



機械学習に基づく 音楽情報処理

吉井 和佳

産業技術総合研究所 (AIST)

発表の内容

○ 音楽情報処理分野の概況

- 産業的・学術的にどうなのか

○ 機械学習の基本的利用法

- 機械学習がどのように取り入れられているか

○ 産総研における研究事例

- デモアプリケーションを4つ紹介

○ これからの音楽情報処理

- 機械学習のプロのみなさまへ

音楽情報処理分野の発展

○「音楽」は産業・文化面で主要なコンテンツ

● 音楽と触れ合う手段が増加

- 定額制聴き放題サービス
- 情報通信・音楽圧縮技術
- 携帯型音楽プレイヤー



● 好きな曲を好きな時に好きな場所で好きなだけ聴ける

近い将来、すべての音楽がデジタルアーカイブ化され
配信・検索・共有・創作・発信されることが確実

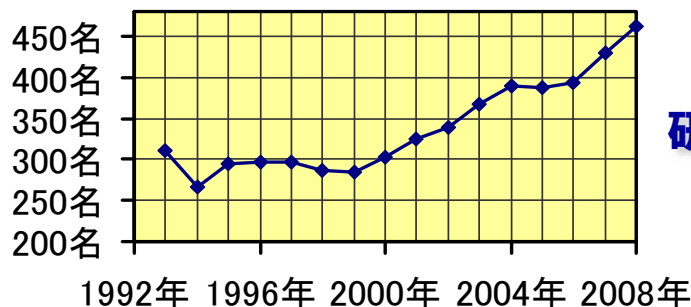


**音楽解析・検索・推薦技術
に対する潜在的な需要は大きい**

学術分野としての成長（国内）

○ 情報処理学会 音楽情報科学研究会 SIGMUS

- 登録者数・発表件数ともに急速に増加
- 年5回開催される研究会の規模が拡大
 - 以前の平均：1日 x 3回 2日 x 2回
 - 07,08年度：2日 x 4回 3日 x 1回



研究会登録者数

音声言語処理研究会
SIGSLPに比肩！



年間発表件数



夏シンポ発表件数

学術分野としての成長 (国際)

○ **国際音楽情報検索学会 ISMIR** International Society for Music Information Retrieval

● **第10回国際会議 ISMIR 2009**

○ **10/26-30@神戸国際会議場**

○ **280人をこえる参加者 (不況なのに過去最大)**

○ **当日参加もwelcome!**

完全なシングルセッション

38件のオーラル発表

85件のポスター発表

Conference Program



10th International Society for Music Information Retrieval Conference
Kobe, Japan
October 26-30, 2009



- HOME
- Call for Papers
- Submission
- Information for Presenters
- Program
- Tutorials
- Invited Presentations
- Special Session: f(MIR)
- MIREX 2009
- Registration
- Conference Venue
- Accommodation
- Access
- Social Program
- Music Performances
- Satellite Events
- Corporate Support
- Conference Committee
- ISMIR Links
- Contact

Welcome to ISMIR 2009

The Tenth International Society for Music Information Retrieval Conference will take place October 26-30, 2009 (Monday through Friday), in Kobe, Japan. Looking at the recent rapid growth of ISMIR conferences since its inception in 2000 and the expansion of the music information retrieval (MIR) community, MIR people may no longer consider acronym MIR as "music information retrieval" in the strict sense, but rather think that MIR essentially fits "music information research" in a broad sense. Correspondingly, MIR also draws considerable attention from other research/application fields. ISMIR therefore aims to open a new horizon of the exchange and discussion of ideas, issues, results, and perspectives for people of widespread areas, such as representatives of academia, industry, entertainment, and education, including researchers, developers, educators, librarians, students, and professional users. Alongside presentations of original theoretical research and practical work, ISMIR provides introductory and in-depth tutorials, and a venue for the showcase of current MIR-related products and systems.

Important Dates

- **Deadline for submission of tutorial proposals: January 31, 2009**
- **Deadline for submission of papers: May 6, 2009**
The deadline was extended to May 15 (Fri), 23:00, JST (GMT +09:00).
The deadline was extended to May 22 (Fri), 23:00, JST (GMT +09:00).
- **Author rebuttal period: June 26 - July 3**
The rebuttal period was changed to July 3 (Fri) - July 6 (Wed), 2009
- **Notification of paper acceptance: July 17, 2009**
The date was changed to July 20 (Mon), 12:00, JST (GMT +09:00).
The date was changed to July 23 (Thu), 12:00, JST (GMT +09:00).
- **Deadline for submission of camera-ready papers: August 7, 2009**
The date was changed to August 10, 2009.
- **Deadline for submission of abstracts for late-breaking / demo: August 31, 2009**
- **ISMIR 2009: October 26-30, 2009**



| | Monday (October 26) | Tuesday (October 27) | Wednesday (October 28) | Thursday (October 29) | Friday (October 30) |
|-------------|------------------------|---|---|---|--|
| 8:00-9:00 | | Registration (8:15) | Registration (8:35) | Registration (8:35) | Registration (8:35) |
| 9:00-10:00 | Registration (9:15) | Opening Remarks Keynote Talk | Oral Session (OS4) | Keynote Talk | Special Session: f(MIR) |
| 10:00-11:00 | Tutorials AM | | | Oral Session (OS7) | Late- Breaking / Demo Session |
| 11:00-12:00 | | Oral Session (OS1) | Oral Session (OS5) | | Oral Session (OS9) |
| 12:00-13:00 | | Poster Madness Poster Session (PS1) with Lunch Box | Poster Madness Poster Session (PS3) with Lunch Box | Poster Madness Poster Session (PS4) with Lunch Box | |
| 13:00-14:00 | | | | | Concluding Remarks |
| 14:00-15:00 | Tutorials PM | Oral Session (OS2) | MIREX Panel Discussion | Oral Session (OS8) | |
| 15:00-16:00 | | Poster Madness Poster Session (PS2) with Coffee | MIREX Poster Session with Coffee | Panel Discussion | |
| 16:00-17:00 | | | Oral Session (OS3) | Oral Session (OS6) | General Meeting |
| 17:00-18:00 | | | Shakuhachi Concert & Reception | | |
| 18:00-19:00 | | | | | |
| 19:00-20:00 | | | | | Banquet |
| 20:00-22:00 | | | | | |

音楽情報処理とは？

- 計算機を使って音楽に関する情報を抽出したり、解析したり、利用したり、付与したりすること
 - 何を扱うのか？
 - 音楽そのもの：音響信号(オーディオ)、MIDI(シンボル) など
 - 関連するもの：アノテーション、タグ、レーティング など
 - 何を行うのか？
 - 認識系
 - さまざまな音楽要素の自動認識
 - ジャンル、ムード、楽器名、基本周波数、発音時刻、拍子、テンポ、調、コード進行、繰り返し構造 など
 - 生成系
 - 作曲や演奏の支援
 - 歌唱音声合成・楽音合成
 - 自動作曲・演奏表情付け
 - 自動伴奏・ジャムセッション

アプリケーション
音楽検索・推薦
プレイリスト生成
楽曲類似度計算

認識系の研究例

○ ジャンル識別：典型的なパターン認識問題

- 音響信号を入力としてジャンル名を出力したい

○ 学習時

- 音響信号とジャンル名のペアから識別器を学習

○ 認識時

- 新たな音響信号のジャンル名を識別



楽器音識別やムード識別も同様の枠組み

認識系の研究例

○コード進行認識：時系列パターン認識問題

- 音響信号を入力として**コード名系列**を出力したい

○学習時

- 音響信号とコード名系列のペアから識別器を学習

○認識時

- 新たな音響信号のコード名系列を出力



コード名 (認識対象クラス)

| | | | | | | |
|---|---|---|----|---|---|---|
| C | D | E | F | G | A | B |
| ド | レ | ミ | ファ | ソ | ラ | シ |

+ maj, min, maj7, min7, dom7, aug, dim, ...

