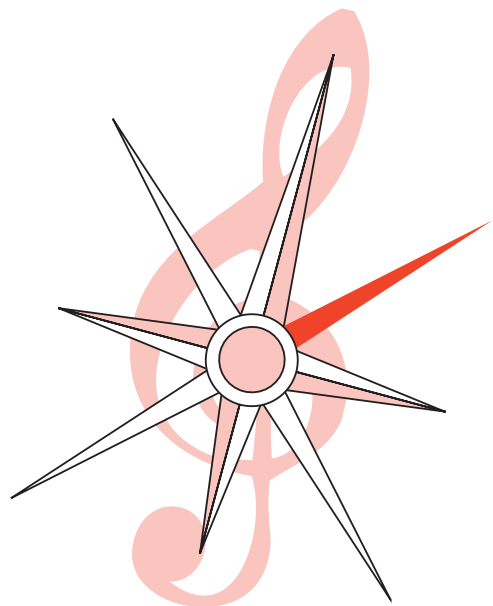
Welcome!

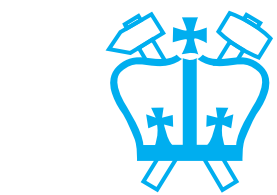
* Adobe Labs
* Amie Street
* Carnegie Mellon
* The College of NJ
* Columbia
* Connecticut College
* Cooper Union
* Dartmouth



North
East
Music
Information
Special
Interest
Group

* Drexel
* The Echo Nest
* Harvard / USC
* McGill
* NYU
* Princeton
* QTrax
* Sun Microsystems
* U de Montréal

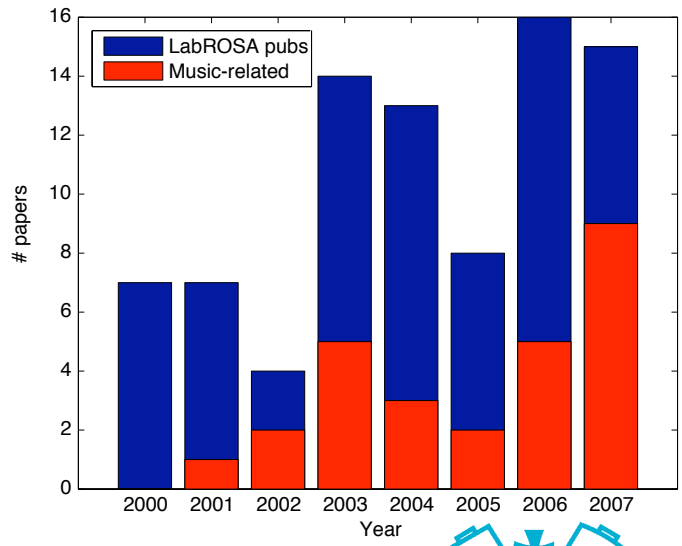
... because talking is good.



Current Music Research at LabROSA

- **The Big Picture:**
  - Lots of data
    + noisy transcription
    + weak clustering
    ⇒ musical insights?

- **History & Support**
  - 2007 first year when majority of LabROSA papers were music-related
  - Support:
    - Columbia Academic Quality Fund
    - Departmental & NSF fellowships
    - NSF grant (3 years from Sep 2007)



