Advances and Challenges in Large-Scale Music Information Retrieval

Colin Raffel and Dan Ellis

DMRN+8, December 17, 2013
What can we do with lots of data?

Pablo H. Rodriguez Zivic et al. “Perceptual basis of evolving Western musical styles”
The Loudness War, four ways

Loudest Artists

Matthias Mauch, “Survival of the Flattest”; Paul Lamere, “The Loudness War Analyzed”
The Million Song Dataset

artist: 'Tori Amos'
release: 'LIVE AT MONTREUX'
title: 'Smells Like Teen Spirit'
id: 'TRKUYPW128F92E1FC0'
key: 5
mode: 0
loudness: -16.6780
tempo: 87.2330
time_signature: 4
duration: 216.4502
sample_rate: 22050
audio_md5: '8'
7digitalid: 5764727
familiarity: 0.8500
year: 1992

100.0 — cover
57.0 — covers
43.0 — female vocalists
42.0 — piano
34.0 — alternative
14.0 — singer-songwriter
11.0 — acoustic
8.0 — tori amos
7.0 — beautiful
6.0 — rock
6.0 — pop
6.0 — Nirvana
6.0 — female vocalist
6.0 — 90s
5.0 — out of genre covers

5.0 — cover songs
4.0 — soft rock
4.0 — nirvana cover
4.0 — Mellow
4.0 — alternative rock
3.0 — chick rock
3.0 — Ballad
3.0 — Awesome Covers
2.0 — melancholic
2.0 — k00l chix
2.0 — indie
2.0 — female vocalistist
2.0 — female
2.0 — cover song
2.0 — american

$5489,4468, Smells Like Teen Spirit
TRTUOUVJ128E078EE10 Nirvana
TRJCHKN12903CDD274 Pleasure Beach
TRELTOJ128F42748B7 The Flying Pickets
TRJKBXL128F92F994D Rhythms Del Mundo feat. Shanade
TRJHLAW128F429BBF8 The Bad Plus
TRKUYPW128F92E1FC0 Tori Amos

12 hello
11 i
10 a
9 and
7 it
6 are
6 we
6 now

6 here
6 us
6 entertain
4 the
4 feel
4 yeah
3 to
3 my

3 is
3 with
3 oh
3 out
3 an
3 light
3 less
3 danger

Thierry Bertin-Mahieux et al. “The million song dataset”
Using Features: Music Evolution

Joan Serrà et al. “Measuring the Evolution of Contemporary Popular Music”
Measuring The Evolution... Results

![Graphs showing the evolution of certain metrics over time and across different conditions.](http://github.com/craffel/music-evolution)
Linking with existing data

McFee and Lanckriet, “Hypergraph Models of Playlist Dialects”
Mahieux and Ellis, “Large-Scale Cover Song Recognition Using the 2D Fourier Transform Magnitude”
McVicar and De Bie, “CCA and a Multi-way Extension for Investigating Common Components between Audio, Lyrics and Tags”
Ground Truth?
Audio? One solution:

Schindler et al. “Facilitating Comprehensive Benchmarking Experiments on the Million Song Dataset”
MIDI + Audio + MSD

artist: 'Tori Amos'
release: 'LIVE AT MONTREUX'
title: 'Smells Like Teen Spirit'
id: 'TRKUYPW128F92E1FCO'
duration: 216.4502
sample_rate: 22050
audio_md5: '8'
7digitalid: 5764727
year: 1992
Matching by Text

artist: 'Tori Amos'
release: 'LIVE AT MONTREUX'
title: 'Smells Like Teen Spirit'
duration: 216.4502

/nirvana/smells_liketeen_spirit.mid

artist: Amos, Tori
release: Live 91/92
title: Smells Like Teen Spirit (Nirvana cover)
duration: 216.3
Matching by Content
Alignment

Turetsky and Ellis, “Ground-Truth Transcriptions of Real Music from Force-Aligned MIDI Syntheses”
Scale and Access

- 300 GB
  - Open access

- 3 GB
  - Open access?

- 200 TB
  - Restricted access
What does aligned MIDI get us?
Thanks!

[Image of The MIDI Organizer advertisement]

CALL NOW TOLL NOT INCLUDED
1-800-555-0199
CINCO
ALLOW 10-12 WEEKS FOR DELIVERY
NO MONEY BACK GUARANTEE

craffel@gmail.com