

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

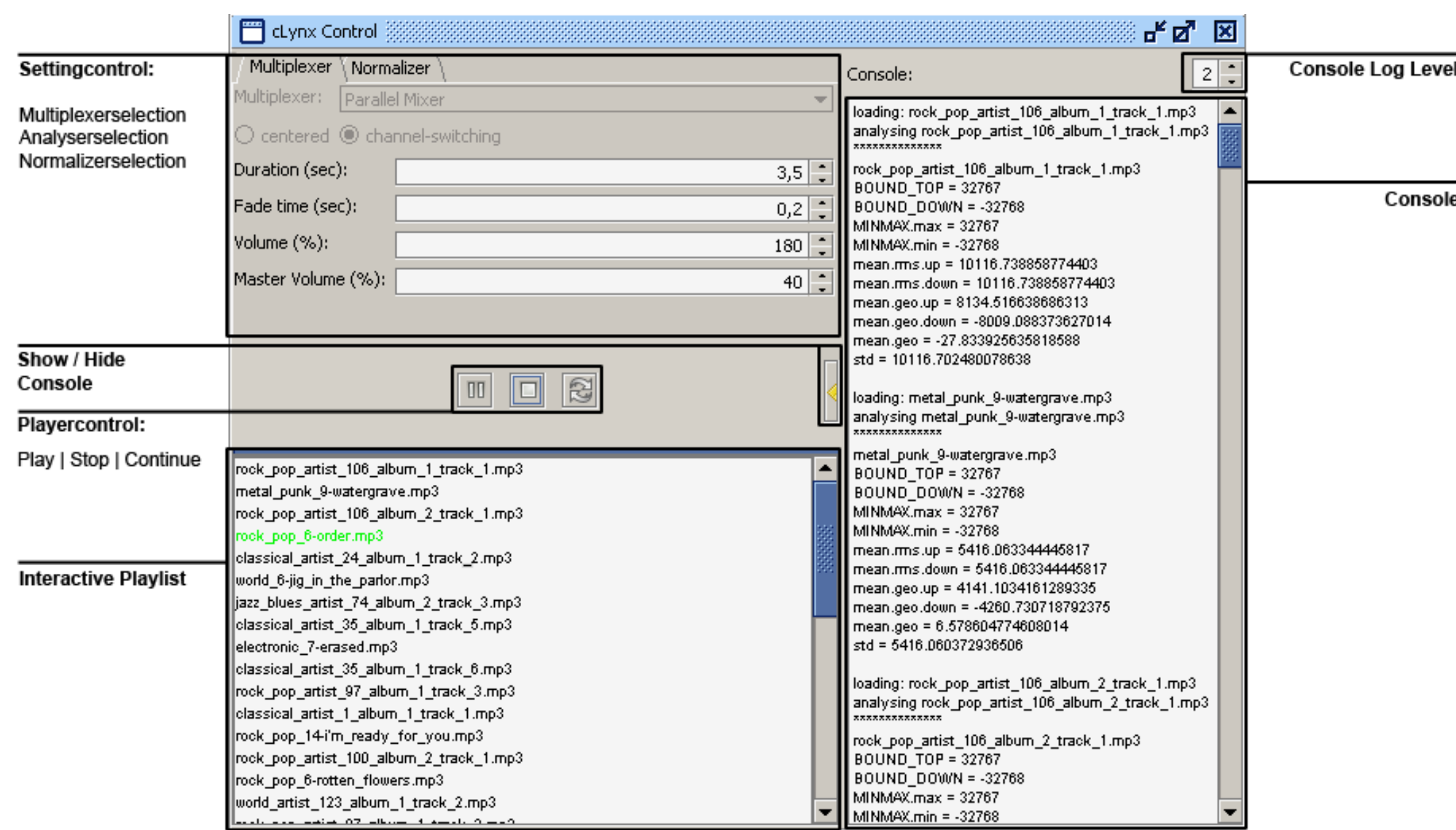


Figure 1: cLynx GUI

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
- 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
- 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

PlaySOM

Distance Mixer (DM)

- every signal is represented by a data point on a plane
- listener is located in the origin
- playback of signal in corresponding channel
- volume level of signal increases with decreasing distance to the origin

