Machine Learning for Signal Processing

Project Ideas

Class 5. 15 Sep 2016

Instructor: Bhiksha Raj

Course Projects

- Covers 30% of your grade
- 10-12 weeks of work
- Required:
  - Serious commitment to project
  - Extra points for working demonstration
  - Project Report
  - Poster presented in poster session
    - 8 Dec 2016
  - Graded by anonymous external reviewers in addition to the course instructors

Course Projects

- Projects will be done by teams of students
  - Ideal team size: 3
  - Find yourself a team
  - If you wish to work alone, that is OK
    - But we will not require less of you for this
  - If you cannot find a team by yourselves, you will be assigned to a team
  - Teams will be listed on the website
  - All currently registered students will be put in a team eventually
- Will require background reading and literature survey
  - Learn about the problem

Projects

- Teams must inform us of their choice of project by 30th September 2016
  - The later you start, the less time you will have to work on the project

Quality of projects

- Project must include aspects of signal analysis and machine learning
  - Prediction, classification or compression of signals
  - Using machine learning techniques
- Several projects from previous years have led to publications
  - Conference and journal papers
  - Best paper awards
  - Doctoral and Masters' dissertations

Projects from past years: 2015

- So you think you can sing? : Fixing Karaoke
- Self-paced learning in multimedia event detection with social signal processing
- Improving intonation in audio book speech synthesis
- Your keyboard is not your friend: reading typed text from audio recordings
- Learning successful strategy in adversarial games
- Gesture phase segmentation
- Electric load prediction for airport buildings
- Unsupervised template learning for birdsong identification
- Realtime keyword spotting in video games

Projects from past years: 2015

- Loop querier – searching the rhythmic pattern
- Vision-based montecarlo localization for autonomous vehicle
- Beatbox to drum conversion
- City localization on flikr videos using only audio
- Facial landmarks based video frontalization and its application in face recognition
- Audioshop: Modifying and editing singing voice
- Predicting and classifying RF signal strength in an environment with obstacles
- Realtime detection of basketball players

Projects from past years: 2014

- IMPROVING SPATIALIZATION ON HEADPHONES FOR STEREO MUSIC
- PREDICTING THE OUTCOME OF ROULETTE
- FACIAL REPLACEMENT IN VIDEOS
- ISOLATED SIGN WORD RECOGNITION SYSTEM
- ACCENTED ENGLISH DIALECT CLASSIFICATION
- BRAIN IMAGE CLASSIFIER
- FACIAL EXPRESSION RECOGNITION
- MOOD BASED CLASSIFICATION OF SONGS TO IDENTIFY ACOUSTIC FEATURES THAT ALLEVIATE DEPRESSION
- PERSON IDENTIFICATION THROUGH FOOTSTEP-INDUCED FLOOR VIBRATION
- DETECT HUMAN HEAD-ORIENTATION BASED ON CONVOLUTIONAL NEURAL NETWORK AND DEPTH CAMERA
- NEURAL NETWORK BASED SLUDGE VOLUME INDEX PREDICTION

Projects from past years: 2014

- 8-BIT MUSIC NOTE IDENTIFICATION - TURNING MARIO INTO METAL
- STREET VIEW HOUSE NUMBER RECOGNITION BASED ON CONVOLUTIONAL NEURAL NETWORKS
- TRAIN-BASED INFRASTRUCTURE MONITORING
- MANIFOLD INTERPOLATION OF X-RAY RADIOGRAPHS
- A SMARTPHONE BASED INDOOR POSITIONING SYSTEM AUGMENTED WITH INFRARED SENSING
- ROCK, PAPER, SCISSORS -- HAND GESTURE RECOGNITION
- LANGUAGE MODELS WITH SEMANTIC CONSTRAINTS
- LEARNING TO PREDICT WHERE A DRIVER LOOKS
- REAL TIME MONITORING OF STUDENT'S LEARNING PERFORMANCE

Projects from past years: 2013

- Automotive vision localization
- Lyric recognition
- Imaging without a camera
- Handwriting recognition with a Kinect
- Gender classification of frontal facial images
- Deep neural networks for speech recognition
- Predicting mortality in the ICU
- Human action tagging
- Art Genre classification
- Soccer tracking
- Image manipulation using patch transforms
- Audio classification
- Foreground detection using adaptive mixture models

Projects from previous years: 2012

- Skin surface input interfaces
  - Chris Harrison
- Visual feedback for needle steering system
- Clothing recognition and search
- Time of flight countertop
  - Chris Harrison
- Non-intrusive load monitoring using an EMF sensor
  - Mario Berges
- Blind sidewalk detection
- Detecting abnormal ECG rhythms
- Shot boundary detection (in video)
- Stacked autoencoders for audio reconstruction
  - Rita Singh
- Change detection using SVD for ultrasonic pipe monitoring
- Detecting Bonobo vocalizations
  - Alan Black
- Kinect gesture recognition for musical control

