Boundaries and novelty: the correspondence between points of change and perceived boundaries

Jordan B. L. Smith, Ching-Hua Chuan and Elaine Chew

ABSTRACT

Data mining tasks such as music indexing, information retrieval, and similarity search, require an understanding of how listeners process music internally. Many algorithms for automatically analyzing the structure of recorded music assume that a large change in one or another musical feature suggests a boundary. However, this assumption has not been tested: while our understanding of how listeners segment melodies has advanced greatly in the past decades, little is known about how this process works with more complex, full-textured pieces of music, or how stable this process is across genres. Knowledge of how these factors affect how boundaries are perceived will help researchers to judge the viability of certain algorithmic approaches with different corpora of music.

We can gain some knowledge about this process by analyzing collections of ground truth descriptions of musical structure. These resources were originally created as tools to evaluate the performance of structural analysis algorithms, but we analyze these tools they have the potential to reveal patterns among listeners for a wide range of music. We present an empirical analysis of a large corpus of recordings whose formal structure was annotated by listeners, and find that the acoustic properties of boundaries in these recordings corroborate findings of previous perceptual experiments.

Nearly all boundaries correspond to peaks in novelty functions, which measure the rate of change of a musical feature at a particular time scale. Moreover, most of these boundaries match peaks in novelty for several features at several time scales. Figure 1 shows a histogram comparison of boundaries (gray) and non-boundaries (white) according to the number of novelty functions with a matching peak. Symmetric boundaries are those between sections with the same label (e.g., between two repetitions of a chorus). The fatter tail of the boundary histogram indicates that these have a greater chance of matching several novelty functions than do non-boundaries. Still, a great many boundaries match novelty functions no better than non-boundaries.

We observe that the boundary-novelty relationship can vary with listener, time scale, genre, and musical feature. Finally, we show that a boundary profile derived from a collection of novelty functions correlates with the estimated salience of boundaries indicated by listeners. Figure 2 compares boundary profiles estimated from annotations (solid line) and from novelty functions (dashed line) for the song "I Close My Eyes" by Shivaree. The correspondence between the highest novelty-function peaks and the annotations indicates that the more importance boundaries are reflected in a greater degree of novelty across the different features.

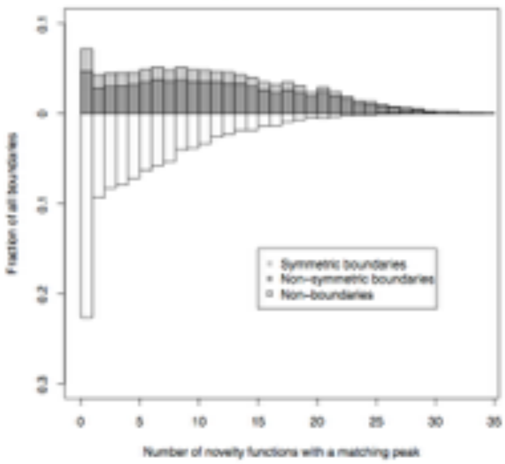

Figure 1. Histogram comparison of boundaries (gray) and non-boundaries (white) according to the number of novelty functions with a matching peak

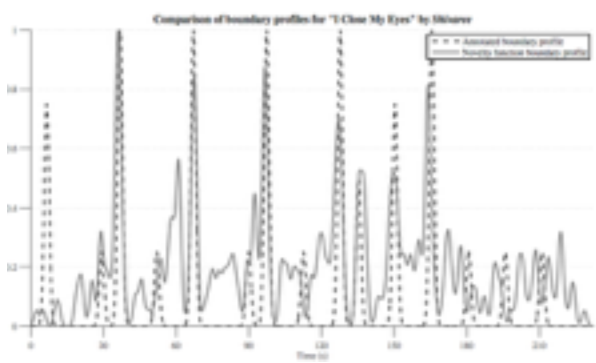

Figure 2. Comparison of boundary profiles estimated from annotations (solid line) and from novelty functions (dashed line) for the song "I Close My Eyes" by Shivaree.