journal of interdisciplinary music studies
fall 2007, volume 1, issue 2, art. #071201, pp. 1-24

Computational Ethnomusicology

Hesaplamalı Etnomüzikoloji

George Tzanetakis[1], Ajay Kapur[2], W. Andrew Schloss[1], Matthew Wright[1]

Department of Computer Science and School of Music, University of Victoria[1]
School of Music, California Institute of the Arts[2]

Abstract. John Blacking said "The main task of *ethno*musicology is to explain music and music making with reference to the social, but *in terms of the musical* factors involved in performance and appreciation" (1979:10). For this reason, research in ethnomusicology has, from the beginning, involved analysis of sound, mostly in the form of transcriptions done "by ear" by trained scholars. Bartók's many transcriptions of folk music of his native Hungary are a notable example.

Since the days of Charles Seeger, there have been many attempts to facilitate this analysis using various technological tools. We survey such existing work, outline some guidelines for scholars interested in working in this area, and describe some of our initial research efforts in this field. We will use the term "Computational Ethnomusicology" (CE) to refer to the design, development and usage of computer tools that have the potential to assist in ethnomusicological research. Although not new, CE is not an established term and existing work is scattered among the different disciplines involved.

As we quickly enter an era in which all recorded media will be "online," meaning that it will be instantaneously available in digital form anywhere in the world that has an Internet connection, there is an unprecedented need for navigational/analytical methods that were entirely theoretical just a decade ago. This era of instantaneously available, enormous collections of music only intensifies the need for the tools that fall under the CE rubric.

We will concentrate on the usefulness of a relatively new area of research in music called Music Information Retrieval (MIR). MIR is about designing and building tools that help us

Özet. John Blacking, "Etnomüzikolojinin temel görevi müziği ve müzik yapmayı toplumsal olana referansla, ancak icra ve temellük bağlamındaki müziksel etkenler açısından açıklamaktır." demiştir (1979:10). Bu nedenle başından bu yana etnomüzikolojideki araştırmalar, daha çok eğitimli araştırmacılar tarafından "kulakla" notaya dökülerek temsil edilen sesin analizini içermiştir. Bartok'un ülkesi Macaristan'ın halk müziği için yapığı bir çok nota yazımı dikkate değer bir örnektir.

Charles Seeger'in zamanından bu yana bu analizi kolaylaştırmak için çeşitli teknolojik araçların kullanılması yönünde girişimler olmuştur. Bu tür çalışmaları gözden geçirirken, bu alanda çalışmaya meraklı araştırmacılar için bazı kılavuz niteliğindeki bilgileri anahatlarıyla sunuyor ve bu alandaki kendi ilk araştırma tecrübelerimizin bir kısmını aktarıyoruz. Etnomüzikoloji araştırmalarına yardım potansiyeli olan bilgisayar araçlarının tasarımı, geliştirilmesi ve kullanımı için "Hesaplamalı Etnomüzikoloji"(CE) terimini kullanacağız. Yeni olmamasına karşın CE oturmuş bir terim değildir ve varolan çalışmalar farklı disiplinlere dağılmış durumdadır.

Hızlı bir biçimde tüm kayıtlı ortamın çevrimiçi hale geleceği, yani internet bağlantısı olan dünyanın herhangi bir yerinden bu verilere anında sayısal bir biçimde ulaşılabilir olacağı bir çağa girdiğimizden, sadece bir on yıl önce tamamen kuramsal olan analitik yöntemler için görülmedik bir ihtiyaç vardır. Devasa müzik kolleksiyonlarının anında erişebilir olduğu bu çağ, CE başlığı altına giren araçlara olan ihtiyacı yalnızca yoğunlaştırmaktadır.

Müzik Bilgi Geri Getirim (MIR) adı verilen göreli olarak yeni bir müzik araştırma alanının kullanışlılığı üzerinde yoğunlaşacağız. MIR

•*Correspondence:* George Tzanetakis, Department of Computer Science, University of Victoria, Engineering/Computer Science Building (ECS), Room 504, PO Box 3055, STN CSC, Victoria, BC, Canada V8W 3P6; phone: +1 250 472-5711; fax: +1 (250) 472-5708; e-mail: gtzan@cs.uvic.ca

organize, understand and search large collections of music, and it is a field that has been rapidly evolving over the past few years, thanks in part to recent advances in computing power and digital music distribution. It encompasses a wide variety of ideas, algorithms, tools, and systems that have been proposed to handle the increasingly large and varied amounts of musical data available digitally. Researchers in this emerging field come from many different backgrounds including computer science, electrical engineering, library and information science, music, and psychology. The technology of MIR is ripe to be integrated into the practice of ethnomusicological research. To date, the majority of existing work in MIR has focused on either popular music with applications such as music recommendation systems, or on Western "classical" music with applications such as score following and query-by-humming.

In addition, as microchips become smaller and faster and as sensor technology and actuators become cheaper and more precise, we are beginning to see ethnomusicological research incorporating both robotic systems and digital capture of music-related bodily gestures; music in general is embodied and involves more than a microphone can record. Our hope is that the material in this paper will help motivate more interdisciplinary and multidisciplinary researchers and scholars to explore these possibilities and solidify the field of computational ethnomusicology.

büyük müzik kolleksiyonlarının düzenlenmesi, anlaşılması ve sorgulanmasında bize yardımcı olan araçların tasarlanması ve gerçekleştirilmesi ile ilgilidir. Bir anlamda hesaplama gücü ve sayısal müzik dağıtımındaki son ilerlemeler sayesinde MIR son birkaç yılda hızla gelişmiştir. MIR, sayısal ortamda giderek büyüyen ve çeşitlenen müzik verisinin ele alınabilmesi için önerilmiş olan çok çeşitli düşünceler, algoritmalar ve sistemleri kapsar. Ortaya çıkan bu alandaki araştırmacılar, bilgisayar bilimleri, elektrik mühendisliği, kütüphane ve enformasyon bilimi, müzik ve psikoloji dahil bir çok farklı disiplinden gelmektedir. MIR teknolojisi etnomüzikoloji araştırma pratiğinde kullanılmak için olgunlaşmıştır. Bugüne kadar MIR alanındaki çalışmaların çoğunluğu ya müzik öneri sistemleri gibi uygulamalarla popüler müzik üzerine ya da partisyon takibi ve mırıldanarak sorgulama gibi uygulamalarla Batı "klasik" müziği üzerine odaklanmıştır.

Ayrıca, mikroyongalar daha küçük ve hızlı, algılayıcı teknolojisi ve erişim düzenekleri daha ucuz ve hassas hale geldikçe etnomüzikoloji araştırmasında hem robotik sistemlerin hem de müzikle ilişkili bedensel jestlerin sayısal olarak tespit edilmesinin kullanıldığını görmeye başlıyoruz. Genel olarak müzik bir mikrofonun kaydedebileceğinden çok daha fazlasını kapsamaktadır. Bu makaledeki malzemenin daha fazla multidisipliner ve disiplinlerarası araşırmacıyı bu olanakları keşfetmek ve hesaplamalı etnomüzikoloji alanını güçlendirmek için motive edeceğini umuyoruz.

Keywords: Ethnomusicology, music information retrieval, automatic transcription, musical gesture, human-computer interface, musical robotics

Anahtar kelimeler: Etnomüzikoloji, müzik bilgi geri getirimi, otomatik notaya ökme, müziksel jest, insan-bilgisayar arabirimi, müzik robotbilimi

1 Introduction

What is now called "computer music" began in the 1950's with the synthesis of sound using computers at Bell Laboratories. During the same decade, musicologist and researcher Charles Seeger made deeply perceptive predictions about the analysis of sound that are only now bearing fruit. Seeger was one of the first researchers in musicology to investigate electronic means of analysis and transcription of orally transmitted music, and the Seeger *Melograph* was one of the earlier attempts to create a graphical representation of sound for musical research (Seeger 1951). He was fifty years ahead of his time; only now, using an array of digital techniques that we consider to be under the general term "Computational Ethnomusicology" (CE), can we begin to realize his dreams, and indeed go beyond them by analyzing and also transforming recorded musical source materials. We therefore consider CE to be the design, development and usage of computer tools that have the potential to assist in ethnomusicological research. Although not new, CE is not an established term and existing work is scattered among the different disciplines involved.

