

音色空間の音高依存性を考慮した 楽器音の音源同定

北原 鉄朗[†] 後藤 真孝^{††} 奥乃 博[†]

[†]京都大学大学院情報学研究科知能情報学専攻

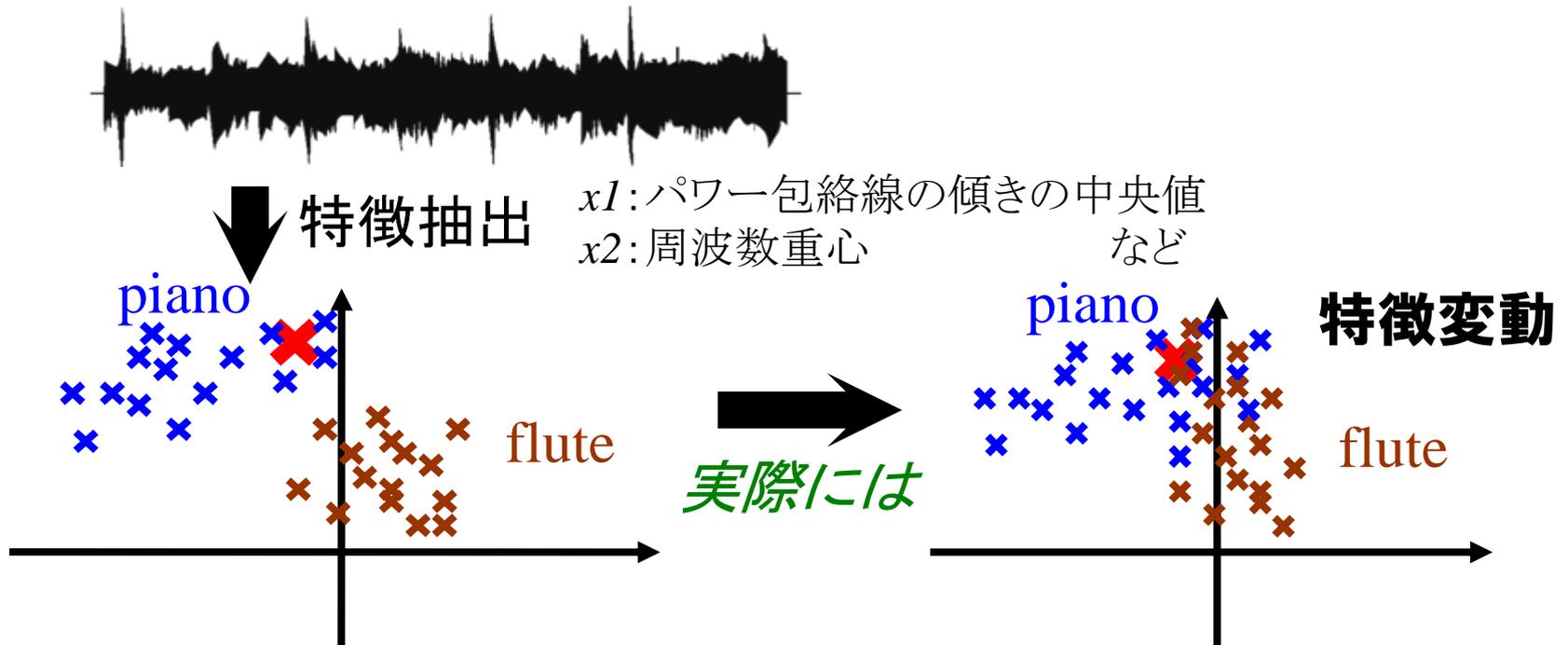
^{††}科技団さきがけ21／産業技術総合研究所

発表の流れ

1. 音源同定とは
2. 音色空間の音高依存性を考慮した音源同定手法
3. 処理の流れ
4. 評価実験
5. まとめ

1. 音源同定とは

- **楽器音の同定** (入力された音は, piano? flute? ...)
- **パターン認識**の一分野
- **自動採譜・メディア検索**などで有用
- 広く研究されるようになったのは**1990年代**から



1. 音源同定とは

- 楽器音における特徴変動の要因：
音高・音の強さ・楽器の個体差・奏法など
- 研究事例：**楽器の個体差に着目した**
「適応型混合テンプレート法」(柏野他, 信学論, '98)
- **音高による音色変化**は扱われてこなかった
特に, 楽器音は音声などに比べて音域が広く,
音高による音色変化が顕著