認識系の研究例

○ テンポ推定

- 基本的な思想：自己相関関数に基づく周期性検出

○ 発音時刻検出

- 基本的な思想：急激なパワー変化を検出

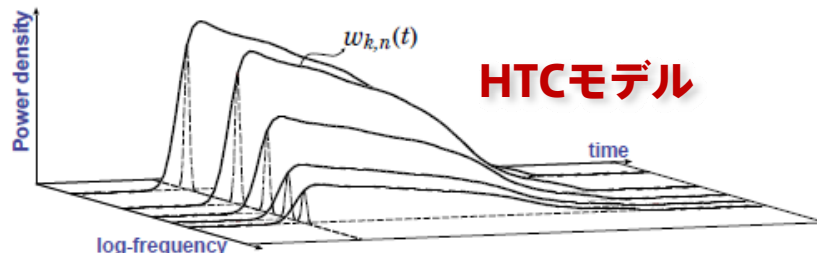
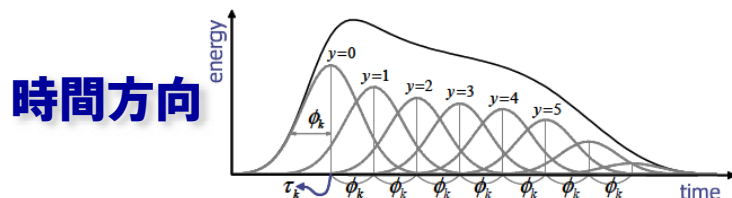
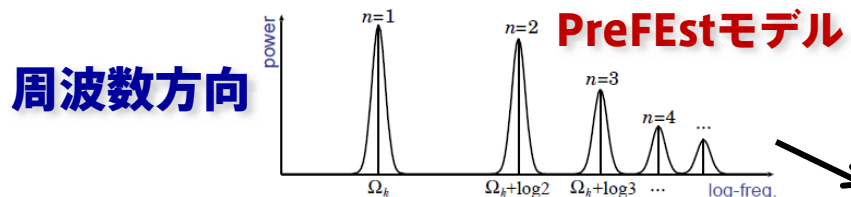
○ 基本周波数推定

- 混合ガウスモデルをスペクトルにフィッティング

- PreFEst [後藤1999]

- HTC [亀岡2005]

ブラインド音源分離
= モデルパラメータの学習



時間-周波数平面における分布

機械学習に基づく音楽情報処理

○ テーマ「何の目的で何を学習するのか」

- **HybridRecommender: 音楽を推薦する**
 - ユーザの好みに合った楽曲を推薦したい
 - 音響信号と評価履歴から**ユーザの好み**を学習
- **MusicThumbnailer: 音楽を可視化する**
 - 楽曲内容が一目でわかるサムネイル画像を生成したい
 - 音響信号から画像への**写像パラメータ**を学習
- **MusicCommentator: 音楽を言語で表現する**
 - 楽曲内容に対して人間のようにコメントを行わせたい
 - 音響信号とコメントとの**対応付け**を学習
- **LyricSynchronizer: 音楽に歌詞を同期させる**
 - 歌詞テキストを楽曲再生に同期させて表示させたい
 - 音響信号と歌詞との**アラインメント**を学習



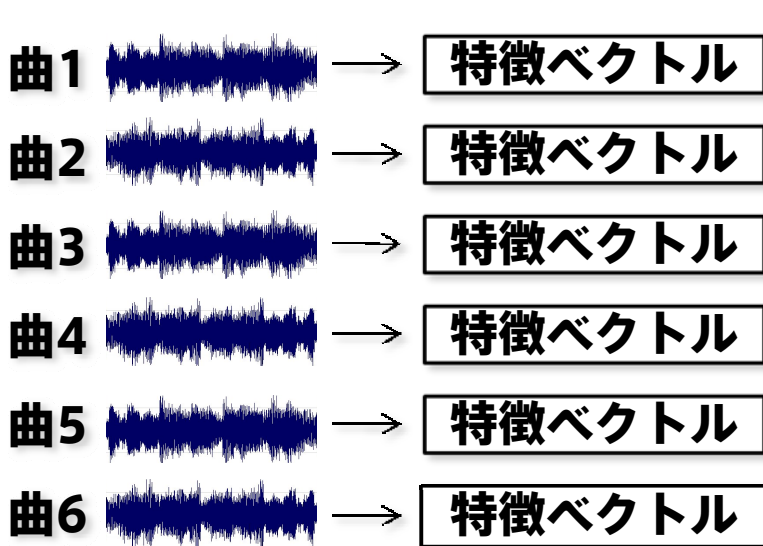
HYBIRD RECOMMENDER

音楽推薦システム

吉井 和佳 後藤 真孝
駒谷 和範 尾形 哲也 奥乃 博

研究アプローチ

- 高精度かつバラエティ豊かな推薦がしたい
 - 「音楽内容」と「ユーザの評価」を同時に考慮



音響的特徴量



| | 曲1 | 曲2 | 曲3 | 曲4 | 曲5 | 曲6 |
|------|----|----|----|----|----|----|
| ユーザ1 | 5 | | 4 | | | 1 |
| ユーザ2 | | 2 | | 3 | 2 | |
| ユーザ3 | | 1 | | | | 5 |
| ユーザ4 | 4 | | | 1 | 4 | 1 |
| ユーザ5 | | 5 | 5 | | 5 | |

5段階評価(行動履歴)

データ生成の不確定性を適切に扱いたい
→ 確率的な枠組みでのデータ統合

デモンストレーション

評価一覧の表示・非表示 楽曲一覧を表示

| タイトル | アーティスト | あなたの評点 |
|--|------------------------|---|
|  ALARM | 安室 奈美恵 | <input type="radio"/> Poor <input type="radio"/> Bad <input type="radio"/> Fair <input type="radio"/> Good <input checked="" type="radio"/> Great 評価を削除 |
|  WANT ME, WANT ME | 安室 奈美恵 | <input type="radio"/> Poor <input type="radio"/> Bad <input type="radio"/> Fair <input type="radio"/> Good <input checked="" type="radio"/> Great 評価を削除 |

