

# 楽器音認識技術を用いた 音楽の可視化

北原 鉄朗<sup>\*1,\*2</sup>, 後藤 真孝<sup>\*1,\*3</sup>,  
奥乃 博<sup>\*1,\*4</sup>, 片寄 晴弘<sup>\*1,\*2</sup>

(<sup>\*1</sup>JST CrestMuseプロジェクト, <sup>\*2</sup>関西学院大学,  
<sup>\*3</sup>産業技術総合研究所, <sup>\*4</sup>京都大学)

# 始まるまえに:CrestMuseプロジェクトとは

## 時系列メディアのデザイン転写技術の開発

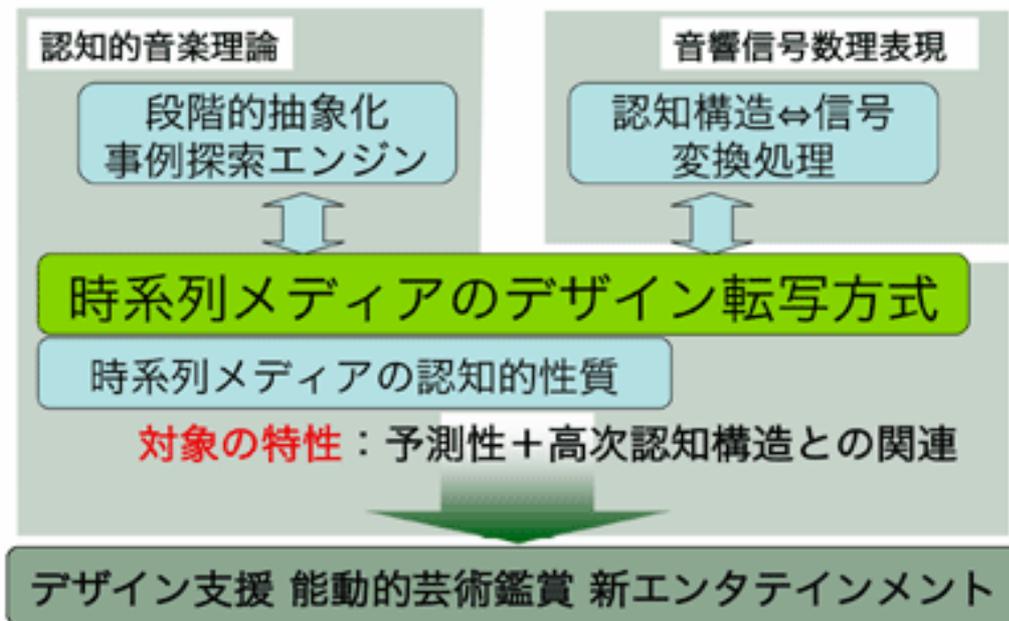
科学技術振興機構

戦略的創造研究推進事業 (CREST / さきがけ)

「デジタルメディア」領域 (研究総括: 原島先生)

