DMRN+1

Digital Music Research Network One-day Workshop 2006

Queen Mary, University of London

Wed 20 December 2006

Programme and Abstracts

Programme

10:30	Registration opens; Tea/Coffee
11:00	**Welcome** and opening remarks **Mark Plumbley**, Queen Mary University of London **KEYNOTE** The Music Information Retrieval Evaluation eXchange (MIREX): Lessons Learned and Future Visions **Dr J Stephen Downie**, University of Illinois at Urbana-Champaign, IL, USA
11:50	**Session 1** Tools for Expert Musicians for Practising and Rehearsing Microtonal Music **Graham Hair, Ingrid Pearson, Nick Bailey and Doug McGilvray**, University of Glasgow / Royal College of Music Music 2.0 at Queen Mary **Mark Sandler and Josh Reiss**, Queen Mary University of London
12:40	**Buffet Lunch**, Networking; **Posters** will be on display
13:50	**Session 2** Digital Music Library Research at McGill University **Ichiro Fujinaga**, McGill University, Canada Creating a musicological toolkit: CHARM and the Mazurkas project **Nicholas Cook**, Royal Holloway, University of London Visual Interfaces to Musical Composition **Jean-Baptiste Thiebaut**, Queen Mary University of London Towards a Grid Composition Environment **John Fitch, James Mitchell and Julian Padget**, University of Bath
15:15	**Tea/Coffee; Posters** will be on display
15:40	**Session 3** Realtime Machine Listening **Nick Collins**, University of Sussex Multi-modal acquisition of performance parameters for analysis of Chopin's B flat minor piano sonata finale Op.35 **Jennifer MacRitchie, Nicholas J.Bailey and Graham Hair**, University of Glasgow Information Dynamics in Music Cognition **Marcus T. Pearce and Geraint A. Wiggins**, Goldsmiths College, University of London
16:25	**Panel Discussion** Funding Digital Music Research
16:55	**Close**

Talks

KEYNOTE:
The Music Information Retrieval Evaluation eXchange (MIREX): Lessons Learned and Future Visions

J Stephen Downie
University of Illinois at Urbana-Champaign

The Music Information Retrieval Evaluation eXchange (MIREX) is a community-based formal evaluation framework coordinated and managed by the International Music Information Retrieval Systems Evaluation Laboratory (IMIRSEL) at the University of Illinois at Urbana-Champaign (UIUC). IMIRSEL has been funded by both the National Science Foundation and the Andrew W. Mellon foundation to create the necessary infrastructure for the scientific evaluation of the many different techniques being employed by researchers interested in the domains of Music Information Retrieval (MIR) and Music Digital Libraries (MDL). The basic components of the MIREX framework are premised on the Text Retrieval Conference (TREC) approach to the evaluation of text retrieval systems:

a) a set of standardized collections;

b) a set of standardized tasks/queries to be performed against these collections; and,

c) a set of standardized evaluations of the results generated with regard to the tasks/queries.

It is the purpose of this presentation to introduce the basics of the MIREX including the unique challenges posed when trying evaluate retrieval systems specialized in working with music information. The talk will also highlight the author's positive and negative experiences in running MIREX as a jumping off point for outlining the motivation for developing a more robust and sophisticated "Do-It-Yourself" (DIY) web service architecture for the sustainability of the MIREX program.

Biography
Dr. J. Stephen Downie is an Associate Professor at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign (UIUC). He is Director of the "International Music Information Retrieval Systems Evaluation Laboratory" (IMIRSEL). Professor Downie is Principal Investigator on the "Human Use of Music Information Retrieval Systems" (HUMIRS) and the "Music-to-Knowledge" (M2K) music data-mining projects. He has been very active in the establishment of the Music Information Retrieval and Music Digital Library communities through his ongoing work with the ISMIR series of MIR conferences as a member of the ISMIR steering committee.

Realtime Machine Listening

Nick Collins
University of Sussex

Recent work in preparing generally usable realtime machine listening plug-ins for SuperCollider 3 will be discussed. Lessons learnt concerning such issues as SC3 plug-in requirements and cross-platform compatibility, processor load amortisation and the practicalities of machine listening for concert purposes will be briefly outlined.