Projects from previous years: 2011

- Spoken word detection using seam carving on spectrograms
  - Rita Singh
- Bioinformatics pipeline for biomarker discovery from oxidative lipidomics of radiation damage
- Automatic annotation and evaluation of solfege
- Left ventricular segmentation in MR images using a conditional random field
- Non-intrusive load monitoring
  - Mario Berges
- Velocity detection of speeding automobiles from analysis of audio recordings
- Speech and music separation using probabilistic latent component analysis and constant-Q transforms

Project Complexity

- Depends on what you want to do
- Complexity of the project will be considered in grading.
- Projects typically vary from cutting-edge research to reimplementation of existing techniques. Both are fine.

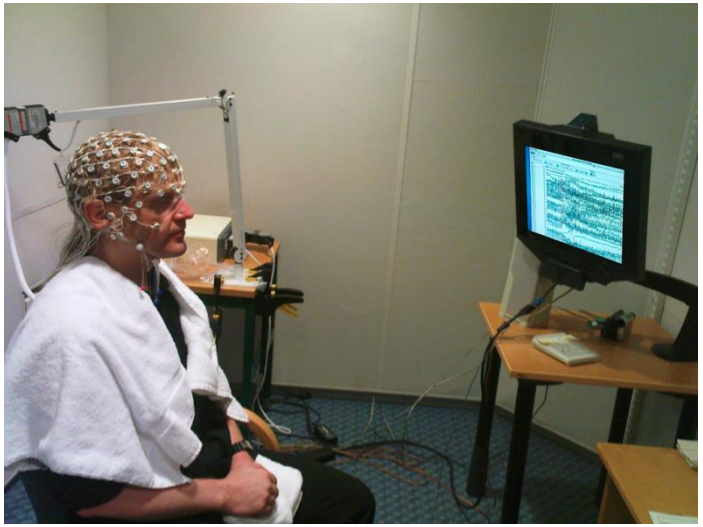
Incomplete Projects

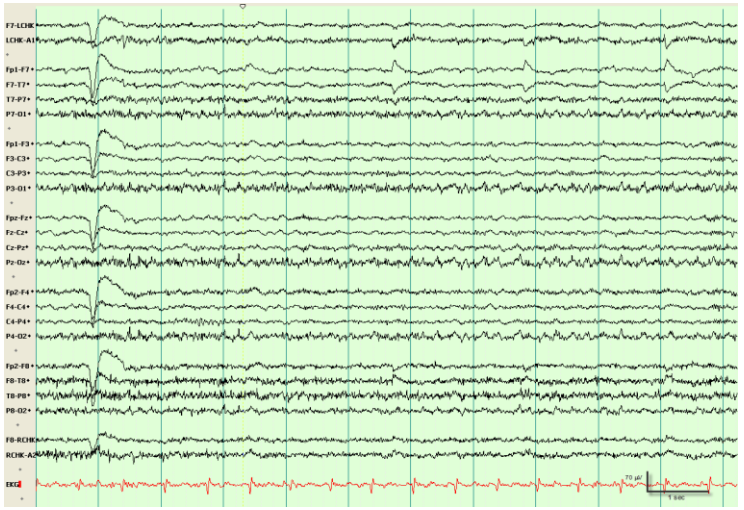
- Be realistic about your goals.
- Incomplete projects can still get a good grade if
  - You can demonstrate that you made progress
  - You can clearly show why the project is infeasible to complete in one semester
- Remember: You will be graded by peers

“Local” Projects..

- Several project ideas routinely proposed by various faculty/industry partners
  - Sarnoff labs, NASA, Mitsubishi, Adobe..
- Local faculty
  - Alan Black is usually good for a project or two
  - LP Morency has fantastic ideas on analysis of multimodal recordings of H-H (and H-C) communication
  - Roger Dannenberg is a world leader in computational music
  - Mario Berges has helped in the past
  - Fernando de la Torre
  - Rita Singh does nice work on speech forensics
  - Others…
- Johns Hopkins: We have several data sources in Hopkins
  - Students may team up with partners from JHU

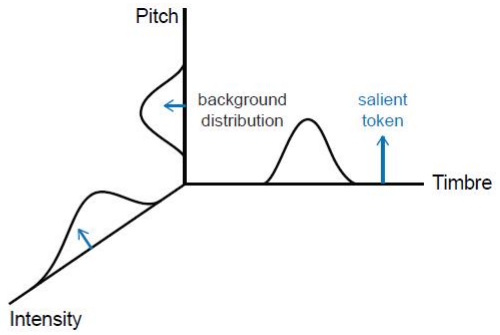
1. Reading the Brain (Hopkins)





