

Instagram

を用いた

楽器構成に基づく類似楽曲検索

北原 鉄朗* 後藤 真孝**

駒谷 和範* 尾形 哲也* 奥乃 博*

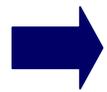
*京都大学大学院情報学研究科

**産業技術総合研究所

13-15 Sept. 2006

類似楽曲検索とは？

- ユーザが指定した楽曲に似た曲を提示
- 内容に基づく音楽情報検索 (Content-based Music Information Retrieval) の一形態
- 誰でも簡単に利用可能
- 課題：「何をもって似ていると判断するか」



楽器構成に着目

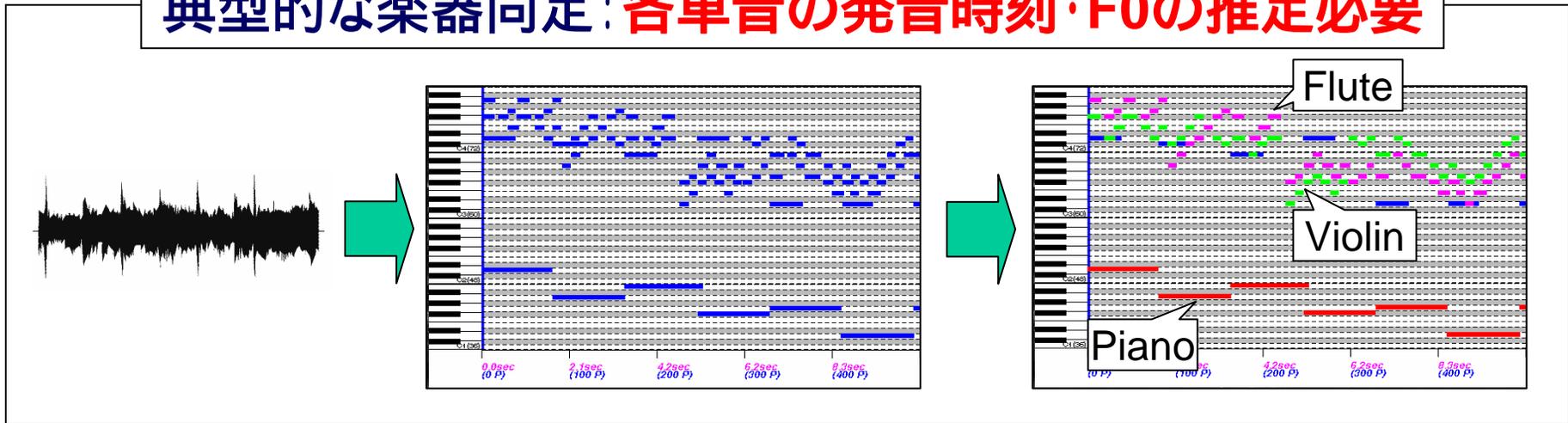
- 「どの楽器か」は楽曲の雰囲気には大きな影響
- ジャンルと楽器構成の関連

どうやって実現するか

i. 楽器同定して結果をマッチング?

➡ 楽器同定は難しい (特に単音単位の楽器同定)

典型的な楽器同定: 各単音の発音時刻・F0の推定必要



ii. 音色特徴量 (e.g. MFCC) を抽出して距離計算?

➡ ボイシングなどからも影響



Instrogramを用いた実現

Instrogramを用いた実現

Instrogram

[音講論'06春, ICASSP'06]

