Fuzzy Audio Similarity Measures Based on Spectrum Histograms and Fluctuation Patterns

Klaas Bosteels and Etienne E. Kerre

Fuzziness and Uncertainty Modelling Research Group
Department of Applied Mathematics and Computer Science
Ghent University, Krijgslaan 281 (S9), B-9000 Gent, Belgium
{Klaas.Bosteels, Etienne.Kerre}@UGent.be

Spectrum histograms and fluctuation patterns are representations of audio fragments. By comparing these representations, we can determine the similarity between the corresponding fragments. Traditionally, this is done using the Euclidean distance. In this chapter, however, we study an alternative approach, namely, comparing the representations by means of fuzzy similarity measures. Once the preliminary notions have been addressed, we present a recently introduced triparametric family of fuzzy similarity measures, together with several constraints on its parameters that warrant certain potentially desirable or useful properties. In particular, we present constraints for several forms of restrictability, which allow to reduce the computation time in practical applications. Next, we use some members of this family to construct various audio similarity measures based on spectrum histograms and fluctuation patterns. To conclude, we analyse the performance of the constructed audio similarity measures experimentally.

1 Introduction

Portable audio players can store several thousands of songs these days, and online music stores currently offer millions of tracks. This abundance of music drastically increases the need for applications that automatically analyse, retrieve or organize audio files. Measures that are able to express the similarity between two given audio fragments, are a fundamental component in many of these applications (e.g. [1, 2, 3, 4, 5, 6]). In particular, many computational intelligence methods for organizing and exploring music collections rely on such an audio similarity measure. The SOM-enhanced JukeBox presented in [6], which uses unsupervised neural networks to build "geographical" maps of music archives, is a noteworthy example.

Usually, audio similarity measures are constructed using a feature-based approach. The audio fragments are represented by real-valued feature vectors, and the similarity is calculated by comparing these vectors. We consider two types of feature vectors in this chapter: Spectrum histograms and fluctuation patterns. So far, the Euclidean distance has always been used for comparing feature vectors of these types. By identifying the feature vectors with fuzzy sets, however, the possibility arises to use fuzzy similarity measures for this task. In this chapter, we investigate this alternative approach.

2 Related Work and Motivation

The audio similarity measure introduced by Aucouturier and Pachet in [1], which can be regarded as an improvement of a technique by Logan and Salomon [7], is well-known in its field. This measure calculates the similarity between two given audio fragments by comparing mixtures of Gaussian distributions that model the spectral information in the fragments. Mandal and Ellis proposed a simplified version of this approach [8]. They use a single Gaussian to model the spectral information, and compute the distance between two of these Gaussians by means of the symmetric Kullback-Leibler divergence. Calculating the Euclidean distance between the spectrum histograms [4] derived from the audio fragments is an alternative spectral approach that is even easier to implement and compute. Nevertheless, the experimental evaluation in [9] indicates that this approach based on spectrum histograms can outperform the above-mentioned more complex techniques in some cases.

Fluctuation patterns, which were orginally called rhythm patterns [5], contain information that is complementary to spectral characteristics. Therefore, Pampalk combined a spectral audio similarity measure with the Euclidean distance between fluctuation patterns, and further optimized this combination by taking into account some additional information derived from the fluctuation patterns [3]. This led to the audio similarity measure that won the MIREX'06 (Music Information Retrieval Evaluation eXchange 2006) audio-based music similarity and retrieval task.[1]

Hence, both spectrum histograms and fluctuation patterns can be considered to be audio representations that play an important role in the current state of the art. Since the Euclidean distance has always been used to compare these representation so far, employing other approaches for the comparison is an interesting research direction that still needs to be explored. As mentioned in the introduction, we propose fuzzy similarity measures as alternatives for the Euclidean distance in this chapter. This does not add any unwanted complexity because many fuzzy similarity measures are very easy to implement and compute, and fuzzy similarity measures offer the additional advantage of being studied extensively and having very solid theoretical foundations.

[1] `http://www.music-ir.org/mirex2006`