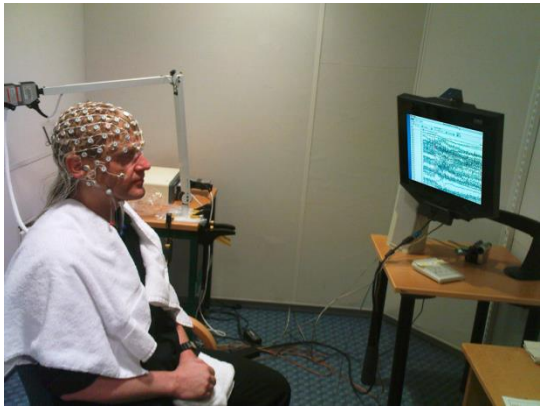
- We have a collection of EEG responses to specific sound stimuli.
- Multiple recordings for each person
  - Mulitple sessions for each stimulus
- Detect stimuli from recordings
  - Mounya Elhilali

Reading the Brain



- Subject watches silent movie while listening to musical notes while paying attention to movie
  - Notes deviate from norm
  - How does the brain respond to deviations
- Also
  - Denoising body signals
  - Denoising electrode connectivity issues
- http://journal.frontiersin.org/article/10.3389/fnhum.2014.00327/full

More brain



- EEG data where the person is listening to two sounds
  - left and right ears listen to two different sounds
- Determine which part of the brain deals with each ear.

2. Hitler Circa 1934

Closing Address To The Nazi Party Congress Nuremberg, Germany, September 14, 1934 Adolf Hitler

- A historical moment that changed the world

What is in the human voice?

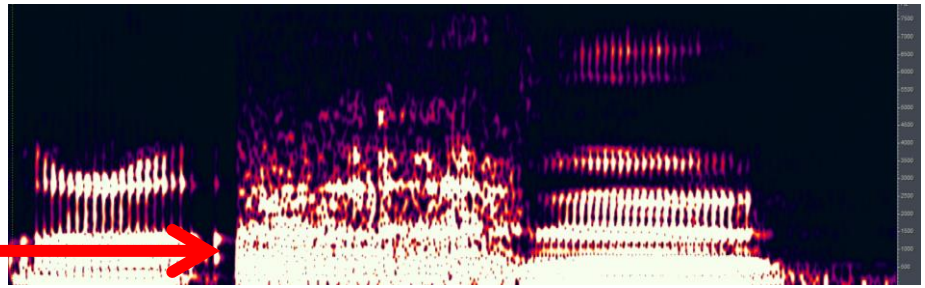
Closing Address To The Nazi Party Congress Nuremberg, Germany, September 14, 1934 Adolf Hitler

- A historical moment that changed the world
- But there's something here that may have prevented it..

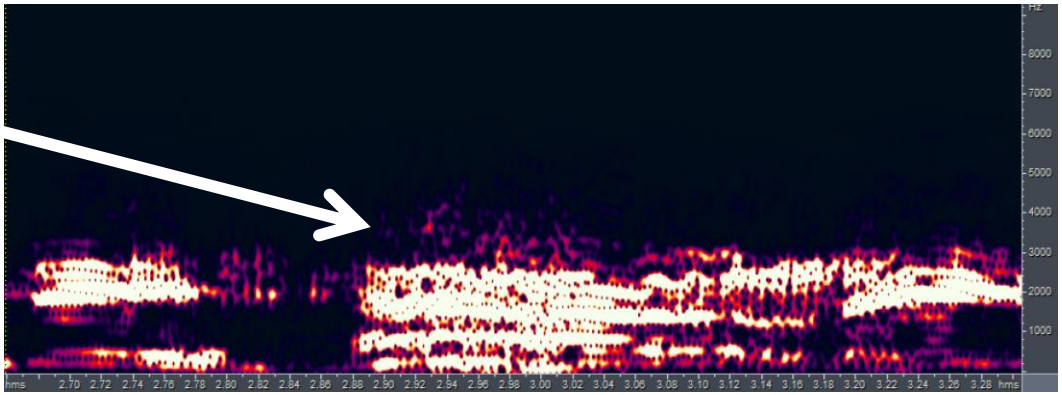
Parkinsons!!



Michael J Fox



Hitler



- Hitler’s voice
- 'Video evidence depicts that Hitler exhibited progressive motor function deterioration from 1933 to 1945.'

