Analyzing lyrics and melody: A MIR approach to linguistic tone and melodic pitch contour similarity in Beijing Opera in a large corpus

Shuo Zhang, Rafael Caro, and Xavier Serra
CompMusic Project, Music Technology Group, DTIC,
Universitat Pompeu Fabra, Barcelona, Spain

Scholars of Beijing Opera have long argued that there exists a correspondence between the pitch contours of linguistic tones and singing[1]. The overall degree and manner of this relationship is not entirely clear, partly due to the limitation that music scholars typically were not able to go beyond analyzing a few arias by hand (Xu 2007). In this project we propose a novel approach to this topic with computational tools in Music Information Retrieval (MIR) and data mining that are able to automatically annotate and extract information from a corpus of 510 recordings (containing 363 arias) at the CompMusic Project[2].

We report a pilot study in the current stage based on a selection of 20 arias of the rhythmic type *yuanban*[3] as an illustration of our methodology. In this study, our first goal is to build a model to represent the average pitch contour shapes in singing for each lingusitic tone category. We're also interested in investigating the long-standing issue of which of the two competing dialectal system of Beijing Opera at work (Beijing and Hu-Guang) appear to be more dominant. The second goal is to explore the effect of four different structural positions of syllables on the similarity: P1B,P2B,UB, and Other positions[4].

A total number of 1,993 tone units (evenly distributed across 4 tones) and relevant metadata tags were (partially) automatically annotated and extracted from the 20 arias[5]. Each pitch contour unit is then normalized in duration and pitch height for comparison. Following data preprocessing, an 'average' contour (with contours of standard error estimates) is computed for each tone category by applying the Smoothing Spline Anova nonparamctric regression models (Gu 2002) in statistical computing language *R*. We also performed various exploratory data analysis on a subset of the 'core' data[6] and computed the 'average' contour for each category in each structural position, generating a 4*4 matrix of contour model plots. Results show that models based on different positions better predict the behavior of tone-melody similarity, whereas cross-category differences may be obscured by the large amount of variation in data when a specific tone category is considered as a whole (possibly due to positional and other effects). In addition, the Hu-Guang tone system is shown as a dominant factor in predicting the melodic contours (especially for tone 3 and 4). Overall, the analysis generated comparable patterns to those derived from manual analysis, while also showing additional unexpected and more fine-grained results than traditional analysis on a large sample size that has not been possible in previous research. Future directions and on-going topics of research include computing the distance/similarity between pitch contours, cluster analysis on pitch contours, pairwise tone-unit contour relations, and extending the analysis to other *banshi*, and to the extent of the entire corpus of 363 arias.

[1]Representative literature on Chinese Operas includes Pian 1972, Yung 1991, Wichmann 1991,Stock 1999.

[2]CompMusic Project seeks to advance the field of MIR by focusing on the computational processing of world music (Chinese, Indian, Turkish, Arab-Andalusian). It is funded by the European Research Council. PI: Xavier Serra.

[3]The 20 arias covers a representative range of melodic modes (*shengqiang*), role types and artists in Beijing Opera. In the pilot study, *yuanban* is chosen as it is the original banshi from which other types of rhythmic patterns are derived.

[4]It is hypothesized (Yu 2009, Pian 1972) that in the poetic structure of lyrics, the boundary or cadential positions (depending on their location in phrase 1 or 2 in a couplet of phrases) carry special musical significance, thus overriding similairty to linguistic tones. In this study we distinguish four types of positions: Phrase1 boundary (P1B), phrase 2 boundary (P2B), unit boundary (UB, first two of the three subdivisions in which a line is typically structured), and others. All boundaries are followed by instrumental interludes of various lengths.

[5]Automatic annotation and extraction tools are developed mainly with a combination of Praat Scripting Language, Java and Python tools. Pitch tracks are computed with MELODIA (Salamon and Gomez 2012) package.

[6]This subset is selected by removing the long tail of outliers of longer tone units so that all unit durations in the subset is represented by a normal distribution, with a mean and sd that are expected to better match the tone durations in speech.

References

[1] Gu,C. 2002. Smoothing Spline ANOVA Models. New York: Springer- Verlag.

[2] Pian, R.C. 1972. “Text Setting with the Shipyi Animated Aria.” In Words and Music: The Scholar’s View, edited by Laurence Berman, 237–270. Cambridge: Harvard University Press.

[3] Salamon, J and Gómez E. 2012. "Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics", IEEE Transactions on Audio, Speech and Language Processing, 20(6):1759-1770..

[4] Stock, J. 1999. “A Reassessment of the Relationship Between Text, Speech Tone, Melody, and Aria Structure in Beijing Opera.” Journal of Musicological Research (18:3): 183–206.

[5] Wichmann, E. 1991. Listening to the theatre: the aural dimension of Beijing opera. Honolulu: University of Hawaii Press.

[6] Xu,Z. 2007. “Jiu shi nian lai jingju de sheng diao yan jiu zhi hui gu”. (Overview of Ninety Years of Research of Tone Relations in Peking Opera). Nankai Journal of Linguistics, 10(2):39-50.

[7] Yu,H. 2009. Qiang ci guan xi yan jiu. (Study on the relationship between tone and melody). Beijing: Central Conservatory Publishing House.

[8] Yung,B. 1991. “The Relationship of Text and Tune in Chinese Opera.” In Music, Language, Speech and Brain: Proceedings of an International Symposium at the Wenner-Gren Center, Stockholm, 5-8 September 1990, edited by Johan Sundberg, Lennart Nord, and Rolf Carlson, 408–418. London: MacMillan Press.