

Inferring Descriptions and Similarity for Music from Community Metadata

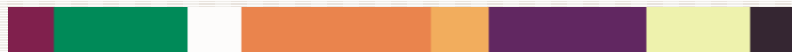
Brian Whitman

MIT Media Lab

Music, Mind and Machine Group
(formerly Machine Listening)

Steve Lawrence

NEC Research Institute

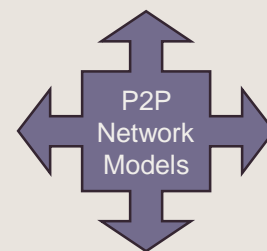


Audio and Audience (1)

Where does music preference come from?

Does the type of music actually matter?

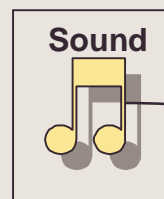
Can an automatic system predict your tastes? Your friends' tastes?



- Daily 'Top 40' for peer-to-peer networks (Napster/Gnutella/etc)
- User models, trend ID

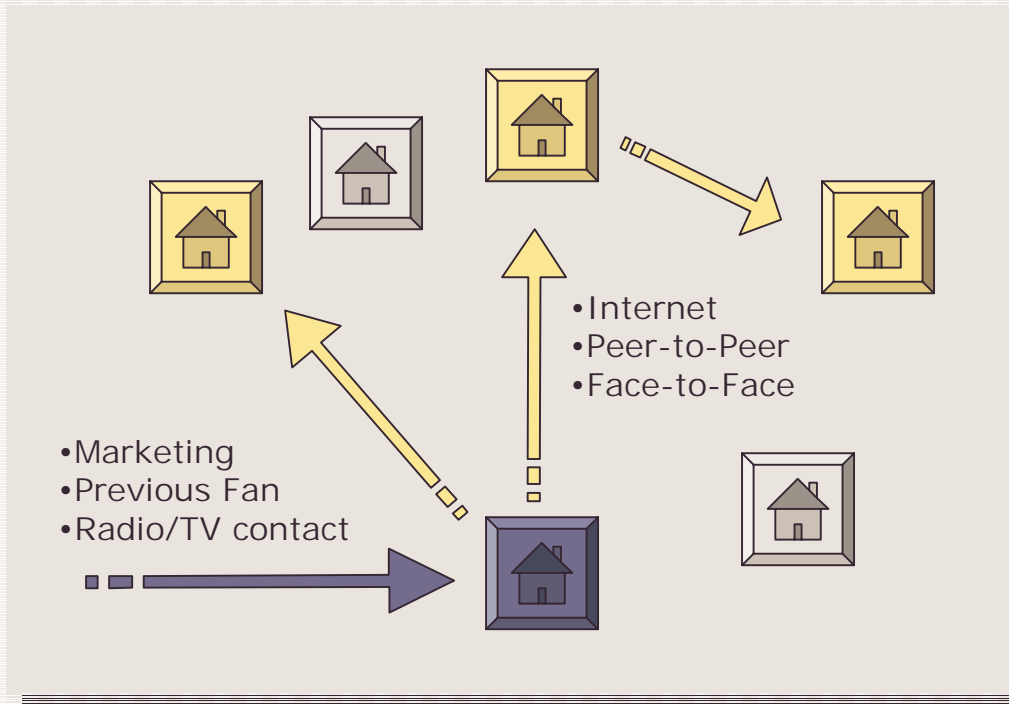


- Automatic music description ("cultural representation")
- Query-by-description
- Time-aware recommendation ('buzz factor' extraction)



- Content-based representation
- Feature extraction (beat, instrument types)

Audio and Audience (2)



How does word get out?

Can we find the 'trendsetter'?

Can we predict future hits? The scope? The path?

Acoustic vs. Cultural Representations

■ Acoustic:

- Instrumentation
- Short-time (timbral)
- Mid-time (structural)
- Usually all we have



Which genre?
Which artist?
What instruments?

■ Cultural:

- Long-scale time
- Inherent user model
- Listener's perspective
- Two-way IR



Describe this.
Do I like this?
10 years ago?
Which style?



