

# Content-based Media Recommender Systems: Are we there yet?

Presentation by Stephen Travis Pope  
stephen@FASTLabInc.com -- Oct 2008



1

## Abstract

Measured in terms of the amount of time they've been heralded as the "next big thing," few technologies (hydrogen cars and cold fusion, perhaps) can rival content-based multimedia search engines. Using data features derived from multimedia content such as sound or images (without requiring human-generated metadata), together with advanced data-mining techniques to deliver user-preference-related similarity metrics (with which to build search engines) has been a central topic in both image processing and music information retrieval for over a decade.



2

## Abstract

The last year has seen the introduction (to great, and largely undeserved, fanfare) of a whole raft of music recommender systems. This presentation will introduce the topic of music recommender systems, and examine the feature extraction and data mining techniques that are at the core of all of these products. Concrete examples will be presented from the author's own 4th-generation "SoLaTi" system, and several products will be compared in terms of the play lists they recommend for given input songs.



3

## Overview

- Context
- MMDB background
- Music information retrieval (MIR)
  - Audio signal analysis for MIR
  - Statistical data processing
  - Database techniques
  - Application examples
- Content similarity and play-list generation



4

## Background/Context

- Then (1ks)
  - MM content larger than computer storage
  - Networks too slow for content delivery
  - All metadata manually generated
- Now (2ks)
  - Local storage of (lossy compressed) AV content
  - WAN streaming of LC AV
  - Need automatic metadata generation for search, segmentation, summarization, content ID, etc.



5

## Multimedia Databases

- Applications
  - Indexing, segmentation, access
  - ID, summarization
  - "User preference matching," recommendation
- Techniques
  - Analysis and feature extraction
  - Numerical/statistical processing
  - Data dimensionality reduction
  - Machine learning (clustering, classification)
  - Data-mining (similarity, search)



6

## Design Variables

- Content format
  - Encoding (loss?), access
  - Dynamicity of content
- Desired metadata
  - Derived, correlated, indexing
  - Required feature extraction
- Application run-time
  - Analysis at run-time?
  - Query generation
  - Access, search/distance-match

## Development & Interactive App

- Database-oriented applications
  - Off-line DB population, data import & analysis, metadata generation
  - Run-time query preparation
- Media databases & RDBMs
- Run-time delivery vehicles
  - Location of metadata
  - Query format

## Content Analysis for MMDBs

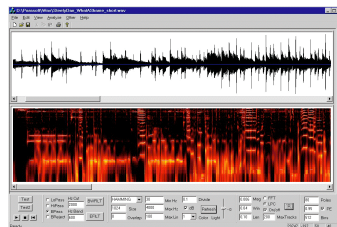
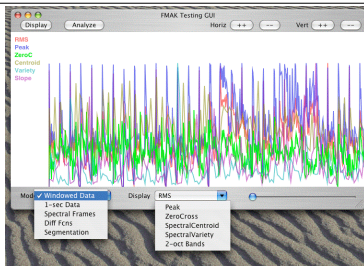
- Feature vector & DB design
- First-pass analysis
  - Direct feature extraction
- Second-pass analysis
  - Smoothing, pruning, reduction
  - Higher-level features
- Numerical/statistical analysis
  - Histogram, GMM, HMM
- Machine-learning, data-mining
  - Clustering, classification, structure-learning

## Music Information Retrieval

- Literature
  - ISMIR, ICMC, CCARH, AES, IEEE, ACM, ...
- Data formats
  - MIDI, audio (tracks, mixed, comp), score (OSR), ...
- Metadata
  - Basic audio features
  - Musical features
  - Text features, ID3
  - "Semantic audio"
- Applications ("MIR" and others)
  - Music DBs: search, access, match, ID, ...

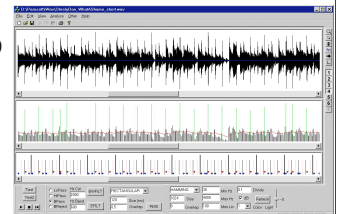
## Audio Feature Extraction

- First-pass analysis (windowed)
  - Time-domain features
  - Frequency/chroma-domain features
- Second-pass analysis
  - Higher-level features, peak tracking
  - Perceptual mapping
  - Smoothing, pruning, reduction



## Audio Feature Extraction 1

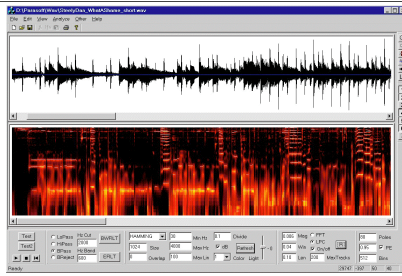
- First-pass (windowed) analysis
  - Time-domain features
    - Windowed RMS/peak amplitude (LF/HF bands)
    - Beats/tempo (AC, fit, model), tempo-changes
    - HF/LF RMS/tempo AC & histogram stats
  - Frequency-domain analysis
    - Spectral coefficients & spectral measures
    - MFCC components
    - LPC coeff, noisiness
    - Pitch-following
    - Hi-Freq bands
- Spatial/surround parameters
- Populate rich/large 1st-stage feature vector



# Audio Feature Extraction 2

## 2nd-pass analysis

- Smoothing, pruning, filtering, reduction
- Perceptual mapping
  - Loudness contour
  - Pitch, harmony and key
- Higher-level features
  - Spectral peaks, tracks, SMS model
  - Spectral track statistics (rate of birth/death)
  - Tempo, tempo changes, tempo curve



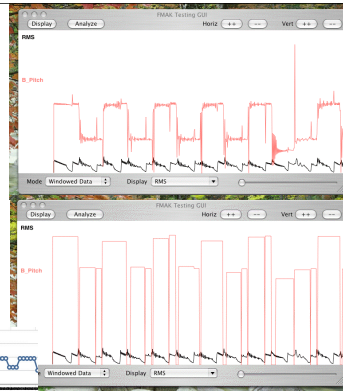
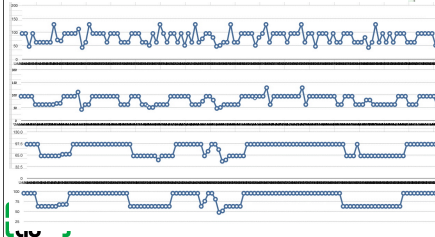
# Spectral Tracker Configurations

# Each entry consists of a line with 4 data values:  
 # peakWidth – closeness measure: peaks this close are considered to be one  
 # minPeakAmplMeanClearanceRatio – the amplitude mean clearance ratio is defined as the ratio of a peaks amplitude to the mean amplitude of the peaks in the containing window. Only peaks with clearance ratios above this parameter are considered when finding tracks  
 # birthFilterLevel – number of extra windows required to consider new peaks "births." A setting of 1 means it takes at least one more window with the peak ( 2 total ) to consider this peak as being born, etc..  
 # deathFilterLevel – number of extra windows required to consider missing peaks "deaths." A setting of 1 means it takes at least one more window missing the peak ( 2 total ) to consider peak dead, etc..  
 # These settings were arrived at after much testing; others are possible  
 SpectralTrackerConfiguration { 1.06, 0.05, 0, 2 }  
 SpectralTrackerConfiguration { 1.04, 0.01, 0, 2 }  
 SpectralTrackerConfiguration { 1.02, 0.01, 0, 3 }  
 SpectralTrackerConfiguration { 1.01, 0.01, 0, 4 }



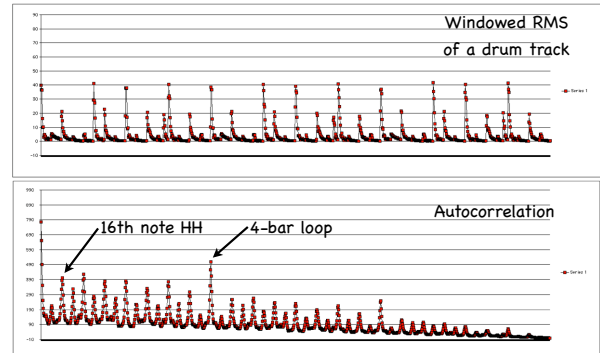
# Data Smoothing Examples

- Bass pitch (sticky value island-builder)
- Tempo est. (multi-pass de-spiker, then GMM)



# Tempo Tracking

- Autocorrelation, IOI GMMs
- Where are the beats?



