

A Matlab Toolbox to Compute Music Similarity

Elias Pampalk, Austrian Research Institute for Artificial Intelligence

Implemented Similarity Measures

Frame Clustering / Cluster Model Similarity (FC/CMS) [1,2,3]

- frame-based (12ms - 40ms) representation of spectral envelopes (e.g. using MFCCs)
- summarize frames through clustering (k-means or GMM)
- compare summaries (cluster models) using Earth Movers Distance or Monte-Carlo sampling

Fluctuation Patterns (FP) [4]

- describe loudness fluctuations per frequency band
- based on FFT over time dimension in sonogram

Periodicity Histogram (PH) [5,6]

- for each possible periodicity count number and strength of beats
- based on comb-filter over time dimension of modified sonogram

Spectrum Histogram (SH) [7]

- for each frequency band: count how many times certain loudness levels are exceeded

Additional Functions for visualizations (feature extraction, Islands of Music) and simple evaluations. See examples on the left.

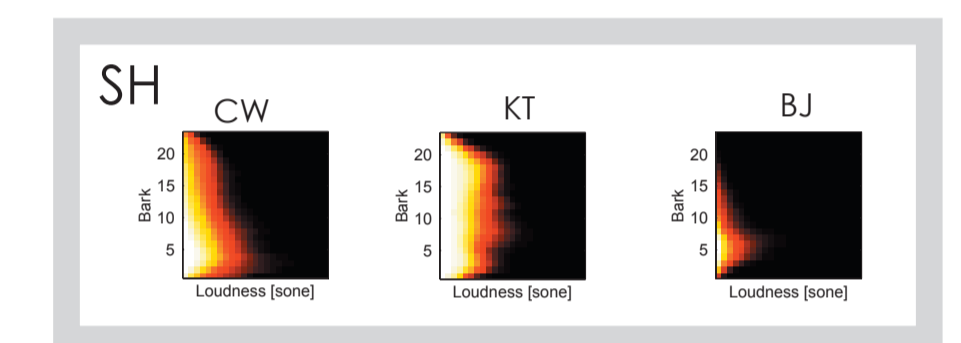
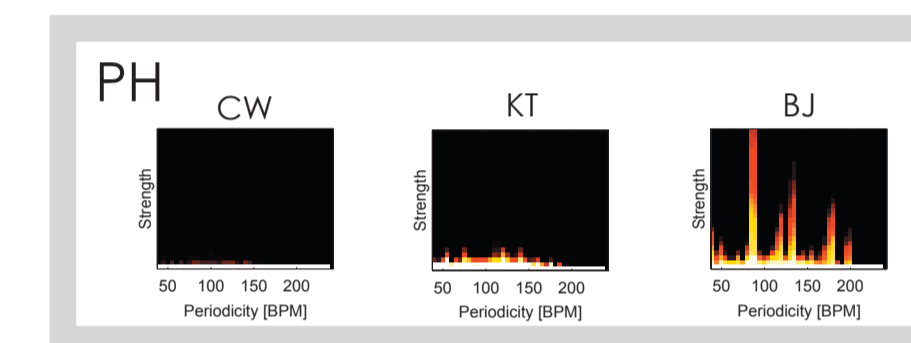
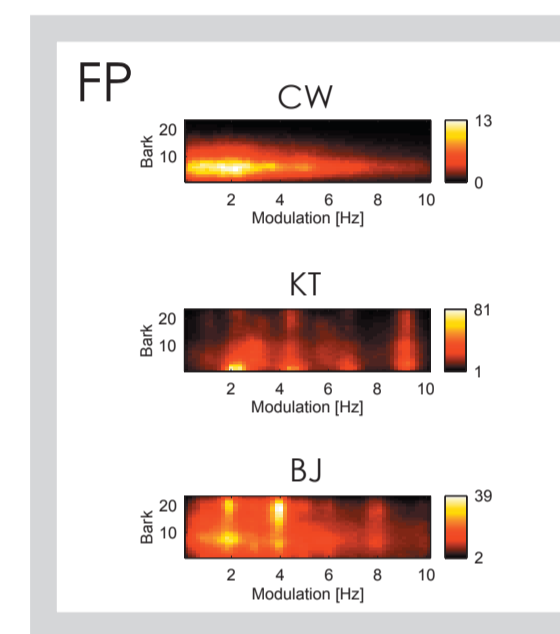
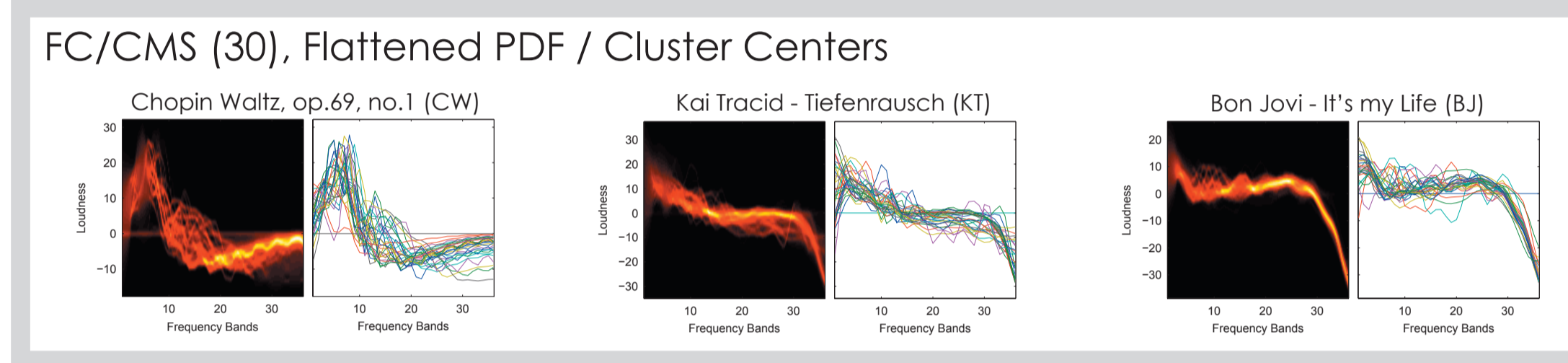
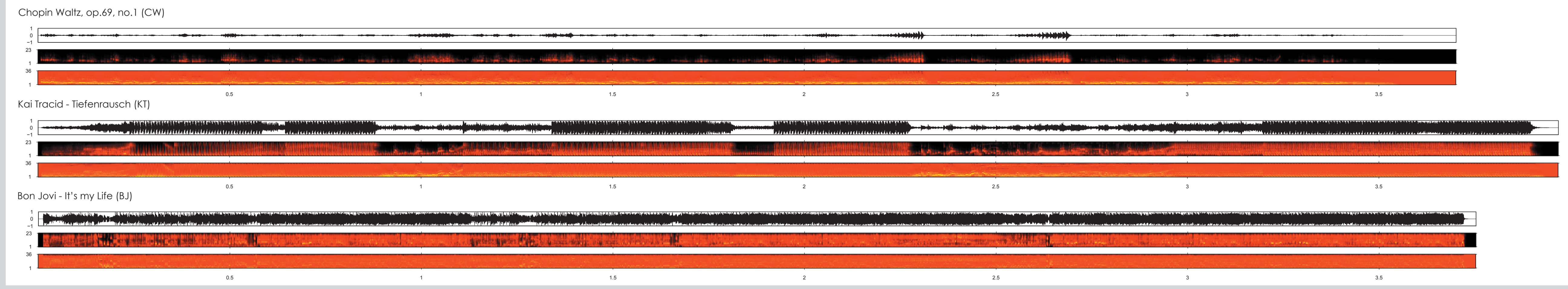
Requirements

