Tutorial 1

Open Source and Reproducible MIR Research

Brian McFee and Thor Kell

Abstract
The goal of this tutorial is to provide hands-on, practical training for MIR researchers to learn modern tools for improving the quality of their research software.

We will provide a project template repository, which we will use as scaffolding to demonstrate practices and techniques including version control, continuous integration, automatic documentation, environments & software dependency tracking, and packaging & distribution. While many of the techniques we will cover are independent of language or programming environment, the practices are best demonstrated by example, and we expect Python to be the most accessible and generally useful language for the expected attendees.

The tutorial will begin with a brief introduction to version control with git and GitHub. We will only cover the basics: creating an account, cloning, pushing, and pulling. We will walk attendees through the process of making their own copy of the repository, verifying that tests and documentation work, implementing some simple functionality, and working with automatic testing. The goal of this exercise is to expose attendees to modern development practices, so that their software is continuously tested during development. Attendees will then learn how to create documentation with Sphinx and automate the process with the ReadTheDocs service. We will teach standard documentation style, and provide an exercise in which attendees write documentation for a previously undocumented function. We will also discuss how to write README files and how to make a package easy for others to use. We note that many of the tools we will demonstrate are provided by web services (Travis, GitHub, etc.), which have freely available counterparts that can run on local servers (Jenkins, 8 GitLab, 9 etc.). We will provide pointers to these alternative implementations, but for ease of exposition and simplicity, we will stick to the web-based services for the tutorial.

Finally, we will close with a tour of the Python/MIR ecosystem. This part of the tutorial will give an overview of the available packages commonly used in MIR research, such as the scipy stack, pandas, librosa, mir eval, and jupyter. We will spend some extra time on Jupyter, including plotting and sonification, as well as discussing how to save and iterate on notebooks.

Brian McFee is a Moore-Sloan Fellow at New York University's Center for Data Science and Music and Audio Research Lab (MARL). He received the B.S. degree (2003) in Computer Science from the University of California, Santa Cruz, and M.S. (2008) and Ph.D. (2012) degrees in Computer Science and Engineering from the University of California, San Diego. His work touches on various topics at the intersection of machine learning, information retrieval, and audio analysis. He is an active open source software developer, and the principal maintainer of the librosa package for music and audio signal processing. His favorite genre is "chip-tune", and he likes dogs.

Thor Kell earned his B.Sc in Computer Science & Music from the University of Victoria in 2011, and an M.A. in Music Technology from McGill in 2015. He has worked for SoundCloud, The Echo Nest, and is currently an engineer at Spotify. His interests include algorithmic music creation as well as track ordering and automatic mixing for DJ sets. His favorite genre is "post", and he likes cats.

Tutorial 2

Computational Approaches for Analysis of Non-Western Music Traditions

Xavier Serra, Martin Clayton, and Barış Bozkurt

Abstract

The main goal of this tutorial is to present an overview for computational analysis applied on non-Western music traditions and within this context, also discuss the role of musicology perspective in Music Information Retrieval. The target audience is researchers and students interested in MIR and computational musicology. The tutorial comprises of three parts.

The first two parts are dedicated to critical overview of recent studies with engineering and musicology perspective. We will start by presenting some of the relevant problems and challenges for analysis of non-Western music traditions that have been studied from an MIR perspective. In the second part we discuss some current and potential challenges posed by musicological research in diverse musical genres, including in rhythmic analysis.

The last part will present resources created during the CompMusic project, that are openly available to the community. It will include demonstrations for accessing non-Western music data and processing these data (focusing on intonation and rhythm analysis) with publicly available tools. We will consider research corpora from various music traditions: Hindustani (North India), Carnatic (South India), Turkish-makam (Turkey), Arab-Andalusian (Maghreb), and Beijing Opera (China). While we will focus on a few culture-specific MIR tasks for our demonstrations, we will also discuss open research problems that can be studied using these datasets. This session would serve as a quick start for students and researchers without prior experience in analysis of non-Western music and will provide them a good entry point for further investigation.

Xavier Serra is a Professor of the Department of Information and Communication Technologies and Director of the Music Technology Group at the Universitat Pompeu Fabra in Barcelona. 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 analysis, description and synthesis of sound and music signals, with a balance between basic and applied research and approaches from both scientific/technological and humanistic/artistic disciplines. Dr. Serra is very active in promoting initiatives in the field of Sound and Music Computing at the local and international levels, being involved in the editorial board of a number of journals and conferences and giving lectures on current and future challenges of the field. He has been awarded an Advanced Grant of the European Research Council to carry out the project CompMusic aimed at promoting multicultural approaches in music computing research which produced the resources this tutorial will present.