Demonstrations will be given of the recently converted Tartini pitch tracking algorithm and some other UGens. A call for collaboration on further development of such a repertoire will be issued.

Creating a musicological toolkit: CHARM and the Mazurkas project

Nicholas Cook
Royal Holloway, University of London

Towards a Grid Composition Environment

John Fitch, James Mitchell and Julian Padget
University of Bath

Grid Problem Solving Envirnments are largely seen as a means to support *in silico* science experiments, with the beginnings of interest in social sciences and now even the arts. The primary building block of Grid PSEs is the web service, supported by mechanisms to find services based on published descriptions and workflow enactment engines that choreograph the execution of combinations of services. We report upon the recent construction of software to enable the creation and deployment of Csound-based web services and the experience of using the Triana and Taverna workflow systems to combine such services. We conclude with an outline of how our related research into the semantic description of mathematical web services and matchmaking and brokerage of such services could be adapted to the discovery and invocation of sound services.

Digital Music Library Research at McGill University
Ichiro Fujinaga, John Ashley Burgoyne, Catherine Lai, Beinan Li, Cory McKay, and Laurent Pugin
McGill University

This talk will present the intersection of various research projects in music information acquisition, preservation and retrieval being conducted by the DDMAL (Distributed Digital Music Archives and Libraries) group at McGill University. These include general-purpose music information retrieval (MIR) software tools, an optical music recognition system for early music scores captured on microfilm, guidelines for the digitisation of phonograph recordings including the creation of metadata dictionaries, an interoperability study of heterogeneous networks of sound recording libraries, and the conversion of scanned 3D microscopic images of LP records to audio.

We have recently developed a suite of software tools that are designed to reduce duplication of work and facilitate research in the field of MIR. The first tool we developed, ACE (Autonomous Classification Engine), is a framework for automatically finding effective classification methodologies for arbitrary supervised classification problems. Classification plays a central role in many MIR tasks including genre classification, instrument identification, and melody recognition. In order to supply the classification engine with a sufficient amount of feature data, we developed two general feature extractors, jSymbolic and jAudio, to extract many possible features from symbolic (e.g., MIDI) and audio data, respectively. We also added jMusicMetaManager to our toolset in order to extract and clean up the metadata associated with music files, as consistent and accurate metadata is essential in classifier training and evaluation.

Once we had these tools, we needed to collect a ground truth dataset, for which we created Codaich, a large database of music. But, acknowledging that truly useful databases for MIR research will be distributed, our tools were refitted to work in a networked environment. This resulted in the development of a system called OMEN (On-demand Metadata Extraction Network), which exploits the local computing resources of libraries. OMEN only exports metadata (including feature values), not audio samples, in order to circumvent copyright restrictions.

The digitisation of music scores (manuscripts or prints) is expensive, and it will be decades before many scores are digitised. It turns out, however, that most existing Western music scores have been microfilmed in the last century. If we use high-speed microfilm scanners, it would be possible to digitise these scores relatively quickly. In order to perform content searches on scores, it is first necessary to convert the scanned images to symbolic representations, which requires optical music recognition (OMR) technology. We are developing an OMR system that specifically targets the challenges posed by the varying quality of microfilms and the various early music notation systems in use before the 18th century.

We have created a digital library of early commercial recordings of George F. Handel's music. These include scanned images of album covers, liner notes, disc labels; audio files in various formats; and over 170 associated metadata fields. We have developed a metadata dictionary that clearly defines each metadata element. Realizing that standardisation of metadata schema in various digital audio libraries in the near future is unrealistic, we are investigating ways to make it possible to used intelligent federated search techniques to process disparate audio collections. The aim is to make searches among these collections transparent to users.

Using a white-light interferometry profiler microscope (Wyko NT8000), we are investigating the possibility of converting scanned 3D images of stereo phonograph record grooves to sound. This will allow us to "play" broken records since the images are stitched together in the visual domain. We are also looking into the possibility of restoring damaged (scratched) recordings using image-processing techniques.

