



Creation and Exploration of Musical Information Spaces

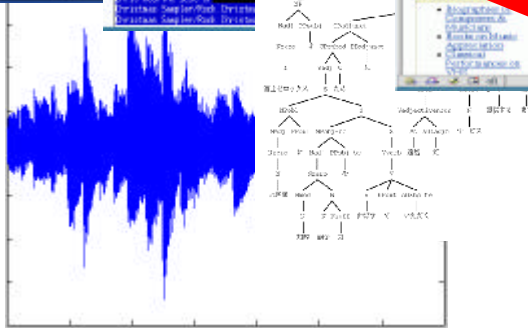
Andreas Rauber

Vienna University of Technology

<http://www.ifs.tuwien.ac.at/~andi>

From Directories to Music Spaces

The top left shows a music directory interface with a list of tracks and a tree diagram. The top right shows an Amazon.com search result for 'Classical Instrumental', listing various albums and artists.



Outline

- Motivation
- Feature Extraction
- Organization: SOM&Co
- Demonstration and Evaluation

Motivation

- Music omnipresent
- Large collections - on small devices
- Increasingly distributed electronically

but

- Difficult to search for
 - textual
 - query by humming

Motivation

- Automatic genre/style - based organization
- Navigational interface
- Playlist generation
- Discovering unknown titles/artists

Challenge:

- How to compute similarity of music?

Outline

- Motivation
- **Feature Extraction**
- Organization: SOM&Co
- Demonstration and Evaluation

Feature Extraction - Overview

Preprocessing

Feature Extraction

Specific Loudness Sensation

Rhythm Patterns

→ raw audio data

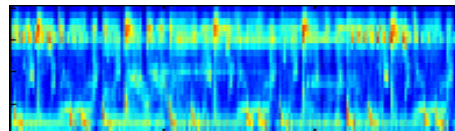
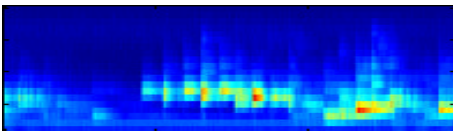
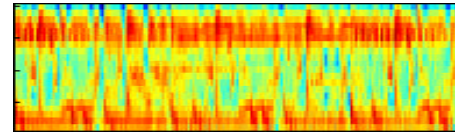
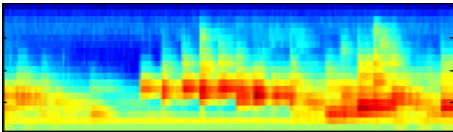
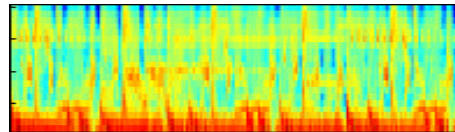
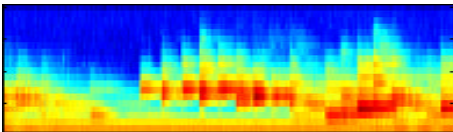
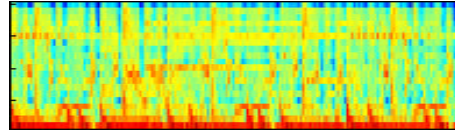
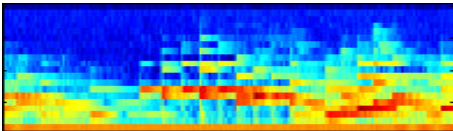
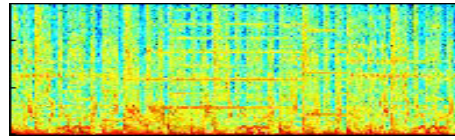
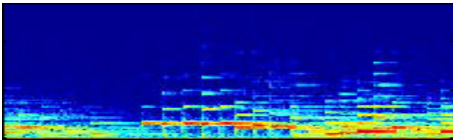
- downsampling: 44kHz to 11kHz
- stereo to mono
- cut into 6sec segments
- remove lead-in and fade out
- keep every 3rd segment

← PCM data segments

Feature Extraction - Loudness

Elise

Freak on a Leash

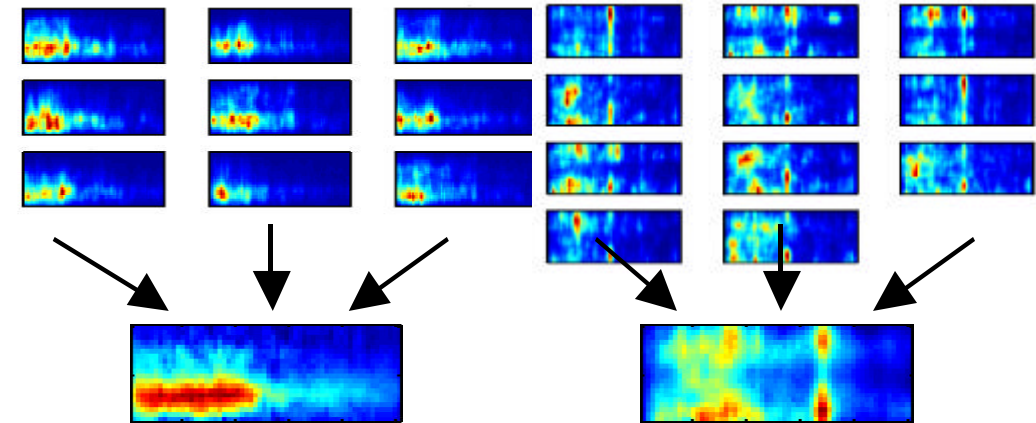
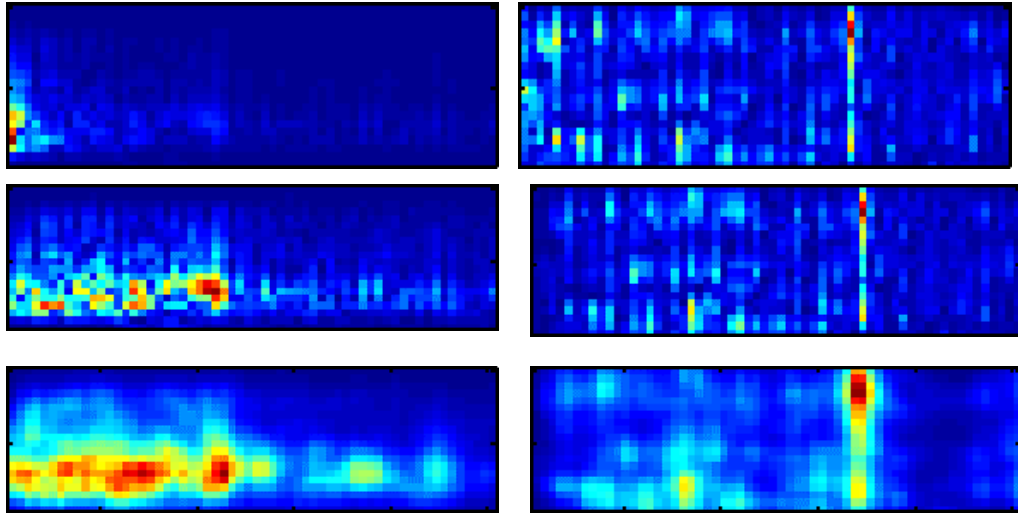


