

R3MO: A Multi-modal Dataset for Emotion Analysis in Orchestral Music

Joann Ching (Johannes Kepler University Linz); Carlos Eduardo Cancino-Chacón (Johannes Kepler University Linz); Gerhard Widmer (Johannes Kepler University Linz)

The R3MO Dataset is a multi-modal dataset with emotion annotations of Rachmaninoff's Piano Concerto No. 3 in D minor, Op. 30. It comprises 11 recordings of the concerto by different pianists and orchestras, a two-piano reduction of the orchestral score, and non-overlapping segment-wise emotion annotations. The segments are created according to the musical structure. Each segment is annotated by, on average, 8.9 students in a CS/AI course, with both dimensional (valence-arousal) and categorical descriptors (GEMS and GEMMES) on a 5-point Likert scale. While Russell's Circumplex [1] provides a general psychological model of affect in a two-dimensional space, the Geneva Emotional Music Scale (GEMS-9) [2] is a music-specific taxonomy designed to capture more nuanced, music-evoked emotions. Furthermore, the Geneva Musical Metaphors Scale (GEMMES) [3] extends this approach through metaphor-based descriptors, allowing intuitive labeling by annotators regardless of musical expertise. A set of low- and mid-level acoustic features extracted from the audio recordings is also provided to complement the annotations. To evaluate annotation reliability, we compute Krippendorff's alpha across emotional measures and quantify inter-annotator variability with the coefficient of variation. We further analyze relationships between categorical and dimensional annotations by correlating the categorical descriptors with valence-arousal ratings, mapping each descriptor onto the circumplex. Despite expected differences in inter-annotator agreement across these emotional measures, the resulting mapping shows consistency with prior neurological findings [4], which classify GEMS descriptors into high- and low- valence-arousal quadrants based on brain activity, thereby showing the non-arbitrariness of the annotations. Furthermore, leveraging the structure-based segmentation, we further explore how emotional trajectories relate to formal sections of the concerto and extracted acoustic descriptors, and differences across performances. This work establishes a multi-modal framework for linking annotated emotions, acoustic features, and musical structure, offering a foundation for both musicological analysis and computational modeling of affect in classical music.

References

- [1] James Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178, 12 1980.
- [2] Marcel Zentner, Didier Grandjean, and Klaus Scherer. Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8:494–521, 08 2008
- [3] Simon Schaerlaeken, Donald Glowinski, and Didier Grandjean. Linking musical metaphors and emotions evoked by the sound of classical music. *Psychology of Music*, 50:030573562199123, 03 2021.
- [4] Wiebke Trost, Thomas Ethofer, Marcel Zentner, and Patrik Vuilleumier. Mapping aesthetic musical emotions in the brain. *Cerebral cortex (New York, N.Y. : 1991)*, 22, 12 2011.