Melody Retrieval and Classification Using Biologically-Inspired Techniques

Dimitrios Bountouridis[1], Dan Brown[2], Hendrik Vincent Koops[1],
Frans Wiering[1], Remco C. Veltkamp[1]

[1] Department of Information and Computing Sciences, Utrecht University
[2] David R. Cheriton School of Computer Science, University of Waterloo

1   Abstract

Assessing the similarity between two musical items is a complex process that requires human intellect. In addition, Marsden [Marsden2012] argues that similarity involves interpretation, which by itself is a personal creative act. As such, similarity modeling using computational tools, naturally fits to the goals of Artificial Intelligence (AI). In addition, the retrieval and classification of music items is a fundamental problem in AI's musical counterpart Music Information Retrieval (MIR), and vital to the music recommendation task that underpins the ever-growing digital music industry. Assessing the similarity between two musical items is at the core of both tasks. However, despite the numerous proposed approaches, the problem is far from being solved. The exact mechanics of perceived similarity, is still unknown or incomplete which comes as no surprise considering music's inherent complex nature. The perceived similarity between two songs is known to be a *subjective* phenomenon, judgments of different individuals can vary significantly. The individual interpretation can be affected also by the *multidimensionality* of music, since similarity between two songs can be a function of timbre, melody, rhythm, structure, or indeed any combinations of those (or other) dimensions. To make matters worse, music similarity is known to be *contextual* thus depending on the circumstances of comparison.

Nevertheless, in the context of symbolically encoded melodies, the most widely used method for similarity assessment has been pairwise alignment via dynamic programming: gaps "-" that represent notes deleted from the melodic sequences are introduced in the sequences, until they have the same length and the amount of "relatedness" between notes at corresponding positions is maximized. Unfortunately, this method suffers from two major drawbacks: first, the dynamic programming technique is slow and fails to efficiently scale to large datasets. Secondly, the final alignment and similarity score are based on pre-defined set of relationships between notes, encoded as a fixed scoring matrix, and pre-defined gap penalties. This fixed scheme is globally applied across the melodies' length and on all pairs of melodies universally. This contradicts our intuition that certain salient parts of a melody are less likely to change than others in a melody rendition. In addition, the simplistic nature of the substitution matrix clearly cannot accommodate for the proper handling of music uncertainty

and variance (or stability), because the matrix focuses only locally on individual notes. This paper aims to enhance melody retrieval and classification by tackling the two aforementioned issues with the use of bioinformatic techniques. Musical and biological sequences are not as detached as one might think: even as early as the 1950's, it has been observed that they share a number of resembling concepts [Bronson1951].

With regard to efficient classification, we employ BLAST (Basic Local Alignment Search Tool) that uses heuristics to filter out unnecessary comparisons between the query sequence and the target database while it only explores a small part of the dynamic programming space by identifying high-matching substrings. On two datasets of folk and pop music or related melodies grouped into classes, we show that music-unaware BLAST can achieve high classification accuracy, comparable to music-aware methods that use alignment via dynamic programming. However, on the harder retrieval task, the performance of BLAST is relatively low due to its lack of incorporated domain knowledge.

We argue that representations that capture salient and robust properties of musical content, called *profiles* or *prototypes*, are a more intuitive approach to enhance and speed up retrieval. Therefore for our second experiment, we employ an approach with major contributions to the field of computational molecular biology [Durbin1998]: profile hidden Markov models (profile HMMs) [Eddy1998]. These are probabilistic automata that take a multiple sequence alignment (MSA) of related sequences and convert it into a position-specific scoring system that can be used for database searching. Profile HMMs are theoretically more appropriate for assessing the similarity between sequences because they provide a powerful framework for dealing with uncertainty and randomness in sequential data. The scoring and penalizing at each position is dependent on the MSA and not on some arbitrary chosen fixed values, e.g. scoring matrix, gap penalties. We use multiple sequence alignment techniques to reveal the locations of high stability or salience among related melodies and profile HMMs to model the MSA, or in other words, to model the intrafamily similarity. On our two datasets of folk and pop music, we show that profile HMMs outperform the typical alignment or other prototype methods in a retrieval scenario.

In summary, addressing the complexity of music similarity altogether is an almost impossible task. Nevertheless, our experiments show that BLAST and profile HMMs can be reliable and efficient solutions for large-scale melody classification and retrieval respectively, without the incorporation of musical heuristics. As such, they can be generalized to other types of music, datasets and real-life scenarios[3].

---

[3] This work appears with the same title in the Proceedings of the 6th International Conference on Computational Intelligence in Music, Sound, Art and Design (2017).

References

[Marsden2012] Alan Marsden. Interrogating melodic similarity: a definitive phenomenon or the product of interpretation? In: Journal of New Music Research 41.4 (2012), pp. 323 - 335.

[Durbin1998] Richard Durbin, Sean R Eddy, Anders Krogh, and Graeme Mitchison. Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge university press, 1998.

[Eddy1998] Sean R. Eddy. Profile hidden markov models. In: Bioinformatics 14.9 (1998), pp. 755 - 763

[Bronson1951] Bertrand H Bronson. Melodic stability in oral transmission. In: Journal of the International Folk Music Council 3 (1951), pp. 50 - 55.