MOODetector: A System for Mood-based Classification and Retrieval of Audio Music

Rui Pedro Paiva

University of Coimbra

# Contents

- ❑ **Research Goals**

- ❑ **Current Work**

- ❑ **Future Work**

MOODetector

Research Goals

- ❑ **Mood models**
  - • **Categorical vs Dimensional, Discrete vs Continuous**
- ❑ **Feature extraction**
  - • **Audio, MIDI, lyrics(?)**
- ❑ **Feature selection and evaluation**
  - • **Feature relevance**
  - • **Feature space reduction**
  - • **Feature combinations**
- ❑ **Knowledge extraction**
  - • **Fuzzy rules**
- ❑ **Mood tracking**

Current Work

❑ **Automatic Creation of Mood Playlists**

    ❑**Mood model: Thayer model**
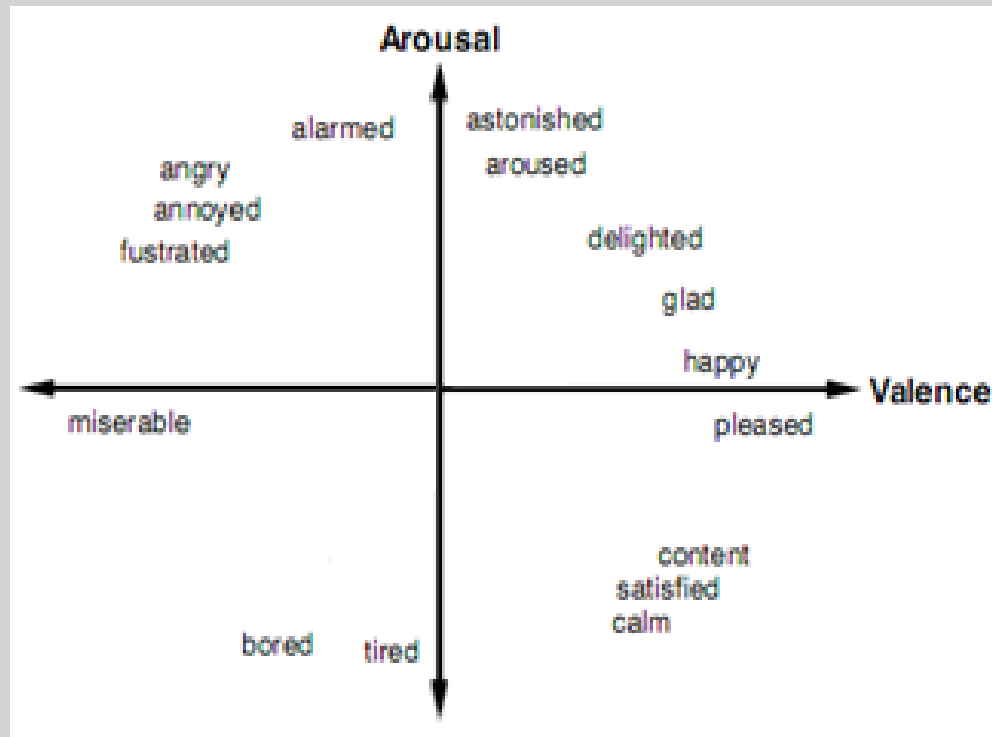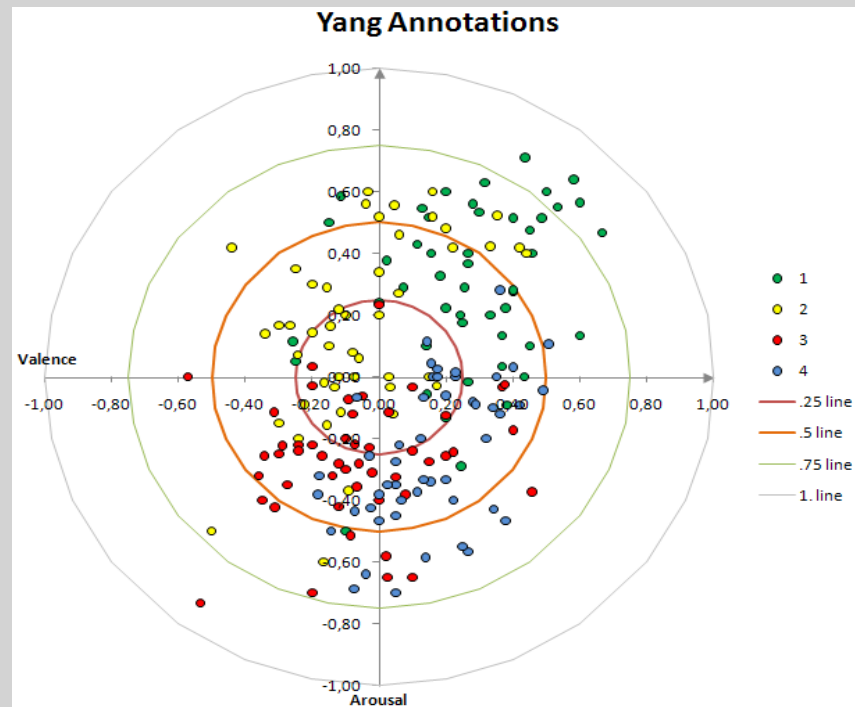
Current Work

❑ **Automatic Creation of Mood Playlists**

    ❑**Ground Truth**

        ❑Yang's annotations

           ❑ 194 25-sec excerpts manually annotated in the Thayer plane

Current Work

❑ **Automatic Creation of Mood Playlists**

    ❑**Feature Extraction and Selection**

        ❑Which features?