- Matlab 6.0 or higher (Mathworks)
- Signal Processing Toolbox (Mathworks)
- Statistics Toolbox (Mathworks)
- Netlab Toolbox (freely available)
- EMD Wrapper (freely available)
- SOM Toolbox (GPL)
- SDH Toolbox (GPL)

MA toolbox is available from <http://www.oefai.at/~elias/ma>
GNU-GPL License

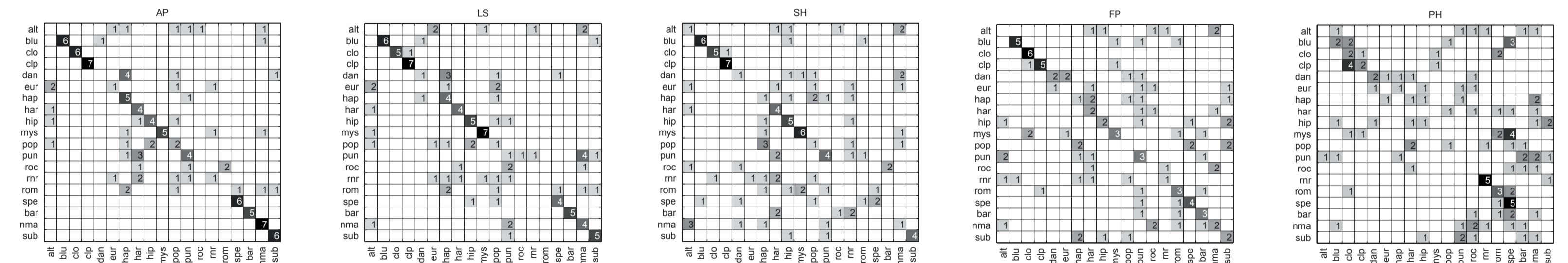
[1] B. Logan and A. Salomon, "A music similarity function based on signal analysis," ICME, 2001.
 [2] J.-J. Aucouturier and F. Pachet, "Music similarity measures: What's the use?," ISMIR, 2002.
 [3] J.-J. Aucouturier and F. Pachet, "Improving timbre similarity: How high's the sky?," in J. of Neg. Research Results in Speech and Audio Sci., vol. 1, no. 1, 2004.
 [4] E. Pampalk, A. Rauber, and D. Merkl, "Content-based organization and visualization of music archives," ACM MM, 2002.
 [5] E. D. Scheirer, "Tempo and beat analysis of acoustic musical signals," JASA, vol. 103, no. 1, 1998.
 [6] E. Pampalk, S. Dixon, and G. Widmer, "Exploring music collections by browsing different views," ISMIR, 2003.

Audio signal (PCM) / Sonogram (23 critical-bands with the unit bark) / MFCC (first 20 coefficients without the first, coding 36 Mel-scaled frequency bands)