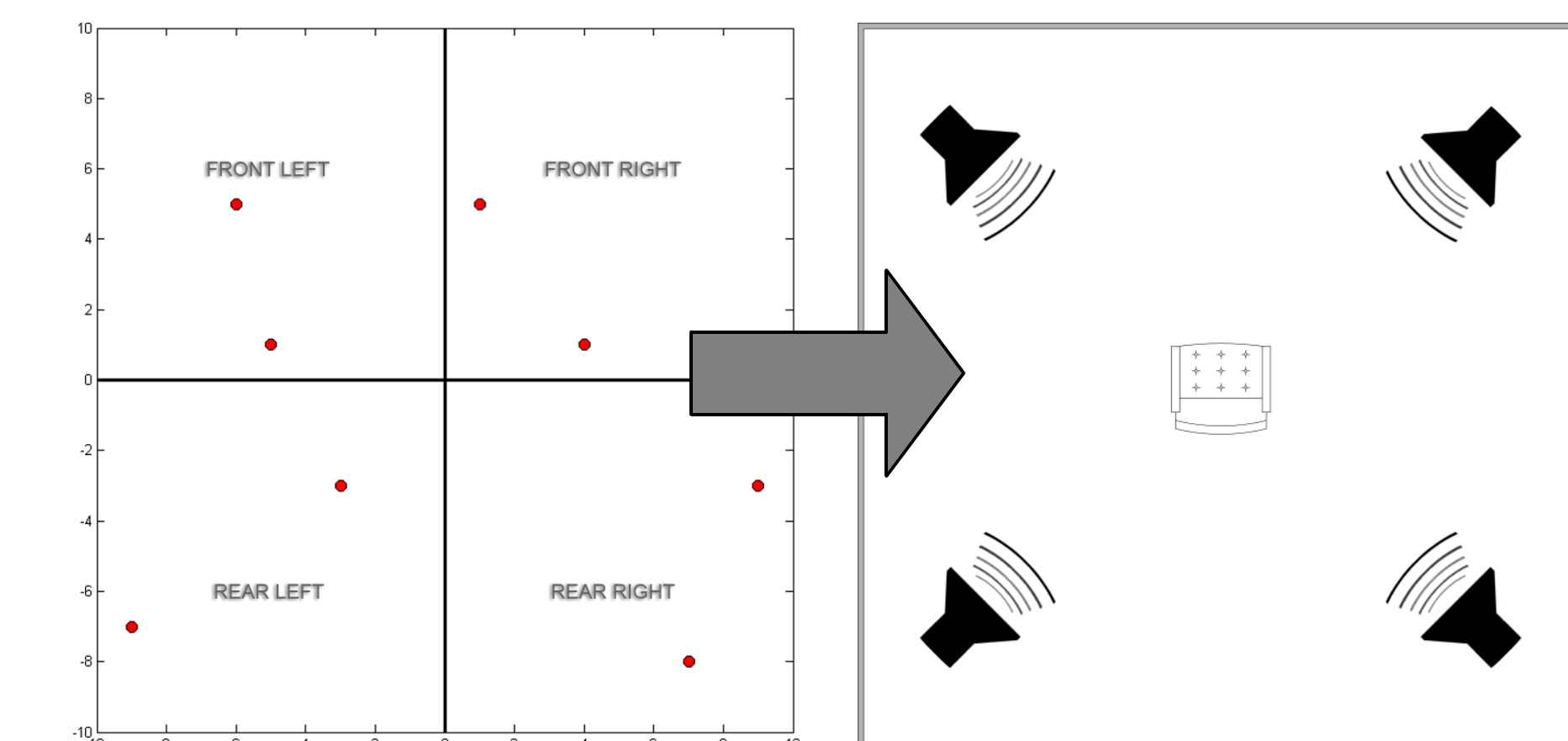


Figure 2a: Distance Mixer

Sequential Playback

Static Sequential Mixer (SSM)