Available Data

Colombian (PC-GITA)	German	Czech
50 PD, 50 HC	88 PD, 88 HC	20 PD, 15 HC
Sound-proof booth	--	--
Age ~ 61	Age ~ 64	Age ~ 60
Speech tasks: Vowels, pa-ta-ka, words, sentences, read text, monologue		

- **Dedicated tests → We know what was said (good for automatic analysis but not for unobtrusive monitoring)**
- **Monologues, e.g. *What did you do yesterday?* (close to unobtrusive monitoring)**

PD Speech: Characteristics

- Reduced loudness
- Monotonic speech
- Breathy voice
- Imprecise articulation
- Accelerated or slowed
- Stutter-like

Hypokinetic dysarthric Speech

Colombian patient
Female, Age: 75
UPDRS-III: 52

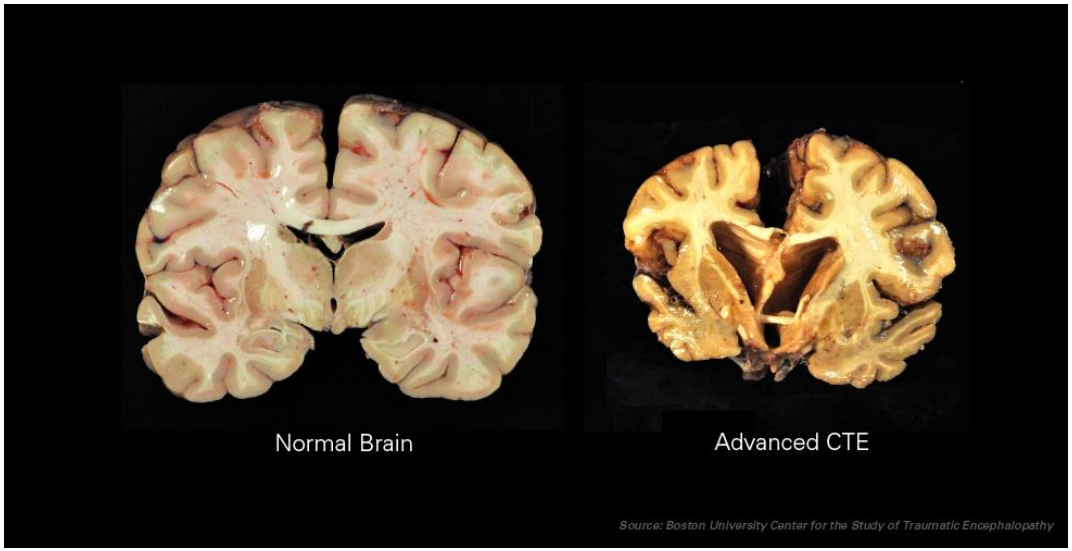
Additional Data

Dataset	Description
Multimodal	Speech, gait, and hand-writing of 30 PD
Longitudinal	Speech of 26 PD recorded in different sessions across 4 years
Genetics	Speech of 3 groups of speakers: 6 PD with the mutation 7 with the mutation but not diagnosed PD 6 non-PD, non-mutation, but relatives
At-home	Speech, gait, and handwriting of 7 PD in 4 all day sessions

Challenge

- Detect Parkinsons from voice
- Bonus – analyze historical figures
  - The Hiter result needs to be published
- Supervisor: Rita Singh

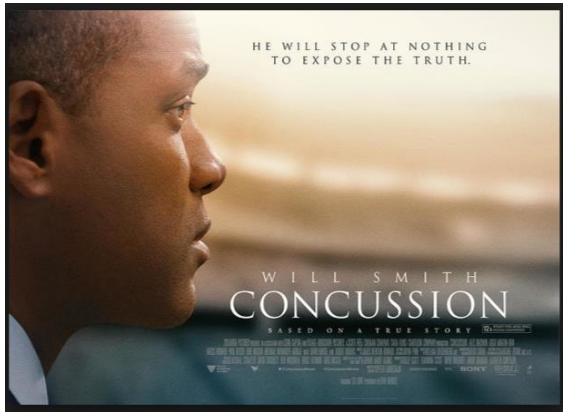
3. Chronic Traumatic Encephelopathy



- Chronic Traumatic Encephelopathy is a progressive neurological disorder that affects the brains of individuals who have experienced repeated blows to the head

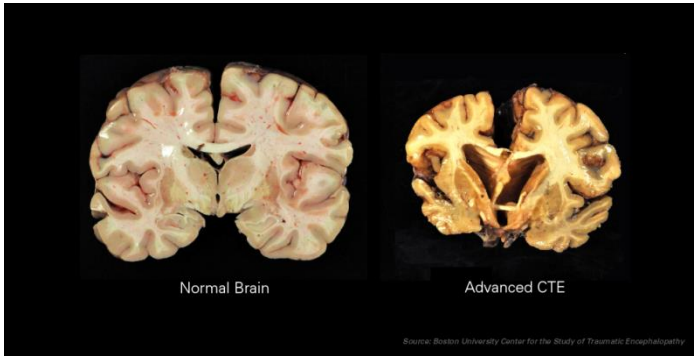
- Increasing evidence that CTE affects athletes of all ages, who are involved in any contact sport
  - American football, boxing, ice hockey, rugby, soccer, professional wrestling..