# Tempo-finder Comparison

Track	FMAK	Marsyas	BPM Studio	Mix Meister	Mixed In Key	Beat Root	Randy
Albert Collins Master Charge	44 124	94/188	96	96	96	194	96
Beethoven - Symphony No. 9 in D minor op. 125 - Choral, 4th Movement	444 135	59/118	102	102	72	135	136
Beethoven - Symphony No. 9 in D Minor (Ode To Joy)	404 123	113/226	115	114	78	-	?
Boston Smokin	47 145	58/174	88	88	88	172	174
Lucinda Williams King of Hearts	437 123	75/150	145	145	146	141	146
John Coltrane Blue Train	404 130	74/148	136	135	139	125	136
Dave Brubeck Take 5	86 128	51/102	102	102	87	170	168
Yanni One Man's Dream	444 114	51/102	113	113	117	116	113
Ray Charles Mess Around	433 146	77/154	146	145	148	145	147
Ottmar Liebert Barcelona Nights	56 115	59/118	116	115	116	115	116
Bob Marley Is This Love?	56 121	59/118	116	115	116	121	116
Beastie Boys Time To Get Ill	45 120	99/198	98	97	98	195	98



# Example (Simple) Feature Vector

## Time-domain features

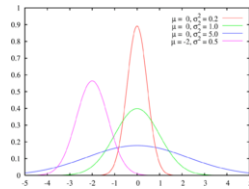
- Windowed RMS amplitude (avg & GMM lobes)
- Count of zero crossings
- Tempo estimate

## Frequency-domain features

- Spectral measures: centroid, slope, variety, flatness, harmonicity
- MFCC coefficients (4–12)
- Chromagram coefficients (4–8)



# Audio Feature Extraction 3



- Feature vector statistics
  - Per-song feature average, mean, variance
  - Feature H-gram/GMM stats (val as main PDF ctr)
  - Feature vector pruning (strip meaningless data)
- Segmentation
  - Locate regularly spaced changes
  - Per-segment statistics, fade-in/out
- Post-segmentation statistics

# Example (Rich) Feature "Vector"

- Time-domain features**
  - Windowed RMS amplitude
  - Max sample amplitude
  - RMS (ratio?) of LP/HP-filtered signal
  - Count of zero crossings
  - RMS dynamic range of sub-windows
  - RMS peak sub-window index
  - Tempo estimates (several)
  - Beat histograms & weights
  - Tempo weight & off-by-2 confidence
  - Time signature guess
- Frequency-domain features**
  - Windowed FFT data (stored?)
  - 1-octave FFT data (10-12 points)
  - 2.5-octave FFT data (4 spectral bands)
  - List of spectral peak indices
  - List of tracked peak frequencies
  - Spectral peak track births/deaths
  - Spectral measures: centroid, slope, variety
  - Relative HF level & spectral variety
  - Corr. between HF and audio-band
  - MFCC coefficients (4-12)
- Spatial features**
  - L/R difference
  - Front/Surround difference
  - Center vs. L/R sum difference
  - Spatial variety
- Pitch estimates**
  - Bass pitch guess in Hz
  - Bass note (MIDI key number) guess
  - Bass note dynamicity (size of histogram)
  - Multi-pitch estimates?
  - Chroma/key data
- LPC features**
  - List of LPC formant peaks
  - List of tracked LPC formants
  - LPC residual level (noisiness)
  - LPC formant track births/deaths
- Fluctuation Pattern features**
  - FP flux
  - FP gravity
  - FP weight
- Segmentation and segment statistics**

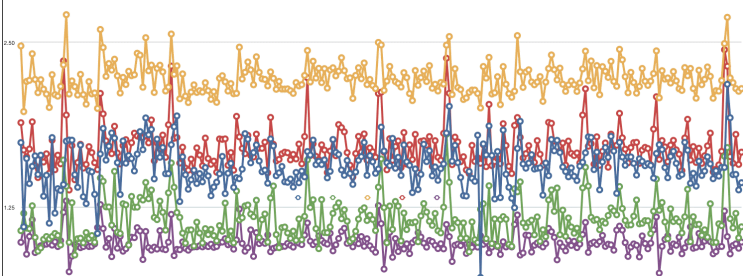
# Audio Segmentation

- Basic (time-domain) procedure
  - Pick a feature vector weighting
  - Calculate inter-window distances (scalar)
  - Identify regular peak spacing
- Challenges
  - Tempo changes
  - Intro/outro
  - Click-track tempo
  - Compressed dynamic range
  - Finding the "1"
- Aggressive (multi-weight, multi-tolerance blackboard) algorithm with confidence measure works ~85% of the time for our (very eclectic) test DB (1691 failed out of 14637) (allowing up to 30 segments)

# Segmentation Techniques/Options

- Distance metrics and inter-segment-boundary detection
- Finding relevant segmentation
  - Grouping short segments
  - Dividing long segments
  - HMMs and Viterbi
  - Similarity regions
  - Simulated annealing
  - Blackboard systems

# 4 Song Segmenter Distance Weightings



- Average, dynamic range, spikiness
- Choose red or green (?)

# Configurable Segmenter

```
#
# Segmenter Configurations
#
# Each block consists of a list of distance-metric weighting maps keyed by feature
# Spectral-/pitch-centric configuration
SegmenterConfiguration {
  HPRMS 0.5
  DynamicRange 0.5
  ZeroCrossings 0.5
  BassPitch 0.5
  SpectralSlope 1
  SpectralCentroid 1
  SpectralVariety 1
  SpectralBandMax 1
}

# MFCC- & tracking-centric configuration
SegmenterConfiguration {
  HPRMS 0.2
  SpectralVariety 1
  ZeroCrossings 0.2
  BassPitch 0.5
  STRackBirths 0.5
  STRackDeaths 0.5
  MFCCCoeff1 1
  MFCCCoeff2 1
  MFCCCoeff3 1
  MFCCCoeff4 1
  MFCCCoeff5 1
  MFCCCoeff6 1
}

# or use PCA or Tree weights
```

## Segmenter Confidence Measures

- How to compare segmentations
  - # of peaks per segment
  - # of segments per song (2-8)
  - % of song accounted for
  - % of peaks accounted for
  - Which weighting was used
  - Which tolerance was used
- Weighted metric of these?

## Song Segmentation Data

```
select artist, title, SegmentWeight, NumSegments, VerseLength,
       TypicalStart, SoloStart, SoloCentroid, SoloVariety, SoloTempo,
       SoloDynRange from fsongs where title = 'I Believe In Love';
```

artist	title	segmentweight	numsegments
Paula Cole	I Believe In Love	0.923772	0.24 (7)
		verselength	typicalstart
		0.631119	0.280232
		solostart	
		0.590672	
		s_centroid	s_variety
		0.4991	0.001422
		s_tempo	s_dynrange
		0.3360	0.654455

## Song/Segmentation Features

### General song metadata

- Title, artist (ID3 data)
- Duration, year
- "Genre" guess(es)

### Tempo features

- Average tempo estimates
- Tempo tracker confidence
- 1-sec dynamic range

### Beat histogram features

- BHSUM1
- BHSUM2
- BHSUM3
- Low/HIPeakAmp
- Low/HIPeakBPM
- HighLowRatio

### Fluctuation-pattern measures

- FP\_gravity
- FP\_bass
- FP\_focus

### Segmentation features

- Segmentation confidence
- NumSegments
- VerseLength

### Verse features

- FirstVerseStart
- SoloIndex
- TypicalIndex
- SoloStart
- TypicalStart

### Section features

- QuietSections
- LoudSections
- FadeIn
- FadeOut

### Solo features

- SoloCentroid
- SoloVariety
- SoloDynRange
- SoloRMS
- SoloTempo

## Advanced Segmentation

- Use derivative of distance vector?
- Adaptive feature weightings/tolerances
- Heuristic techniques
- Confidence calculus (multi-D)
- Robust tree-based segment percolation methods
- Post-segmentation statistics (can be quite valuable, when present)

## FV Pruning/Storage

- How to handle invalid data
  - If song is silent (set x/y/z to NULL)
  - If tempo guess invalid (BH sums the same)
  - If MFCC/LPC data not reasonable
  - If SegmentConfidence < s\_threshold
- DB output: SQL or file-based
- Write 1-4 FV records to DB
  - Avg, var, solo, typical FV records
- Write 1 FC record
  - Top-level metadata, ptrs to FV data
- Normalize DB?