- one track played on all channels consecutively

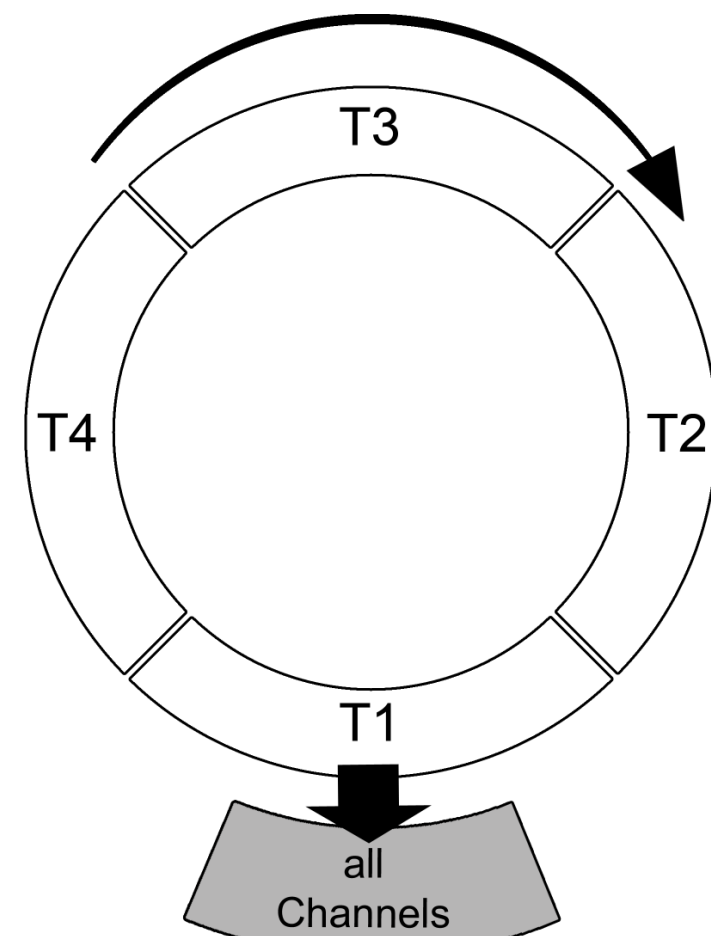


Figure 2b: SSM

Dynamic Sequential Mixer (DSM)

- one track played on each channel

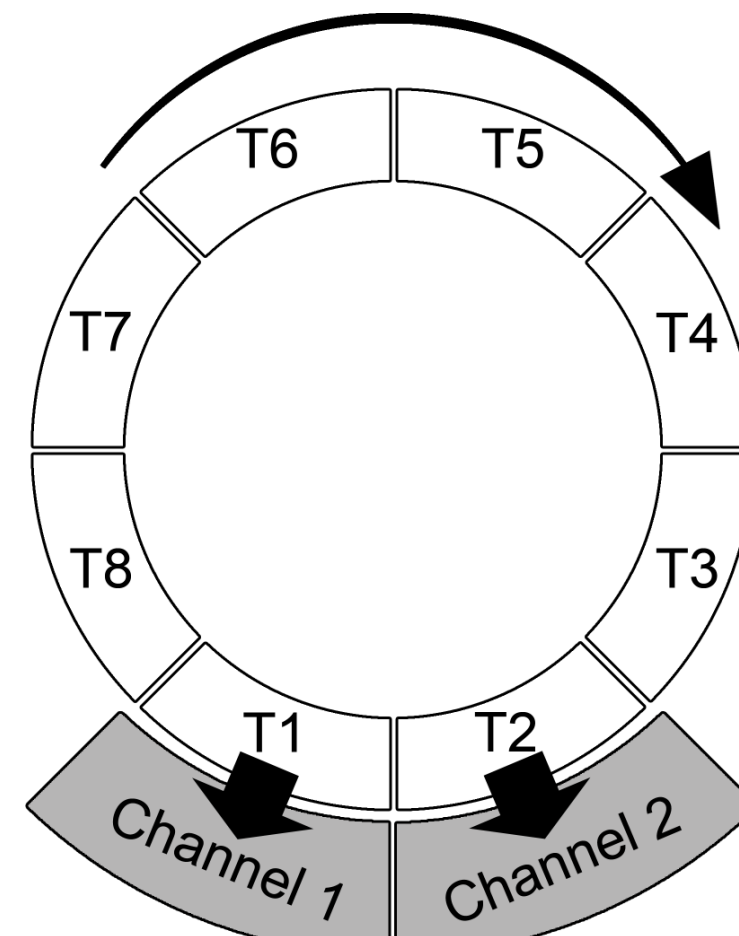


Figure 2c: DSM

Simultaneous Playback

Static Parallel Mixer (SPM)

