ISMIR 2021 Tutorial

Teaching Music Information Retrieval

George Tzanetakis

University of Victoria

2021

Table of Contents I

Table of Contents II

Introduction

Music Information Retrieval as a research field is now more than 20 years old. I have been involved in teaching MIR in undergraduate, graduate, and tutorial courses to students from a variety of disciplines for over 15 years. In this tutorial we will explore different aspects of teaching MIR and share what I have learned over the years about how to make the teaching of MIR topics more effective.

Tutorial Structure

There are 6 modules each with a duration of approximately 25 minutes followed by 5 minutes of questions. Each unit consists of a set of slides and in some cases some associated hands-on demonstrations in the form of Jupyter/Python notebooks. The target audience is anyone (professor, researcher, postdoc, graduate student) who is interested in teaching MIR. I assume that participants are familiar with the main ideas and tasks of MIR.

Modules

- Overview and organization
- Adapting to a target audience and format
- Online learning
- Assessment
- Projects
- Resources

Relevant Background

- Main focus of my research has been Music Information Retrieval (MIR)
- Involved from the early days of the field
- Have published papers in almost every ISMIR conference
- Organized ISMIR in 2006
- Tutorials on various MIR topics in several conferences
- Taught MIR as a 4th year CS course for 12 times
- Kadenze MIR program (3 courses) approximately 3000 participants

Education and Academic Work Experience

- 1997 BSc in Computer Science (CS), University of Crete, Greece
- 1999 MA in CS, Princeton University, USA
- 2002 PhD in CS, Princeton University, USA
- 2003 PostDoc in CS, Carnegie Mellon University, USA
- 2004 Assistant Professor in CS, Univ. of Victoria, Canada
- 2010 Associate Professor in CS, Univ. of Victoria, Canada
- 2016 Professor in CS, Univ. of VIctoria, Canada
- 2010-2020 Canada Research Chair (Tier II) in Computer Analysis of Audio and Music
- Music theory, saxophone and piano performance, composition, improvisation both in conservatory and academic settings

Work Experience beyond Academia

Many internships in research labs throughout studies. Several consulting jobs while in academia. A few representative examples:

- Moodlogic Inc (2000). Designed and developed one of the earliest audio fingerprinting systems (patented) - 100000 users matching to 1.5 million songs
- Teligence Inc (2005). Automatic male/female voice discrimination for voice messages used in popular phone dating sites - processing of 20000+ recordings per day.
- Smule (2015-present) - various MIR related projects

Software - Marsyas

- Music Analysis, Retrieval and Synthesis for Audio Signals
- Open source in C++ with Python Bindings
- Started by me in 1999 - core team approximately 4-5 developers
- Approximately 400 downloads per month
- Many projects in industry and academia
- State-of-the-art performance while frequently orders of magnitude faster than other systems
- Not actively developed for the last 5-6 years

Visiting Scientist at Google Research

Six month study leave. Things I worked on (of course as part of larger teams):

- Cover Song Detection (applied to every uploaded YouTube video).
- Audio Fingerprinting
- Named inventor on 6 pending US patents related to audio matching and fingerprinting

Table of Contents I

Table of Contents II

UVic MIR Course - Learning outcomes

- Basic knowledge of DSP
- Basic knowledge of Machine Learning (ML)
- Basic knowledge of Music Theory
- Familiarity with the basic tasks that have been explored in MIR research and the algorithms used to solve them
- **Being able to read and understand the majority of published literature in ISMIR**
- Experience with designing and developing MIR algorithms and systems

UVic MIR Course - Lectures

University of Victoria Computer Science

- 2 lectures 1.5 hours /week
- 3 hours of associated homework expected for each lecture
- Total weekly commitment approximately 9-10 hours
- 6 assignments each worth 10% of final grade and done invididually.
- Assignments typically will combine some reading and understanding of MIR literature as well as some programming of MIR algorithms
- 1 final group project (2-3 students per group) 40%

History of MIR before computers

How did a listener encounter a new piece of music throughout history ?