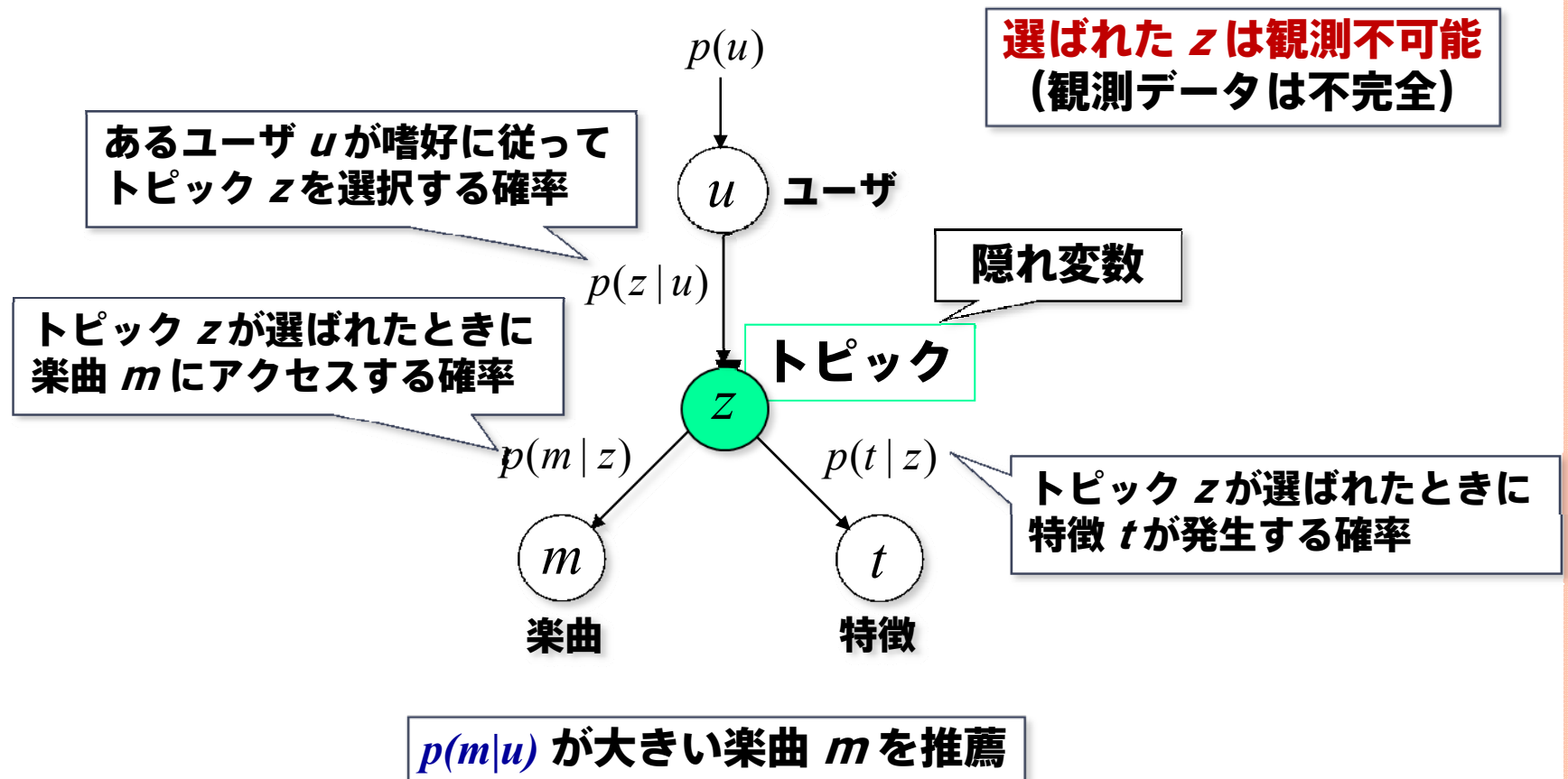
[評価を更新](#)

| 協調フィルタリング | 内容に基づくフィルタリング | ハイブリッド型フィルタリング |
|--|---|---|
| 1  GIRL TALK / 安室 奈美恵 |  REEEWIND! / M-FLO loves Crystal Kay |  DO THE MOTION / BoA |
| 2  Put 'Em Up / 安室 奈美恵 |  Put 'Em Up / 安室 奈美恵 |  GIRL TALK / 安室 奈美恵 |
| 3  SO CRAZY / 安室 奈美恵 |  SO CRAZY / 安室 奈美恵 |  メリクリ / BoA |
| 4  I WILL / 安室 奈美恵 |  To All Tha Dreamers / SOUL'd OUT |  VALENTI / BoA |
| 5  ALL FOR YOU / 安室 奈美恵 |  Can You Keep A Secret? / 宇多田 ヒカル |  Shine We Are! / BoA |
| 6  DO THE MOTION / BoA |  恋におちたら / Crystal Kay |  SO CRAZY / 安室 奈美恵 |
| 7  SMILY / 大塚 愛 |  Far away / 浜崎 あゆみ |  REEEWIND! / M-FLO loves Crystal Kay |
| 8  メリクリ / BoA |  make a secret / BoA |  miss you / m-flo loves melody. & 山本領平 |
| 9  LEGEND / 中島 美嘉 |  QUINCY / BoA |  Put 'Em Up / 安室 奈美恵 |
| 10  & / 浜崎 あゆみ |  THE マンパワー!!! / モーニング娘。 |  YOUR NAME NEVER GONE / CHEMISTRY |

ハイブリッド型推薦モデル

トピックモデルに基づく嗜好推定

- 文書-単語の共起モデルであるpLSIの拡張版を利用
 - 音楽内容と他人の評価を同時にモデル化





MUSIC THUMBNAILER

音楽を可視化するシステム

吉井 和佳 後藤 真孝

研究の背景

○音楽は聴いてみるまで中身が分からない

- 逐一試聴していくには時間がかかる
- 気分にあわない曲も聴くことになる

どんな曲？



あなたに薦めたい曲は…

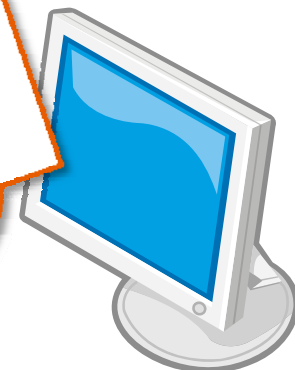
Lost in My Dreams: J. Manning

21st Century: D. Burke

Guess Again: D. Burke

My Faith: 22 Project Band

Rocket in My Pocket: G. Irwin



研究アプローチ

○ 音楽の中身がすぐに把握できるようにしたい

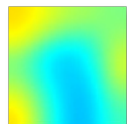
- 音響信号からサムネイル画像を生成する
- 「耳で聴く」から「目で見える」へ

ユーザの経験

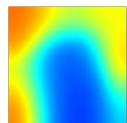
既知のヒップホップ曲
は  だったから…

3・5曲目は
ヒップホップ？

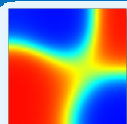
あなたに薦めたい曲は…



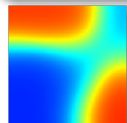
Lost in My Dreams: J. Manning



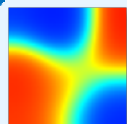
21st Century: D. Burke



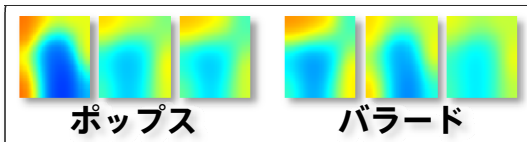
Guess Again: D. Burke



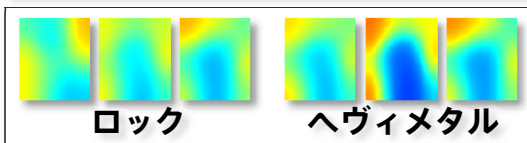
My Faith: 22 Project Band



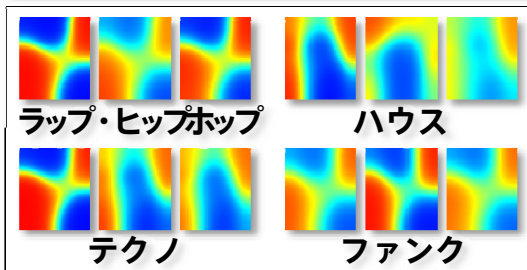
Rocket in My Pocket: G. Irwin



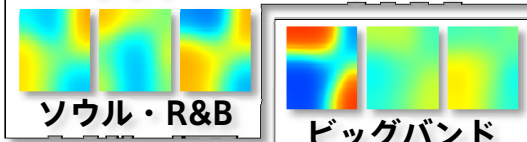
ポップス
#01~#06



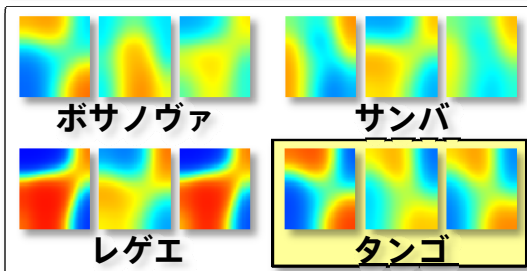
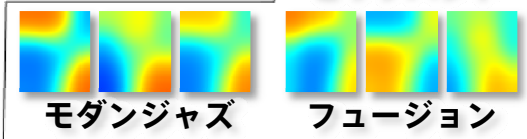
ロック
#07~#12