- all tracks played on all channels simultaneously

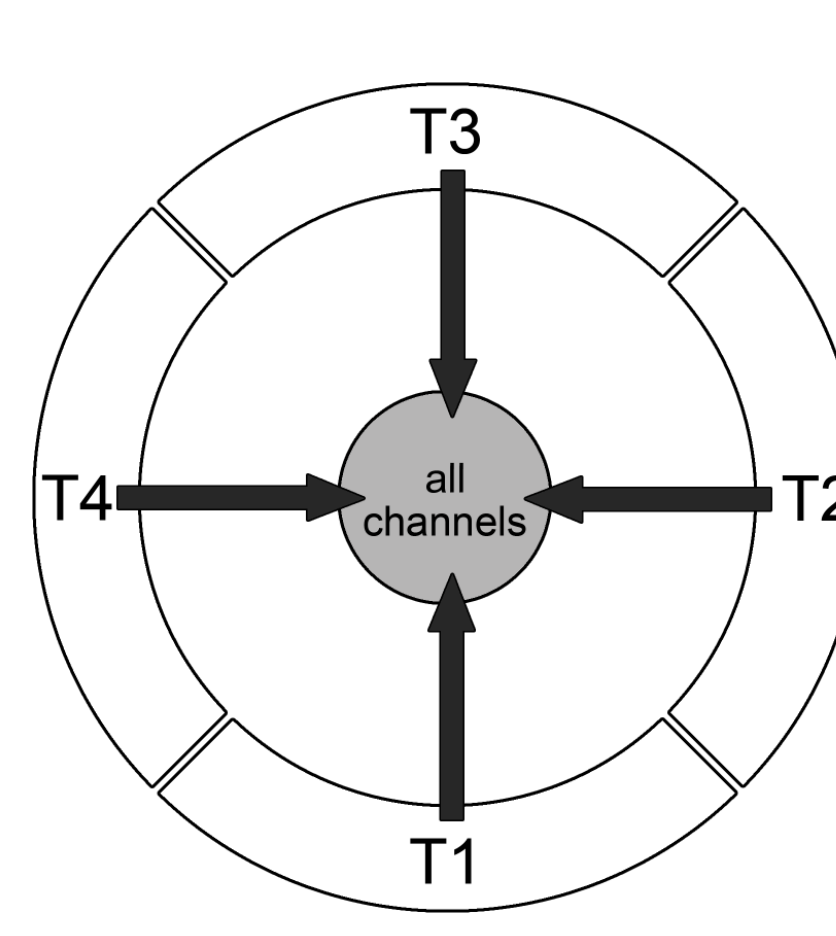


Figure 2d: SPM

Dynamic Parallel Mixer (DPM)