- 楽器存在確率を全時刻・全周波数にわたって計算したもの

楽器存在確率 $p(i; t, f)$

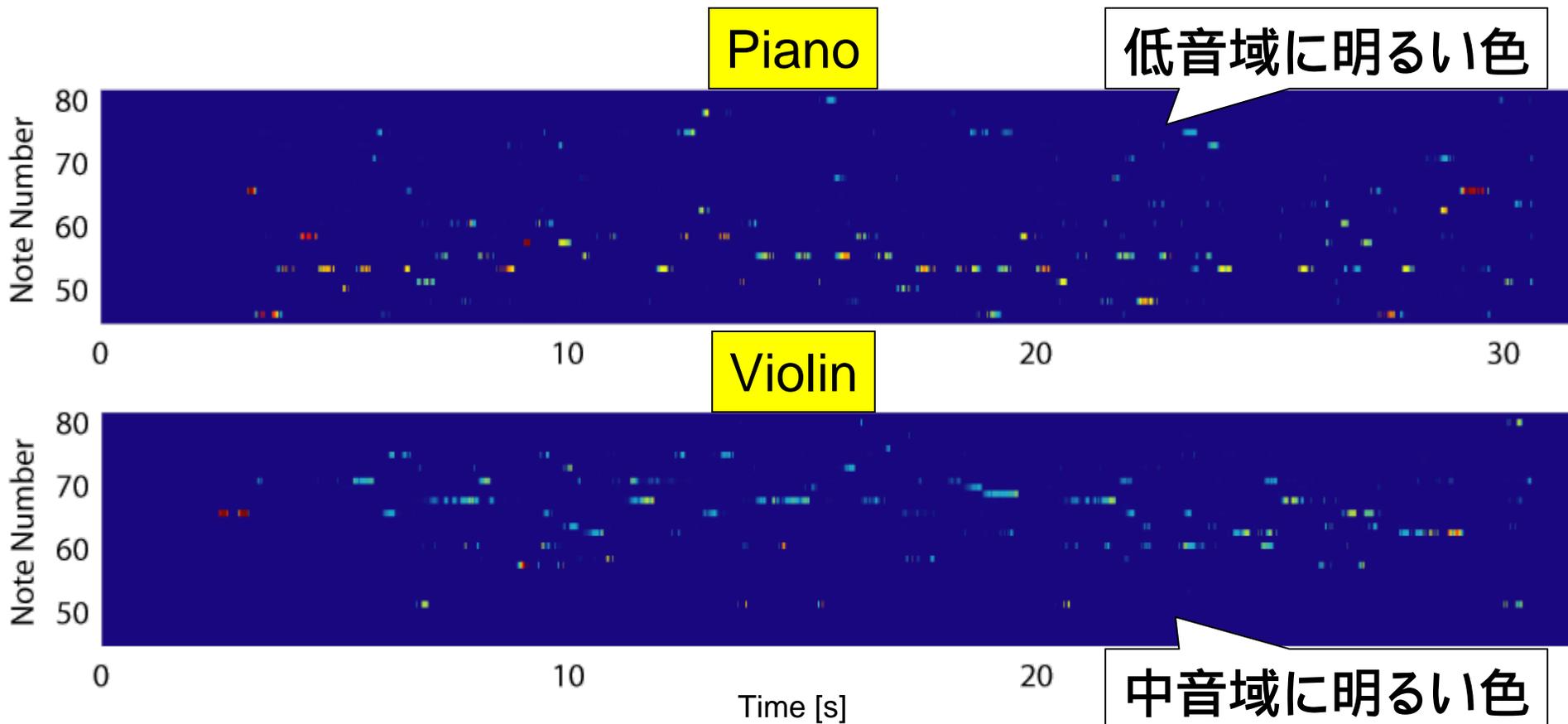
時刻 t において周波数 f をF0とする楽器 i の音が存在する確率

- 発音時刻・音高推定相当処理を確率計算に包含し, 明示的な前処理不要
- Instrogram間の類似度を定義

Instrogram (1/2)