Martin Clayton is Professor in Ethnomusicology in Durham University. He studied at the School of Oriental and African Studies (SOAS) in London, where he obtained degrees in Music and Hindi (BA, 1988) and Ethnomusicology (PhD, 1993). His research interests

include Hindustani (North Indian) classical music, rhythmic analysis, musical entrainment and embodiment, comparative musicology and early field recordings, British-Asian music and Western music in India. He currently directs the research project Interpersonal Entrainment in Music Performance, a collaborative, multidisciplinary effort to understand ensemble sychronisation and coordination cross-culturally.

Barış Bozkurt is a Post-doctoral researcher in the Music Technology Group at the Universitat Pompeu Fabra in Barcelona. In the first phase of his research career, he dedicated to development of signal processing algorithms for speech analysis and synthesis. He has obtained his PhD degree in 2005 from Faculte Polytechnique De Mons, Belgium and also worked in speech industry after his PhD. Since 2007, he has been teaching in Electrical Engineering and Computer Science departments, and carrying research in the fields of audio signal processing and computational musicology.

Tutorial 3

Statistical Analysis of Results in Music Information Retrieval: Why and How

Julián Urbano and Arthur Flexer

Abstract

Nearly since the beginning, the ISMIR and MIREX communities have promoted rigor in experimentation through the creation of datasets and the practice of statistical hypothesis testing to determine the reliability of the improvements observed with those datasets. In fact, MIR researchers have adopted a certain way of going about statistical testing, namely non-parametric approaches like the Friedman test and multiple comparisons corrections like Tukey's. In a way, they have become a standard of reporting and judging results for researchers, reviewers, committees, journal editors, etc. It is nowadays more frequent to require statistically significant improvements over a baseline with a well-established dataset.

But hypothesis testing can be very misleading if not well understood. To many researchers, especially newcomers, even the simpler analyses and tests are seen as a black box where one puts performance scores and gets a p-value which, as they are told, must be smaller than 0.05. Therefore, significance tests are in part responsible of determining what gets published, what research lines to follow, and what project to fund, so it is very important to understand what they really mean and how they should be carried out and interpreted. We will also focus on experimental validity, and will show how a lack of internal or external validity, even if experiments are reliable and repeatable and hypothesis testing is done correctly, can render even your best results invalid. Problems discussed include adversarial examples or the lack of inter-rater agreement when annotating ground truth data.

The goal of this tutorial is to help MIR researchers and developers get a better understanding of how these statistical methods work and how they should be interpreted. Starting from the very beginning of the evaluation process, it will show that statistical analysis is always required, but that too much focus on it, or the incorrect approach, is just harmful. The tutorial will attempt to provide better insight into statistical analysis of results, present better solutions and guidelines, and point the attendees to the larger but ignored problems of evaluation and reproducibility in MIR.

Julián Urbano is an Assistant Professor at Delft University of Technology, The Netherlands. His research is primarily concerned with evaluation in IR, working in both the music and text domains. Current topics of interest are the application of statistical methods for the construction of datasets, the reliability of evaluation experiments, statistical significance testing for IR, low-cost evaluation and stochastic simulation for evaluation. He has published over 50 research papers in related venues like Foundations and Trends in IR, the IR Journal,

the Journal of Multimedia IR, ISMIR, CMMR, ACM SIGIR, ACM CIKM and ECIR, winning two best paper awards and a best reviewer award. He has been active in the ISMIR community since 2010, both as author and PC member, and is also co-organizer of the MediaEval AcousticBrainz task. He is reviewer for other conferences and journals, such as ACM CIKM, HCOMP, IEEE TASLP, IEEE MM, ACM TWEB, IEEE TKDE or the Information Sciences journal.

Arthur Flexer is a senior researcher, project manager and vice-head at the `Intelligent Music Processing and Machine Learning Group' of the Austrian Research Institute for Artificial Intelligence (OFAI). He has twenty-five years of experience in basic research on machine learning with an emphasis on applications to music in the last twelve years. He holds a PhD in psychology which provides him with the necessary background concerning design and evaluation of experiments. He has published comprehensively on the role of experiments and on problems of ground truth and inter-rater agreement, all in the field of MIR. He is author and co-author of more than 80 peer-reviewed articles. He has been active in the ISMIR community since 2005 and has also published in related venues like DAFx, SMC, ECIR, Journal of Machine Learning Research and Journal of New Music Research. He is a member of the editorial board of the Transactions of the International Society for Music Information Retrieval (TISMIR).