- Timing: Tempo, tempo variation, duration contrast
- Dynamics: overall level, crescendo/decrescendo, accents
- Articulation: overall (staccato/legato), variability
- Timbre: Spectral richness, onset velocity, harmonic richness
- Pitch (high/low)
- Interval (small/large)
- Melody: range (small/large), direction (up/down)
- Harmony (consonant/complex-dissonant)
- Tonality (chromatic-atonal/key-oriented)
- Rhythm (regular-smooth/firm/flowing-fluent/irregular-rough)
- Mode (major/minor)
- Loudness (high/low)
- Musical form (complexity, repetition, new ideas, disruption)

Current Work

- ❑ **Automatic Creation of Mood Playlists**
    - ❑ **Feature Extraction and Selection**
        - ❑ Audio features: 3 frameworks
        - ❑ Forward Feature Selection
        - ❑ PCA
    - ❑ **Arousal-Valence Modeling**
        - ❑ Support Vector Regression
   - ❑ Grid parameter search

Framework	Feature	Description
PsySound 2	44	Extracts features like loudness, sharpness, volume, spectral centroid, timbral width, pitch multiplicity, dissonance, tonality and chord, based on psycho acoustic models
MIR Toolbox	177	Among others: Root mean square (RMS) energy, rhythmic fluctuation, tempo, attack time and slope, zero crossing rate, rolloff, flux, high frequency energy, Mel frequency cepstral coefficients (MFCCs), roughness, spectral peaks variability (irregularity), inharmonicity, pitch, mode, harmonic change and key.
Marsyas	237	Extracts features like spectral centroid, rolloff, flux, zero cross rate, linear spectral pair, linear prediction cepstral coefficients (LPCCs), spectral flatness measure (SFM), spectral crest factor (SCF), stereo panning spectrum features, MFCCs, chroma, beat histograms and tempo.

Current Work

❑ **Automatic Creation of Mood Playlists**

　❑**Playlist Generation**

　　❑ Closest songs to the seed in the Thayer plane

　　❑ Mood trajectory
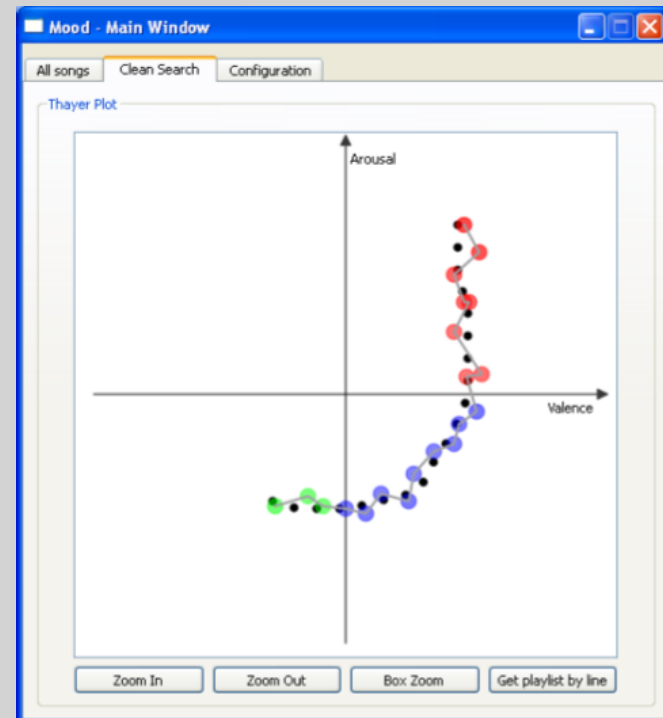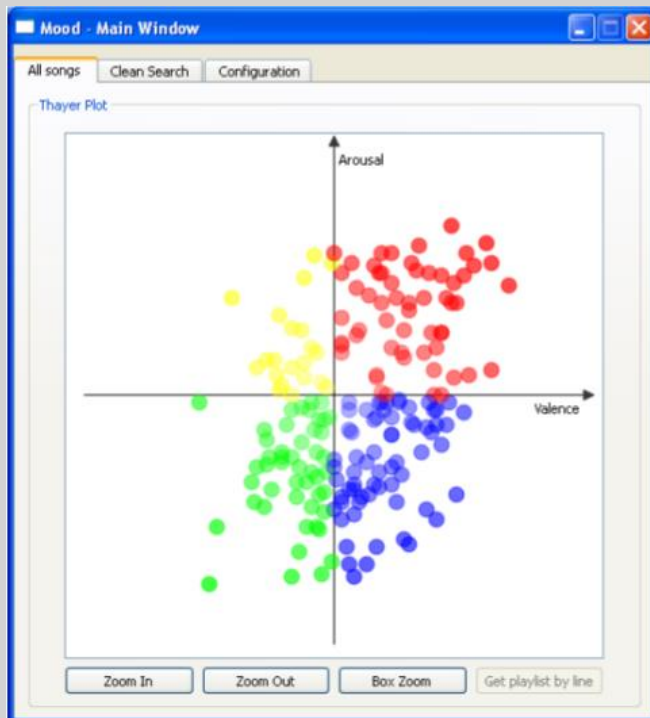
Current Work

