

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

Presentation by Stephen Travis Pope
stephen@FASTLabInc.com -- Oct 2008
Updated Mar 2014



1

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 6th-generation SndsLike] and several products will be compared in terms of the play lists they recommend for given input songs.



2

Music/Sound Database Projects

- ARA/DoubleTalk/HyperScore/MODE/Siren (1980–present)
 - Composer's tools: metadata, persistency, data-mining
 - Paleo (1996–9) MIDI performance expression data-mining
 - NOLib (1998–9) Feature extraction framework in MATLAB
 - 8S Speech segmenter & database in Smalltalk (comps)
 - FASTLab MusicAnalysisKernel (MAK) 1 (1999–2003)
 - MusicMagic, MusicIP, LibOFA, AmpliFind, GraceNote
 - OMNI/LoCAA Network-based access, recommender (2001)
 - FASTLab 2: Expert Mastering Assistant (EMA) (2002–4)
 - FASTLab 3: Locus animation system (MUGI) (2006–7)
 - FASTLab 4: SoLaTi recommender (Catalyst) (2007–8)
 - FASTLab 5: Imagine Research/iZotope (SndObjRec) (2008–11)
 - SndsLike & PlayListMgr (2012–3)



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 (for search engines) has been a central topic in both image processing and music information retrieval for over a decade.



1

Overview

- Introduction
 - MMDB background
 - Feature Extraction & Processing
 - Segmenting and Seg-derived Features
 - Dimensionality-reduction and Mapping
 - Examples
 - SoLaTi (2007)
 - SndLike (2012)



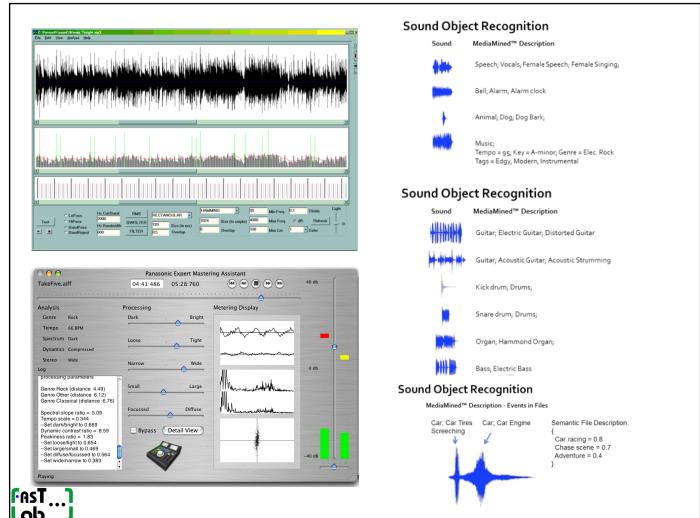
1

Music/Sound Database Projects

- ARA/DoubleTalk/HyperScore/MODE/Siren (1980–present)
 - Composer's tools: metadata, persistency, data-mining
 - Paleo (1996–9) MIDI performance expression data-mining
 - NOLib (1998–9) Feature extraction framework in MATLAB
 - 8S Speech segmenter & database in Smalltalk (comps)
 - FASTLab MusicAnalysisKernel (MAK) 1 (1999–2003)
 - MusicMagic, MusicIP, LibOFA, AmpliFind, GraceNote
 - OMNI/LoCAA Network-based access, recommender (2001)
 - FASTLab 2: Expert Mastering Assistant (EMA) (2002–4)
 - FASTLab 3: Locus animation system (MUGI) (2006–7)
 - FASTLab 4: SoLaTi recommender (Catalyst) (2007–8)
 - FASTLab 5: Imagine Research/iZotope (SndObjRec) (2008–11)
 - SndsLike & PlayListMgr (2012–3)



