

Cross-Correlation of Beat-Synchronous Representations for Music Similarity

Dan Ellis, Courtenay Cotton, and Michael Mandel
Laboratory for Recognition and Organization of Speech and Audio
Dept. Electrical Eng., Columbia Univ., NY USA

{dpwe,cvcotton,mim}@ee.columbia.edu

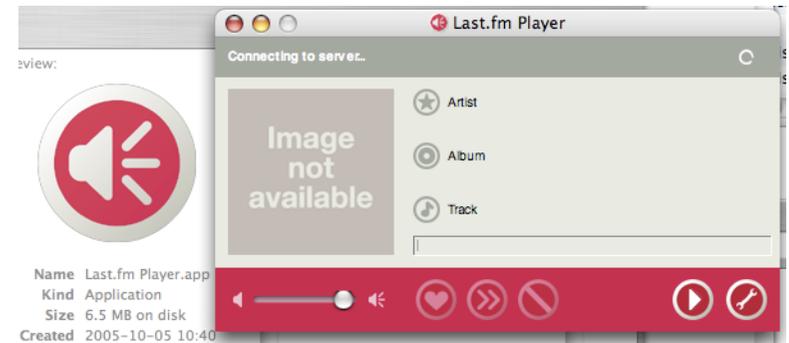
<http://labrosa.ee.columbia.edu/>

1. Music Similarity
2. Beat-Synchronous Representations
3. Cross-Correlation Similarity
4. Subject Tests



I. Music Similarity

- **Goal: Computer predicts listeners' judgments of music similarity**
 - e.g. for playlists, new music discovery

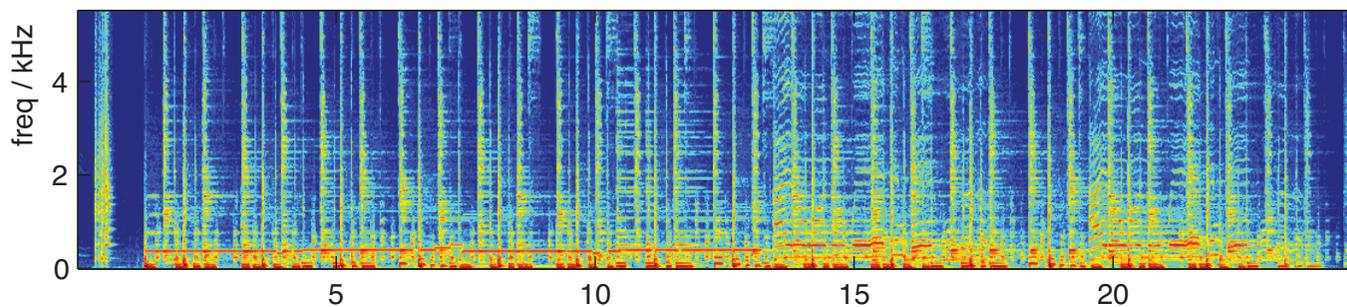


- **Conventional approach**
 - **statistical** models of broad spectrum (MFCCs)
- **Evaluation?**
 - MIREX: 2004 onwards
 - proxy tasks: Genre classification, artist ID ...
 - direct evaluation: **subjects** rate systems' hits

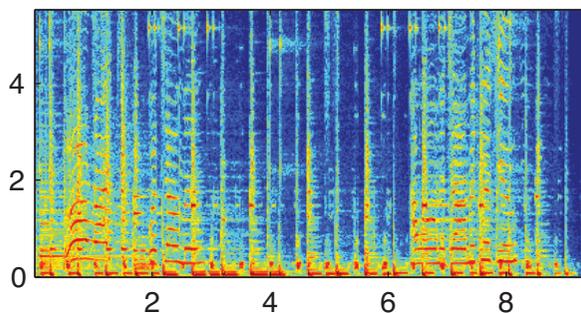
Which is more similar?