## SQL Table Creation

```
CREATE TABLE FTables (
  RMS, Peak, LPRMS, HPRMS real,
  ZeroCrossings, DynamicRange, Tempo real,
  StereoWidth real,           — spatial-domain features
  SpectralBands real[4],      — 4-band spectrum as vector or values
  SpectralCentroid real,     — Spectral measures, MFCC, BassPitch, Tracks,
  [...]
);
CREATE TABLE FSongs (
  Name, Album, Artist, Title, Genre, SubGenre text,
  AvgFT, WeightedFT integer, — average and peak (or solo/typical) feature tables
  RMS, RMSVar real,         — Variance values for main columns
  [...]
  TempoAvg, TempoWeight, TempoDistr real, — various tempo stats
  SegmentWeight real,      — segmentation confidence
  NumSegments, TypicalIndex, SoloIndex integer, — # segments, index of verse and solo
  SoloCentroid, SoloVariety, SoloDynRange real, — ratios of the solo verse vs normal
  ClusterId integer DEFAULT 0, — the cluster it's in
  ClusterDistance real,    — distance to closest cluster representative
  ClusteringBatch integer DEFAULT 0, — cluster run of target
  ClusteringStatus char(1) DEFAULT 'O' CHECK (ClusteringStatus in ('O', 'X', 'P', 'R', 'F', 'M'...)), — make table copies
);
CREATE TABLE FSongStats AS SELECT * FROM FSongs;
```

# SQL Query Examples

— Count of all records  
 SELECT count(\*) from FSongs; → 14996

— Basic data quality tests  
 SELECT count(\*) from FSongs where TempoWeight < 0.1; → 84  
 SELECT count(\*) from FRawSongs where NumSegments > 30; → 648

SELECT min(SpectralCentroid), avg(SpectralCentroid), max(SpectralCentroid),  
 stddev(SpectralCentroid) FROM FDTables WHERE SpectralCentroid > 0;  
 → 0.492183 | 0.757860648821136 | 9.62927 | 1.20454739256262

SELECT min(TempoAvg), avg(TempoAvg), max(TempoAvg) FROM FSongs  
 WHERE TempoAvg > 0;  
 → 10.4896 | 93.7434194806919 | 193.838

SELECT min(SoloVariety), avg(SoloVariety), max(SoloVariety) FROM FSongs  
 WHERE SoloVariety > 0;  
 → 0.00309152 | 2.0699358320286 | 867.099

# DB Processing Techniques

- Machine-Learning, data-mining, AI
  - Many techniques
  - Many apps
- Dimensionality reduction
  - PCA, ISA, SOM, SVM, trees, nets, ...
- Clustering, classification
  - Fixing incomplete/noisy classification
- Similarity metrics & matching



# Feature Rank (InfoGain)

0.22089	8	SpectralVariety	0.06447	47	RMSVar
0.20219	55	SpectralVarietyVar	0.06263	44	fp_gravity
0.19994	46	fp_focus	0.05001	12	SpectralBandMax
0.17689	20	MFCCcoeff4	0.03816	38	HighPeakAmp
0.16469	18	MFCCcoeff2	0.03233	41	BHSUM1
0.16457	21	MFCCcoeff5	0.02625	37	LowPeakBPM
0.16404	19	MFCCcoeff3	0.02625	36	LowPeakAmp
0.12209	22	MFCCcoeff6	0.02625	40	HighLowRatio
0.12151	17	MFCCcoeff1	0.02625	39	HighPeakBPM
0.11668	45	fp_bass	0.02455	35	SoloRMS
0.08726	43	BHSUM3	0.02414	34	SoloTempo
0.08701	24	STrackDeaths	0.02105	42	BHSUM2
0.08701	23	STrackBirths	0.02001	33	SoloDynRange
0.08701	16	SpectralBand4	0.01918	9	SpectralFlux
0.08475	7	SpectralSlope	0.01775	15	SpectralBand3
0.08387	10	SpectralRollOff	0.01772	14	SpectralBand2
0.08115	51	ZeroCrossingsVar	0.01715	13	SpectralBand1
0.07597	4	ZeroCrossings	0.01026	25	TempoAvg
0.07	2	LPRMS	0.01001	32	SoloVariety
0.06621	3	HPRMS	0.00991	31	SoloCentroid

# 1st 2 PCA Dimensions

0.18 MFCCcoeff6 + 0.18 MFCCcoeff5 + 0.18 MFCCcoeff4 + 0.18 MFCCcoeff3 + 0.18 MFCCcoeff2 + 0.18 SpectralFlux + 0.18 SpectralRollOff + 0.18 SpectralFluxVar + 0.18 SpectralSlopeVar + 0.18 SpectralRolloffVar + 0.18 MFCCcoeff5Var + 0.18 MFCCcoeff4Var + 0.18 MFCCcoeff3Var + 0.18 MFCCcoeff2Var + 0.179 SpectralBand2Var + 0.179 SpectralBand1Var + 0.179 STrackBirthsVar + 0.179 STrackDeathsVar + 0.179 SpectralBand4Var + 0.179 SpectralBand3Var + 0.179 SpectralBand2

- 0.378 BHSUM3 - 0.345 LowPeakAmp - 0.323 BHSUM1 - 0.309 BHSUM2 - 0.298 HighPeakAmp - 0.294 fp\_bass - 0.269 ZeroCrossingsVar - 0.267 ZeroCrossings - 0.237 HPRMS + 0.161 TempoWeight + 0.161 TempoAvg - 0.15 fp\_gravity - 0.12 LPRMSVar + 0.114 HighPeakBPM + 0.114fp\_focus + 0.093 LowPeakBPM + 0.086 HPRMSVar - 0.058 SpectralVarietyVar + 0.054 QuietSections + 0.053 LPRMS - 0.049 SpectralBandMaxVar - 0.049 RMSVar - 0.045 PeakVar - 0.041 SoloTempo

SpectralVarietyVar <= 0.021884  
 | fp\_focus <= 0.415299  
 | | MFCCcoeff1Var <= 0.127492  
 | | | fp\_bass <= 0.412635  
 | | | | BassDynamicity <= 0.698656  
 | | | | | DynamicRangeVar <= 0.370633  
 | | | | | SpectralCentroidVar <= 0  
 | | | | | | HPRMS <= 0.867819: **Rock-Alternative (17.0/3.0)**  
 | | | | | | HPRMS > 0.867819  
 | | | | | | | SoloStart <= 0.015131: **Rock (3.0)**  
 | | | | | | | SoloStart > 0.015131: **Rock-Alternative (3.0/1.0)**  
 | | | | | | | SpectralCentroidVar > 0: **Rock (4.0/2.0)**  
 | | | | | | | DynamicRangeVar > 0.370633: **Rock (5.0/1.0)**  
 | | | | | | | BassDynamicity > 0.698656  
 | | | | | | | LPRMS <= 0.852023: **unknown (2.0/1.0)**  
 | | | | | | | LPRMS > 0.852023: **Pop-BritPop (2.0/1.0)**  
 | | | | | | | fp\_bass > 0.412635  
 | | | | | | | LPRMS <= 0.687344: **Comedy (2.0/1.0)**  
 | | | | | | | LPRMS > 0.687344: **unknown (2.0/1.0)**  
 | | | | | | | MFCCcoeff1Var > 0.127492  
 | | | | | | | FadeOut <= 0.1  
 | | | | | | | NumSegments <= 0: **Jazz-Big Band\_Swing (3.0/1.0)**  
 | | | | | | | NumSegments > 0  
 | | | | | | | SoloVariety <= 0.001143: **Blues (3.0/1.0)**  
 | | | | | | | SoloVariety > 0.001143: **Rock (3.0/1.0)**

## \*ART-Tree-learning

# Tree-training

- Example: Mulcher CART trees

The screenshot shows the FasMulcher 2.1.04 interface. The main window displays a decision tree for song classification. The tree structure is as follows:

- Root Node: SoloStart <= 0.015131: **Rock (3.0)**
  - Left Child: SoloStart > 0.015131: **Rock-Alternative (3.0/1.0)**
  - Right Child: SoloStart <= 0.015131: **Rock (3.0)**
    - Left Child: SpectralCentroidVar > 0: **Rock (4.0/2.0)**
    - Right Child: SpectralCentroidVar <= 0: **Rock (3.0)**
      - Left Child: DynamicRangeVar > 0.370633: **Rock (5.0/1.0)**
      - Right Child: DynamicRangeVar <= 0.370633: **Rock-Alternative (17.0/3.0)**
        - Left Child: HPRMS > 0.867819: **Rock-Alternative (17.0/3.0)**
        - Right Child: HPRMS <= 0.867819: **Rock-Alternative (17.0/3.0)**
          - Left Child: BassDynamicity > 0.698656: **Rock-Alternative (17.0/3.0)**
          - Right Child: BassDynamicity <= 0.698656: **Rock-Alternative (17.0/3.0)**
            - Left Child: LPRMS > 0.852023: **Pop-BritPop (2.0/1.0)**
            - Right Child: LPRMS <= 0.852023: **unknown (2.0/1.0)**