楽器ごとに時間・周波数平面上に楽器存在確率を可視化

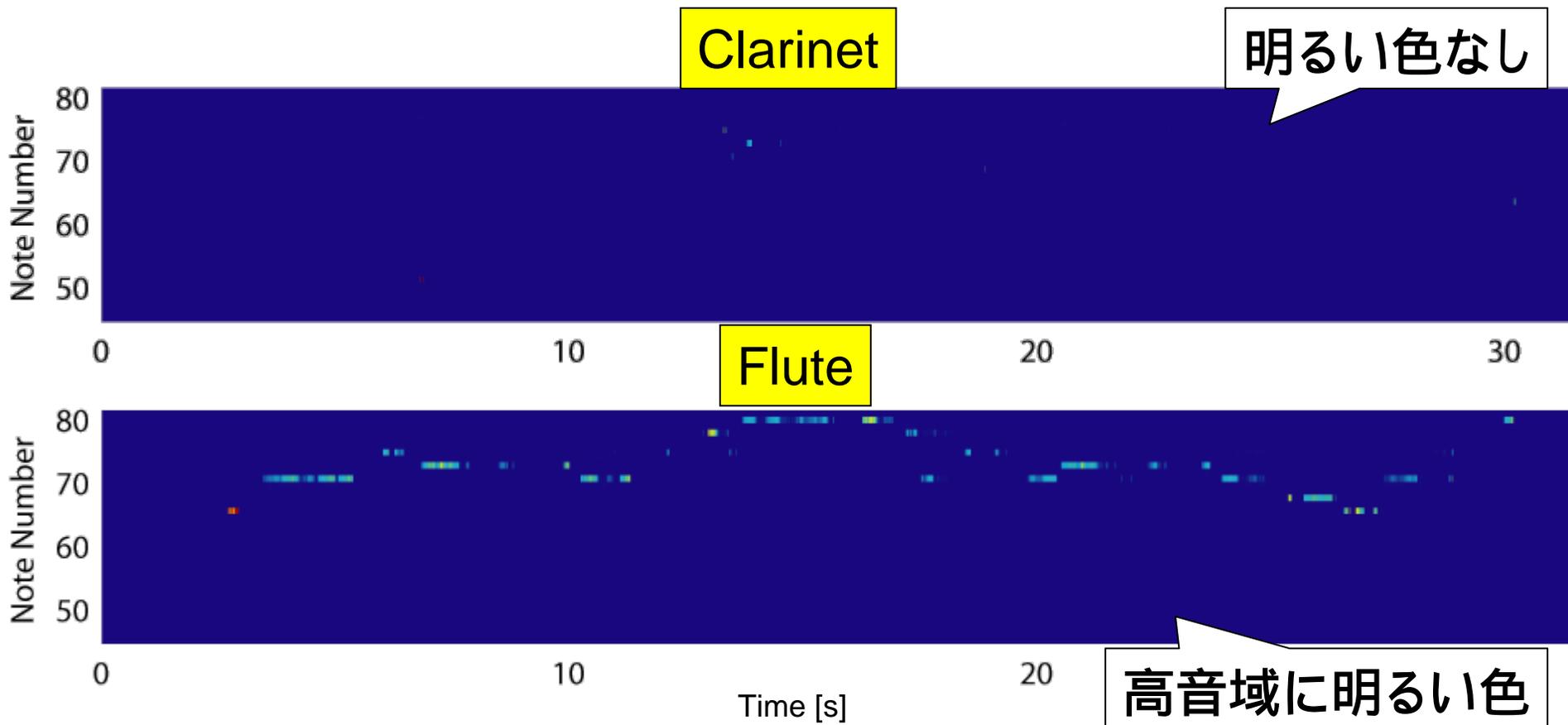
例：蛍の光 (Flute-Violin-Piano)



Instrogram (2/2)

楽器ごとに時間・周波数平面上に楽器存在確率を可視化

例：蛍の光 (Flute-Violin-Piano)



楽器存在確率の定式化

楽器存在確率 $p(\omega_i; t, f)$ を以下のように分解できる

$$p(\omega_i; t, f) = p(X; t, f) p(\omega_i | X; t, f)$$

$$\omega_i \cap X = \omega_i$$

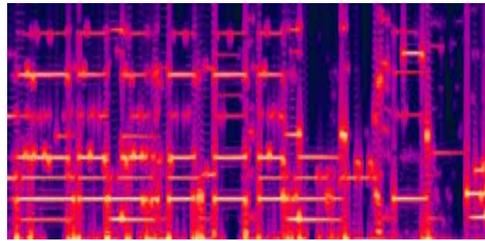
$X = \omega_1 \cup \dots \cup \omega_m$ (何らかの楽器が発音している事象)

- $p(X; t, f)$ **不特定楽器存在確率**
(t, f)に何らかの楽器音が存在する確率
- $p(\omega_i | X; t, f)$ **条件つき楽器存在確率**
(t, f)に楽器音が存在するとすると, それが楽器 i である確率

