Vol.2009-MUS-82 No.1
2009/11/2

特別セッション：ISMIR コミュニティからの招待講演

帆足 啓一郎　KDDI 研究所
藤島 琢哉　ヤマハ株式会社
後藤 真孝　産業技術総合研究所

2009 年 10 月 26 日〜30 日に，神戸で開催された国際会議 ISMIR 2009 (10th International Society for Music Information Retrieval Conference) との連携企画として，国際的に活躍されている下記 5 名の研究者 (敬称略，アルファベット順) による招待講演セッションを開催する．

- Simon Dixon (Queen Mary University of London, UK)
- Anssi Klapuri (Tampere University of Technology, Finland)
- Geoffroy Peeters (IRCAM, France)
- Xavier Serra (Universitat Pompeu Fabra, Spain)
- Malcolm Slaney (Yahoo! Research, USA)

本稿では，各招待講演の題目と概要，ならびに講演者のプロフィール情報を示す．

Special Session: Invited Talks from the ISMIR Community

Keiichiro Hoashi　KDDI R&D Laboratories Inc.
Takuya Fujishima　Yamaha Corporation
Masataka Goto　National Institute of Advanced Industrial Science and Technology (AIST)

The Special SIGMUS Invited Talk Session is held in conjunction with ISMIR 2009 (10th International Society for Music Information Retrieval Conference, October 26-30, 2009) at Kobe, Japan. In this session, we have invited the five following speakers (listed in alphabetical order), who all are established researchers in the field of music information research.

- Simon Dixon (Queen Mary University of London, UK)
- Anssi Klapuri (Tampere University of Technology, Finland)
- Geoffroy Peeters (IRCAM, France)
- Xavier Serra (Universitat Pompeu Fabra, Spain)
- Malcolm Slaney (Yahoo! Research, USA)

This paper lists the titles and abstracts of all invited talks, as well as the biographies of the speakers.

Vol.2009-MUS-82 No.1
2009/11/2

Invited Talk 1:
"The Centre for Digital Music at Queen Mary University of London"

Simon Dixon (Queen Mary University of London, UK)

Abstract

The Centre for Digital Music (C4DM) at Queen Mary University of London is a world-leading multidisciplinary research group in the field of audio and music technology, consisting of around 50 full-time members (academic staff, postdoctoral researchers and postgraduate students). Our research covers a wide range of topics related to music synthesis, analysis, processing, production, delivery and retrieval, including the fields of music informatics, music signal processing, audio engineering, machine listening, interactive music systems and auditory display. Computational techniques at the heart of the Centre's research include time-frequency and time-scale analysis, neural networks, hidden Markov models, dynamic Bayesian networks, matching pursuits, transient analysis, independent component analysis, blind source separation, sparse representations and knowledge discovery, which have been applied to problems as diverse as automatic music transcription, beat tracking, audio alignment, music segmentation, automatic mixing, music recommendation, instrument identification and automatic musical accompaniment. Another particular focus of the group has been on semantic audio, and more recently the semantic web for music, where we have played a leading role in the development of the Music Ontology. In this talk I will give an overview of research at C4DM, making particular mention of recent results from the OMRAS-2 (Online Music Recommendation and Searching 2) project.

Biography

Simon Dixon is a lecturer in the Centre for Digital Music at Queen Mary University of London. He received BSc and PhD degrees in computer science from the University of Sydney, and AMusA and LMusA diplomas in classical guitar. He was a lecturer at Flinders University of South Australia (1994-1999) and then a research scientist at the Austrian Research Institute for Artificial Intelligence (1999-2006). His research interests focus on the extraction and processing of musical (particularly rhythmic and harmonic) content in audio signals, including beat tracking, onset detection, alignment, automatic transcription, and the measurement and visualisation of expression in music performance.