Tools for Expert Musicians for Practising and Rehearsing Microtonal Music

Graham Hair*, Ingrid Pearson+, Nick Bailey* and Doug McGilvray
*University of Glasgow, +Royal College of Music

The objective is to demonstrate some ways in which musicians (especially singers) with a advanced Conservatory-based musical training can learn to adapt their thinking, perception and performance skills to music composed with "microtonal" scales. For practical purposes, consideration is restricted, in the present case, to music composed with one particular microtonal scale: the one with 19 tones per octave.

The authors have drawn on the collaborative expertise of 4 musicians and 2 engineers in prosecuting this project. Composer Graham Hair has written some very short (60") 19-ET songs for voice, clarinet and keyboard accompaniment, drawing on some rough music-cognitive "rules of thumb" suggested by Richard Parncutt (Professor of Systematic Musicology, University of Graz). Clarinettist Ingrid Pearson and soprano Amanda Morrison (BBC Singers, Steve Reich & Musicians, Scottish Voices et al) rehearse this music and related practice exercises with the composer (keyboard accompanist).

Rehearsal of this material is aided by 3 pieces of special technology developed by the engineers:

(1) Pitch-tracking software ("the Rosegarden Codicil"), which "listens" to the performers and provides feedback on how they're doing.

(2) A special "scordatura" (re-tuned) keyboard, which makes it possible to play 19-ET music on a regular ("12-ET") keyboard

(3) A special "tablature" score, which re-types a 19-ET score telling the keyboardist "where to put his fingers" (tablature) on a 12-ET keyboard to get the 19-ET sounds he wants (instead of telling him "what sounds are to emerge").

The presentation will demonstrate all of these.

Relevant aspects of tuning and intonation theory, existing microtonal repertoire in the West, microtonal practices in non-Western cultures and practical performance issues are sketched in, in order to establish a context for a discussion of the issues of perception and cognition which composers and performers have to confront and to contextualise the approach which the project has adopted.

This paper is one of the fruits of an inter-disciplinary research project, funded by the Arts and Humanities Research Council.

Multi-modal acquisition of performance parameters for analysis of Chopin's B flat minor piano sonata finale Op.35

Jennifer MacRitchie, Nicholas J.Bailey and Graham Hair
University of Glasgow

The finale of Chopin's B flat minor piano sonata Op.35 has confounded traditional musicological approaches and there exists little serious analysis of the piece. However, there are many recorded and live performances of the work, so a fruitful approach to compensating for the lacunae is to use engineering techniques to proceed directly to the sound of the piece.

Previous approaches to performance analysis measure variables such as tempo and dynamics, and this is usually done without much consideration of the musical structure. This paper will discuss an approach which involves the collection of audio, video and MIDI data over several performances with different interpretations in performance and how they illuminate aspects of the musical structure, demonstrated live on the piano. Collected multi-variate performance data is correlated, with care being taken that the analysis exposes the musical semantics in a quantifiable manner. These results of this are then to be compared with traditional score-based music analysis to establish a better understanding of the piece, as well as developing and innovating technology for the better representation of music by computational means.

Information Dynamics in Music Cognition

Marcus T. Pearce and Geraint A. Wiggins
Goldsmiths College, University of London

Research on music perception is often inspired by psychological theories developed in other domains. This practice may simply reflect a simplifying methodological approach or it may imply that the two domains make shared use of common cognitive components and processes. For example, the Gestalt principles of similarity and proximity, originally developed to describe phenomena in visual perception, have been adapted and incorporated into theories of the perception of melodic groups (Lerdahl & Jackendoff, 1983) and the cognitive processes involved in melodic expectation (Narmour, 1990).

Here we examine the use of information theoretic methods for modelling the cognitive processing of music. In particular, we focus on the statistical induction of sequential regularities in music and the use of these regularities in prediction. There is existing evidence that statistical learning and prediction are involved in the segmentation of tone sequences (Saffran et al., 1999) and generation of melodic expectations (Pearce & Wiggins, 2006).

