Content Management for EMD: What is “Interesting” Music?

Francois Pachet, Head Music Team
SONY Computer Science Laboratory - Paris
EMD: a Mapping problem
Millions of music titles, billions of users

20 million titles

1 billion users

• Sony Music: 
  500 K titles
• Western Music labels: 
  4 M
• Recorded Music: 
  10 M
• Indies: 
  10 M
• 1000 new CDs each week
EMD: a Mapping problem
Traditional Solution

20 million titles 1 billion users

- Sony Music: 500 K titles
- Western Music labels: 4 M
- Recorded Music: 10 M
- Indies: 10 M
- 1000 new CDs each week

Small subset of catalogue actually distributed

- Tangible support (CD)
- Complex Dynamics Labels / Radios / Retailers
- Hits phenomenon
EMD: a Mapping problem
New Solution?

20 million titles
Satellite, DAB, cable, internet, HDTV

1 billion users

- Sony Music: 500 K titles
- Western Music labels: 4 M
- Recorded Music: 10 M
- Indies: 10 M
- 1000 new CDs each week
EMD: What are the Issues?

Attention has been focused on enabling technologies:

- network transmission,
- audio compression,
- copy protection,
- copyright management.

Now remains the hard questions:

how to provide users with what they want?
should EMD systems provide only music they want?
how to attract attention of users?
how to exploit fully catalogues?
what is “interesting” music?
Is User Requirement Possible?

Many Individual Listening Behaviors

• For accessing known music:
  support – based (SACD)
  genre based, artist based, album-based feature-based,
• For accessing unknown music: exploratory vs exclusive.
• People usually cannot explicit musical features

• Listening Behaviors change with time

  Most people (90%) do not like novelty – at first.

• Some behaviors are unexpected
  • Accumulation of music never listened to (Napster)
  • Compulsive zapping
User Requirements?

Focus

genre
feature
artist
title

Exploration

Conservative

Exploratory
Research Areas at CSL

Features extraction - one title
- From the signal: Rhythm Extraction
- From data mining: Genre Extraction

Features of *sequences of titles*
- PathBuilder

User models and *exploration*
- PersonalRadio
Extracting Features

2 main methods to extract features/descriptors:
- signal analysis : the hard way
- data mining & collaborative filtering : the easy way

Signal well adapted to low level features:
- segmentation
- fundamental frequency
- beat tracking

Statistical analysis well adapted to high level features:
- genre
- clusters and similarity at the title level
Rhythm Extraction

Goal:
Extract an representation the rhythmic structure of a title

Rhythm structure made up of repeated occurrences of percussive sounds

Problem:
- Percussive sounds are infinite, a priori unknown
- Percussive sounds do not fit well with spectral methods
- Temporal synchronization between different sounds contribute to rhythm: need to differentiate between sound classes (e.g. snare and bass drum)
First Step
Extract occurrences of percussive events

Audio signal samples

Symbolic data: time indexes series

Descriptors
Extraction

Extract occurrences of two different percussive timbres

Two classes:

- Snare-like sounds
- Bass Drum-like sounds

Classification embedded in a detection scheme
Extracting Features: Genre

Genre taxonomy is necessary

No standard genre classification (Pachet 2000, RIAO)

Very difficult to extract from the signal, easier to extract from data mining

Experiments on 3 sources of data using co-occurrence techniques
Tracks on this CD
Tears For Fears—Everybody Wants To Rule The World
Split Enz—Message To My Girl
Suzanne Vega—Marlene On The Wall
Fine Young Cannibals—Suspicious Minds
Robert Cray—Don't Be Afraid Of The Dark
The Bluebells—Young At Heart
James Brown—I Got You (I Feel Good)
The Christians—Harvest Of The World
Big Country Fields Of Fire
Roger Daltrey—Giving It All Away
Curiosity Killed The Cat—Down To Earth
Was Not Was—Papa Was A Rolling Stone
D.N.A. Featuring Suzanne Vega—Tom's Diner

CDDB
40 M titles

Google
billions pages

Sources of data

http://www.amazon.com/...

1962-1966: The Red Album by the Beatles

Disc: 2
12. Eleanor Rigby

Comment from a customer:
“Still, with just about every song here an absolute classic (the remainder are simply "great"), this essential album is as important to pop as Beethoven's symphonies and Mozart's Requiem are to classical music.”

Radio (Fip)
BDS (100K titles)
Data Mining and Music Genre

Data sources:
- CDDB
- Radio Programs
- The Web (Google)

Co-occurrence: two items appear in a same context.

