

Armonique: A framework for Web 2.0 audio archiving, searching, and metadata extraction

Bill Manaris, Thomas Zalonis, J.R. Armstrong, Brys Sepulveda, College of Charleston, Charleston, SC, USA

The cultural legacy of our society is being captured and increasingly preserved in digital transcriptions of audio, text, images, and video. Organizations ranging from national archives, to libraries, to museums, to Internet repositories all have to deal with massive amounts of digital material. This digital growth demands innovative ways of processing archival data and usable management tools, which can help users make sense of and navigate through large data collections to discover what they really want. In the area of data management, we present, Armonique, a Web2 framework for the management of audio material, including searching, archiving, and metadata extraction. This framework allows users to navigate large audio collections based solely on the similarity of audio content itself (as opposed to listening patterns/preferences of other users, as in systems like Pandora, Last.fm, and Genius). Within our framework, an archive of unnamed/untagged audio samples can be easily added into the system, and may have metadata extracted automatically (e.g., artist, genre, or timbre classification). The main advantages of using Armonique with large music collections is that (a) it requires no human pre-processing, and (b) it allows users to discover songs of interest that are rarely listened to and are hard to find otherwise. We have designed a web-based search and archiving portal designed for simplicity and usability. Our framework may be used by organizations that require music searches based on content (i.e., patterns within the music). It includes user search capabilities based on arbitrary MP3 file uploads (i.e., find songs similar to a given song), album art, audio streaming capabilities, title and artist information, a link to the media's page, and "find similar" search refinement. The framework and its search capabilities will be demonstrated at the conference (an early demo of the search mechanism is available at <http://www.armonique.org>). Our framework is based on many years of research involving the extraction of power-law based features from MP3 audio files. These features capture aspects of balance and proportion within audio content. In particular, they measure proportions of music-theoretic and other attributes, such as pitch, duration, melodic intervals, and chords; they also capture various proportions of timbre within FFT power spectra. Over the years we have performed numerous experiments (psychological and computational), which validate Armonique's ability to capture musical similarity. These include experiments with human subjects, where we have measured human emotional and physiological responses to the music chosen by the search engine. The results demonstrate that power-law metrics are very promising for content-based music querying and retrieval, as they appear to correlate with aspects of human emotion and aesthetics. These results will also be presented at the conference.