Representation Uses

■ Acoustic:

- Artist ID
- Genre ID
- Audio similarity
- Copyright Protection

■ Cultural:

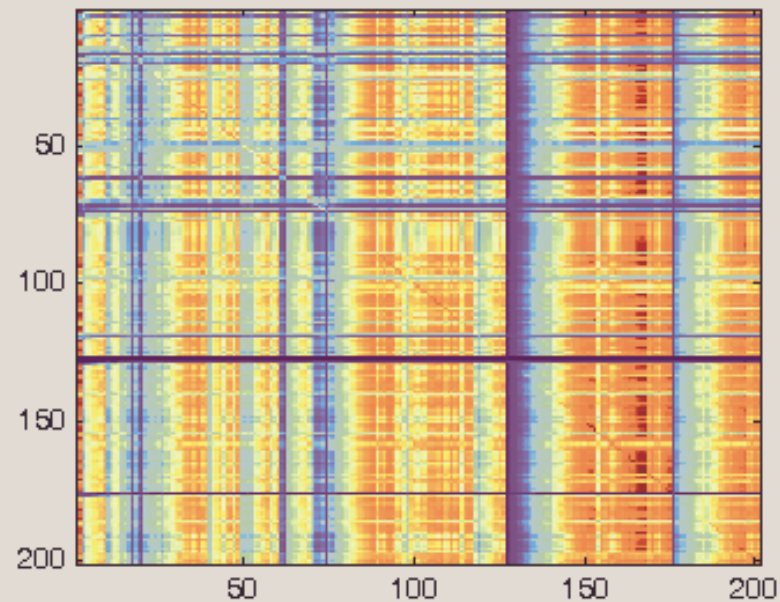
- Style ID
- Recommendation
- Query by Description
- Cultural similarity

Combined:

Auto description of music
Community Synthesis
High-accuracy style ID
Acoustic/Cultural User Knob

“Community Metadata”

- Combine all types of mined data
 - P2P, web, usenet, future?
- Long-term time aware
- One comparable representation via gaussian kernel
 - Machine learning friendly



What's On

- Explain CM as cultural representation
- Evaluate model
- CM in action:
 - Recommendation / Buzz Factor Extraction
 - Query by Description
 - Style ID
 - Community Synthesis

Artists, History

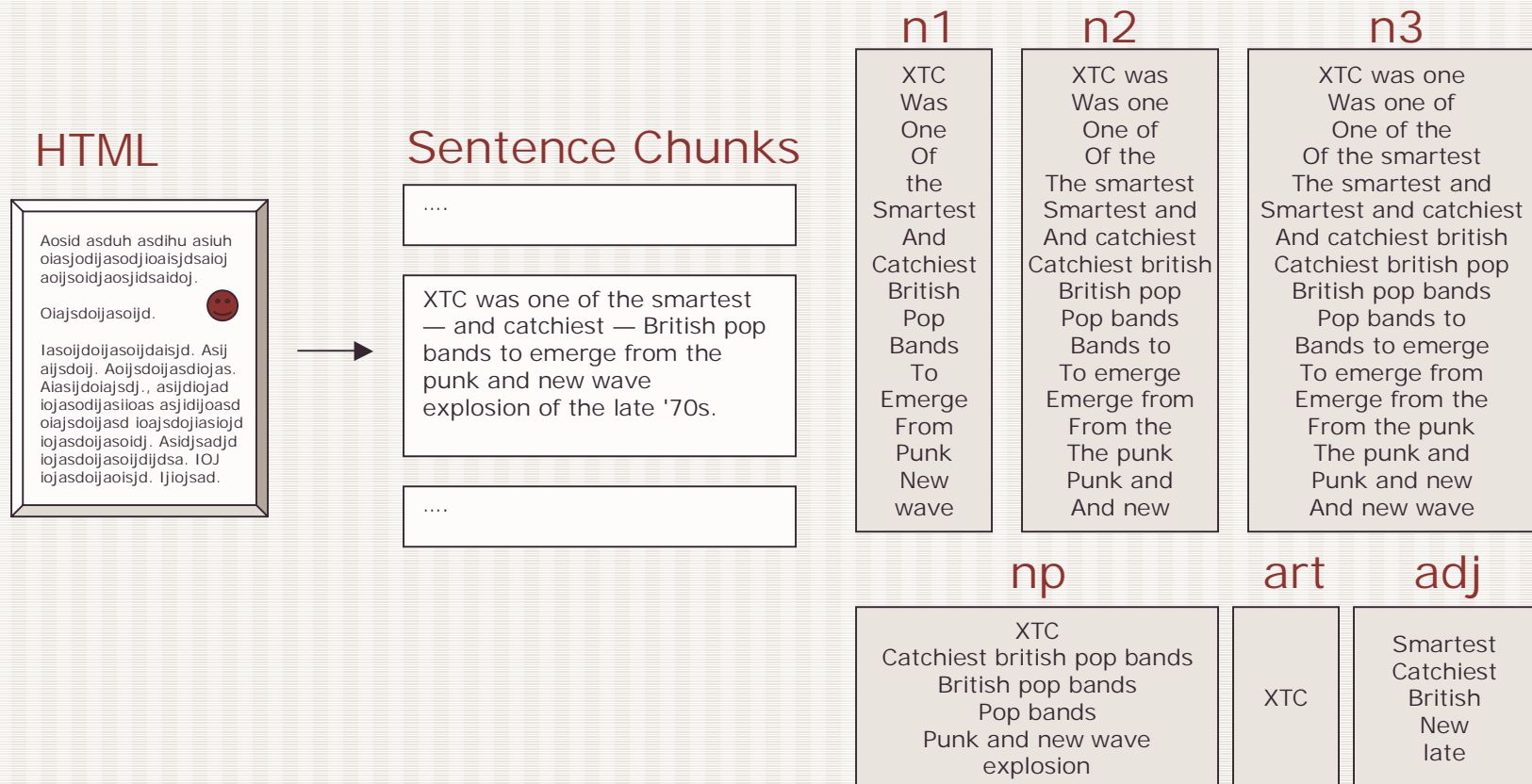
- NEC Minnowmatch Testbed
 - Audio+CM
 - 400 artists, 1000 albums
 - Top albums on OpenNap, August '01
 - Testbed content used elsewhere:
 - Whitman/Flake/Lawrence artist ID
 - Ellis/Berenzweig vocal artist ID
 - Ellis/Whitman/Lawrence/Berenzweig artist survey
 - Kim/Whitman vocal artist ID
 - Whitman/Smaragdis style ID
 - Whitman/Rifkin description generation
 - Mostly pop&rock, some jazz/classical