We extend these approaches in a study of Two Pages, a minimalist composition by Philip Glass, which is strictly monodic, isochronous and monotimbral allowing for a tightly controlled study without compromising ecological validity. The investigation is conducted using a statistical model (Pearce & Wiggins, 2006) of melodic structure that operates on a symbolic musical surface consisting of sequences of note-like events. It is capable of representing and processing regularities in a number of different representations present in or derived from this surface. The model derives its knowledge of sequential melodic structure entirely from a corpus of existing music but is also capable of dynamically learning the structure of the composition it is currently predicting.

The results demonstrate that the formal structure of the piece (York, 1981) is indicated by characteristic changes in information content, while information-theoretic measures of surprise successfully predict the judgements of an expert musicologist.

We argue that while Gestalt theories merely describe the perceptual behaviour of listeners, the information dynamic approach yields functional models of the cognitive processes underlying that behaviour. In order to further corroborate this claim, we are currently devising experiments to examine which organising principles are used in the generation of expectations and how they are implemented in the brain. Specifically, we plan to compare empirically the predictions of the information-dynamic model with behavioural measures of expectation, physiological measures of arousal and spatiotemporal brain dynamics, recorded while listening to Two Pages.

Kostelantz, R. (Ed.). (1997). Writings on Glass. Berkeley, CA: University of California Press.

Lerdahl, F. & Jackendoff, R. (1983). A Generative Theory of Tonal Music. Cambridge, MA: MIT Press.

Narmour, E. (1990). The Analysis and Cognition of Basic Melodic Structures: The Implication-realisation Model. Chicago: University of Chicago Press.

Pearce, M. T. & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. Music Perception, 23(5), 377–405.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. Cognition, 70(1), 27–52.

York, W. (1981). Form and process in Two Pages of Philip Glass. Sonus, 1(2), 28–50. Reprinted in Kostelantz (1997, pp. 60–79).

Music 2.0 at Queen Mary

Mark Sandler and Josh Reiss
Queen Mary University of London

OMRAS2

The OMRAS2 project represents several years of discussion, preparation and hard work - and that was just to get it funded. It follows on from the seminal OMRAS project, which claimed the world's first symbolic music retrieval from a polyphonic audio query over a meaningful test collection.

OMRAS2 will last 42 months and funds 132 man-months of research effort. Its aim is to provide a distributed research environment for Music Informatics scientists, Music Information Retrieval researchers and Computational Musicologists.

This brief talk will outline the project and will attempt to kindle interest in the community, hoping that this will lead to an even bigger effort from UK and other researchers.

The EASAIER Project (Enabling Access to Sound Archives through Integration, Enrichment and Retrieval)

Many digital sound archives still suffer from tremendous problems concerning access. Materials are often in different formats, with related media in separate collections, and with non-standard, specialist, incomplete or even erroneous metadata. Thus, the end user is unable to discover the full value of the archived material. To expose the inherent value of the archived material, powerful multimedia mining techniques are needed, in combination with content extractors, meaningful descriptors, and visualization tools.

Existing retrieval systems often do not take into account the specific nature of the media content. Search by speech or musical feature functionality is rare. Thus retrieval techniques are restricted and inflexible. To address this, multiple retrieval techniques need to be merged and deployed, and similarity and structure must be conceptualised in order to provide a usable service. An efficient and effective retrieval system needs to be grounded in semantic description, similarity, and structure in order to provide rich functionalities related to the exploration of sound archives.

Another issue is that of providing appropriate interaction with and presentation of material for the end-users. An archive used by musicians and music students, for instance, requires that the material can be manipulated or modified appropriately at playback; archives of recorded broadcasts need to emphasise appropriate segmentation and interactive speech recognition features.

This is the motivation for the two and a half year European project, Enabling Access to Sound Archives through Integration, Enrichment and Retrieval (EASAIER). EASAIER will enable access to sound archives through enrichment of materials, integration of multimedia sources, and creation of advanced retrieval systems. It implements recent advances in machine learning, music and speech processing, and information retrieval. Furthermore, it addresses a growing demand for interactive electronic materials.

EASAIER allows archived materials to be accessed in different ways and at different levels. The system will be designed with libraries, museums, broadcast archives, and music schools and archives in mind. However, the tools may be used by anyone interested in accessing archived material; amateur or professional, regardless of the material involved. Furthermore, it enriches the access experience as well, since it enables the user to experiment with the materials in exciting new ways.