- PCM Audio Signal
- Power Spectrum
- Frequency Bands
- Masking Effects
- Phon
- Sone

Feature Extraction - Rhythm

Elise

Freak on a Leash



- Loudness Modulation Amplitude
- Fluctuation Strength
- Filter (Gradient, Gauss)
- Median
- 1200-dim feature vec.

Outline

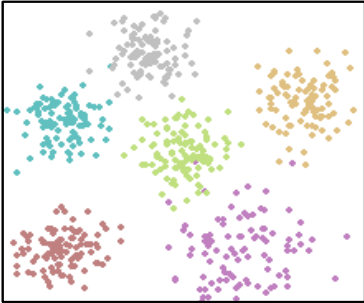
- Motivation
- Feature Extraction
- Organization: SOM&Co
- Demonstration and Evaluation

SOM Basics

- Self-Organizing Map, Kohonen Map
- determines mapping from high-dimensional input space to 2-dim output space ("map")
- such that neighborhood relationships in data are preserved
- "spatially smooth k-means"

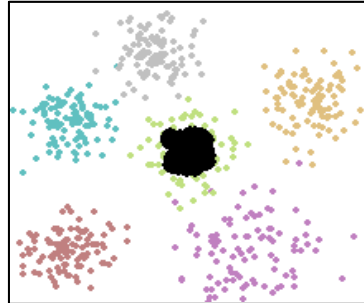
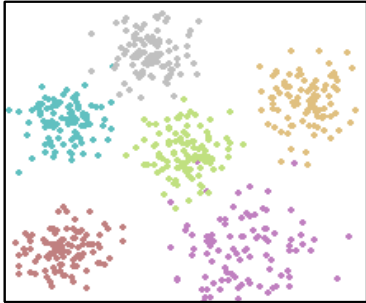
Self-Organizing Map

Self-Organizing Map (SOM)



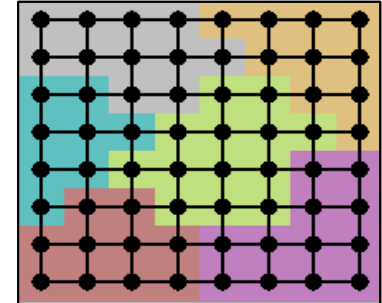
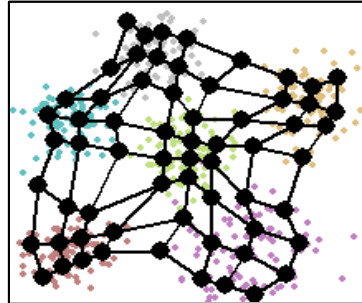
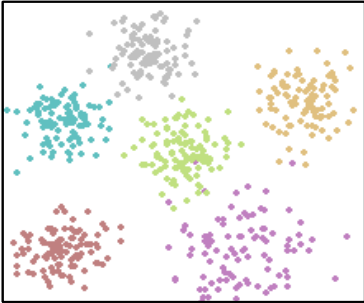
Self-Organizing Map

Self-Organizing Map (SOM)



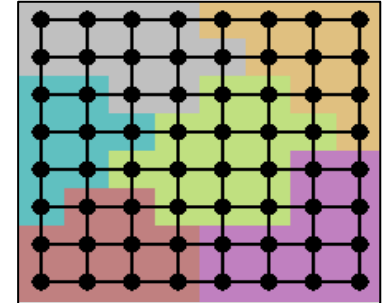
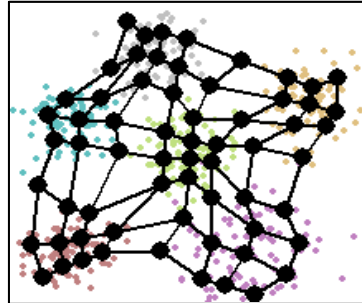
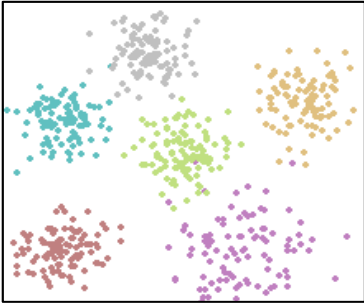
Self-Organizing Map

Self-Organizing Map (SOM)

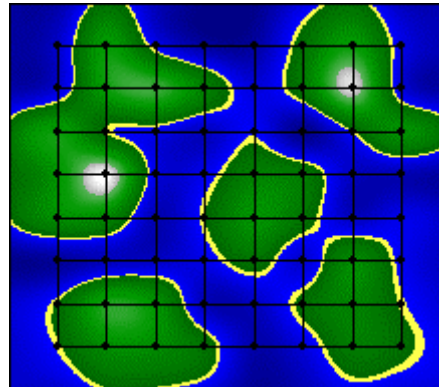


Self-Organizing Map

Self-Organizing Map (SOM)



Smoothed Data Histograms (SDH)