- Live performance
- Music Notation
- Physical recording
- Radio

Brief History of computer MIR

- Pre-history (< 2000): scattered papers in various communities. Symbolic processing mostly in digital libraries and information retrieval venues and audio processing (less explored) mostly in acoustics and DSP venues.
- The birth 2000: first International symposium on Music Information Retrieval (ISMIR) with funding from NSF Digital Libraries II initiative organized by J. Stephen Downie, Time Crawford and Don Byrd. First contact between the symbolic and the audio side.
- 2000-2006 Rapid growth of ISMIR
- 2006-2014 Slower growth and steady state
- 2015-2021 Maturation - robust industry involvement

Teaching Organization

One of the main challenges of teaching MIR is how to structure/organize the material to be taught. A good organization should be comprehensive (i.e cover most MIR tasks/topics), incremental (concepts should be introduced gradually in a logical fashion), flexible (it should be possible to emphasize or skip different topics), and balanced.
In the following slides I will discuss some possible ways to organize MIR for teaching purposes based on my extensive experience and experimentation.

Organize by stages/specificity

- Stages
  - Representation/Hearing
  - Analysis/Learning
  - Interaction/Action
- Specificity
  - Audio fingerprinting
  - Common score performance
  - Cover song detection
  - Artist identification
  - Genre classification
  - Recommendation ?

Organize based on data

University of Victoria Computer Science

Data sources:

- Audio
- Track metadata
- Score
- Lyrics
- Reviews
- Ratings
- Download patterns
- Micro-blogging

MIR Tasks

- Similarity retrieval, playlists, recommendation
- Classification and clustering
- Tag annotation
- Rhythm, melody, chords
- Music transcription and source separation
- Query by humming
- Symbolic MIR
- Segmentation, structure, alignment
- Watermarking, fingerprinting and cover song detection

Audio content analysis - A. Lerch

- Fundamentals (signals, sampling, quantization, convolution, blocking, fourier transform, correlation)
- Instantaneous features (statistical features, spectral features, post-processing, dimensionality reduction)
- Intensity and Loudness
- Tonal Analysis (pitch, monophonic pitch detection, polyphonic pitch tracking, tuning estimation, key recognition, chord detection)
- Temporal Analysis (onset detection, tempo and beat detection, downbeat, rhythm description)
- Audio Alignment (dynamic time warping, audio/score alignment)
- Music Classification (genre, similarity, mood)
- Audio Fingerprinting

Fundamentals of Music Processing - Meinard Muller

- Music Representations
- Fourier Analysis of Signals
- Music Synchronization
- Music Structure Analysis
- Chord Recognition
- Tempo and Beat Tracking
- Content-based Audio Retrieval
- Musically Informed Audio Decomposition

Kadenze online MIR program

University of Victoria Computer Science

- Extracting Information from Music Signals
- Machine Learning for MIR
- Music Information Rertieval Systems

Extracting Information from Music Signals

- Time, frequency, and sinusoids
- DFT and Time-Frequency Representations
- Monophonic Pitch Detection
- Audio Feature Extraction
- Rhythmic Analysis

Machine Learning for MIR

- Supervised Learning and Naive Bayes Classification
- Discriminative Classifiers
- Genre Classification
- Emotion Recognition and Regression
- Tags
- Music Visualization

Music Retrieval Systems

- Query Retrieval
- Polyphonic Alignment and Structure Segmentation
- Chord Detection and Cover Song Identification
- Transcription and Sound Source Separation
- Audio Fingerprinting and Watermarking

Other topics

- Optical Music Recognition (OMR)
- Symbolic Music Retrieval
- MIR for live music performance
- Computer-assisted music pedagogy
- Computational Ethnomusicology
- MIR for music production
- Natural language processing for MIR

Notebook break I

An important pedagogical advice is to use multiple alternating modes of delivery when teaching a concept. Notebooks combine text and snippets of code structured in cells that can be executed individually. Originally popularized by Mathematica, they are nowdays very popular as ways of introducing concepts. For the last couple of years I frequently utilize Python notebooks in my teaching and throughout this tutorial I will be showing some representative examples.