確率計算アルゴリズムの概要

pitchwise processing

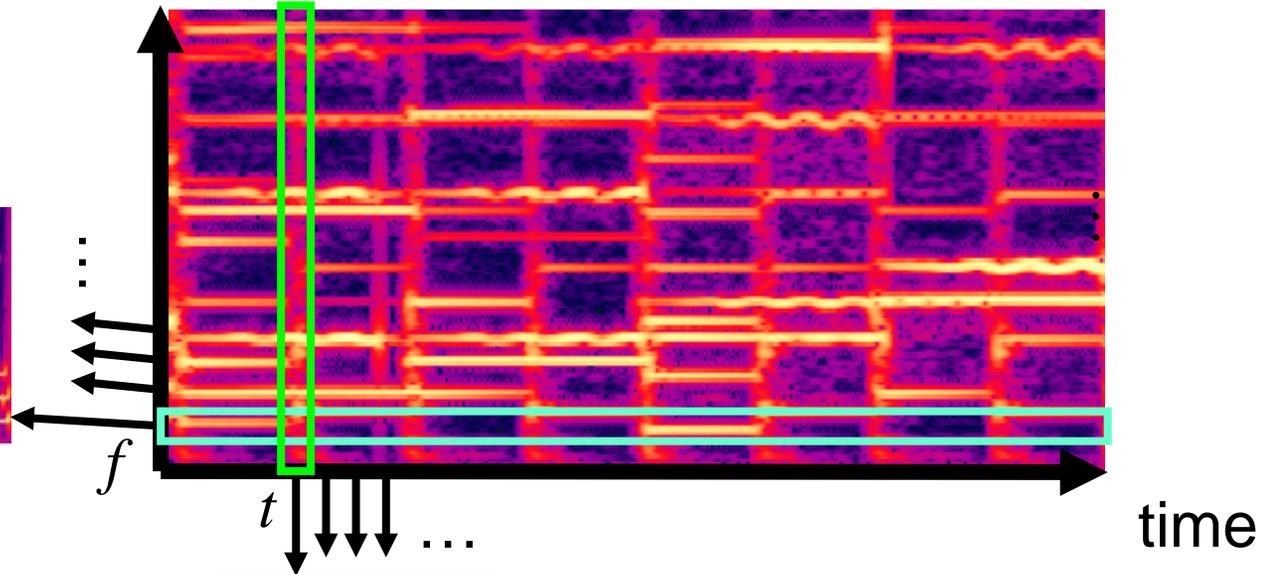
for each f
調波構造時系列



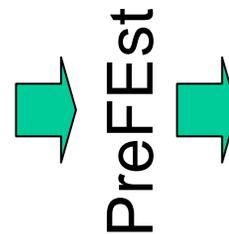
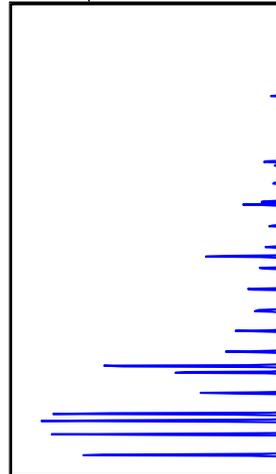
HMM

条件つき
楽器存在確率

freq. 入力音響信号のスペクトログラム



for each t
パワースペクトル



PreFest

不特定
楽器存在確率

timewise processing

Instrogram間の類似度計算

- 類似度計算のキーアイデア:

各時刻の楽器存在確率を特徴ベクトルとみなして、この時系列同士に対してDTW (DPマッチング)

時刻 t におけるベクトル p_t : 全楽器存在確率の結合

$$p_t = (p(\omega_1; t, I_1), p(\omega_1; t, I_2), \dots, p(\omega_m; t, I_N))'$$

2ベクトル p, q 間の距離: コサイン距離 $(p, q) = p' R q$
 $\text{dist}(p, q) = 1 - (p, q) / \|p\| \cdot \|q\|$ $\|p\| = \sqrt{(p, p)}$

上記距離尺度を用いてDTW

R : 正定値行列
要素間の関連性考慮可能
単位行列なら通常の内積

Instrogram作成実験 (実演奏)