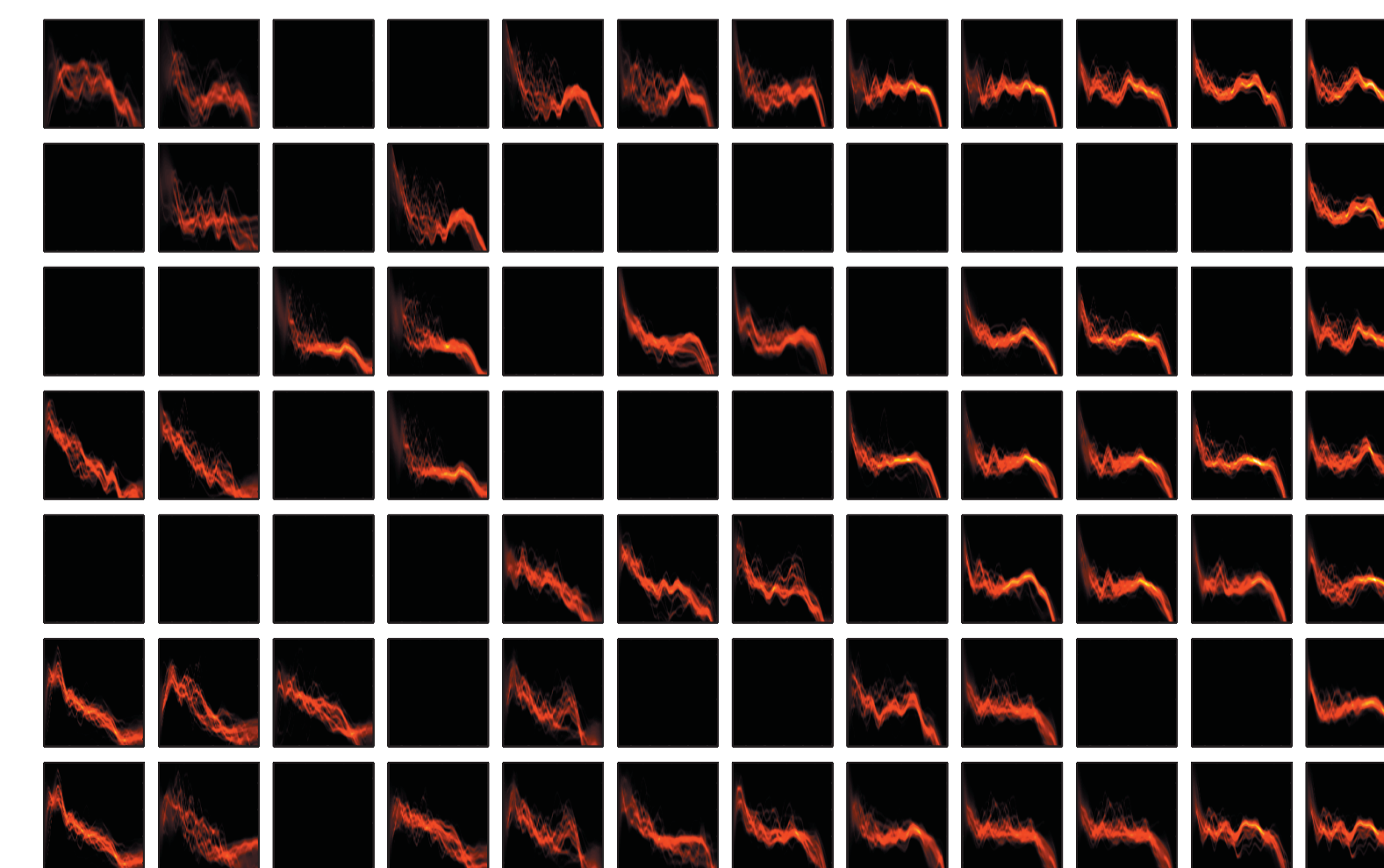
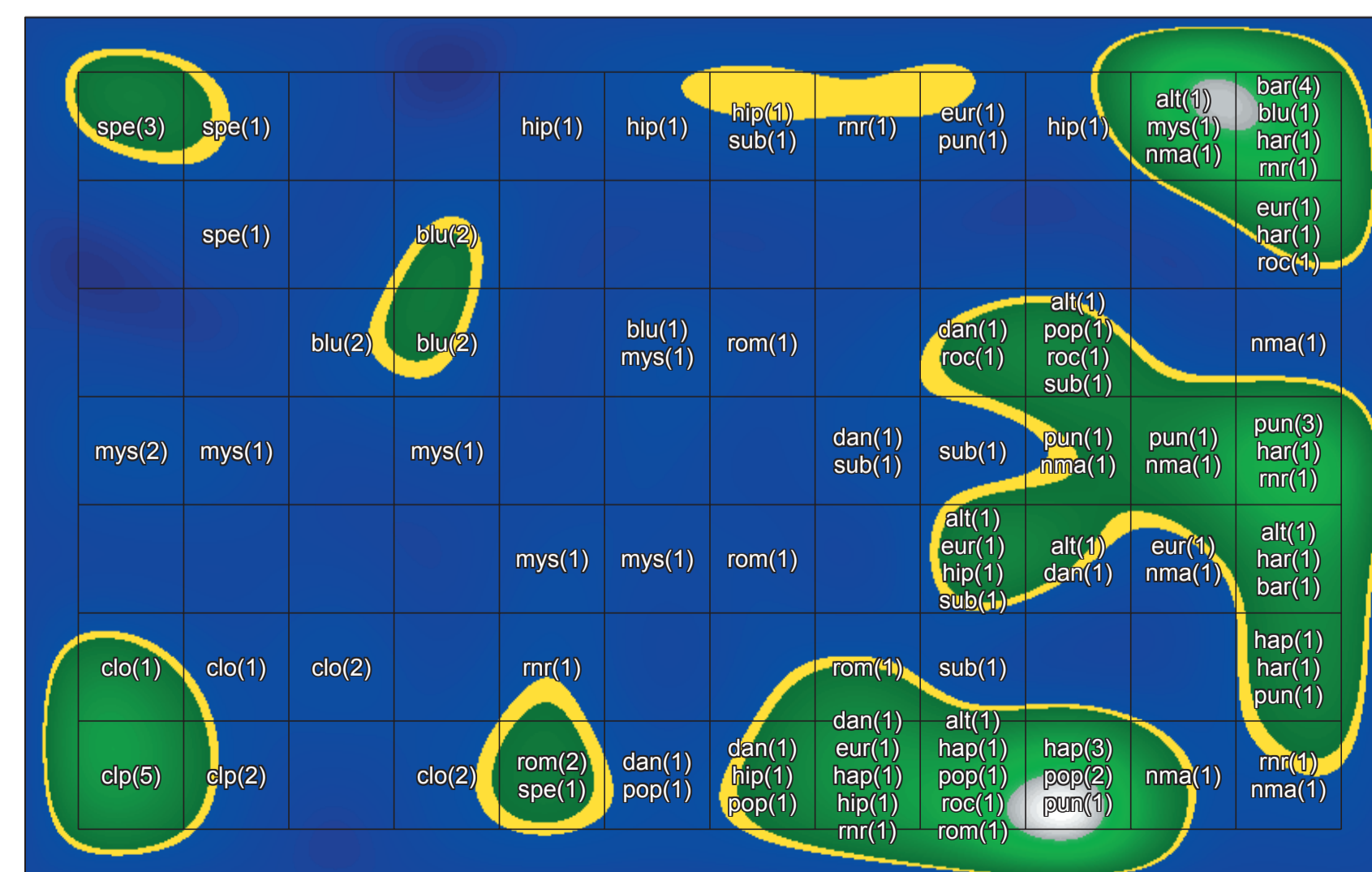


Simple Evaluation Results / Confusion Matrices

	AP 30 [2,3]	LS 30 [1]	SH	PH	FP
Sone			6 min	6 min	6 min
MFCC 20 -1	7 min	7 min			
Features O(N)	132 min	15 min	48 sec	12 min	23 sec
Distances O(N ²)	74 min	4 min	<1 sec	<1 sec	<1 sec
Total	213 min	26 min	7 min	18 min	6.5 min
R-Precision	0.43	0.38	0.21	0.13	0.25
k-NN (k=1)	0.58	0.50	0.41	0.19	0.37



Islands of Music (12 x 7 SOM, trained on collection with 118 pieces from 19 categories) / Codebook (FC/CMS 30, GMM+MC aka AP 30 [2,3])



alt ... Alternative
 blu .. Blues
 clo .. Classic Orchestra
 clp .. Classic Piano
 dan .. Dance
 eur .. Eurodance
 hap .. Happy Sound
 har .. Hard Pop
 hip .. Hip Hop
 mys .. Mystera
 pop .. Pop
 pun .. Punk Rock
 roc .. Rock
 mr ... Rocker
 rom .. Romantic Dinner
 spe .. Talk
 bar .. Bad Religion
 nma New Model Army
 sub .. Sublime