Vol.2009-MUS-82 No.1
2009/11/2

Invited Talk 2:
"Audio Research Group at Tampere University of Technology, Finland"

Anssi Klapuri (Tampere University of Technology, Finland)

Abstract

Audio Research Group at Tampere University of Technology, Finland, consists of a bit less than 20 people working on various aspects of audio signal processing. This talk introduces the group's activity on music signal processing and analysis. First, some applications of music transcription are introduced. A semiautomatic transcription tool is demonstrated which allows the user to write the score of a piece with the help of automatic analysis tools. Other applications include music retrieval and a karaoke system where the melody of a music piece is removed and the user's voice is tuned to replace that. Another topic discussed is music structure analysis. A method is introduced which segments a piece into parts and tries to recognise their names, such as "verse" or "chorus". In the latter half of the talk, the group's recent work on sound source modeling and separation is discussed. This comprises both spectral modeling and modeling of the temporal evolution of musical sounds. For sound separation, a method is presented which combines a structured source model with multipitch estimation and non-negative matrix factorization to handle complex polyphonic music signals.

Biography

Anssi Klapuri received the M.Sc. and Ph.D. degrees from Tampere University of Technology (TUT), Tampere, Finland, in 1998 and 2004, respectively. In 2005, he spent six months at the Ecole Centrale de Lille, Lille, France, working on music signal processing. In 2006, he spent three months visiting the Signal Processing Laboratory of Cambridge Univerisity, Cambridge, U.K. He is currently heading the Audio Research Group at the Department of Signal Processing, TUT. His research interests include audio signal processing, auditory modeling, and machine learning.

Invited Talk 3:
"Sound analysis/synthesis team and music indexing activities at IRCAM"

Geoffroy Peeters (IRCAM, France)

Abstract

IRCAM (Institute of Research and Coordination in Acoustic/Music) was created in 1977 by the French composer and conductor Pierre Boulez. Ircam is both a research and a music creation centre. The research and development department welcome over 90 researchers in various fields related to music: instrument acoustic, room acoustic, digital signal processing, computer science (language, real-time system, user interface and database), musicology, music representation, music cognition and perception, sound design.

Ircam is both connected to university – by welcoming two Masters diploma and numerous PHD students – and to the industry – by producing over 6 professional softwares – and by collaborating with many private companies (Renault, Orange, Creative Labs, MakeMusic, Thomson, EMI, Sony). The sound analysis/synthesis team focus on the use of digital signal processing for music creations. Algorithms are developed for signal re-synthesis (sinusoidal modelling), synthesis from scratch (Chant synthesis), signal modifications (phase vocoder, p-sola) or text-to-speech (corpus-based concatenative synthesis). These technologies are currently used by composers, by sound designers, in movie or games production.

Activities related to music indexing started in 1998, with the Studio-On-Line project, the first online large-sound-database with search-by-content facilities. Over the years and the successive national or European projects (Cuidado, SemanticHIFI, Quaero) Ircam has developed technologies for content-description related to most music facets: multi-pitch, beat, down-beat, meter, chords, key/ mode, music structure, singing voice, music genre/ mood/ tag, sound/music similarity using various methods ranging from low-level, to score estimation and source separation methods. Content-description relies on a large part on annotated data. Recent research activities focus on the development of annotation concepts and collection suitable for audio indexing and musically relevant for nowadays music. Content-description is commonly used for facilitating data access,

Vol.2009-MUS-82 No.1
2009/11/2

it can also be used for sound creation or modification as was studied in the Orchestration project or is used in corpus-based concatenative music synthesis (musaicing). In this talk, we will review these various activities of Ircam.

Biography

Geoffroy Peeters is a researcher at IRCAM (Institute of Research and Coordination in Acoustic and Music) in Paris, France and is currently leading the music indexing research activities in the Quaero project. He received his Ph.D. in computer science from the University Paris VI, France in 2001 during which he developed new signal processing algorithms for speech and audio processing. Since then, his research focus on signal processing and pattern matching applied to audio and music indexing: timbre description, sound classification, audio identification, rhythm description, music structure discovery, audio summary and music genre / mood recognition. He is co-author of the ISO MPEG-7 audio standard.