Data Collection Overview

- Web/Usenet crawl:
 - Crawls for music information
 - Retrieved documents are parsed for language constructs
- P2P crawl:
 - Robots watch OpenNap et al networks for shared songs on collections.

Language Processing for IR

■ Web page to feature vector



What's a Noun Phrase?

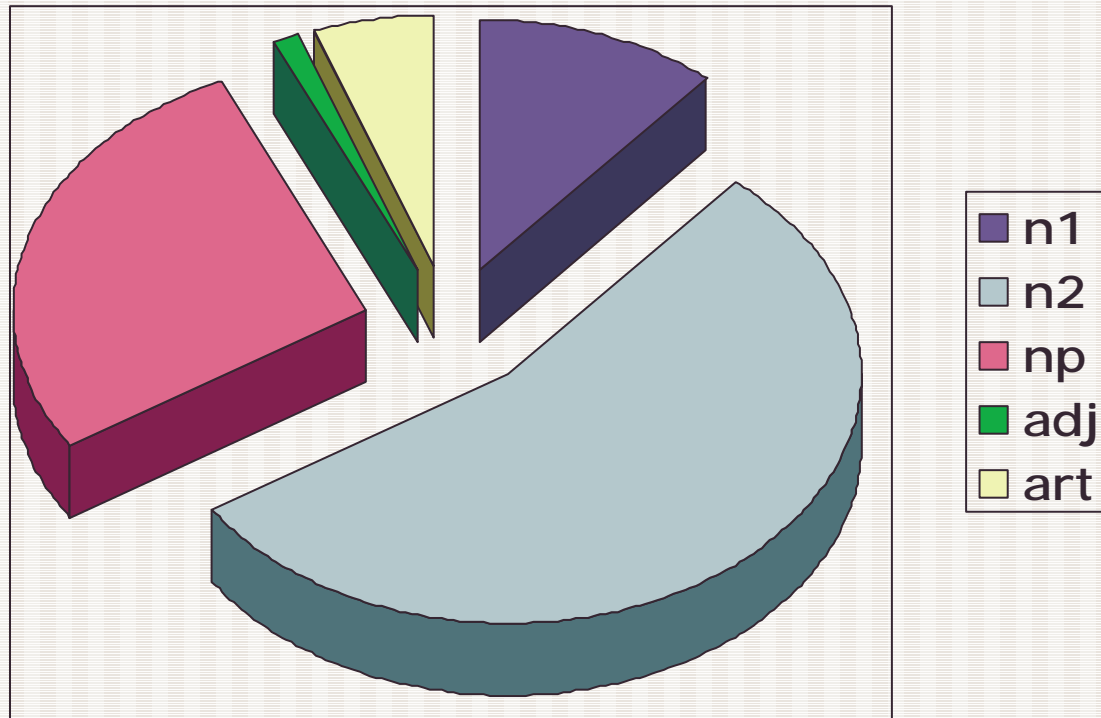
- Simple NP chunking algorithm:
 - Find a noun
 - Extend selection to the most description relating to the noun
- "loud rock," "rock," "tons of loud rock," "rock lobster," "not rock"
- Far more descriptive than an n-gram, but far more rare (less overlap)

Counting Terms

- Each term (n1,n2,np,art,adj) has two outputs: f_d and f_t
- f_d is 'document frequency:'
 - How often does this term occur everywhere? (High for 'the', 'loud')
- f_t is 'term frequency:'
 - How often does this term occur for our artist?