The techniques of Music Information Retrieval (MIR) are particularly useful, powerful, and ripe for applications in this domain. The MIR community has been, for the past decade, designing and building tools that help us organize, understand and search large collections of music. Historically, the majority of work in MIR has focused on either popular music with applications such as music recommendation and personalized radio systems or on Western "classical" music with applications such as score following and query-by-humming. In this paper, we explore the application of these ideas and techniques to the study of non-Western music for which there is no standardized written reference (which is a large percentage of the music of world if not of album sales). In many cases the relevant work is preliminary and proof-of-concept without yet having matured enough to have impact in actual musicological research, hence we include the phrase "*potential* to assist" in our definition of CE. Our main goal is to survey existing work in CE both inside and outside the rubric of MIR, provide some guidelines for researchers interested in exploring it, and describe some specific concrete examples highlighting our ideas.

1.1 Musicology, Ethnomusicology, Comparative Musicology, Systematic Musicology, Empirical Musicology

The discipline now known as "ethnomusicology" was originally called "comparative musicology" (or, in German, *vergleichende Musikwissenschaft*). Cook explains the change:

> The middle of the twentieth century saw a strong reaction against the comparative methods that played so large a part in the disciplines of the humanities and social sciences in the first half of the century, and musicology was no exception. The term 'comparative musicology' was supplanted by 'ethnomusicology,' reflecting a new belief that cultural practices could only be understood in relation to the particular societies that gave rise to them… Perversely, this meant that the possibility of computational approaches to the

> study of music arose just as the idea of comparing large bodies of musical data – the kind of work to which computers are ideally suited – became intellectually unfashionable" (Cook 2004: 103).

The term "ethnomusicology" is problematic in many ways. What we mean by it is really the study of all music, for which the accurate term would be simply "musicology." Unfortunately for us, those who study European and European-derived art music traditions have already claimed the term "musicology". In order to include other musics we must therefore add the prefix "ethno-", implying either that we are studying "ethnic" music, whatever that means (does it mean that Beethoven didn't have an ethnicity?)[1], or that we are using ethnographic methods borrowed from anthropology, particularly fieldwork. Ethnographic methods can be valuable for the study of the aforementioned art music traditions just as for other music (Stock 2004), while at the same time we can gain insight about non-Western music by studying recordings without doing any ethnographic research (as we hope to show with some of the examples in this paper). We therefore use "ethnomusicology" to mean "the study of all the world's musics," without implying any particular methodology.

Table 1. Naïve view of some of the distinctions between Musicology and Ethnomusicology

Discipline	*Musicology*	*Ethnomusicology*
Music studied	"Notated music of Western cultural elites" (Parncutt 2007: 4)	Everything else
How the music studied is transmitted	Notation	Oral transmission
Methodology	Analysis of scores and other documents	Fieldwork, ethnography

Table 1 summarizes a simplistic idea of the distinction between Musicology and Ethnomusicology. Although in general it reflects the common usage of the terms, close scrutiny calls almost every detail into question. For example, one important technical issue that strongly affects how we can use computers to study a given kind of music is the question of notation. There are many non-Western cultures with their own indigenous notation systems, e.g., Chinese, Indian, and Indonesian; furthermore, musicians from many non-Western cultures, e.g., Turkish, Iranian, and Arabic, adopted Western notation (sometimes with slight modifications such as extra accidentals for notes outside the 12-tone chromatic scale) over a century ago (see Marcus 1989: 123-142). We can study music with or without a score by directly examining audio recordings from a signal processing and acoustics perspective. This could include the generation of visual representations of the audio material, including attempts at *automatic transcription* (see section 2.2) into some form of written notation. For music that does have a score, we can also study just the score itself (which is the easiest and therefore most common use of computers in music analysis),

---

[1] Blacking put it this way: "We need to remember that in most conservatories they teach only one particular kind of ethnic music, and that musicology is really an ethnic musicology" (1973: 3).

or the relationship between the score and one or more performances, e.g., study of intonation or expressive timing.

We also want to acknowledge the term *systematic musicology*, as used by Parncutt to refer to "subdisciplines of musicology that are primarily concerned with music in general, rather than specific manifestations of music" (2007: 1). From a computational point of view, the distinction between "music in general" and "specific manifestations" is just a question of amount of data, which brings us to *empirical musicology*, "a musicology that embodies a principled awareness of both the potential to engage with large bodies of relevant data, and the appropriate methods for achieving this" (Cook and Clarke 2004: 5).

1.2 Ethnomusicology and Technology

The origin of ethnomusicology is variously attributed to Carl Stumpf's 1886 study of Bella Coola Indian songs or Alexander Ellis' 1885 quantitative descriptions of musical scales (Nettl 1964). Less than 20 years later ethnomusicology started to rely on technology, especially various methods of audio recording. The *phonogram*[2] was one of the earliest portable sound-recording technologies, used as early as 1901 for field research in Croatia, Brazil, and on the isle of Lesbos in Greece. Practical field recordings radically changed the field of ethnomusicology by providing a way to preserve sound and music beyond a particular location and time. In 1928 Von Hornbostel, one of the founders of ethnomusicology, wrote "As material for study, phonograms are immensely superior to notations of melodies taken down from direct hearing; and it is inconceivable why again and again the inferior method should be used" (Von Hornbostel 1928: 32, quoted by Carterette and Kendall 1999). Bartok agreed: "The only true notations are the sound-tracks on the record itself" (Bartok and Lord 1951: 7). These points are correct in the sense that audio recordings contain much more information than transcriptions and also because the process of musical transcription is in many ways inherently subjective; on the other hand, visual representations of musical material have certain advantages over audio recordings, a point we will take up in our discussion of automatic transcription in section 2.2.

1.3 Music Information Retrieval

Music Information Retrieval (MIR) is an emerging interdisciplinary research area that encompasses all aspects of accessing digital music material. Its recent increase in visibility and prominence reflects the tremendous growth of music-related data available and the consequent need to search within it to retrieve music and musical information efficiently and effectively. During the past six years a variety of MIR problems have been identified and techniques for solving them have been proposed, including query-by-humming, automatic musical genre classification, structural analysis, computer accompaniment, score following, and tempo tracking. Researchers in this emerging field come from many different backgrounds. These include

[2] http://www.pha.oeaw.ac.at/phawww/geschichte_e.htm

computer science, electrical engineering, library and information science, music, and psychology (Futrelle and Downie 2002). They all share the common vision of designing and building tools that help us organize, understand and search large collections of music. As we quickly enter an era in which all recorded media will be "online" (meaning that it will be instantaneously available in digital form anywhere in the world that has an Internet connection) there is an unprecedented need for navigation/analytical methods that were entirely theoretical just a decade ago.

To make the ideas of MIR more concrete we briefly describe two scenarios based on ideas and algorithms that have been proposed in the literature. In query-by-humming, the user sings or hums a melody or melodic fragment (Dannenberg et al. 2004). The system then searches a database of songs and returns to the user the songs that contain the given melodic pattern. More specifically, the sung query is automatically analyzed using pitch extraction algorithms and is converted to a symbolic representation of pitches and durations similar to a score. All the songs in the database are also converted into a similar representation and then efficient search algorithms are used to find the closest matching song to the query. Challenges include variations and errors in the singing of the query, invariance to transposition and rhythmic variation, and scalability to large datasets. The second scenario we describe is content-based similarity retrieval. In this scenario, the user tries to find music that is "similar" to particular piece of music (the *query*). The only information used as "input" to the system is the actual recording. By analyzing information related to timbre, pitch distribution and rhythmic characteristics, the system can identify other pieces of music that are similar in various ways to the query.

MIR techniques utilize state-of-the-art signal-processing, machine-learning and perception-modeling algorithms. Because MIR is a relatively new research area, existing systems are still not ready for general use and require some technical expertise to set up and use. However, they do offer unique capabilities to the researchers who decide to use them. The two main advantages they offer, which are particularly relevant to CE, are:

- The ability to accurately detect and analyze information that would be difficult to do by hand. For example by automatically extracting onsets from different recordings of a particular rhythm one can analyze minute variations in timing between different performers. Although theoretically the onset information could be collected manually, the process would be very tedious and probably error prone.
- MIR algorithms scale to large datasets. For example, a musicologist can easily find occurrences of a particular melodic/rhythmic pattern "by hand/ear" in a small collection of up to approximately one hundred folk songs. Although an automatic system might not perform as accurately as a human, it has the advantage that it can find occurrences of the pattern in huge collections of thousands or even millions of songs.

