

Masterstudium: Intelligente Systeme Diplomarbeitspräsentationen der Fakultät für Informatik

Tool Support for Acoustic Evaluation of Music Similarity



Andreas Senfter andreas@senfters.com

Technische Universität Wien Institut für Softwaretechnik und interaktive Systeme Information & Software Engineering Group Betreuer: ao.univ.Prof. Dr.techn. Andreas Rauber, Dipl.Ing. Thomas Lidy

Abstract

The evaluation of the quality of similarity measures between pieces of music is a non-trivial, sophisticated task because human cognition of music and perception of similarity inherently is biased by subjective interpretation and reasoning. Therefore, the evaluation of similarity between music tracks typically requires human judgment.

This work focuses on the acoustic evaluation of music similarity. It presents a novel tool called cLynx which offers several different features for an acoustic evaluation of music similarity and additionally enables exploring and analyzing structured audio repositories much faster and more efficiently. An evaluation based on a listening test studies the applicability of cLynx in various listening situations. It analyses connections between subjective interpretation and individual preferences and the subjects' perception of music similarity and susceptibility to outlier detection.

Contributions

- Development of a novel tool, namely cLynx, for:
 - acoustic validation of similarity between pieces of music
 - analyzing, browsing and exploring music data of structured audio collections
- Listening test based on cLynx tool:
 - study the human cognition of music and subjective interpretation of music similarity
 - analyze the performance of cLynx's various features

• Implementation:

- Clynx plugged in a Music Similarity System called PlaySOM
- required input passed by an underlying SOM

cLynx: A Novel Form of Music Similarity Evaluation

	🛗 cLynx Control	- C Z	1
Settingcontrol:	Multiplexer \ Normalizer \	Console: 2	Console Log Level
	Multiplexer: Parallel Mixer 🗸	loading: rock_pop_artist_106_album_1_track_1.mp3	
Multiplexerselection Analyserselection		analysing rock_pop_artist_106_album_1_track_1.mp3	
Normalizerselection			
Normalizerselection	Duration (sec): 3,5	Frock_pop_artist_106_album_1_track_1.mp3	
	Fade time (sec): 0,2		Console
		MINMAX.max = 32767	
	Volume (%): 180		
	Master Volume (%): 40 🖸	mean.ms.up = 10116.738858774403 mean.ms.down = 10116.738858774403	
	10	mean.geo.up = 8134.516638686313	
		mean.geo.down = -8009.088373627014	
Ch / 115-1-		mean.geo = -27.833925635818588	
Show / Hide		std = 10116.702480078638	
Console		loading: metal_punk_9-watergrave.mp3	
Playercontrol:		analysing metal_punk_9-watergrave.mp3	
Play Stop Continue		metal_punk_9-watergrave.mp3	
	rock_pop_artist_106_album_1_track_1.mp3	BOUND_TOP = 32767	
	metal_punk_9-watergrave.mp3	BOUND_DOWN = -32768	
	rock_pop_artist_106_album_2_track_1.mp3	MINMAX.max = 32767 MINMAX.min = -32768	
	rock_pop_6-order.mp3 classical_artist_24_album_1_track_2.mp3	mean.rms.up = 5416.063344445817	
Interactive Playlist	world_6-jig_in_the_parlor.mp3	mean.ms.down = 5416.063344445817	
	jazz_blues_artist_74_album_2_track_3.mp3	mean.geo.up = 4141.1034161289335	
	classical_artist_35_album_1_track_5.mp3	mean.geo.down = -4260.730718792375 mean.geo = 6.578604774608014	
	electronic_7-erased.mp3	std = 5416.060372936506	
	classical_artist_35_album_1_track_6.mp3		
	rock_pop_artist_97_album_1_track_3.mp3	loading: rock_pop_artist_106_album_2_track_1.mp3	
	classical_artist_1_album_1_track_1.mp3	analysing rock_pop_artist_106_album_2_track_1.mp3	
	rock_pop_14-i'm_ready_for_you.mp3	rock_pop_artist_106_album_2_track_1.mp3	
	rock_pop_artist_100_album_2_track_1.mp3	BOUND_TOP = 32767	
	rock_pop_artist_100_album_2_track_1.mp3 rock_pop_6-rotten_flowers.mp3 world_artist_123_album_1_track_2.mp3	BOUND_TOP = 32767 BOUND_DOWN = -32768 MINMAX.max = 32767	

Figure 1: cLynx GUI

PlaySOM

What does it do?

