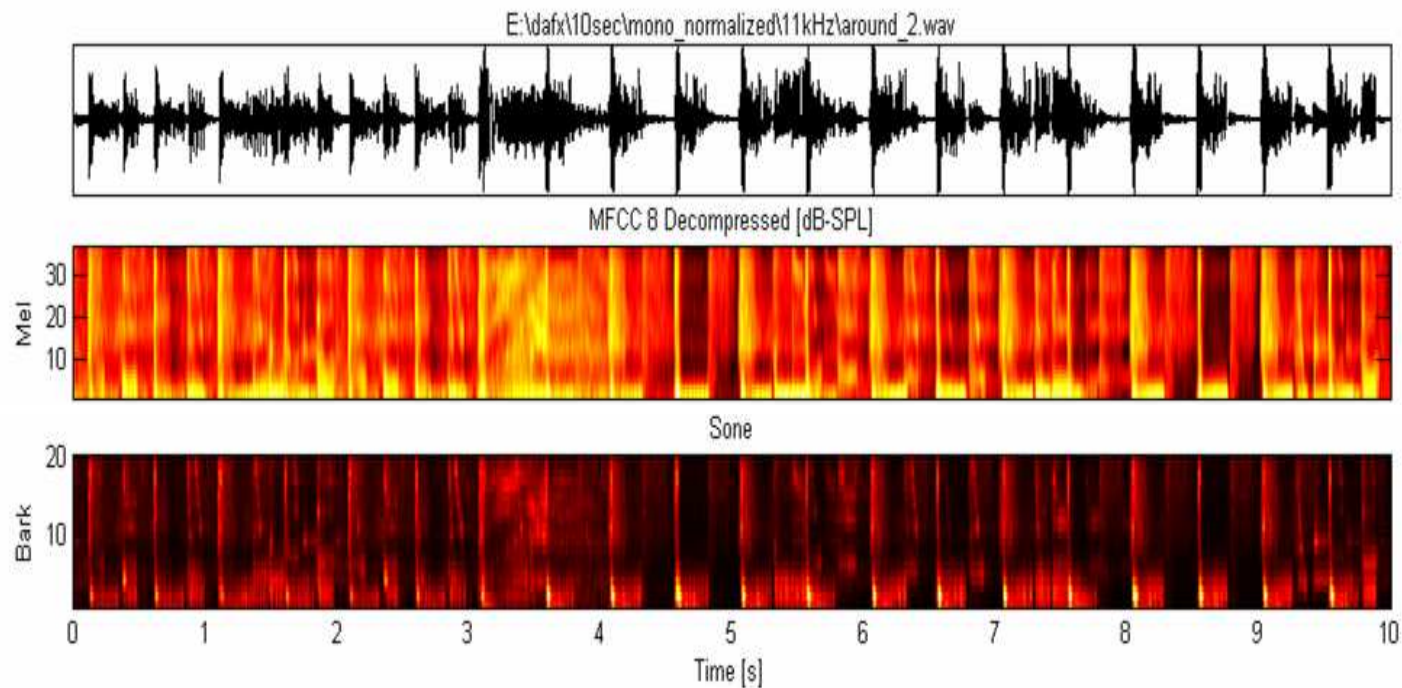http://www.oefai.at/~elias/aligned-soms/

Preprocessing

Zwicker-based Model

- Outer & middle ear
- Perceptual frequency scale: Bark
- Spectral masking
- Perceptual loudness scale: Sone

MFCC Model

- -
- Mel Scale (optimized for speech)
- Spectral smoothing through DCT
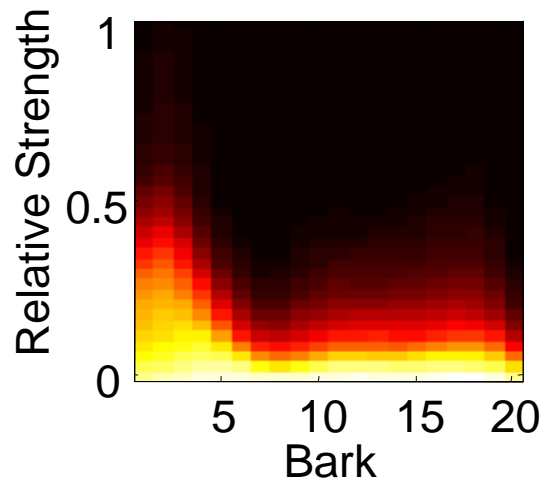- dB-SPL
- + Compressed representation