Organization: Discussion/Questions

Any thoughts/questions/discussion regarding the first module "Organization" ? (5-10 minutes)

Table of Contents I

Table of Contents II

Adaptation

University of Victoria Computer Science

Another challenge when teaching MIR is the diversity of the students. MIR is an interdisciplinary field and that can mean that the students can come from different disciplines and can have very different backgrounds. In this module we will go over some strategies for adapting the material for different audiences and providing supports to students with different backgrounds.

Interdisciplinary Research

Inherently inter-disciplinary and cross-disciplinary work. Connecting theme: making computers better understand music to create more effective interactions with musicians and listeners.

- Music Information Retrieval
- Digital Signal Processing
- Machine Learning
- Human-Computer Interaction
- Software Engineering
- Artifical Intelligence
- Multimedia
- Robotics
- Visualization
- Programming Languages

EDI

- **Equity** - everyone is treated fairly and equally
- **Diversity** - value differences
- **Inclusion** - everyone feels supporterd and integrated

As instructors how can we support EDI?

Some ideas for supporting EDI

- Focus and value diversity - for example go over a list of past projects that are as diverse in terms of topic as possible
- Consent form to share student material with future offerings/students
- Highlight examples of researchers from underrepresented groups
- Encourage student participation
- Model constructive and respectful feedback
- Optional educational supports, different style deliverables
- Interdisciplinary field means everyone will be outside their comfort zone
- Peer mentorship

Backgrounds

MIR students can come from many different backgrounds. Let's look at some representative examples and what they might lack:

- **Computer Science:** good programming and ML skills, not as comfortable with mathematical notation and numerical stuff, no formal music knowledge
- **Electrical and Computer Engineering:** good mathematical notation and numerical programming, maybe ML, not as experienced with general programming, no formal knowledge
- **Music:** little experience with programming, intimidated by mathematical notation, good knowledge of music, writing skills

Music 101

Some basic knowledge about music notation and theory is useful in understanding a lot of MIR research.

- Absolute and relative encoding of pitch/intervals
- Common western music relative rhythm notation
- Chords/Tonality/Key
- Very briefly scales, harmony, counterpoint, instrumentation
- Excellent tutorial ISMIR 2021: Practical Music Theory for MIR researchers

Programming 101-A

The biggest challenge with teaching programming to musicians is to convince them that it is not that hard. There are many excellent tutorials available for any programming language/environment. Some thoughts:

- Ask the students to change the tutorial examples they are encountering to be more about music. For example instead of the finding the maximum of a list of numbers change it to finding the highest note in a melody
- Music21 is a great environment to introduce Python programming to music students as they can easily see and hear the results of symbolic music manipulations
- Try to provide skeleton code with increasing amounts of student involvement

Programming 101-B

- Jupyter notebooks provide means of introducing interactivity and programming
- Some music students are familiar with visual programming environments like Max/MSP. This can be leveraged to get them going with text-based programming

Mathematics 101

Understanding of mathematics is a long process and in many cases music students or even CS students feel that it is not necessary. Some thoughts:

- Decouple notation from concepts
- Three views: concrete toy example, computer code, mathematical expression
- Read papers with emphasis on writing and notation not content
- Basic vector/matrix notation from linear algebra, probabilities/stats
- Khan Academy

Active Learning

The term **active learning** is used to describe a variety of approaches that are different than the traditional lecture mode of delivering courses. Active learning interventions include:

- Group problem-solving
- Worksheets completed in class
- Use of polls/clickers
- Studio/workshop course design
- Mock conference/peer review
- Live coding

Class Size

Class size can have a significant effect on how to deliver a course. Smaller class sizes (¡30-40 students) allow more discussion/interaction and do not requireas much preparation for assessment. Bigger class sizes (¿40 students) pose different challenges.
Ideas for large class sizes:

- Labs can function as smaller classes
- More scaffolding in learning materials
- More standarized assignments
- Organized communication in public forums with themes
- Detailed weekly workplans

Evolving over time

It is important to constantly evolve both the content and how it is delivered to reflect the rapid changes in MIR research and the surrounding software context.