CTE in the news



- Currently one of the most prominent sports-related health problems
  - Gained prominence recently from the death of several well-known athletes
  - Although known for a long time as the “punch-drunk” syndrome

The CTE Problem



- Problem: CTE can only be confirmed through dissection of patients' brains post-mortem
- No clinical recordings at all
  - Since we don't know if the patients have CTE
- On the other hand:
  - Several famous personalities were found to have CTE
  - Many recordings of them on YouTube etc.

Detecting CTE

- Hypothesis: CTE is exhibited through behavior
  - Speech, gaze, gesture, gait
- Proposal: Gather data of famous CTE patients from public sources
- Attempt to develop diagnostic from behavioral charactistics

- If you succeed, they may make a movie about you

Potential Projects from Alan Black

- 4. Find F0 in story telling
  - F0 is easy to find in isolated sentences
  - What about full paragraphs
  - Storytellers use much wider range
- 5. Find F0 shapes/accent types
  - Use HMM to recognize “types” of accents
  - (trajectory modeling)
  - Following “tilt” and Moeller model

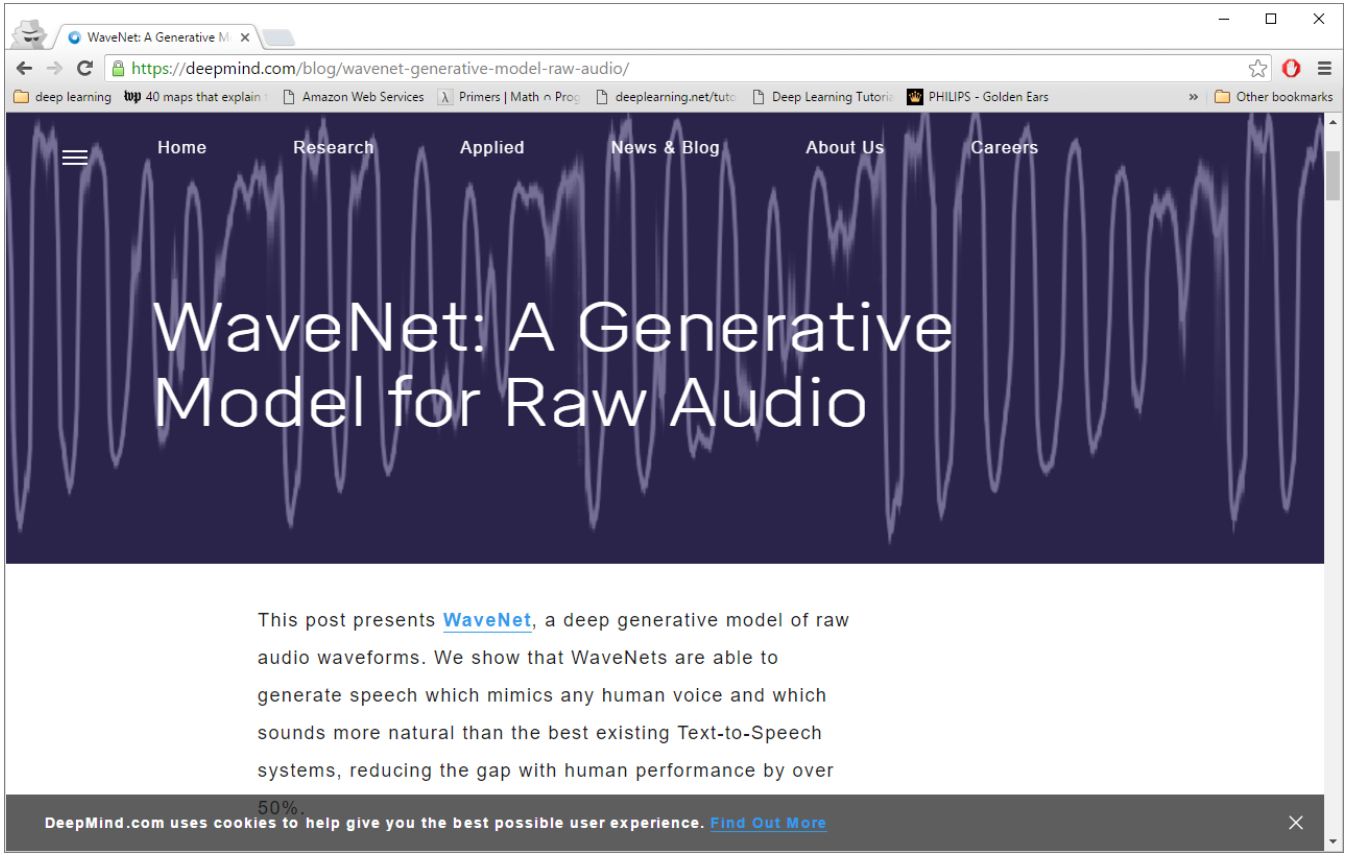
Parametric Synthesis

- 6. Better parametric representation of speech
  - Particularly excitation parameterization
- 7. Better Acoustic measures of quality
  - Use Blizzard answers to build/check objective measure
- 8. Statistical Klatt Parametric synthesis
  - Using “knowledge-base” parameters
  - F0, aspiration, nasality, formants
  - Automatically derive Klatt parameters for db
  - Use them for statistical parametric synthesis

Speech without Text