By enriching and organizing the materials in the archive, we will create new and innovative methods of access. This enhanced access stimulates use, because multimedia resources will be connected, and archived material will be exploited and made visible to the right users. This will be deployed in connection with strong evaluation studies using communities of users and content managers.

Visual Interfaces to Musical Composition

Jean-Baptiste Thiebaut
Queen Mary University of London

External representations, especially visual representations, provide a common means of distributing the cognitive and computational effort involved in a variety of human activities (Scaife and Rogers, 1986). In the context of music the most salient visual representations are notations. These developed initially as an aid to memorise and subsequently teach songs. However, as music increased in complexity (i.e. with polyphony), notation acquired a compositional function: scores facilitated the simultaneous representation of interactions between multiple voices.

The potential of notations to foster new musical practices has been exploited throughout the 20th century. For example, Morton Feldman introduced with a graphical notation designed to foster freedom of interpretation. Iannis Xenakis developed the UPIC, a program that transforms graphical representations into sound parameters.

It is clear for domains such as architecture that design sketches are strategically vague (Suwa and Tversky, 1996; Neilson and Lee, 1994). The creative process depends to some degree on the underspecification or ambiguity characteristic of a sketch. In this paper we make two points. First, that sketching is also common practice among contemporary composers in the early stages of creating a piece. Second, that contemporary music technologies do not provide support for this function.

Our argument is developed through examples of contemporary composers who exploit music technology to realise their compositions but not to envisage them. These composers use sketches to provide successive approximations of a musical idea. We discuss what is required to bridge the gap between creative practice and visual interfaces to music.

Neilson, I. and Lee, J., 1994, Conversations with graphics: implications for the design of natural language/graphics interfaces, International Journal of Human-Computer Studies, vol. 40, pp 509 -541.

Scaife, M. and Rogers, Y., 1996, External cognition: How do graphical representations work?, International Journal of Human- Computer Studies, vol. 45, pp 185-213.

Suwa and Tversky, 1996: What architects see in their sketches: implications for design tools, Conference on Human Factors in Computing Systems, Vancouver, British Columbia, Canada , pp 191-192.

Posters

Information dynamics and musical structure

Samer Abdallah and Mark Plumbley
Queen Mary University of London

Given a probabilistic model of a temporal random process, certain well-known information-theoretic quantities like entropy and entropy rate can be used to characterise the process as a whole. In additional to global measures such as the average information rate, it is also possible, for a given sequence drawn from the model, to formulate time-varying information-theoretic quantities such as the optimal coding length of each element in the sequence and the information supplied by each element about other elements. Combined with the possibility that the model of the process is being learned online as it is observed, we are lead to consider a rather dynamic process in which the learning the model and the processing of each observation can be tracked and assessed in information-theoretic terms which are not themselves specific to the particular model or the medium by which the sequence is realised.

Authors such as L. B. Meyer and D. E. Berlyne have argued that such an information-theoretic assessment is fundamental to the cognition of music, and that quantities such as entropy and information are closely related to concepts which are more directly relevant to a human observer, such as uncertainty, complexity, tension, implication, resolution, and so on. These are what Berlyne called ‘collative variables’ since they are to do with patterns of occurrence rather than medium-specific details.

Our work focuses on (a) how these ideas can be implemented in the context of specific dynamic probabilistic models, (b) which information-theoretic variables can usefully be computed, and (c) on how these relate to subjective experiences in human listeners. In particular, we introduce the concept of ‘predictive information’, which might provide an explanation for the ’inverted-U’ relationship (cf the Wundt curve) between entropy and certain aspects of subjective experience. We also present an analysis of Philip Glass’s Two pages and show how even a simple Markov chain model yields a structural analysis of the piece which agrees with that of an expert listener.

Aspects of Similarity

Hamish Allan and Geraint Wiggins
Goldsmiths College, University of London

We present a novel method for combining musical similarity measures.

It is often assumed that similarity in music can be distilled into a single measure, despite the existence of various different foci (e.g., melodic, timbral, harmonic, rhythmic) in the literature. These may be combined using some weighting trained from a ground truth based on expert or community opinion, but in trying to be jack of all trades, this kind of single measure remains master of none.

