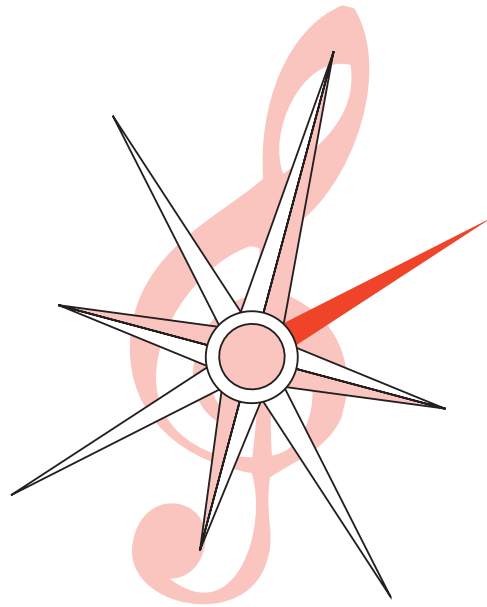


Welcome!



*North
East
Music
Information
Special
Interest
Group*

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

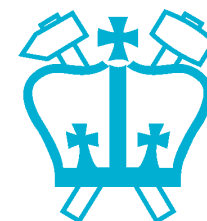
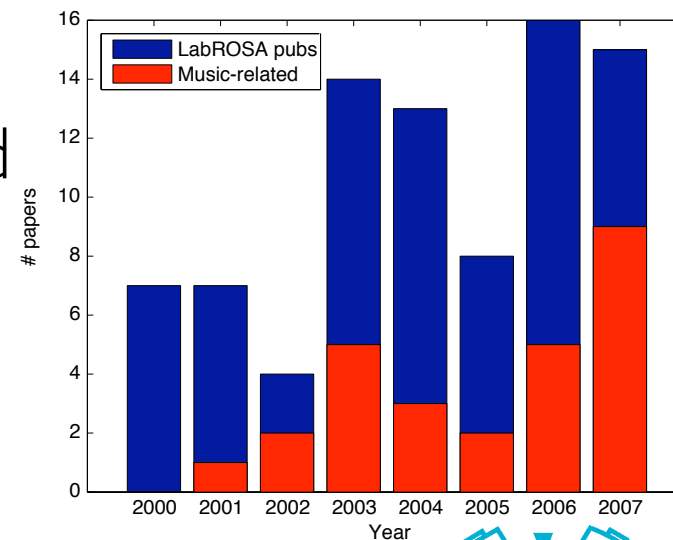
- * Drexel
- * The Echo Nest
- * Harvard / USC
- * McGill
- * NYU
- * Princeton
- * QTrax
- * Sun Microsystems
- * U de Montréal

... 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.



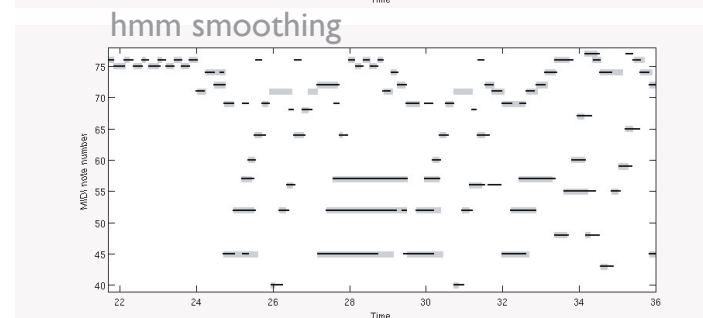
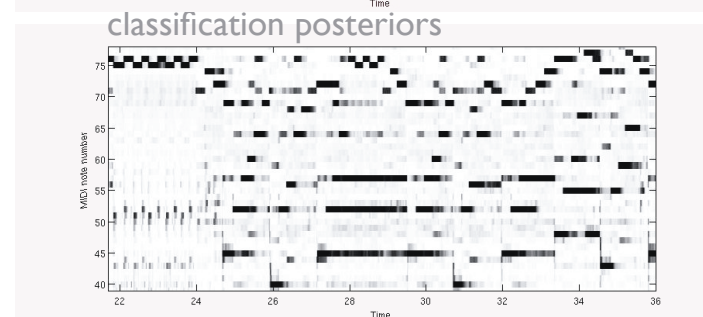
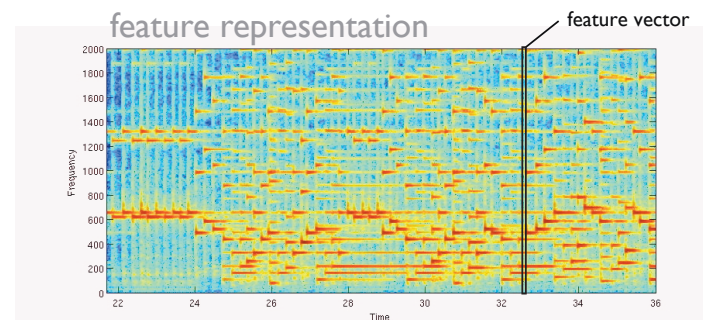
Classification:

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



Temporal Smoothing:

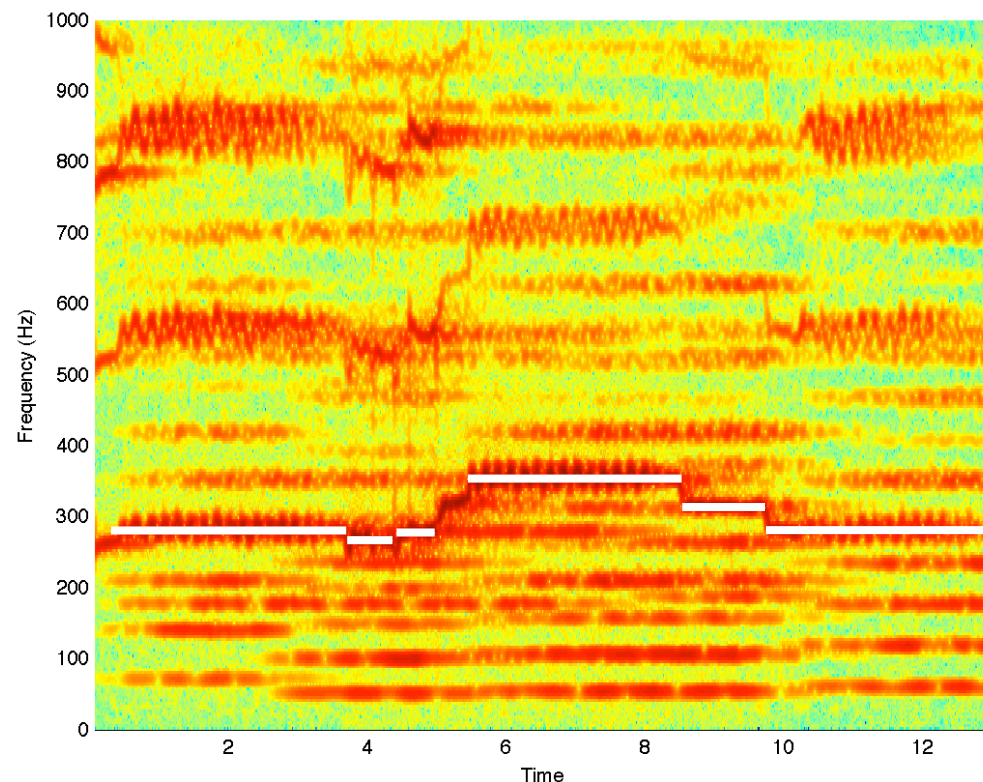
- Two state (on/off) independent HMM for ea. note. Parameters learned from training data.
- Find Viterbi sequence for ea. note.