E:\dafx\10sec\mono_normalized\11kHz\around_2.wav

MFCC 8 Decompressed [dB-SPL]

Sone

Spectrum Histogram

Count how often loudness level is exceeded
for each frequency band
20 bands, 50 loudness levels –> 1000 bins
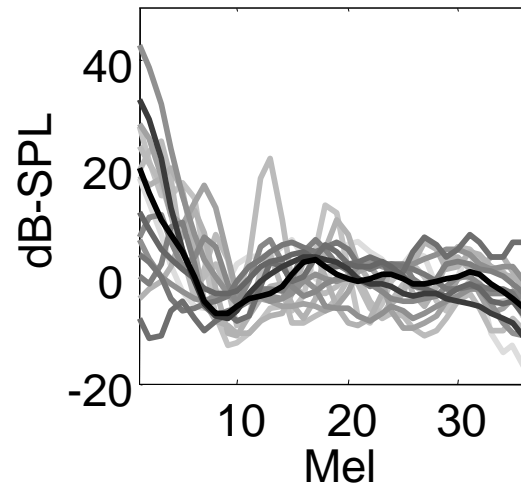Interpret as 1000 dimensional vectors
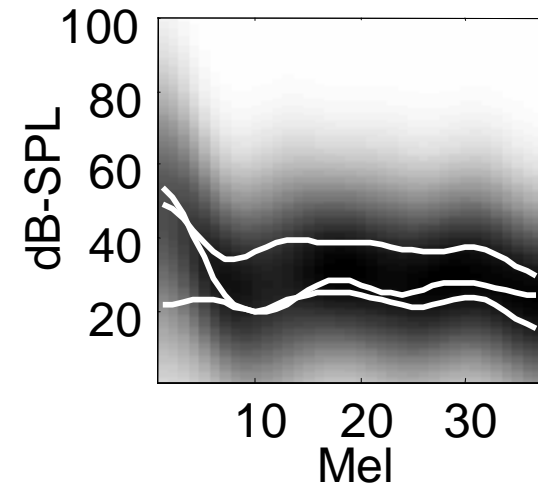(PCA compression possible)
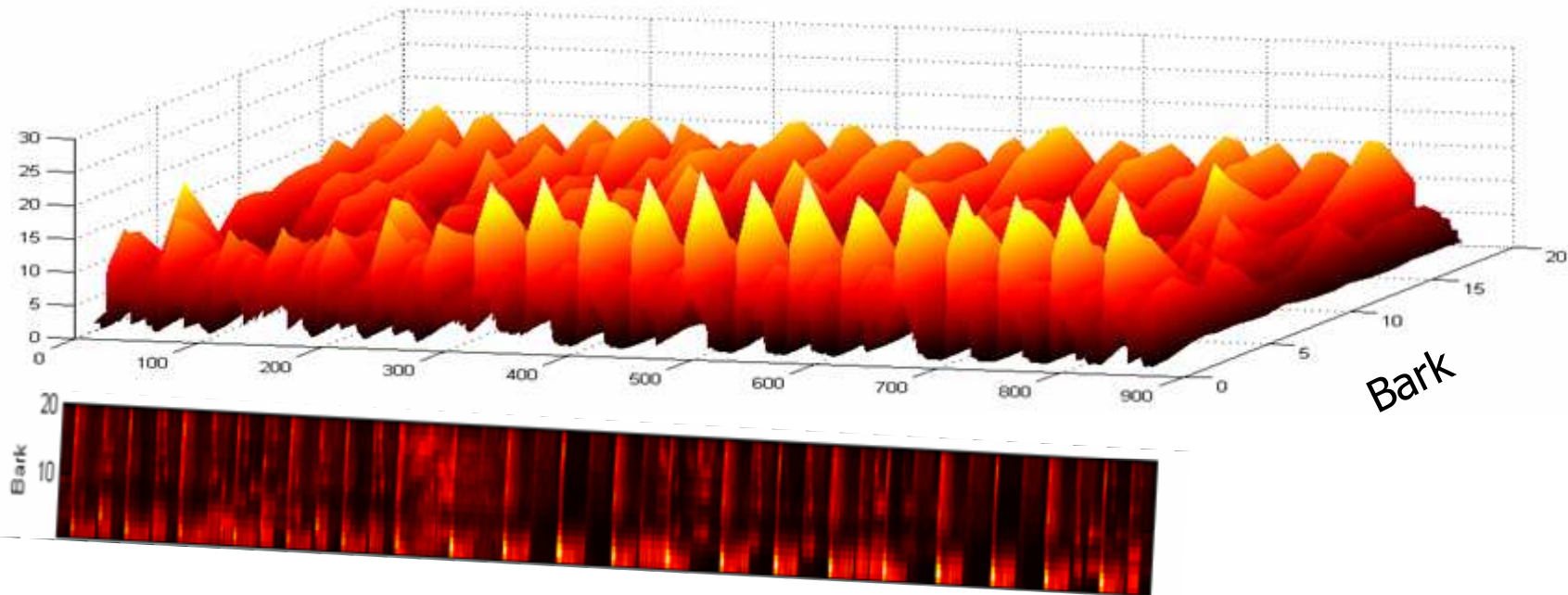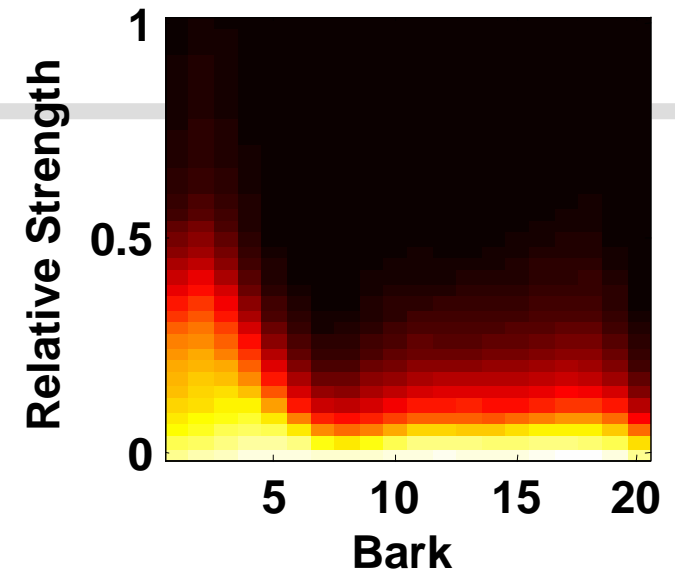
Spectrum Histogram – Alternatives