Term Coverage

- n2,np far greater in size



Experiment

■ AMG similar lists:

end of the year. Critics praised the angular yet melodic pop, and the album reached number 38 in the U.K. charts. However, none of the singles released from the album charted (including "This Is Pop"), nor did "Are You Receiving Me?," the teaser single for their second album, *Go 2* (1978).

After returning from a brief U.S. tour, Andrews quit the band; he would eventually form the League of Gentlemen with ...

[click here for more](#)

RELATED ARTISTS

Major Contributors to the Group: Todd Rundgren ★ John Leckie ★ Steve Lillywhite ★ Hugh Padgham

Similar Artists: Squeeze ★ Wire ★ The Posies ★ Duffy ★ World Party ★ The Police ★ Talking Heads
★ The Rezillos ★ Matthew Sweet ★ Hoodoo Gurus ★ Aztec Camera ★ Squares ★ Blur ★ Undertones
★ Split Enz ★ The Specials ★ The Soft Boys ★ The Only Ones ★ Madness ★ The La's ★ The Jam

■ "Soft" GT: relative comparisons are useful

What's a good scoring metric?

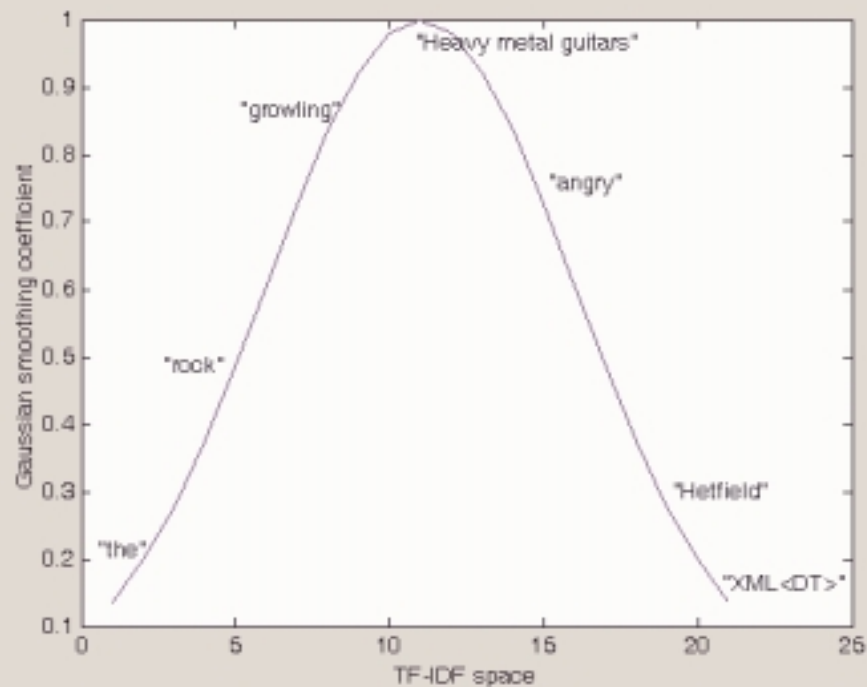
- Count # of shared terms?
 - What about 'the'? 'music?' 'web?'
 - Shouldn't more specific terms be worth more? "Electronic gamelan rock"

What's a good scoring metric?

- TF-IDF provides natural weighting
 - TF-IDF is $s(f_t, f_d) = \frac{f_t}{f_d}$
 - More 'rare' co-occurrences mean more.
 - i.e. two artists sharing the term "heavy metal banjo" vs. "rock music"
- But...

Curse of Specificity

- Typos, names, technical terms
- Reward mid-ground terms



Smoothing Function

- Inputs are term and document frequency with mean and standard deviation:

$$s(f_t, f_d) = \frac{f_t e^{-(\log(f_d) - \mu)^2}}{2\sigma^2}$$

- We use mean of 6 and stdev of 0.9

Example

■ For Portishead:

n1 Term	Score	n2 Term	Score	np Term	Score	adj Term	Score
gibbons	0.0774	beth gibbons	0.1310	beth gibbons	0.1648	cynical	0.2997
dummy	0.0576	sour times	0.0954	trip hop	0.1581	produced	0.1143
displeasure	0.0498	blue lines	0.0718	dummy	0.1153	smooth	0.0792
nader	0.0490	17 feb	0.0675	goosebumps	0.0756	dark	0.0583
tablets	0.0479	lumped into	0.0665	soulful melodies	0.0608	particular	0.0571
godrich	0.0479	which come	0.0635	rounder records	0.0499	loud	0.0558
irks	0.0467	mellow sound	0.0573	dante	0.0499	amazing	0.0457
corvair	0.0465	in together	0.0519	may 1997	0.0499	vocal	0.0391
durban	0.0461	musicians will	0.0494	sbk	0.0499	unique	0.0362
farfisa	0.0459	enough like	0.0494	grace	0.0499	simple	0.0354