The vast majority of existing work in MIR has focused on Western music. Computer technology itself has no intrinsic bias towards Western music, but since using advanced computer tools to study music tends to happen in universities, and since universities tend to be oriented towards European and European-derived art music traditions, these biases have crept into the field. We hope that this paper will illustrate the potential of using computer technology for the study of non-Western music.

2 Survey of Existing Work

Although CE is not an established term, there is existing work that fits our definition. In this section we collect and survey some of this existing work, which tends to be scattered among the different disciplines involved. Although by no means complete, we have tried to make the survey balanced and comprehensive. All presented works share the common thread of potential application to non-Western music that relies on aural transmission.

2.1 Work Based on a Score or Other Symbolic Notation

Some ethnomusicologists have analyzed the structure of symbolic representations of musical material using the notion (from linguistics and computer science) of a grammar that generates stylistically "correct" sequences of symbols, for example, Baily's (1989) "motor grammar" for right-hand rhythmic patterns for the Afghan Rubâb.

The ethnomusicologist James Kippen teamed up with the computer scientist Bernard Bel to look for and test grammatical rules for generating stylistically correct variations of compositions for the *tabla* drums (Kippen and Bel 1992). They brought an Apple IIc computer to a master *tabla* player to try a unique form of fieldwork: the computer took the role of the student, "learning" the unspoken rules of the musical style by example and then "improvising" new sequences which were then tested for "correctness" and "goodness" by checking them with the teacher. This process was iterative: generated sequences judged "incorrect" or "bad" resulted in modifications to the set of grammatical rules, as did new examples dictated by the teacher. "[A] prominent aim of the research has been to create a human-computer interaction where musicians themselves respond to the output of [these] grammars, and grammars are in turn modified to account for the input of musicians. Thus, in theory at least, analytical control over the models lies with the musicians, and it is they, not us, who are the sole arbiters of the correctness of computer-generated data" (Kippen and Bel 1992: 17). There already exists a set of spoken quasi-onomatopoeic *bol* syllables for the *tabla* (such as "dha", "tin", "ge", and "na") each corresponding to a possible combination of strokes for the two hands, so they were able to use this symbolic representation for the "words" of their grammars. Given a fixed *tabla* composition and one or more examples of "correct" variations, they hand-coded grammatical rules that would allow the transformation of the input composition into each of the variations. In later work they used *inductive inference*, a machine-learning technique, to generate the grammatical rules directly from the information provided by their informants, eliminating the "dependen[ce] upon the skill and intuition of the analysts in making generalisations and inferring knowledge without strict empirical justification" that characterized their earlier work (Bel and Kippen 1989: 3).

Lieberman encoded a large collection of Indonesian *gamelan* music scores, consisting of his transcriptions of field recordings of 38 *gendèr* improvisations and 82 representative pieces from traditional notation of *Djogjakarta kraton* (palace) music. Once these data were in machine-readable form he was able to compute various statistics on the use of pitch material as well as search for patterns such as cadential formulae (Lieberman 1970).

A research group from Finland (Toiviainen and Eerola 2001) started with a symbolic score representation of 6,252 mainly Germanic-region folk songs from the Essen collection plus 2,226 folk songs from Northern China. They computed histograms for pitch class and score duration of each note, melodic interval between each successive pair of notes, and also for the transitions for these three parameters. Then they used a form of artificial neural network called the *self-organizing map* to translate from the high-dimensional space of these input histograms to a two-dimensional array, with the property that similar melodies are mapped to nearby spaces in the output. Their envisioned applications include finding similar melodies from geographically distant areas.

A large variety of tools for such types of symbolic data have been developed for the analysis of Western music. The ethnomusicological application of such tools however is limited to music traditions in which such symbolic representations are meaningful. For the remainder of this paper we focus more on tools and computational approaches that are not adaptations of existing tools for western music.

2.2 Work Related to Automatic Transcription

In its broadest sense, *automatic transcription* is any process that converts recorded sound into a symbolic and/or visual representation with no or relatively little work on the part of a human supervisor. Naturally the institutional biases in the study of music mentioned in section 1.1 have affected the history of automatic transcription by privileging the output of Western music notation. Aside from the many possible discrete symbolic representations of musical material, another large category is graphs in which one axis is time, which can avoid some of the thorniest technical and epistemological issues in discretizing musical sound.[3] Well-crafted visual representations of any type have advantages over audio recordings such as being able to "read" them faster than real-time and the ease of visual comparison (Tufte 1990).

Cooper and Sapiro (2006) provide a survey of early analog technologies for automatic transcription used in ethnomusicology including such amazing contraptions as the array of 54 tuning forks each separated by 4 Hz from its neighbors created in 1834 by the German acoustician Johann Heinrich Scheibler. In 1951 Seeger described a "melograph" capable of taking audio input and outputting a graph of pitch against time "traced upon suitable graph paper" for "a single melodic line"; unfortunately it appears that this project was not completed. Askenfelt (1976) described a functioning computer-based notation system designed for transcription of monophonic Swedish folk melodies, consisting of a hardware "rest detector" and "pitch extractor" front-end, graphical display of a *fundamental frequency histogram* for the operator to select peaks corresponding to notes of the scale (outputting these frequencies in cents relative to the tonic), graphical display of a *duration histogram*

---

[3] The intersection of these two sets is the category known as *proportional notation*, in which the discrete symbols are spaced horizontally according to the exact time placement of the sonic events they represent.

for the operator to define the duration ranges for each notated duration value, and outputting printed Western music notation (without bar lines and always in the key of G major or G minor). Schloss (1987) described the "digital melograph" in one of the first presentations of digital technology and computation to the ethnomusicology community.

Klapuri (1998) reviewed the history of automated transcription by computer (without any focus on ethnomusicology). In recent years there have been a few automatic transcription systems optimized for specific non-Western musics. Nesbit et al. (2004) describe a system designed for the automatic transcription of Australian aboriginal music. The automatic labelling of *tabla* signals using an approach inspired by speech recognition techniques is described in (Gillet 2003). Chordia (2006) also built a transcription system for *tabla*, using a variety of signal-processing and machine-learning techniques and incorporating domain-specific prior knowledge such as patterns of strokes that are ergonomic at high speed.

New technologies for automatically producing visual representations of recorded sound hold great promise for ethnomusicology, such as the *Rhythmogram* (Todd 1994) and the *Beat Spectrum*, *Beat Spectrogram* and *self-similarity matrix* (Foote and Cooper 2001). Cooper et al. (2006) surveyed other such visualizations.

2.3 Work Based Strictly on Audio

2.3.1 Pitch Analysis

Arom and Fürniss brought a DX7 digital synthesizer to Central Africa to investigate the pitches of the pentatonic scale of the contrapuntal vocal polyphony of the Aka Pygmies. "Such abstract concepts as musical scales… are not subject to verbalization" in this culture (Arom and Fürniss 1993: 7), so these researchers rendered selections of Pygmy vocal music on the synthesizer using various collections of pitches, then asked for feedback in the form of "cultural judgment," with the surprising result that all 10 pitch collections they tried were accepted equally. Dehoux and Voisin (1993) also brought a DX7 synthesizer to Central Africa, to study the interaction of pitch and timbre in xylophones' scalar systems. They developed an interface so that local musicians could both play and retune the digital instruments as if they were xylophones, and built models of scales as well as pitch/timbre interaction for four Central African ethnic groups. Voisin (1994) applied this same methodology to the study of Javanese gamelan scales in Surakarta. A team of collaborating ethnomusicologists and instrumental acoustics researchers applied this same method in North Cameroon to the study of scales used by the flute bands of the Ouldémé tribe; they built a "synthetic flute" with a continuous differential pressure sensor embedded in a conical bore of bakelite, which they connected via MIDI to a real-time physical model of a recorder-like flue instrument (Cuadra, Vergez, and Causse 2002, Marandola 2003, Arom, Fernando, and Marandola 2007). Many of the same researchers also researched intonation in the vocal polyphony of the Bedzan Pygmies in Cameroon: they made multitrack recordings of these pieces in the field, then used software spectral analysis tools to produce sinusoidal additive synthesis models of each vocal part, then modified the pitch of certain scale degrees without changing rhythm, loudness, or timbre. By seeing how pitch could be altered without making the result unacceptable to their informants they were able to make conclusions about "the conception of the scales." (Arom, Fernando, and Marandola 2007).