Similarity depends very much on the perspective of the beholder and the context of the exercise. We present a method for allowing users to create example sets of pieces of music which are used to determine a weighting for various existing similarity measures. This allows for users lacking a musical vocabulary to make queries like "I think all these pieces of music are similar in a way that appeals to me. Find me music that is similar to them in the same way."

A Locality Sensitive Hashing Algorithm for Fast Song Similarity

Michael A. Casey* and Malcolm Slaney+
* - Goldsmiths, University of London, + - Yahoo! Research Inc.

Today's million-song music databases require careful consideration of algorithms to perform computations within reasonable expectations of time, memory usage and other computing resources.

We demonstrate a new algorithm for radius-based nearest neighbor searching in high-dimensional data. We test the algorithm in an experiment on identification of remixed songs in a database of pop songs. We use subsequences of features from each song, called audio shingles, to perform a sequence search using the Euclidean metric. The distances are sorted to find the k-nearest neighbors between each song pair in the database.

Choosing a pair-wise algorithm in a 2M+ song database results in order 10^{18} multiplications. Instead, we employ a new algorithm that uses Locality Sensitive Hashing, which scales sub-linearly with the number of songs, to find nearest neighbors without pair-wise distance computations. We present results from audio song similarity experiments along with some of the details of the algorithms.

Splunge: An Extensible HCI and Audio Network Model

William Evans
Univeristy of Glasgow

Splunge is a new model for recording, processing, editing and performing music. It is based on a network model, employing an interface operating system to allow a wide range of interactive models.

Current DAW systems primarily model the analog recording process. Although this modeling has been accomplished with great success, it is based on software implementations of physical devices and their relationships — neither of which are required in the digital domain. Mixers, instruments, effects processors, tracks, busses, and other traditional constructs remain the dominant elements of interaction, software architecture, and hardware design.

The main aspects of Splunge can be divided into 4 categories:

1) Separation of Dynamic and Static data.

Central to Splunge is separation of “dynamic” and “static” data in the hardware/software architecture and user interface. I define dynamic data as linear digital information such as digital audio, MIDI, video, and automation instructions (streams). Static data refers to the configuration of the DSP (Digital Signal Processing) environment, which processes the dynamic data.

2) Static and Dynamically Generated Interfaces

Likewise, there are two parts to interface generation in Splunge: static (standard, predefined GUI) and dynamically generated. I use a static method to generate the interface for dynamic data, and a dynamic method for static data. Splunge employs GoldMark a user interface operating system I am developing for abstracting and modeling a program's functionality, to dynamically generate the interfaces.

3) DigitalFire, an Audio Network Protocol

DigitalFire is a multi-user audio network protocol for encapsulating digital music processes and devices into Nodes—standardized, programmable entities. Nodes can contain other nodes, and use other nodes' services to create conglomerate nodes, mixing hardware and software services. Nodes can communicate with other nodes independently using Node Aspects.

(5) Flexible Modeling of Audio Generation, Processing Units, and Signal flow.

Users create a MAX/MSP-like network to model a custom music system containing the processes they wish to use. They then choose an interaction style and interface with which to use the system.

Digital Processing and the Assessment of Violin Sound Quality

Colin Gough
University of Birmingham

The sound of a violin is controlled by the skill of the performer, the quality of the instrument and the acoustic in which it is played. Any attempt to assess the quality of sound is therefore dependent on all three factors. In this project, an electric violin is played by an expert player and the force exerted by the strings on the bridge is first recorded for later reproduction or in real time. Using a Signal Wizard sound processor, the signal is then convoluted with the previously recorded transient response of a number of violins of varying quality both close to the instrument, when the resulting sound is dominated by the response of the instrument, and at a distance, when the sound is dominated by the performing acoustic. This can be done using previously recorded samples or in real time. In this way, the relative importance of player, instrument and performing acoustic in determining the overall sound quality can be independently assessed Preliminary progress will be described and demonstrated.