- 9. Speech processing without written form
  - Derive symbolic form from speech (done-ish)
  - Discover “words”/”syllables”
  - Derive speech translation models
- 10. Build a cross linguistic synthesizer
  - Hindi text in, but speaks in Konkani
- 11. Audio only in target language
  - Speech to speech translation
  - Dialog System

12. Wavenet



- Latest from deepmind
- The biggest advance in speech synthesis this millennium

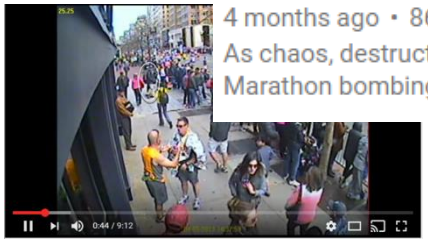
Wavenet challenge

- Uses RNNs
- Duplicate Wavenet
  - With fewer resources
- Other DNN/RNN formalisms
- Advisors: A. Black, B. Raj

13. Largescale audio retrieval

- 

A challenge



About 455,000 results

Boston Bombing Day 1 | The Stunning Stop the Killers Made After the Attack

ABC News

4 months ago • 86,708 views

As chaos, destruction erupts, authorities jumpstart investigation into Boston Marathon bombings. Day 2: http://bit.ly/1XKMH2J [This ...

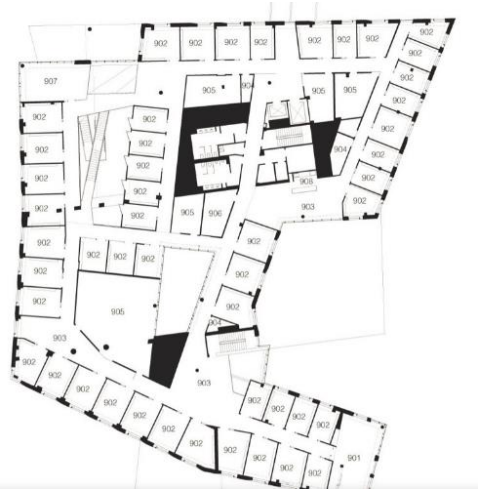
- Hundreds of people recorded the event and uploaded to YouTube
  - Each of the recordings is recording the *exact same* sounds
- FBI has one recording
- They want to recover all the other recordings
  - To compose an entire timeline and get evidence

Audio Fingerprinting Challenge

- Given huge collection of multimedia recordings
- Given a snippet of a recording of an event

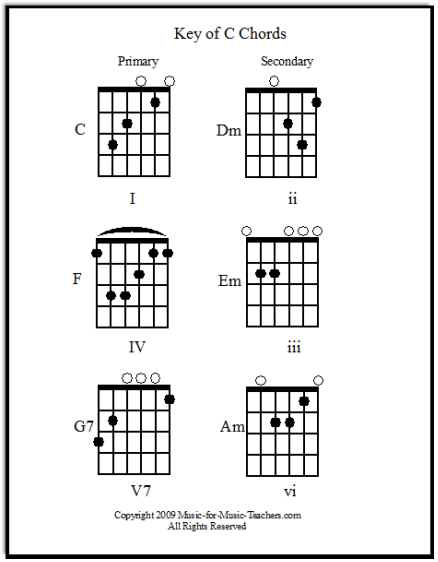
- Recover *all* other recordings of *exactly the same* event
  - Not *similar* events
  - Recordings may have been taken from different perspectives, different locations etc.
  - Video may not match at all
  - Matching video does not indicate identical event
  - Evidence in audio