1



Content Analysis for MMDBs

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

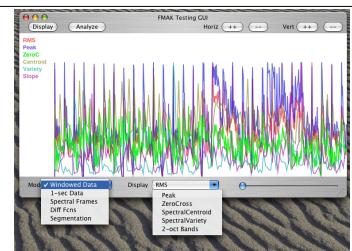
FasT...Lab

7

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

FasT...Lab



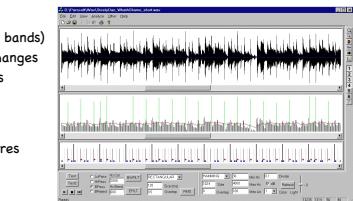
8

Audio Feature Extraction 1

First-pass (windowed) analysis

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

FasT...Lab

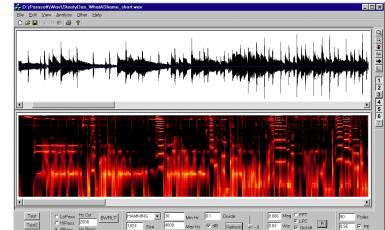


9

Audio Feature Extraction 2

- 2nd-pass analysis
 - Smoothing, pruning, 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
- Multi-pass: stage-configs and confidences

FasT...Lab



10

Spectral Tracker Configurations

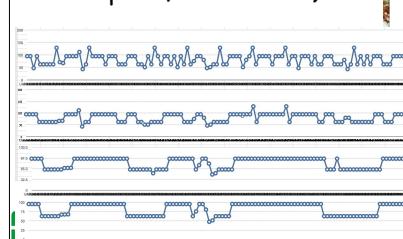
```
# Each entry consists of a line with 4 data values:
#   peakWidth    - closeness measure: peaks this close are considered to be one
#   minPeakAmpMeanClearanceRatio - the amplitude mean clearance ratio is defined
#                               as the ratio of a peaks amplitude to the mean amplitude 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 }
```

GRINNELL

11

Data Smoothing Examples

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



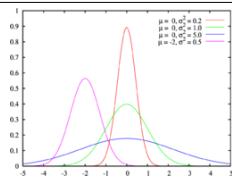
12



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

FastLab

13

Music Segmentation

- Detect onsets
- Find regular hierarchy of onsets
- Segment track into verses
- Detect intro/outro
- Detect “solo” verse or bridge
- Calculate segmentation-related features (excellent genre/style correlation)

FastLab

15

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

FastLab

17

Example (Rich) Feature “Vector”

Time-domain features

- Windowed RMS amplitude
- Max sample amplitude
- RMS (ratio?) of LP/HF-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

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

CREATE

14

PIRT

Audio Segmentation

Basic (time-domain) procedure

- Pick a feature vector weighting
- Calculate inter-window distances (scalar)
- Identify regular peak spacing

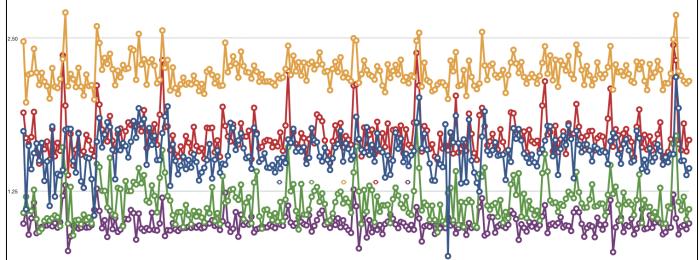
Challenges

- 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)

FastLab

16

4 Song Segmenter Distance Weightings



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

FastLab

18

Configurable Segmenter

```

# Segmenter Configurations
#
# Each block consists of a list of distance-
# metric weighting maps keyed by feature
#
# Spectral/pitch-centric configuration

SegmenterConfiguration {
    HPRMS 0.2
    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

```

GRiANTE

19

MIR Media Arts and Technology Graduate Program

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	solostart	
	0.631119	0.280232	0.590672	
	s_centroid	s_variety	s_tempo	s_dynrange
	0.4991	0.001422	0.3360	0.654455

