Towards a cognitively-based transcription algorithm for non-Western vocal melody

Kaustuv Kanti Ganguli and Carlos Guedes

Division of Arts & Humanities
New York University Abu Dhabi, Abu Dhabi, UAE
{kaustuvkanti, carlos.guedes}@nyu.edu

Abstract

For the predominantly oral non-Western tonal music traditions, there exist text resources for musicology, wherein the lexicon finds its place in a rather prescriptive manner compared to a written music score. Alongside addressing the issue of subjectivity, scalability, and reproducibility in musicological research; we propose novel methods for relevant music information retrieval (MIR) applications in the light of artificial intelligence (AI). While a general engineering approach is to optimize certain evaluation metrics, we aim to align our findings to be informed by human judgment. We propose computational representations that robustly capture the particular features while being sensitive enough to the differences between differently labeled concepts, for the task of melodic similarity in content-based retrieval scenario; tested within a sizable, representative music corpus. Coming to algorithms, we take up a case study for vocal melody transcription in Indian art music. While dimensionality reduction of time-series to discrete symbol strings is a standard approach that can exploit computational gains from the data compression as well as the availability of efficient string matching algorithms, the compressed representation of the pitch time-series itself is not well understood given the pervasive-ness of pitch inflections in the melodic shape. A melodic stylization algorithm combining domain knowledge- and data-driven optimization is shown to achieve a retrieval performance comparable to the time-series based matching in an ``audio search by query'' task at significantly lower computational cost. This also facilitates a decomposition of a performance melody into ``prototypical'' or signature shapes invariant across artists; and the residue as artist-specific ornamentations. Apart from being an efficient ``search enabler'', this can also facilitate as a supplement to enthnomusicological studies.