- like SPM and one track is emphasized consecutively

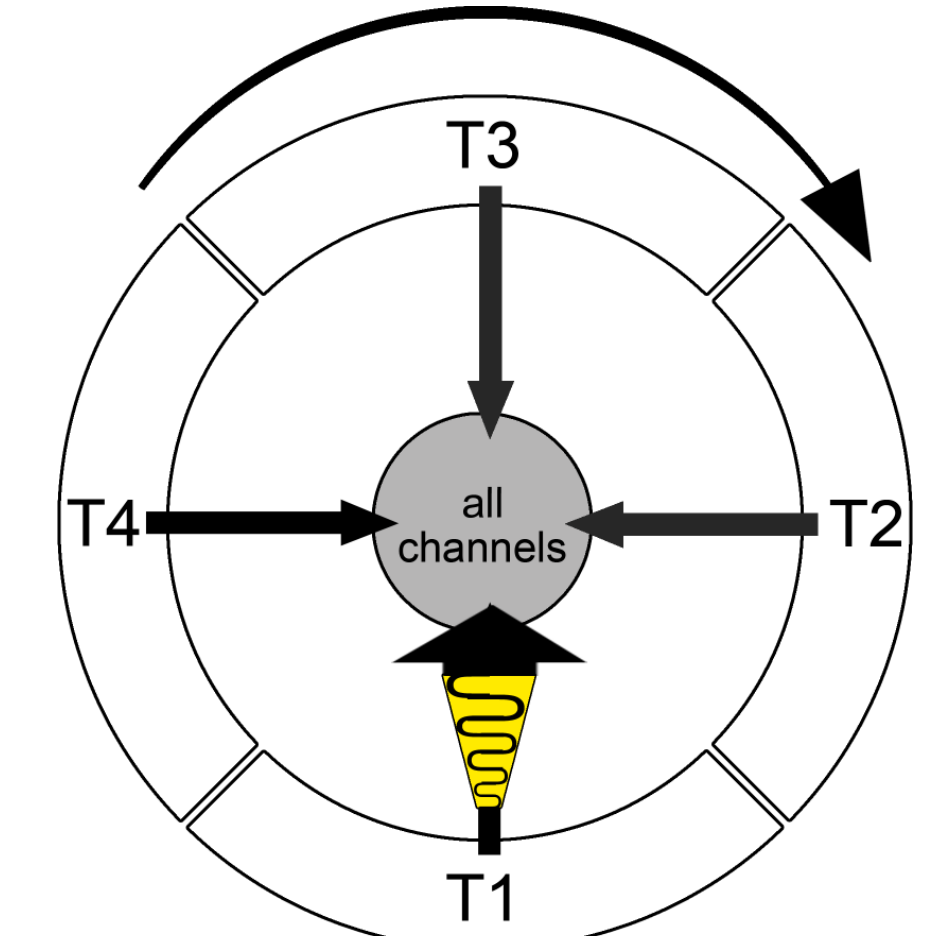


Figure 2e: DPM