Tutorial 4

Music Separation with DNNs: Making It Work

Antoine Liutkus and Fabian-Robert Stöter

Abstract

This tutorial concerns music source separation, that we also call music demixing, with a resolute focus on methods using DNN.

- In an introductory part, we will motivate the tutorial by explaining how music separation with DNN emerged with data-driven methods coming from machine-learning or image processing communities. This comes with machine-learning tricks to make methods work in practice. Meanwhile, many audio processing good practices are often forgotten or not correctly applied, although they are mandatory for good performance.

- In a second part, we present and discuss the few concepts that are mandatory to design a source separation method. Each point will firstly be the focus of screencasting from an interactive notebook session that all the audience will be invited to, and then will also be explained with a theoretical presentation when appropriate. The whole tutorial will be thus split into practical hands-on sessions using online interactive Python sessions and more classical theoretical insights.

- The third part of the tutorial provides some feedback on what seems to be important to get good performance in practice, with a focus on the training stage. On the one hand, many of the tricks discussed there are not often discussed in papers because a lot of them are negative results that are hard to publish: some interesting ideas that turn out ineffective yet. On the other hand, we also show how some very simple things make a huge difference in practice.

- In the following part, we pick one single system, resulting from the previous discussion, and show how its performance can be dramatically improved by using just a few simple tricks at test time, including resynthesis methods, filtering tricks, and how to go stereo.

This tutorial is first targeted at PhD students and at engineers, that want to implement audio demixing methods in practice and to achieve state of the art performance while keeping highly readable code. Second, by showing how pytorch enables easy design and debugging, including new cost functions, architectures, etc., it will hopefully be of interest to researchers wondering how to do actual investigations on audio with DNNs, without being just users of high-level black-box systems.

Antoine Liutkus received the State Engineering degree from Télécom ParisTech, France, in 2005, and the M.Sc. degree in acoustics, computer science and signal processing applied to music (ATIAM) from the Université Pierre et Marie Curie (Paris VI), Paris, in 2005. He worked as a research engineer on source separation at Audionamix from 2007 to 2010 and obtained his PhD in electrical engineering at Télécom ParisTech in 2012. He is currently researcher at Inria, France. His research interests include audio source separation and machine learning.

Fabian-Robert Stöter received the diploma degree in electrical engineering in 2012 from the Leibniz Universität Hannover and worked towards his Ph.D. degree in audio signal processing in the research group of B. Edler at the International Audio Laboratories Erlangen,

Germany. He is currently researcher at Inria, France. His research interests include supervised and unsupervised methods for audio source separation and signal analysis of highly overlapped sources.

Tutorial 5

Deep Learning for MIR

Alexander Schindler, Thomas Lidy, and Sebastian Böck

Abstract

Deep Learning has become state of the art in visual computing and continuously emerges into the Music Information Retrieval (MIR) and audio retrieval domain. To bring attention to this topic we provide an introductory tutorial on deep learning for MIR. Besides a general introduction to neural networks, the tutorial covers a wide range of MIR relevant deep learning approaches. Convolutional Neural Networks are currently a de-facto standard for deep learning based audio retrieval. Recurrent Neural Networks have proven to be effective in onset detection tasks such as beat or audio-event detection. Siamese Networks have shown to be effective in learning audio representations and distance functions specific for music similarity retrieval. We introduce these different neural network layer types and architectures on the basis of standard MIR tasks such as music classification, similarity estimation and onset detection. We will incorporate both academic and industrial points of view into the tutorial. The tutorial will be accompanied by a Github repository for the presented content as well as references to state of the art work and literature for further reading. This repository will remain public after the conference.

Alexander Schindler is member of the Music Information Retrieval group at the Technical University since 2010 where he actively participates in research, various international projects and currently finishes his Ph.D on audio-visual analysis of music videos. He participates in teaching MIR, machine learning and DataScience. Alexander is currently employed as scientist at the AIT Austrian Institute of Technology where he is responsible for establishing a deep learning group. In various projects he focusses on deep-learning based audio-classification, audio event-detection and audio-similiarity retrieval tasks.

