EXPLORING AUTOMATIC TRANSLATION BETWEEN AFFECT REPRESENTATION SCHEMES OF MUSIC AFFECTIVE CONTENT

Author: Alissia Rugina | a.rugina@student.tudelft.nl

affiliates: Bernd Dudzik, Chirag Raman

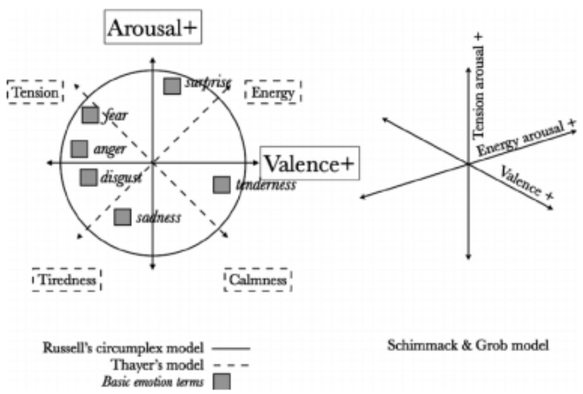
1. Background

Music Affect Content Analysis
is the analysis of affect states induced by music content as stimuli
Applications:

Therapy, Filmmaking, Marketing

Representation Schemes: Discrete vs. Dimensional

Figure comparing Rusell's model and Schimmack & Grob's model taken from [2]



2. Research Questions

- **Is it feasible to translate between representation schemes of Music Affect Content using Machine Learning Models?**
- **What variables impact the performance of the translation?**

3. Methodology

FIND DATASET → DATA ANALYSIS → SELECT MODELS → OPTIMIZE MODELS → EVALUATE

Soundtracks as Stimuli
- 360 excerpts of movie soundtracks
- scores from 1 - 9
- Basic Emotions: **Happy, Sad, Fear, Anger, Tender.**
- Schimmack & Grob model **Valence, Energy, and Tension**.

- imbalanced classes
- Fear & Anger close in dimension means

Classification task:

Valence, Energy, Tension → Anger Fear Happy Sad Tender

Dimensional → Categorical

3 Classifiers:
- Logistic Regressor
- Decision Tree
- K-Nearest Neighbor

- Accuracy of classifying validation set averaged over 20 times repeated Stratified 5-fold Cross-Validation
- t-test comparison to Dummy classifier performance

4. Findings

Models Performance

Model	Accuracy			Significantly different from Baseline
	Average	Min	Max	
Logistic Regressor	0.66	0.62	0.69	✓
Decision Tree	0.70	0.61	0.75	✓
K-Nearest Neighbors	0.67	0.61	0.74	✓

Models Performance
- An ANOVA test finds no significant difference across the accuracies of the different models

Performance across emotion classes
- ANOVA test confirms a significant difference in models' performances across the emotion classes
- Models perform the worst on either *Fear* or *Anger*

Analysis of Fear and Anger
- Tukey test results: Anger & Fear are significantly different in :

Valence	Energy	Tension
✗	✓	✗

Discussion
- Cannot conclude which model is better
- Similarity of Fear and Anger causes low accuracy

Past studies suggest...
- music as stimuli can only induce a restricted range of unpleasantness and activation (Valence & Arousal), causing a small separation of emotions [3]
- Dimensional schemes are more representative of states induced by music than categorical emotions [1]

5. Limitations

Interdisciplinary topics | Finding datasets | Time constraint

6. Conclusion

Feasible?
- Yes!

Variables that impact performance?
- proximity of Fear & Anger in Valence & Tension dimensions
- using categorical emotions to describe music affect states

Future research ideas:
- using acoustic properties
- using more dispersed categorical emotions

Note: these conclusions have been made based on this specific Music Affect database using the specific emotion schemes, it is unknown how this model would generalize to different datasets, as well as whether translation between different schemes are feasible

References

[1] Julian Cespedes-Guevara and Tuomas Eerola. Music communicates affects, not basic emotions - a constructionist account of attribution of emotional meanings to music. Frontiers in Psychology, 9, 2 2018.

[2] Tuomas Eerola and Jonna K. Vuoskoski. A comparison of the discrete and dimensional models of emotion in music. Psychology of Music, 39:18–49, 2011.

[3] Klaus R. Scherer. Which emotions can be induced by music? what are the underlying mechanisms? and how can we measure them? Journal of New Music Research, 33:239–251, 2004.

TU Delft