# Rule-learning

```
fp_focus > 0.519652
AND fp_bass <= 0.273051
AND LPRMS <= 0.858613
AND ZeroCrossings <= 0.197117
AND FadeOut > 0.1 AND FadeOut <= 0.7
AND SpectralVarietyVar <= 0.021011
AND MFCCcoeff1Var <= 0.393386
AND SpectralBandMaxVar <= 0.003615
→ Classical (78.0/1.0)
```

# Applications

- ContentID, finger-printing, cover-song detection
  - Generate unique key from song data
- Summarization, thumb-nailing
  - Find most "representative" snippet
- Segmentation, indexing
  - Store splices, manual labeling, speaker ID, classify segments, etc.
- Mapping (e.g., for DJing or mastering)
- Similarity-based queries

# MIREX Tests

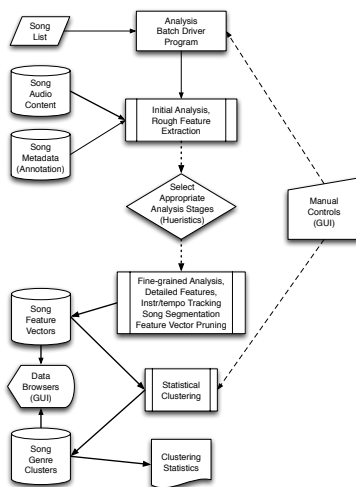
- Genre classification
  - General collection: 7000 30-second clips, 22.05kHz uncompressed, 10 genres
  - Latin collection: 3227 files from 10 genres
- Artist ID
- Cover song ID
- FO/chord/melody ID
- QbHumming results

# Examples: MIREX Results

MIREX Task Results			
Task	Participant	Accuracy	Rank
Audio Genre Classification (GenreMixer)	1. Tsunashima, S. (Japan)	48.11%	1
	2. Muralidharan, S. (India)	47.25%	2
	3. Tsunashima, S. (Japan)	46.11%	3
	4. Tsunashima, S. (Japan)	45.47%	4
	5. Tsunashima, S. (Japan)	45.21%	5
	6. Tsunashima, S. (Japan)	44.85%	6
	7. Tsunashima, S. (Japan)	44.85%	7
	8. Tsunashima, S. (Japan)	44.85%	8
	9. Tsunashima, S. (Japan)	44.85%	9
	10. Tsunashima, S. (Japan)	44.85%	10
Audio Artist Identification	1. Tsunashima, S. (Japan)	48.11%	1
	2. Muralidharan, S. (India)	47.25%	2
	3. Tsunashima, S. (Japan)	46.11%	3
	4. Tsunashima, S. (Japan)	45.47%	4
	5. Tsunashima, S. (Japan)	45.21%	5
	6. Tsunashima, S. (Japan)	44.85%	6
	7. Tsunashima, S. (Japan)	44.85%	7
	8. Tsunashima, S. (Japan)	44.85%	8
	9. Tsunashima, S. (Japan)	44.85%	9
	10. Tsunashima, S. (Japan)	44.85%	10
Audio Chord Detection (Pre-Train)	1. Tsunashima, S. (Japan)	48.11%	1
	2. Muralidharan, S. (India)	47.25%	2
	3. Tsunashima, S. (Japan)	46.11%	3
	4. Tsunashima, S. (Japan)	45.47%	4
	5. Tsunashima, S. (Japan)	45.21%	5
	6. Tsunashima, S. (Japan)	44.85%	6
	7. Tsunashima, S. (Japan)	44.85%	7
	8. Tsunashima, S. (Japan)	44.85%	8
	9. Tsunashima, S. (Japan)	44.85%	9
	10. Tsunashima, S. (Japan)	44.85%	10

# Example Framework: (F)MAK4

- Batch analysis
- Segmentation
- Clustering
- Matching/play-list generation



# Music/Sound Database Projects

- DoubleTalk/HyperScore/MODE/Siren (1980s - present)
  - Composer's tools: persistency, data-mining & versioning
- Paleo/SMS/CSP (1996-8)
  - MIDI performance analysis, expression data-mining
- NOLib (1998)
  - Comprehensive feature extraction framework
- 8S Speech segmenter & database
- FASTLab 1 (-> MusicMagic, MusicIP, 1999 - present)
  - Finger-printing (libOFA)
- OMNI/LoCAA (2001)
  - Network-based music access, recommender systems
- FASTLab 2: Expert Mastering Assistant (EMA) (2002-3)
- FASTLab 3: Locus animation system (2006-7)
- FASTLab 4: MAK4 core & recommender system (2007-8)

## MAK MIR Utilities

- make\_list.sh - create song lists using find in the shell
- fbatcher - parse list file and make analysis script
- fanalyzer - run 1 song, generate 3 records
- fcleanup.sql - prune DB data, setting NULLs
- fnormalizer - join & normalize song table
- fclusterer - find/label DB genre clusters
- fdistance - song set distance matrices
- frecommender - create playlists based on similarity
- MATLAB scripts - tempo/chroma features
- Smalltalk scripts - numerics, WEKA/SQL bridge
- WEKA scripts - DB clustering, PCA



43

## Batch Analysis in a Shell

```
# make song list (use your own files)
find ~/music -name \*.mp3 -print > song_list.txt
```

```
# make batch analyzer script
fbatcher song_list.txt ~/music > populate.sh
```

Each line in the script is a multi-stage analyzer command such as

```
fanalyzer -s "/Volume/Genre/SubGenre/Artist/Album/Title.mp3"
-i "Genre/SubGenre" "Artist" "Album" "Title"
-x Configuration.txt;
ftempostats "/Volume/Genre/Artist/Album/Title.aiff" tempo_stats_sql.txt;
fcleanup "/Volume/Genre/Artist/Album/Title.aiff"
"/Volume/Genre/Artist/Album/Title.mp3" failed_song_list.txt
```

Split populate.sh script and run in parallel on many cores



44

## Test Data Sets

- EMA data set
  - 3500 "well-produced" songs
- STP data set
  - 5000 random
  - Lots of new/abstract music & spoken word
  - Large holes (no country, new wave, etc.)
- Catalyst data set
  - 15000, widely-distributed (few holes)
  - Less classical, spoken word



45

## Simple SQL Processing

```
SELECT min(SpectralCentroid), avg(SpectralCentroid), max(SpectralCentroid),
stddev(SpectralCentroid) FROM FDTables;
```

```
SELECT min(SoloTempo), avg(SoloTempo), max(SoloTempo) FROM FRawSongs;
# Analysis post-processing — filter outliers before normalization
# Pruning the analysis database-set IgnoreMe for silent songs...
```

```
sql user/password@database < Cleanup.sql
```

```
fnormalizer # run the DB normalizer, create/join FSongs table; dump stats
# Now use SQL to check the normalized song tables
```

```
SELECT count(*) from FSongs WHERE TempoWeight < 0.1;
```

```
SELECT count(*) from FSongs WHERE SegmentWeight < 0.1;
```

```
SELECT title, album, fc_id FROM FSongs WHERE artist like '%Beatles%';
```