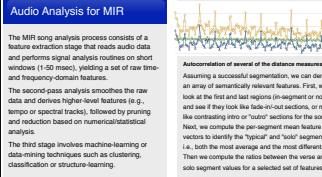
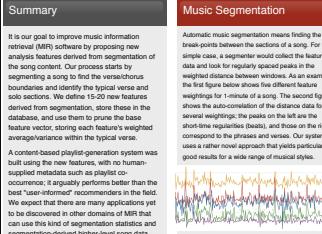
GRiANTE

21

MIR Media Arts and Technology Graduate Program

Improving Music Information Retrieval using Segmentation

Stephen Travis Pope
 UCSB Graduate Program in Media Arts and Technology, UCSB Dept. of Music



23

MIR Media Arts and Technology Graduate Program

CREATE

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?

Fast... Lab

20

Song/Segmentation Features

General song metadata

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

Segmentation features

- Segmentation confidence
- NumSegments
- VerseLength

Tempo features

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

FirstVerseStart

- SoloIndex

Beat histogram features

- BHSUM1
- BHSUM2
- BHSUM3
- Low/HipPeakAmp
- Low/HipPeakBPM
- HighLowRatio

QuietSections

- LoudSections
- Fadeln
- FadeOut

Fluctuation-pattern measures

- FP_gravity
- FP_bass
- FP_focus
- SoloCentroid
- SoloVariety
- SoloDynRange
- SoloRMS
- SoloTempo

MIR Media Arts and Technology Graduate Program

22

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)

GRiANTE

24

MIR Media Arts and Technology Graduate Program

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?

FAST
Lab...

25

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
[...]								
MFCCFirst	min:	-118.78	max:	169.699	avg:	56.441	var:	2818.36
								dev: 53.0883

CREATE

MRT Media and Technology

26

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

FAST
Lab...

27

1st 2 PCA Dimensions

```

0.18 MFCCCoeff6 + 0.18 MFCCCoeff5 + 0.18 - 0.378 BHSUM3 - 0.345 LowPeakAmp -
MFCCCoeff4 + 0.18 MFCCCoeff3 + 0.18 0.323 BHSUM1 - 0.309 BHSUM2 - 0.298
MFCCCoeff2 + 0.18 SpectralFlux + 0.18 HighPeakAmp - 0.294 fp_bass - 0.269
SpectralRolloff + 0.18 SpectralFluxVar + 0.18 ZeroCrossingsVar - 0.267 ZeroCrossings -
SpectralSlopeVar + 0.18 SpectralRolloffVar + 0.237 HPRMS + 0.161 TempoWeight + 0.161
0.18 SpectralSlope + 0.18 MFCCCoeff6Var + TempoAvg - 0.15 fp_gravity - 0.12 LPRMSVar
+ 0.14 HighPeakBPM + 0.114fp_focus + 0.093
0.18 MFCCCoeff5Var + 0.18 MFCCCoeff4Var +
0.18 MFCCCoeff3Var + 0.18 MFCCCoeff2Var +
0.179 SpectralBand2Var + 0.179 SpectralVarietyVar + 0.054 QuietSections +
SpectralBand1Var + 0.179 STrackBirthsVar +
0.179 STrackDeathsVar + 0.179 SpectralBand4Var +
0.179 SpectralBand3 + 0.179 SpectralBand2 SoloTempo

```

CREATE

MRT Media and Technology

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

CREATE

MRT Media and Technology

*ART-Tree-learning

```

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)

```

CREATE

MRT Media and Technology

30

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

FAST...
Lab

37

Music Recommender Systems

- (selected, in approx. order of release)
- MusicIP MyDJ (FMAK0++)
- QMUL SoundBite
- MIT/EchoNest MusicBrain/API
- FMAK/SoLaTi
- iLike
- Apple/Gracenote MusicGenius
- MS Zune 3.0

FAST...
Lab

38

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

FAST...
Lab

39

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

FAST...
Lab

40

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 Klezmatics -- 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 -- 3a -- Vocal
84: Hootie & The Blowfish -- Fairweather Johnson -- Rock/Alternative
84: The Art Of Noise -- Kiss (Featuring Tom -- Electronic

