Chapter 15
Acoustic Analysis of Music Albums

Kristoffer Jensen
Aalborg University Esbjerg, Denmark

ABSTRACT
Most music is generally published in a cluster of songs, called an album, although many, if not most people enjoy individual songs, commonly called singles. This study proposes to investigate whether or not there is a reason for assembling and enjoying full albums. Two different approaches are undertaken in order to investigate this, both based on audio features, calculated from the music, and related to the common music dimensions rhythm, timbre and chroma. In the first experiment, automatic segmentation is done on full music albums. If the segmentation is done on song boundaries, which is to be expected, as different fade-ins and –outs are employed, then songs are seen as the homogenous units, while if the boundaries are found within songs, then other homogenous units also exist. A second experiment on music sorting by similarity reveals findings on the sorting complexity of music albums. If the sorting complexity is high, then the albums are unordered; otherwise the album is ordered with regards to the features. A discussion of the results of the evaluation of the segment boundaries and sorting complexity reveals interesting findings.

INTRODUCTION
Music can be enjoyed on different time scales, going from the individual notes, to the riffs, as popularized in for instance ring tones, through choruses and full songs, which are popularized through the single format, and to albums, that many consider cannot be listened to other than at full length. This applies in particular to the concept albums. The investigations of albums will lead to the analysis of segmentation of music, to the analysis of sorting of music, and to the analysis of the theories of music perceptions.
Theories of what homogenous units are to be found in music can be found in the music theory, for instance by the grouping theory of Lerdahl & Jackendoff (1983). Results from memory research (Snyder 2000) can also be used as the ground reference. Snyder refers to echoic memory (early processes) for event fusion, where fundamental units are formed by comparison with 0.25 seconds, the short-term memory for melodic and rhythmic grouping (by comparison up to 8 seconds), and long-term memory for formal sectioning by comparison up to one hour. Snyder (2000) relates this to the Gestalt theory grouping mechanisms of proximity (events close in time or pitch will be grouped together). Proximity is the primary grouping force at the melodic and rhythmic level (Snyder 2000, p 40). The second factor in grouping is similarity (events judged as similar, mainly with respect to timbre, will be grouped together). A third factor is continuity (events change in the same direction, for instance pitch). These grouping mechanisms give rise to closure, that can operate at the grouping level, or the phrase level, which is the largest group the short-term memory can handle. When several grouping mechanisms occur at the same time, intensification occurs, which gives rise to higher-level grouping. Other higher-level grouping mechanisms are parallelism (repeated smaller groups), or recurrence of pitch. The higher-level grouping demands long-term memory and they operate at a higher level in the brain, as compared to the smaller time-scale grouping. The higher-level grouping is learned while the shorter grouping is not. Snyder (2000) further divides the higher level grouping into the objective set, which is related to a particular music, and the subjective set, which is related to a style of music. Both sets are learned by listening to the music repeatedly. Snyder (2000) also related the shorter grouping to the 7±2 theory (Miller 1956), that states that the short-term memory can remember between five to nine elements.

Recently, the chunk has been appointed as an important element of music (Kühl 2007, Godøy 2008). A chunk is a short segment of a limited number of sound elements, corresponding to the working memory of approximately 3 seconds. A chunk consists of a beginning, a focal point (peak) and an ending. Both Kühl and Godøy seems to believe that the chunk is fundamental in music, while Kühl mainly relates thechunking to the cognition, in particular the memory, Godøy also relates chunking to the action, i.e. physical gestures. Kühl (2007) extends the chunks to include microstructure (below 1/2 second), mesostructure (the present, approximately 3 seconds) and macrostructure (approximately. 30-40 seconds).

Automatic segmentation using dynamic programming has been proposed previously (Jensen et al 2005, Jehan 2005). In Jensen (2007), the dynamic programming is done of self-similarity matrices, created from the original features (rhythm, chroma or timbre) by comparing each time vector to all other time vectors. The dynamic programming will cluster the time vectors into segments, as long as the vectors are similar. By varying the insertion cost of new segments, segment boundaries can be found at different time scales. A low insertion cost will create boundaries corresponding to micro-level chunks, while a high insertion cost will only create few meso-level chunks. Thus, the same segmentation method can create segments of varying size, from short to long.

With the advent of downloaded music on personal computers, the necessity of assisting users choosing music among thousands of songs has arisen. Such a choice can be random (Shuffle play), by automatic playlist generation, or relate to a degree of similarity between songs. Playlist generation can be done on audio features (Foote 1997), for instance based on one song, or audio input, as in the query-by-humming systems (McNab 1996, Rolland et al 1999, Ghias et al 2001). Playlist generation can also be based on meta-data (Pauws & Eggen 2002), and collaborative filtering.
Related Content

Computer-Aided Detection and Diagnosis of Breast Cancer Using Machine Learning, Texture and Shape Features
www.igi-global.com/chapter/computer-aided-detection-diagnosis-breast/56174?camid=4v1a

Fuzzy Sets and Analytical Hierarchical Process for Manufacturing Process Choice
M. Reza Abdi (2008). Handbook of Computational Intelligence in Manufacturing and Production Management (pp. 63-82).
www.igi-global.com/chapter/fuzzy-sets-analytical-hierarchical-process/19353?camid=4v1a

Common Problems Faced When Developing Applications for Mobile Devices
www.igi-global.com/chapter/common-problems-faced-when-developing-applications-for-mobile-devices/121769?camid=4v1a

Translation of Biomedical Terms by Inferring Rewriting Rules
www.igi-global.com/chapter/translation-biomedical-terms-inferring-rewriting/56204?camid=4v1a