❑ **Automatic Creation of Mood Playlists**
- ❑**Preliminary Results**
  - ❑Regression
    - ❑ 10-fold cross validation, 20 repetitions
    - ❑ Forward Feature Selection: 53 features for Ar., 80 for Va.
    - ❑ R2 statistics

	All features		FFS	
	Arousal	Valence	Arousal	Valence
PsySound (15)	58.7%	12.7%	60.3%	21.0%
PsySound (44)	57.3%	7.9%	57.3%	19.1%
MIR Toolbox	58.2%	8.5%	58.7%	25.7%
Marsyas	52.9%	2.7%	56.0%	4.6%
ALL + PCA	56.5%	23.4%	61.8%	27.2%
ALL	57.4%	19.4%	62.9%	35.6%

❑ Yang: Ar = 58.3%, Va = 28.1%

Current Work

❑ **Automatic Creation of Mood Playlists**

  ❑**Preliminary Results**

    ❑Playlist quality

      ❑ SVR training and distance to the seed

      ❑ Previously selected features

      ❑ 4-fold cross validation (75-25%)

|	Top1		Top5		Top20	
	All	*FFS*	*All*	*FFS*	*All*	*FFS*
PsySound (15)	4.2	5.6	21.1	21.5	61.9	62.0
PsySound (44)	4.2	3.8	20.9	21.2	60.5	61.9
MIR Toolbox	3.6	5.2	22.8	22.0	62.7	62.5
Marsyas	4.0	4.4	18.1	19.8	58.5	60.0
ALL	4.2	6.2	21.0	24.8	60.7	62.3

MOODetector

Current Work

❑ **Mood Tracking**

❑ **Ground Truth**

❑ Annotations by 2 subjects: quadrants only

❑ **Preliminary Results**

❑ Classification: SVM

❑ 4 classes (quadrants)

❑ Using only Marsyas: 48%
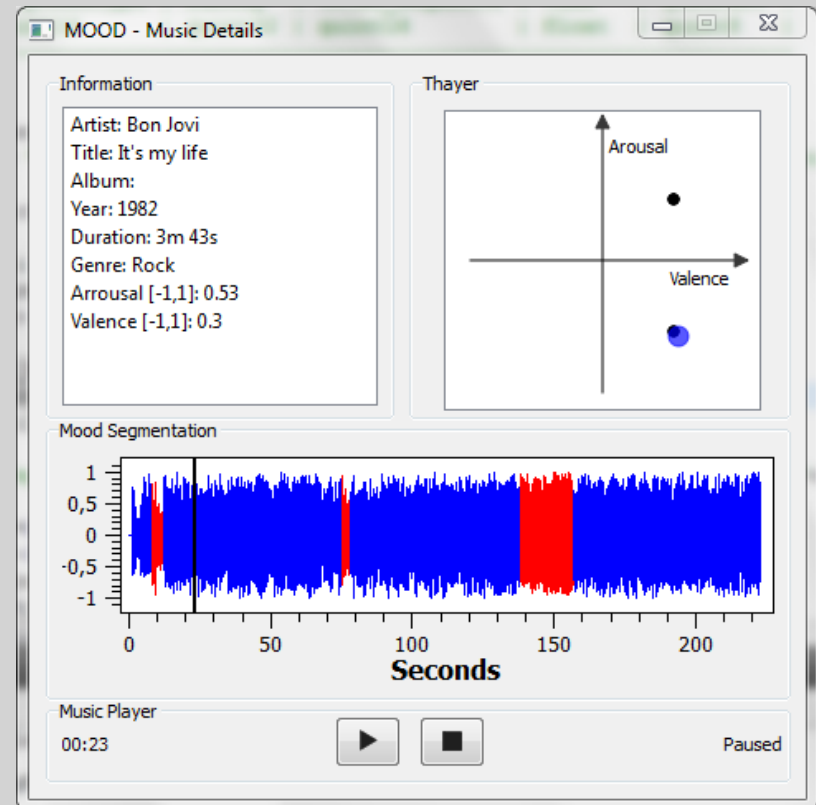
Future Work

❑ **Ground Truth**

   ❑**Clip-level and mood-tracking**

❑ **Feature Extraction**

   ❑**Propose good features for valence**

   ❑**Extract symbolic features from MIDI files**

      ❑MIDI versions of the audio songs

      ❑Automatic music transcription

❑ **Feature Selection**

   ❑**Feature combination: e.g., RRelieff**

   ❑**Genetic Algorithms**

❑ **Knowledge Extraction**

   ❑**Neural-Fuzzy Systems**

MOODetector