46

## Normalization Table

FNormalizer update loop

RMS	max:	0.638273	avg:	0.202503	var:	0.016578	dev:	0.128756		
Peak	max:	1.70248	avg:	0.561167	var:	0.129063	dev:	0.359253		
LPRMS	max:	2.41038	avg:	1.0482	var:	0.141003	dev:	0.375504		
HPRMS	max:	1.66038	avg:	0.504619	var:	0.0632947	dev:	0.251584		
ZeroCrossings	max:	122	avg:	46.6459	var:	517.531	dev:	22.7493		
DynamicRange	max:	2.74092	avg:	2.2039	var:	0.331195	dev:	0.575495		
BassDynamicity	max:	0.761808	avg:	0.0486035	var:	0.0270738	dev:	0.164541		
StereoWidth	max:	5.90478	avg:	1.47027	var:	0.82816	dev:	0.910033		
SpectralCentroid	max:	377.162	avg:	219.996	var:	5399.51	dev:	73.4814		
SpectralSlope	max:	6.0231	avg:	0.976817	var:	0.456675	dev:	0.675777		
SpectralBandMax	max:	3.86842	avg:	1.41041	var:	0.49113	dev:	0.700806		
STrackBirths	max:	0.37037	avg:	0.0522794	var:	0.00191105	dev:	0.0437156		
STrackDeaths	max:	1.16993	avg:	0.0603716	var:	0.0145724	dev:	0.120716		
[ ... ]										
MFCFfirst	min:	-118.78	max:	169.699	avg:	56.441	var:	2818.36	dev:	53.0883



47

```
UPDATE FSongs SET ClusteringStatus = 'X';
fclusterer -p 4 -q 3 -k 20 -c 5 -a 0.7 # run the CURE multi-stage clusterer
-p [NUM] - Number of pre-clustering partitions (2 - 10)
-q [NUM] - Pre-clustering partition factor (3)
-k [NUM] - Desired number of clusters (4 - 50)
-c [NUM] - Representatives per cluster (2 - 15)
-a [NUM] - Representative scaling factor (0.2 - 0.7)
```

## Clustering & Distances

```
# SQL examples to view clusters
SELECT min(ClusterDistance), avg(ClusterDistance), max(ClusterDistance),
FROM FSongs WHERE ClusterID = 6 AND ClusterStatus = 'M';
SELECT genre, subgenre FROM FSongs
WHERE FC_ID in (select Song from FSongs WHERE ClusterID = 1);
SELECT count(*) FROM FSongs WHERE ClusterID = 1 AND ClusterStatus = 'M'; (or 'R')
```

```
# distance measure on 6 Blondie songs - they look similar (d < 0.2)
fdistance 45851 45953 45941 45929 45902 45962
```

45851	0.00000	0.18307	0.14211	0.12747	0.13292	0.12819
45953	0.18307	0.00000	0.12112	0.16571	0.13695	0.17453
45941	0.14211	0.12112	0.00000	0.16523	0.08877	0.12869
45929	0.12747	0.16571	0.16523	0.00000	0.18698	0.19223
45902	0.13292	0.13695	0.08877	0.18698	0.00000	0.12974
45962	0.12819	0.17453	0.12869	0.19223	0.12974	0.00000



48



## Commercial Music-access Players



55

## Application: Recommender Systems

- Search-by-user-preference and automatic play-list generation
- Content access, play-list generation
  - Song ID, feature extraction
  - Similarity search/sort
  - Play-list sequencing (arch, cresc, tempo, energy)
- Multimedia-related tools
  - Human-supplied metadata
  - Automatic metadata only



56

## Music Discovery in the Small

- (Celma & Lamere Tutorial: ISMIR 2007)
- Personal music players
  - No experts to guide you
  - No social network
  - Music discovery is random
  - Shuffle play doesn't scale
- Study of 5,000 iPod users:
  - 80% of plays in 23% of songs
  - 64% of songs never played



57

## The Value of Recommendation

- Netflix
  - 2/3 of movies rented were recommended
  - Recommendation is "absolutely critical to retaining users"
- Google News
  - Recommendations generate 38% more click-throughs
- Amazon
  - 35% of product sales result from recommendations



58

## What is Recommendation?

- A kind of data-mining (or information-filtering) problem
- Task
  - Select items that match the user's preferences
  - Predict items the user will like
  - Selection problem separate from sequencing
- May involve (content- or metadata-based) similarity metric



59

## Recommender Options

- Manual or automatic metadata?
  - Reasonable to assume manual processing?
- Are all songs known?
  - Reasonable to assume 1 site knows all songs?
- One or several key songs
  - From/to sequencing
- Which similarity metric?
- Which sequencing policy?



60

## Similarity (again)

- Match text metadata
  - Artist, year, label
- Match "genre" (or genre distance)
- Play-list co-occurrence, social
  - Other customers also bought...
  - Web search of play lists
  - Most common
- Content-based (FV distance)
  - Most interesting, hardest
- Finger-print distance (meaningful?)

## Example Play-list



Glenn Gould -- WTC/I Prelude No. 2 in C m

30:	Andrés Segovia	--	Pavana No. 3	--	Classical
41:	Glenn Gould	--	Variations XXVI and	--	Classical/Symphony
48:	Händel; Pinnock	--	Accompagnato (Bass)-	--	Classical/Symphony
59:	Andrés Segovia	--	Pavana No. 6	--	Classical
63:	Andrés Segovia	--	Prelude	--	Classical
64:	unknown	--	Raindrop Prelude	--	Soundtrack
64:	James Edwards	--	Plaisir D'Amour-Gi	--	Classical/Symphony
66:	unknown	--	Raindrop Reprise	--	Soundtrack
67:	Bing Crosby Frank Si	--	O Little Town Of Bet	--	Religious/Holiday
70:	Spencer the Gardener	--	LuLu Interlude	--	Rock
70:	Hootie & The Blowfis	--	Sometimes I Feel	--	Rock/Alternative
72:	Glenn Gould	--	Prelude in D minor,	--	Classical/Symphony

## Music Recommender Systems

- (selected, in approx. order of release)
- MusicIP MyDJ (FMAKO++)
- QMUL SoundBite
- MIT/EchoNest MusicBrain/API
- FMAK/SoLaTi
- iLike
- Apple/Gracenote MusicGenius
- MS Zune 3.0

## Apple/Gracenote MusicGenius

- Finger-print: take 1st 15 sec of song and make a unique key (finger-print)
- Get song metadata from Gracenote's DB of "all" music
- Compare with metadata for other songs in your personal DB (Apple process)
- Generate similarity-based play-list using artist/genre metadata & your own play-list co-occurrence

## MusicGenius Play-list (19kSongs)

Name	Artist	Genre
A Case Of You	Joni Mitchell	Folk
The Needle And The Damage Done	Neil Young	Rock
Angel From Montgomery	Bonnie Raitt W/ John Prine	Rock
The Wind	Cat Stevens	Folk
Crazy Love	Van Morrison	Pop
Alison	Elvis Costello	New Wave
For Emily, Whenever I May Find Her [Live]	Simon & Garfunkel	Folk
Sweet Baby James	James Taylor	Folk
Helplessly Hoping	Crosby, Stills, Nash & Young	Folk
Shelter From The Storm	Bob Dylan	Folk
The Only Living Boy In New York	Simon & Garfunkel	Soundtrack
Something happening	Flickie Lee Jones	Rock/Pop
This Must Be The Place (Naive Melody)	Talking Heads	New Wave
Songbird	Fleetwood Mac	Rock
Carey	Joni Mitchell	Folk
Can't Find My Way Home	Blind Faith	Rock
Long Time Gone	Crosby, Stills & Nash	Folk
America	Simon & Garfunkel	Folk
Sweet Jane	Cowboy Junkies	Folk
Sugar Magnolia	Grateful Dead	Rock
Tangled Up In Blue	Bob Dylan	Folk
Carry On	Crosby, Stills, Nash & Young	Folk
The Last Time I Saw Richard	Joni Mitchell	Folk
Moonshadow	Cat Stevens	Folk
Wooden Ships	Crosby, Stills & Nash	Folk



## MusicGenius Play-list 2

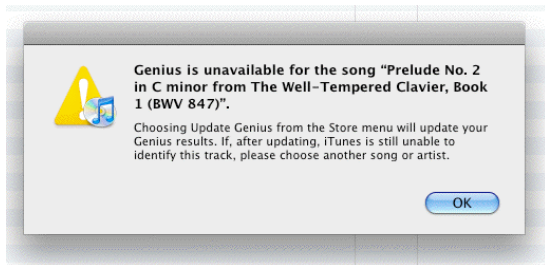
Based On: Rapture by Blondie

	Name	Time	Artist	Genre
	<input checked="" type="checkbox"/> Rapture	3:37	Blondie	New Wave
2	<input checked="" type="checkbox"/> Let's Dance	4:10	David Bowie	Rock
3	<input checked="" type="checkbox"/> Voices Carry	4:21	'Til Tuesday	New Wave
4	<input checked="" type="checkbox"/> Burning Down The House	4:01	Talking Heads	New Wave
5	<input checked="" type="checkbox"/> Head Over Heels	5:02	Tears For Fears	New Wave
6	<input checked="" type="checkbox"/> People Are People	3:52	Depeche Mode	New Wave
7	<input checked="" type="checkbox"/> The One Thing	3:24	INXS	Rock
8	<input checked="" type="checkbox"/> You Might Think	3:05	The Cars	New Wave
9	<input checked="" type="checkbox"/> Call Me	3:33	Blondie	New Wave
10	<input checked="" type="checkbox"/> Take Me To The River	5:03	Talking Heads	New Wave
11	<input checked="" type="checkbox"/> Tempted	3:59	Squeeze	New Wave
12	<input checked="" type="checkbox"/> Chains Of Love	3:44	Erasure	Brit Pop
13	<input checked="" type="checkbox"/> Eminence Front	5:40	The Who	Hard Rock
14	<input checked="" type="checkbox"/> Take The Long Way Home	5:10	Supertramp	Rock
15	<input checked="" type="checkbox"/> Golden Years	3:30	David Bowie	Rock