Evaluation

- Will two known-similar artists have a higher overlap than two random artists?
- Use 2 metrics
 - Straight TF-IDF sum
 - Smoothed gaussian sum
- On each term type
- Similarity is: $S(a,b) = \sum s(f_t, f_d)$
for all shared terms

Predicting Similarity

- Accuracy: % of artist pairs that were predicted similar correctly ($S(a,b) > S(a,\text{random})$)
- Improvement = $S(a,b)/S(a,\text{random})$

	N1	N2	Np	Adj	Art
Accuracy	83%	88%	85%	69%	79%
Improvement	3.4x	2.7x	3.0x	6.8x	6.9x

P2P Similarity

- 1000s of 'Napster Robots' currently crawling P2P networks
- Download user->song relations
- Similarity inferred from collections?
- Similarity metric:

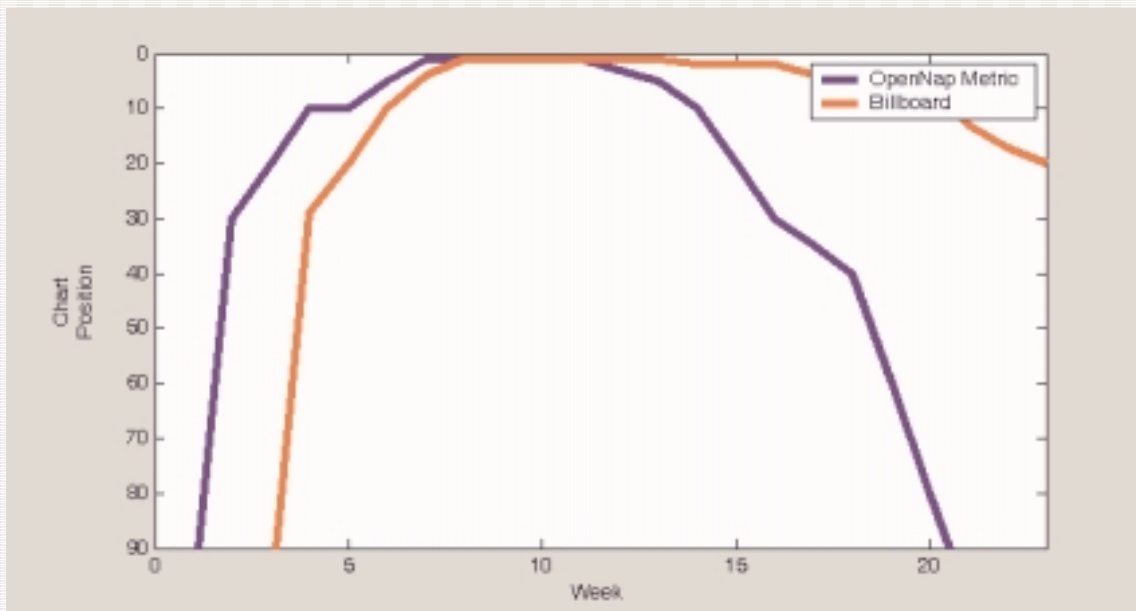
$$S(a,b) = \frac{C(a,b)}{C(b)} \left(1 - \frac{|C(a) - C(b)|}{C(c)}\right)$$