片寄チーム  
の愛称

**CrestMuse**  
project



# 音楽の可視化

楽曲の中身を目でも確認できるようにする技術

## 音楽可視化の意義

エンターテインメント性の向上

- 見ていて楽しい

鑑賞能動性の向上

- 楽曲の中身を発見しながら聴く
- 音楽教育効果

検索効率性の向上

- 一覧性の付与, 楽曲サムネイルとしての効果

# 古典的可視化事例1：五線譜

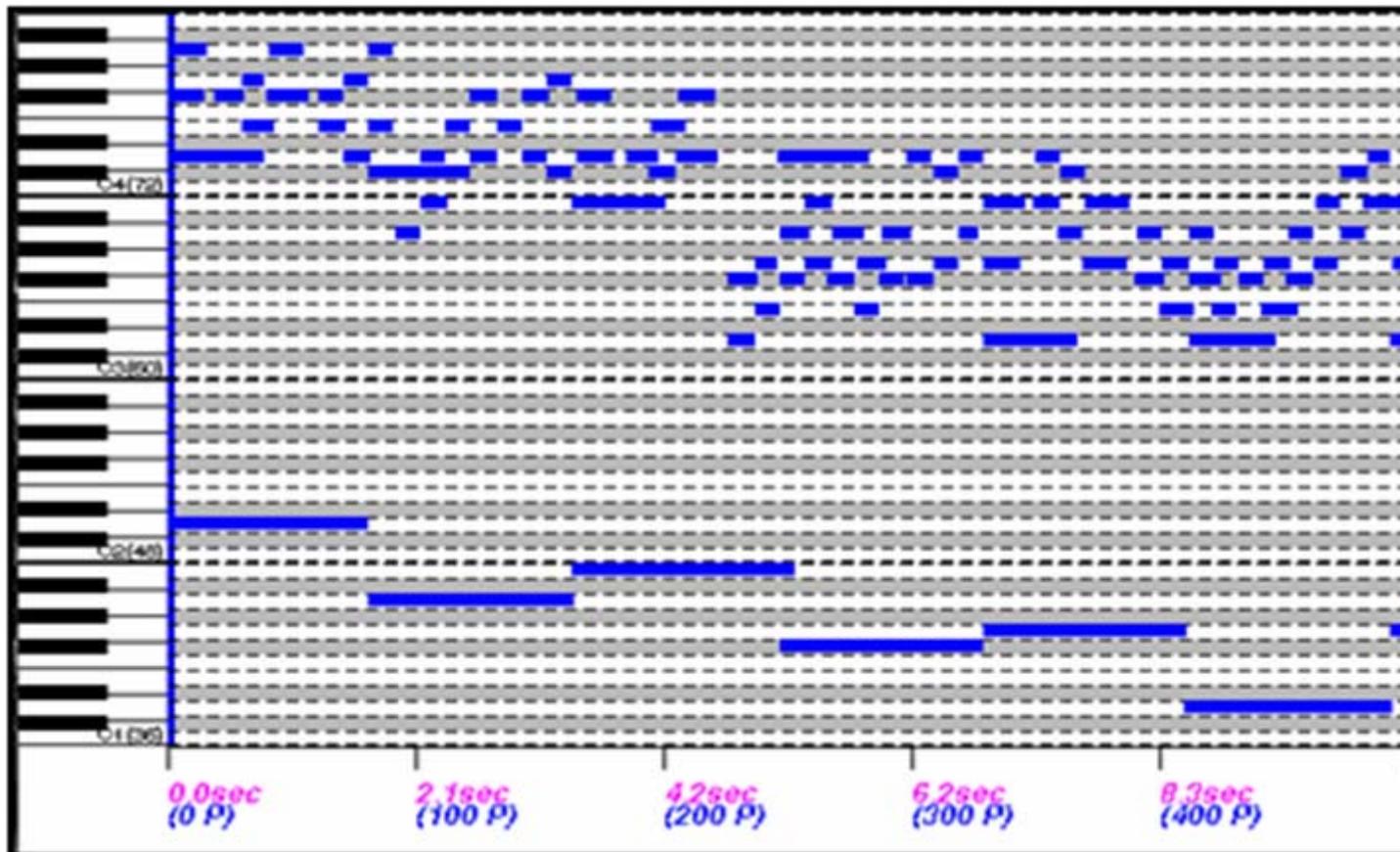
- 楽曲の記録・伝達向けでエンターテインメント性低い
- 音楽家同士の意思伝達手段としてポピュラー
- 音響信号からの自動作成困難



The image displays a musical score in 4/4 time, consisting of two systems of three staves each. The first system includes chord symbols: B♭, B♭, Cm, F9, Gm, B♭, Cm, and E♭. The second system includes chord symbols: B♭, B♭7, Cm, Cm, B♭, F9, B♭, and E♭. A box containing the number '5' is positioned at the beginning of the second system. The notation includes treble and bass clefs, a key signature of two flats, and various rhythmic values such as quarter, eighth, and sixteenth notes.