Logan & Salomon
(ICME 2001)

Aucouturier & Pachet
(ISMIR 2002)

+ Euclidean vector space
+ Fast to compute
− Too simple?

MFCC,
k-means,
EMD.

MFCC,
GMM-EM,
Sampling +
Log Likelihood.
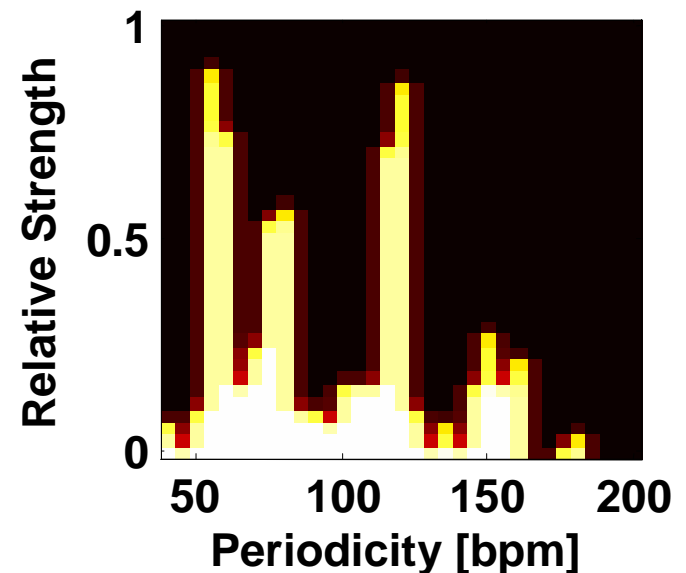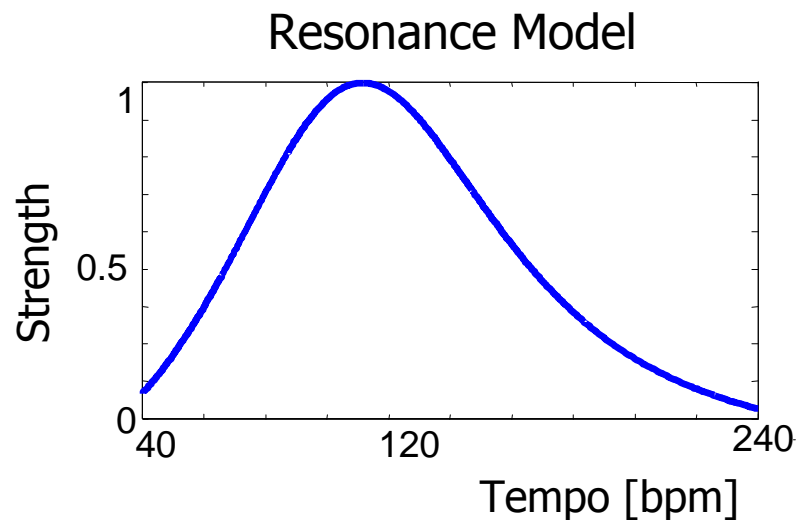
Periodicity Histogram