FAST...
Lab

41

SndsLike

42

SndsLike Goals

- Similarity-based "recommender" system aimed at production music data sets (why?)
 - Written from scratch (I sold the old code to iZotope)
 - Use the “latest features” (> 400 features)
 - Use the “latest statistics” (sophisticated de-noising)
 - Use the “latest distance metrics” (learned)
 - Use existing noisy/partial labels to train clustering, labels and distance metrics
 - Simple, fast, portable, embeddable
 - Python + C++, octave, java, MySQL

FAST...
Lab

43

The “Latest Features”

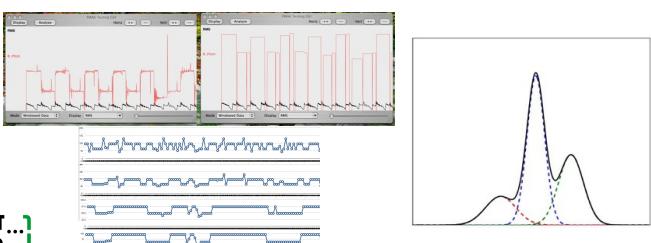
- Standard time- and freq-domain features
 - HPF/LPF versions
 - Many freq bands
 - Chroma, harmonicity, MFCCs & spectral measures
 - Sp-slope, spread, bandwidth, variety, kurtosis, roll-off...
 - Spectral tracking and track birth/death stats (useful)
 - Fluctuation pattern features (E Pampalk)
 - Beat histograms (G Tzanetakis)
 - Statistical Spectrum Descriptors (Lidy & Rauber)
 - Several tempo estimates (BH + stats)
 - Several bass pitch estimates (+ stats) + tracking
 - Several chord/key pitch estimates (+ stats)
 - Musical segmentation and segment-related features

FAST... Lab

49

The Latest Statistics

- Lots of feature-dependent smoothing
 - Data mode: noisy, bi-modal, clicky, etc.
 - Take Gaussian Mixture Models (GMM) of all features
 - Save gmm-avg, main-lobe width/weight, bi-modality...
 - Also save dev, del, del2



1

SndsLike “Demo”