- “Waiting in Vain”
by Bob Marley & the Wailers

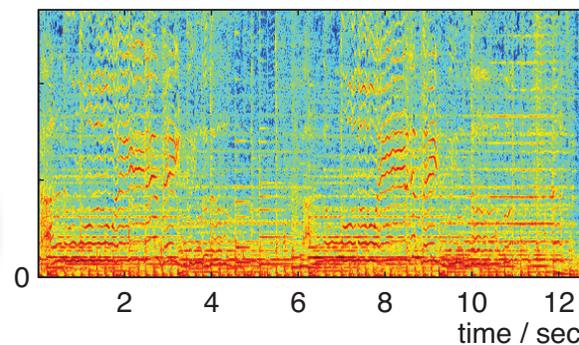
Waiting in Vain - Bob Marley



Jamming - Bob Marley



Waiting in Vain - Annie Lennox

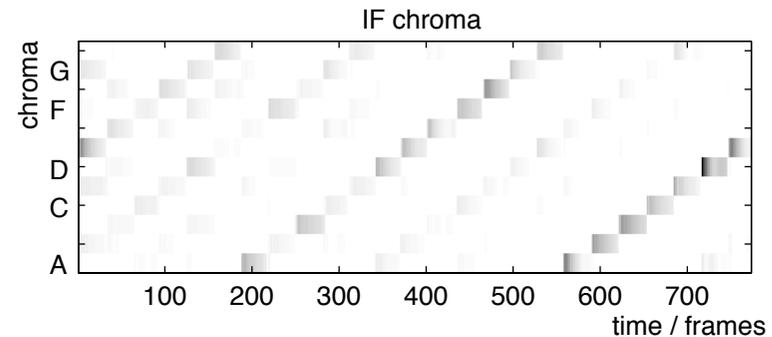
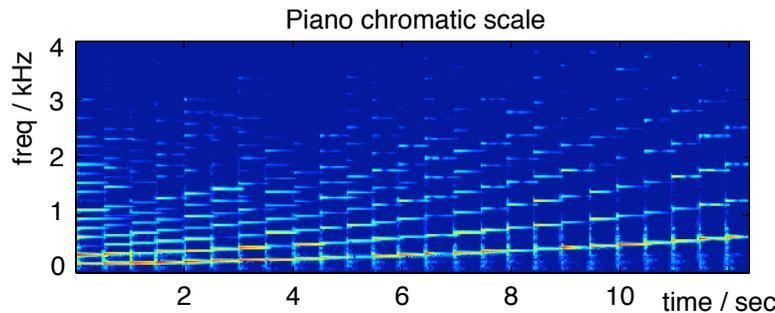


- Different **kinds** of similarity

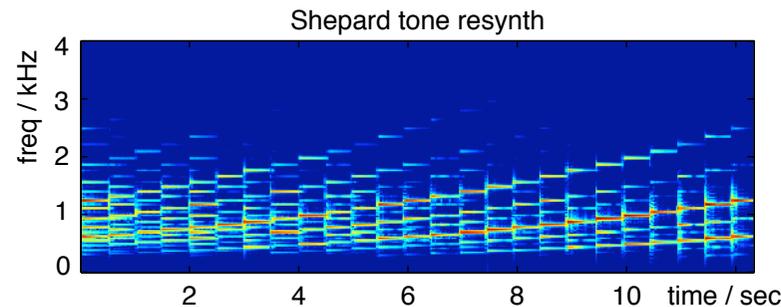
2. Chroma Features

- Chroma features map spectral energy into one **canonical octave**
 - i.e. 12 semitone bins