Causal Contexts, Cognitive Cartoons and Spatial Sound

Peter Lennox* and Tony Myatt+
* - University of Derby, + - University of York

Based on previous work the proposal here is that spatial perception problems in artificial environments (e.g. spatial music displays) can be cast as a subset of the problems of cognitive mapping of the causal context that surrounds and supports the perceiver.

The intuitively available distinctions in these contexts of foreground and background, previously couched in terms of perceptual significance exist as externally valid causal distinctions; the task of perception is to cognitively represent these distinctions sufficiently for appropriate interaction. Effectively, this means that some items will "naturally" occupy attention, whilst others should equally naturally appeal to background, inattentive processes. Hence, aspects of the causal context will be accorded differing cognitive resources according to their significance, and some may be very sparsely represented in cartoon form. That is, perception engages in sophisticated information reduction in

cognitive representation in order to capitalise on available resources.

This poster outlines how causal contexts (including spatial matters) can be physically cartoonified in reciprocal manner to the dedicated perceptual mechanisms' operations, to economically and intuitively appeal to perception.

Performance Markup Language: Combining the score and the performance

Douglas McGilvray
University of Glasgow

A significant amount of information can be extracted from the audio recording of a musical performance. However, this information tells us little of the nature of musical performance and expression unless it can be analysed in the context of the score. For example, it is interesting to extract expressive gestures such as variations in tempo, timing or dynamics, but to understand the use of such gestures, they must be examined in the context of the structural information which can only be extracted from the score. Performance Markup Language (PML) is a specification for the representation of musical performance. It provides the framework necessary for the thorough and rigorous representation and analysis of musical performance, and the investigation of musical expression. It consists of an XML-based specification for the coordination and representation of three domains of symbolic musical information: score, performance & analysis. This presentation describes the design of the specification and the current implementation based on MusicXML.

Review of Musically Useful Audio Distortion Effects

Mark V. Oliver
Coventry University

This poster paper reviews musically useful audio distortion effects, in relationship to their cause, synthesis, utility and ergonomics. In audio engineering systems distortion is normally seen as an undesirable effect, which reduces the quality of music or speech. However, there is a well-known musical use of distortion. Non-linear distortion of a signal can create extra harmonics which can be used raw or subsequently processed. There are many audio processing system imperfections that have historically been used to create these harmonics. These are things like signal limiting/shaping that can occur as a result of overdriving instrument pickup, preamplifier, power amplifier, transformer or loudspeaker. Different technologies perform differently in this respect: for example valve distortion is often preferred because of its higher proportion of musically pleasing even harmonics.

The musician and recording community has for many years used analogue or digital units that emulated amplifier overdrive. Equally there is a market in novel distortion effects. The techniques involved in these are examined (including re-use of the original technology and analogue/digital/dsp replications). Examples are use of non-linearity modes of analogue components, clipping (hard and soft), nonlinear waveshaping (with/without memory, symmetrical or non-symmetrical, with/without pre/post processing), multiband waveshaping distortion. Issues that can arise are intermodulation and aliasing. The analysis of the distortion modes of existing components using conventional non-linear methods (similar to those used in control theory).

There are ‘traditional’ combinations of other effects with distortion. The use of wah-wah, flanging,

compression and feedback are common. Some of these combinations are used to increase the length/dramatic effect of a musical phrase or increase its harmonic interest.

The physical construction of the distortion effect is also important. They are normally portable floor units (with a couple of chunky rotary knobs), or rack-mount units. They may also be incorporated in amplifiers, recording equipment or the musical instrument itself. Consideration needs to be given to control placing (remembering that the musician who is using a distortion unit is often using both hands to play an instrument).

The exact signal that is fed into a distortion unit can affect the musical utility of a unit. Major chords often distort more pleasantly. The amount of a distortion effect is also important. A smaller quantity of distortion can be used for rhythm guitar and a larger amount for lead guitar. The more notes that are played simultaneously the less distortion is needed. The relationship between the distorted sound and the surrounding instruments in the ensemble is also important. Multiple distorted sounds can create a messy mixture that has little musical meaning.