Tsahalinas and colleagues (1997) took measurements of the physical dimensions of the *Elgin Auloi*, an unplayable but mostly intact specimen of an ancient Greek wind instrument from the 5th century BC. They then used these measurements as the input parameters to a computational/numerical model of the acoustics of a wind instrument with a cylindrical bore, which determined the pitches that the instrument might have produced as a function of unknown parameters such as the size of the reed.

2.3.2 Rhythm and Timing Analysis

Computer analysis of digital audio excels at precise scientific measurement of the timing of recorded music. In the past this kind of analysis required highly inaccurate methods such as stopwatch measurement or highly labor-intensive methods such as marking important time points with pencil on a magnetic tape and then measuring distance between marks with a ruler (Epstein 1985).

Schloss' early work on automatic transcription of percussive music (1985) output musical notation for a given audio input and also a detailed "tempo line" showing the variation of tempo over time. Bilmes (1993) worked with multitrack audio recordings of Afro-Cuban percussion, extracting note timing, modelling note timing with respect to a metric context as tempo variation plus per-note deviations, and finally applying machine learning techniques to produce stylistically correct expressive timing for new phrases. Lindsay (2006) used computational techniques to measure "swing" in recordings of Afro-Brazilian, Reggae, and Western popular music.

Clayton's (2000) wide-ranging theoretical and empirical study of the temporal aspect of North Indian music includes many graphs of tempo as a function of time, though the data they display came from manual marking of time points in recordings. It's a good example of the use of empirical timing information for ethnomusicologial analysis of a large corpus of recordings. Clayton et al. (2004) promote "chronometric analysis" as a method for ethnomusicologists in an article about the significance of *entrainment* to ethnomusicology.

2.3.3 MIR

The use of MIR techniques for general analysis of music in a style-independent way is explored in (McKay and Fujinaga 2007) which also provides a good overview of different types of features that can be extracted and classifiers that can be used. It also motivates the use of MIR techniques for analyzing large collections that would be hard to perform manually. The need of MIR research to include other domains of inquiry than Western pop and "classical" music has been argued by Futrelle and Downie (2002).

Although limited, there is some existing CE work in the field of MIR and it seems to be slowly growing as evidenced by relevant papers in the International Society for Music Information Retrieval (ISMIR) conference.[4] There are several archives of music from around the world that could potentially benefit from MIR access techniques (Proutskova 2007). In addition specific archives from particular music

[4] http://ismir.net

cultures are being created (Lanzelotte et al. 2007). Frequently rhythmic information is important for understanding and characterizing non-western music. Retrieval based on rhythmic information has been explored in the context of Greek and African traditional music (Antonopoulos et al. 2007). Automatic genre or style classification has been a well-explored problem for popular music. It has been explored in the context of Malay music styles (Norowi 2005). Some research involves designing algorithms for specific instruments. For example Persian music and the *santur* instrument are explained and analyzed by Heydarian and Reiss (2005). The classification of melodic patterns in Greek folk *clarinet* music using variable duration Hidden Markov Models (HMM) has been explored by Pikrakis et al. (2006). One of the most interesting current attempts at using MIR technologies in ethnomusicological archives involves the digitization of the ethnomusicological sound archive of the Royal Museum for Central Africa (Cornelis et al. 2005, Moelants et al. 2006).

Probably the most explored non-Western music tradition using MIR approaches is from India. There have been several works relevant to the study of traditional Indian music. The extraction of melodic information has been used for Raag recognition in (Chordia and Rae 2007).

Unlike Western music (and particularly Western "classical" music) where the notion of a composition as a well-defined "work" is relatively easy to define (but see Smiraglia 2001), in other music cultures the boundaries between composition, variation, improvisation and performance are more blurred and factors such as cultural context have to be taken into account. This can affect the design of digital MIR collections and libraries as described for the case of Carnatic (South Indian) ragas (Geekie 2002). For Carnatic music the main melodic "units" can be quite different than western pitches and scales; automatic pitch extraction techniques can be used to identify these "melodic atoms" (Krishnaswamy 2004).

3 Guidelines for CE Research

One of the main challenges facing CE researchers is the lack of established guidelines, principles and methodologies. In this section, we highlight some of the guidelines we have discovered through surveying existing literature in CE and our own experiences in this area. In addition, these guidelines emphasize areas in which MIR techniques can have significant impact beyond traditional approaches. To make the guidelines more concrete we provide specific examples from our own work. We intentionally focus on ideas and examples that are not obvious extensions of existing approaches to computer analysis, such as the analysis of symbolic representations similar to Western scores.

3.1 Collaboration with Music Scholars, Not Just Engineers

Much of the prior work cited above involves collaboration between ethnomusicologists and technical researchers. However, in the majority of existing MIR work that could potentially be used for CE purposes the authors are primarily engineers or computer scientists, which is not surprising given the early exploratory nature of this area. Unfortunately, frequently this results in a blind application of existing techniques typically to some specific music culture without having a clear musicological goal or motivation. This sometimes results in “solutions in search of a problem”. We believe the best way to overcome this is to actively seek interdisciplinary collaborations that include music scholars and technical researchers. Experimental results should generally be interpreted by music scholars with a understanding of the specific music(s) involved, similar to how scientific empiricism and musicological insight can complement each other, as Huron argued (1999).

Another good way to initiate such collaborations is through integrating MIR tools in the teaching of relevant courses. This provides an initial test-bed for ideas and approaches and helps the collaborating parties become familiar with each other and their respective vocabularies, tools, and methodologies. Frequently students are more open and receptive to new ideas and they can provide valuable feedback to help improve the application of MIR techniques to CE. Application of MIR techniques for original research can then follow.

3.2 Large Collections

“[T]here would be grounds for legitimate criticism if musicologists working in data-rich fields did not take full advantage of the methods available under such conditions, instead restricting themselves to traditional ‘humanities’ approaches developed for data-poor fields” (Cook and Clarke 2004: 5).

Humans are extremely effective at recognizing patterns, so it is important to emphasize that automatic approaches are far away from replacing them. However the amount of data a human can process is limited whereas machines are much better at dealing with enormous collections of audio and music. Large collections are also important for developing automatic approaches to analysis as frequently these approaches utilize machine-learning techniques that require large amounts of data for training. In practice a constant interaction between human and machine is required to build effective tools. Converting a large number of analog recordings to digital form is only the first step in creating effective collections for MIR purposes. Annotation and organization by experts are frequently also required.

3.3 Design of Domain-specific Techniques

One advantage of using computers for ethnomusicological research is that it is easy to apply existing general-purpose tools to new musical material, for example the plotting of spectrograms. However, we believe that the most interesting CE research involves implementing custom software specific to both the musical style(s) being studied and the aims of the research. Typically the first attempts at applying MIR techniques to ethnomusicological collections use generic techniques developed for Western music. However many music traditions or styles exhibit particular characteristics and constraints that enable the design of more effective automatic and semi-automatic analysis tools that are specific to the domain under consideration. We describe three specific examples that we are currently exploring and exhibit domain specific characteristics: Afro-Cuban *clave*, chant research, and tempo analysis of Kazakh *dombra* music.

Music involves periodicity at many levels, from the signal itself to high-level structure. In the case of African and Cuban music, there is a very fertile middle-ground temporal level that is based on what we call *meso-periodicity*, typically a 1-4 second long pattern. The pattern is repeated thousands of times, with very small variations in two modes:

- Rational deviations from the pattern (embellishment or improvisation)
- Minute timing deviations from the canonical pattern.

The Cuban *clave* is a small collection of rhythms (and also the name of the two-stick instrument usually used to play them) found embedded in virtually all Cuban music. *Clave* is a repeated rhythmic pattern that is often explicitly played, but often only implied; it is the essence of periodicity in Cuban music. Clave is found in three forms: *rumba clave* and two variants of *son clave*: 2-3 and 3-2. Son clave's two variants are simply phase-shifted versions of each other; it is the same pattern but starting in the second phrase instead of the first. It can be quite difficult sometimes to tell these variants of clave from each other. We are experimenting with different computational methods of understanding, identifying and generating *clave* in Cuban music. Since there are several variants of clave in Cuban music, we will attempt to automatically categorize them based on machine learning techniques. In testing our methods in the context of Cuban clave, we will further this kind of periodicity analysis in music that can be applied by musicologists in many other contexts.