Piano scale

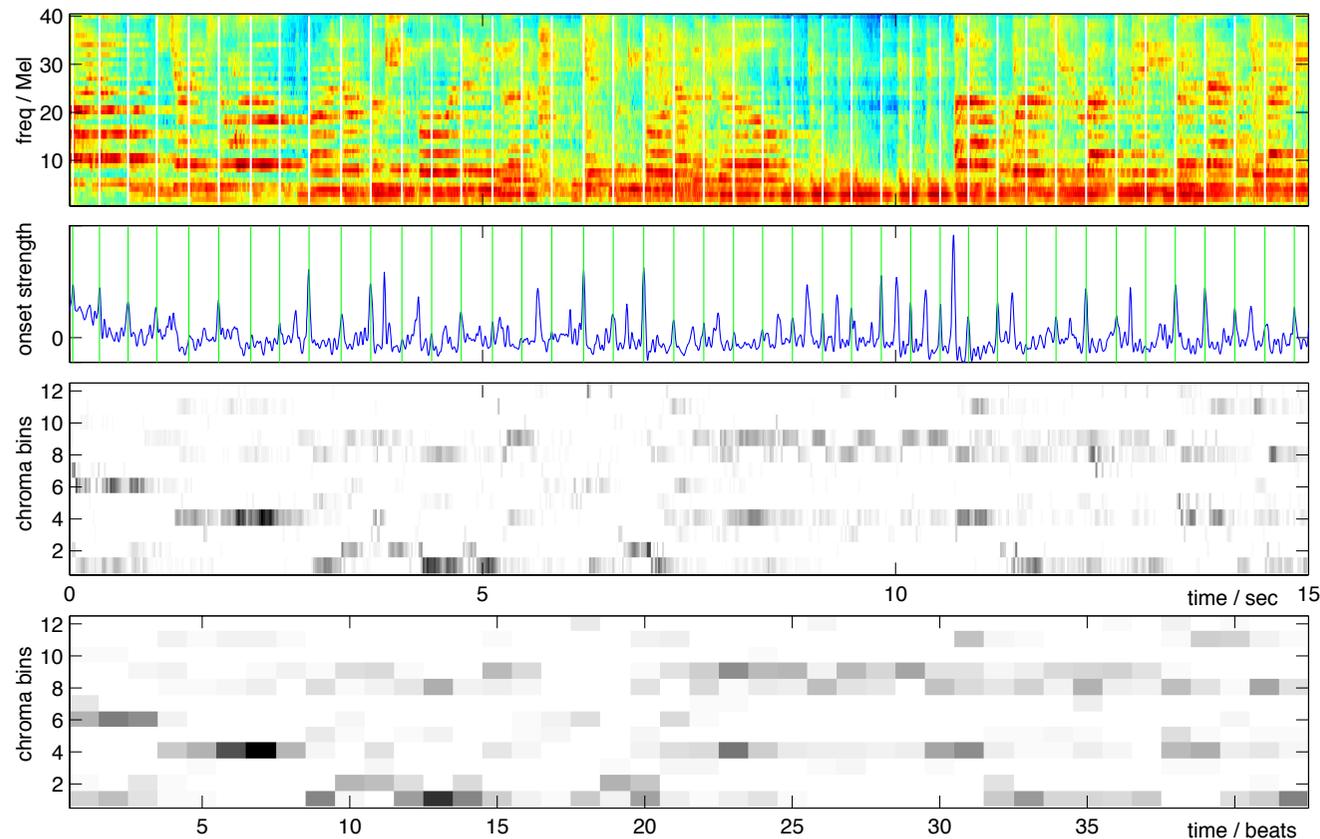


- Can resynthesize as “Shepard Tones”
 - all octaves at once



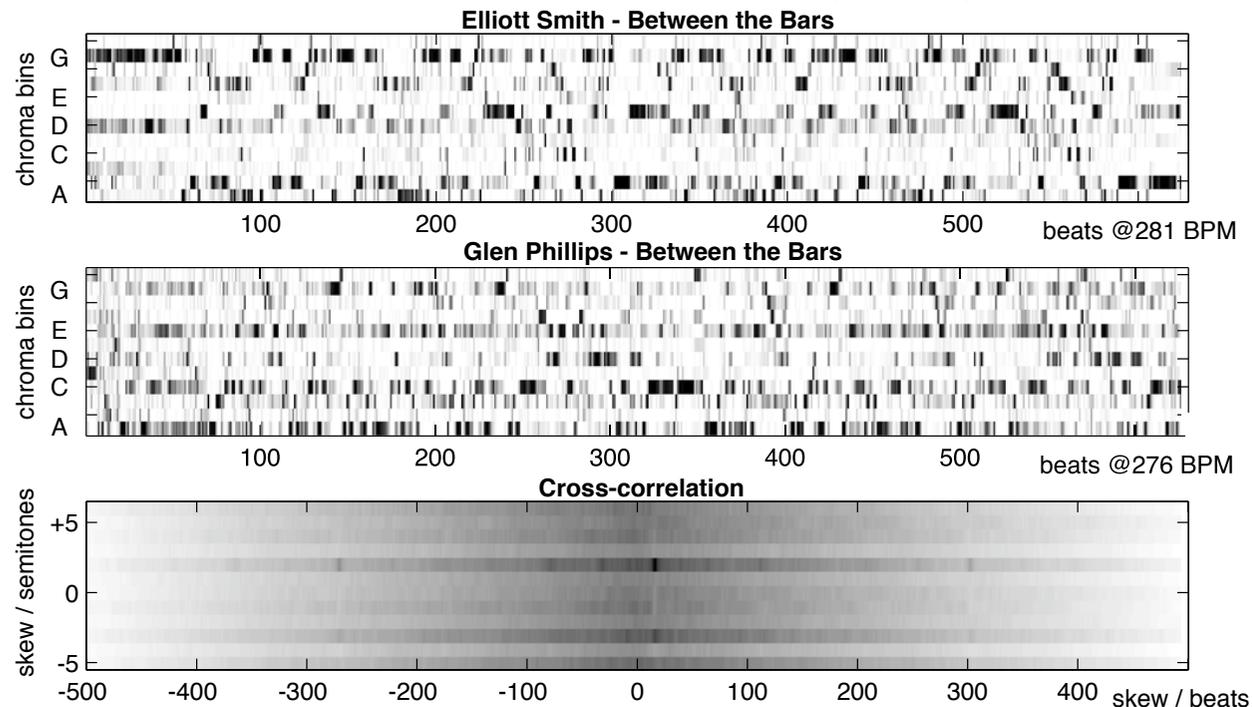
Beat-Synchronous Chroma Features

- **Beat + chroma features** / 30ms frames
 - **average chroma** within each beat
- compact; sufficient?



3. Cross Correlation

- Cross-correlate **entire** beat-feature matrices
 - ... including all **transpositions** (for chroma)
 - implicit **combination** of match quality and duration

