

# Audio Lifelogging

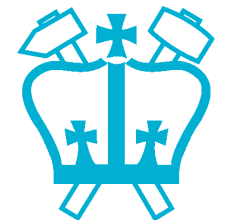
Dan Ellis

Laboratory for Recognition and Organization of Speech and Audio  
Dept. Electrical Eng., Columbia Univ., NY USA

[dpwe@ee.columbia.edu](mailto:dpwe@ee.columbia.edu)

<http://labrosa.ee.columbia.edu/>

1. Audio Lifelogging
2. Speech
3. Environmental Sound Analysis
4. Medical Applications



COLUMBIA UNIVERSITY  
IN THE CITY OF NEW YORK

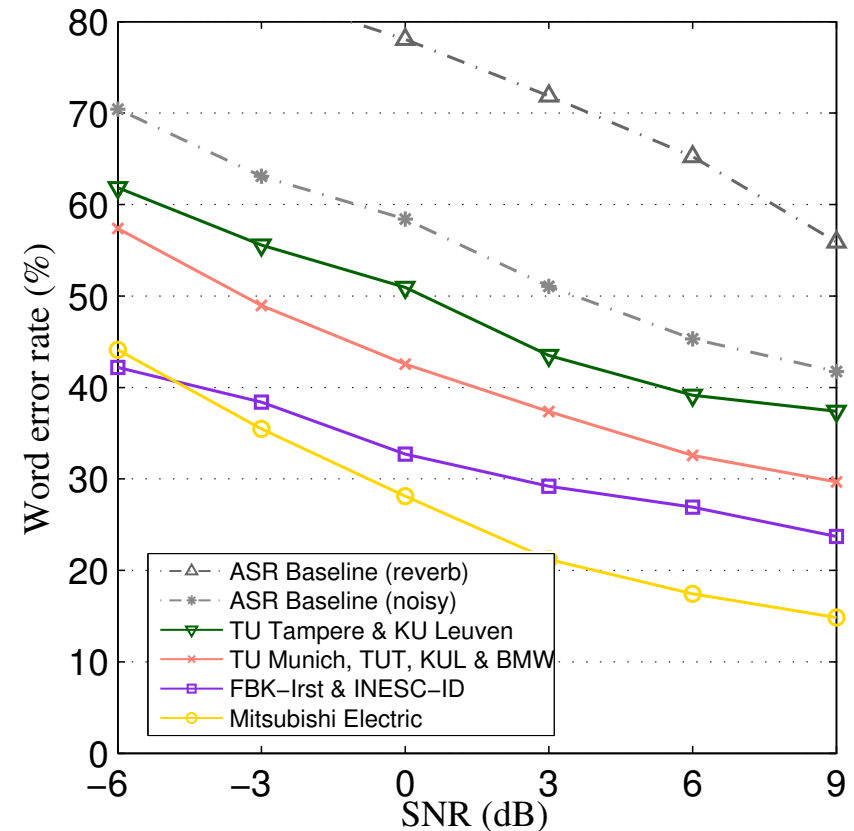
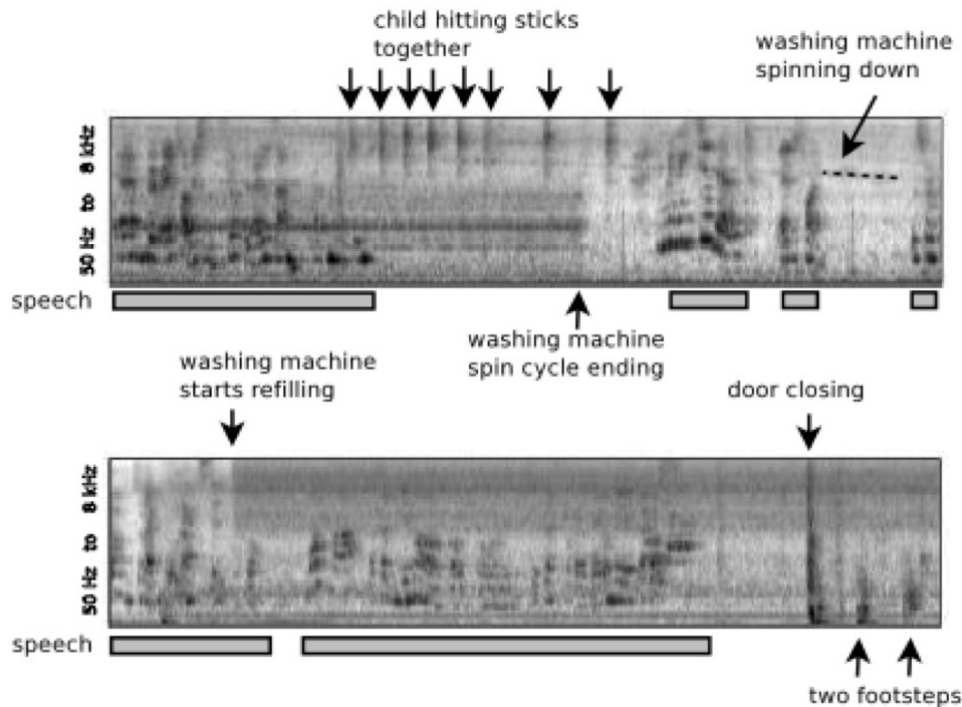
# Personal Audio Lifelogs

