Strictly ballroom analysis: Computers get to know their rumba from their cha-cha-cha

August 26 2009

Computer scientists in Taiwan have devised a neural network program that can successfully classify a computerized music file based on its beat and tempo. The system could be a boon for music archivists with large numbers of untagged recordings and for users searching through mislabeled mp3 libraries. Details of tests on ballroom dancing music are reported this month in the *International Journal of Intelligent Information and Database Systems*.

Mao-Yuan Kao and Chang-Biau Yang of National Sun Yat-sen University, in Kaohsiung and Shyue-Horng Shiau of the Chang Jung Christian University, in Tainan, explain that most music fans can put a tune into a particular genre even on a first listen. However, for archivists and others with large collections of unclassified music an automated approach that assigns the main genre to each tune would save a lot of time and effort.

Until now, there were two main approaches to classifying music - the Ellis and the Dixon methods - named for their inventors. These methods work reasonably well by analyzing the audio signal but Yang and colleagues hope to combine the strengths of each and to use a neural network to do the initial classification.

An artificial neural network, is a type of computer model that mimics the behavior of clusters of brain cells. The researchers "play" the music file to the neural network, which analyses the beat and tempo and outputs a general musical genre. Additional music files are played on

after the other and genres assigned.

In this initial learning phase the researchers correct the misses and feed the hits back into the neural network so that it builds up an audio profile of how different music files sound in each different genre. Once the neural network has been trained it can then classify a whole collection of music files.

The next step is then to use the Ellis and Dixon methods to further confirm the genre of each neurally classified group of music files. These methods using different signal processing approaches to "listen" to the music file and to determine the position of peaks that correspond to the musical beat. They can be used to estimate tempo and beat pattern.

The Taiwanese team has so far tested their approach on a collection of several hundred ballroom dance music files. Their system has classified different music styles, such as cha-cha-cha, jive, quickstep, and tango, with varying degrees of success, the cha-cha-cha being the most accurately categorized in tests of dozens of music files. The American rumba, unfortunately failed classification. Paradoxically, the neural network could successfully classify a Viennese waltz with relative ease, but not a standard waltz.

Overall, the neural network approach is never 100% accurate, but in all musical genres tested it outperformed both the Ellis and Dixon methods. They suggest that further training of the neural network with classical, jazz and pop music files will ultimately allow them to classify much more diverse music collections automatically.