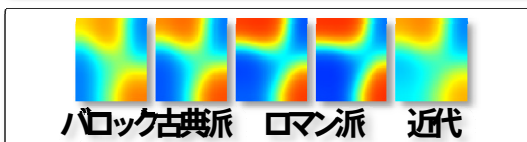
ダンス
#13~#27



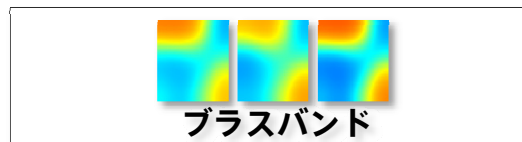
ジャズ
#28~#36



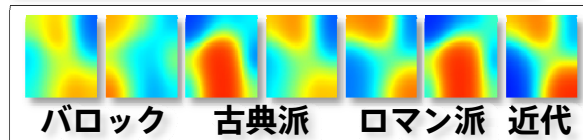
ラテン
#37~#48



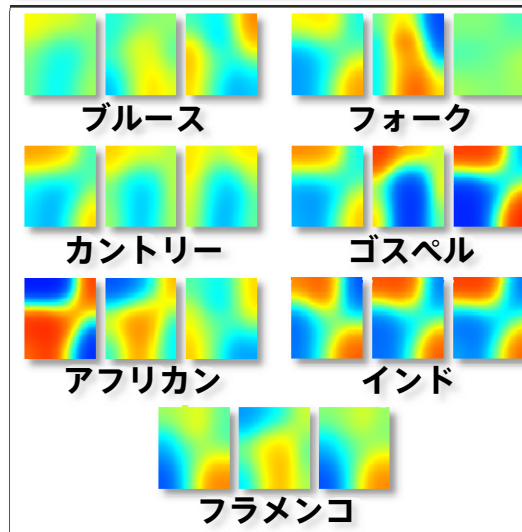
クラシック
(管弦楽)
#49~#53



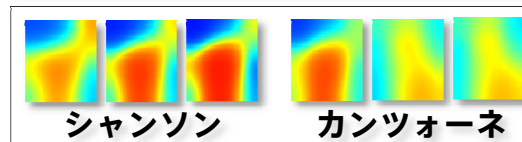
行進曲
#54~#56



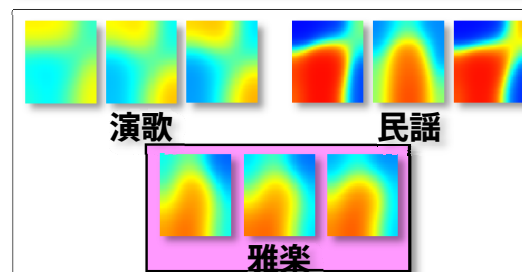
クラシック
(器楽・室内楽)
#57~#63



ワールド
#64~#84



声楽
#85~#90



邦楽
#91~#99
ア・カペラ
#100

最適な写像の学習と運用

1. 写像モデルを設計

$$Y = \text{Sigmoid}(AX)$$

- 未知パラメータ A を含む

2. コスト関数を定式化

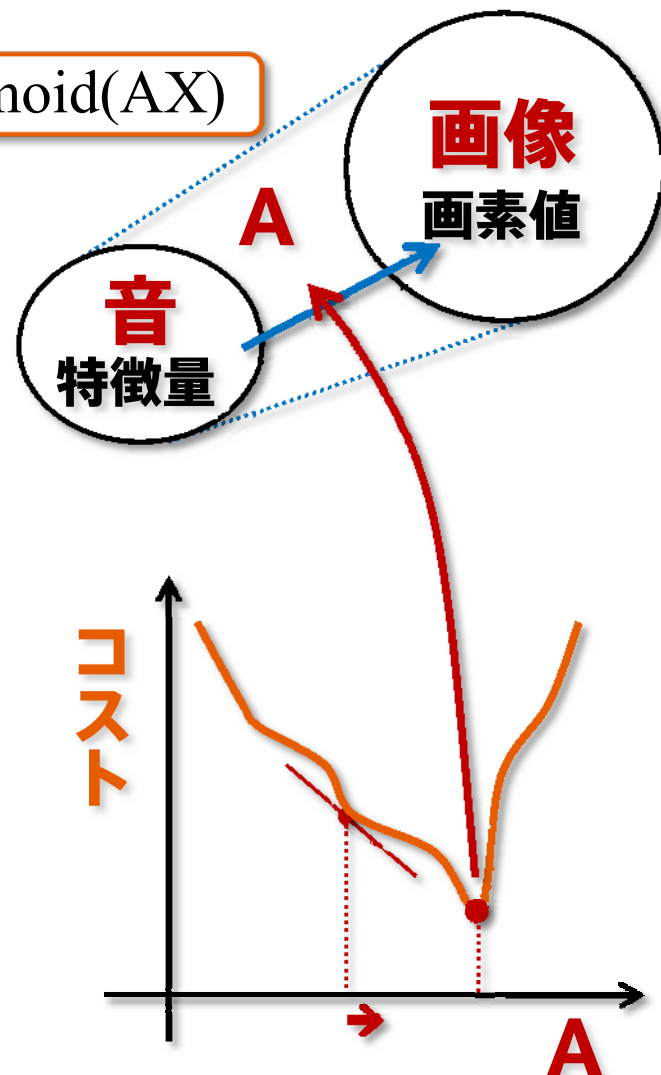
- A の関数

3. コスト関数を最小化

- 最急降下法による反復最適化
 - 傾き計算
 - 減少する方向へ更新

4. サムネイル画像の生成

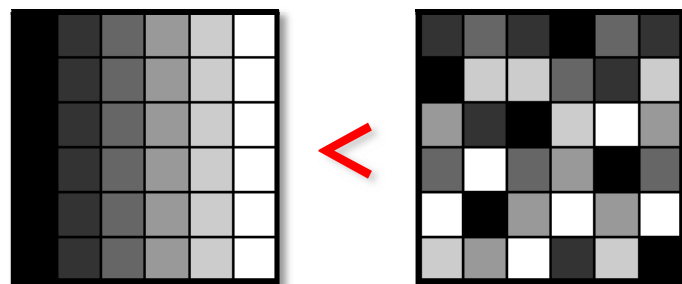
- 最適な A を写像モデルに適用



サムネイルコストの評価尺度

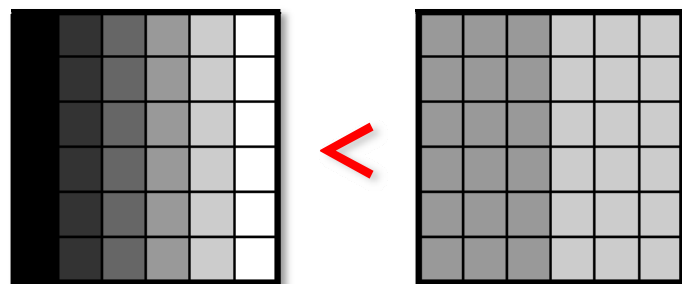
○記憶性コスト

- 各サムネイルの覚えにくさ
 - 隣接する画素値の差の合計



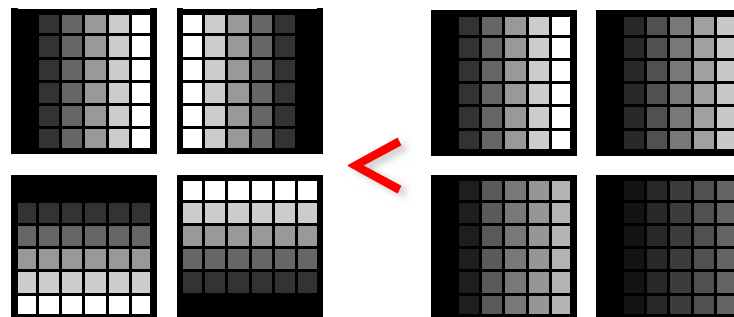
○表現力コスト

- 各サムネイルの情報不足度
 - サムネイル内分散 $\times -1$



○区別性コスト

- サムネイル群の見分けにくさ
 - サムネイル間分散 $\times -1$





MUSICCOMMENTATOR

音楽にコメントするシステム

吉井 和佳 後藤 真孝

研究の背景

○ 「音楽」を「言語」で表現する重要性

● 言語：音楽を介したコミュニケーションの手段

楽曲全体に対するタグ

The screenshot shows a YouTube video player interface. The video title is "どこのでもいっしょ公式チャンネル「どこのでもいっしょ@公式チャンネル」とろちゃ”オープン!! 「トロ」たち人気キャラクターの素材をニコニコ・コモンズにて好評配信中". The video content shows a person playing a guitar. The video player has a progress bar at the bottom showing 00:39 / 04:33. On the right side, there is a comment list with the following entries:

| 再生時 | コメント |
|-------|------------------------------|
| 02:15 | きいてきもちいいiiiiiiiiiiiiiiiiiiii |
| 02:14 | ...(°▽)o≡°ヤ☆パ☆!ヤ☆パ☆! |
| 02:13 | W!S!E!D!A! |
| 02:13 | 腕反則!www |
| 02:12 | 上手いですね |
| 02:09 | すきw |
| 02:04 | きもちいいiiiiiiiiiiiiiiiiiiii |
| 02:03 | 服のチャック付いてる^^ |
| 01:59 | アレンジすごくシンプルなのに耳に残る |
| 01:52 | カッコいい ギターが |
| 01:51 | ニコにまたあげてくれ--- |
| 01:48 | あこがれます!!!!!!!!!!!!!!!!!!!! |
| 01:47 | 弾幕いいぞ |
| 01:44 | ...(°▽)o≡°ヤ☆パ☆!ヤ☆パ☆! |
| 01:43 | 肉質いいなw |

ある時刻に対して付与されたコメント

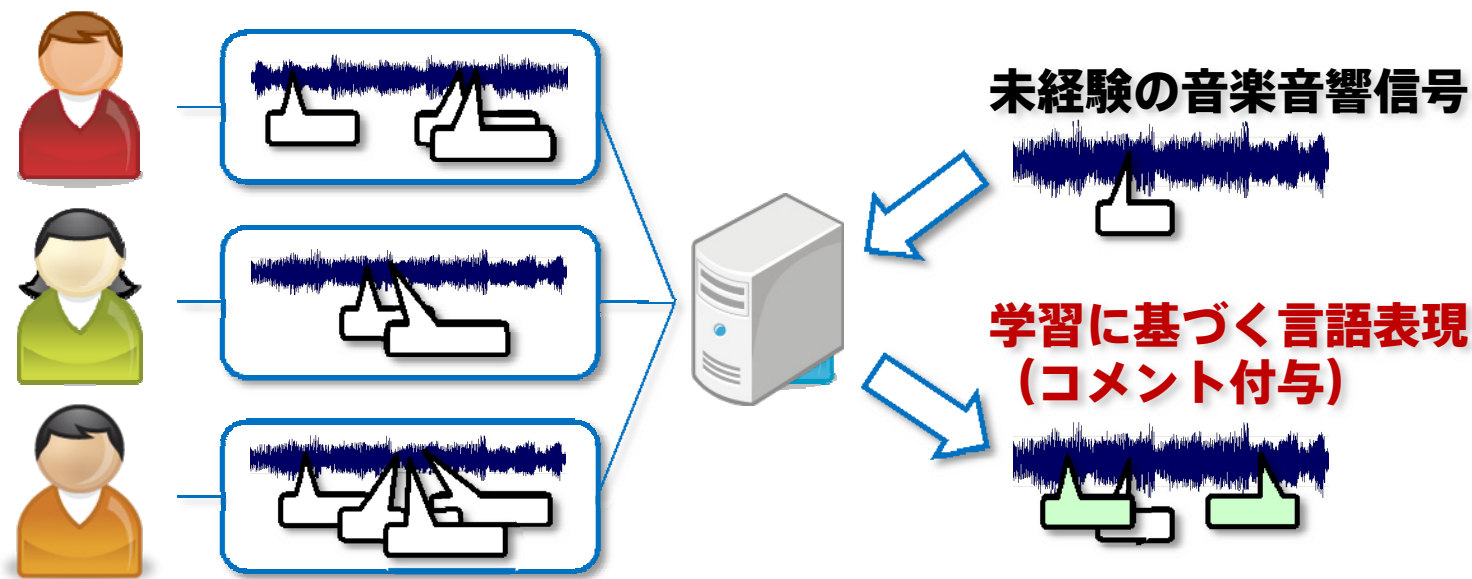
多数のユーザが同じ楽曲にコメントしあいながら一緒に鑑賞している感覚

研究アプローチ

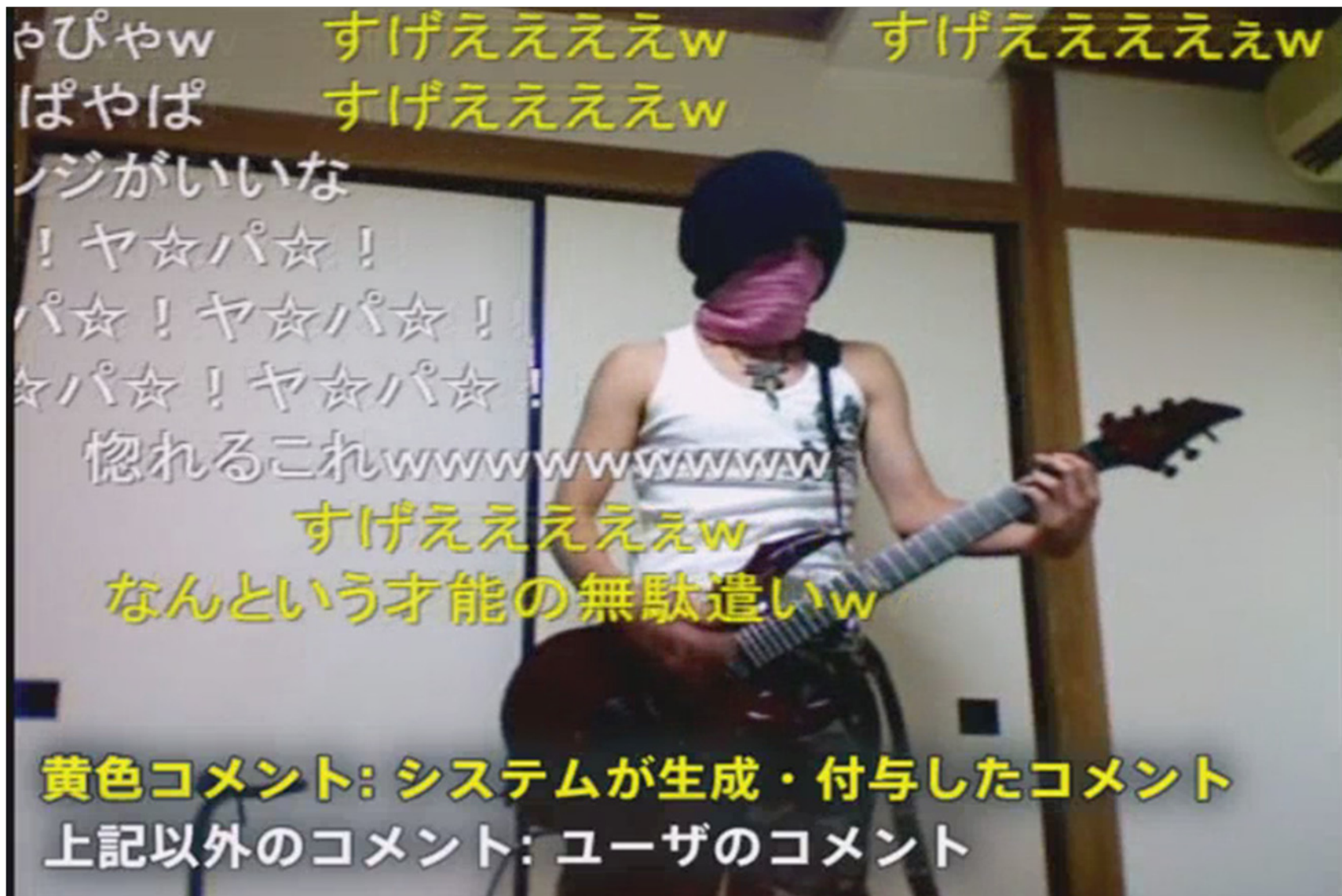
○ 音楽と言語を対応付ける数理モデルの構築

● 機械学習に基づく音楽音響信号へのコメント付与

- コメント付きの多数の音楽音響信号から音楽と言語の対応付けモデルを学習
- 新たに与えられた音楽音響信号に対して適切な時刻を選んで適切なコメントを付与



デモンストレーション



わぴやw

すげええええw

すげえええええw

ぱやぱ

すげええええw

マジがいいな

！ヤ☆パ☆！

パ☆！ヤ☆パ☆！

☆パ☆！ヤ☆パ☆！

惚れるこれwwwwwwwwwwwwwwww

すげえええええw

なんという才能の無駄遣いw