- Easy to record **everything** you hear
  - ~250GB / year @ 64 kbps
- Very hard to **find anything**
  - how to scan?
  - how to visualize?
  - how to index?



# Speech Recognition

- Transcripts of discussions would be useful...
  - distant-mic / noisy ASR still quite limited
  - CHiME (Comp. Hearing in Multisource Env.s):



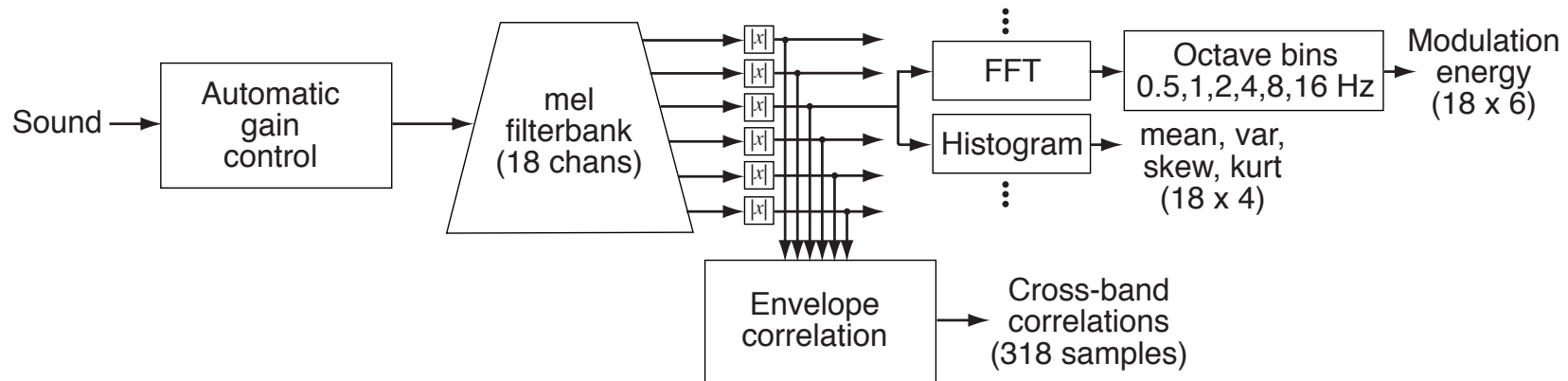
Barker et al. '13

[http://spandh.dcs.shef.ac.uk/chime\\_workshop/slides/CHiME13\\_overview.pdf](http://spandh.dcs.shef.ac.uk/chime_workshop/slides/CHiME13_overview.pdf)

# Environmental Sound Classification

Ellis, Zheng, McDermott '11

- Classify soundtracks with “**texture**” features



- Mixed results...

The screenshot shows a web browser window with the address bar displaying 'file:///u/drspeech/data/aladdin/code/videoSndtrkClass/html/Dog-max.html'. The page title is 'Dog-max'. Below the title, there are four video thumbnails, each with a title and a score:

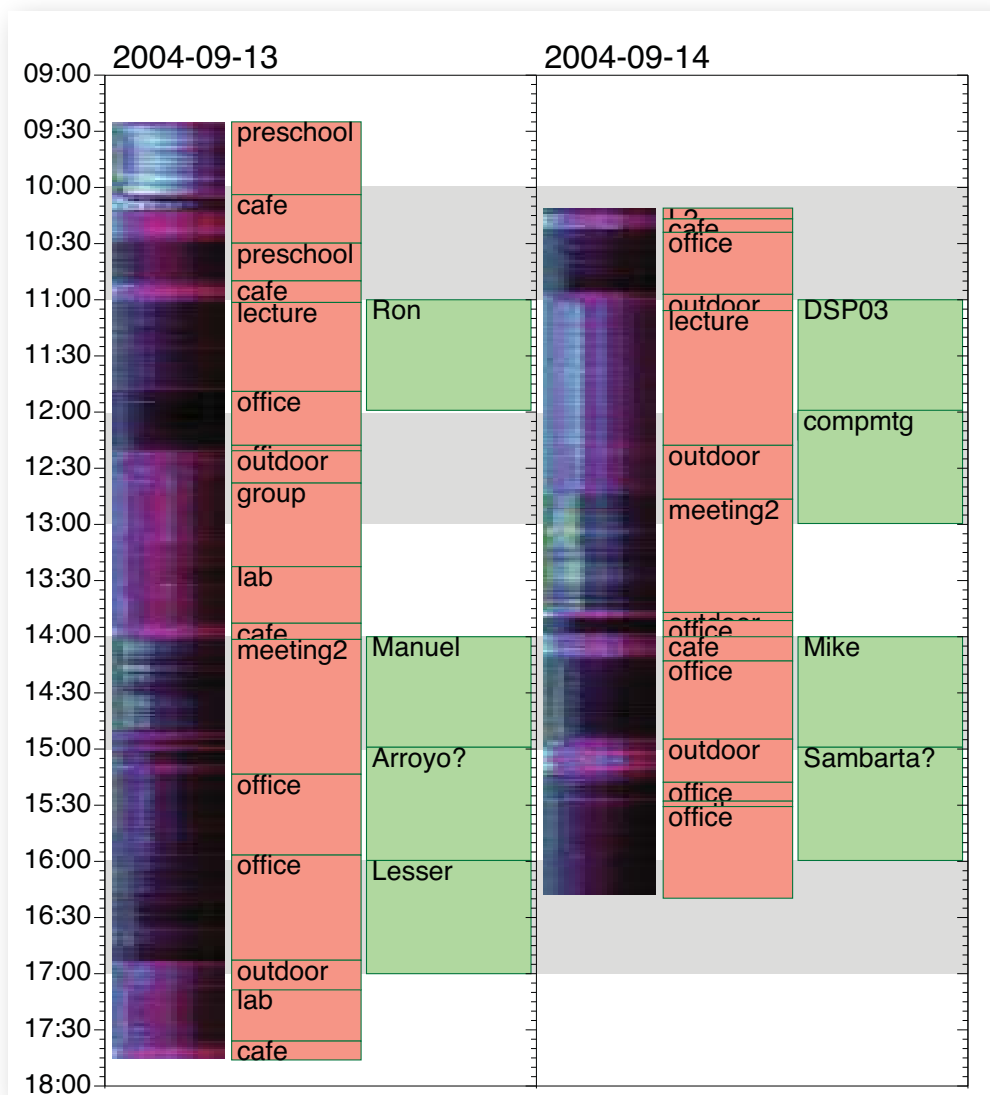
- HVC862001 - 0.33779
- HVC229331 - 0.29265
- HVC386054 - 0.24994
- HVC205402 - 0.18894

The first thumbnail shows a person lying on the ground. The second shows a person holding a white kitten. The third shows a dog running in a field. The fourth shows a close-up of a brown dog's face.