Vivaldi

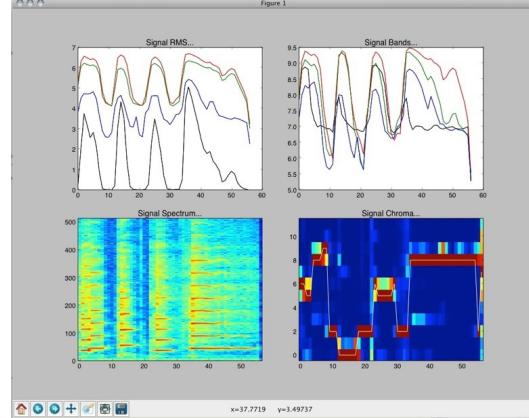
FASTLab SndList 5.0 Player						FASTLab SndList 5.0 Player					
ID	Title	Artist	Album	Genre	Duration	ID	Title	Artist	Album	Genre	Duration
108	Shout It Out Loud	Green Day	Insomniac	Alternative	4:04	108	Shout It Out Loud	Green Day	Insomniac	Alternative	4:04
118	Stairway to Heaven	The Who	Tommy	Highway Rock	7:07	118	Stairway to Heaven	The Who	Tommy	Highway Rock	7:07
119	Smells Like Teen Spirit	Nevermind	Nevermind	Alternative	3:07	119	Smells Like Teen Spirit	Nevermind	Nevermind	Alternative	3:07
120	Smells Like Teen Spirit	Nevermind	Nevermind	Alternative	3:07	120	Smells Like Teen Spirit	Nevermind	Nevermind	Alternative	3:07
121	Forget You	Ray Charles	Leavin' On The Bus	Alternative	5:22	121	Forget You	Ray Charles	Leavin' On The Bus	Alternative	5:22
122	Hotel California	Hotel California	Hotel California	Rock	4:03	122	Hotel California	Hotel California	Hotel California	Rock	4:03
123	Wishbone Salad	Green Day	Insomniac	Alternative	2:13	123	Wishbone Salad	Green Day	Insomniac	Alternative	2:13
124	The Young & the Hopeless	Good Charlotte	Good Charlotte	Alternative	3:36	124	The Young & the Hopeless	Good Charlotte	Good Charlotte	Alternative	3:36
125	Smash Mouth	Smash Mouth	Smash Mouth	Pop	3:37	125	Smash Mouth	Smash Mouth	Smash Mouth	Pop	3:37
126	21st Century Breakdown	Green Day	Insomniac	Alternative	1:50	126	21st Century Breakdown	Green Day	Insomniac	Alternative	1:50
127	Chemical Smile	Blink-182	Spanish Fly	Alternative	3:07	127	Chemical Smile	Blink-182	Spanish Fly	Alternative	3:07
128	Moanin'	Bill Withers	The Young & the Hopeless	Alternative	3:37	128	Moanin'	Bill Withers	The Young & the Hopeless	Alternative	3:37
129	Eye of the Tiger	Eye of the Tiger	Eye of the Tiger	Pop	3:37	129	Eye of the Tiger	Eye of the Tiger	Eye of the Tiger	Pop	3:37
130	Break It Up	Green Day	Insomniac	Alternative	2:16	130	Break It Up	Green Day	Insomniac	Alternative	2:16
131	Eye of the Tiger	Eye of the Tiger	Eye of the Tiger	Pop	3:37	131	Eye of the Tiger	Eye of the Tiger	Eye of the Tiger	Pop	3:37
132	But You're Good On The Dance Floor	Arctic Monkeys	Whatever People Say... Is True To You	Pop	3:34	132	But You're Good On The Dance Floor	Arctic Monkeys	Whatever People Say... Is True To You	Pop	3:34
133	Don't You Worry Bout a Thing	Donna Summer	Donna Summer	Pop	3:57	133	Don't You Worry Bout a Thing	Donna Summer	Donna Summer	Pop	3:57
134	Don't Cry Baby	Donna Summer	Donna Summer	Pop	3:57	134	Don't Cry Baby	Donna Summer	Donna Summer	Pop	3:57
135	Far Out	OutKast	Daydream	High	4:03	135	Far Out	OutKast	Daydream	High	4:03
136	Midnight Drive	OutKast	Daydream	High	4:03	136	Midnight Drive	OutKast	Daydream	High	4:03
137	Don't Cry Baby	OutKast	Daydream	High	4:03	137	Don't Cry Baby	OutKast	Daydream	High	4:03
138	Tragedy	My Chemical Romance	The Tragedy	Rock	2:07	138	Tragedy	My Chemical Romance	The Tragedy	Rock	2:07
139	My War	My Chemical Romance	The Tragedy	Rock	2:07	139	My War	My Chemical Romance	The Tragedy	Rock	2:07

FastLab

Artist	Title	Date
0	You've Got A Friend	1969
12	Walking Man	1970
172	The Ballad Of Uncle Charlie	1971
191	Communicator	1971
207	Dreamin'	1972
210	Rocky Mountain Way	1972
217	Believe Me	1973
218	Versatile	1973
230	If I Loved You (Live)	1974
231	The Crane	1974
234	Sitter On Top Of The World	1974
236	Don't Let Me Be Lonely Tonight	1974
241	Iron Hand	1975
242	Rockin' Around The World	1975
248	Same Thing	1976
250	The Croup	1976
252	Speech Of Angels	1976
254	Winter Wheat	1977
255	Group Life	1977