Transcription as Classification

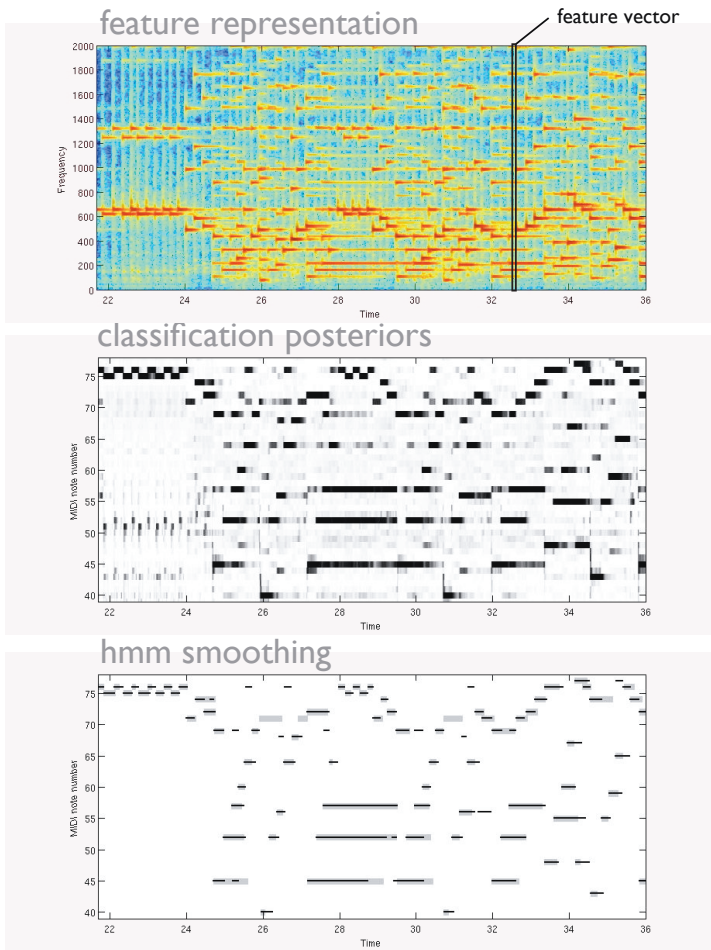
Graham Poliner

- Exchange signal models for data
  - transcription as pure classification problem:

Training data and features:
- MIDI, multi-track recordings, playback piano, & resampled audio (less than 28 mins of train audio).
- Normalized magnitude STFT.

Classification:
- N-binary SVMs (one for ea. note).
- Independent frame-level classification on 10 ms grid.
- Dist. to class bndy as posterior.

Temporal Smoothing:
- Two state (on/off) independent HMM for ea. note. Parameters learned from training data.
- Find Viterbi sequence for ea. note.

