

Sound Retrieval and Searching Technique for Web-Based and Application Specific Search Engines

M.M.A.V. Thilakarathne and U.A.J. Pinidiyaarachchi

*Department of Statistics and Computer Science, Faculty of Science,
University of Peradeniya*

Recent improvements of multimedia technologies made legacy text based search techniques inefficient due to the fact that the content of those data units or files could not be represented in textual form. To solve this problem, Content Based Search techniques were introduced.

In this study, we propose an efficient method for audio data searching and retrieval where users can input queries as audio data (melody, sound slice etc.) and retrieve similar sound patterns from a reference index using an Artificial Neural Network (ANN).

An audio database containing 22 different audio patterns including male and female voices, sounds of musical instruments, classical and rock music etc. were created. A reference relation was created in a database storing a unique file ID and the reference to each audio file. Using these references, each audio file was retrieved and sliced into one second (1s) pieces and these slices were transferred into two wavelet domains, "Haar" and "Daubechies3" (Db3), with five scale levels to obtain the base feature matrix for analysis. This feature matrix was used as the input for a Feed Forward ANN by taking five means of five scale levels and four ratios between means that sum up to nine different feature values per second of audio data. These inputs and corresponding file IDs were passed as training data to the ANN. The query audio inputs were decomposed and transformed into a wavelet domain and used as the input to the trained ANN where the audio reference (file ID) that contains a similar pattern was produced as the output.

The results for each slice of audio data show an average of 71% of correlation with originally fed audio slices on both Haar and Db3 wavelets. Noise-added slices show 69% of similarity and scaled slices show 45% similarity on Db3 wavelets. Each test output was produced with maximum 5 s result latency.

It was observed that the proposed approach shows significantly accurate results and it could be used to search audio data efficiently for both web-based and standalone purposes with a certain tolerance limit. System improvements could be done by adding a Hopfield Network for preliminary classification and then proceeding with the current method or by converging multiple ANNs each of which is trained with different wavelet type.