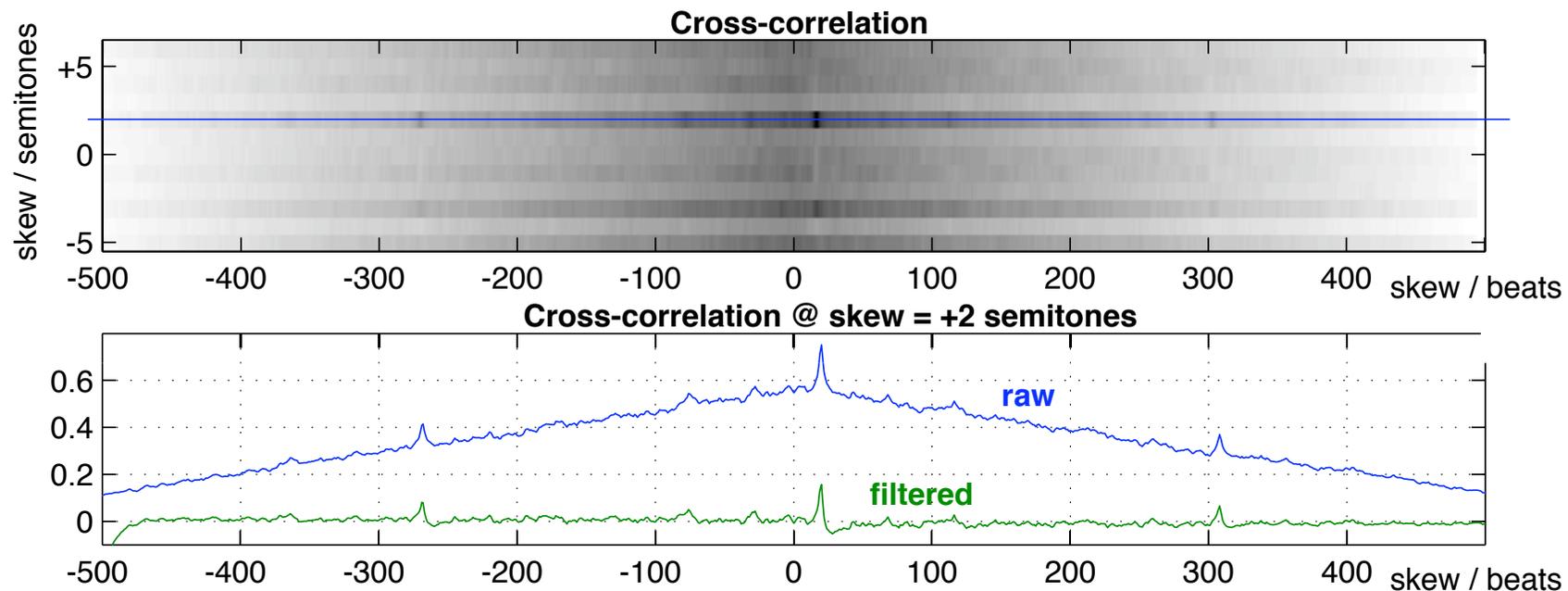


- One good matching fragment is sufficient...?



Filtered Cross-Correlation

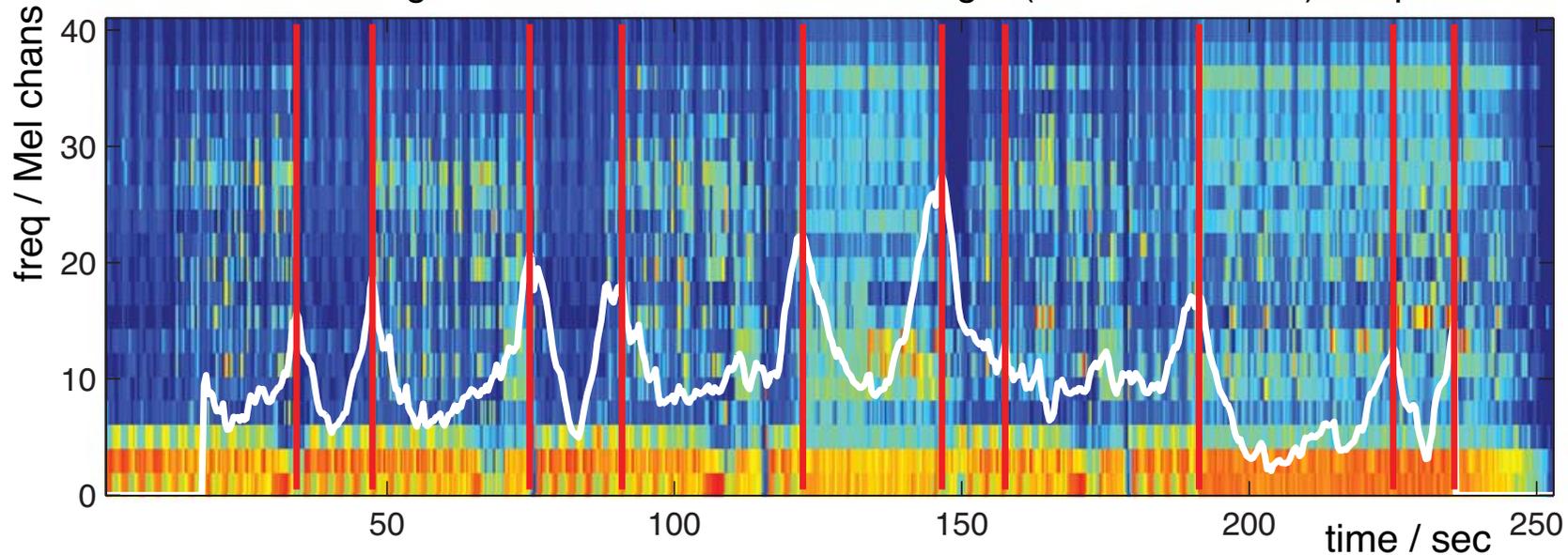
- Raw correlation not as important as precise **local match**
 - looking for large **contrast** at ± 1 beat skew
 - i.e. **high-pass filter**