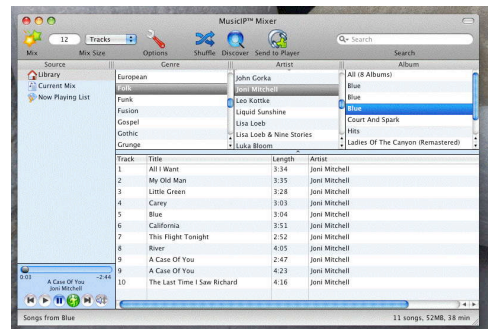
## MusicGenius Play-list 3

- If the FP's not in GraceNote's DB (and this is Bach's WTK...)



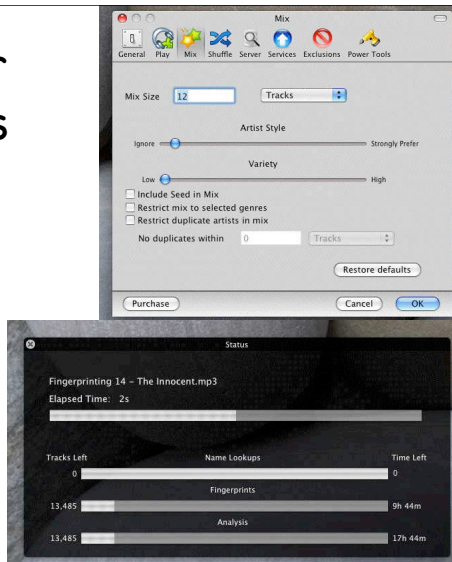
## MusicIP MyDJ Mixer

- Similar to Apple/Gracenote
- Takes middle 2 minutes of song for FP, then matches full FV data

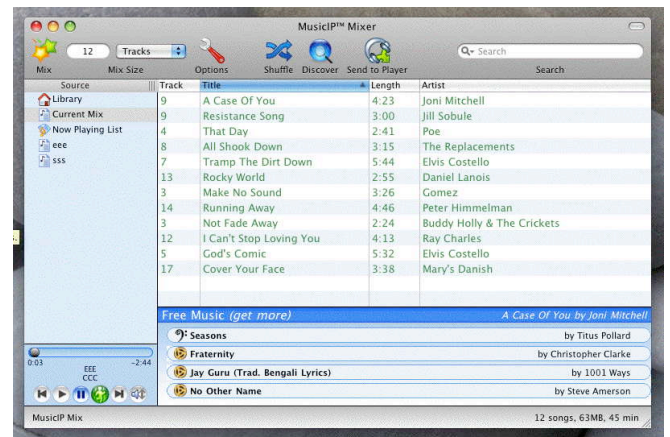


## MyDJ Mixer Prefs/Status

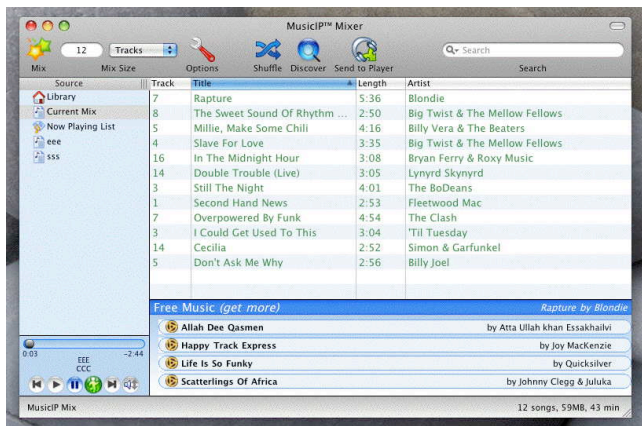
- Batch fingerprinting
- OpenFP server



## MusicIP MyDJ Play-list (15 kSongs)



## MusicIP MyDJ Play-list 2



## QMUL/Isophonics SoundBite

- Interfaces to iTunes DB
- Batch analysis/FP



```
<track tID="4813c80743a3013f">
  <feature ftype="fv_0">-6.62858 8.42714 -2.4534400 4.58270 6.998010
  2.8057200 3.4099400 5.003010 2.346510 5.2242900 5.0869200 -3.3310500
  -8.658210 -2.74836e-05 2.846600 -4.950010 -8.588510 -2.876410 -7.7040400
  9.3793100 6.58215 6.229230 9.675120 8.3485600 6.437400 7.2308300 3.4815300
  2.4267300 8.9972300 5.5670300 2.2319200 3.6605200 5.1908200 3.9213200
  8.8951200 2.757200 1.6446300 9.601400 9.1293300 8.8187200 </feature>
</track>
```

# SoundBite XML Metadata

```
<key>36150</key>
<dict>
  <key>Track ID</key><integer>36150</integer>
  <key>Name</key><string>SlutsFromSpace_1</string>
  <key>Artist</key><string>NatLampoonRadioHour</string>
  <key>Genre</key><string>Comedy</string>
  <key>Kind</key><string>MPEG audio file</string>
  <key>Size</key><integer>28224556</integer>
  <key>Total Time</key><integer>1263516</integer>
  <key>Date Modified</key><date>2005-08-22T21:48:14Z</date>
  <key>Date Added</key><date>2008-06-28T20:52:00Z</date>
  <key>Bit Rate</key><integer>178</integer>
  <key>Sample Rate</key><integer>44100</integer>
  <key>Persistent ID</key><string>4813C80743A32E93</string>
  <key>Track Type</key><string>File</string>
  <key>File Type</key><integer>1297106739</integer>
  <key>File Creator</key><integer>1752133483</integer>
  <key>Location</key><string>file://Volumes/.../SlutsFromSpace_1.mp3</string>
  <key>File Folder Count</key><integer>5</integer>
  <key>Library Folder Count</key><integer>2</integer>
</dict>
```

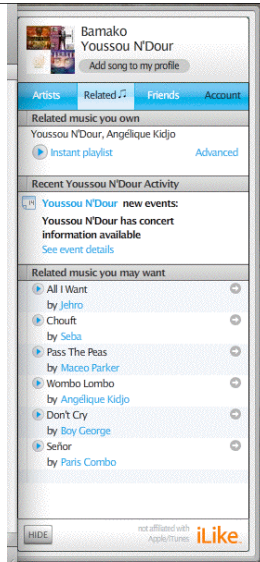
# SoundBite Play-lists

Name	Time	Artist	Genre
1 A Case Of You	4:23	Joni Mitchell	Folk
2 A Case Of You	4:23	Joni Mitchell	Folk
3 A Case Of You	3:20	Joni Mitchell	Audiophile
4 Do Nothing Till You Hear From Me	7:44	Ella Fitzgerald	Jazz
5 EEE	2:48	CCC	AAA
6 April In Paris	6:37	Ella Fitzgerald & Louis Armstrong	Jazz
7 California	3:51	Joni Mitchell	Folk
8 The Nearness Of You	5:43	Ella Fitzgerald & Louis Armstrong	Jazz
9 Autumn In New York	5:58	Ella Fitzgerald & Louis Armstrong	Jazz
10 All Too Soon	3:49	Sarah Vaughan	Big Band S...
11 California	3:51	Joni Mitchell	Folk
12 Stars Fell On Alabama	3:36	Ella Fitzgerald & Louis Armstrong	Jazz

Name	Time	Artist	Genre
1 Rapture	5:37	Blondie	New Wave
2 rapture	5:27	Blondie	Rock/Pop
3 Nightrider	4:23	Electric Light Orchestra	Rock
4 Octopus's Garden	2:51	The Beatles	Pop
5 Red Hill Mining Town	4:54	LJ2	Rock/Pop
6 Red Hill Mining Town	4:53	LJ2	Rock
7 Rapture	5:34	Blondie	Rock/Pop
8 He Don't Live Here Anymore	5:15	Robert Cray	Blues
9 A Holly Jolly Christmas	2:09	The Silver And Gold Singers	Holiday
10 True Love Is Hard To Find	3:39	Bonnie Raitt	Rock
11 Octopus's Garden	2:51	The Beatles	Rock
12 Mother Hold the Candle Steady While I...	11:19	King Crimson	Rock

# iLike

- iTunes add-on drawer
- (Apparently) based on text metadata
- Includes human recommendations, specials, concert info, iLike web site, etc.