- 2000 - The birth of Marsyas - self-contained code base for MIR, DSP and ML, C++ from scratch
- 2003 - Weka (and SVMs) comes to the scene for ML
- MIREX 2007 - Marsyas submissions (92 citations)
- 2010 - sklearn
- 2014 - librosa,mir_eval
- 2016 - shift to notebooks for teaching
- 2018 - more and more datasets
- 2019 - Spleeter

Notebook break II

Let's look at a notebook for explaining the basics of rhythm notation to studens without a formal music background and another one explaining matrix factorization for students without previous linear algebra skills.

Adaptation: Discussion/Questions

Any thoughts/questions/discussion regarding the second module “Adaptation” ? (5-10 minutes)

Table of Contents I

Table of Contents II

Online learning

Online learning has been around for a long time. Massive Open Online Courses rose to prominence around 2012. Despite initial claims of signalling the death of brick and mortar university teaching they have simply become another mode of learning. Covid and the resulting transition to online learning has resulted in a renewed interest in online learning. In this module we will discuss online learning strategies for effective online teaching of MIR.

MOOC critiques

In fact, the absence of serious pedagogy in MOOCs is rather striking, their essential feature being short, unsophisticated video chunks, interleaved with online quizzes, and accompanied by social networking. - Moshe Vardi
A lot of online learners seek very specific answers to particular problems they encounter rather than more general learning of concepts. For example they might look up the syntax of Python list comprehensions but do not care about the more general ideas of functional programming.

My personal favorite online learning experience

Open Studio is an online learning resource for Jazz that I discoverd and have followed regularly throughout the pandemic. It is very different than a MOOC. Some observations:

- As much as I would love to I can't find the time to properly practice
- Regular live activities - guided practice sessions
- Sense of community
- Good production quality but also informal interaction
- Feeling that everyone is part of a journey
- Regular free social media presence to attract new members

Categories of online learners

- Auditors - watch all videos, few quizzes/exams
- Completers - most videos, most quizzes/exams
- Disengaged - quickly drop the course
- Sampling - occassionally watch specific lectures

Flipped classroom

Traditionally students are introduced to new content in class and then work on assignments and projects independently at home. A flipped classroom works by introducing the content at home and practice working through the material at school. It is a type of blended learning.
Flipping a classroom requires providing video recording of lectures as well as other support material for the students to learn concepts at their own pace. Common issue is students not viewing the material. In 2018 I taught MIR completely as a flipped course with 2 hours of contact per week in which I answered questions, had discussions, and did some tutorials. The outcome was mixed: students were much more engaged and asked lots of questions but they said they would have liked some more traditional lecturing.

Recording video

University of Victoria Computer Science

Video recording is essential for online learning. Unfortunately, the standards for video quality and production are becomning higher and higher with all the content creator stuff. Some pragmatic advice:

- Video recording full lectures can be useful for review but does not work very well for online learning
- Either edit well or not at all
- Shorter (5-7 minutes) videos of concepts are better
- Change of camera - close up, zoom out
- Alternate talking, hand-writing, live-coding, slides
- Scripting and tele-prompting
- Repeated recording

Research-enriched teaching

Research-enriched teaching is a term used to describe university level teaching in which the experience of conducting research is interleaved with the teaching of concepts.

- Present papers by previous students in the course
- Invite as guest lectures previous students
- Invite industry researchers
- Present industry-led papers
- Connect to commercial products
- Use paper templates for project reports
- Conference style peer reviewing

Scaffolding

University of Victoria Computer Science

Scaffolding provides multiple supports in the form of worked examples, incremental increase of difficulty, live worked examples, and experimentation to assist with learning of concepts. It also refers to breaking a complicated concept into smaller units. At the extreme it becomes hand holding or spoon feeding.

- Show and tell - active coding
- Tap into prior knowledge
- Give time to talk
- Pre-teach vocabulary/notation
- Pause, pause, review, pause

Expanding audience - Initial experiment

In the Spring of 2014 I started video recording my MIR lectures (simply doing screencasting and Google Hangouts/YouTube) and invited external participants to the course. Overall it was a positive experience despite the low production quality with more than 300 external participants engaging with the course by viewing videos and participating in discussion. No assessment/work was offered to the external participants.

Expanding audience - Kadenze course

This experience led to the development of a full online MIR program consisting of 3 courses offered through Kadenze Inc. `https://www.kadenze.com/programs/music-information-retrieval` The videos were recorded in the Spring of 2016 but the course was finally made public in 2020.