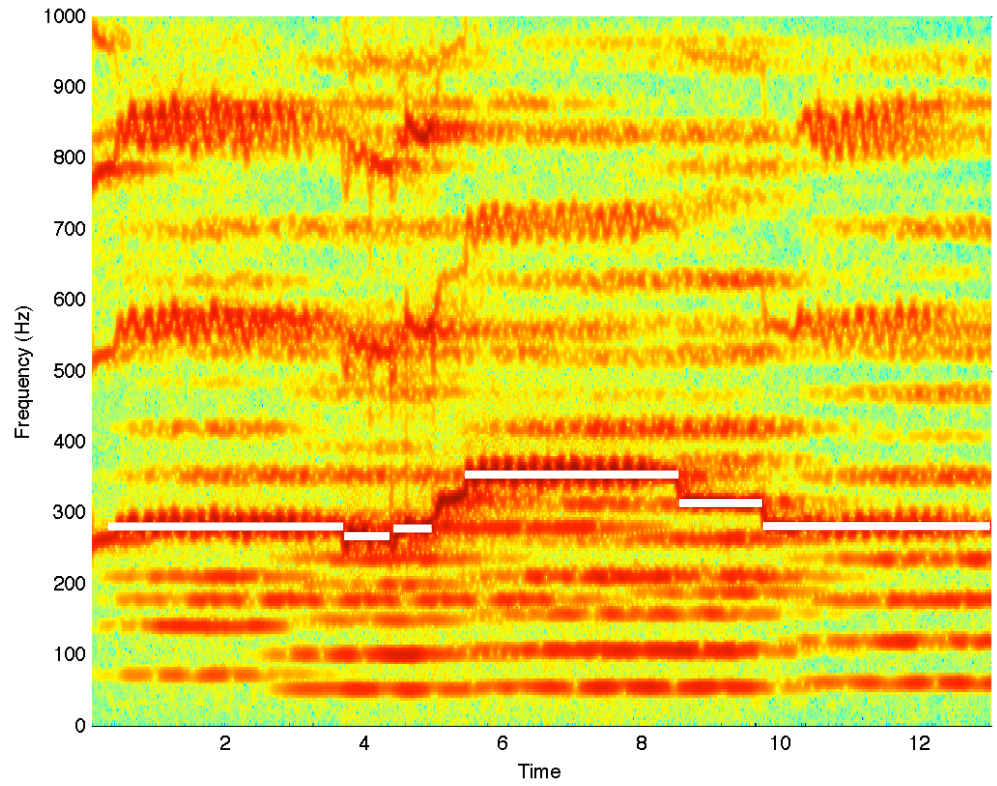


Singing Voice Modeling & Alignment

Christine Smit
Johanna Devaney

- How are phonemes sung?
  - e.g. "vowel modification" in classical voice

- Collect the data
  - .. by identifying solos
  - .. by aligning libretto to recordings
  - e.g. align Karaoke MIDI files to original recordings

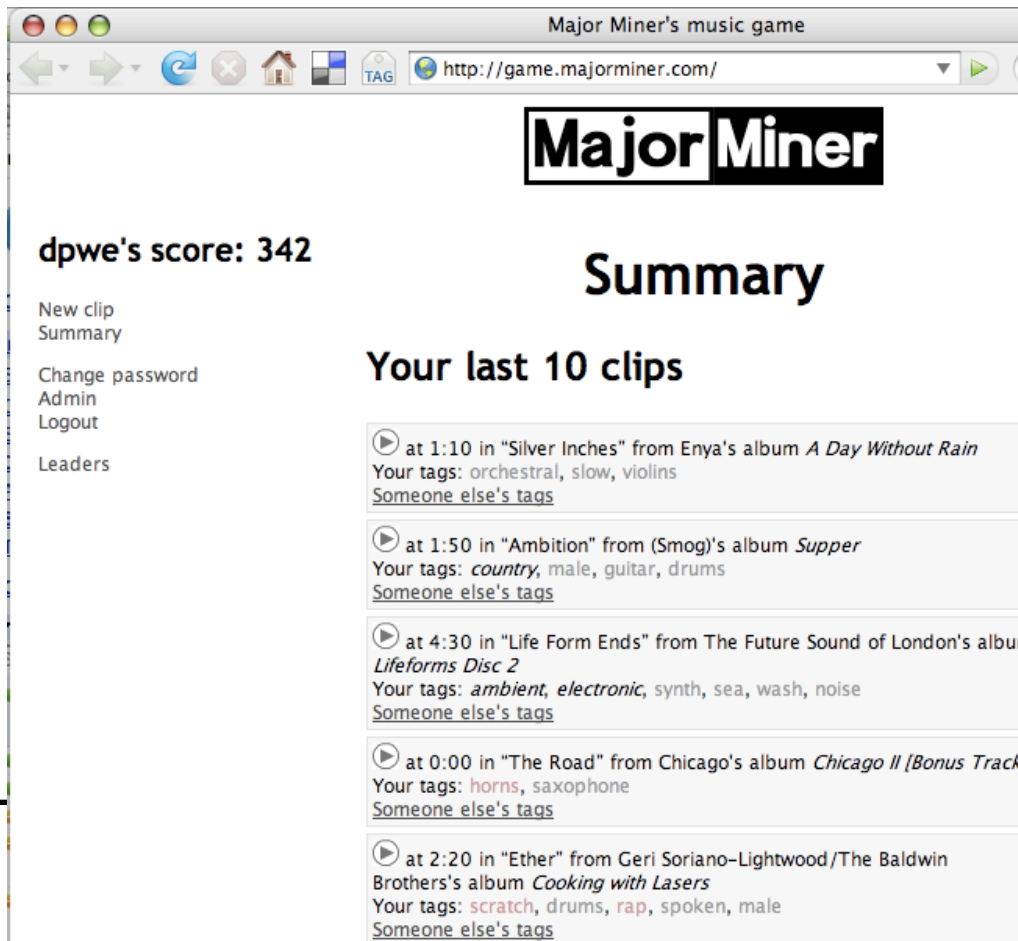


- Lyric Transcription?