Poisson Point Process Modelling for Polyphonic Music Transcription

Paul Peeling, Chung-fai Li, Simon Godsill
Cambridge University Engineering Department

We present a novel approach to modelling of polyphonic music through the use of Poisson point processes. The basic idea is that peaks detected in the frequency domain spectrum of a musical chord can be modelled as realisations of a non-homogeneous Poisson point process, such that peaks are most likely to be detected around the fundamental frequency of a note and its harmonics. This allows us to write down in a straightforward way the likelihood function for the detected frequency domain peaks conditioned on the (unknown) fundamental frequencies of multiple notes. A principal advantage of the approach is that when several notes are superimposed to make a chord, their individual Poisson process models combine to give another Poisson process, whose likelihood is still easy to compute. This contrasts with other peak modelling approaches, which require some kind of data association step linking individual harmonics explicitly with detected peaks in the spectrum, a problem that leads to a combinatorial explosion of terms when many harmonics and many notes are simultaneously present. The likelihood function in our approach is in an ideal form for implementation of likelihood-based or Bayesian inference about the unknown note frequencies in a chord, and for incorporation in more elaborate multi-frame approaches that model the evolution of chords and notes in music over time. Here we present proof-of-principle results for maximum likelihood estimation of fundamental frequencies in polyphonic piano music, showing very promising performance on real recorded piano music excerpts.

A Case Study in Digital Music Preservation With Calliope and DFDL

Stuart Pullinger and Nicholas J. Bailey
University of Glasgow

The rapid progress in the power of computer systems has led to equally rapid obsolescence of less powerful systems. The field of digital music preservation has arisen out of the necessity to prevent the loss of digitally-stored music when the system used for accessing it becomes obsolete. The open, descriptive qualities of eXtensible Markup Language (XML) make it an attractive choice for data preservation and many XML dialects exist for describing and preserving music. This paper introduces

a method for digital music preservation using a new XML-based dialect: the Data Format Description Language (DFDL) which provides a view over binary data from an annotated XML-Schema. This approach allows the underlying binary data to be preserved exactly whilst exposing it to standard XML tools such as the query language (XQuery) and conversion language (XML Stylesheet Transformations – XSLT).

A case study is presented which applies DFDL to the preservation of music notation created with the Calliope music notation editor, which became obsolete along with the NeXT platform. PorkPy, a tool to assist in the reverse-engineering of binary file formats, has been constructed and was successfully deployed in this case study.

Single Channel Audio Separation by Non-negative Matrix Factorization

Beiming Wang and Mark Plumbley
Queen Mary University of London

Musical audio is often a combination of human voice, sound of various instruments and more or less environmental noise.

The aim of our research is to develop a methodology for separating musical audio into streams of individual sound sources. Being able to so, it will provide widely enhancement to automatic music transcription, object coding, musical information retrieval and so on. Here we propose a novel approach for separating musical instruments in a single-channel audio recording using the Non-negative Matrix Factorization (NMF) algorithm and human interaction. A Graphical User Interface is also developed to aid this processing. The system has been tested on both artificially mixed audio and real musical recording with encouraging results.

Ontology-Based Annotation of Harmonic Structure

Jens Wissmann and Tillman Weyde
City University London

We present an OWL ontology for annotating the harmonic structure of music and rule-system for automatically generating the annotations on given harmony symbols. Our approach is implemented for the musical idiom of jazz. The ontology defines layers of harmonic analysis from chords symbols to the level of a complete piece. The ontology is complemented by a rule system that produces the annotations automatically. In connection with manual annotation, this approach can support a number of scenarios in music production, education, retrieval and in musicological research.

Key words: Annotation, Music, Harmony, Ontology, Rule System, Automatic Music Analysis

---

Adaptive Spectrogram

Wen Xue
Queen Mary University of London

This is a small introduction to the adaptive spectrogram, which represents an audio signal with adaptive resolution, i.e. a different resolution is selected for each area in the time-frequency plane so that the signal is optimally represented. The poster takes an easy-to-understand way of splitting the time-frequency space, and derive the adaptive spectrogram from this framework, in comparison with wavelet packets and local cosines. While the two provides frequency-dependent/time-dependent resolution, the adaptive spectrogram allows a resolution both time- and frequency- dependent.