14. Layout Mapping



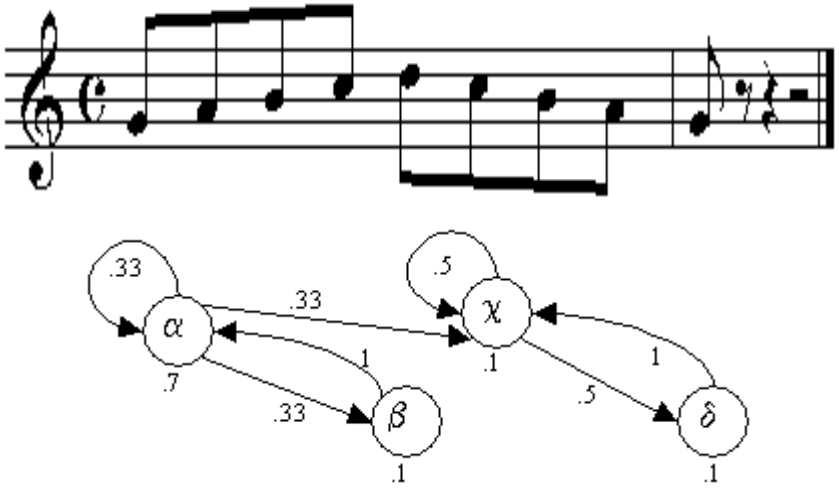
- You walk around these spaces all day, yet you are lost!
- Your phone walks with you.
- Use sensor (accelerometer, other sensors) readings to build up a *layout* of the space and label it

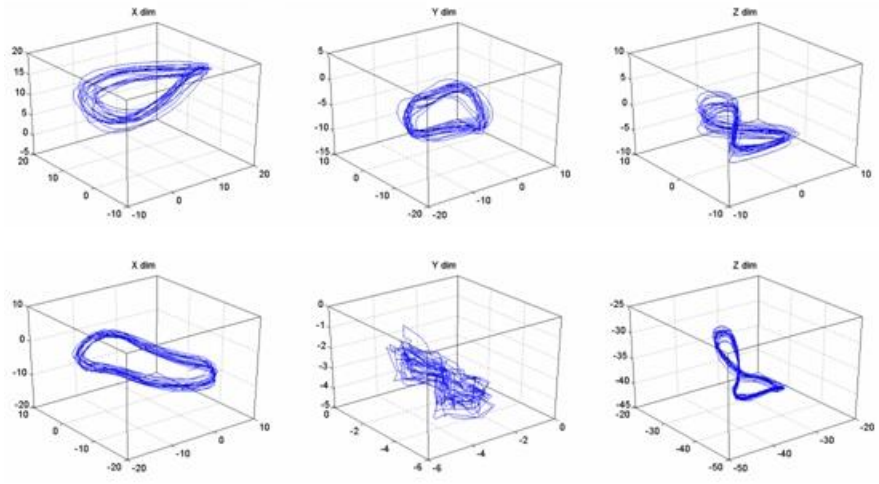
Music Ideas: Roger Dannenberg
15: Finding Chords



- Build a classifier to find all C-major chords in music recordings. Build a collage from the discovered sounds.