New SoundScan for P2P

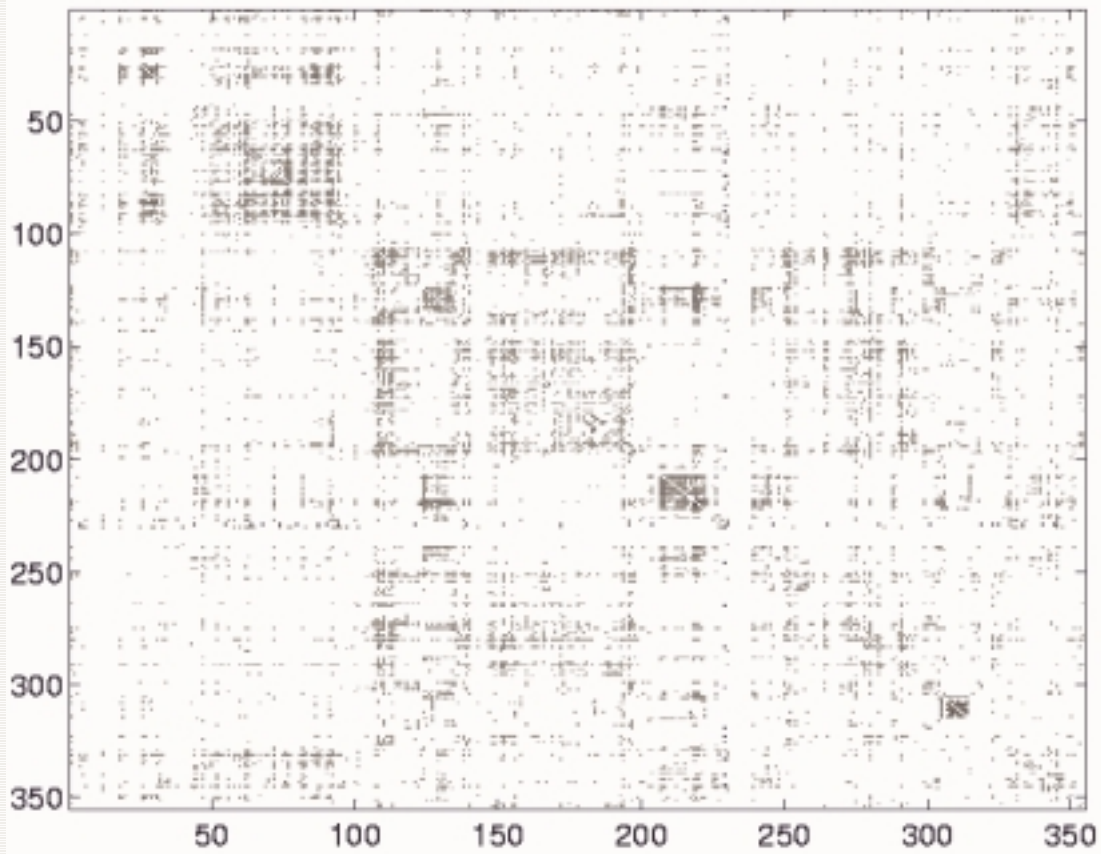
THIS WEEK	LAST WEEK	2 WKS. AGO	WEEKS ON	TITLE PRODUCER (SONGWRITER)	IMPRINT & NUMBER/PROMOTION LABEL	Artist	PEAK POSITION
1	1	1	1	Only Time - Enya			77240
2	2	2	1	Hotel California - Eagles			72555
3	3	3	1	Can'T Get You Out Of My Head - Kylie Minogue			70876
4	4	4	1	Smoke On The Water - Deep Purple			69949
5	5	5	1	Smooth Criminal - Alien Ant Farm			67992
6	6	6	1	In The End - Linkin Park			66538
7	7	7	1	Thank You - Dido			66603
8	8	8	1	Hero - Enrique Iglesias			65343
9	9	9	1	Get The Party Started - Pink			63849

p2p Trend Maps

- Far more #1s/year than 'real life'
- 7-14 day lead on big hits
- No genre stratification



P2P Similarity



OpenNap Comparison

- Instead of using AMG for GT, try the OpenNap data
- Similar results

	N1	N2	Np	Adj	Art
Accuracy	80%	82%	84%	68%	72%
Improvement	2.6x	2.1x	2.4x	7.1x	4.5x

Long-Scale Time-Aware

- CM is 'Time-Aware:'
 - Artists change over time
 - So does audience perception
- Gauges buzz
 - Parsable content goes up during album releases, major news
- Avoids 'stale' recommendations
- Captures that non-audio 'aboutness'

Query by Description

- “Play me something fast with an electronic beat!” “I’m tired tonight, let’s hear some romantic music.”
- We can use CM vectors to achieve time-aware QBD.
- We don’t need to label any data—the internet does that for us.

tuna-Mima Music Server - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Page Info Up Highlight

Google Search Web Search Site Page Info Up Highlight

Address http://tuna.xenakis.media.mit.edu/query_bd.php?move=1 Go

I want to hear something romantic to listen to with my girl near the fire.

