Cognitively Based Music Information Retrieval

August 9, 2011
Program Information and Abstracts

Description:
The seminar will cover topics relevant to music information retrieval including musical similarity, crosscultural perception, and cognitive modeling of musical expectations.

Speakers:
Prof. David Huron, School of Music, Ohio State University
Prof. Geraint Wiggins, Goldsmiths, University of London
Mr. David Beckford, co-founder and CSO of waveDNA, Toronto, ON
Mr. Glen Kappel, CIO of *wave*DNA, Toronto, ON
Dr. Naresh Vempala, SMART Lab, Ryerson University

Schedule:

09:30 am-10:30 am	- Coffee/Tea/Snacks
10:30 am-10:45 am	- Introduction: Prof. Frank A. Russo, Director of SMART Lab
10:45 am-11:45 am	- Talk 1: Prof. Geraint Wiggins
11:45 am-12:45 pm	- Talk 2: Prof. David Huron
12:45 pm-02:00 pm	- Lunch followed by tour of the SMART lab
02:00 pm-03:00 pm	- Talk 3: Prof. Geraint Wiggins
03:00 pm-03:45 pm	- Talk 4: Mr. David Beckford/Mr. Glen Kappel, *wave*DNA
03:45 pm-04:00 pm	- Coffee
04:00 pm-04:45 pm	- Talk 5: Dr. Naresh N. Vempala, SMART Lab
04:45 pm-05:00 pm	- Concluding remarks: Dr. Naresh N. Vempala

Talk 1: Prof. Geraint Wiggins

What's Cognitive about Music Information Retrieval?

In this talk, I consider what it might mean to study music information retrieval (MIR) from a cognitive perspective. Information retrieval falls broadly into two categories: that which is done on the basis of content, for example, when Google searches for a phrase in a text document; and that which is done on the basis of meta-data, for example, when we look up a book using a known shelf-mark. Music information retrieval too can make this distinction (though of course, there is an intermediate point where we retrieve using meta-data derived from content).

The question begged, then, is how to analyze musical signals or scores in such a way that retrieval done using the resulting data is meaningful to human listeners. This question is more difficult than it might appear, since not all of musical meaning is present in the audio signal – indeed, one might argue that the audio signal is merely a stimulus that gives rise to musical meaning. Substantial progress can be made using surface-form matching; but this does not generally explain the core question of why different pieces of music “sound similar”.

This talk will involve audience participation, though no one will be asked to sing.

Talk 2: Prof. David Huron

Session 1: What Did They Know? And When Did They Know It? (20 minutes)

Listeners are able to extract a tremendous amount of information from very brief listening experiences. They can recognize sound sources, decipher meter, tempo, mode, and texture, process lyrics, and dynamics, identify style and genre, perceive performance nuance, and apprehend emotional character. Two empirical studies focus on the first 3 seconds of the listening experience, and chronicle when the various types of information become available to consciousness. A third ("name-that-tune") study similarly chronicles the point-of-recognition for well-known tunes.

Session 2: Where Do We Go From Here? Arts and Entertainment in the 21st Century (40 minutes)

For two millennia, music scholars have asked "What is good music?" and a related question: "What makes for a good musical culture?" Until the 1960s, these questions had a ready answer in the defense of high-brow classical music from the presumed corrupting effects of pop music. Those days have thankfully gone. However, there remains the unfinished business of addressing the fundamental question: What makes for a good musical culture? For example, can people listen to too much music? Is copyright good or bad? Can we minimize the destructive cultural effects of globalization? With the rapid changes in technology, and the effects of globalization, such questions are arguably more pressing now than they have ever been. Unfortunately, unlike our counterparts in Education, Social Work, Law, Engineering, or Medicine, arts and humanities scholars have virtually no expertise in policy. Even if lawmakers asked our opinions, we have little coherent to say, no concrete policy advice, and what we say has virtually no research basis to support one view over another. At a time when special interests are shaping legislation whose repercussions will influence musical culture for generations, who is representing the public interest? How do we maximize the benefits of the new musical order?

Talk 3: Prof. Geraint Wiggins

How Cognitive Models Can Help with Musical Similarity

In this talk I will present a cognitive model, which is capable both of pre-processing symbolic (i.e. note-based) representations of music, to prepare them for use in music information retrieval, and also to perform similarity analysis in its own right. I will illustrate this via a rather unusual example: applying the paradigmatic analysis technique of Nicolas Ruwet to Claude Debussy’s flute piece, *Syrinx*.

Notwithstanding the discrete nature of the data used here, the mathematical principles involved, based in information theory, are applicable also to continuous musical features, and so may be applied directly to audio recordings.

Talk 4: Mr. David Beckford (co-founder and CSO)/Mr. Glen Kappel (CIO), *wave*DNA

Towards Cross-over Toolkits for MIR and Music Production

We will present a novel MIR framework and discuss its potential applications in (a) music production environments, and (b) as a tool-set for academic research in empirical musicology and music cognition. Topics touched on will include issues and problems with traditional/current music representation systems in the context of computational analysis, and the potential benefits of a language/system that could facilitate ongoing interdisciplinary research, including validation studies and improvements of the model/system itself. The talk will include brief software demonstrations that will serve to suggest some of these possibilities.

Talk 5: Dr. Naresh N. Vempala, SMART Lab

Session 1: The Effect of Culture on Rhythmic Perception

When listening to a piece of music, listeners use their existing mental template as shaped by their musical experience. Their enculturated representation of rhythm interacts with available sensory information to influence the perception of structure. Our goal was to understand how musical enculturation could affect a listener's sensitivity to rhythmic structure conveyed by intensity accents. We examined goodness-of-fit judgments of probes in different rhythmic contexts across two different cultural groups, Canadians and Ecuadorians. We presented rhythmic stimuli comprising symmetric and asymmetric rhythmic groupings. Our hypothesis was that culture would influence perception of rhythmic groupings leading to an interaction of probe with culture. Because Canadians would have predominantly been enculturated to symmetric groupings, we predicted them to show sensitivity to the intensity accent structure present in symmetric rhythmic stimuli. In contrast, because of their enculturation to asymmetric rhythmic groupings, we predicted that Ecuadorians would show sensitivity to the intensity accent structure present in asymmetric rhythmic stimuli. Our results showed that Canadian but not Ecuadorian participants were more sensitive to surface structure in symmetric rhythmic groupings than to asymmetric rhythmic groupings. Our results indicated the strong effect of internalized rhythmic schemas as a product of enculturation, in perceiving rhythmic structure. Based on the results of this study, we propose an initial set of rules for developing a theoretical model of rhythmic perception that focuses on the influence of previous enculturation.

Session 2: A Cognitively Informed Method for Measuring Melodic Similarity

We describe an empirical method for determining the similarity between two melodies, a standard and a comparison using multiple linear regression. *Pitch distance, pitch direction, tonal stability, rhythmic salience,* and *melodic contour* were used as the five predictors. Eight standard melodies, four in major and four in minor, were created. Twelve comparison melodies were created for each standard by systematically manipulating one note with respect to pitch distance, pitch direction and rhythmic salience. Our regression analysis showed that for non-transposed comparisons, pitch distance, pitch direction and melodic contour were the similarity determining predictors, while for transposed comparisons, tonal stability and melodic contour were the similarity determining predictors. We believe that the implementation of a cognitively informed method, such as this, can be scaled to account for more than single note changes and that it has potential to extend the domain of similarity-finding methods by instantiating human-like characteristics.