Challenges: contract agreement with university, privacy issues, grading issues

Notebook break III

An alternative to video recordings is providing a more complete self-contained detailed notebook. The FMP (Foundations of Music Processing) notebooks are a great example of such a teaching resource. Let's look at the notebook for music syncronization as an example.

Online Learning: Discussion/Questions

Any thoughts/questions/discussion regarding the second module “Online Learning” ? (5-10 minutes)

Table of Contents I

Table of Contents II

Assessment

Assessment is a big and unavoidable challenge in teaching. Ideally, assessment should be comprehensive and properly capture the knowledge and ability of students in a particular topic. The more multi-faceted this assessment is, the more likely it is that it will correlate better with real world performance. It is important to talk to students about assessment and share how difficult it is. It is also important to differentiate between assessment and feedback although obviously they are related.

Assessment Overfitting

Standarized assessment especially face-to-face timed exams can be poor indicators of actual knowledge or performance. At the same time they are less prone to academic integrity violations and perceived as more objective and fair. The current trend in education research is to emphasize learning outcome and well defined rubrics there is a dange of putting too much effort into teaching for the assessment not for the knowledge. Motivation, curiosity, and creativity are not fosted when only material that is going to be in the exam is covered. I call this assessment overfitting in analogy to ML overfitting. It is more likely to happen with a flipped classroom style of instruction.

Academic Integrity

Academic integrity violations are increasing with the shift to online teaching.

- Contract cheating
- Public code repositories (github)
- Honest and pragmatic discussion impossible
- Build trust and respect rather than blame
- Copy but understand - self respect is critical

Degree of Difficulty (my magic bullet)

An important and fundamental challenge with assessment is calibration to student ability. If the tasks are too hard or too easy that leads to frustration. It is important to stress that each student is unique and has their own story. Degree of difficulty labeling refers to explicity labeling assessed work (exam questions, assignments, quizzes) and even reading materials in terms of difficulty.

- Basic (40%), Expected (50%), Advanced (10%)

Self-reflection

University of Victoria Computer Science

A good way to ensure personalized course work is to include some self-reflection activities. These could be in the form of a blog, journal, discussion board. Reading papers and code and writing about what was read is another good strategy. One of the challenges with self-reflection is that it takes more time to grade and does not scale easily to large classes.

Peer review

Peer review is a great way to make the students take more ownership of their work and learn to provide constructive feedback. It can be challenging to administer in large courses. Also some institutions have restrictions on using student feedback for grading.

From Model solution to peer exposure and review

The classic approach for assignments is to provide a model solution. A better approach is to critique (anonymous) student work and contrast different possible solutions. It also useful for students to contrast their work with other students especially when the assessment is subjective.

Auto-grading

Auto-grading is trivial with simple questions such as multiple choice or keyword matching quizzes. However, it becomes quite more complex and challenging as the deliverables get more complicated. A more complex setup is to have various unit tests for a piece of code (some shared with students and some withheld). My ideal situation would be to have a problem generator that personalizes assignments to students followed by more flexible similarity checking. For example a monophonic melody could be randomly generated using samples, then automatic pitch extraction could be applied, followed by comparing the results with a state-of-the art pitch detectors. The similarity will have to be calibrated but MIR techniques can be used for this purpose.

Rubrics and feedback

A well defined rubric defines how a deliverable will be graded in a detailed way. It can help students, instructors, and teaching assistants be fair and consistent. It also can be used to provide feedback to the student if it is well designed. It is extremely challenging to design good rubrics.

Gameification

Awards, prizes, badges - recycling of journals, student voting

Notebook break IV

For assignment a effective possibility is to base them on existing notebook/code resources. Let's look at a particular example for monophonic pitch detection from the Kadenze program.

Assessment: Discussion/Questions

Any thoughts/questions/discussion regarding the second module “Assessment” ? (5-10 minutes)

Table of Contents I

Table of Contents II

Group projects

The most interesting for the students and most challenging aspect of teaching a MIR course is the completion of a group project. One of my most satisfying experiences as a MIR teacher has been to see some of these projects evolve into successful ISMIR papers. In this module, we will go over some of the logistics of group projects, overview some examples, and discuss instructor strategies for making them successful.