黄色コメント: システムが生成・付与したコメント

上記以外のコメント: ユーザのコメント

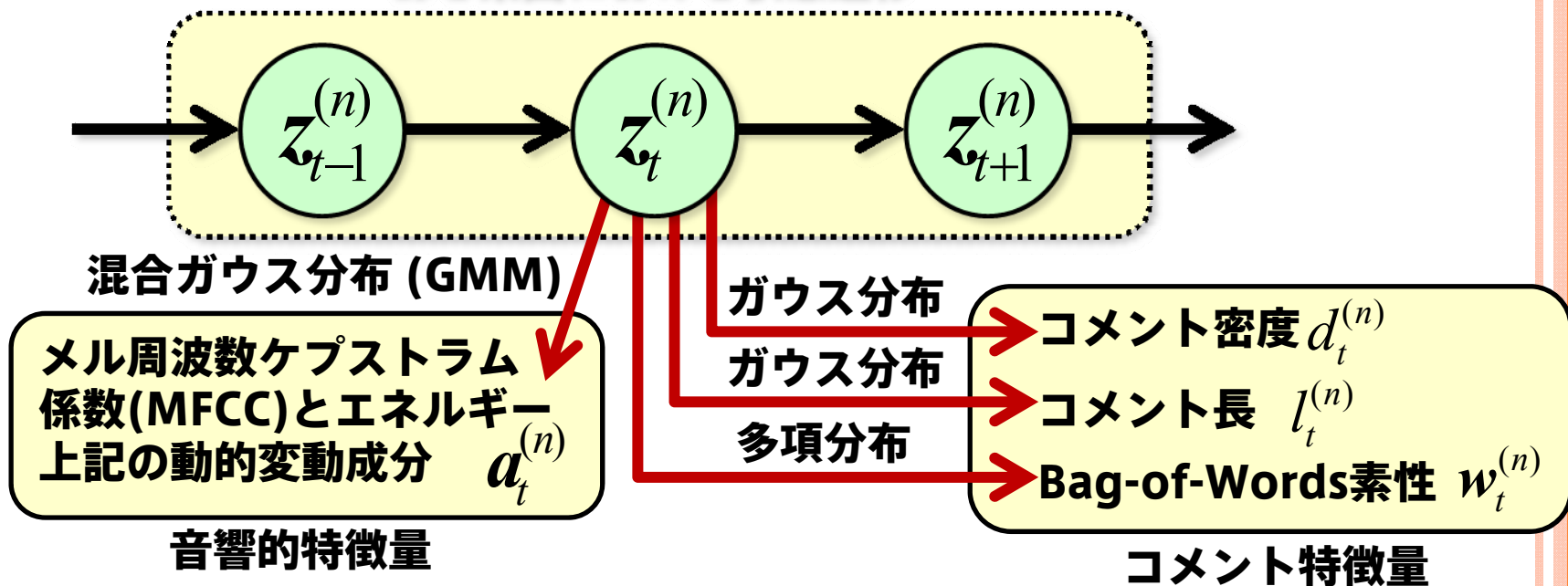
コメント生成モデルの学習

○モデルが備えるべき要件

- 4つの特徴量を同時にモデル化できること
- 特徴量の時系列をモデル化できること
- 4つの特徴量の動的なふるまいは共通であること

→ 隠れマルコフモデル(HMM)の利用

ある楽曲における状態遷移



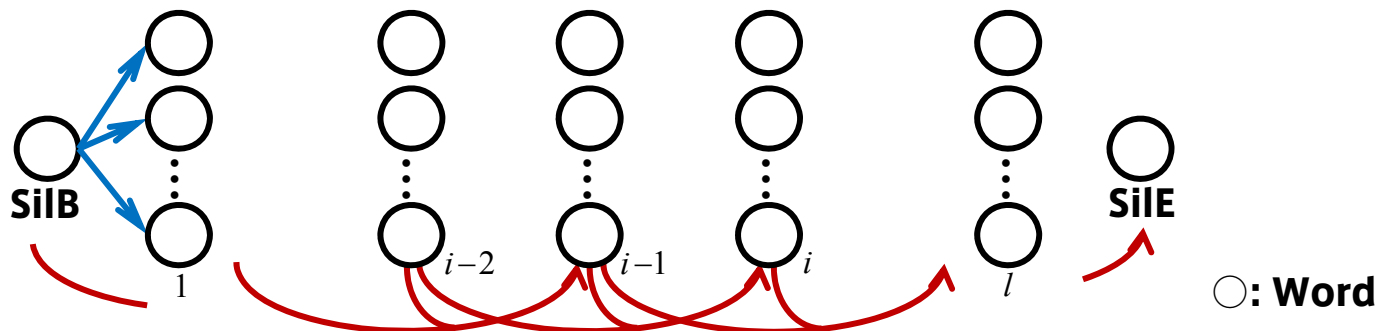
尤度最大化原理に基づく学習と生成

○学習フェーズ

- HMMパラメータを最尤推定
 - EMアルゴリズムを利用

○生成フェーズ

- 与えられた音響信号に対して状態系列をデコード
 - 最尤のBag-of-Words素性が得られる
- コメント文を生成
 - 上記Bag-of-Words確率と言語モデルとを統合
 - ビタビ探索





LYRICSYNCHRONIZER

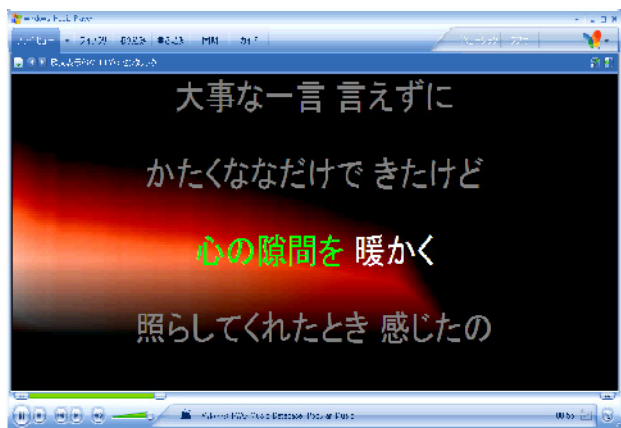
音楽と歌詞の対応付けシステム

藤原 弘将 後藤 真孝 緒方 淳
駒谷 和範 尾形 哲也 奥乃 博

研究の背景

○ 歌詞は音楽の魅力にかかわる重要な要素

- 歌詞カードを眺めつつ音楽を鑑賞
- カラオケのような歌詞表示を行うプラグインを利用
 - 人手で音楽と歌詞の対応付けが必要だった



■ 掲示板に戻る ■ 全部 1- 最新50
タイムタグ付き歌詞ファイル総合スレ 10曲目@アニソ

1: ななしはさお@オマエモナゆりかご会 : 2005/12/31(土) 12:06:39 ID:DTi+sBX1
自作や拾い物を持寄って楽しみましょう。

前スレ
タイムタグ付き歌詞ファイル総合スレ 9曲目@アニソ
<http://anime2ch.net/test/read.cgi/asong/1114600876/>
関連スレは>>2-4あたり

