Genre Determination

Seminar: Music Information Retrieval, WS 16/17

Małgorzata Orłowska
Paul Wódz

Motivation

- GOSIA: Napisz tu moze cos ciekawego. Albo dopasuj tytul. Chodzi o jakis ogolny, krotki wstep
- Przykladowe zastosowania
- Dla czego nie wystarczy sie zapytac czlowieka?

Key Terms

- Genre
  - High-level concept, e.g. Classik, Pop, Rock, Jazz
- Style
  - Sub-classes of genres
  - Human concept <- to raczej dla obu
- Both defined by:
  - Audio features
  - Social, cultural or historical context
  - Relation between artists

Main Aspects

- Ontologies
- Data Source(s)
- Feature Extraction Methods
- Classification Methods

Ontologies / Taxonomies

Strona 3, C:

https://infoscience.epfl.ch/record/87348/files/Scaringella2006_1436.pdf?version=1

Opisuje co to znaczy “ill-defined” genre labels. Dodac do ostatniego slajdu jesli uzywane.

Nagrac film z ekranu i pokazac w slajdzie:

http://www.thomson.co.uk/blog/wp-content/uploads/infographic/interactive-music-map/index.html

Pokazuje jak mocno powiazane te wszystkie style’sy sa

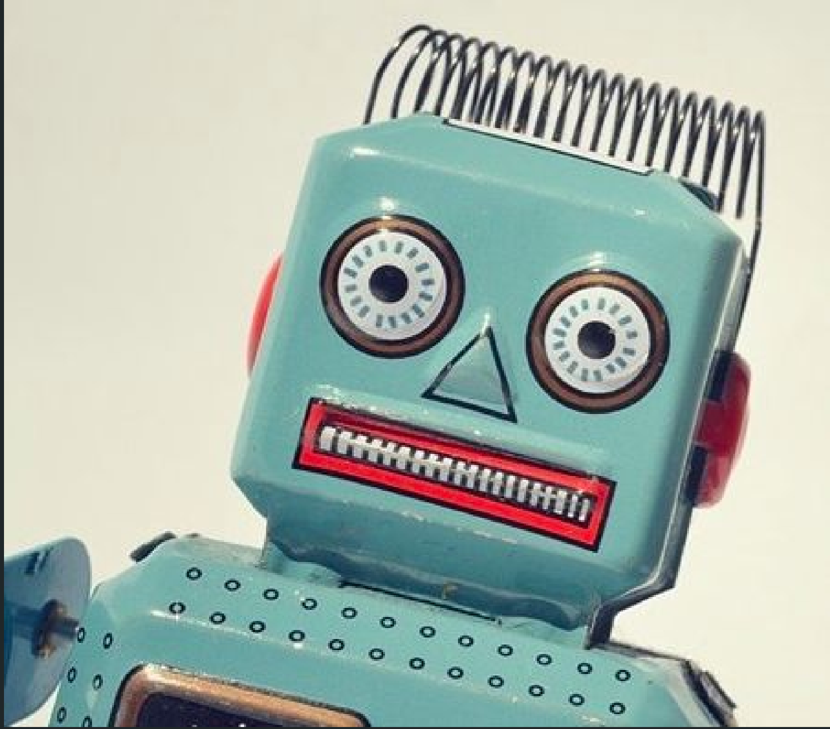
Mam jeszcze przyklad na cykliczna/nie logiczna definicje genre/style’u:

Bylo to: Trip-Hop/Downtempo/Boombap etc. Musze wyszykac na wikipedii. Ty przedstawisz

Do machines perform better than humans ?

YES!

... they do.



[2]

Experiment[x] Setup

- 24 persons
  - Aged 20-40
  - No specific musical background
  - typical mainstream music consumers
- 19 genres
  - Each represented by 10 songs
- Well-known & totally unknown artists
- Explicitly assigned genre label by artists
- Collaborative evaluation of votes