Invited Talk 4:
"The Music Technology Group of the Universitat Pompeu Fabra in Barcelona"

Xavier Serra (Universitat Pompeu Fabra, Spain)

Abstract

The Music Technology Group (MTG) of the Universitat Pompeu Fabra in Barcelona, part of its Department of Information and Communication Technologies and of its Audiovisual Institute, is specialized in sound and music computing. With more than 40 researchers coming from different and complementary disciplines, the MTG carries out research on topics such as sound processing and synthesis; music content description; interactive music systems; computational models of perceptual and music cognition; and the technologies related with music social networks. The MTG wants to contribute to the improvement of the technologies related to sound and music communication, carrying out competitive research at the international level and at the same time transferring its results to society. To that goal, the MTG aims at finding a balance between basic and applied research and at the same time promotes interdisciplinary approaches that incorporate knowledge from both scientific/technological and humanistic/artistic disciplines.

In this talk I will first overview the different research lines of the MTG and then I will talk about one of our latest projects: Freesound.org, which is a platform for the open exchange of sounds. With more than 1 million registered users, 75thousand sounds in the database, and averaging 30 thousand visitors a day, this site is becoming much more than just a repository of sounds. By including social networking services, technologies for automatic tagging and for content based searching, support for research and artistic uses, and by promoting projects that built on that to experiment with collaborative production ideas, Freesound.org is a great platform to explore and develop new social networking concepts. In this talk I will analyze and describe Freesound.org, both from a technical and a social perspective, and I will present ideas for its future development. In particular I will introduce the idea of Music 3.0, concept based on the integration of the most recent technologies of the Web 2.0, advance on-line tools for music creation, and large sound and music repositories.

Vol.2009-MUS-82 No.1
2009/11/2

Biography

Xavier Serra is the head of the Music Technology Group of the Universitat Pompeu Fabra in Barcelona, Spain. After a multidisciplinary academic education he obtained a PhD in Computer Music from Stanford University in 1989 with a dissertation on the spectral processing of musical sounds that is considered a key reference in the field. His research interests cover the understanding, modeling and generation of musical signals by computational means, with a balance between basic and applied research and approaches from both scientific/technological and humanistic/artistic disciplines.

Invited Talk 5: "Finding Music on the Web: A Yahoo Perspective"

Malcolm Slaney (Yahoo! Research, USA)

Abstract

Without a doubt the Internet has changed the way people consume music. But it also brings a wealth of data and new opportunities for music-information retrieval services. Our goal is to connect users with their entertainment and information needs.

The data is both plentiful and noisy. We have billions of ratings by users about their musical interests. One one hand the the large amount of data means we can build robust models. On the other hand, the data does come from people, with all their idiosyncratic behavior and opinions. This wealth of personal data — we have to assume it is all correct — sometimes means what we think it means, and other times represents personal behaviors unrelated to anybody else's opinion. Separating out the signal from the noise is the new frontier for web sciences.

I'll illustrate my talk with several kinds of technologies we find interesting, drawing from successes we have had from all types of multimedia. These approaches impact recommendations, tagging, and search. The frontiers of web science are wonderful.

Biography

Malcolm Slaney is a principal scientist at Yahoo! Research Laboratory. He received his PhD from Purdue University for his work on computed imaging. He is a coauthor, with A. C. Kak, of the IEEE book "Principles of Computerized Tomographic Imaging." This book was recently republished by SIAM in their "Classics in Applied Mathematics" Series. He is coeditor, with Steven Greenberg, of the book "Computational Models of Auditory Function." Before Yahoo!, he has worked at Bell Laboratory, Schlumberger Palo Alto Research, Apple Computer, Interval Research and IBM's Almaden Research Center. He is also a (consulting) Professor at Stanford's CCRMA, where he organizes and teaches the Hearing Seminar. His research interests include auditory modeling and perception, multimedia analysis and synthesis, music similarity and audio search, and machine learning. For the last several years he has lead the auditory group at the Telluride Neuromorphic Workshop.