46: 名無し3.EXE: 03/09/15 21:40 ID:xKGnbaI3
ネタ吉いですが、てゆか既出ですか？

BUMP OF CHICKEN 天体観測 (2001.3.14 on sale)

[00:00:00]天体観測[00:00:00][00:10:77]
[Single 「天体観測」収録曲]
[00:10:78]作詞・作曲:藤原基央/編曲:BUMP OF CHICKEN[00:10:86][00:22:37]
歌:BUMP OF CHICKEN
[00:22:38]
[00:30:95]午前[00:31:13][00:31:61]二時[00:31:86][00:32:37]フミキリに[00:33:30] [0
[00:35:75]だった[00:36:23][00:36:57]
[00:36:76]へ[00:36:89]ルトに[00:37:81]結んだ[00:38:86]ラジオ[00:39:27] [00:39:78]
[00:42:01][00:42:39]

非常に手間ひまがかかる

研究アプローチ

○ 自動で音楽と歌詞を対応付けたい

● 音声認識技術との関連

- 「発話区間」において「話声」と「テキスト」を対応付ける技術は存在

- 単独発話による単独音

- 「歌声区間」において「歌声」と「テキスト」を対応付けるにはどうすれば？

- 複数音源による混合音

● 本研究での処理の流れ

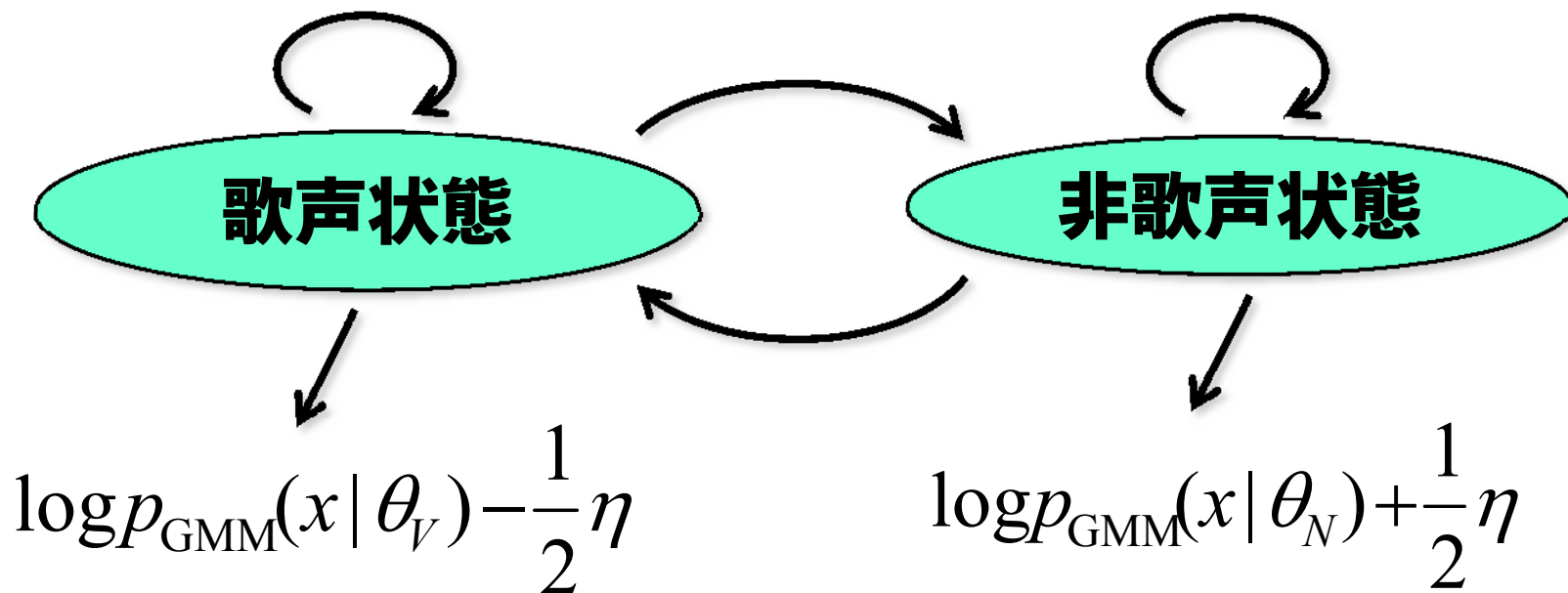
1. 混合音からメロディを分離
 - PreFEst [後藤] でF0推定+正弦波重畳合成
2. **メロディ中の歌声区間を推定**
3. **歌声と歌詞テキストとをアラインメント**
 - 話声モデルを歌声モデルに適応

デモ

歌声区間推定

○ 歌声状態と非歌声状態を行き来するHMM

- 出力パラメータはラベル付きデータから学習
- 遷移確率は実験的に設定



η を調整することでバランスをコントロール

適合率重視 or 再現率重視



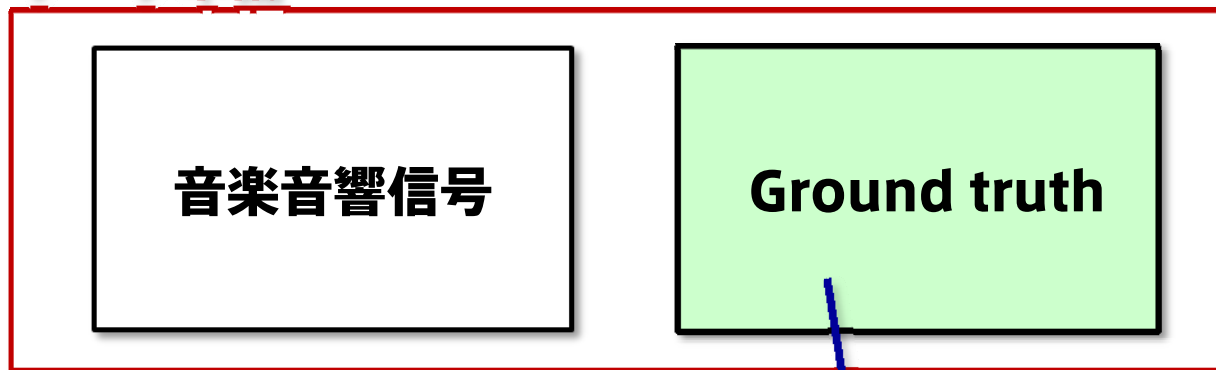
音楽情報処理のすすめ

困ったぞ

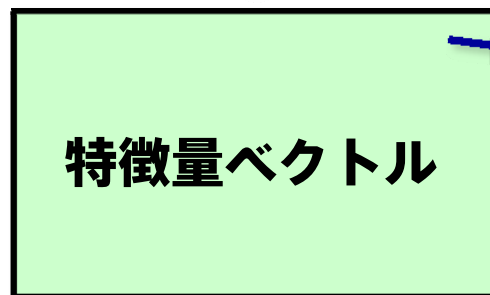
○ 音楽情報処理の研究はおもしろそうだが
門外漢なところが多くて難しそう

- 音楽データを対象にして自分の得意技を試したいが、
「データ準備」や「データ解析」が大変

データ準備



データ解析

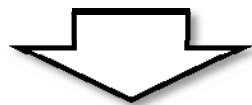


これらを扱うのは得意なのに…
機械学習, パターン認識,
多変量解析, 検証技術 etc

さあ、始めよう

○新規参入者に対する間口が広い分野

- **お金も時間もかからない**
 - **音楽データの準備**
 - **研究用データベースが無償で公開**
 - **Ground truthつき**
 - **共通データによるコンテストの開催**
 - **音楽データの解析**
 - **音楽特徴量が抽出できるフリーツールが公開**
- **自分の得意技が必ずどこかで生きる**
 - **さまざまな研究テーマが存在**



