

Gabriel Chrisman

Literature Review for Jin Ha Lee, February-March 2010

The field of music similarity research spans the disciplines of information science, computer science and engineering, electrical engineering, cognitive psychology, hearing in health science, music theory, and communications. Researchers in these disparate fields have evolved significantly different approaches to the basic problems of the field, and I have attempted to identify the significant facets of this body of research in this literature review.

In searching for articles, I used the Academic Search Complete database, Web of Science, Google Scholar, and the website for the International Conference on Music Information Retrieval (ISMIR). I located 43 articles, the bulk of them published between 2002 and 2010, but also including a few outliers from the 1990s and a foundational earlier work that was frequently cited in the later articles (Tversky, 1977).

Reviewing the articles, I divided them into 7 groups which represent the major research approaches I have identified. I categorized the articles, allowing some articles to be included in two categories rather than forcing them into one. What follows is a summary of these categories, with explanations of the general research approaches taken by the authors in that category.

The largest category, with 16 articles, consists of computational comparison methods utilizing internal structural or signal data. This approach is characterized by data analysis of digital music files, and mathematical or algorithmic comparison of the results. The articles are authored by researchers from a range of backgrounds, including computer science, artificial intelligence, information science, electrical engineering, and communications. They tend to be highly technical, and require a high baseline of existing knowledge in order to understand their results. Major contributors in this category include Elias Pampalk, Jean-Julien Aucouturier, and Francois Pachet. The publication dates of articles that utilize this approach indicate that it emerged in the late 1990s, and articles with this approach have increased in number throughout the last decade. Several of the articles date from within the past year, and this aspect of the field appears active and growing. Most of these articles were published in conference proceedings: ISMIR, the International Workshop on Multimedia Information Retrieval, the MIREX Annual Music Information Retrieval eXchange, SIGIR, and the international ACM Multimedia Conference. The major journals represented in this category are *Signal Processing* and *Information Sciences*.

The next major category, with 7 articles, is the use of computational comparison methods which utilize metadata instead of internal signal data. These approaches usually apply data harvesting methods to community generated metadata sources on the internet, such as blogs, music sharing sites, and similar data environments. These articles are similar to the computational methods of the previous category, but differ from those methods in the data to which they apply these algorithms. Also, because the harvested metadata is significantly less complex than the acoustic data files, these articles tend to be more approachable and understandable to a wider range of potential audiences. The authors frequently acknowledge the more computational angle, but highlight the closer relationship of community metadata to subjective approaches, and the value of avoiding copyright issues by not requiring the use of the audio

data. They focus primarily on the end results for the user, and not as much on the value of certain algorithmic approaches over others. Authors come from a range of fields, including artificial intelligence, digital music, and information science, and there are also more articles published by researchers in commercial organizations (Yahoo labs, etc.). These articles were published between 2003 and 2009, with more publications towards the end of that period. Most of them come from the ISMIR proceedings, with one earlier article originating from the International Conference on Digital Audio Effects.

Another significant approach to music similarity research, and that was commonly referenced by authors in the previous category, is subjective listener similarity measurements. Essentially, these researchers are interested in assessing and understanding human judgments of music similarity, using both surveys and experiments. The research still tends to be focused on usable outcomes, and not purely on the theory of similarity. This category is also comprised of 7 articles, and it seemed to be an equally common approach when compared to the previous category. They date from the mid-2000s to the present. These researchers come from a wide range of fields, including musicology, computer science, applied mathematics, psychology, and more. They also come from a wide range of institutions, including both commercial and academic environments. ISMIR is well represented in this category, as well as the *Journal of Psychology*, the *Journal of New Music Research*, and the *Journal of the American Society for Information Science and Technology*.

A fourth category consists of researchers who are primarily concerned with comparisons between the three methods of assessment described above. I refer to this category as comparisons of different methods, and I located 6 articles that fell into this category. Again, the largest source for articles in this category was the ISMIR conference.

Another category was formed by five articles which focused on the psychology of similarity, and were largely based in cognitive psychology. Researchers in this category are frequently concerned with understanding the ‘why?’ of differing human music perceptions, and the focus of these papers is necessarily more theoretical. Authors in this category come from backgrounds in cognitive psychology, health sciences, and music psychology. Included in this category is a foundation work by A. Tversky, which was frequently cited by publications in all of the other categories.

There were a group of three articles which focused on the classification of genres, categorization and similarity. These were interested in learning how to leverage human methods of categorization to improve or understand either subjective or computational methods of determining music similarity. These articles date mostly from the late 2000s, and came exclusively from the ISMIR proceedings.

Comprising the final category were two articles on speech-music discrimination and music similarity. These very recent articles used computational approaches which are very similar to those described in the first category, but applied to a more basic level of identifying music compared with speech. The authors come exclusively from an artificial intelligence background. These articles may or may not represent a new direction for the field, but at the very least research on this facet of music similarity may allow for improved methods which can be utilized by the researchers publishing articles in the first category.

Overall, the multifaceted field of music similarity seems to have an active future, and there are many possibilities for future research in all of these categories. The greatest potential for researchers in

information science may be in the area of metadata harvesting and combining computational and subjective methods.