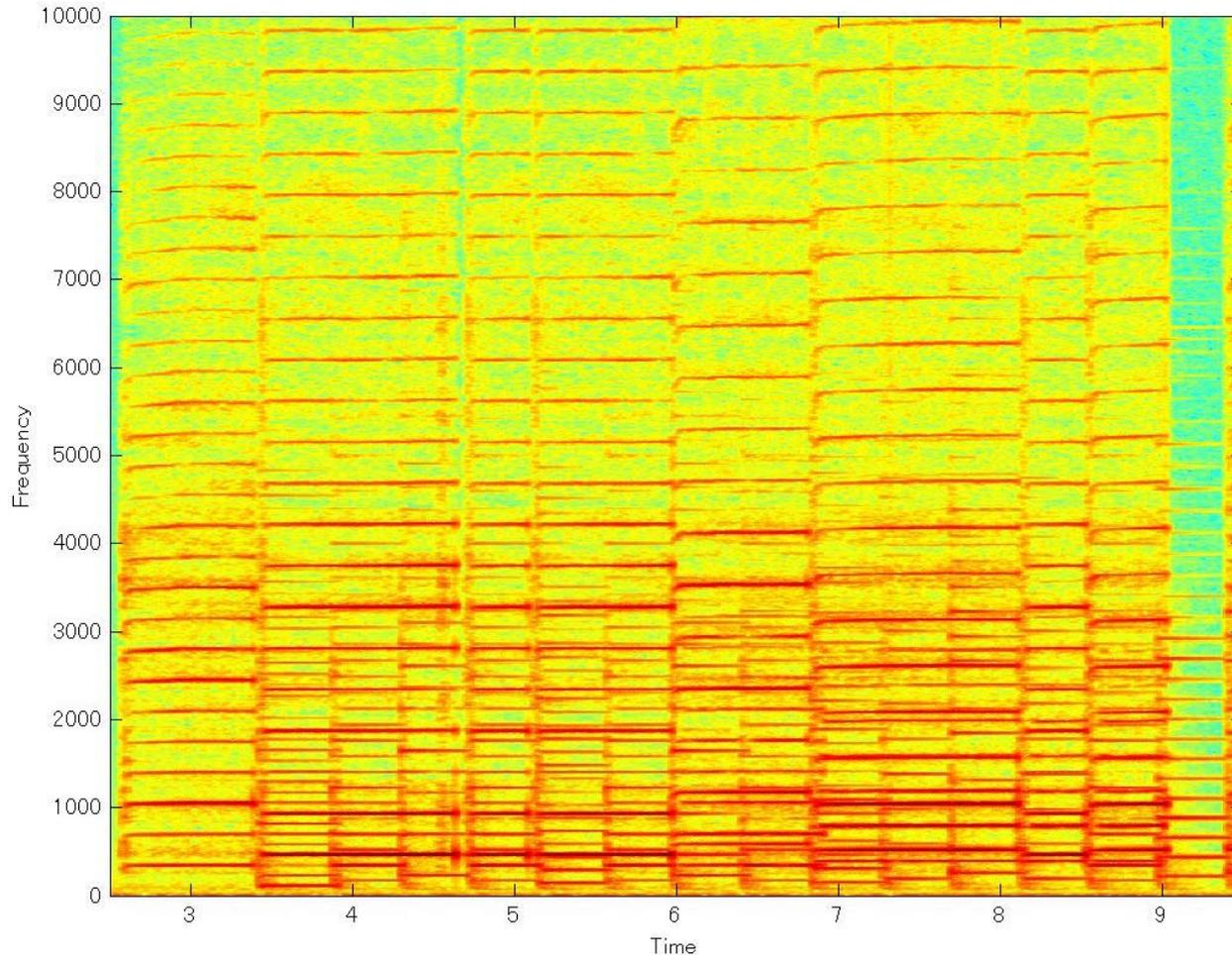
# 古典的可視化事例2:ピアノロール

- DTMで一般的
- 楽譜同様エンターテインメント性低い
- 楽譜同様自動作成困難



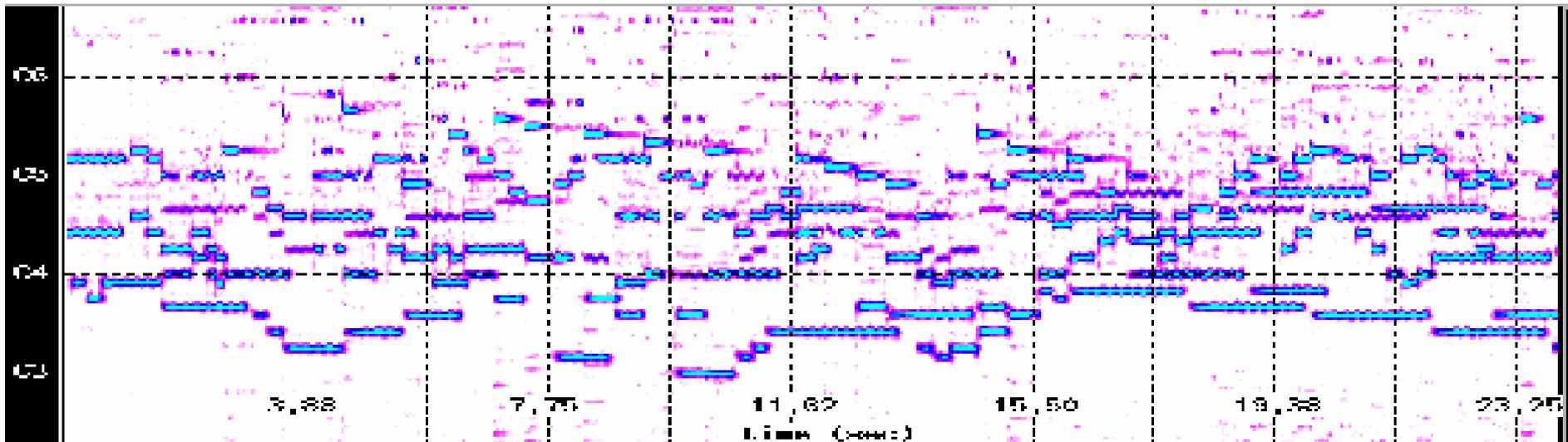