# iLike Play-lists

Name	Time	Artist	Genre
1 A Case Of You	3:20	Joni Mitchell	Audiophile
2 No More	2:48	Billie Holiday	Jazz
3 A Case Of You	4:23	Joni Mitchell	Folk
4 The Ghost Of Tom Joad	4:23	Bruce Springsteen	Unclassifiable
5 Them There Eyes	2:51	Billie Holiday	Jazz
6 All Grown Up	4:18	Elvis Costello	New Wave
7 California	3:51	Joni Mitchell	Folk
8 There's No Greater Love	2:59	Billie Holiday	Jazz
9 River	4:05	Joni Mitchell	Folk
10 Carey	3:04	Joni Mitchell	Folk
11 Georgie And Her Rival	3:38	Elvis Costello	New Wave
12 People's Parties	2:15	Joni Mitchell	Folk
13 You're My Thrill	3:24	Billie Holiday	Jazz
14 My Man	2:58	Billie Holiday	Jazz
15 Invasion Hit Parade	5:34	Elvis Costello	New Wave

Name	Time	Artist	Genre
1 Sunday Girl	3:00	Blondie	Rock/Pop
2 Hanging On The Telephone	2:33	Blondie	New Wave
3 Start	1:21	Peter Gabriel	Rock
4 Games Without Frontiers	4:07	Peter Gabriel	Rock
5 Rapture	5:34	Blondie	Rock/Pop
6 Mamma Mia	3:34	ABBA	Soundtrack
7 Heart Of Glass	4:36	Blondie	New Wave
8 The Tide Is High	3:53	Blondie	Rock
9 Call Me	3:33	Blondie	New Wave
10 Heart Of Glass	4:30	Blondie	Rock/Pop
11 One Way Or Another	3:36	Blondie	New Wave
12 No Self Control	3:56	Peter Gabriel	Rock
13 (I'm Always Touched By Your) Presenc...	2:43	Blondie	New Wave
14 union city blues	4:26	Blondie	Rock/Pop
15 Dreaming	3:08	Blondie	New Wave

# EchoNest MusicBrain API

- HTTP requests for search and recommendation
- Early results questionable

**developer nest** API Data Feeds Forums Showcase Your API Account

---

**get\_similar**

Returns similar artists given one or more artists for comparison. The Echo Nest provides up-to-the-minute artist similarity and recommendations from their real-time musical and cultural analysis of what people are saying across the internet and what the music sounds like.

Example: [http://developer.echonest.com/api/v1/similar?api\\_key=52A08B38B047E7&mbid=radiohead&format=json&size=10](http://developer.echonest.com/api/v1/similar?api_key=52A08B38B047E7&mbid=radiohead&format=json&size=10)

**Input Parameters**

Name	Required	Multiple	Data Type	Description
api_key	yes	no	APIKey	Your Echo Nest API key
id	any of id, mbid, or name	yes	Artist ID	The Echo Nest ID for the artist
mbid	any of id, mbid, or name	yes	MusicBrain ID	The MusicBrain ID for the artist
name	any of id, mbid, or name	yes	Artist Name	The name of the artist
size	optional	no	Integer >= 0	The row number to begin with, only available for paid customers
rows	optional	no	Integer >= 11	The number of results to return

**Response**

The response format is described by a RELAX-NG schema. The schema specific to this method's response describes the `similar` element.

**Example Response**

```
<?xml version="1.0" encoding="UTF-8"?>
<response version="2">
  <status>
    <code>200</code>
    <message>Success</message>
  </status>
  <query>
    <parameter name="api_key">52A08B38B047E7</parameter>
    <parameter name="mbid">radiohead</parameter>
    <parameter name="format">json</parameter>
    <parameter name="size">10</parameter>
    <parameter name="size">10</parameter>
    <parameter name="size">10</parameter>
    <parameter name="size">10</parameter>
  </query>
  <similar>
```

# SoLaTi System

- FASTLab, Inc + Catalyst
  - Based on FMAK 4.2 analysis kernel
- Assume only audio-derived metadata
  - To be augmented with other sources in Rev 2
- FV Statistics
  - Aggressive smoothing, histograms, GMM
  - Store mean and variance FVs for "typical" and "solo" verses (or mean/var for song)
- Multiple similarity metrics
  - Configurable/PCA/FV weighting
  - Euclidean/Earth-mover's/Mahalanobis distance

## Example SoLaTi Play-list

Key: Joni Mitchell -- A Case Of You -- Folk

(Song set 1)

98:	Stephen Lynch	--	Walken III	--	Comedy
102:	Joni Mitchell	--	California	--	Folk
106:	David Sanborn	--	Carly_s Song	--	Jazz/Cool
107:	Mazzy Star	--	Wasted	--	Rock/Alternative
110:	The Art Of Noise	--	Opus 4	--	Electronic
114:	Joni Mitchell	--	California	--	Folk
114:	Billy Joel	--	Just The Way You Are	--	Pop
117:	Bonnie Raitt	--	Have A Heart	--	Rock
119:	Harry Connick, Jr	--	It Had To Be You	--	Jazz
120:	Crosby, Stills, Nash	--	Teach Your Children	--	Folk
122:	unknown	--	Veinte Anos	--	Soundtrack
123:	Queen	--	Body Language	--	Rock/Hard Rock

ASST...  
Lab

79

## Example SoLaTi Play-list 2

Key: Joni Mitchell -- A Case Of You -- Folk

(Song set 2, semi-PCA-weighted EMD)

68:	Maurizio Pollini	--	Sonata No.13 II. All	--	Classical/Symphony
70:	Bing Crosby Frank Si	--	White Christmas	--	Religious/Holiday Music
76:	Antonio Vivaldi	--	Con In C, RV 537- 2.	--	Classical/Symphony
78:	Stevie Ray Vaughan	--	LENNY	--	Blues
79:	Steely Dan	--	Black Cow	--	Rock
79:	David Sanborn	--	Carly_s Song	--	Jazz/Cool
80:	Various Artists	--	La Forza Del Destino	--	Classical/Symphony
81:	Bonnie Raitt	--	Have A Heart	--	Rock
82:	Mark and Brian	--	James Woods Advance	--	Comedy
83:	The Fabulous Baker B	--	Makin_ whoopee	--	Soundtrack
84:	unknown	--	Stupid Halloween_Po	--	Soundtrack
84:	Dead Can Dance	--	The Wind That Shakes	--	Rock/Alternative

ASST...  
Lab

80

## Example SoLaTi Play-list 3

Joni Mitchell -- River

38:	Daniel Johnston	--	Love Defined	--	Rock/Alternative
40:	unknown	--	Part Of Your World	--	Children
41:	Robert Drake & Ruben	--	Berimbau	--	Classical/Symphony
41:	unknown	--	Claire_s Theme	--	Soundtrack
41:	Everything But The G	--	The Road	--	Rock/Alternative
42:	The Bobs	--	Cowboy Lips	--	Vocal
43:	10,000 Maniacs	--	Verdi Cries	--	Rock/Alternative
45:	Spencer the Gardener	--	LuLu Interlude	--	Rock
46:	Susan Voelz	--	Red River Waltz Over	--	Rock
49:	The Beatles	--	You_ll Be Mine	--	Rock
50:	Billy Joel	--	Where_s The Orchestr	--	Pop
50:	Cowboy Junkies	--	Powderfinger	--	Folk