# Segmentation & Clustering

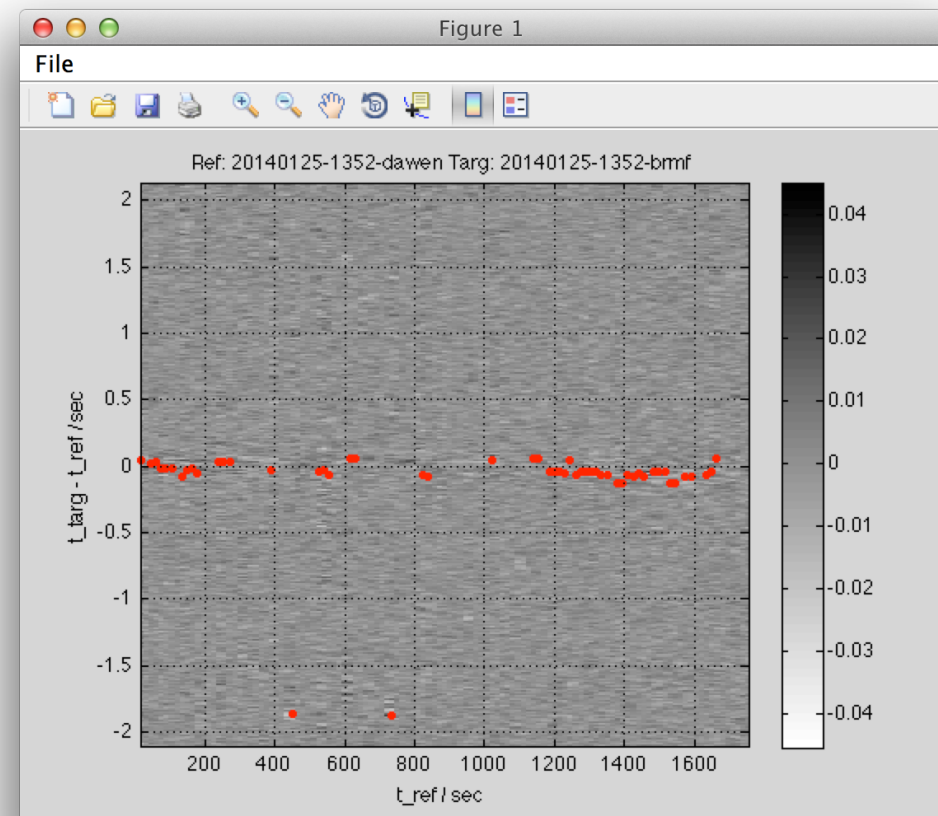
Lee & Ellis '04

- Features from 1 min windows
- Segmentation by local statistics
- Cluster across whole set
- Manual labels of each class



# Medical Applications

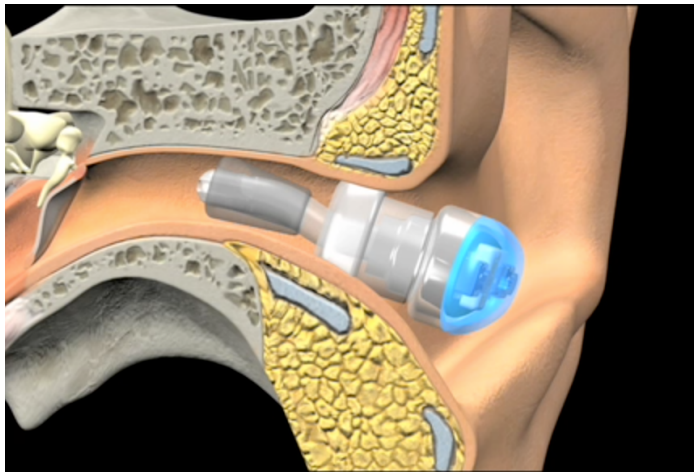
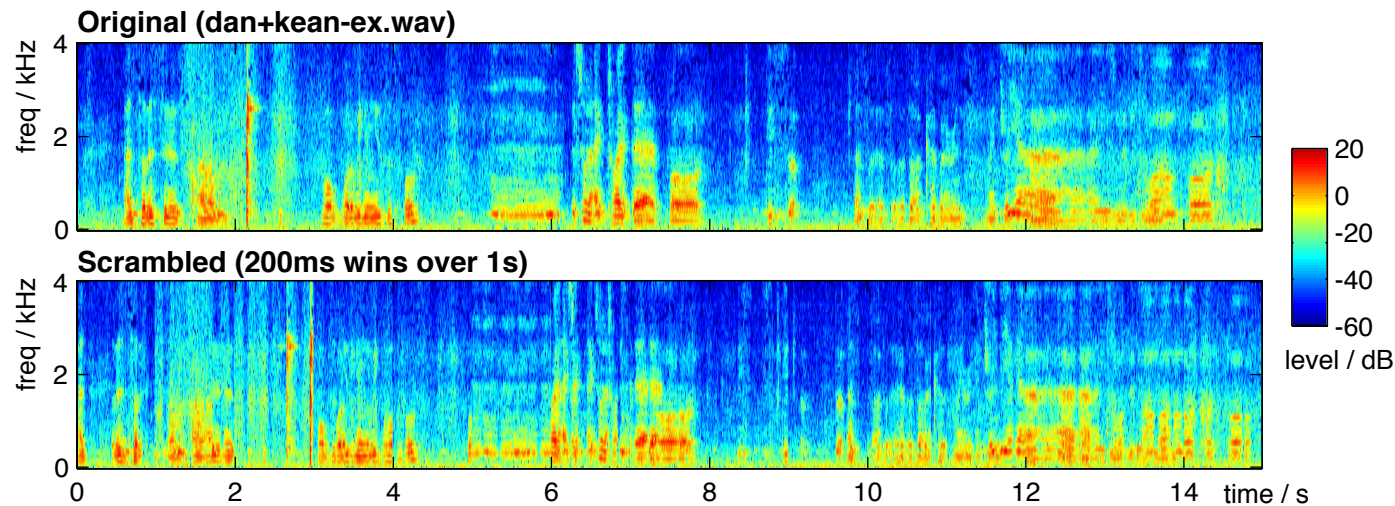
- Tracking behavior & symptoms
  - coughing
  - sleep disturbances
- Tracking hospital interactions
  - correlate recordings



# Privacy

- Privacy-preserving features

- scramble audio over 200ms windows



Hearium

- Self-audio only
- augmented-reality earphones/mics

# Summary

- **Personal Audio Lifelogs**  
Too good to waste!
- **Speech & Nonspeech Content**  
Both useful in different ways
- **Applications**  
Personal information, behavior measurement