Automatic pitch extraction techniques work relatively well on monophonic signals such as chant. Many chant scholars and ethnomusicologists have analyzed melodic formulas in examples of oral, semi-oral and notated chant. Attempting to identify such formulas, scholars must take a variety of conditions into consideration. How does the scholar identify such a formula? What are the main musical parameters used in identifying melodic formulas? How can melodic formulas overlap? Might formulas in an oral chant practice correspond to formulas in a notated or semi notated chant practices? How do these formulas relate to the textual syntax of a given chant example? How do melodic formulas in one version of plainchant compare with those in another example employing the same text?

MIR algorithms and systems can help to more objectively locate these formulas and to analyze their functionality in regard to musical and textual syntax. Being able to more objectively categorize melodic formulas would allow for a larger database of their musical parameters, their textual functionality as well their regional traits and relations. With such a database, the complexities and cross relations within chant traditions might be better understood. Scholars would then be able to better assess the functionality of such melodic formulas in both oral and written cultures. Comparing the functionality of melodic formulas in oral, semi-oral and written chant traditions might also provide for better explanations as to the development of notation within plainchant. For our study the primary examples are taken from archival recordings of various interpretations of early plainchant, Eastern European and Sephardic (Iranian and Syrian) Torah trope and laments from Eastern Europe (Hungary and Rumania).

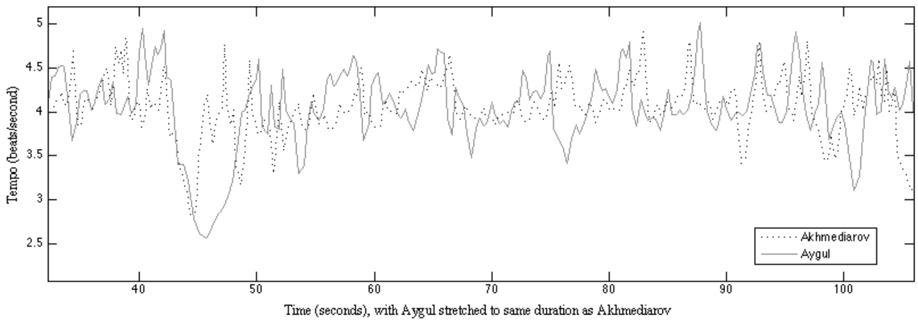

Figure 1. Instantaneous tempo curves of two performances of a composition for the Kazakh *dombra* (time-stretched for alignment as they have different durations).

Wright (2005) used a computer to compare the spectral content and note timings for two recorded performances of a composition for the Kazakh *dombra*, a long-necked, fretted, two-stringed plucked lute. He implemented a simple custom beat-tracking algorithm optimized for the extreme short-term tempo changes of this style: given hand-marked note onset times for both performances it produced continuous curves showing local tempo as a function of time, as shown in Figure 1. Comparison of these curves indicated that both performers varied tempo in much the same way at specific points in the composition (supporting the idea that specific tempo modulations are an intrinsic part of this musical style and this particular piece) while also showing stylistic differences between the two musicians. For example, though both performances tended to have "bumps" in the tempo curve (i.e., a brief change in local tempo followed by a return to the original speed) at the same places in the piece, those of the younger musician tended to spend more time at the second tempo before returning to the first, while those of the older musician tended to return more quickly to the original tempo. This difference is very apparent for the sudden slowing down near the 45-second point in Figure 1.

3.4 Content-and-Context Interfaces for Active Preservation

Due to globalization and various other factors, frequently all that is left from a particular music culture is an ethnomusicological archive consisting of recordings, artifacts, notes, photographs and other materials. Digitizing the archive is the obvious first step for preservation and analysis by specialists. In addition digital technology offers the possibility of allowing anyone in the world to explore the archive. Content and context interfaces for exploring digital ethnomusicological archives can be used to make the archive more accessible and interesting to explore for non-specialists who are not necessarily as familiar with the material as are researchers and musicologists.

Figure 2 shows an example of a content-aware user interface for exploring melodic structure. An excerpt from a recording of the *duduk,* an Armenian double-reed instrument, is shown. In addition to the *duduk* playing the main melodic line the recording contains background accompaniment consisting of percussion and plucked instruments. In order to separate the predominant melodic line from the audio signal we used the method of Lagrange et al. (2007). The resulting spectrogram is shown at the bottom of the figure and is used to detect repeating melodic ornamentations shown in the middle of the figure. The boundaries of the repeated ornaments are detected automatically; however their labeling as a particular pattern (in this case, "P7") is performed by an expert. The open source Audacity (http://audacity.sourceforge.net) audio editor is used as the graphical interface.

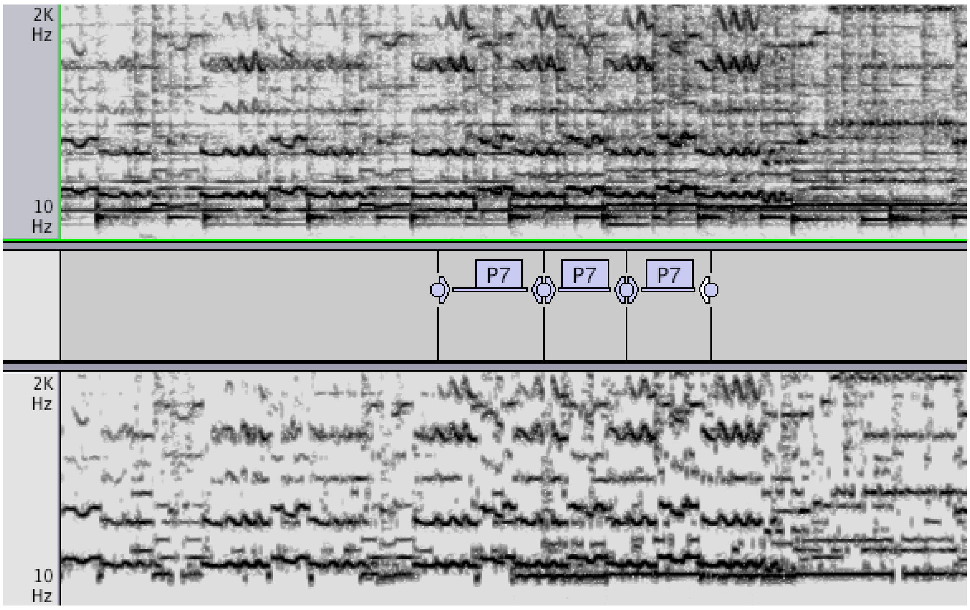

Figure 2. Spectrogram (showing amount of energy (darkness) as a function of time (horizontal) and frequency (vertical)) of an excerpt from a recording of *duduk* (traditional wind instrument of Armenian origin) with accompaniment (top), spectrogram of the automatically separated *duduk* melody (bottom), semi-automatic structural analysis of melodic ornaments (middle)

Another interesting application is the browsing of large ethnomusicological collections by listeners who frequently are unfamiliar with the material. Interactive content and context aware exploratory interfaces can be used for this purpose. A self-organizing map (SOM) is a automatic technique for mapping a high-dimensional space (in our case a collection of music) to a 2-dimensional discrete grid in such a way that objects that are similar in the high-dimensional space are close together in the grid. By automatically extracting features that describe the audio content it is possible to construct such maps for effectively browsing large music collections. As an example Figure 3 shows how different music traditions are spatially distributed over a self-organized map of music from around the world. The pattern of dark squares on the left is the distribution of Middle Eastern music and the pattern of dark squares on the right is the distribution of African and Afro-Cuban music. What is important to understand is that this layout is discovered automatically by analyzing the audio signals without any use of information about the geographic origin of the music in each recording. Essentially the structure is discovered by the system. In the middle we show how such music maps can be browsed on a collaborative tabletop interface (Hitchner, Murdoch and Tzanetakis 2007).