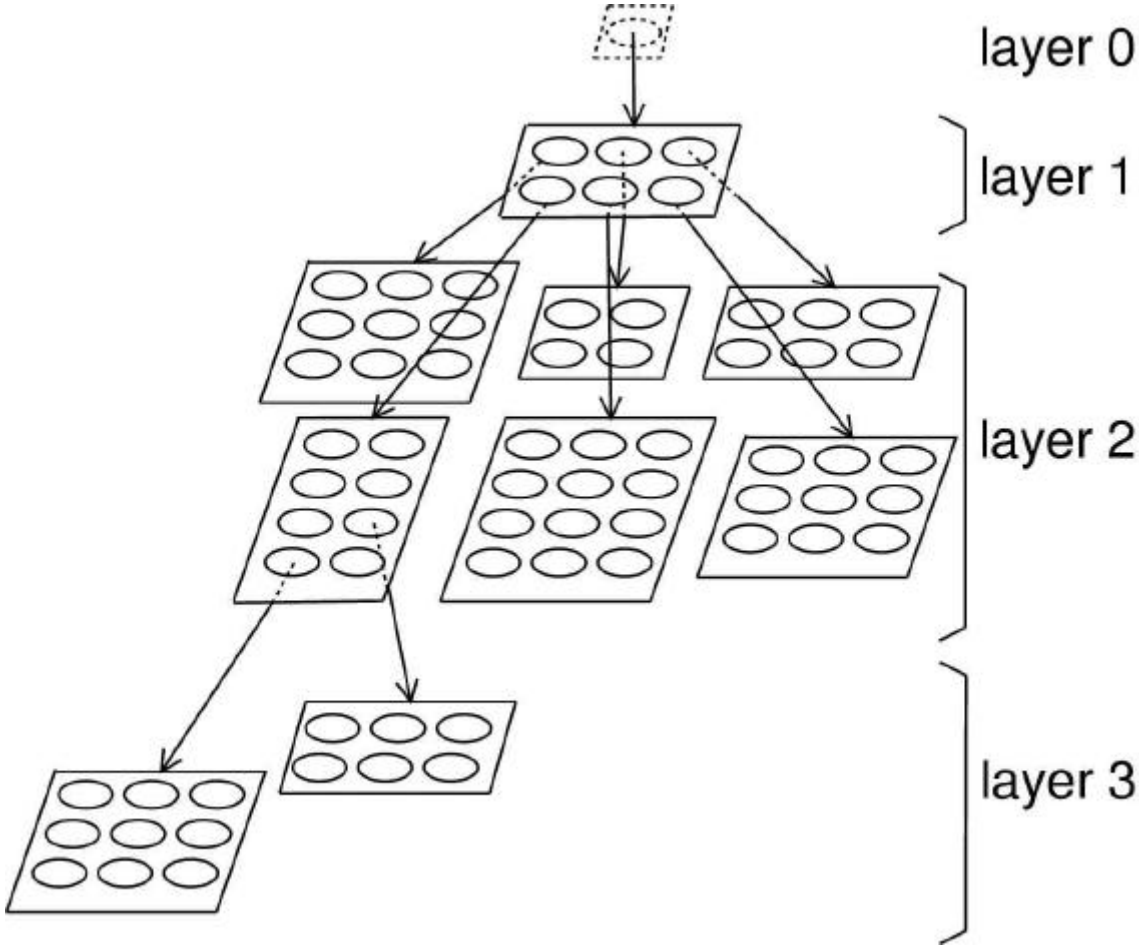


GHSOM:

Growing Hierarchical SOM

- based on SOM model (GG & HierSOM)
- dynamic hierarchical growth
(divisive alg., top-down refinement)
- dynamic horizontal growth
(granularity gain per layer)
- unbalanced structure according to
data requirements

GHSOM Architecture



Outline

- Motivation
- Feature Extraction
- Organization: SOM&Co
- Demonstration and Evaluation

Experiments - Coll359

- 359 pieces of music (~24 hrs)
- variety of different genres

- 2x4 top-layer map
- all units expanded as layer 2 maps
- 25 out of 64 units expanded on layer 3



Experiments - Coll359



<p>QE: 163.839 beethoven.mp3 fuguedminor.mp3 lovemetender.mp3 sml-icecream.mp3 therose.mp3 vm-brahms.mp3</p>	<p>QE: 89.6064 elvira.mp3 schwan.mp3 stormsinafrica.mp3 tell.mp3</p>	<p>QE: 70.602 jurassicpark.mp3 leavingport.mp3 merry.mp3 pachelbl.mp3</p>
<p>QE: 162.628 branden.mp3 vm-bach.mp3 zarathustra.mp3</p>	<p>QE: 88.6057 forelle.mp3 minuet.mp3 schindler.mp3 walzer.mp3 zapfenstreich.mp3</p>	<p>QE: 112.733 adagio.mp3 air.mp3 avemaria.mp3 elise.mp3 flute.mp3 fortuna.mp3 funeral.mp3 kidscene.mp3 mond.mp3 mountainking.mp3 nachtmusik.mp3</p>

<p>QE: 1800.03 bfmc-freestyler.mp3 bongobong.mp3</p>	<p>QE: 630.134 sl-summertime.mp3</p>	<p>QE: 3094.98 bfmc-rock.mp3 cocojambo.mp3</p>
<p>QE: 881.358 bfmc-fashion.mp3 rhcp-world.mp3</p>	<p>QE: 2370.83 bfmc-others.mp3 limp-n2gether.mp3 rhcp-scartissue.mp3</p>	<p>QE: 1618 bfmc-1234.mp3 kiss.mp3</p>
<p>QE: 1573.18 singalongsong.mp3</p>	<p>QE: 418.188 seeyouwhen.mp3</p>	<p>QE: 4493.34 rhcp-dirt.mp3 rhcp-getontop.mp3</p>

Experiments - Coll359



Tocatta and Fuge in D Minor



Vanessa Mae:

Scherzo in D Minor

Partita #3 in E for Solo Violin

Tequila Mockingbird

Tocata and Fuge in D Minor

The 4 Seasons

Red Violin

Classical Gas



Experiments - Dance Sport

- 1129 pieces of dance music (~56 hrs)
- 10 dances (IDSF):