# 古典的可視化事例3:スペクトログラム

- 音響信号から容易に作成可能
- 音楽的内容の把握は困難



# 関連研究1: Specmurt [Sagayama'04]

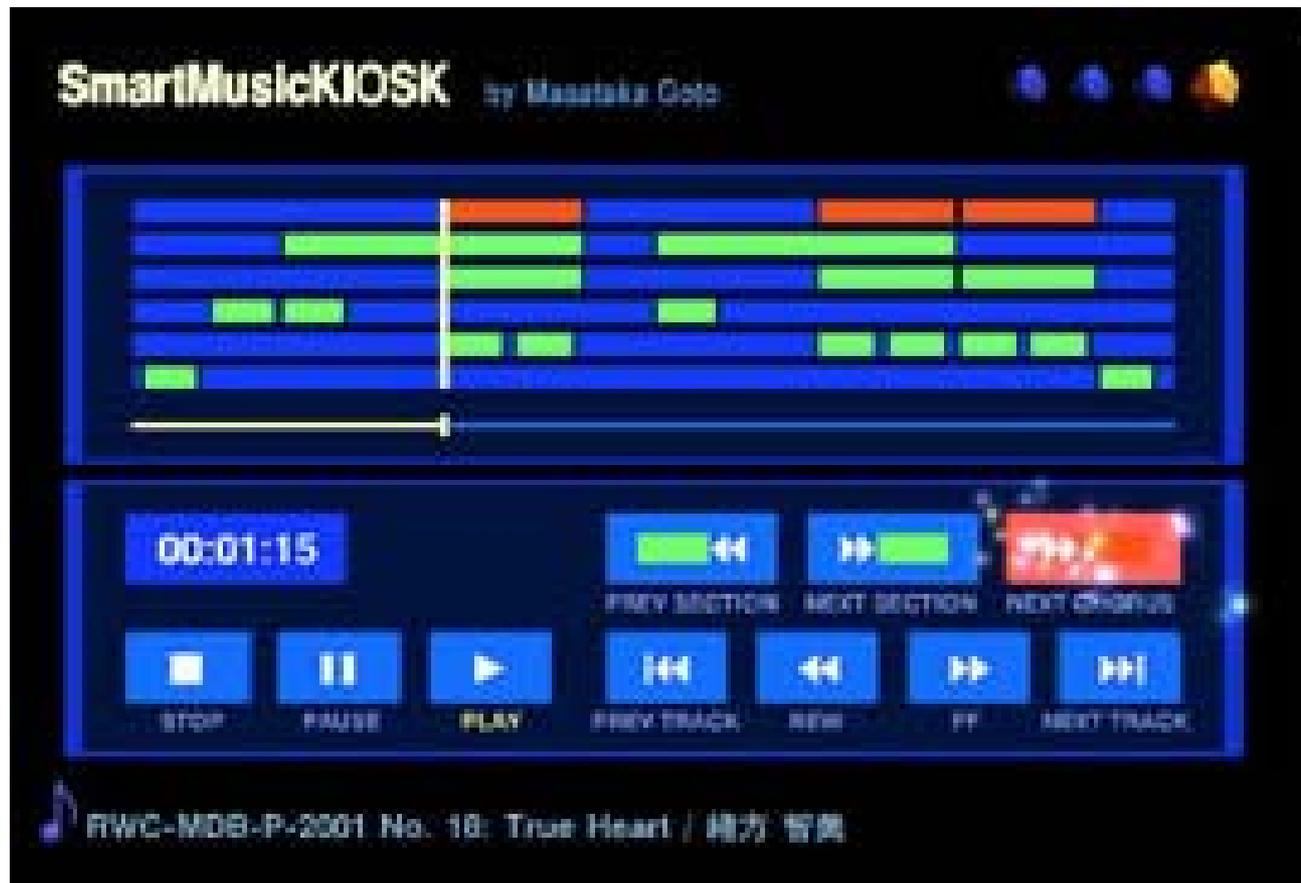
- スペクトログラムに典型的な調波構造スペクトルを逆畳み込みすることで倍音成分を抑制