Figure 3. Pattern of distribution of Middle Eastern music (left) and African and Afro-Cuban (right) on a self-organized map of a world music collection. Collaborative browsing of the collection on a tabletop display (middle)

3.5 Beyond the Audio Recording – Analysis of Performance Gestures

Historically, the majority of music traditions were preserved by oral transmission of rhythms and melodies from generation to generation. Indian culture in particular is well known for its musical oral traditions that are still prevalent today. For the Western "classical" music tradition, however, the use of written notation has allowed more people to learn from the masters, not just those who had the opportunity to sit with them face to face. Then in the 1900's the age of audio recordings dawned using a progression of media including phonograms, vinyl, analog tapes, digital compact disks, and multi-channel audio systems, with each step improving the quality and the accuracy of the recordings. The invention of visual recording of musical performances on film, video, DVD, or online QuickTime and You Tube clips, has given musicians a closer look at the masters' performances in order to aid emulation. However, both audio and visual records turn performed music into an artifact, ignoring *the process* of making music that is so important to learn and preserve tradition.

There is more information in music performances than that which audio and video recordings alone can capture. A lot of this information becomes evident when observing the traditional oral teaching of music from teacher to student. Such lessons frequently involve detailed physical instructions and adjustments regarding how a musical instrument is played. One interesting possibility is to capture these physical gestures directly using sensors either placed on an instrument or the body of a performer in order to extract information about a music performance.

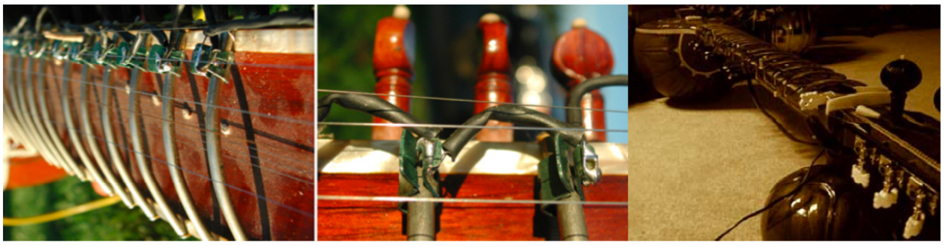

Figure 4. Fret-detection resistor network on the ESitar for capturing left-hand position.

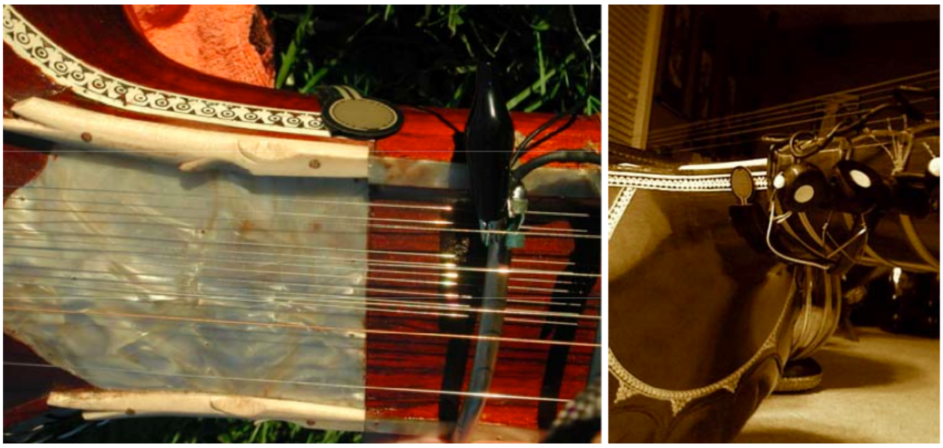

Figure 5. Thumb sensor on the ESitar for capturing plucking information.

As specific examples of digitally capturing gestures, we describe some of our work in the preservation of performance information for North Indian instruments. More details can be found in (Kapur et al. 2005). The ESitar is an Indian *sitar* retrofitted with a variety of sensors for capturing gestures of the performer while still producing sound acoustically and being playable as a traditional *sitar*. Figure 4 shows how the frets are connected using an exponentially distributed set of resistors in order to detect what fret is played by the performer. Figure 5 shows the location of a sensor that is placed under the right hand thumb and is used to deduce the direction and patterns of plucking. A force sensing resistor captures the applied force, which varies based on the stroke direction. A *Dha* stroke (upwards stroke) produces more pressure on the thumb than a *Ra* stroke (downwards stroke).

Figure 6 shows thumb pressure and fret detection data (middle plots) for a short excerpt of a music performance. Based on that information rhythmic onsets can be extracted (bottom plots). As expected, the figure shows that the melodic rhythmic pattern of the fret data is quite different from the rhythmic pattern of up-strokes and down-strokes based on the thumb pressure data. This information supports analysis of *sitar* performance that would be challenging to extract and decouple from the audio signal. Moreover it can be combined with audio analysis for improved real-time beat tracking (Benning et al. 2007a, Kapur et al. 2007a) and pitch tracking (Kapur et al. 2007a). The resulting tempo curves can be used for comparative studies of the dynamics of tempo changes by different musicians. Similar experiments have also been carried on the North Indian *tabla* percussion instrument (Benning et al. 2007b).

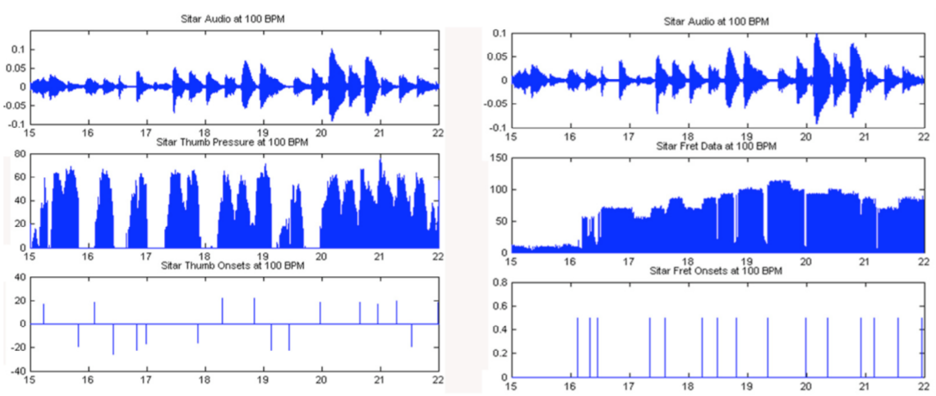

Figure 6. ESitar rhythmic onset detection using thumb pressure and fret detection data

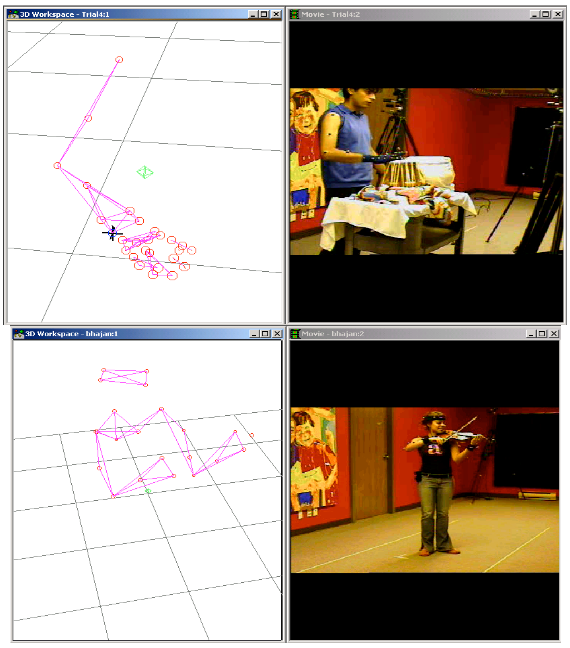

Figure 7. Screenshot of motion data capturing process for *tabla* performance (top) and violin (bottom)

Figure 7 shows an even more extensive data-capturing paradigm, using a VICON motion-capture system to gather full body skeleton models. Markers are placed on the body of the performers and multiple infrared cameras are used to track the motion in 3D. Data like this can aid in pedagogical applications to provide students particular information such as real-time posture detection and revealing the intricacies of a master's hand techniques (Kapur et al. 2005).