Boundary Detection

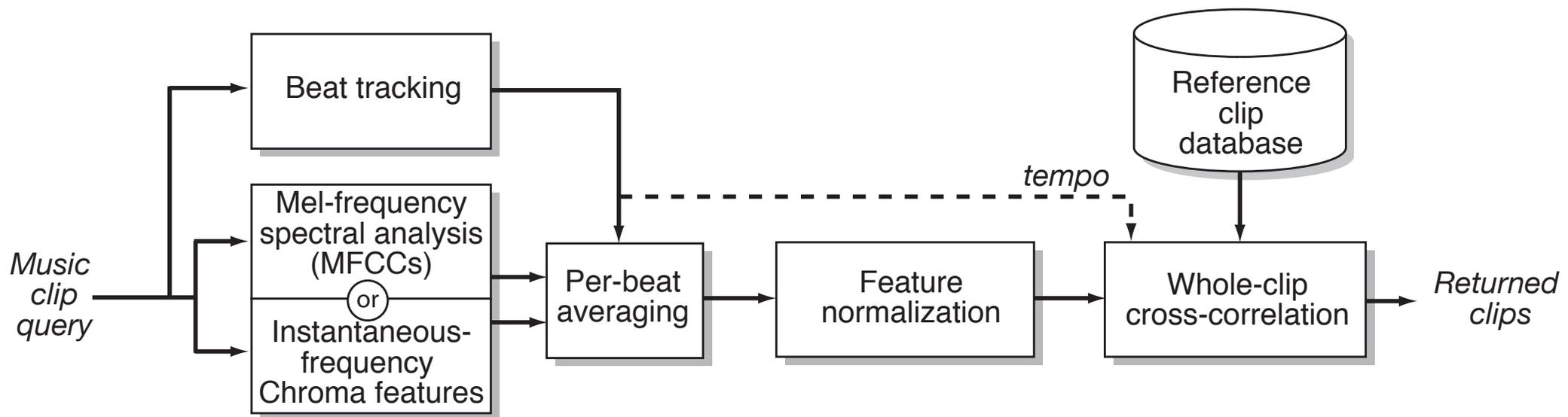
- If we had **landmarks**, no need to correlate
 - save time - **LSH** implementation
- Use single Gaussian model likelihood ratio to find point of greatest **contrast**

Come Together - The Beatles - localchange3(48 beat window) + top10



Correlation Matching System

- Based on **cover song** detection system
- Chroma and/or MFCC features
 - **chroma** for melodic/harmonic matching
 - **MFCCs** for for spectral/instrumental matching