MajorMiner: Semantic Tags

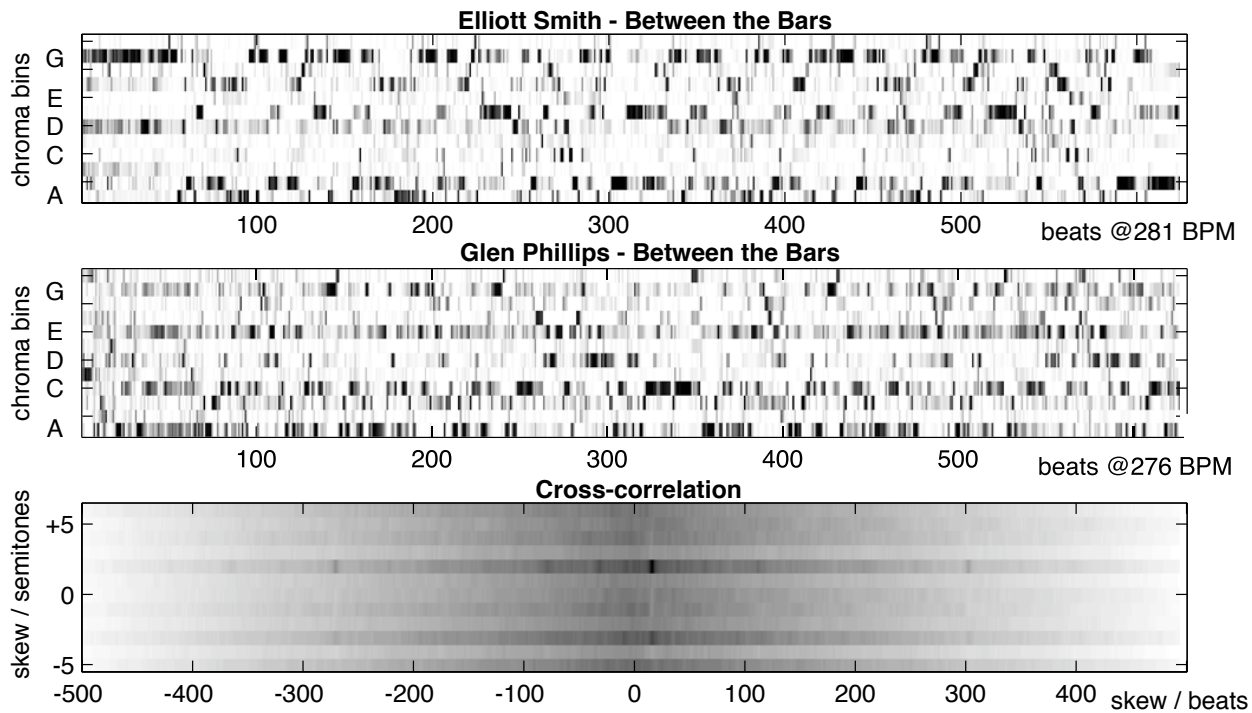
Mike Mandel

- Describe segment in human-relevant terms
  - e.g. anchor space, but more so
- Need ground truth...
  - what words to people use?
- MajorMiner game:
  - 400 users
  - 7500 unique tags
  - 70,000 taggings
  - 2200 10-sec clips used
- Train classifiers...



Cover Song Matching: Correlation

- Cross-correlate *entire song* beat-chroma matrices
  - ... at all possible transpositions
  - implicit combination of match quality and duration