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

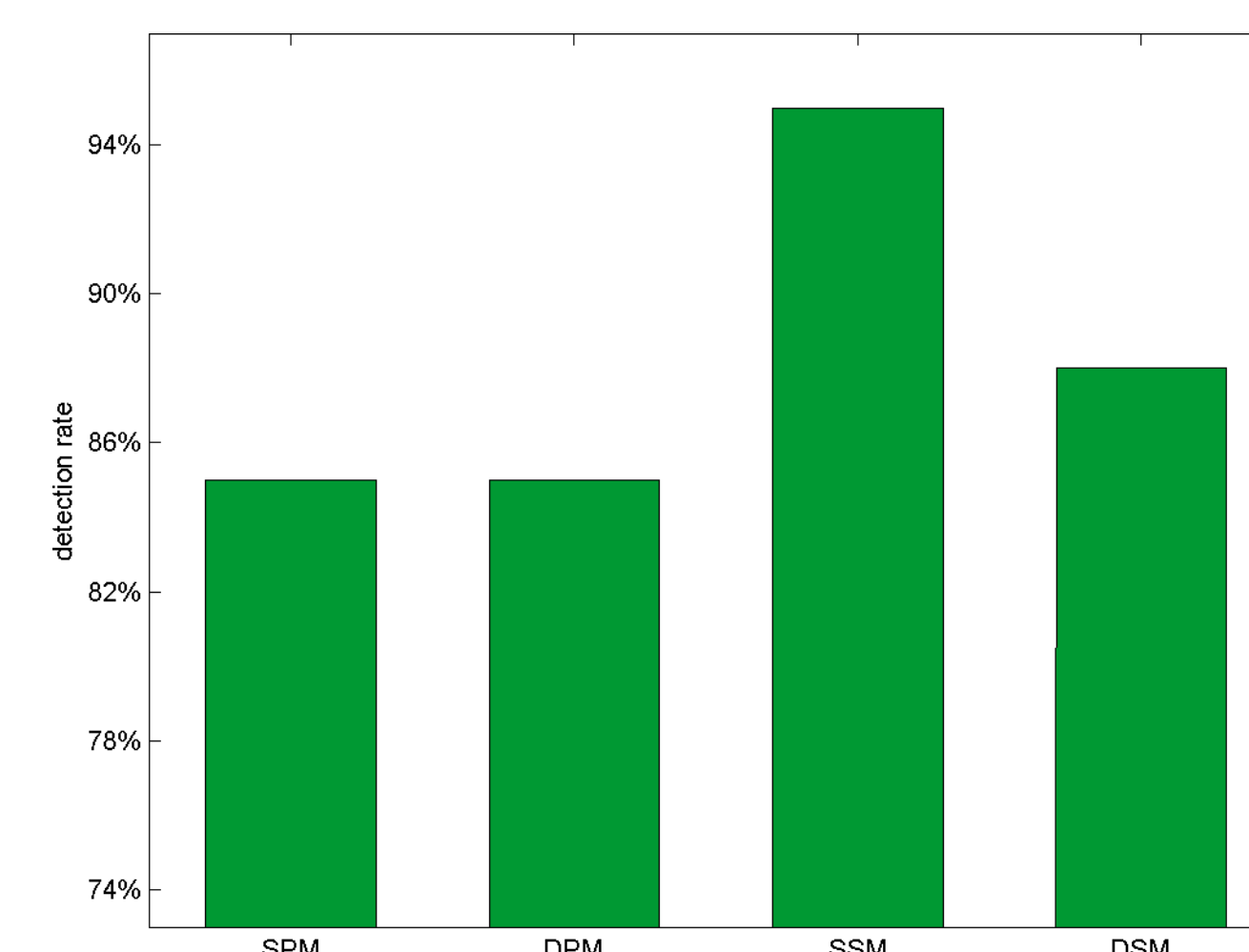


Figure 3a: Genre detection