- 元のスペクトログラムに比べピアノロールに近い



(b) *Specmurt Analysis showing fundamental frequencies*

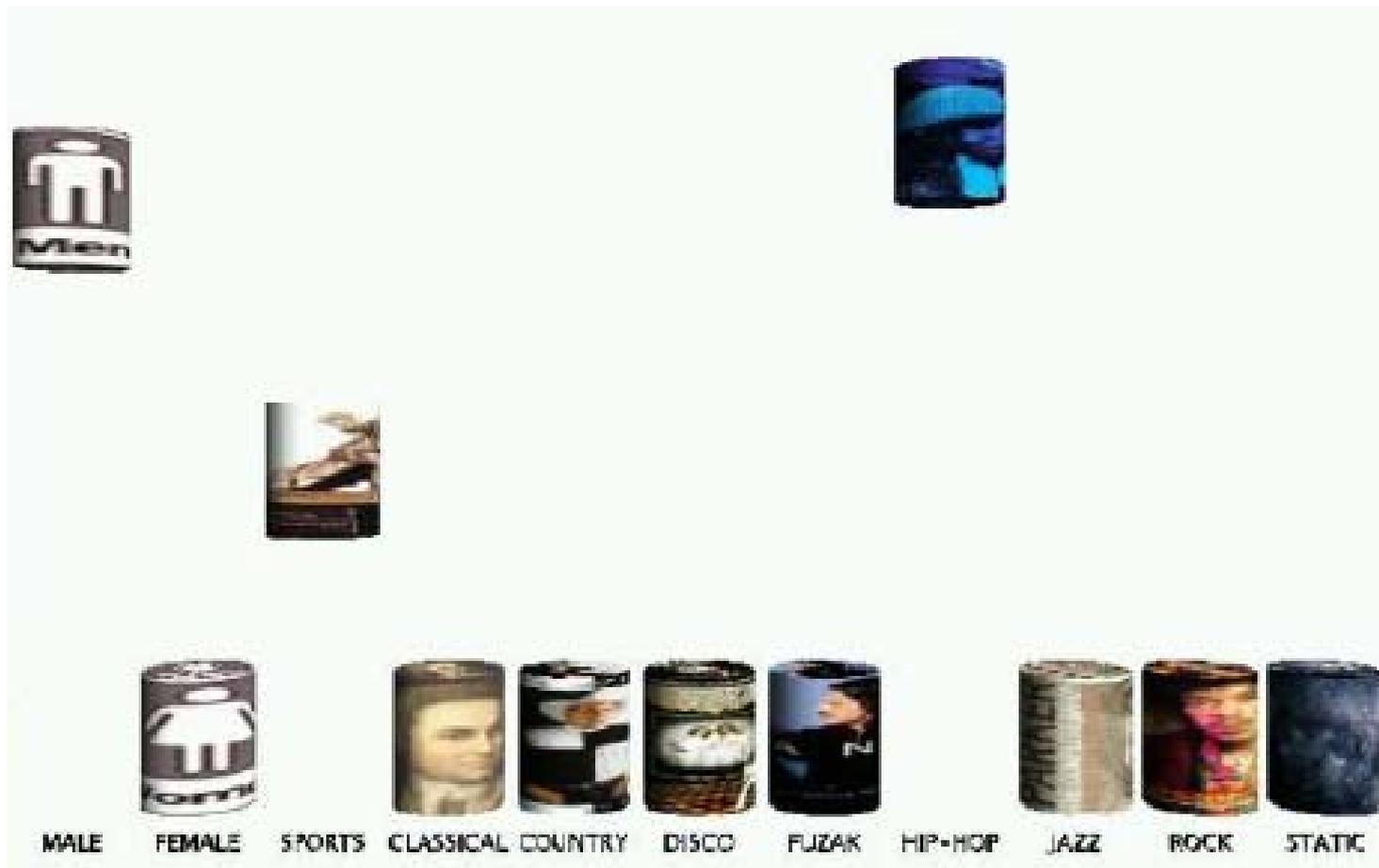
# 関連研究2: SmartMusicKIOSK [Goto'03]