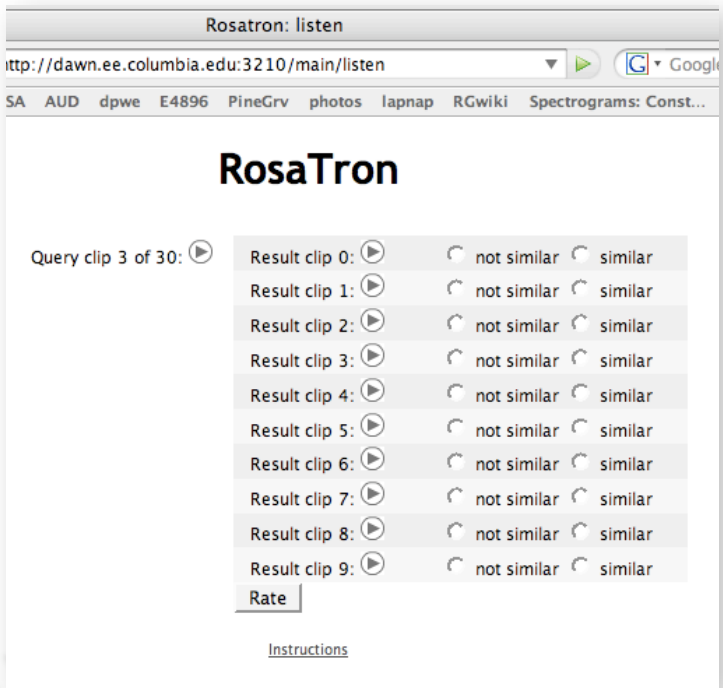


- One good matching fragment is sufficient...?

Cross-Correlation Similarity

Courtenay Cotton
Mike Mandel

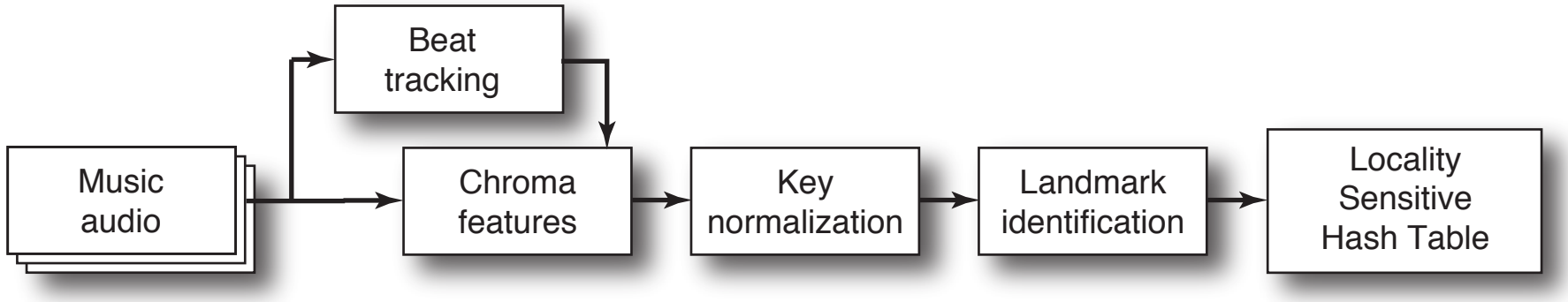
- Use correlation to find similarity?
  - e.g. similar note/instrumentation sequence
  - may sound very similar to judges
- Evaluate by subjective tests
  - modeled after MIREX similarity



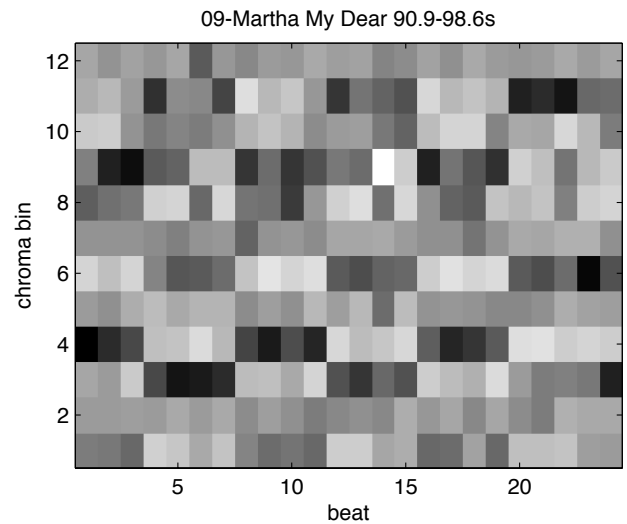
Algorithm	Similar count
(1) Xcorr, chroma	48/180 = 27%
(2) Xcorr, MFCC	48/180 = 27%
(3) Xcorr, combo	55/180 = 31%
(4) Xcorr, combo + tempo	34/180 = 19%
(5) Xcorr, combo at boundary	49/180 = 27%
(6) Baseline, MFCC	81/180 = 45%
(7) Baseline, rhythmic	49/180 = 27%
(8) Baseline, combo	**88/180 = 49%**
Random choice 1	22/180 = 12%
Random choice 2	28/180 = 16%