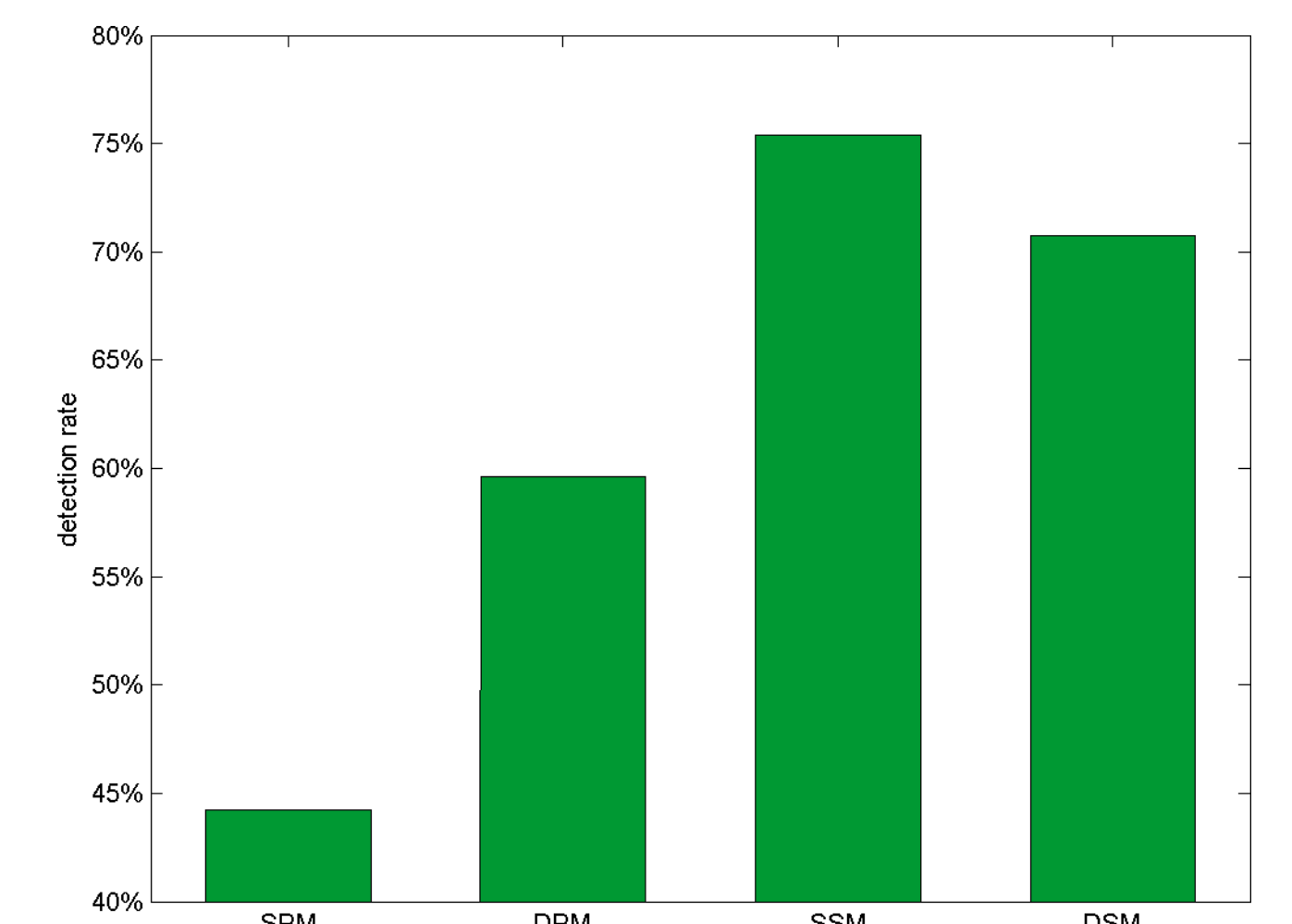


Figure 3b: Outlier detection

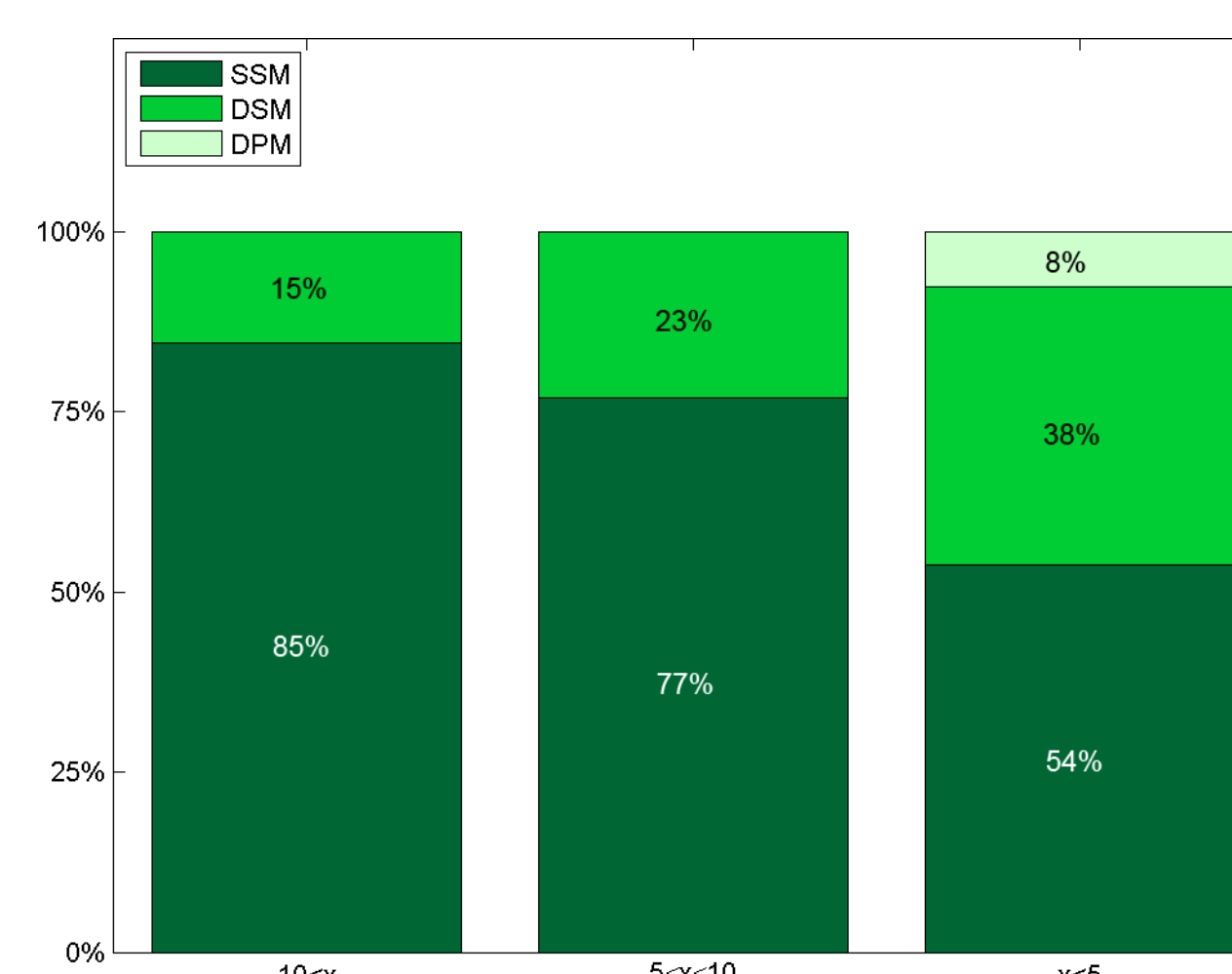