- 楽曲の繰り返し構造 (e.g. Aメロ, Bメロ, サビ) を可視化
- 1ボタンでサビへジャンプ可能



# 関連研究3: GenreGramモニター [Tzanetakis'02]

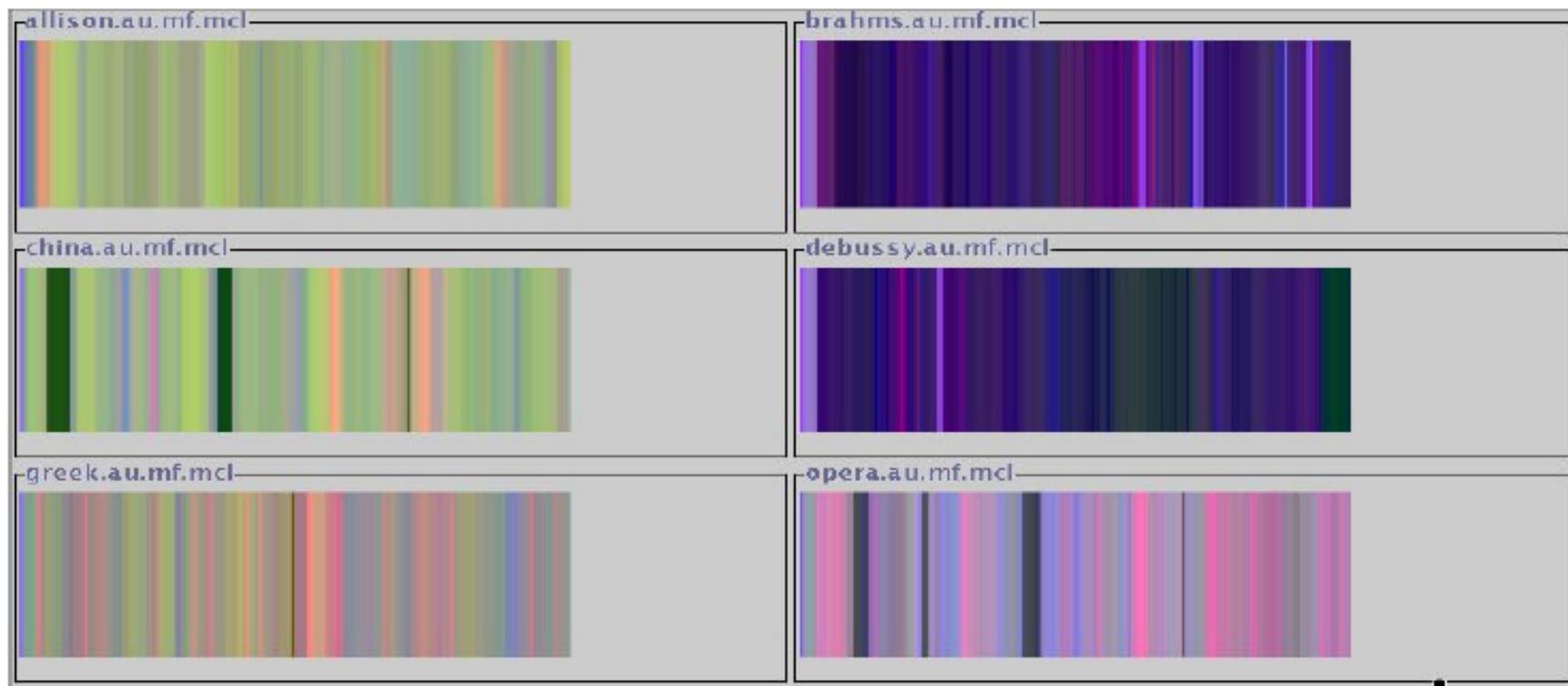
- ジャンル解析を行い, 各ジャンルの尤度に合わせて円柱が上下にリアルタイムに移動