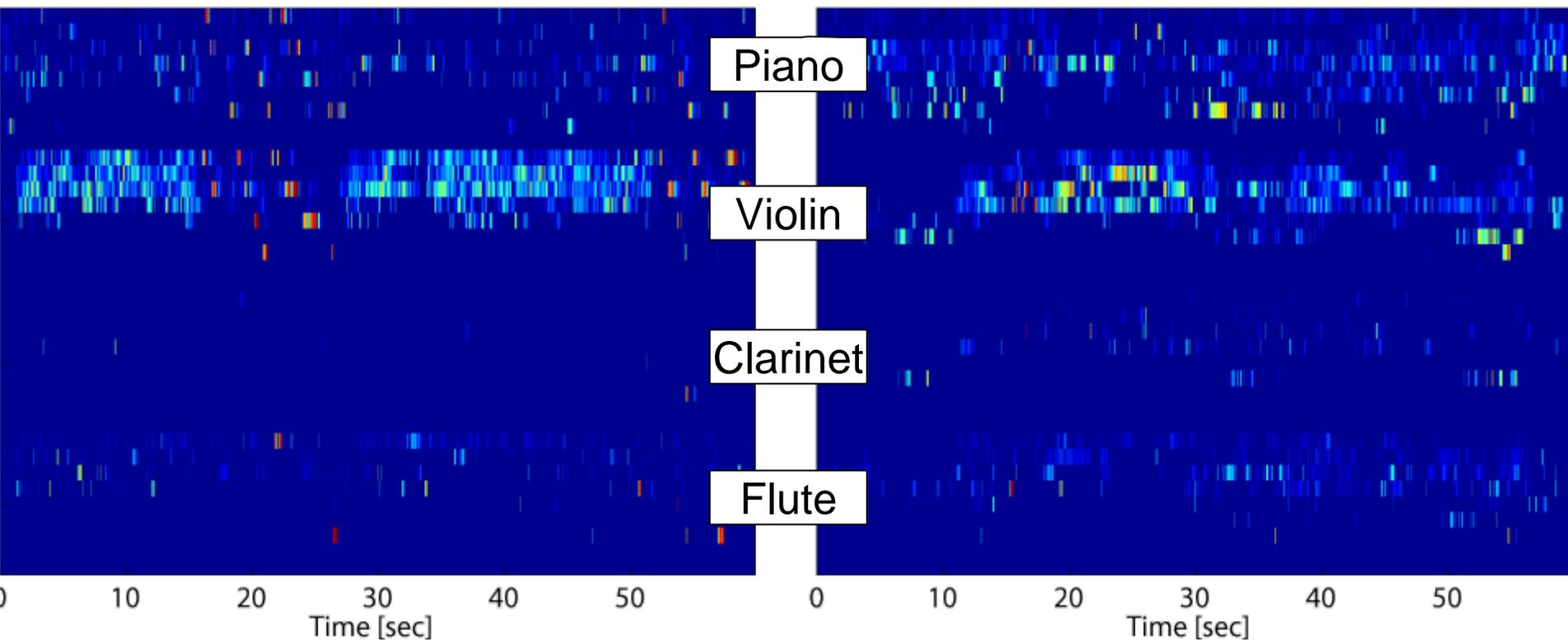
- 1～4重奏程度の実演奏(クラシック, ジャズ)
(RWC音楽データベースより)
- 学習データ: RWC-MDB-I-2001, NTTMSA-P1
で作成した切り貼り3重奏
- 対象楽器: Piano, Violin, Clarinet, Flute

Classical	(i) No. 12, 14, 21, 28	Strings
	(ii) No. 19, 40	Piano+Strings
	(iii) No. 43	Piano+Flute
Jazz	(iv) No. 1, 2, 3	Piano solo

実験結果 (1/2)

RM-C No.14 (Str.)

RM-C No.19 (Pf.+Str.)