Alternative & Punk	59.2							5.0		0.4	0.4		2.1	2.9	0.4		30.4	0.4		0.4
Blues	2.5	46.3	0.4	0.4	0.8	13.3	1.3	0.4	0.8		15.4	0.4	1.3	4.2	2.9		6.7	0.4	1.7	0.8
Childens's	0.8	0.4	51.2	4.2	2.1	3.3	1.7	1.7	1.7			2.1	0.4	12.1			10.0	0.8	5.8	3.3
Classical		0.8		66.3	0.4		2.9	0.8	9.2		0.8	3.8	0.8	4.2				4.6	1.3	5.0
Comedy&Spoken Word			0.8	1.3	91.7	0.4	0.4	0.4	0.4	0.4				1.7	0.4			2.1		
Country	0.4	2.9	0.8	0.4		56.7	1.7		2.9			0.4	0.4	7.9	0.8	0.4	7.1	0.8	16.3	2.1
Easy Listening	0.8	2.9	0.4	5.0	0.8	2.5	32.1	1.3	0.4		4.2	0.8	23.8	8.3	1.7		7.1	4.2	1.3	3.3
Electronic & Dance	0.8	0.4						94.6		0.8			0.4	0.4	0.4		1.7	0.4		0.4
Vocals	0.4	0.4	1.3	0.8	0.4	17.1	5.8	0.4	5.0	0.4	0.4	1.3	1.7	10.8	0.4		6.3	1.7	39.6	5.8
Hip-Hop									1.3	90.0			0.4		7.9	0.4				
Jazz		9.6		1.3			7.5				70.4	2.9	0.4	2.9	2.9			0.8		2.9
Latin		0.8		0.8		0.4	12.1	0.8	0.8		5.4	45.8	0.8	6.3	1.3	0.4	6.3	1.7	5.0	14.2
New Age	0.4	0.8		5.8		0.4	6.7	13.3	0.4			0.4	43.3	9.6			0.4	9.6		10.8
Other																				
R&B & Soul		1.3				0.4	2.9		2.9	2.1	1.3		0.8	2.9	77.5		7.1	0.4		0.8
Reggae	0.4	0.8				0.4	4.2	1.3	0.8	18.8	0.4	3.8		0.4	5.0	57.1	4.2			3.3
Rock & Pop	6.3			0.4	0.4	2.1	1.7		2.9	1.3			3.3	5.0	6.3	0.4	67.5	1.3	0.8	1.3
Soundtracks & More		0.4	0.8	12.1			6.7	5.0	0.4		0.8	1.3	15.8	8.3	0.4			45.8		2.9
Folk	0.4	1.7	2.1	8.3		2.1	5.4	0.4	33.3		0.8		1.3	15.0	9.2		11.3	2.1	3.3	5.0
World			0.8	0.8	0.8	2.1	2.9	0.4	1.7		0.8	2.5	6.3	9.2	1.3	0.8	2.5	1.7	17.1	50.4

Human vs. Machine Performance

- Worst individual success rate: 26%
- Best individual success rate: 71%
- average accuracy: 55%
- median accuracy: 55%

- Best method -> MIREX 2016
- <Result of other method(s)>

Classification Methods

Classification Methods

Supervised

- Pre-defined classes
- Clustering according to similarity
- Typical algorithms:
    - k-Nearest-Neighbours (KNN)
    - Neural Networks (NN)
    - Support Vector Machines (SVM)
    - Rocchio’s Relevance Feedback

Unsupervised

- No taxonomy: classification emerges from data
- Typical algorithms:
    - k-Means
    - Self-Organizing Map (SOM)
    - Hidden Markov Model (HMM)

Classification Methods

Supervised

- Labor-intensive & error-prone
  - Explicit modelling of classes
  - Requires Training of Classifier
  -

Unsupervised

-

Classification Methods

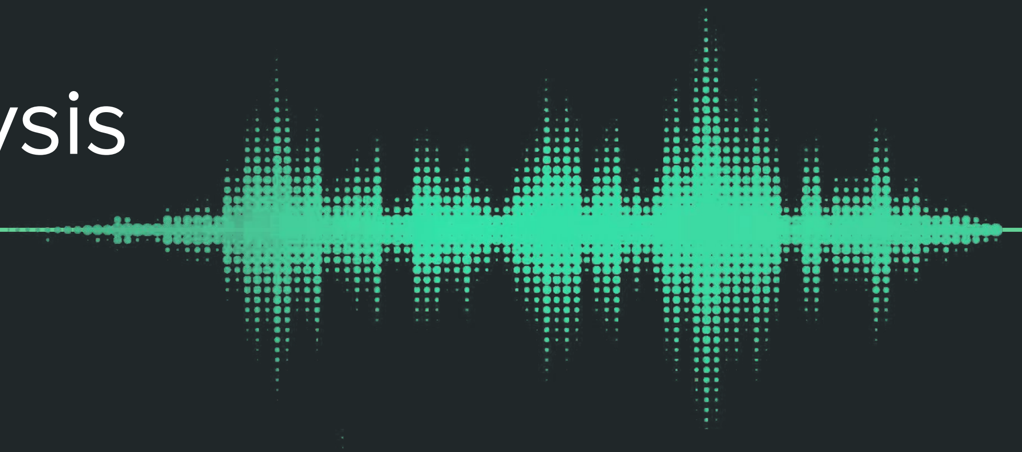
Audio Signal Analysis

- MFCCs
- ...