ASST...  
Lab

81

## Example SoLaTi Play-list 4

Blondie -- Rapture

53:	Talking Heads	--	Once In A Lifetime	--	Rock/New Wave
56:	Roxy Music	--	The Space Between	--	Rock
63:	Ben Harper	--	Homeless Child	--	Rock/Alternative
63:	Alison Krauss & Union	--	It Won_t Work This T	--	Country/Bluegrass
70:	August Campbell And	--	The I-95 Song	--	Country
73:	The Klezmatiks	--	Clarinet Yontev	--	Religious
74:	unknown	--	Ev_rybody Has A Laug	--	Children
75:	Daniel Johnston	--	I Remember Painfully	--	Rock/Alternative
75:	They Might Be Giants	--	Whistling In The Dar	--	Rock/Alternative
80:	Professor Michael DC	--		--	Vocal
84:	Hootie & The Blowfis	--	Fairweather Johnson	--	Rock/Alternative
84:	The Art Of Noise	--	Kiss (Featuring Tom	--	Electronic

ASST...  
Lab

82

## Example SoLaTi Play-list 5

Metallica -- The House Jack Built

43:	The Breeders	--	S.O.S	--	Rock/Alternative
46:	The Beatles	--	Carry That Weight	--	Rock
48:	Denzil	--	Bastard Son Of Elvis	--	Rock
48:	JET	--	L_esprit D_escalier	--	Rock
49:	Paul McCartney & Win	--	The Lovely Linda	--	Rock
51:	Smashing Pumpkins	--	Sweet Sweet	--	Rock/Alternative
52:	Sky Cries Mary	--	Joey_s Aria	--	Rock/Alternative
52:	Sublime	--	What I Got (Reprise)	--	Rock/Alternative
57:	The Breeders	--	Roi (Reprise)	--	Rock/Alternative
59:	Oasis	--	(Untitled)	--	Rock
61:	Spinal Tap	--	Cups And Cakes	--	Rock
62:	KT Tunstall	--	Another Place To Fal	--	Pop

ASST...  
Lab

83

## Example SoLaTi Play-list 6

Glenn Gould -- WTC/I Prelude No. 2 in C m

30:	Andrés Segovia	--	Pavana No. 3	--	Classical
41:	Glenn Gould	--	Variations XXVI and	--	Classical/Symphony
48:	Händel; Pinnock	--	Accompagnato (Bass)-	--	Classical/Symphony
59:	Andrés Segovia	--	Pavana No. 6	--	Classical
63:	Andrés Segovia	--	Prelude	--	Classical
64:	unknown	--	Raindrop Prelude	--	Soundtrack
64:	James Edwards	--	Plaisir D_Amour-Gi	--	Classical/Symphony
66:	unknown	--	Raindrop Reprise	--	Soundtrack
67:	Bing Crosby Frank Si	--	O Little Town Of Bet	--	Religious/Holiday
70:	Spencer the Gardener	--	LuLu Interlude	--	Rock
70:	Hootie & The Blowfis	--	Sometimes I Feel	--	Rock/Alternative
72:	Glenn Gould	--	Prelude in D minor,	--	Classical/Symphony

ASST...  
Lab

84

## See Also

- Celma & Lamere ISMIR Tutorial: Evaluating Recommender Systems
  - Accuracy
  - Relevance
  - Coverage
  - **\*\*Quality of recommendation list\*\***
  - The play-list Turing test

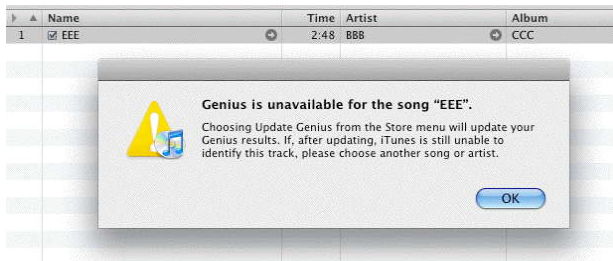
## Song Degrading/Spoofing

- To break finger-print
  - Or detect bootleg or in-stream data
  - Related to cover song ID
- Degrاداتions
  - Clip/extend start/finish
  - Scale, compress, mono
  - Transpose, change rate
  - Tape, FM hiss
  - Remove/scramble text metadata



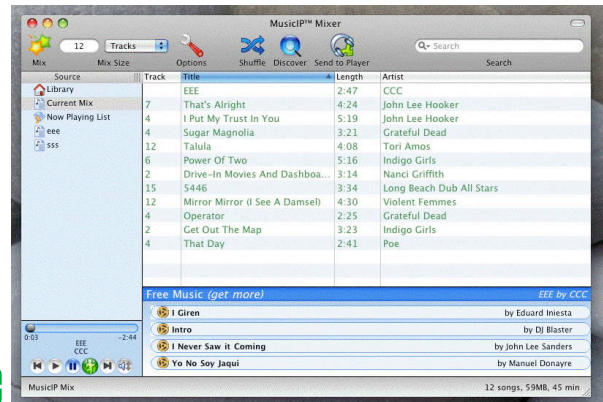
## MusicGenius on "A Case of You2"

- No exact finger-print
- iLike also fails



## MusicIP MyDJ on EEE

- Lost?



## SndBite on EEE

- FP found (list still strange)

Name	Time	Artist	Genre
1 EEE	2:48	CCC	AAA
2 A Case of You	4:23	Joni Mitchell	Folk
3 A Case Of You	4:23	Joni Mitchell	Folk
4 Do Nothing Till You Hear From Me	7:44	Ella Fitzgerald	Jazz
5 A Case of You	3:20	Joni Mitchell	Audiophile
6 Crazy	2:44	Patsy Cline	Pop
7 Day-Dream	4:01	Ella Fitzgerald	Jazz
8 Little Green	3:28	Joni Mitchell	Folk
9 The Nearness Of You	5:43	Ella Fitzgerald & Louis Arm...	Jazz
10 A Foggy Day	4:34	Ella Fitzgerald & Louis Arm...	Jazz
11 Little Green	3:28	Joni Mitchell	Folk
12 Wish I Could	4:18	Norah Jones	Jazz

## SoLaTi on "A Case of You2" (eee)

- Dist = 0.18301 (i.e., very close)

Matches for: CCC -- eee -- AAA/BBB

13: Spencer the Gardener	--	LuLu Outro	--	Rock
19: Hootie & The Blowfis	--	Sometimes I Feel	--	Rock/Alternative
20: Robert Drake & Ruben	--	Berimbau	--	Classical/Symphony
22: Ben Folds Five	--	CIGARETTE	--	Rock
22: Billie Holiday	--	Don_t Explain	--	Jazz
22: Everything But The G	--	Walking To You	--	Rock/Alternative
22: Cory Sipper	--	Strange World	--	New Age
22: Crosby, Stills & Nas	--	Haven_t We Lost Enou	--	Folk
23: Billie Holiday	--	Crazy He Calls Me	--	Jazz
24: Patsy Cline	--	Lonely Street	--	Country
25: Jesse Colin Young	--	Morning Sun	--	Folk
26: Elvis Costello	--	Shipbuilding	--	Rock/New Wave

## Next Steps

- Better analysis/processing
  - Chroma features
  - Segmentation
- Genre cloud & statistics
- Data-mining
- Add web play-list co-occurrence
- Generate ID key

## The Future of MIR

- Analysis & FV design
  - Spectral, harmonic, melodic, rhythmic features
  - Mood, emotion, genre
- Music DB apps
  - Play-list recommenders/personal radio
  - QbH song ID
    - Noisy source
    - Streaming source
- New MIREX tests

## For More Info

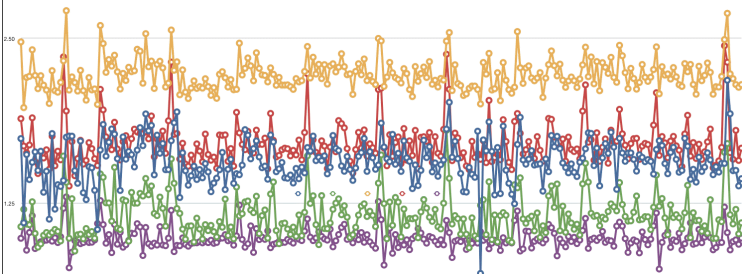
- Proc. ISMIR, ICMC, AES
- Celma/Lamere IRMIR 2007 tutorial
- MAT 240F web site
- Open-source FMAK 3

## Summary

- Context
- MMDB background
- Music information retrieval (MIR)
- Audio signal analysis for MIR
- Statistical data processing
- Database techniques
- Application evaluation

## Thank You

- Q&A



Stephen Travis Pope  
stephen@FASTLabInc.com