Abstract

Melody is central to music perception. The melody can be defined as a monophonic sequence of music notes, the tune of a song that the listener remembers and would sing back when listening to a polyphonic music. There exist several algorithms that attempt to extract melody. However, there is a gap in knowledge about whether these state-of-art algorithms extract perceived melody. Nonetheless, the initial step to achieving this goal is to understand the melodic complexity within the existing algorithms. Therefore, a literature review was conducted to select two state of art melody extraction algorithms, one based on merely signal processing while the latter based on deep learning. A computational pipeline was implemented in this study to understand the melodic complexity, this pipeline iteratively removes the maximally correlated track with the extracted melody. A pilot behavioural analysis was also conducted with 9 participants to validate if the melodies extracted at each iteration from the computational pipeline are perceptually coherent. The framework that was resulted from this study can identify the different instrument tracks correlated with the extracted melody at each iteration. The results show that the extracted melody is a combination of various instrument tracks. A melodic contour was still extracted at different iterations even after removing the maximum correlated track, whether this is the perceived melody was tested in the pilot behavioural analysis. An interesting pattern emerged for the pop music pieces in the behavioural analysis, even though the signal based algorithm extracted melody in different iterations, the perceived melody decreased as the number of iterations increased. However, more behavioural data is required to make any statistically significant statement.