Tzanetakis 博論'02 より引用

# 関連研究4: TimbreGramモニター [Tzanetakis'02]

- 音色(音響的特徴)の類似性に合わせて着色
- 「近さ」だけが可視化され「どの楽器か」は分ならず



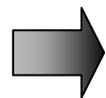
# 楽器構成可視化機能付き音楽プレイヤー

「どんな楽器による演奏か」を時々刻々と表示しながら楽曲を再生

我々の仮説

楽器構成は楽曲を特徴づける重要な要素

- 同一楽曲でも異なる楽器の演奏では印象変化



楽器構成は可視化でも重要な要素のはず

本システムの特徴

楽曲の楽器構成のアニメーション表示

楽器アイコンを見ながらのプレイリスト編集 etc.

# 実現方法

詳しくはIPSJ Journal  
Vol.48, No.1 (2007)で

基本的アイデア

対象楽器毎に各時刻・周波数における  
**楽器存在確率**を計算

**楽器存在確率**を2種類の確率の積に分解

$$p(\omega_i; t, f) = \underbrace{p(X; t, f)}_{\text{不特定楽器存在確率}} \underbrace{p(\omega_i | X; t, f)}_{\text{条件付楽器存在確率}}$$

**不特定楽器存在確率**

何らかの楽器音が  
存在する確率

**PreFEst** [Goto'99] (混合音  
用F0推定手法)を利用

**条件付楽器存在確率**

何らかの楽器音が存在する  
とすると, それが $\omega_i$ の確率

**隠れマルコフモデル** (音声  
認識で一般的)を利用

# 考察

## エンターテインメント性の向上

- イラストのアニメーションによる見て楽しい表現  
→飽きずに聴ける(聞き流しの防止)

## 鑑賞能動性の向上

- どんな楽器が演奏しているかを発見しながら聴く  
→音楽教育効果

## 検索効率性の向上

- 楽器アイコンによる楽器構成の一覧表示  
→楽曲の内容を想像しながらプレイリスト編集

# まとめ

## 楽器構成に基づく音楽可視化システムを紹介

- 楽器構成を自動的に推定
- アニメーションやアイコンとして可視化

最後に提言

もっとコラボを！

- 可視化などのエンターテインメント分野はコアな信号処理技術にとって絶好の応用先
- 高度な信号処理技術を利用できればエンターテインメントの幅が広がる
  - 両分野の研究者にとってWin-Winな関係