Creation and Exploration of Musical Information Spaces

> Andreas Rauber Vienna University of Technology http://www.ifs.tuwien.ac.at/~andi

ICDL 2004, New Delhi

From Directories to Music Spaces



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

had a child to for a find a shift to have a the lock of a state of the second particular sound of the second of th

- 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, Gaus

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 (SOM)



Self-Organizing Map (SOM)





Self-Organizing Map (SOM)







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

77 pieces of music (~5 hrs) variety of different genres



bine-specing bi		irro-inchenso Irro-sodsing Irro-cliqiimi	scoprito inpunization			corgs siled 5 mindle/s Borntew		
hurungsh	•	bangabang	naze e adcij			kveini tihani t	grannek mania menday 1970 supertrouper	
si-summetime		ro-insulyikr	sam Imp-rokodj imp		Imp-polution	dawinggues		
rhep -cuill ornio rhep-work sti-what go	salikari di N	nauto mb	ga-doadai ga-kennak Q3-jipton mma-bigbilae	ga-norpeach	kom-linuk. pt-dantasi pt-swenge	Bootwaryoung Rootwaryoung		
ailiomidean Mingun dominyteat	missarhing	irend peskerdap-b	oenalia teo ing	n d	ummetoy Insurdaon Isono	lakana kavensile rober kite some	been towen flugued minor wm-blach wm-blachms	
bigworld	ang elir	rangicik ami-acia	-		Terroy	2004	ar averata	
nddion (13-lie	la americanpie le Cristierran		Mediad	110	erinectedy	aantaan	Nidsoene mond	

- 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

File Edit View Ga Coamanicatar								-	
and Back	Forum	Relat	N	At Home	Search	Netscape	State Print	See 1	
1 800	ARRES .	§ 60	TH: NEW	1.77999	ifs toxi	en ac at/-	7 🚺 WT	afi Arisk	
Re M									
	-	1	-	1					
	iller.	B	N.					8	
	-								
	trin.	iiic.	W.,			112. <u>2</u>	£ (£		
		-	-		10			*	
	0000	E.	崖						
				1	-		-		
					100	. 12	THE .	- II	
	1	-	Sillin.	-	1 📟		B- 1	695.	
Same .	and an	-	192	28	500	- E.	time 1	-	
-	-	tin-	Her-	12	臣	- 19	are a		
-		E SHE	386	Anet	1	: 12	1	-	
-					iller.	~ 199	Sec.		
-	1	-	-	. 1	-		-		
-	-		x 1			100 x 100	- 24		
2.04		-				822. SE	ter Wite		
1.00	5. S			-		100 M	ä. 1888		
	97 H	10 III	-	5r		100er	(Willess		
	1	-		1			apure	1	
	Balan -	535	100				E. Efforts		
	P.C.R.	MAL	100			Anna Art	tion Street	- 1	
1 0	1		- and -		1 1			- 1	

The Back Communication Back Forward Related Hone Searce Helicapo Piret Se	QE: 163.839 beethoven.mp3 fuguedminor.mp3 lovemetender.mp3 sml-icecream.mp3 therose.mp3 vm-brahms.mp3	QE: 89.6064 elvira.mp3 schwan.mp3 stormsinafrica. tell.mp3	QE: 70.602 jurassicpan leavingpor merry.mp2 pachelbl.m	2 r <u>k.mp3</u> r <u>t.mp3</u> <u>3</u> 1 <u>p3</u>
	QE: 162.628 branden.mp3 vm-bach.mp3 zarathustra.mp3	QE: 88.6057 forelle.mp3 minuet.mp3 schindler.mp3 walzer.mp3 zapfenstreich.n	QE: 112.7 adagio.mp air.mp3 avemaria.n elise.mp3 flute.mp3 fortuna.mj funeral.mj kidscene.n mond.mp3 mountainl nachtmusi	33 <u>3</u> <u>mp3</u> <u>p3</u> <u>p3</u> <u>p3</u> <u>p3</u> <u>king.mp3</u> <u>k.mp3</u>
	QE: 1800.03 <u>bfmc-freestyler.r</u> <u>bongobong.mp3</u>	np3 QE: 630	.134 nertime.mp3	QE: 3094.98 bfmc-rock.mp3 cocojambo.mp3
State State State State State State State State State State State State State State State State	QE: 881.358 bfmc_fashion.mp rhcp_world.mp3	QE: 237 bfmc-on limp-n2 rhcp-sc	0.83 thers.mp3 gether.mp3 artissue.mp3	QE: 1618 <u>bfmc–1234.mp3</u> <u>kiss.mp3</u>
	QE: 1573.18 <u>singalongsong.m</u> J	p3 QE: 418.188 QE: 44 seeyouwhen.mp3 PE: 44 rhcp-c		QE: 4493.34 rhcp-dirt.mp3 rhcp-getontop.mp3



Toccata 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











Slow Waltz









Rumba



Paso Doble

-



Jive



🐠 JI63









(] [[RU61



Cha-Cha-Cha

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:
 - B Jazz titles
 - plus: Text/Speech: 4 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