まずは自分の得意な部分のみに専念できる！

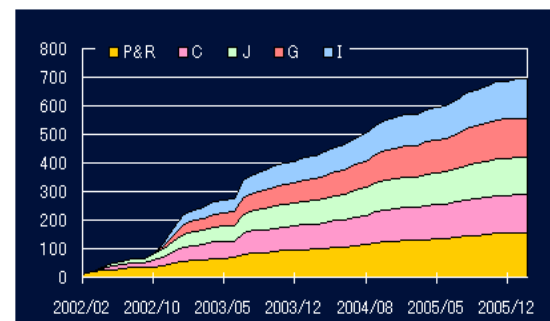
音楽データの準備

○ 研究目的で自由に利用可能なものが存在

● RWC研究用音楽データベース

著作権問題なし

- ポピュラー音楽データベース (100曲)
- 著作権切れ音楽データベース (15曲)
- クラシック音楽データベース (50曲)
- ジャズ音楽データベース (50曲)
- 音楽ジャンルデータベース (100曲)
- 楽器音データベース (50楽器・約150個体)



配布件数

● AIST Annotation for the RWC Music Database

- 上記データベースに対する人手によるアノテーション情報
 - ビート構造
 - メロディライン
 - サビ区間
- 音響信号に同期したMIDIデータ

極めて貴重!

音楽データの解析

○ 特徴量を抽出できるツールが存在

• Marsyas [Tzanetakis 2002]



- プラットフォーム非依存
- 低次の音響的特徴量が抽出可能
 - スペクトル重心, ゼロ交差率など

• MIRtoolbox [Olivier 2007]



- MATLAB用のツールボックス
- 低次の音響的特徴量が抽出可能
- 認識系のアルゴリズムも実装
 - 発音時刻検出
 - テンポ推定

**低次の音響的特徴量がよく利用される
→ 楽曲内容をBag-of-Featuresで表現**

腕試しがしてみたいなら

○ 認識精度コンテストMIREX



- 共通データセット上で手法の性能を比較
 - TRECの成功を受けて2005年から開始

TREC: テキスト検索手法コンテスト

- 大規模なデータを利用して性能を測定
- 6年間で検索システムの性能が2倍に向上

- さまざまなトラックが存在
 - ジャンル識別・ムード識別
 - 発音時刻検出
 - 基本周波数推定
 - カバーソング同定
 - コード認識
 - スコアアラインメント など

| Participant | Average Classification Accuracy |
|-------------|---------------------------------|
| GP1 | 63.67% |
| GT1 | 55.00% |
| GT2 | 52.50% |
| GT3 | 58.20% |
| HW | 30.33% |
| KL | 49.83% |
| LRPPI1 | 56.00% |
| LRPPI2 | 55.50% |
| LRPPI3 | 54.50% |
| LRPPI4 | 55.50% |
| ME1 | 50.33% |
| ME2 | 50.00% |
| ME3 | 49.67% |

WEB上で結果公開
(ムード識別結果)

基礎理論に強い研究者がたくさん

研究者⁺

目次: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A⁺

- 安倍 直樹 得意技:機械学習, 計算論的学習理論
- Agrawal, Rakesh 奥義:Apriori, 得意技:バスケットデータ, プライバシー保護データマイニング
- 赤穂 昭太郎 (あかほ) 朱鷺の社Wiki管理人, 得意技:カーネル, EMアルゴリズム
- 有村 博紀 得意技:データマイニング, 頻出パターンマイニング
- 麻生 英樹 得意技:ニューラルネット, 知能システム
- 甘利 俊一 必殺技:情報幾何, 得意技:特異モデル, ニューラルネット
- 安藤 晋 得意技:遺伝的アルゴリズム, はずれ値検出

B⁺

- Barto, Andrew 得意技:強化学習
- Bishop, Christopher M. 得意技:ニューラルネット, パターン認識
- Blei, David M. 必殺技:latent Dirichlet allocation, 得意技:ベイズ推定
- Blum, Avrim 必殺技:共訓練, 得意技:半教師あり学習, 計算論的学習理論
- Bottou, Léon 得意技:大規模計算, 確率的勾配降下法, SVM
- Breiman, Leo 必殺技:CART, バギング, 得意技:機械学習, アンサンブル学習
- Burges, Chris 得意技:自然言語処理

C⁺

- Caruana, Rich 得意技:機械学習, 帰納轉移, アンサンブル学習
- Chakrabarti, Soumen 得意技:Webマイニング
- Clifton, Christopher W. 必殺技:プライバシー保護データマイニング
- Cohen, William W. 得意技:自然言語処理

アプリケーションとして
音楽情報処理分野に
食指を伸ばし始めた!

HDP-GMMによる
楽曲類似度計算
@ISMIR 2008

強者ども、きたれ

○ 機械学習が活用できる場面は多い

● 最近の優れた機械学習法による大幅な精度改善の期待

- Bag-of-Features + GMM の限界の打破
- 時系列のモデル化・類似度計算

○ 新たな研究テーマの開拓にも期待

● 優れた機械学習法で初めて実現可能なものがあるはず

- 学習データが不十分で汎化能力のあるモデルが学習できなかった問題
- 単純なクラス分類問題以外に適用できないか？



**いまこそ機械学習という武器を手に
音楽情報処理に殴りこむチャンス！**

参考文献

- 吉井 和佳, 後藤 真孝, 駒谷 和範, 尾形 哲也, 奥乃 博: "ユーザの評価と音響的特徴との確率的統合に基づくハイブリッド型楽曲推薦システム", 情報処理学会 音楽情報科学研究会 研究報告 2006-MUS-066, Vol. 2006, No. 90, pp. 45-52, August 2006.
- 吉井 和佳, 後藤 真孝, 駒谷 和範, 尾形 哲也, 奥乃 博: "楽曲推薦システムの効率性とスケーラビリティの改善のための確率的推薦モデルのインクリメンタル学習法", 情報処理学会 音楽情報科学研究会 研究報告 2007-MUS-071, Vol. 2007, No. 81, pp. 19-26, August 2008.
- 吉井 和佳, 後藤 真孝: "MusicThumbnailer: 音響的特徴に基づく楽曲のサムネイル画像生成手法", 情報処理学会 音楽情報科学研究会 研究報告 2008-MUS-076, Vol. 2008, No. 78, pp. 57-62, August 2008.
- 吉井 和佳, 後藤 真孝: "MusicCommentator: 音楽に同期したコメントを自動生成するシステム", 情報処理学会 音楽情報科学研究会 研究報告 Vol. 2009-MUS-81, No. 20, July 2009.
- 藤原 弘将, 後藤 真孝, 緒方 淳, 駒谷 和範, 尾形 哲也, 奥乃 博: "音楽音響信号と歌詞の時間的対応付け手法: 歌声の分離と母音のViterbiアラインメント", 情報処理学会 音楽情報科学研究会 研究報告 2006-MUS-066, Vol. 2006, No. 90, pp. 37-44, August 2006.

連絡先

○ 吉井 和佳

- **k.yoshii@aist.go.jp**
 - **ご質問・ご意見・ご感想等ありましたら
メールしてくださるとうれしいです**
- **産業技術総合研究所 情報技術研究部門**
- **メディアインタラクション研究グループ 研究員**