

Every Feature that Rises will Converge?

Towards incorporating notions of feature shape in music information retrieval

Yi-Yun Cheng (Jessica)

Features describing aspects of a musical audio signal can approximate semantic descriptions of interest to musicologists, but understanding and making good use of these features is not straightforward. Imagine a context where a musicologist wishing to conduct a harmonic analysis could be guided toward features sharing a “harmonic shape” (operating in the spectral domain), without requiring extensive signal processing background knowledge. We propose to address this issue by conveying information about feature *shapes*, the characteristics of the feature extraction process that are shared between different subsets of features.

The Audio Feature Ontology and Vocabulary (AFO/AFV) surveyed existing MIR feature taxonomies, enumerating a comprehensive list of audio features, and presenting process descriptions specifying the operation sequence of each feature extractor. For example, for the chromagram feature, AFO/AFV describes its operation sequence as Windowing, Discrete Fourier Transform, Logarithm, and Sum.

In this research, we use the operation sequences of the AFV/AFO to inform analytical workflows on feature data exhibiting different types of processes. We also explore the feasibility of feature shape-based filtering and querying within the Internet Live Music Archive, a large collection of audio recordings. We further consider the commonalities and divergences between the operational sequences defined by AFO/AFV and analogous processes within the Extracted Feature Dataset of the HathiTrust, a digital library containing a large collection of OCR-digitized volumes, to gain a more generic understanding of feature-shape based explorations in information retrieval.