

Can't trust the feeling?

How open data reveals unexpected behavior of high-level music descriptors

Cynthia C. S. Liem

Chris Mostert

Multimedia Computing Group

Delft University of Technology

The Netherlands

AcousticBrainz

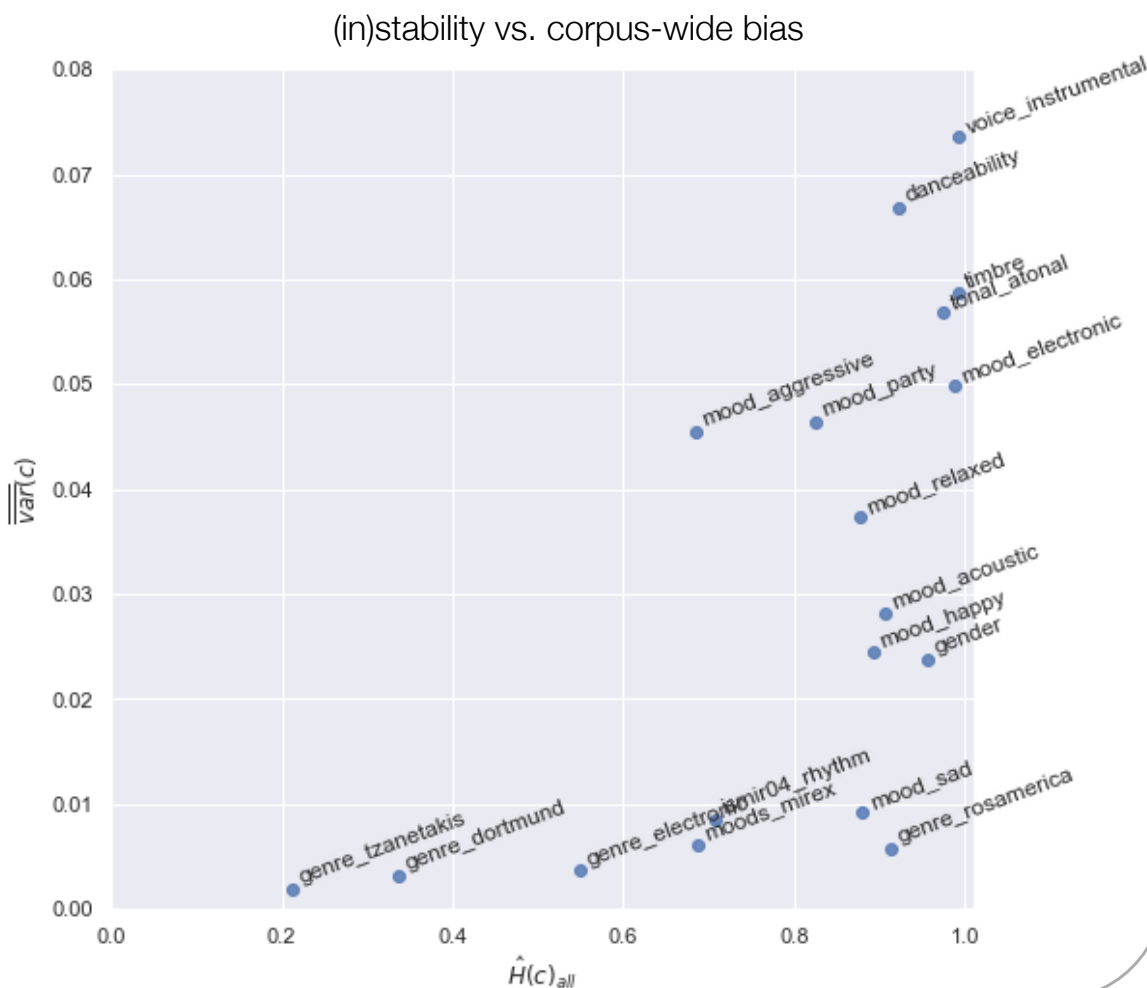
- Large-scale music audio descriptors, extracted on community-chosen repertoire
- Multiple submissions per MusicBrainz Recording ID allowed
- Well-known, well-documented open-source extractors/classifiers + extensive metadata
- **Does ‘in-the-wild’ descriptor behavior differ from performance in the lab?**
- **Do descriptors describe what they are supposed to describe?**
- **How can we know, if we don’t have ground truth?**

Stability in resubmissions

- Inspired by derived oracles in Software Testing
- For the same classifier c and same MBID m , one would expect

`classify_c(my_preprocessing(m))`
==
`classify_c(your_preprocessing(m))`

- Consider (in)stability within resubmissions and bias across corpus (pooled variance/entropy)