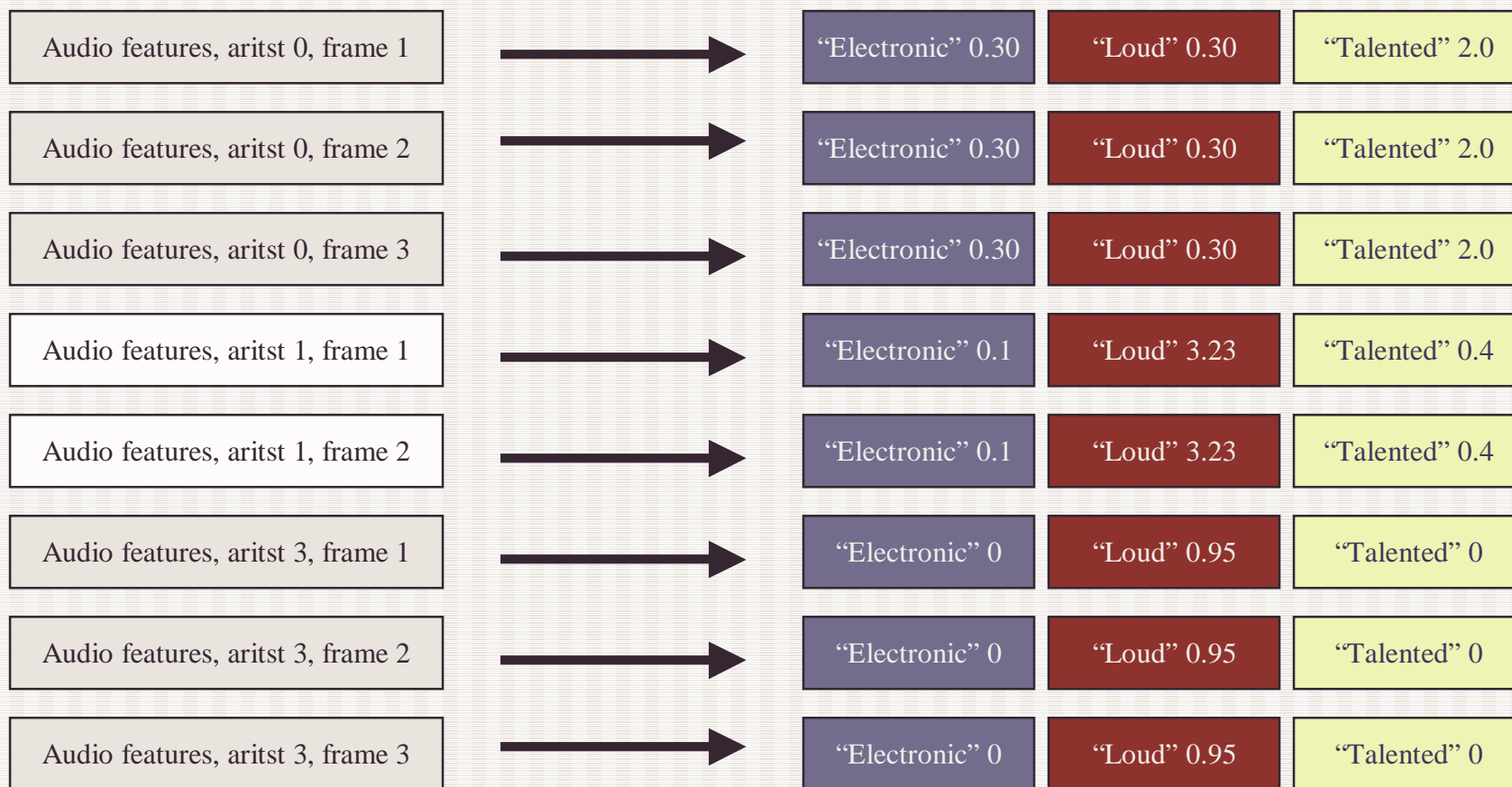
Click Me

I want to hear something romantic to listen to with my girl near the fire.

Album title	Artist	Songs	Average Rating
Enrique	Iglesias, Enrique	13	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Album title	Artist	Songs	Average Rating
One Woman's Live Journey	Newton John, Olivia	6	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Album title	Artist	Songs	Average Rating
Live at the London Palladium	Gaye, Marvin	3	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Live [Motown]	Gaye, Marvin	4	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Let's Get It On	Gaye, Marvin	1	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Album title	Artist	Songs	Average Rating
My Love Is Your Love [Australia Bonus Remix...]	Houston, Whitney	11	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
My Love Is Your Love	Houston, Whitney	1	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Whitney	Houston, Whitney	7	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Whitney Houston	Houston, Whitney	6	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Album title	Artist	Songs	Average Rating
Part III [#1]	112	10	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Room 112	112	13	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
112	112	10	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Album title	Artist	Songs	Average Rating
Aida	Sisqo	1	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>
Enter the Dru	Sisqo	1	<div style="width: 100%; height: 10px; background: linear-gradient(to right, #4a5568, #c63800);"></div>