- Co-occurrence matrix
  \( C(i, j) = \text{nb of times } T_i \text{ and } T_j \text{ appeared together} \)

- Similarity matrix
  \[ \text{Sim}\left(T^1, T^2\right) = \sum_i T^1_i \cdot T^2_i \]

- PCA and clustering performed
Data Mining and Music Genre (Ctd)

Preliminary results

• CDDB and Google produce almost systematically:
  • *Artist clustering*
  • *Main genre clustering*

• Need large databases to produce relevant information (over 5000)

• Radio does not produce artist and genre clustering, but does produces unexpected (interesting) relations between titles:
  • *Covers:*
    “Lady Madonna” / Baroque ensemble  close to
    “Ticket to Ride” / The Beatles
  • *Timbral similarity*
    “Eleanor Rigby” and a Haydn String quartet
  • *Content of lyrics*
    “Kiss – Prince” and “Le baiser / A. Souchon”
<table>
<thead>
<tr>
<th>Google clustering</th>
<th>CDDB clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>** (25, 97)</td>
<td>** (0, 100)</td>
</tr>
<tr>
<td>**** (51, 75)</td>
<td>**** (5, 100)</td>
</tr>
<tr>
<td>****** Midnight Mood / Wes Montgomery</td>
<td>****** (13, 13)</td>
</tr>
<tr>
<td>****** (51, 51)</td>
<td>****** Music / Madonna</td>
</tr>
<tr>
<td>******** Body and Soul / Jim Hall</td>
<td>******** Goldeneye / Tina Turner</td>
</tr>
<tr>
<td>********* Girl from ipanema / Stan getz</td>
<td>********* All you need is love / The Beatles</td>
</tr>
<tr>
<td>**** (25, 90)</td>
<td>**** (5, 28)</td>
</tr>
<tr>
<td>***** (34, 38)</td>
<td>******** Michelle / The Beatles</td>
</tr>
<tr>
<td>******** california girls / The Beach Boys</td>
<td>******** Goldeneye / Tina Turner</td>
</tr>
<tr>
<td>********* God only knows / The Beach Boys</td>
<td>********* All you need is love / The Beatles</td>
</tr>
<tr>
<td>***** (25, 79)</td>
<td>**** (0, 100)</td>
</tr>
<tr>
<td>******** (25, 54)</td>
<td>***** (14, 72)</td>
</tr>
<tr>
<td>********* (25, 25)</td>
<td>******** Girl from ipanema / Stan getz</td>
</tr>
<tr>
<td>********** Michelle / The Beatles</td>
<td>********* (14, 19)</td>
</tr>
<tr>
<td>********** Music / Madonna</td>
<td>********* God only knows / The Beach Boys</td>
</tr>
<tr>
<td>********** (34, 34)</td>
<td>********** (14, 14)</td>
</tr>
<tr>
<td>*********** All you need is love/The Beatles</td>
<td>*********** Good vibrations / The Beach Boys</td>
</tr>
<tr>
<td>*********** Eleanor rigby / The Beatles</td>
<td>*********** california girls / The Beach Boys</td>
</tr>
<tr>
<td>*********** (67, 67)</td>
<td>***** (0, 100)</td>
</tr>
<tr>
<td>*********** Goldeneye / Tina Turner</td>
<td>********* Requiem / Mozart</td>
</tr>
<tr>
<td>*********** Requiem / Mozart</td>
<td>******** (0, 0)</td>
</tr>
<tr>
<td></td>
<td>******** Body and Soul / Jim Hall</td>
</tr>
<tr>
<td></td>
<td>******** Midnight Mood / Wes Montgomery</td>
</tr>
</tbody>
</table>
Music retrieval: the sequence approach

**Rationale:**
- Music listening is rarely performed on individual titles but rather by sequences: radio, concert, album, etc.
- Querying titles is possible only for titles you know how to describe: the *language mismatch* problem

**Idea:** specify properties of the music *sequences* instead of properties of individual titles. Solve the related combinatorial problem and let the system chose the items

Which properties? How to solve the combinatorial problem?
Musical descriptors?