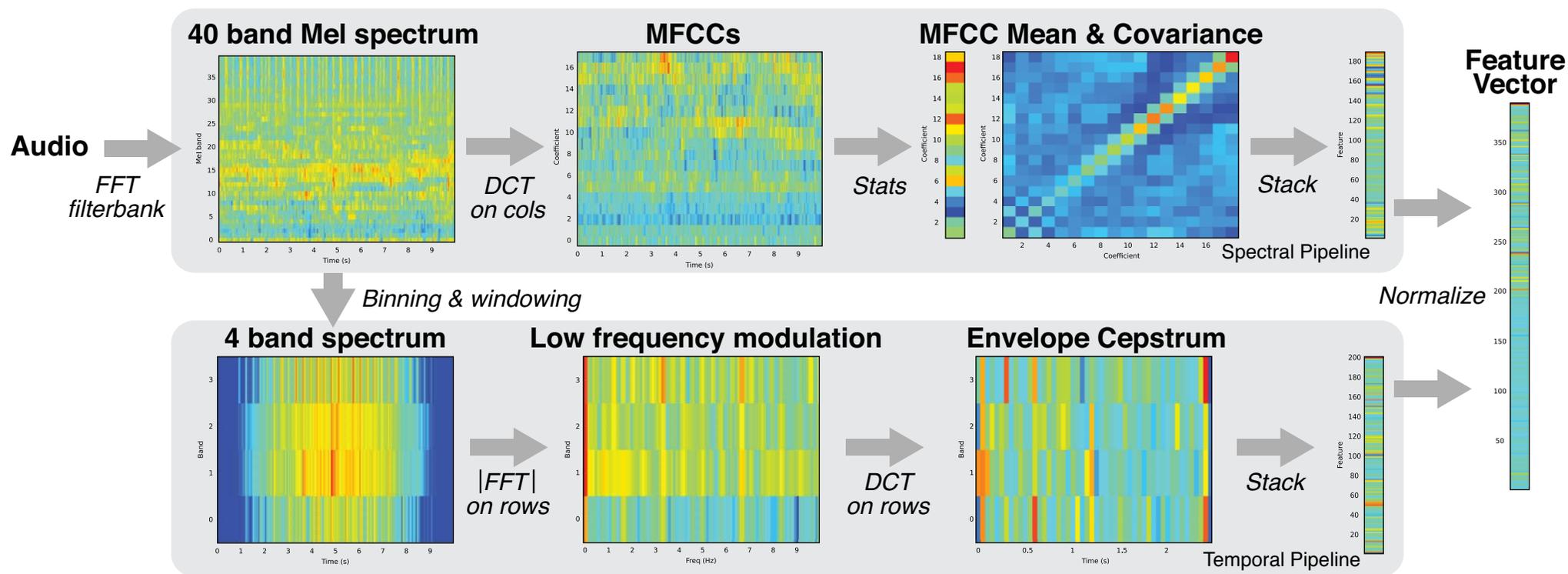
4. Experiments

- **Subject data** collected by listening tests
 - 10 different algorithms/variants
 - binary similarity judgments
 - 6 subjects \times 30 queries = 180 trials per algorithm



Baseline System

- From Mandel & Ellis MIREX'07
 - 10 sec clips (from 8764 track uspop2002)
 - spectral and temporal paths



- classification via SVM

Results

- Traditional (**baseline**) system does best:

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%

- **Cross-correlation** better than **random**...