James Taylor, different song				
Title	Artist	Album	Genre	Duration
Sweet Baby James	James Taylor	James Taylor	Folk	4:02
Long Time Running	The Tragically Hip	Broad Asiles	Rock	4:22
U2's A Lesson	Shawn Colvin	U2, Me, & Life's A Lesson	Religious	3:41
Gaudy Luck Charm	Elliott Murphy	One More Nine O'Clock	Folk	3:54
Fernent Blue	Chris Isaak	Forever Blue	Folk	2:43
Don't Think Twice	Dixie Chicks	Home	Folk	3:20
Fat Boy	Tom Waits	Spunk	Folk	3:00
Playin' Possum	Tom Waits	Under the Volcano	Rock	4:00
It's All Begun	Elli Fitzgerald	The Best Of What Happens	Jazz	3:03
The Lee Shore	Crosby, Stills & Nash	Crosby Stills & Nash (2x)	Folk	5:33
Saints Of The Sun	Paul Simon	From The Beginning	Afrobeat	3:56
Spirit Voices	Paul Simon	Mystrophe II	Folk	3:56
Heavy Pet	The Beatles	The Beatles (10th Anniversary) (10x)	Folk	2:41

Feature Extractor Development



4

1

The latest distance metrics

- Using noisy labels
- Dimensionality reduction vs clustering
 - PCA
 - SVMs
 - CURE
 - **FLDA**
- FLDA training and clusterer app
- Train on a couple dozen well-known genres

FASTLab...

49

SndsLike Development Process

- Smalltalk prototype in Siren
- Analysis core in C++: RMS & FFT features
- Wrapper in Python
- Call-outs to Java (SSD) and Octave (FP) code
- Higher-level features
 - Rhythm, key, bass line, SSDs, etc.
- Simple tests
 - Feature extraction
 - DB populate batches

FASTLab...

50

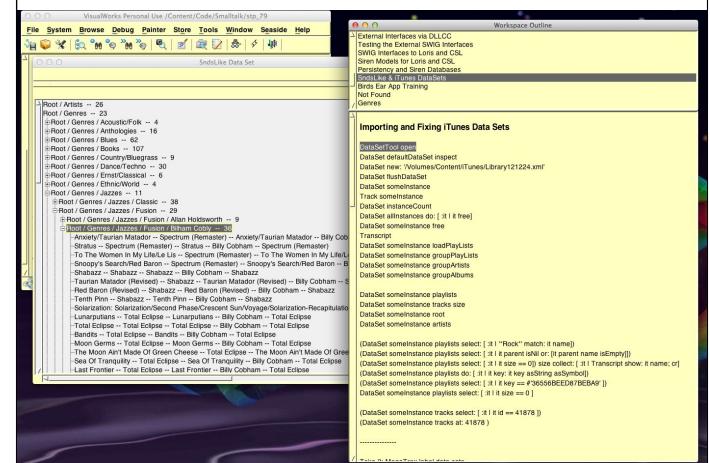
Data Sets

- FASTLab - 14 kSongs, very diverse, “high-quality,” well-encoded
- LikeZebra - 250 kSongs pop/rock
- MegaTrax - 160 kSongs + stems
- AudioNet - 200 kSongs + stems
- Others (not public)

FASTLab...

51

Smalltalk Tools for SndsLike



52

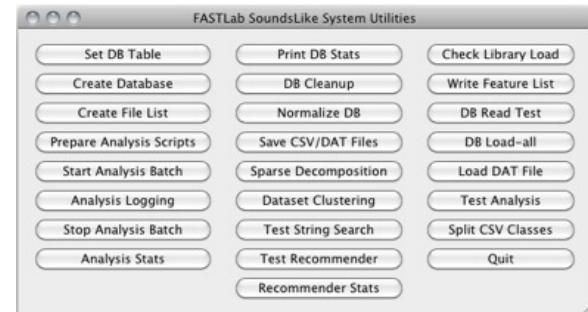
Testing/Demo GUIs

- Test GUI button panel
 - In Python & Qt
- Various data plots
 - Several tools: gnuplot, XL, etc.
- MySQL tests
- Demo "Player" GUI
 - C++/JUCE

FASTLab...

53

OPs/Test GUI



54