Figure 3c: Listening comfort

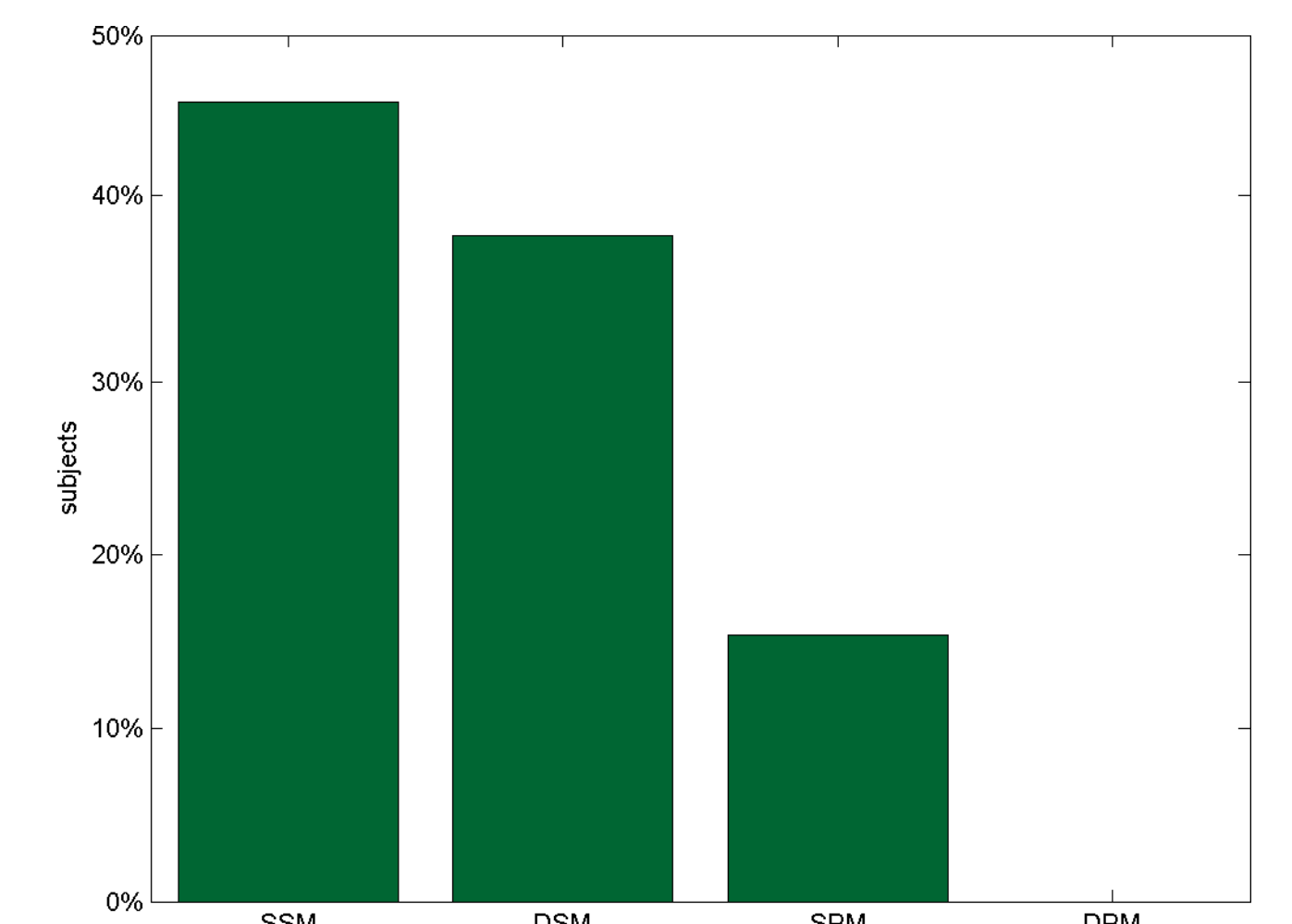


Figure 3d: Mixer's effectiveness