Examples - Baseline

Dbasetf.html

file:///Users/dpwe/projects/musicsim/uspop2002/snips/Dbasetf.html

DX.html DXmhpf.html Dbasetf.html

Query	Result 1	Result 2	Result 3	Result 4	Result 5	Result 6	Result 7	Result 8	Result 9
See Saw Aretha Franklin	See Saw Aretha Franklin 9.00	All At Once Whitney Houston 8.00	soul deep Roxette 7.00	Earn While You Learn Bonus Track Elton John 6.00	Bed Of Nails Alice Cooper 5.00	Too Much Time On My Hands Styx 4.00	After the Glitter Fades Stevie Nicks 3.00	Just Another Woman Air Supply 2.00	Tra Paul 1.00
Dammit blink 182	Dammit blink 182 9.00	Last Caress Metallica 8.00	Thing That Should Not Be The Metallica 7.00	Green-tinted Sixties Mind Mr Big 6.00	Redundant Green Day 5.00	Dancing Queen Abba 4.00	Hummer Smashing Pumpkins 3.00	Voices Disturbed 2.00	I Ca Radi 1.00
One Wild Night Bon Jovi	One Wild Night Bon Jovi 9.00	Where Is Everybody Nine Inch Nails 8.00	Rocket to the Moon Culture Beat 7.00	Soul Surfing Fatboy Slim 6.00	You Janet Jackson 5.00	The Fragile Nine Inch Nails 4.00	History Repeating Propellerheads 3.00	Pretty Baby Blondie 2.00	Dig Sce 1.00
Touch the Hand Bryan Adams	Touch the Hand Bryan Adams 9.00	Rude Awakening 2 Creedence Clearwater Revival 8.00	Photograph Weezer 7.00	16 - Better Be Good To Me Extended Version Tina Turner 6.00	Tripping Billies Dave Matthews Band 5.00	Soolaimon Brother Love's Travelling Salvation Show Neil Diamond 4.00	Dynamite Scorpions 3.00	The Conqueror Genesis 2.00	Ign Rem 1.00
His Freinds Are More Than Fond Of Robin Carly Simon	His Freinds Are More Than Fond Of Robin Carly Simon 9.00	Jezebel Sade 8.00	Love Is Such A Crazy Thing Pink 7.00	I Remember Deborah Morgan 6.00	Dear Diary Britney Spears 5.00	One Sweet Day Mariah Carey 4.00	Giving Him Something He Can Feel En Vogue 3.00	Gypsy Queen Santana 2.00	Bea Dru 1.00
						Pleasure			



Examples - Xcorr Chroma

DX.html	DXmhp.html			
Last Ride In Green Day	Last Ride In Green Day -0.03	Dirty Day U2 -0.06	Who Wrote Holden Caulfield Green Day -0.06	Don t Tell Me Madonna -0.07
Don t Tell Me Madonna	Don t Tell Me Madonna -0.03	Dirty Day U2 -0.06	Tomorrow U2 -0.06	Original Sinsuality Tori Amos -0.07
Cry Freedom Dave Matthews Band	Cry Freedom Dave Matthews Band -0.03	Are We the Waiting Green Day -0.06	The Outlaw Torn Metallica -0.07	Motion Picture Soundtrack trimmed Radiohead -0.07
Lolita Suzanne Vega	Lolita Suzanne Vega -0.02	Dirty Day U2 -0.06	From the Edge of the Deep Green Sea Cure -0.07	Ninety-Nine And A Half Won t Do Creedence Clearwater Revival -0.07
Battery Metallica	Battery Metallica -0.02	The Outlaw Torn Metallica -0.06	Are We the Waiting Green Day -0.07	Rikki Don t Lose That Number Steely Dan -0.07
I Call Your Name Roxette	I Call Your Name Roxette -0.05	I Call Your Name Frank Mono-Mix 1987 Roxette -0.06	Stuart And The Ave. Green Day -0.09	Secrets That She Keeps Tits Ass Demo 1986 Roxette -0.09
Who Wrote Holden Caulfield Green Day	Who Wrote Holden Caulfield Green Day -0.04	Dirty Day U2 -0.06	Don t Tell Me Madonna -0.07	Boulevard of Broken Dreams Green Day -0.07



Conclusions and Future Work

- Music **similarity** is complicated
 - no single, simple, signal-processing model
- **Cross-correlation can detect ‘covers’**
 - or similar melodic-harmonic content
 - how common is this in practice?
- **Future work**
 - finding common 8-24 beat ‘**fragments**’
 - better analysis of song **structure**
- **Code available!**
Google “matlab cover songs”