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

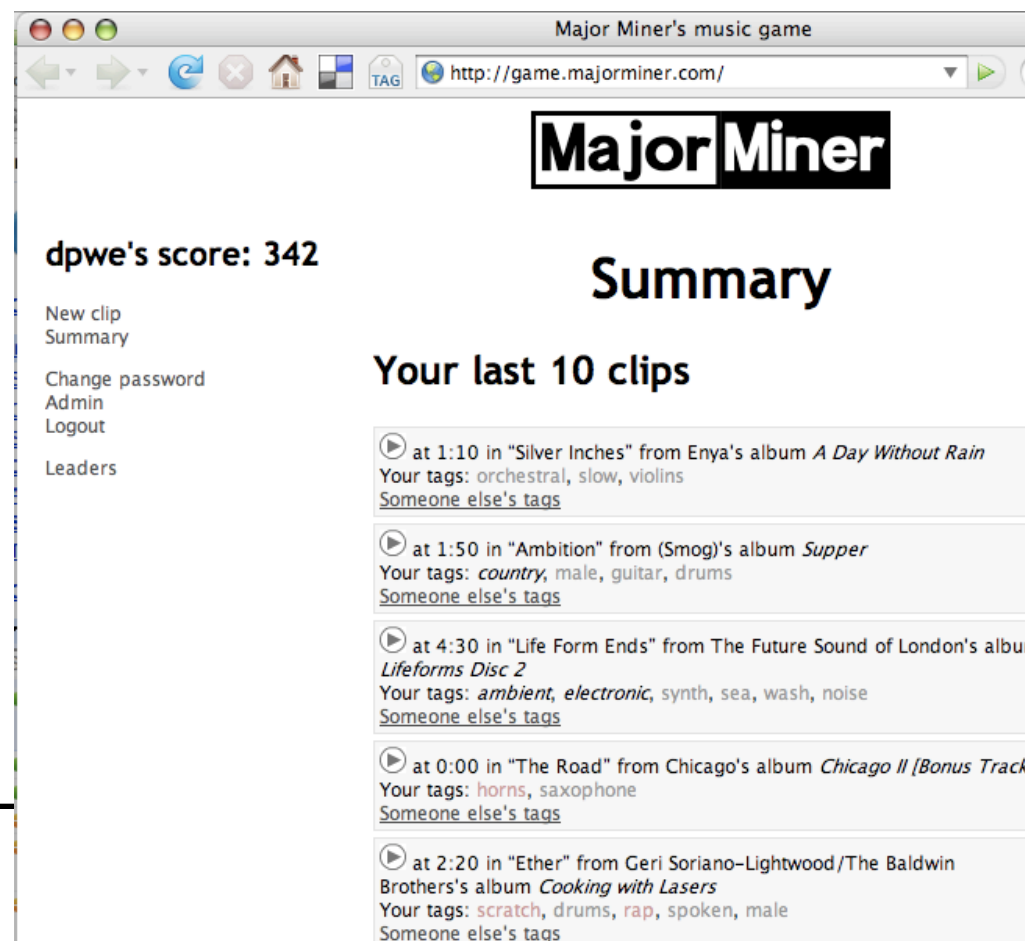


- Lyric Transcription?

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**...