Thomas Lidy has been a researcher in music information retrieval in combination with machine learning at TU Wien since 2004. Since 2015, he has been focusing on how Deep Learning can further improve music & audio analysis, winning 3 international benchmarking contests. He is currently the Head of Machine Learning at Musimap, a company that uses Deep Learning to analyze styles, moods and emotions in the global music catalog, in order to create emotion-aware search & recommender engines that empower music supervisors to find the music for their needs and music streaming platforms to deliver the perfect playlists according to people's mood.

Sebastian Böck received his diploma degree in electrical engineering from the Technical University in Munich in 2010 and his PhD in computer science from the Johannes Kepler University Linz. He continued his research at the Austrian Research Institute for Artificial Intelligence (OFAI) and recently also joined the MIR team at the Technical University of Vienna. His main research topic is the analysis of time event series in music signals, with a strong focus on artificial neural networks.

Tutorial 6

Fundamental Frequency Estimation in Music

Rachel Bittner, Alain de Cheveigné, and Johanna Devaney

Abstract
This tutorial will cover the core concepts in fundamental frequency estimation, starting with the monophonic case and building up to the polyphonic case. Topics will include pitch perception, common algorithms, data and annotation, and evaluation metrics. This tutorial aims to provide the audience with an understanding of the challenges, the common approaches, and the open problems in f0 estimation, as well as the potential applications of f0 estimation to transcription, music performance analysis, and source separation problems.

Rachel Bittner is a Research Scientist at Spotify in New York City, and recently completed her Ph.D. at the Music and Audio Research Lab at New York University under Dr. Juan P. Bello. Previously, she was a research assistant at NASA Ames Research Center working with Durand Begault in the Advanced Controls and Displays Laboratory. She did her master's degree in math at NYU's Courant Institute, and her bachelor's degree in music performance and math at UC 2 Irvine. Her research interests are at the intersection of audio signal processing and machine learning, applied to musical audio. Her dissertation work applied machine learning to various types of fundamental frequency estimation.

Alain de Cheveigné trained in physics, maths and neuroscience at Université Pierre et Marie Curie, and is currently Senior Scientist with the Centre National de la Recherche Scientifique (CNRS) in France, affiliated with Ecole normale supérieure (Paris) and UCL (London). He is active in psychophysics and modeling of auditory perception, and data processing for audio and electrophysiological signals. He is author of the widely used YIN method for f0 estimation, and of several widely cited chapters on pitch perception. He currently heads a H2020 project on the cognitive control of hearing aids.

Johanna Devaney is an Assistant Professor of Music Technology at Brooklyn College, City University of New York and the speciality chief editor for the Digital Musicology section of Frontiers in Digital Humanities. Previously she taught in the Music Technology program at NYU Steinhardt and the Music Theory and Cognition program at Ohio State University. Johanna completed her post-doc at the Center for New Music and Audio Technologies (CNMAT) at the University of California at Berkeley and her PhD in music technology at the Schulich School of Music of McGill University. She also holds an MPhil degree in music theory from Columbia University, as well as an MA in composition from York University in Toronto. Johanna's research seeks to understand how humans engage with music, primarily through performance, with a particular focus on intonation in the singing voice, and how computers can be used to model and augment our understanding of this engagement.

Tutorial 7

Optical Music Recognition for Dummies

Jorge Calvo-Zaragoza, Jan Hajič jr, Alexander Pacha, and Ichiro Fujinaga

Abstract

Optical Music Recognition (OMR) is a field of research that investigates how to computationally decode music notation in documents. As most musical compositions in the Western tradition have been written rather than recorded, bringing this music into the digital domain can significantly diversify the sources for MIR, digital musicology, and more broadly lower the costs of introducing previously unheard works to audiences worldwide. While OMR has been regarded as a largely unsolved problem, this situation has recently shifted: new large-scale datasets and tools have been released, methods based on deep learning are successfully dealing with musical symbol detection and partial end-to-end recognition, and applications of OMR such as retrieval have started migrating from article introductions to the Results sections.

Our tutorial will present this new and rather exciting state of the art in OMR. We will demonstrate recent methods and results, introduce the audience to the tools and datasets used to achieve them, and showcase the opportunities for using OMR. Finally, we will introduce the current challenges in OMR.

After the tutorial, the participants should be familiar with state-of-the-art OMR research, and should be able to start using existing tools to integrate OMR into their own work, whether in MIR or (digital) musicology. For those interested in working on OMR themselves, the tutorial should provide a head start. The tutorial will be hands-on: if you wish to get the most out of it, be ready to follow jupyter notebooks.