3.6 Unfreezing the frozen – Audio manipulations and musical robots

Until the advent of audio recording, music was completely flexible, but at the same time, completely volatile, in that it ceased to exist beyond the moment of performance. When recording technology appeared, the reverse situation happened: music could be preserved, but the preserved object, the recording, was un-malleable or frozen in time. This means that even though the invention of recording enabled the preservation of music of all cultures around the world, it turned music into a fixed artifact rather than the flexible ephemeral process it is. Finally, we can now "unfreeze the frozen" – we can not only search and analyze recordings, but we can make them malleable, according to our specifications. Examples of this malleability include changing the tempo without changing pitch, changing the pitch without changing the tempo, removing individual parts, adding more parts, rearranging materials, etc.

As a concrete example of how audio manipulations can be useful, consider the problem of detecting a particular rhythmic pattern in a large collection. The MIR system searches the database and returns the ten examples that it "believes" contain that particular pattern. Each song might be at a different tempo, and each performance might have variations of the query pattern. By time-stretching the rhythmic pattern to align, it is possible to analyze the variations between them visually and aurally at the same time.

Going beyond playback of audio signal manipulations, we can automatically "perform" manipulated or computer-generated symbolic representations of musical performances with a mechanical device controlled by a computer, i.e., a robot. Initial anecdotal evidence indicates that some traditional musicians are more interested in participating in performance analysis experiments when interacting with a robot compared to a disembodied computer playing through a loudspeaker.

In 1920, C.V. Raman, designed an automatic mechanical playing violin (Raman 1920) in order to conduct detailed studies of its acoustics and performance. Motivated by such work, Kapur et al. (2007b) built the *MahaDeviBot*, a mechanical musical instrument that extends North Indian musical performance scenarios and is also a pedagogical tool, keeping time and portraying complex rhythmic cycles and patterns in a way that no audio speakers can ever emulate. The development of the MahaDeviBot (see Fig. 8) serves as a paradigm for various types of solenoid-based robotic drumming techniques, striking 12 different percussion instruments gathered from around India, including frame drums, bells, finger cymbals, wood blocks, and gongs. The machine even has a bouncing head that can convey tempo to the human performer. In the future, the data collected from archival methods described above could be used to perform robotic musical instruments such as the MahaDeviBot.

Figure 8. MahaDeviBot communicating in real-time with a performer using sensor on the ESitar and audio analysis

4 Conclusions

We believe that the time is ripe for integrating Music Information Retrieval techniques into the area of Computational Ethnomusicology. After surveying representative existing work in this direction we described several guidelines for work in this area and provide specific examples from our own work to make them more concrete. MIR techniques have enormous potential to assist ethnomusicological research and it is our hope that this article helps to motivate researchers and music scholars explore this direction.

The globalization of music distribution has been a mixed blessing for traditional music cultures around the world. On the one hand the increasing popularity of "world" music has enabled some traditional musicians to have their music recorded and performed around the globe. On the other hand traditional music is being hybridized with music influences from around the world and frequently is only preserved in recorded archives. We hope that digital tools that enable more interaction with ethnomusicological archives than just listening can help us better understand the rich diversity of music traditions on this planet, as well as enable the ethnomusicologist to engage in studies that were not possible before.

Looking far into the future we envision a day when all recordings of every music culture are preserved and accessible using multiple intelligent interfaces and it is possible to capture and analyze every aspect of the incredible complexity of music performances. It might someday even be possible for the students of a music tradition to practice with robots that have been "taught" by a master a long time ago.

Acknowledgements

We thank the Social Sciences and Humanities Research Council of Canada and the Canada Foundation for Innovation for their financial support.

References

Antonopoulos, I. Pikrakis, A., Theodoridis, S., Cornelis, O., Moelants, D., and Leman, M. 2007. Music Retrieval by Rhythmic Similarity Applied on Greek and African Traditional Music. Proceedings of the International Conference on Music Information Retrieval (Vienna, Austria): 297-300.

Arom, S., and Fürniss, A. 1993. An interactive experimental method for the determination of musical scales in oral cultures: Application to the vocal music of the Aka Pygmies of Central Africa. *Contemporary Music Review* 9(1) and 9(2): 7-12.

Arom, S., Fernando, N., and Marandola, F. 2007. An Innovative Method for the Study of African Musical Scales: Cognitive and Technical Aspects. *Proceedings of the 4th Sound and Music Computing Conference* (Lefkada, Greece): 107-116.

Askenfelt, A. 1976. Automatic notation of played music (status report). *Dept. for Speech, Music, and Hearing: Quartetly Progress and Status Report* 17(1). Stockholm: KTH.

Baily, J. 1989. Principles of Rhythmic Improvisation for the Afghan Rubâb. *International Council for Traditional Music, UK Chapter Bulletin* 22: 3-16.

Bartok, B. and Lord, A. B. 1951. *Serbo-Croatian Folk Songs*. New York: Columbia University Press.

Bel, B. and Kippen, J. 1989. The identification and modeling of a "percussion" language and the emergence of musical concepts in a machine-learning experimental setup. *Computers and the Humanities*, 23(3): 199-214.

Benning, M., Kapur, A., and Tzanetakis, G. 2007a. Multimodal Sensor Analysis of Sitar Performance: Where is the Beat? *Proceedings of the IEEE International Workshop on Multimedia Signal Processing* (Chania, Greece).

Benning, M., Kapur, A., Till, B., Tzanetakis, G., and Driessen, P. 2007b. A Comparative Study on Wearable Sensors for Signal Processing of the North Indian Tabla. *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing* (Victoria, Canada).

Bilmes, J. 1993. *Timing is of the Essence: Perceptual and Computational Techniques for Representing, Learning, and Reproducing Timing in Percussive Rhythm.* Masters thesis, Massachusetts Institute of Technology.

Blacking, J. 1973. *How Musical is Man?* Seattle: University of Washington Press.

_________. 1979. The Study of Man as Music-Maker. In Blacking, J. and Keali'inohomoku, J. W. (Eds.) *The Performing Arts: Music and Dance* (pp. 3-15). The Hague: Mouton De Gruyter.

Carterette, E. and Kendall, R. 1999. Comparative Music Perception and Cognition. In D. Deutsch (Ed.) *The Psychology of Music (2nd ed.)* (pp. 725-792) San Diego: Academic Press.

Chordia, P. 2006. *Automatic Transcription of Solo Tabla Music*. Ph.D. diss., Stanford Univ.

Chordia, P., and Rae, A. 2007. Raag Recognition Using Pitch-Class and Pitch-Class Dyad Distributions. *Proceedings of the International Conference on Music Information Retrieval* (Vienna, Austria).

Clayton, M. 2000. *Time in Indian Music: Rhythm, Metre, and Form in North Indian Râg Performance*. Oxford Monographs on Music. Oxford: Oxford University Press.

Clayton, M., Sager, R., and Will, U. 2004. In time with the music: The concept of entrainment and its significance for ethnomusicology. *ESEM CounterPoint* 1: 1-45.

Cook, N. 2004. Computational and Comparative Musicology. In: Clarke, E. and Cook, N. (Eds.). *Empirical Musicology: Aims, Methods, Prospects* (pp. 103-126) Oxford: Oxford University Press.

Cook, N. and Clarke, E. 2004. Introduction: What is Empirical Musicology? In: Clarke, E. and Cook, N. (Eds.). *Empirical Musicology: Aims, Methods, Prospects* (pp. 3-14) Oxford: Oxford University Press.

Cooper, D. and Sapiro, I. 2006. Ethnomusicology in the Laboratory: From the Tonometer to the Digital Melograph. *Ethnomusicology Forum*, 15(2), 301-313.

Cooper, M., Foote, J., Pampalk, E., and Tzanetakis, G. 2006. Audio Visualization for Music Information Retrieval. *Computer Music Journal*, 30(2): 42-62.

Cornelis, O., De Caluwe, R., De Tré, G., Hallez, A., Leman, M., Matthé, T., Moelants, D., and Gansemans, J. 2005. Digitization of the Ethnomusicological Sound Archive of the Royal Museum for Central Africa, *IASA Journal,* 26: 35-43.

Cuadra, P. de la, Vergez, C., and Caussé R. 2002. Use of physical-model synthesis for developing experimental techniques in ethnomusicology: The case of the Ouldémé flute. *Proceedings of the International Computer Music Conference*: 53-56.

Dannenberg, R., Birmingham, W., Tzanetakis, G., Meek, C., Hu, N., and Pardo, B. 2004. The MUSART testbed for query-by-humming evaluation. *Computer Music Journal* (Special Issue on Music Information Retrieval) 28(2): 34-48.