- supports users in acoustic evaluations of audio-based music similarity
- measures

- helps users discovering cluster structures of audio data much faster and more efficiently
- offers users several different approaches for accessing the underlying data
- assist users in genre identification and outlier detection etc.

How does it work?

Analysis:

- processes all passed sound files entirely
- calculates a global minimum and maximum sample value for the
- normalization

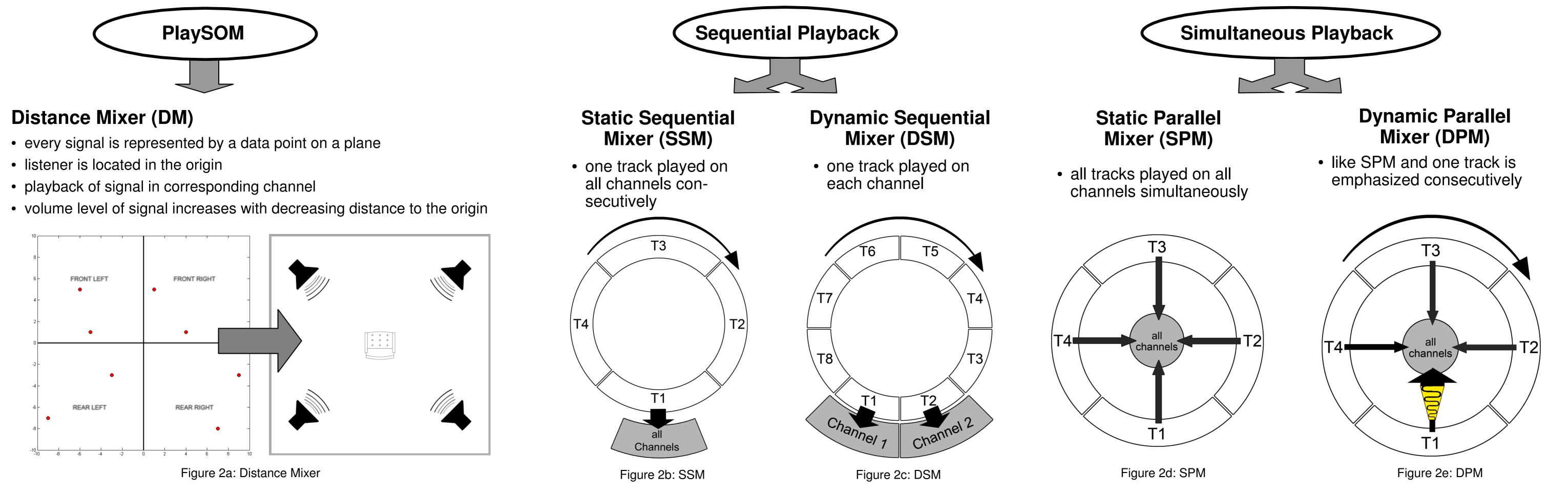
2. Normalization:

• combines the reduction of the dynamic range and the smoothing of the dynamic

- adjusts volume levels to common value for all musical pieces
- linear and dynamic approaches