16. Computational Creativity

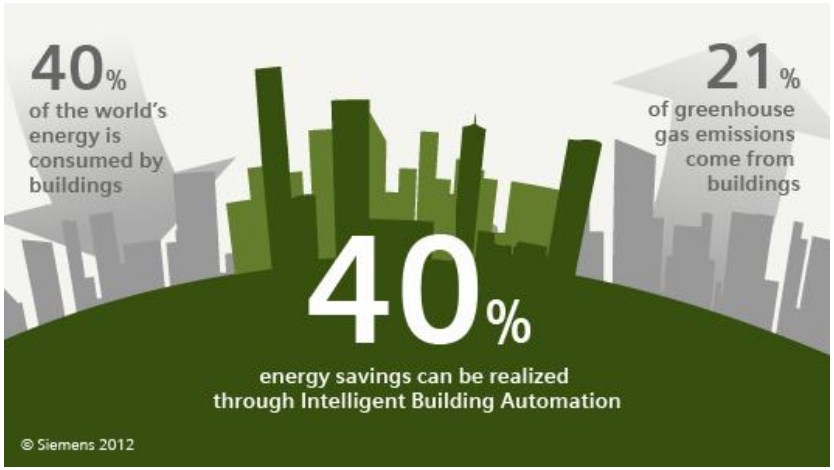


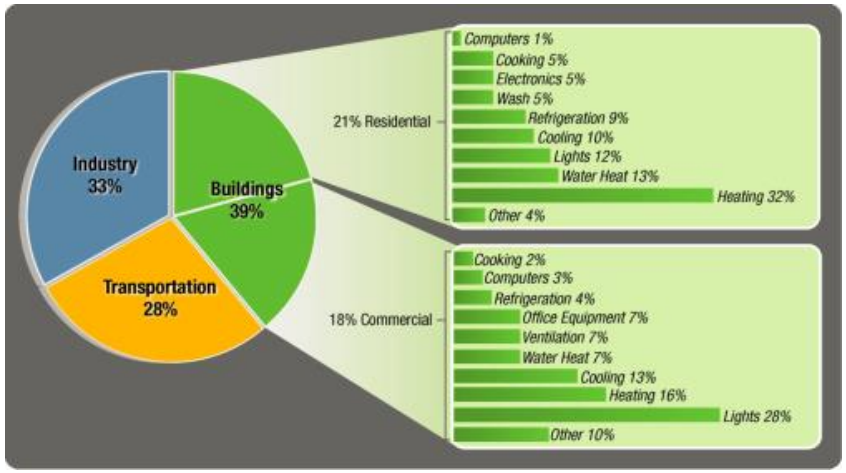


- "Create" music from existing pieces
  - Model ensembles of music through graphical models. *Generate* new music from the snippets
  - Model music trajectories as low-dimensional trajectories in embedding space.

Ideas from Mario Berges

17. Energy disaggregation





- Energy disaggregation as a binary matrix factorization problem, approximated via deep nets (http://nilmworkshop.org/2016/slides/HenningLange.pdf)
- Based only on trajectory of current / power levels, disaggregate consumption of individual devices

18. Anomaly detection

- Anomaly detection on whole-building energy consumption data for campus buildings
  - Data sets available

- Determine anomalous events in energy consumption
  - Can be hard to find
  - Could have serious consequences

19. Room Occupancy Traces

- Analysis of per-room occupancy traces (# of people in every room, every second) for an office building throughout 6 months.

- Important to optimize energy consumption

20. Classifying sensor type

- Classifying sensor type from just raw measurement time-series (i.e., is this the time series of temperature measurements, or is it humidity?).
  - See, for example:https://dl.acm.org/citation.cfm?doid=2821650.2821670

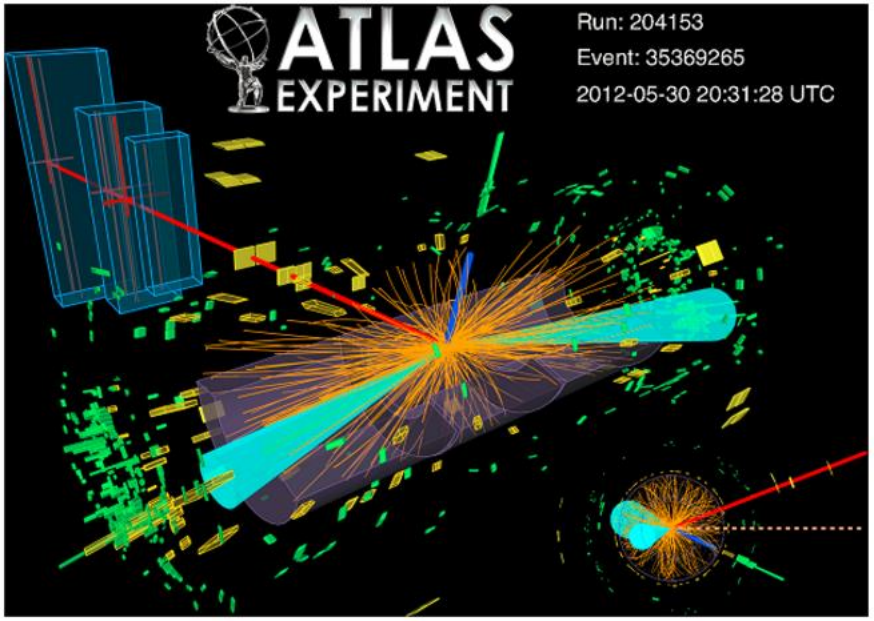
Najim Dehak

21. *Signature verification*

Is this really X?

Najim Dehak

Use the ATLAS experiment to identify the Higgs boson



22. *Higgs Boson Machine Learning Challenge*
    https://www.kaggle.com/c/higgs-boson/data
    - *Machine learning to find properties of the boson*

Najim Dehak

23. DNNs based on LDA or NCA

24. *Discriminative training for generative models.*

25. *PLDA, Gaussian classifier for face recognition and speaker verification.*

You get the idea

- You may pick any of these problems or come up with a fun one of your own
- They *must* exercise your MLSP skills

- Please form teams and inform me and TAs of teams asap
  - Or we will assign you to a team
- Please send us project proposals before 25th
  - Try to break down the steps in solving your problem in your proposal
  - Needed to evaluate feasibility