Low-level: Mpeg-7 (e.g. spectral centroid, segmentation)

High-level:

- Energy
- Tempo
- Genre
- Rhythm type
- Voice type
- Main instruments
- Configuration
- Danceability ...
Mpeg7 : Spectral Descriptors

- Spectral Envelope
- Spectral Envelope of Sinusoidal Component
- Spectral Envelope of Stochastic Component
- Spectral Centroid
- Spectral Tilt
- Noise Shape
- Harmonic Distortion
- Harmonics Odd/Even ratio
- Harmonic/noise cutting frequency
# The Database (excerpt)

<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Artist(s)</th>
<th>Genre</th>
<th>Location</th>
<th>Record Length</th>
<th>Tempo</th>
<th>Vocal(s)</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>It's only make believe</td>
<td>McDowell / Twi</td>
<td>Country Crooner</td>
<td>Calif</td>
<td>231s</td>
<td>slow</td>
<td>low-pitched</td>
<td>Two men, Calif guitar, strings</td>
</tr>
<tr>
<td>2</td>
<td>Take it to the limit</td>
<td>W. Nelson / Jen</td>
<td>Country Blues</td>
<td>Calif</td>
<td>229s</td>
<td>nasal</td>
<td>Two men</td>
<td>Piano, Harmonica</td>
</tr>
<tr>
<td>3</td>
<td>Sunrise</td>
<td>Atkins Chet</td>
<td>Jazz Calif</td>
<td>Calif</td>
<td>250s</td>
<td>slow fast</td>
<td>Instrumental</td>
<td>Instrumental, Jazz guitar, strings</td>
</tr>
<tr>
<td>4</td>
<td>Cricket Ballet (The)</td>
<td>Atkins Chet</td>
<td>Jazz Calif</td>
<td>Calif</td>
<td>211s</td>
<td>fast slow</td>
<td>Instrumental</td>
<td>Instrumental, Jazz guitar, Funk bass</td>
</tr>
<tr>
<td>5</td>
<td>Don’t give up</td>
<td>Nelson W. / O’</td>
<td>Country Pop</td>
<td>Calif</td>
<td>290s</td>
<td>nasal</td>
<td>Man/Women</td>
<td>Acoustic guitar, harmonica</td>
</tr>
<tr>
<td>6</td>
<td>Still is still moving to</td>
<td>Nelson Willie</td>
<td>Country Calif</td>
<td>Calif</td>
<td>210s</td>
<td>fast</td>
<td>Man</td>
<td>Calif guitar, Calif guitar</td>
</tr>
<tr>
<td>7</td>
<td>Jukebox Junkie</td>
<td>Mellons Ken</td>
<td>Country Swing</td>
<td></td>
<td>162s</td>
<td>very fast</td>
<td>Low-pitched</td>
<td>Violin, Calif guitar</td>
</tr>
<tr>
<td>8</td>
<td>Little Rock</td>
<td>Raye Colin</td>
<td>Country Pop</td>
<td></td>
<td>237s</td>
<td>slow</td>
<td>Normal</td>
<td>Piano, Strings</td>
</tr>
<tr>
<td>9</td>
<td>Not a moment too soon</td>
<td>Mac Graw Tim</td>
<td>Country Calif</td>
<td>Calif</td>
<td>222s</td>
<td>slow fast</td>
<td>Hoarse</td>
<td>Calif guitar, Piano</td>
</tr>
<tr>
<td>10</td>
<td>Lovin’ all night</td>
<td>Crowell Rodney</td>
<td>Country Pop</td>
<td></td>
<td>227s</td>
<td>Fast</td>
<td>Normal</td>
<td>Calif guitar, Brass</td>
</tr>
<tr>
<td>11</td>
<td>Hard way (the)</td>
<td>Carpenter Mary</td>
<td>Country Pop</td>
<td></td>
<td>262s</td>
<td>slow fast</td>
<td>Normal</td>
<td>Woman, Calif guitar, Piano</td>
</tr>
<tr>
<td>12</td>
<td>Girls with guitars</td>
<td>Judd Winonna</td>
<td>Country Pop</td>
<td></td>
<td>194s</td>
<td>Fast</td>
<td>Powerful</td>
<td>Pop guitar, Piano</td>
</tr>
</tbody>
</table>
Classification of Music Genres

A graph of genres with 2 relations:
isA
closeTo
Example of Music Program

“Liner note-like” description of desired Music Program:

- 12 Titles, all different
- Styles are close to their neighbors
- No slow/very slow tempos (Cardinality Constraint)
- At least 30% female-type voice
- At least 30% purely instrumental pieces
- At least 40% brass
- Authors are all different
- At most 20% “Country Pop” style
- One song by “Harry Connick Jr”
<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Artist</th>
<th>Genre</th>
<th>State</th>
<th>Duration</th>
<th>Tempo</th>
<th>Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Sunrise</td>
<td>Atkins Chet, Jazz</td>
<td>Calif</td>
<td>250s</td>
<td>slow fast</td>
<td></td>
<td>instrumental, instrumental, jazz guitar, strings</td>
</tr>
<tr>
<td>21</td>
<td>Surrounded</td>
<td>Kreviazuk Chant</td>
<td>Pop, Calif</td>
<td>238s</td>
<td>slow fast</td>
<td></td>
<td>powerful, Woman, piano, strings</td>
</tr>
<tr>
<td>6</td>
<td>Still is still moving to</td>
<td>Nelson Willie, Country</td>
<td>Calif</td>
<td>210s</td>
<td>fast</td>
<td></td>
<td>nasal, Man, calif guitar, calif guitar</td>
</tr>
<tr>
<td>9</td>
<td>Not a moment too soon</td>
<td>Mac Graw Tim, Country</td>
<td>Calif</td>
<td>222s</td>
<td>slow fast</td>
<td></td>
<td>hoarse, Man, calif guitar, piano</td>
</tr>
<tr>
<td>10</td>
<td>Lovin' all night</td>
<td>Crowell Rodney, Country, Pop</td>
<td>227s</td>
<td>fast</td>
<td>normal</td>
<td></td>
<td>Man, calif guitar, brass</td>
</tr>
<tr>
<td>11</td>
<td>Hard way (the)</td>
<td>Carpenter Mary, Country, Pop</td>
<td>262s</td>
<td>slow fast</td>
<td>normal</td>
<td></td>
<td>Woman, calif guitar, piano</td>
</tr>
<tr>
<td>17</td>
<td>Point of rescue (the)</td>
<td>Ketchum Hal, Country</td>
<td>Calif</td>
<td>265s</td>
<td>fast</td>
<td>normal</td>
<td>Man, calif guitar, calif guitar</td>
</tr>
<tr>
<td>50</td>
<td>At seventeen</td>
<td>Ian Janis, Pop, Folk</td>
<td>Calif, Folk</td>
<td>281s</td>
<td>slow fast</td>
<td></td>
<td>soft, Woman, acoustic guitar, brass</td>
</tr>
<tr>
<td>27</td>
<td>Dream on</td>
<td>Labounty Bill, Pop</td>
<td>Calif</td>
<td>298s</td>
<td>slow fast</td>
<td></td>
<td>broken, Man, keyboard, brass</td>
</tr>
<tr>
<td>106</td>
<td>Another time another place</td>
<td>Steely Dan, Jazz</td>
<td>Calif</td>
<td>245s</td>
<td>fast</td>
<td></td>
<td>instrumental, instrumental, piano, keyboard</td>
</tr>
<tr>
<td>112</td>
<td>Learn to love you</td>
<td>Connick Harry J, Jazz, Crooner</td>
<td>279s</td>
<td>slow fast</td>
<td>muffled</td>
<td></td>
<td>Man, brass, strings</td>
</tr>
<tr>
<td>137</td>
<td>Heart of my heart</td>
<td>Elgart Les, Jazz, Swing</td>
<td>151s</td>
<td>slow fast</td>
<td></td>
<td></td>
<td>instrumental, instrumental, double bass, brass</td>
</tr>
</tbody>
</table>
Problem Complexity

sequence: 1 2 3 4 ... 19 20

domains:

• Build a sequence of 20 titles each title in database of 100,000 items,

• satisfy a set of global constraints

• naive search space = $100,000^{20} = 10^{100}$
Constraint Satisfaction

• Set of variables: item1, item2, item3 ...
• Variables have a finite domain (e.g. the database of titles)
• Set of constraints holding on variables

Exhaustive search with pruning = constraint filtering: At each step, constraints are considered individually and used to reduce the domains of future variables.

Filtering depends on the constraint, and on the variables already instantiated.
Global Constraints

- Running constraints
- Cardinality constraints
- Difference constraints
- Pattern constraints

Each constraint class has efficient filtering procedures
Music Exploration: the Personal Radio Project

Aim:
provide music for all possible listening behaviors
collect user feedback and behaviours

Focus on the *exploration* aspect:
System provides ability to shift explicitly between exploratory and non exploratory modes: the *Magic slider*
Personal Radio

• User Workshop in progress in Europe in 4 locations.
  10,000 titles

Analysis of Listening Behaviors change with time

  Most people use the magic slider intensively, after an initial period of conservatism.
Conclusion

• Music can be “interesting” for users for many reasons
  • grounded features yielding similarity functions
  • sequences
  • exploration / discovery of unknown music

• Cuidado European IST project in progress (Ircam, UPF, Oracle) to develop other extractors and perform scale-up on very large catalogues (1 million)