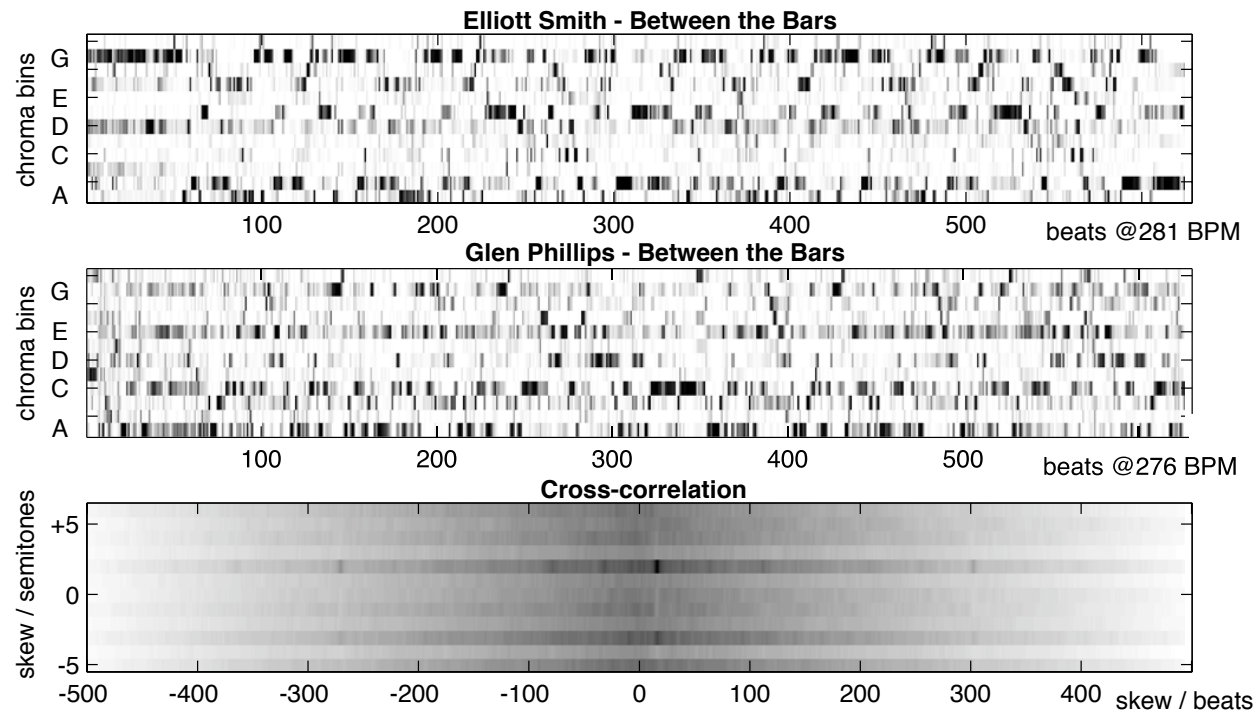


The screenshot shows a web browser window titled "Major Miner's music game" with the URL "http://game.majorminer.com/". The page features the "Major Miner" logo at the top. On the left, a sidebar lists navigation options: "New clip", "Summary", "Change password", "Admin", "Logout", and "Leaders". The main content area displays "dpwe's score: 342" and a "Summary" section titled "Your last 10 clips". This section lists five clips with their timestamps, album information, and semantic tags. For example, the first clip is "Silver Inches" from Enya's album "A Day Without Rain", with tags "orchestral, slow, violins".

Timestamp	Album	Your tags	Someone else's tags
at 1:10	"Silver Inches" from Enya's album <i>A Day Without Rain</i>	orchestral, slow, violins	Someone else's tags
at 1:50	"Ambition" from (Smog)'s album <i>Supper</i>	country, male, guitar, drums	Someone else's tags
at 4:30	"Life Form Ends" from The Future Sound of London's album <i>Lifeforms Disc 2</i>	ambient, electronic, synth, sea, wash, noise	Someone else's tags
at 0:00	"The Road" from Chicago's album <i>Chicago II [Bonus Track]</i>	horns, saxophone	Someone else's tags
at 2:20	"Ether" from Geri Soriano-Lightwood/The Baldwin Brothers's album <i>Cooking with Lasers</i>	scratch, drums, rap, spoken, male	Someone else's tags

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

Rosatron: listen
http://dawn.ee.columbia.edu:3210/main/listen

SA AUD dpwe E4896 PineGrv photos lapnap RGwiki Spectrograms: Const...

RosaTron

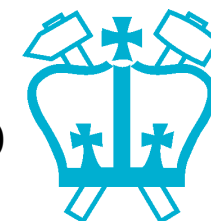
Query clip 3 of 30: ▶

Result clip 0: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 1: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 2: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 3: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 4: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 5: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 6: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 7: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 8: ▶	<input type="radio"/> not similar	<input type="radio"/> similar
Result clip 9: ▶	<input type="radio"/> not similar	<input type="radio"/> similar

Rate

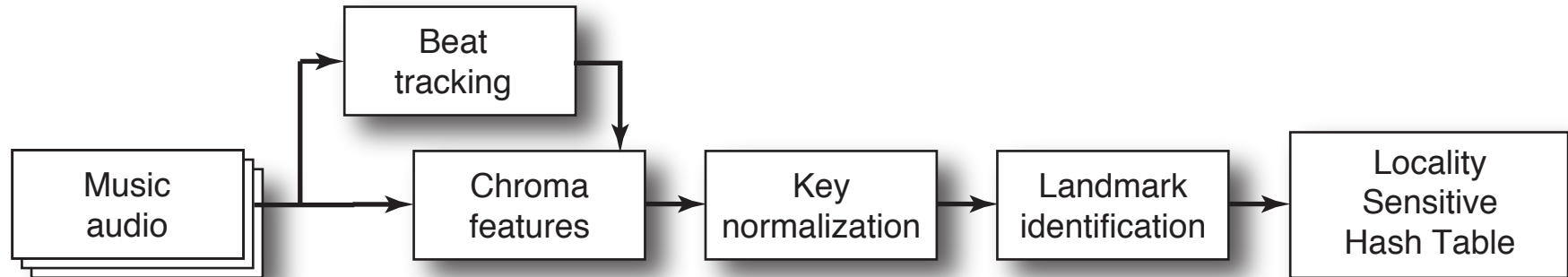
[Instructions](#)