Natural Language Analysis

- Cultural features
- Collaborative Filtering
- Co-Occurence analysis

Audio Signal Analysis



Audio Features

- Timbre
- Harmony
- Melody
- Rhythm

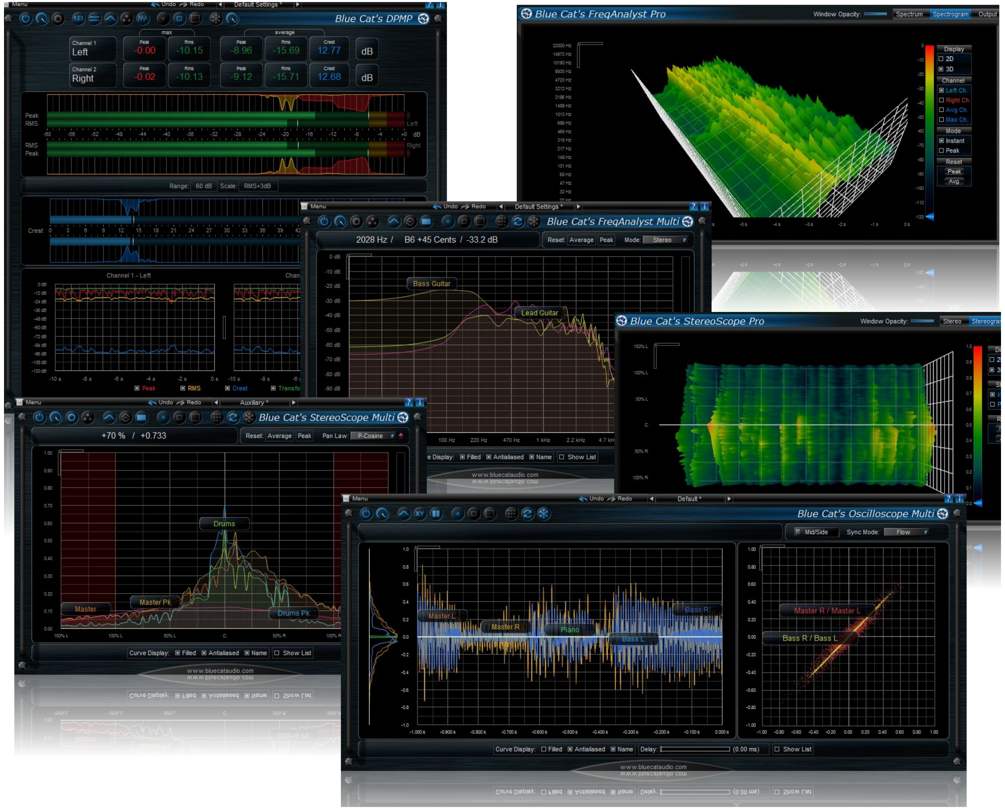


Image: www.bluecataudio.com

Feature Extraction / Algorithms

- Moze byc nie potrzebne

Natural Language Analysis

Collaborative Filtering

- Co-occurence analysis
  - user profiles
  - playlists

Analysis of Cultural Features

- Complements audio driven classification
  - Dissimilar music, but same style
  - Similar music, but different style
- Time-awareness
  - Dynamics of cultural context
  - Artists’ relations
- Instant availability
  - No bootstrapping phase

Data Sources

- Music File
  - ID3 Tags
  - Artist
  - Album name

Search

- Structured
  - Music Stores
  - Compact Disc Database
  - DBpedia
- Unstructured
  - Lyrics
  - Community meta data
    - Forums
    - Social Media
    - Reviews

Feature Extraction

From unstructured Data

Natural Language Analysis

1. Part-of-speech tagging (POS)
2. Correction of misspellings
    - Typos, phonetic
3. Term weighting and similarity calculation

Questions you have?

Ask now, you can.

References

[1] http://www.thomson.co.uk/blog/wp-content/uploads/infographic/interactive-music-map/index.html

[2] https://cdn.searchenginejournal.com/wp-content/uploads/2015/07/shutterstock_169002578-760x400.jpg

[3] Klaus Seyerlehner, Gerhard Widmer, Peter Knees, “A Comparison of Human, Automatic and Collaborative Music Genre Classification and User Centric Evaluation of Genre Classification Systems”, 2010

[x] http://vignette2.wikia.nocookie.net/starwars/images/d/d6/Yoda_SWSB.png/revision/latest?cb=20150206140125

[x] https://udemy-images.udemy.com/course/750x422/753140_f3f3_2.jpg