Comb-filter approach (E. Schreirer, 1998)
+ Weighting using resonance model of preferred tempo (Moelants, 2002)
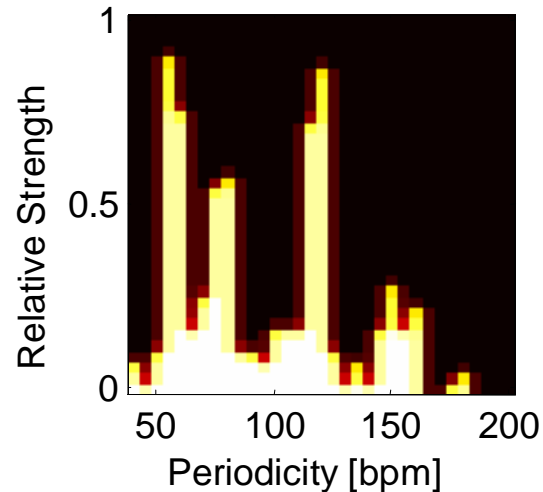
Each piece described by histogram
  Count how often strength level is exceeded for each periodicity
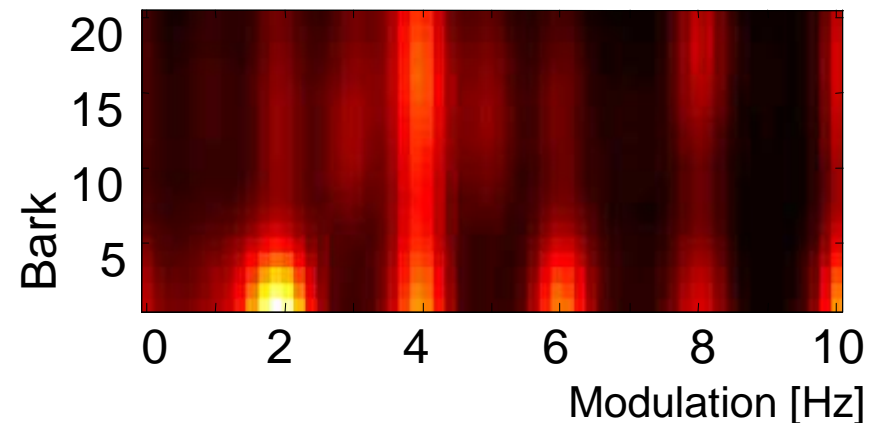 (40 periodicities, 50 strength levels –> 2000 bins)

Periodicity Histogram – Alternative

Fluctuation Pattern
(static measure used for previous Islands of Music, 2001)

+ Contains no spectral information
+ Closer related to "Rhythm"
− Slower to compute

Zwicker-based model,
FFT, Fluctuation Strength
(10bpm – 300bpm)

Demonstration

http://www.oefai.at/~elias/aligned-soms/

Conclusion

- First results are promising

- Limitations & Challanges
  - Quality of similarity measures
  - Zoom function to deal with larger collections (currently constrained by plain HTML)

  - Integration of meta-data?
  - Usability?