Done Internet

ML Task for Audio QBD

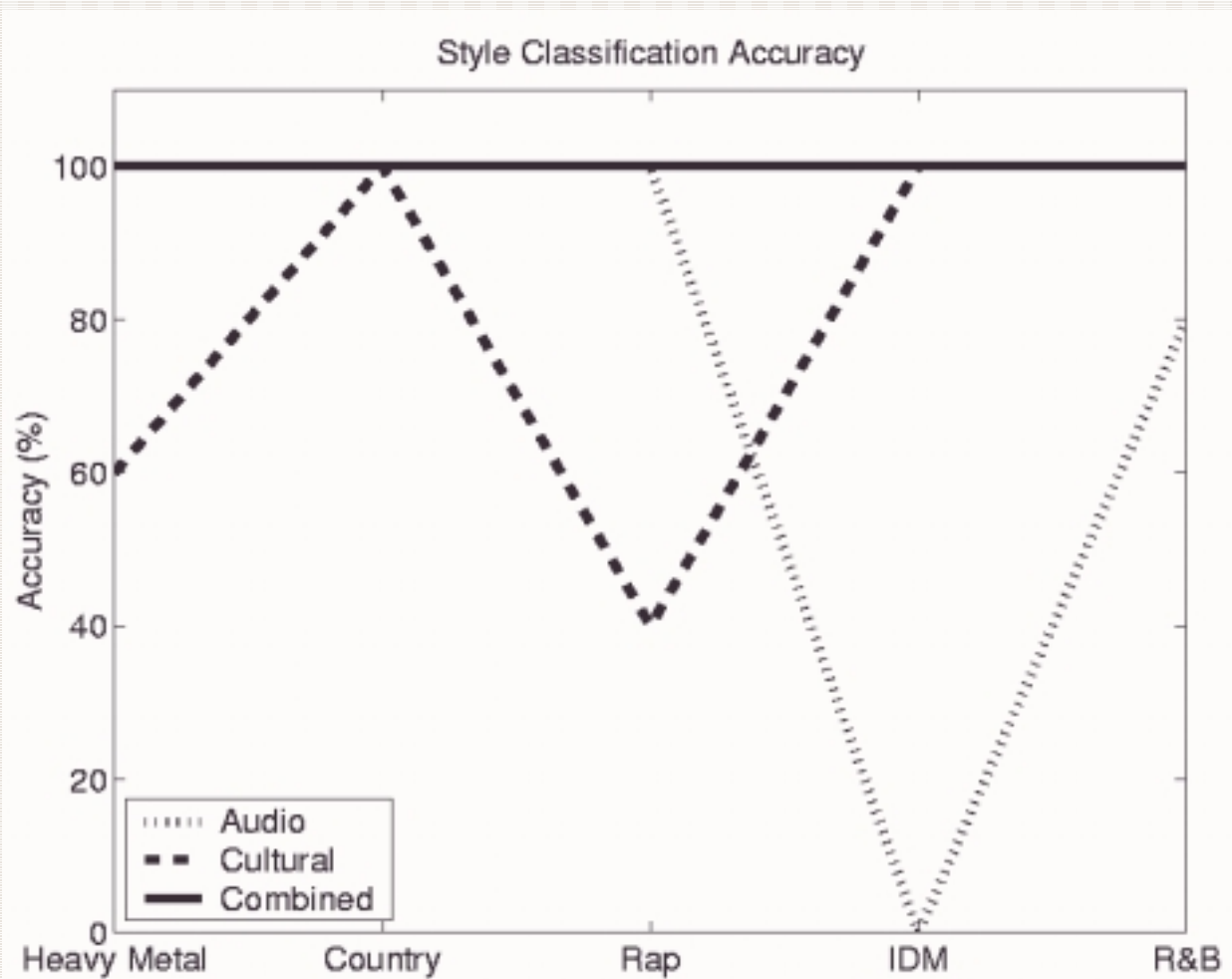


(3 of 20,000 possible terms)

Style ID

- Styles are specific small-set genres
 - "No Depression," "IDM," "Christian Metal," "Micro-house"
- Some are culturally defined, some are more acoustic
- Combine CM and acoustic representation for 100% style ID:

Style ID



Community Synthesis

- “What does romantic sound like?”
- Training machines for eigen-synthesis of linguistic concepts
- Current pass uses HMM
'regularization' for low-dimension spectral re-synthesis of high confidence terms

Conclusions

- CM a powerful cultural representation for music IR tasks
- Missing link for “bi-modal” music perception
- Next steps: standardizing, more accurate scoring, more applications
- Data publicly available