Group formation and size

Group projects are challenging but rewarding. Optimal group size can vary. In small classes groups of 2 are feasible. Groups of 3/4 students are more common with larger classes. Bigger groups might be a necessity with really large classes but almost always create work balance tensions. Some ideas:

- Random group formation
- Student-led group formation
- Project-based group formation
- Multi-phase with reshuffling
- Pre-defined roles/responsibilities
- Trade-off between detailed specification and creative freedom

Project deliverables

- Design Specification 15%
- Progress Report 15%
- Final report 20%
- Presentation/video 10%

Project stages

At least one of these stages should be not trivial. With more software frameworks and datasets it is easier to create more sophisticated projects.

- Problem specification, data collection and ground truth annotation
- Information extraction and analysis
- Implementation and interaction
- Evaluation

Example projects

Several projects in the UVic MIR course evolved into ISMIR publications. Some examples:

- Query-by-beatboxing - 2006
- Stereo Panning - 2008
- Examining DJ ordering of playlists - 2013
- Curriculum learning for automatic chord recognition (2021)

Notebook break V

Matrix factorization notebook

Projects: Discussion/Questions

Any thoughts/questions/discussion regarding the second module “Projects” ? (5-10 minutes)

Table of Contents I

Table of Contents II

Teaching Resources

As MIR evolves as a research field, there are more and more resources such as books, overview articles, software frameworks and tools, and datasets that can be used to support teaching. In this module, we go over some of my personal favorites. This is by no means a comprehensive list of resources but rather some representative examples from what I use in my own teaching with some discussion of what I like about them.

Books

- Fundamentals of Music Processing - Meinard Muller
- Audio content analysis - Alexander Lerch
- Music Data Mining - Li, Ogihara, Tzanetakis (editors)
- Digital Signal Processing Primer - Ken Steiglitz
- Music Similarity and Retrieval - Markus Schedl

Overview papers

Good overview papers can be a great resources for MIR students. A favorite example: “Automatic chord estimation from audio: A review of the state of the art”, 2014. M McVicar, R Santos-Rodríguez, Y Ni, T De Bie

MIR Software Frameworks

- Essentia
- MIR Toolbox
- librosa
- Marsyas
- Music21

MIR Task Software

- Spleeter
- musicnn
- madmom
- melodia

MIR Tools

- mirdata
- jams
- mireval

MIR GUIs

- Sonic Visualizer/Vamp plugins
- Audacity
- Tony
- DAWs such as Ableton Live, Reaper, Plugins
- Max/MSP, PureData
- Game Engines

Other Software

- Weka
- scikit-learn
- tensorflow
- pytorch
- pandas,numpy,scipy
- Juce

Datasets

- GTZAN (historic)
- FMA (large, medium, small)
- Magnatagatune
- Music4All
- MTG-JAMENDO
- mir_data

Reproducible research

- Provide code, data, plots
- Work on multiple machines and with multiple people
- It takes a lot of effort and time but can be worth it
- Do not and do re-invent the wheel
- Be part of an ecosystem
- Open source license
- Public source repository
- Documentation, tutorials, examples
- Marsyas: 581 citations, librosa: 1251 citations, scikitlearn: 47321, essentiaL: 431

Teaching Materials

- https://musicinformationretrieval.com/
- https://ismir.net/resources/tutorials/
- https://www.audiolabs-erlangen.de/resources/MIR/FMP/C0/C0.html
- https://www.audiocontentanalysis.org/

Environments

- Jupyter notebooks
- Hosted environments: JupyterHub, Google Colab
- Github
- Overleaf

Notebook break VI

THX Logo Notebook

Resources: Discussion/Questions

Any thoughts/questions/discussion regarding the second module “Resources” ? (5-10 minutes)

Table of Contents I

Table of Contents II

Conclusions

- Teaching is a journey not just a destination
- Teaching is garden not just a journey
- Experiment and question assumptions
- Be honest and pragmatic with students
- Focus on intuition and motivation
- Teaching is more about pacing and filtering and less about content delivery