Jorge Calvo-Zaragoza received his PhD degree in computer science from the University of Alicante (Spain) in 2016. He joined the Single Interface for Music Score Searching and Analysis (SIMSSA) project as Postdoctoral Fellow in 2017. He is currently the recipient of a Juan de la Cierva postdoctoral grant from the Spanish government. His scientific contribution has focused so far on ancient manuscripts, in notations like neumatic or mensural. He has also made important contributions to the pre-processing of music score images such as the development of algorithms to separate the elementary graphic layers of the document (i.e., staff, notes, text, or background). He has authored more than 20 papers about Optical Music Recognition in peer-reviewed journals and international conferences.

Jan Hajič jr. is a Ph.D. student at Charles University (Czech Republic), where he is working on Optical Music Recognition and Deep Learning since 2016 at the Center of Excellence for Multimodal Data Interpretation project. He has especially contributed to infrastructure for OMR, creating e.g. the first full-pipeline OMR dataset (MUSCIMA++). He also has a strong musical background, having studied composition at the Janáček Academy of Music and Performing Arts.

Alexander Pacha received his M.Sc. with honors in computer science from the TUM, University of Augsburg and LMU Munich (Germany) in 2013. He is a professional software engineer, trainer for clean code programming and a passionate musician and composer. Since 2017 he is a PhD student at the TU Wien (Austria) working on Optical Music Recognition and Deep Learning.

Ichiro Fujinaga is an Associate Professor and the Chair of the Music Technology Area at the Schulich School of Music at McGill University. He has Bachelor's degrees in Music/Percussion and Mathematics from University of Alberta and a Master's degree in Music Theory and a Ph.D. in Music Technology from McGill. He is currently directing a large optical music recognition and analysis project called Single Interface for Music Score Searching and Analysis (SIMSSA).

Tutorial 8

Overview and New Challenges of Music Recommendation Research in 2018

Markus Schedl, Peter Knees, and Fabien Gouyon

Abstract

The current revolution in the music industry represents great opportunities and challenges for music recommendation systems. Recommendation systems are now central to music streaming platforms, which are rapidly increasing in listenership and becoming the top source of revenue for the music industry. It is increasingly more common for a music listener to simply access music than to purchase and own it in a personal collection. In this scenario, recommendation calls no longer for a one-shot recommendation for the purpose of a track or album purchase, but for a recommendation of a listening experience, comprising a very wide range of challenges, such as sequential recommendation, or conversational and contextual recommendations. Recommendation technologies now impact all actors in the rich and complex music industry ecosystem (listeners, labels, music makers and producers, concert halls, advertisers, etc.).

To acknowledge these developments, we give an introductory tutorial providing an overview of music recommendation research, as well as the challenges it faces today. We focus on three use cases: automatic playlist generation, context-aware music recommendation, and recommendation in the creative process of music making.

Markus Schedl is an Associate Professor at the Johannes Kepler University Linz / Department of Computational Perception. He graduated in Computer Science from the Vienna University of Technology and earned his Ph.D. from the Johannes Kepler University Linz. Markus further studied International Business Administration at the Vienna University of Economics and Business Administration as well as at the Handelshögskolan of the University of Gothenburg, which led to a Master's degree. His main research interests include web and social media mining, information retrieval, multimedia, and music information research. He has been an active member of the MIR community since 14 years and since then co-authored almost 200 peer-reviewed research articles.

Peter Knees is Assistant Professor of the Faculty of Informatics, TU Wien, Austria. For over a decade he has been an active member of the ISMIR community, reaching out to the related areas of multimedia, text IR, and recommender systems. Apart from serving on the program committees of major conferences in the field, he has organized several workshops on topics of media retrieval. He is an experienced teacher of graduate-level courses on recommender systems and information retrieval and has given tutorials on music information retrieval at RecSys, SIGIR, ECIR, RuSSIR, and the Indonesian Summer School on MIR.

Fabien Gouyon is Principal Scientist at the music streaming service Pandora, where he does applied research on personalized music recommendation, and works with the Music Genome Project. Before joining Pandora, he received a PhD in Computer Science while working in the Music Technology Group in the University Pompeu Fabra in Barcelona, and was a co-founder of Barcelona Music and Audio Technologies (BMAT), worked in the Austrian Research Institute for Artificial Intelligence in Vienna, and started and led the Sound and Music Computing Group while teaching at the University of Porto. He was President of ISMIR in 2016-2017.