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%



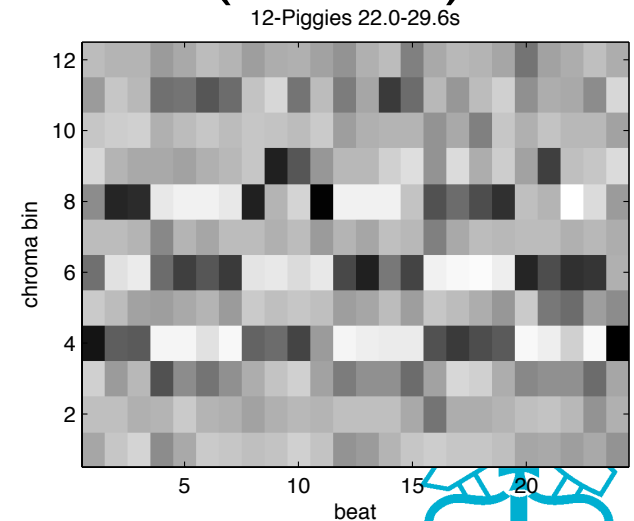
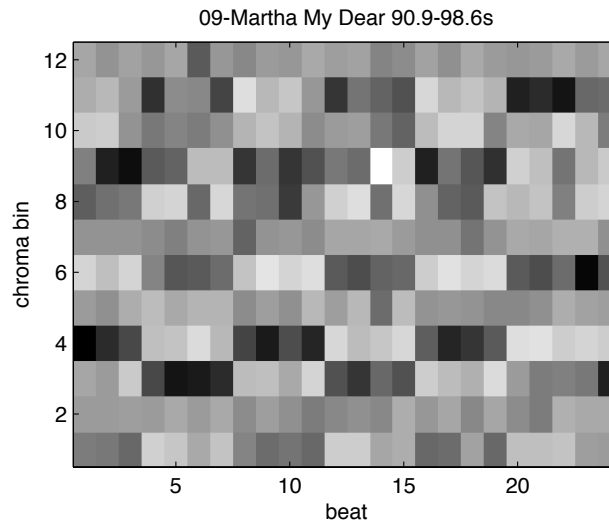
Beat Chroma Fragment Clustering

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



- 86 Beatles tracks \Rightarrow 41,705 patches (12x24)

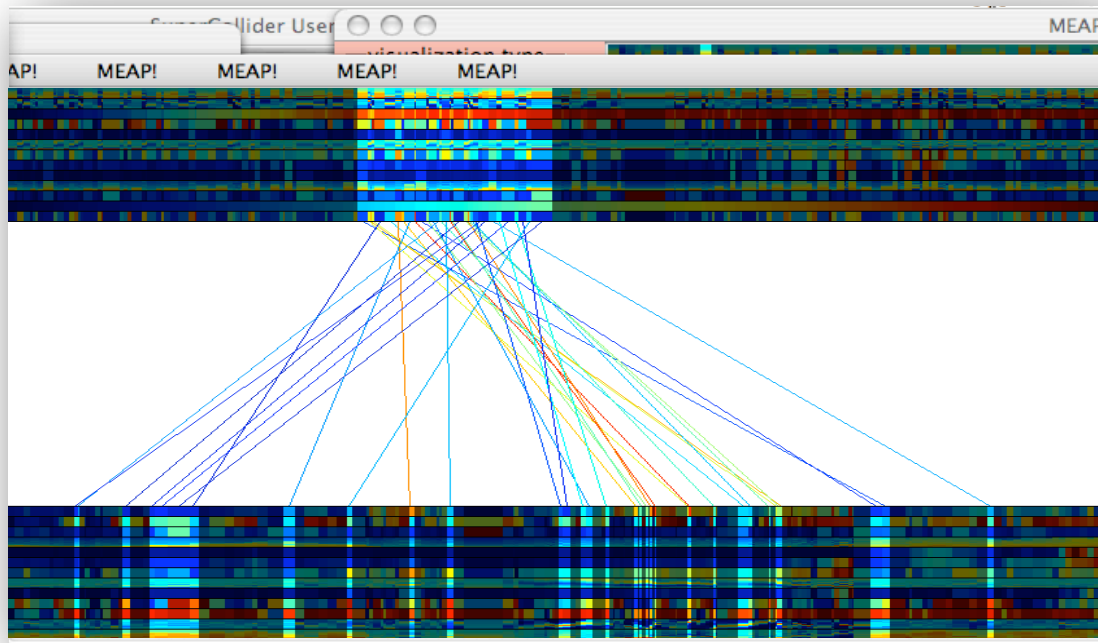
- LSH takes ~300 sec
- High-pass along time
- Song filter



MEAPsoft

- **M**usic **E**ngineering **A**rt **P**rojects
 - collaboration between EE and Computer Music Center
- MEAPsoft combines **music IR** analysis with wacky **resequencing** algorithms
 - also some neat visualizations...

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



Summary

- What is made **possible** by **so much data**?