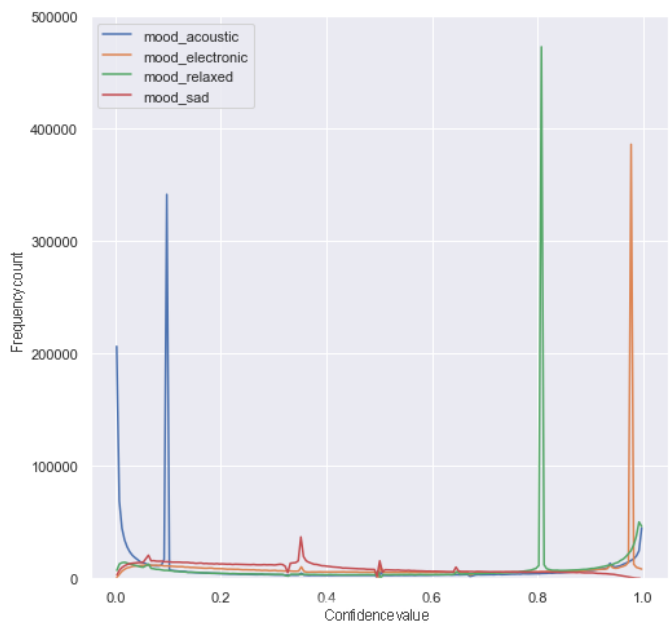
Correlations between constructs

- Inspired by approaches to construct validity in Psychology
- Redundancy in constructs (e.g., multiple genre classifiers with overlapping labels)
- Theorized relations between constructs (e.g., happy opposite of sad (?))

Classifier, label A	Classifier, label B	Pearson's r
genre_rosamerica, cla	genre_tzanetakis, cla	.29
genre_dortmund, rock	genre_rosamerica, roc	.24
genre_dortmund, jazz	genre_rosamerica, jaz	.22
genre_dortmund, pop	genre_rosamerica, pop	.11
genre_dortmund, jazz	genre_tzanetakis, jaz	.08
genre_rosamerica, pop	genre_tzanetakis, pop	.06
genre_rosamerica, hip	genre_tzanetakis, hip	.05
genre_rosamerica, jaz	genre_tzanetakis, jaz	.02
genre_dortmund, blues	genre_tzanetakis, blu	.01
genre_dortmund, pop	genre_tzanetakis, pop	-.05
genre_dortmund, rock	genre_tzanetakis, roc	-.06
genre_rosamerica, roc	genre_tzanetakis, roc	-.07
mood_aggressive, aggressive	mood_relaxed, not_relaxed	.59
mood_acoustic, acoustic	mood_electronic, not_electronic	.58
danceability, danceable	mood_party, party	.53
mood_electronic, electronic	genre_dortmund, electronic	.48
danceability, danceable	genre_rosamerica, dan	.33
mood_happy, happy	mood_party, party	.20
mood_happy, happy	mood_sad, not_sad	.13

Strange distributions

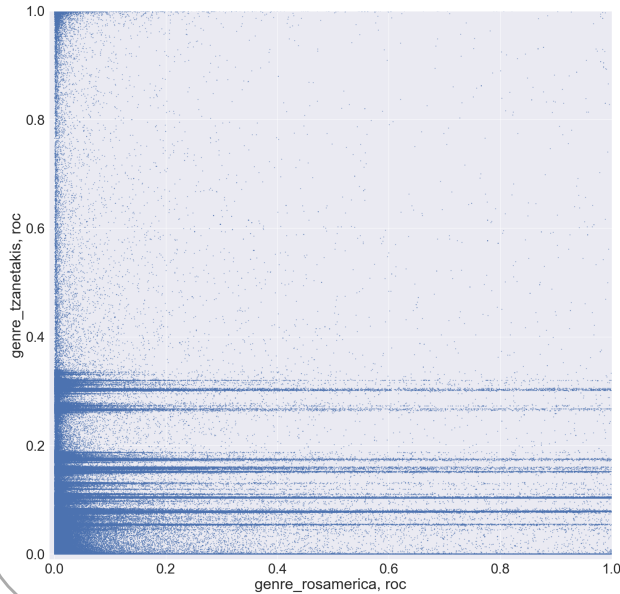
confidence distributions: acoustic, relaxed, electronic and sad moods



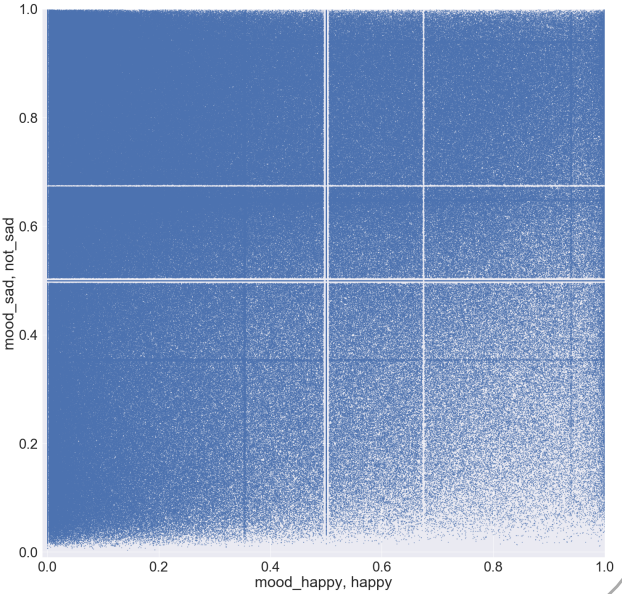
- Compare metadata (through Jensen-Shannon distance) for anomalies vs. non-anomalies

	acoustic	relaxed	electronic	sad
bit_rate	.42	.32	.39	.17
codec	.34	.26	.32	.06
length	.15	.15	.15	.32
lossless	.28	.21	.27	.02
essentia_low	.61	.52	.59	.15
essentia_git_sha_low	.67	.58	.66	.23
essentia_build_sha_low	.70	.62	.69	.24

rock: genre_tzanetakis vs. genre_rosamerica



mood: not_sad vs. happy



- largest distances for **codec, bit rate & software version of low-level feature extractors.**
- **these aspects are normally overlooked!**