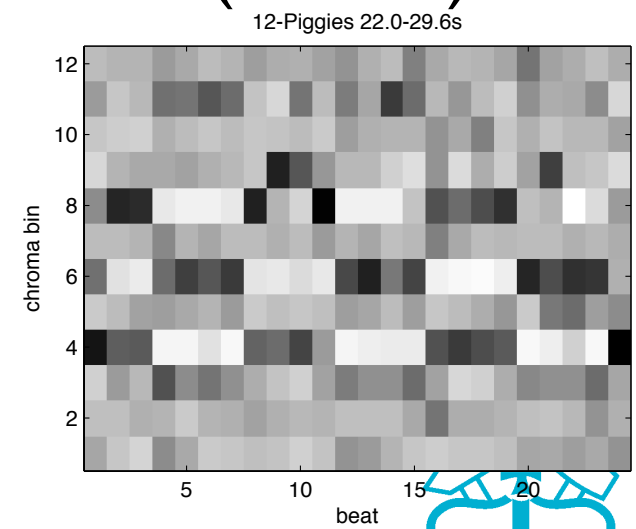
Beat Chroma Fragment Clustering

- Idea: Build a dictionary of harmonic/melodic fragments by clustering a large corpus



- 86 Beatles tracks ⇒ 41,705 patches (12x24)
  - LSH takes ~300 sec
  - High-pass along time
  - Song filter



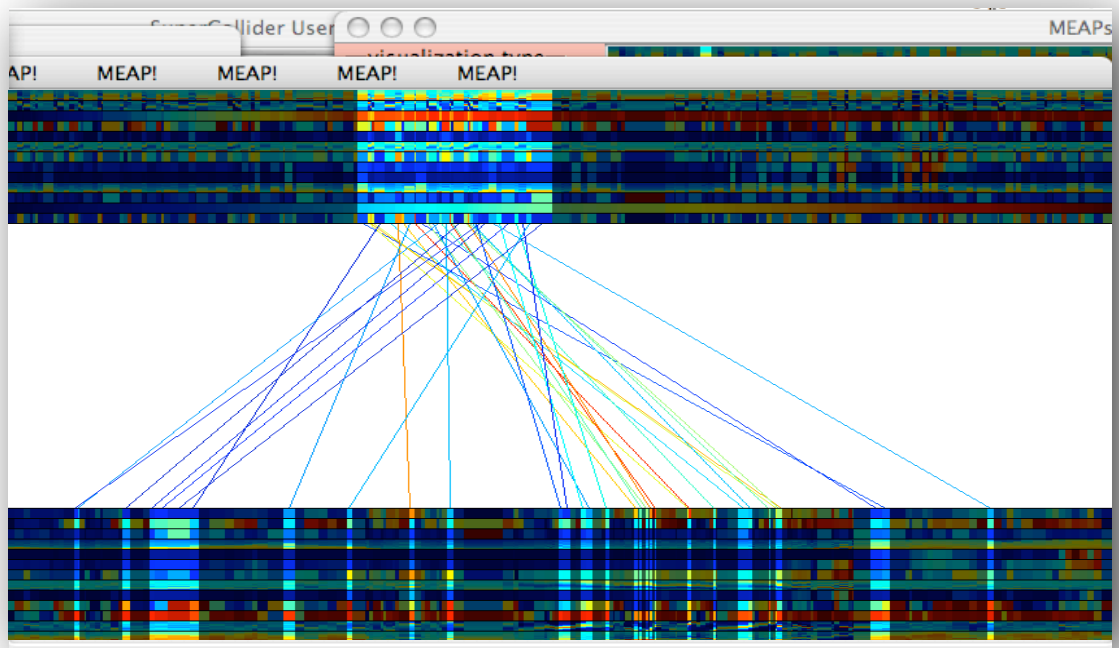


MEAPsoft

- Music Engineering Art Projects
  - collaboration between EE and Computer Music Center

with Douglas Repetto, Ron Weiss, and the rest of the MEAP team

- MEAPsoft combines music IR analysis with wacky resequencing algorithms
  - also some neat visualizations...



Summary

- What is made possible by so much data?