LATIN

Samba

Cha-Cha-Cha

Rumba

Paso Doble

Jive

BALLROOM

Slow Waltz

Tango

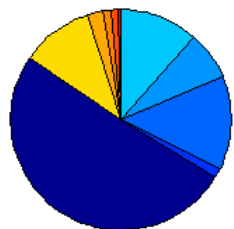
Viennese Waltz

Slow Foxtrott

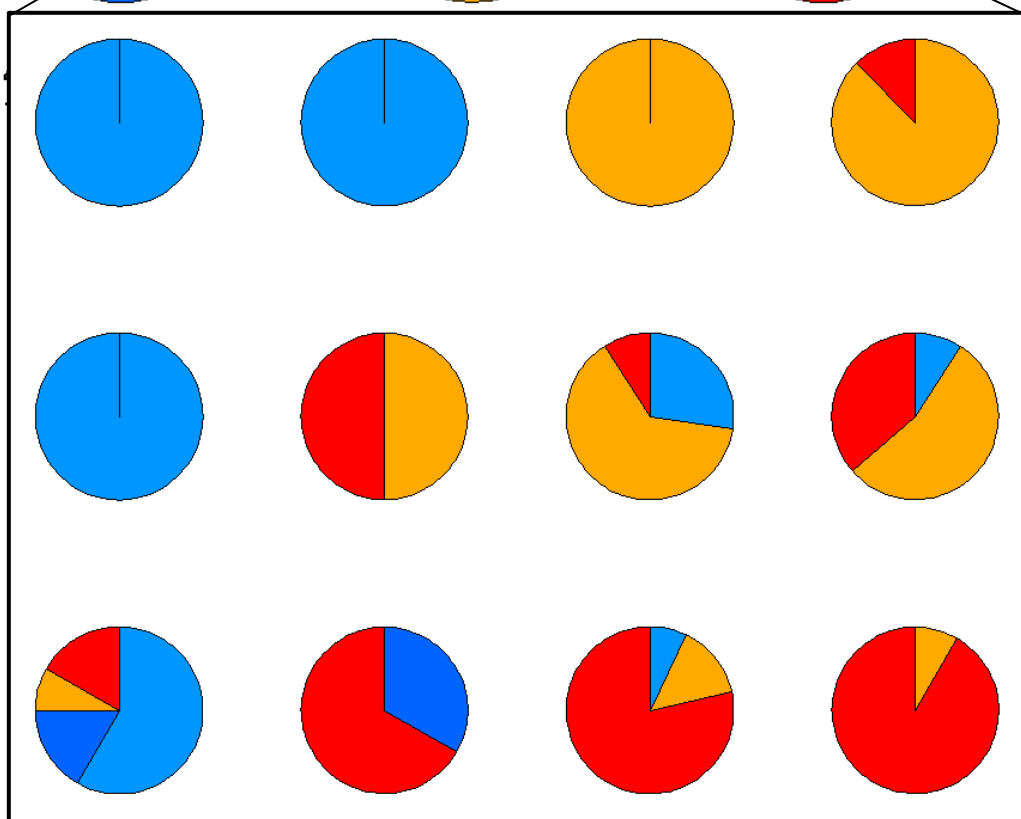
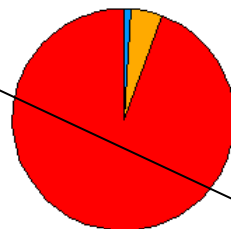
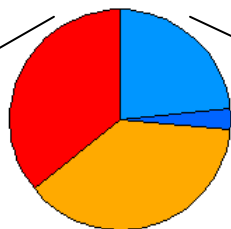
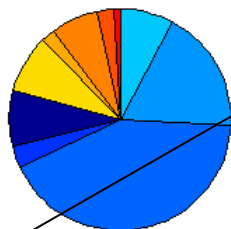
Quickstep

- GHSOM top 4x2, all expanded on layer 2

Experiments - Dance Sport



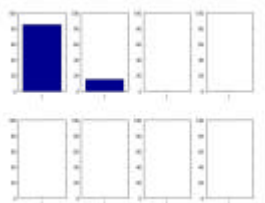
51% Waltz
13% Slow Foxtrott
11% Viennese Waltz



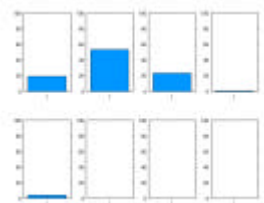
32% Quickstep
22% Rumba
14% Slow Foxtrott
12% Samba