3. Mixer:

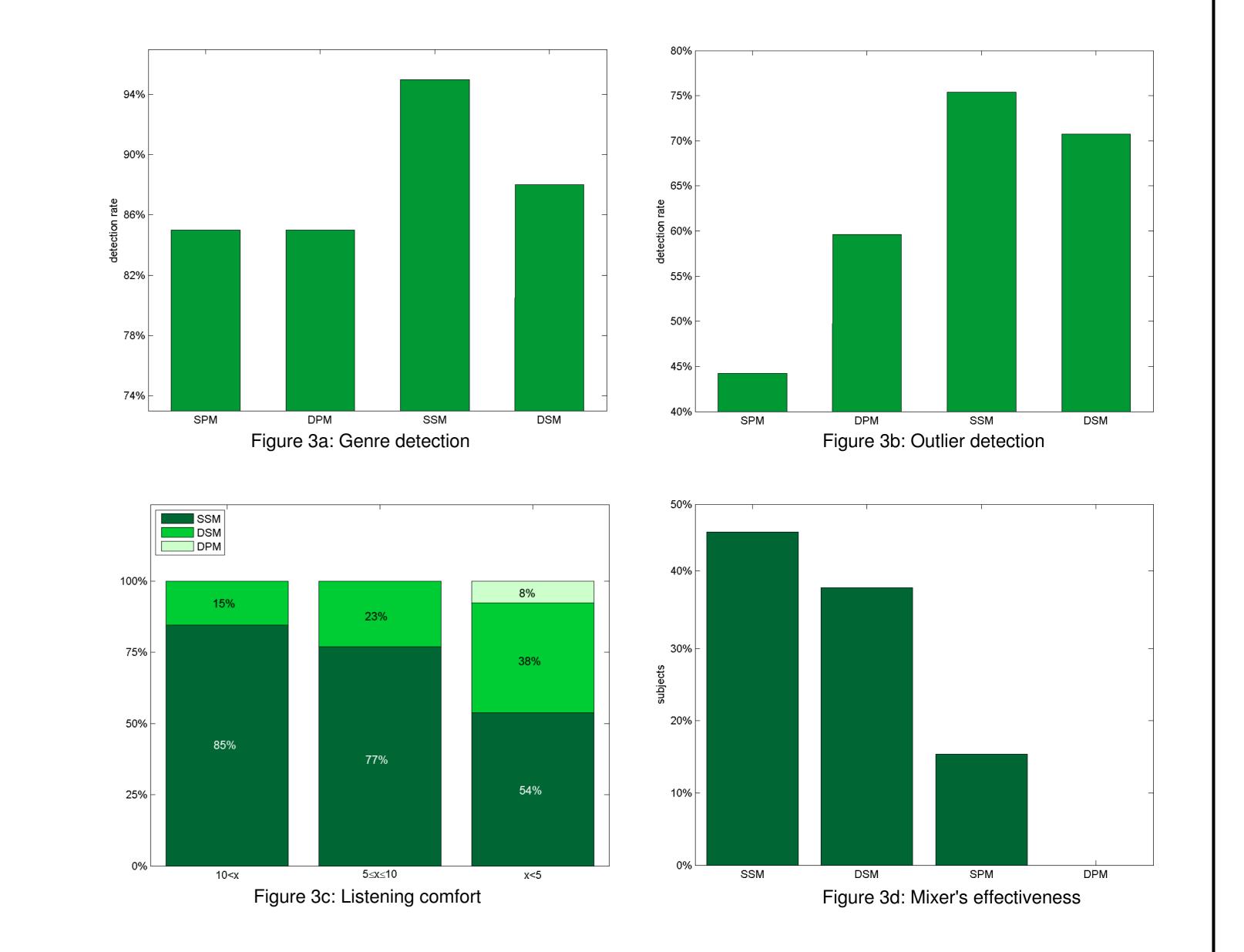
- mixes passed audio signals in different ways
- multiplexes input signals to one single output signal
- supports simultaneous and sequential playback with different emphasis
- supports playback based on spatial information



User Study

Goal and Setup

- study the applicability of the cLynx tool for music similarity judgment and outlier detection
- analyze human cognition of music and subjective interpretation of perceived audio data
- detect connections between subjective interpretation and individual preferences and the perception of music similarity
- System:
- PlaySOM application organized the ISMIRgenre test set
- playback by cLynx plug-in with different mixer settings
- 13 subjects
- 23 listening situations covering 11 different regions of a PlaySOM music map



Genre and Outlier Detection

- Figure 3(a): genre detection harder using parallel playback (SPM, DPM)
- 85% correctly detected genres using parallel playback, 95% using SSM
- more detailed identification of:
 - classical genres by subjects who enjoyed classical instrumental education • regions which contained pieces of preferred genre
- more outliers detected in areas consisting of tracks of preferred genre (stricter judgment)
- Figure 3(b): rate of outlier detection increases by switching from parallel playback to sequential playback
- DPM: outlier detection rate of 60%

cLynx Applicability

- Figure 3(c): subjects felt more comfortable with sequential playback (SSM, DSM)
- Figure 3(d): 15% rated SPM as the most effective mixer

Voice Detection

- voice identification rate irrespective of the type of used mixer component
- higher rate of voice detection than lyrics understood