

# An Industrial-Strength Content-based Music Recommendation System

Pedro Cano  
Music Technology Group  
IUA-Universitat Pompeu Fabra  
08003, Barcelona, Spain  
pcano@iua.upf.es

Markus Koppenberger  
Music Technology Group  
IUA-Universitat Pompeu Fabra  
08003, Barcelona, Spain  
koppi@iua.upf.es

Nicolas Wack  
Music Technology Group  
IUA-Universitat Pompeu Fabra  
08003, Barcelona, Spain  
nwack@iua.upf.es

## ABSTRACT

We present a metadata free system for the interaction with massive collections of music, the MusicSurfer. MusicSurfer automatically extracts descriptions related to instrumentation, rhythm and harmony from music audio signals. Together with efficient similarity metrics, the descriptions allow navigation of multimillion track music collections in a flexible and efficient way without the need of metadata or human ratings.

## Categories and Subject Descriptors

H.5.5 [Information Interfaces and Presentation]: Sound and Music Computing – *signal analysis, synthesis, and processing, systems.*

## General Terms

Algorithms

## Keywords

Music content management, music information retrieval, content-based audio retrieval, music recommendation.

Nowadays access to online music is possible by querying artist or song names (or other types of editorial data such as genre), or by browsing recommendations generated by collaborative filtering, i.e. using recommender systems that exploit information of the type "users that bought this album also bought this album." An obvious drawback of the first approach is that consumers need to know the name of the song or the artist beforehand. The second approach is only suitable when a number of consumers has heard and rated the music. This situation makes it difficult for users to access and discover the vast amount of music composed and performed by unknown artists which is available in an increasing number of sites and which nobody yet rated nor described. As an alternative, content-based methods are currently being pursued (see [www.music-ir.org](http://www.music-ir.org)).

We present the MusicSurfer, a content management system based on music descriptions computed directly from audio data. These descriptions account for several perceptually and musically meaningful dimensions. In our system, musical similarity relies first of all on the *Timbre*. Based on spectral characteristics of the audio signals, this dimension represents aspects of the instrumentation, type of voice as well as sound quality

Copyright is held by the author/owner(s).

*SIGIR '05*, August 15–19, 2005, Salvador, Brazil.

ACM 1-59593-034-5/05/0008.

characteristics. Rhythm is another important musical aspect and is represented by descriptions such as *Tempo*, *Meter* or *Rhythm patterns*. Our system also describes harmonic and melodic aspects of musical pieces as well as some other musical dimensions that include genre membership estimation, perceived loudness or the dynamics of a track. For a more complete description of the dimensions we refer to [1].

Evaluating the quality similarity metrics is a difficult task that can easily become highly subjective. Besides human evaluation, we assess our similarity measures in metadata prediction tasks, e.g. how good is the system at recognizing genre, detecting cover versions, identifying instruments or predicting mood. In the ISMIR Contest [2,3] a comparison of state-of-the-art algorithms for music description took place. One of the tasks was artist identification, which consisted in identifying the artist of a song (from 30 and 40 author sets). While the best participant systems scored one third (34%) over the 30-author test set, the system we present outperformed this results in almost a factor of two (60%). Moreover, the system has also been evaluated on a larger database test case consisting of 273,751 songs from 11,257 different artists. The experimental results yielded a 24% artist identification rate. This is comparable to the accuracy of the best state-of-the-art algorithm over a test set of 40 artists.

We will demonstrate the system online performance in query-by-example tasks on a musical repository of over a million songs: retrieval takes tenths of a second while extraction of the aforementioned set of musical features runs 20 times faster than playing time, both operations carried on using an off-the-shelf PC.

The system functionalities can be accessed and evaluated on a database of 5000 items from Magnatune (only this smaller legal database from <http://www.magnatune.com> is publicly available for copyright reasons). See <http://musicsurfer.iua.upf.edu>.

**Additional authors:** José Pedro Garcia, Jaume Masip, Òscar Celma, David García, Emilia Gómez, Fabien Gouyon, Enric Guaus, Perfecto Herrera, Jordi Maseguer, Beesuan Ong, Miguel Ramirez, Sebastian Streich and Xavier Serra.

This work is partially funded by the European Union to the SIMAC IST-FP6-507142 project (<http://www.semanticaudio.org>).

## 1. REFERENCES

- [1] <http://www.iua.upf.edu/mtg/publications.php>
- [2] [http://ismir2004.ismir.net/ISMIR\\_Contest.html](http://ismir2004.ismir.net/ISMIR_Contest.html)
- [3] Cano, P., Gómez, E., Gouyon, F., Koppenberger, M., Herrera, P., Ong, B., Serra, X., Streich, S., Wack, N. "ISMIR 2004 Audio Description Contest", in press.