Dehoux, V., and Voisin, F. 1993. An interactive experimental method for the determination of musical scales in oral cultures: Application to the xylophone music of Central Africa. *Contemporary Music Review* 9(1) and 9(2): 13-19.

Epstein, D. 1985. Tempo Relations: A Cross-Cultural Study. *Music Theory Spectrum* 7: 34-71.

Foote, J. and Cooper, M. 2001. Visualizing Musical Structure and Rhythm via Self-Similarity. *Proceedings of the International Computer Music Conference* (Havana, Cuba): 419-422.

Futrelle, J. and Downie, S. 2002. Interdisciplinary Communities and Research Issues in Music Information Retrieval. *Proceedings of the International Conference on Music Information Retrieval* (Paris, France) : 215-221.

Geekie, G. 2002. Carnatic Ragas as Music Information Retrieval Entities. *Proceedings of the International Conference on Music Information Retrieval* (Paris, France): 257-258.

Gillet, O. and Richard, G. 2003. Automatic Labeling of Tabla Signals. *Proceedings of the International Conference on Music Information Retrieval* (Baltimore, MD): 117-124

Heydarian, P. and Reiss, J.D. 2005. The Persian music and the santur instrument. *Proceedings of the International Conference on Music Information Retrieval* (London): 524-527.

Hitchner, S., Murdoch, J., and Tzanetakis, G. 2007. Music Browsing using a Tabletop Display. *Proc. of the International Conference on Music Information Retrieval* (Vienna): 175-176.

Huron, D. 1999. *The new empiricism: Systematic musicology in a postmodern age*. 1999 Ernst Bloch Lecture. University of California, Berkeley.

Kapur, A., Davidson, P., Cook, P.R., Driessen, P.F., and Schloss, W.A. 2005. Preservation and Extension of Traditional Techniques: Digitizing North Indian Performance. *Journal of New Music Research*: 227-236.

Kapur, A., Percival, G., Lagrange, M., and Tzanetakis, G. 2007a. Pedagogical Transcription for Multimodal Sitar Performance. *Proceedings of the International Conference on Music Information Retrieval* (Vienna, Austria): 351-352.

Kapur, A., Trimpin, Singer, E., Suleman, A., and Tzanetakis, G. 2007b. A Comparison of Solenoid-Based Strategies for Robotic Drumming. *Proceedings of the International Computer Music Conference* (Copenhagen, Denmark).

Kippen, J., and Bel, B. 1992. Modelling music with grammars: formal language representation in the Bol Processor. In: A. Marsden and A. Pople (Eds.). *Computer Representations and Models in Music*. London: Academic Press.

Klapuri, A. 1998. *Automatic Transcription of Music*. Master's Thesis, Tampere University of Technology.

Krishnaswamy, A. 2004. Melodic Atoms for Transcribing Carnatic Music. *Proceedings of the International Conference on Music Information Retrieval* (Barcelona, Spain): 345-348.

Lagrange, M., Martins, L. G., Murdoch, J., and Tzanetakis, G. 2007. Normalized Cuts for predominant melodic source separation. *IEEE Transactions on Audio, Speech and Language Processing* (to appear).

Lanzelotte, R. S. G, Balleste, A. O., and Ulhoa, M. 2007. A Digital Collection of Brazilian Lundus. *Proceedings of the International Conference on Music Information Retrieval* (Vienna, Austria): 163-164.

Lieberman, F. 1970. Computer-Aided Analysis of Javanese Music. In: Lincoln, H. (ed.). *The Computer and Music* (pp. 182-192) Ithaca: Cornell University Press. See also: http://arts.ucsc.edu/faculty/lieberman/Gamelan.html

Lindsay, K. A. 2006. *Rhythm Analyzer: A Technical Look at Swing Rhythm in Music*. Master's thesis, Southern Oregon University.

Marandola, F. 2003. The Study of Musical Scales in Central Africa: The Use of Interactive Experimental Methods. *Proceedings of the International Symposium Computer Music Modeling and Retrieval*: 34-41.

Marcus, S. L. 1989. *Arab music theory in the modern period*. Ph.D dissertation, UCLA.

McKay, C. and Fujinaga, I. 2007. Style-Independent Computer-Assisted Exploratory Analysis of Large Music Collections. *Journal of Interdisciplinary Music Studies*, 1(1): 63-86.

Moelants, D., Cornelis, O., Leman, M., Gansemans, J., De Caluwe, R., De Tré, G., Matthé, T., and Hallez, A. 2006. Problems and Opportunities of Applying Data- & Audio-Mining Techniques to Ethnic Music. *Proceedings of the International Conference on Music Information Retrieval*: 334-336.

Nesbit, A., Hollenberg, L., and Senyard, A. 2004. Towards Automatic Transcription of Australian Aboriginal Music. *Proceedings of the International Conference on Music Information Retrieval* (Barcelona, Spain): 326-330.

Nettl, B. 1964. *Theory and Method in Ethnomusicology*. New York: Macmillan.

Norowi, M., Doraisamy, S., and Wirza, R. 2005. Factors Affecting Automatic Genre Classification: An Investigation Incorporating Non-Western Musical Forms. *Proceedings of the International Conference on Music Information Retrieval* (London): 13-20.

Parncutt, R. 2007. Systematic Musicology and the History and Future of Western Musical Scholarship. *Journal of Interdisciplinary Music Studies*, 1(1): 1-32.

Pikrakis, A., Theodoridis, S., and Kamarotos, D. 2006. Classification of Musical Patterns using Variable Duration Hidden Markov Models. *IEEE Transactions on Audio, Speech and Language Processing*, 14(5): 1795-1807.

Proutskova, P. 2007. Musical Memory of the World – Data Infrastructure in Ethnomusicological Archives. *Proceedings of the International Conference on Music Information Retrieval* (Vienna, Austria): 161-162.

Raman, C. V. 1920. *Experiments with Mechanically played Violins*, presented at Indian Association of Cultivation of Science, India.

Schloss, W. A. 1985. *On the Automatic Transcription of Percussive Music: From Acoustic Signal to High-Level Analysis*. Ph.D dissertation, Stanford University.

_________. 1987. *The Digital Melograph: Digital Technology Meets Ethnomusicology*. (abstract only) Presented at the annual meeting of the Northeast chapter of the Society for Ethnomusicology.

Seeger, C. 1951. An Instantaneous Music Notator. *Journal of the International Folk Music Council III* 103-106.

Smiraglia, R. 2001. Musical Works as Information Retrieval Entities: Epistemological Perspectives. *Proceedings of the International Conference on Music Information Retrieval* (Bloomington, IN): 85-91.

Stock, J. P. J. 2004. Documenting the Musical Event: Observation, Participation, Representation. In: Clarke, E. and Cook, N. (Eds.). *Empirical Musicology: Aims, Methods, Prospects* (pp. 15-34) Oxford: Oxford University Press.

Todd, N. 1994. The Auditory “Primal Sketch”: A Multiscale Model of Rhythmic Grouping. *Journal of New Music Research* 23: 25-70.

Toiviainen, P. and Eerola, T. 2001. A Method For Comparative Analysis Of Folk Music Based On Musical Feature Extraction And Neural Networks. *Proceedings of the 7th International*

Symposium on Systematic and Comparative Musicology and 3rd International Conference on Cognitive Musicology (Jyväskylä, Finland): 41-45.

Tsahalinas, K., Tzedaki, K., Cook, P., Psaroudakes, S., Kamarotos, D., and Rikakis, T. 1997. Physical Modeling Simulation of the Ancient Greek Elgin Avloi. *Proceedings of the International Computer Music Conference* (Thessaloniki, Greece): 454-457.

Tufte, E. 1990. *Envisioning Information.* Cheshire, CT: Graphics Press.

Voisin, F. 1994. Musical Scales in Central Africa and Java: Modeling by Synthesis, *Leonardo Music Journal* 4: 85-90.

Von Hornbostel, E. M. 1928. African Negro music. *Africa*, 1: 30-62.

Wright, M, 2005. *Empirical Comparison of Two Recordings of the Kazakh Dombra piece 'Akbai'*. Electronic publication: http://ccrma.stanford.edu/~matt/dombra