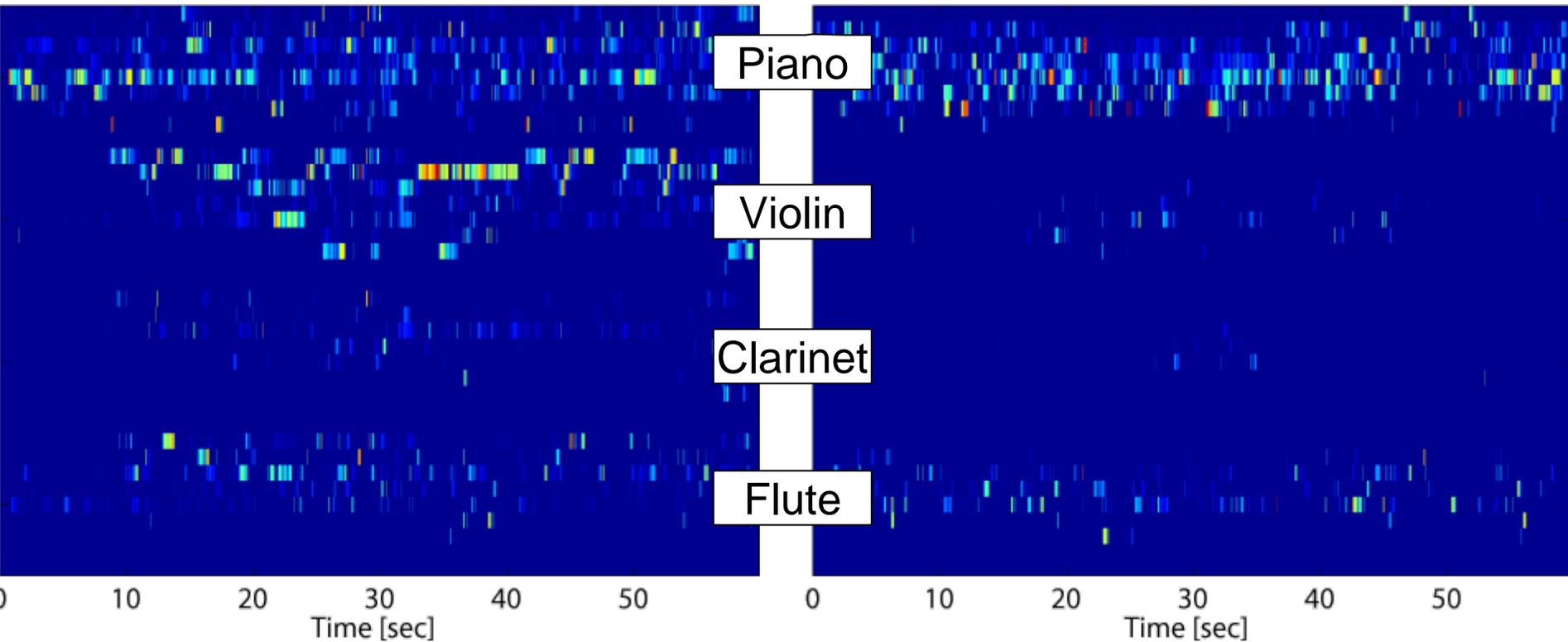


- Violinの存在確率が高い(明るい色)
- 左に比べて右のPianoの存在確率が高い

実験結果 (2/2)

RM-C No.40 (Pf.+Str.)

RM-J No.1 (Piano)



- 左：最初Pfだけで途中からStrが入ってくる
- 右：Pianoだけ存在確率が高い

Instrogram間類似度計算結果

- グループ内非類似度は多くが7000以下
- 弦楽曲-ピアノ曲間の非類似度はおよそ9000以上
 - ⇒ 実際の楽器構成を適切に反映
- 類似楽曲ベスト3
 - Instrogramの場合
 - 弦楽曲の類似楽曲ベスト3はすべて弦楽曲
 - 非弦楽曲の類似楽曲ベスト3はすべて非弦楽曲
 - MFCCの場合
 - 弦楽曲と非弦楽曲の混同しばしば

まとめ

- 目的：楽器構成に基づいた類似楽曲検索
 - 誰でも使える内容に基づく音楽情報検索
- Instrogramを用いて実現
 - 楽器構成を確率的に表現
 - 楽器存在確率間の連続的類似度尺度を定義
- 今後の課題
 - 検索システムとしての有用性の評価など