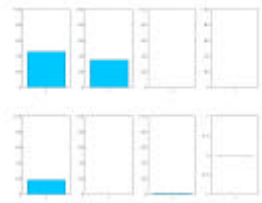
Experiments - Dance Sport



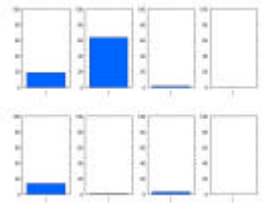
Slow Waltz



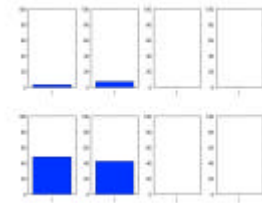
Slow Foxtrott



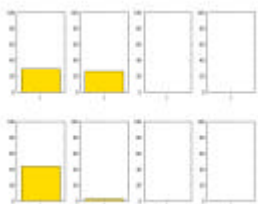
Viennese Waltz



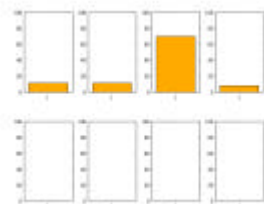
Tango



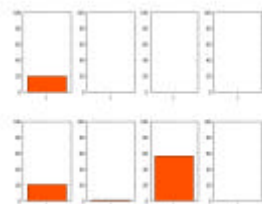
Quickstep



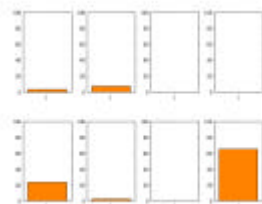
Rumba



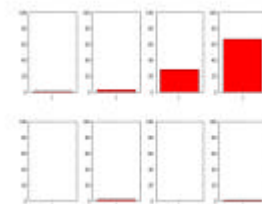
Paso Doble



Jive



Samba



Cha-Cha-Cha



JI37



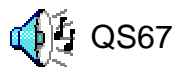
JI63



SA44



SA01



QS67



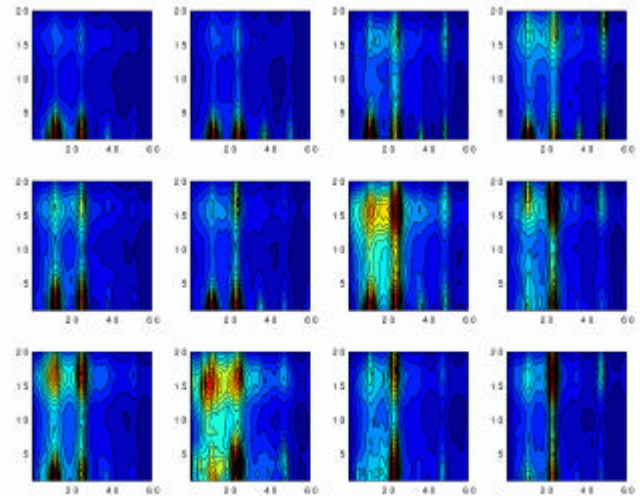
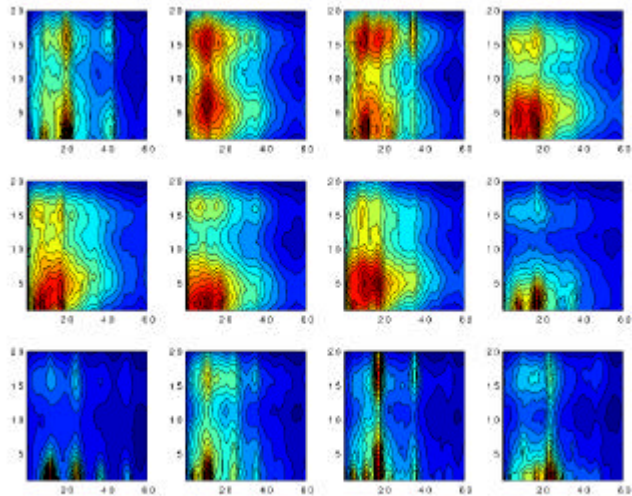
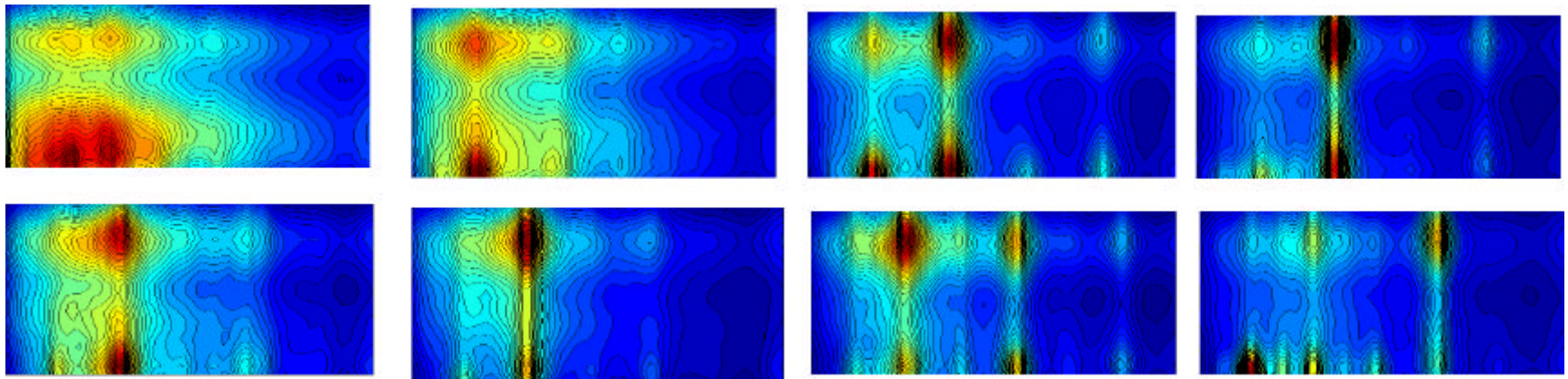
RU61

Conclusions

- Organization of music by sound similarity
- Content-based access and retrieval
- SOMeJB prototype available at

<http://www.ifs.tuwien.ac.at/~andi/somejb>

Experiments - Dance Sport



Experiments - Dance Sport

knn, k=5:

CC	JI	LW	PD	QS	RU	SA	SF	TG	WW
110	1	0	8	1	0	0	2	2	0
0	93	2	0	1	4	0	0	1	2
0	0	144	0	1	1	0	5	0	6
6	0	0	36	0	0	0	2	9	0
0	0	3	0	113	8	0	2	0	0
0	0	10	0	5	80	0	2	0	0
0	0	0	0	4	3	92	0	0	1
0	0	11	0	0	0	0	180	0	4
1	0	1	0	1	0	0	7	89	0
0	1	13	0	2	1	0	9	0	40

Experiments - Others

On automatic genre-based evaluation:

■ Classic-Cluster:

- 14 classical orchestra pieces

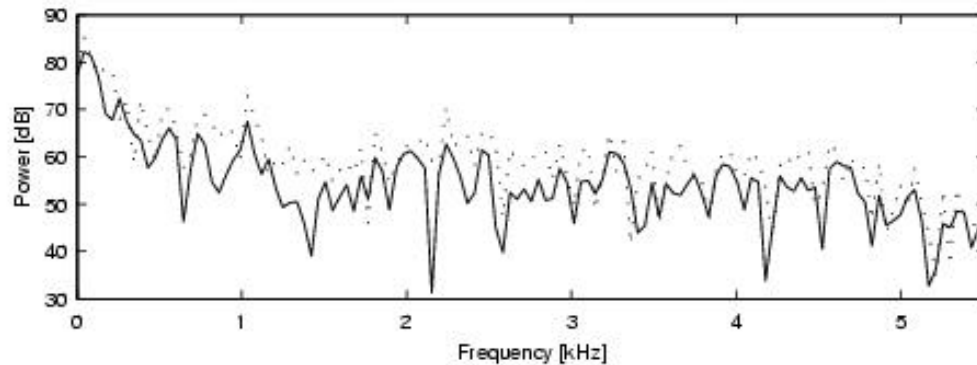
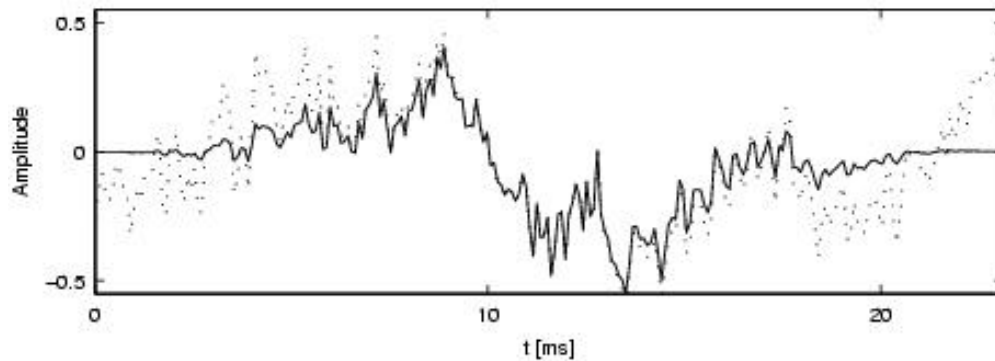
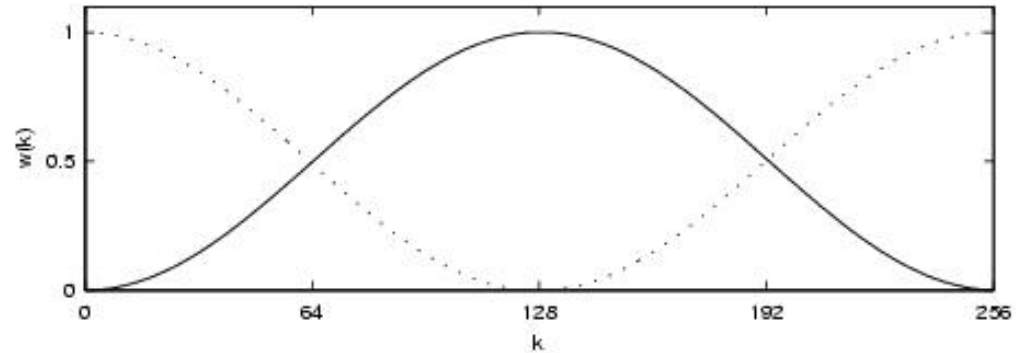
- plus: Heavy Metal:  Metallica: The Extasy of Gold

■ Jazz-Cluster:

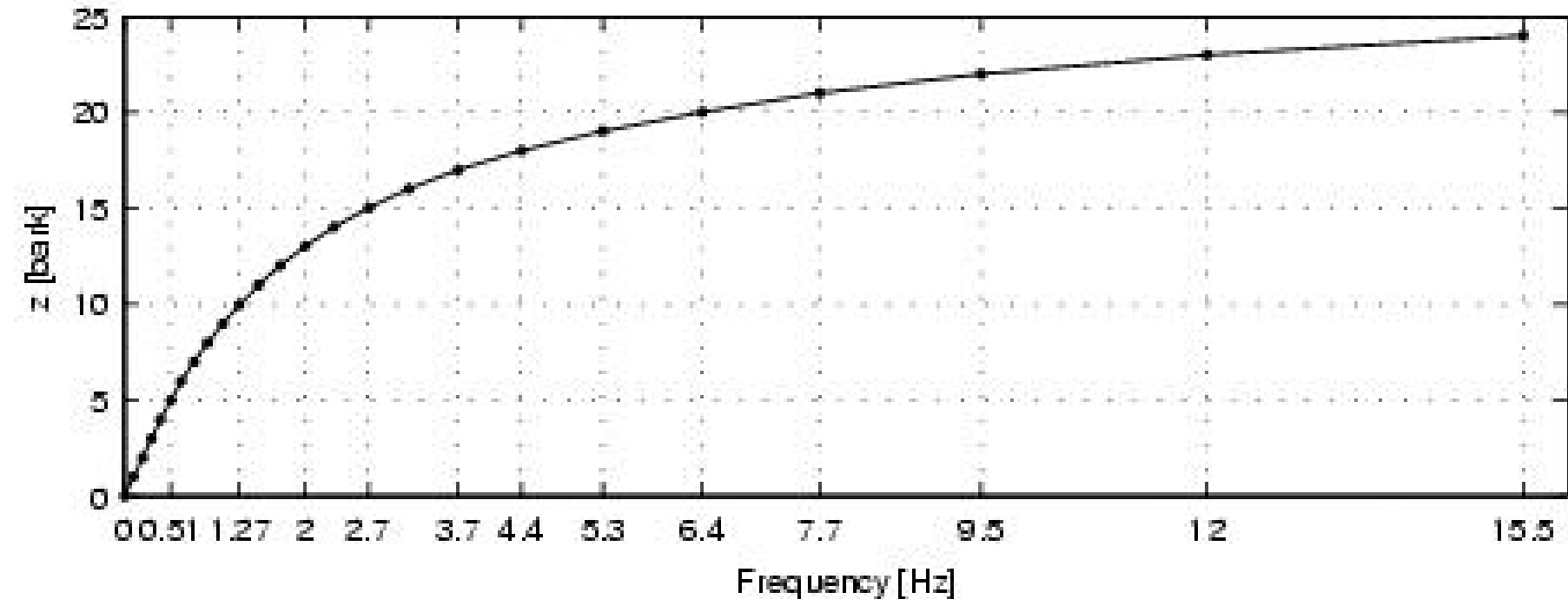
- 8 Jazz titles

- plus: Text/Speech:  Dorfer: Freispiel

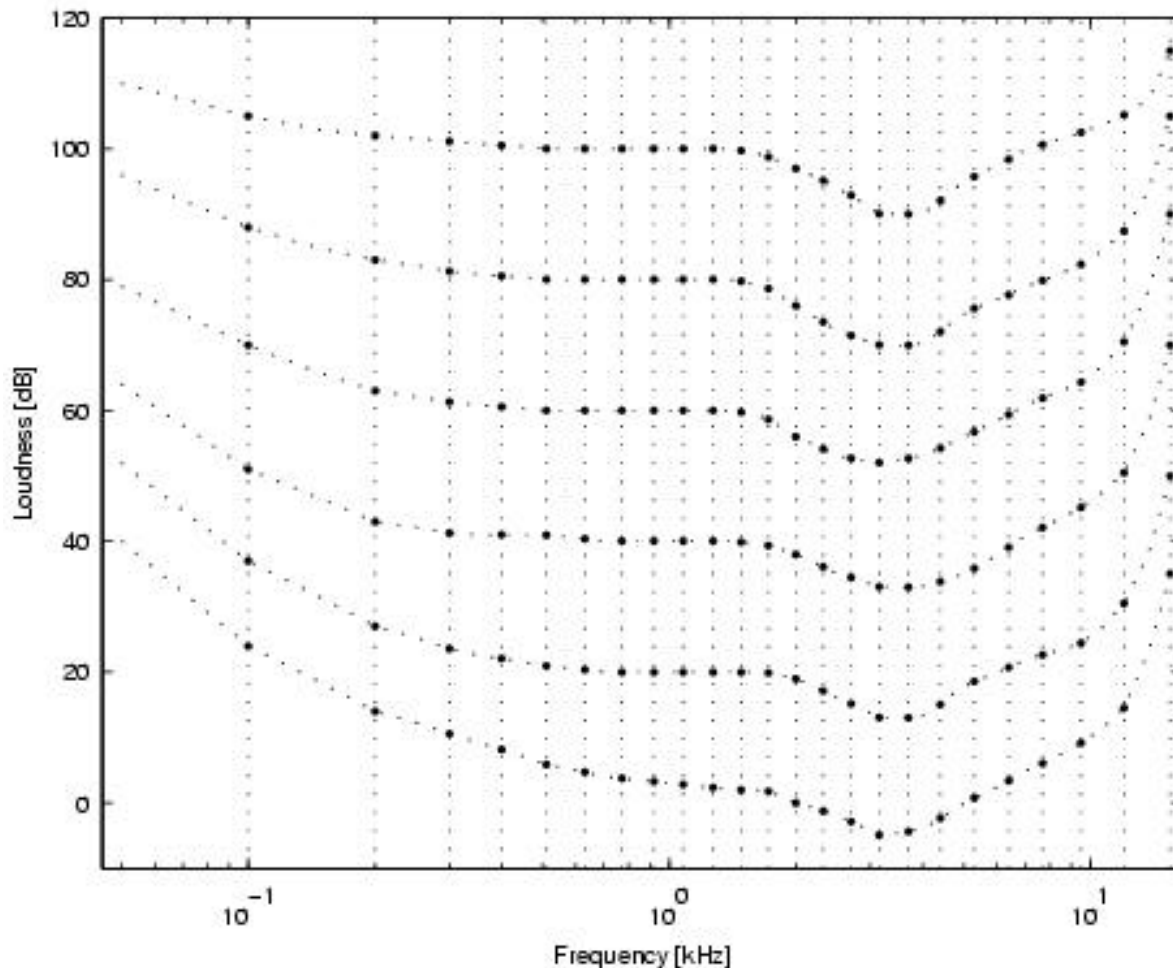
Feature Extraction - Hanning



Feature Extraction - Bark Scale

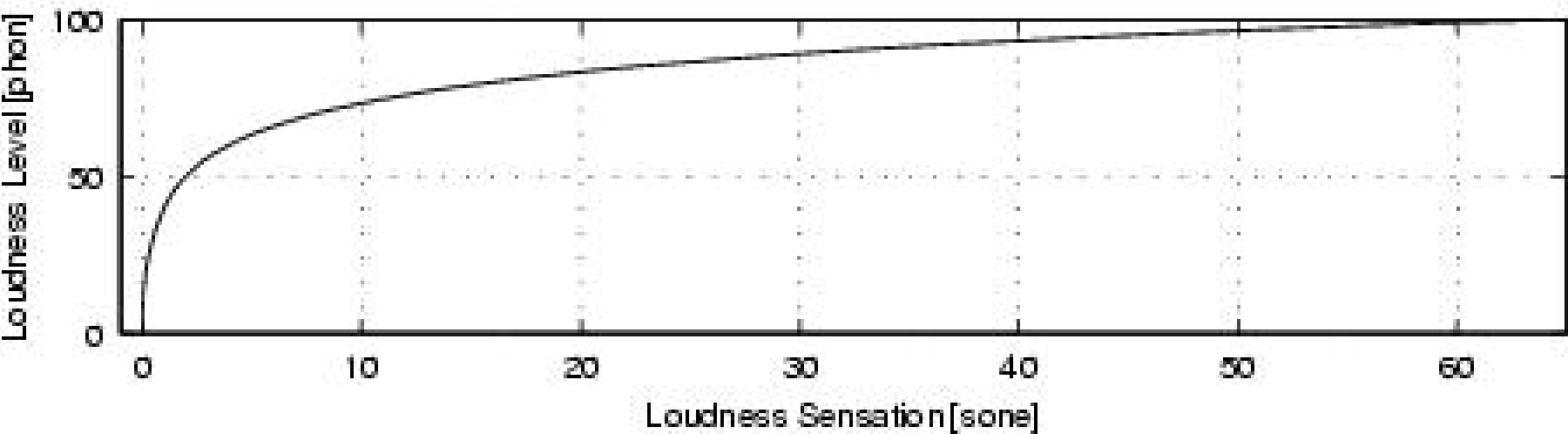


Feature Extraction - Phon

