

# CMPE 58P - Lecture 1.

## Machine Listening

Introduction, Course Structure, Applications of audio and music  
signal processing, Audio Representations



Department of Computer Engineering,  
Boğaziçi University, Istanbul, Turkey

Instructor: A. Taylan Cemgil

# What is machine listening ??

- ▶ Audio and Music processing
  - ▶ Musical audio and music modelling
  - ▶ Environmental sound
  - ▶ Source separation
  - ▶ "Like machine vision but as applied to audio"
  - ▶ "Things which are less covered in a speech processing course"
- ▶ Active research topic
  - = partially understood,
  - = mostly we don't know what we are talking about
  - = don't expect well formulated problems and established solution strategies
  - ... (well, then it wouldn't be called a research topic)

# Layers

- ▶ Audio Processing Practicalities:
  - ▶ Restoration, Enhancement, Coding, Synthesis
- ▶ Extracting structure and information from sound
  - ▶ Music
  - ▶ Natural sounds
- ▶ Ultimate Goal:
  - ▶ Developing computational methods for simulating on a computer the perceptual mechanisms underlying human hearing
  - ▶ Designing computer systems that display a human level performance in sound understanding

# Applications

- ▶ Hearing Aids
- ▶ Monitoring and Surveillance
  - detection and recognition of environmental sounds,  
distributed sensing
- ▶ Music Information Retrieval
  - content-based querying and retrieval, automatic  
classification and transcription
- ▶ Interactive Music Performance systems
- ▶ Music Education, Musicology, Music Perception and  
Cognition Research

# Approaches

- ▶ Computational modelling of sound and human listening
- ▶ Blend Data driven, statistical methods with prior knowledge
- ▶ Multi-disciplinary
  - ▶ Computer Science, Acoustics, Signal Processing, Machine Learning, Statistics, Musicology, Cognitive Science, Neuroscience

# Goals of this Course

- ▶ Familiarity with aspects of audio and music signals
  - ▶ Extraction, Manipulation, Recognition via signal processing
  - ▶ Statistical Machine Learning
  - ▶ Practical applications, create a test-bed for computational methodologies
- ▶ Review of the relevant literature, understand state of the art
- ▶ Many paradigms, personal bias towards a **model based approach**,
  - ... develop models and associated inference algorithms rather than describing algorithms
- ▶ There will be some maths
  - Signal processing
  - Statistical Machine Learning, Pattern Recognition
  - Calculus and Linear Algebra,

# Topics

- ▶ Audio and Music Representations
- ▶ Transforms, Spectrogram representations
- ▶ Basic sound synthesis techniques
- ▶ Hidden Markov models and Linear Dynamical systems
- ▶ Rhythmic structure: Tempo, Rhythm, Time Signature
- ▶ Frequency and Pitch detection
- ▶ Audio/Score Matching,
- ▶ Audio source separation multichannel/single channel
- ▶ Polyphonic music transcription
- ▶ Music information retrieval, similarity search, classification

# Course Structure

- ▶ Web page

<http://www.cmpe.boun.edu.tr/courses/cmpe58P/spring2009/>

- ▶ Required Work

- ▶ Biweekly assignments (Mostly Programming, analytic derivations)
- ▶ Presentations (every week one or two persons, about 30 - 50 mins)
  - ▶ Reading group style
  - ▶ Choose a topic, read a few papers, explain it
  - ▶ Graded also by fellow students
- ▶ Final Project: Implementation, Report and Presentation

- ▶ Grading

- ▶ Relative weights
  - ▶ % 25 Literature Presentation(s)
  - ▶ % 35 Assignments
  - ▶ % 40 Final Project Report and Presentation

# Study materials

- ▶ Papers, Tutorials
  - ▶ Links will be made available on the web page
- ▶ Slides, Handouts
- ▶ Some Reference Books all available via BU library:
- ▶ Klapuri, A. and Davy, M. (Editors), "Signal Processing Methods for Music Transcription," Springer-Verlag, New York, 2006 <http://www.cs.tut.fi/~klap/amt.html>
- ▶ Ben Gold and Nelson Morgan, Wiley 2000 (ISBN: 0-471-35154-7) Speech and Audio Signal Processing: Processing and perception of speech and music
- ▶ Perry R. Cook (Editor), Music, cognition, and computerized sound [electronic resource] : an introduction to psychoacoustics.

# Sources of extra material

## ▶ Conference Proceedings

- ▶ IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),
- ▶ IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA),
- ▶ International Conference Music Information Retrieval (ISMIR),
  - ▶ also check: Music Information Retrieval Evaluation eXchange (MIREX)
- ▶ International Computer Music Conference (ICMC)
- ▶ Neural Information Processing Conference (NIPS)
- ▶ Uncertainty in Artificial Intelligence (UAI)

## ▶ Journals

- ▶ IEEE Transactions on Speech, Audio and Language Processing
- ▶ Journal of New Music Research
- ▶ Computer Music Journal

# Audio and Music Specific

- ▶ Chris Raphael's papers and Automatic Accompaniment project

[http://xavier.informatics.indiana.edu/~craphael/music\\_plus\\_one/index.html](http://xavier.informatics.indiana.edu/~craphael/music_plus_one/index.html)

<http://www.musicplusplus.net/examples.html>

- ▶ Dan Ellis' Lab in Columbia (LabRosa) research pages contain a wealth of material

- ▶ <http://labrosa.ee.columbia.edu/>
- ▶ Matlab Code
- ▶ Tutorials, slides

- ▶ Julius Smith's homepage in Stanford, CCRMA

<http://ccrma.stanford.edu/~jos/>

- ▶ Centre for Digital Music, Queen Mary, University of London

<http://www.elec.qmul.ac.uk/digitalmusic/index.html>

- ▶ Paris Smaragdis' demos and papers

[http://web.media.mit.edu/~paris/Paris\\_Smaragdis/Paris\\_Smaragdis\\_Audio\\_Demos.html](http://web.media.mit.edu/~paris/Paris_Smaragdis/Paris_Smaragdis_Audio_Demos.html)

# Remarks

- ▶ Feel free to propose a topic related to audio and music processing, even it is not on our topic list
- ▶ If you have **not** chosen research topic but roughly have something in mind or simply don't know yet
  - ▶ Come and talk to me to clarify a topic/technique
  - ▶ May be a good experience for choosing a research topic for your thesis
- ▶ Ideally, a good report could be presented with some extensions at a national or international conference
  - ▶ Many open problems, so perhaps **you** will find a new approach. (More likely than you think)

# Remarks

- ▶ Any programming language or other system for computation and visualisation
  - ▶ Matlab (preferred)
  - ▶ Octave
  - ▶ Java,
  - ▶ C/C++, BLAS, ATLAS, GNU Scientific Library