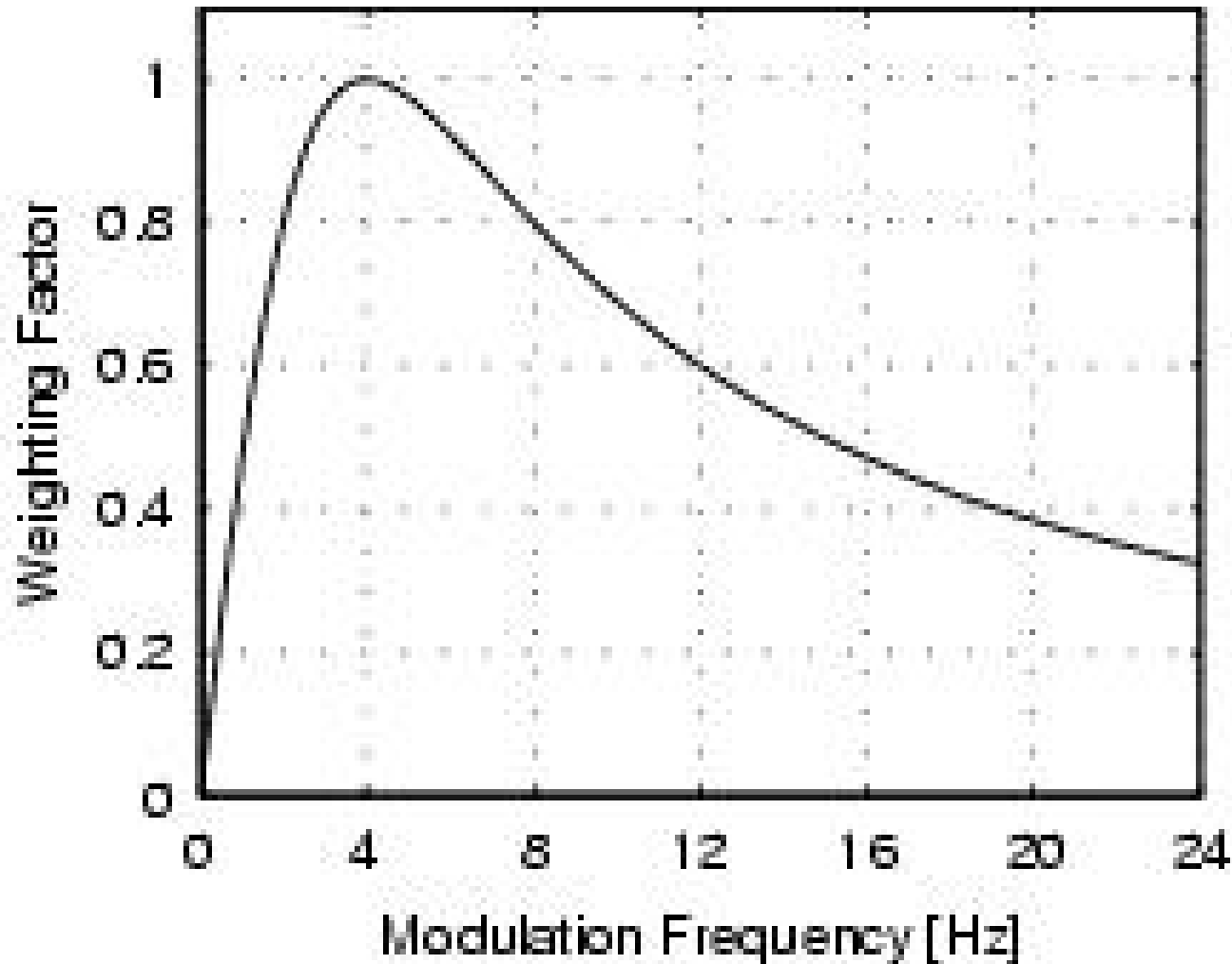


equal loudness contours for 3, 20, 40, 60, 80, 100 phon

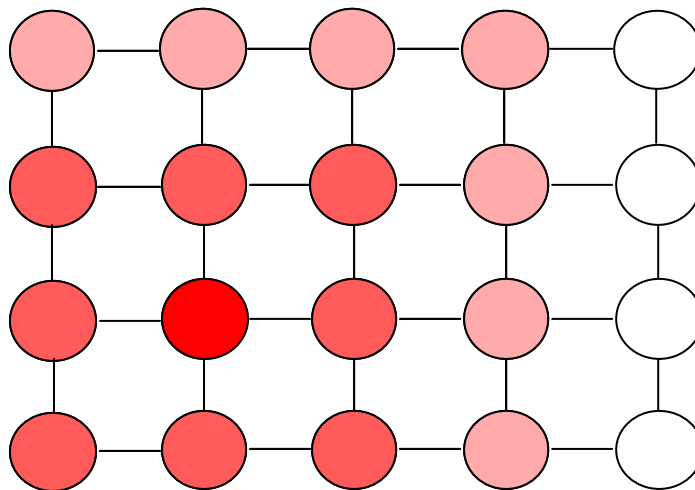
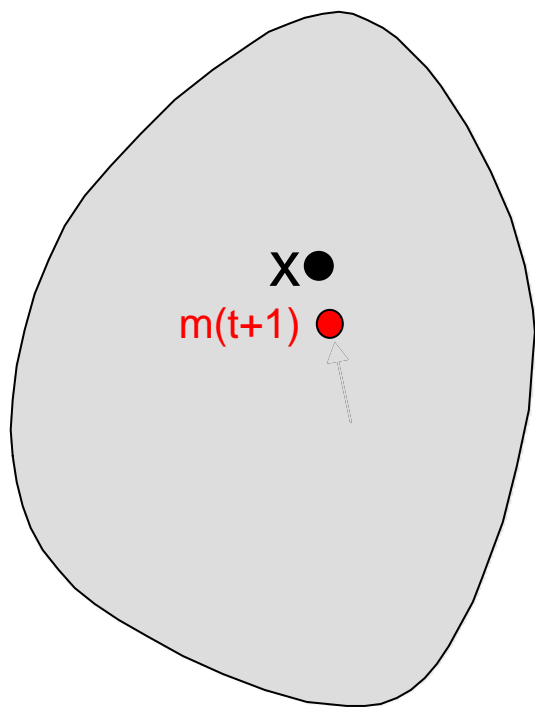
Feature Extraction - Sone



Feature Extraction - Fluctuation



SOM Training



GHSOM Demonstration