音高による音色変化を考慮した
音源同定手法を提案

2. 音色空間の音高依存性を考慮した 音源同定手法

音高依存性を考慮する方法:

たとえば, 音域をいくつかのブロックに分割して
入力信号と同じブロックのデータで学習



細かく分割 ⇒ 音高依存性○, 学習データ数△
粗く分割 ⇒ 音高依存性△, 学習データ数○
(学習データが少ないと性能低下の可能性)



学習データを減らさずに音高依存性を考慮するため,
基本周波数の関数をパラメータに持つ分布を提案

2. 音色空間の音高依存性を考慮した音源同定手法 F0依存多次元正規分布

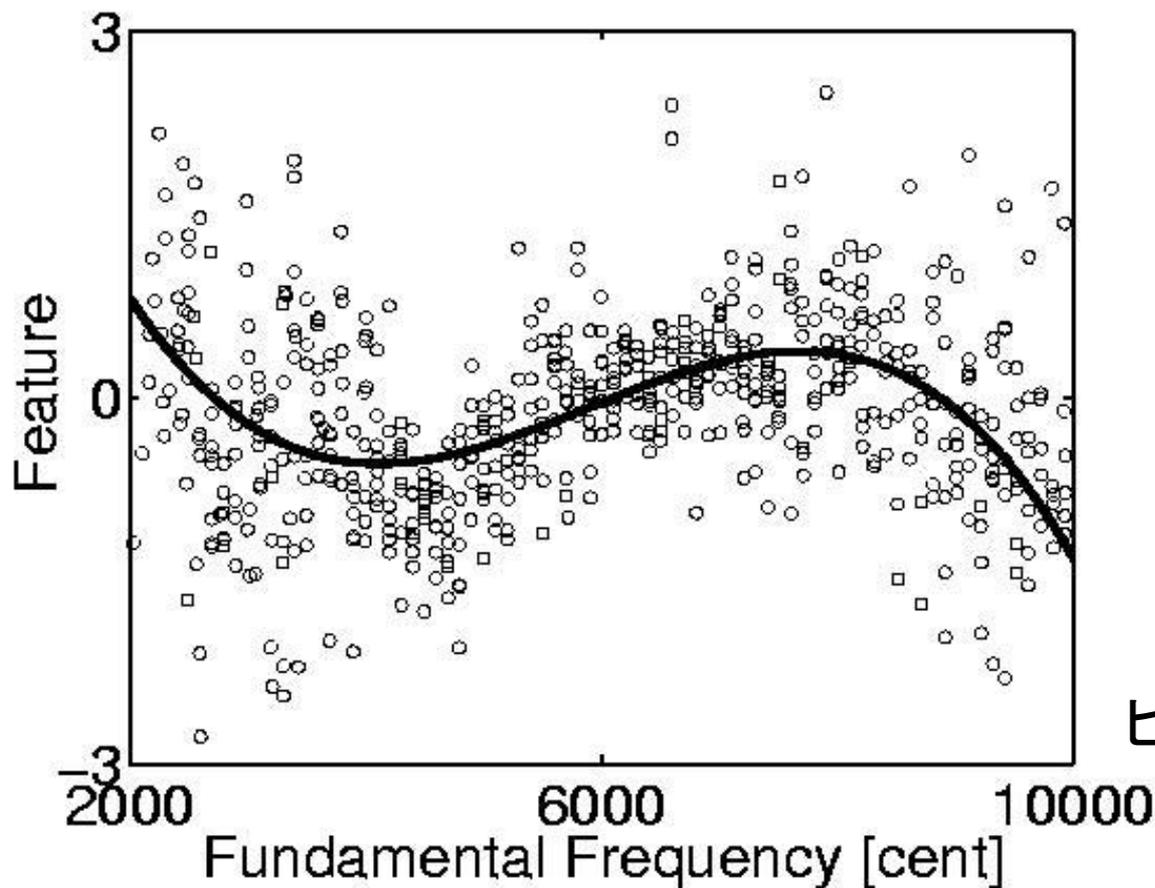
多次元正規分布を以下のように拡張：

- **平均**：基本周波数の関数として定義
⇒音高によって分布の平均がどのように変化するか(音高による音色変化)を表す
- **共分散**：上記の関数からのちらばりの程度を表す(音高非依存)
⇒音高以外の要因による音色変化を表す

これにより、音高による音色変化を考慮しながら、**全音域**の音色パターンを**1つの分布**で表現可能

2. 音色空間の音高依存性を考慮した音源同定手法 代表値関数

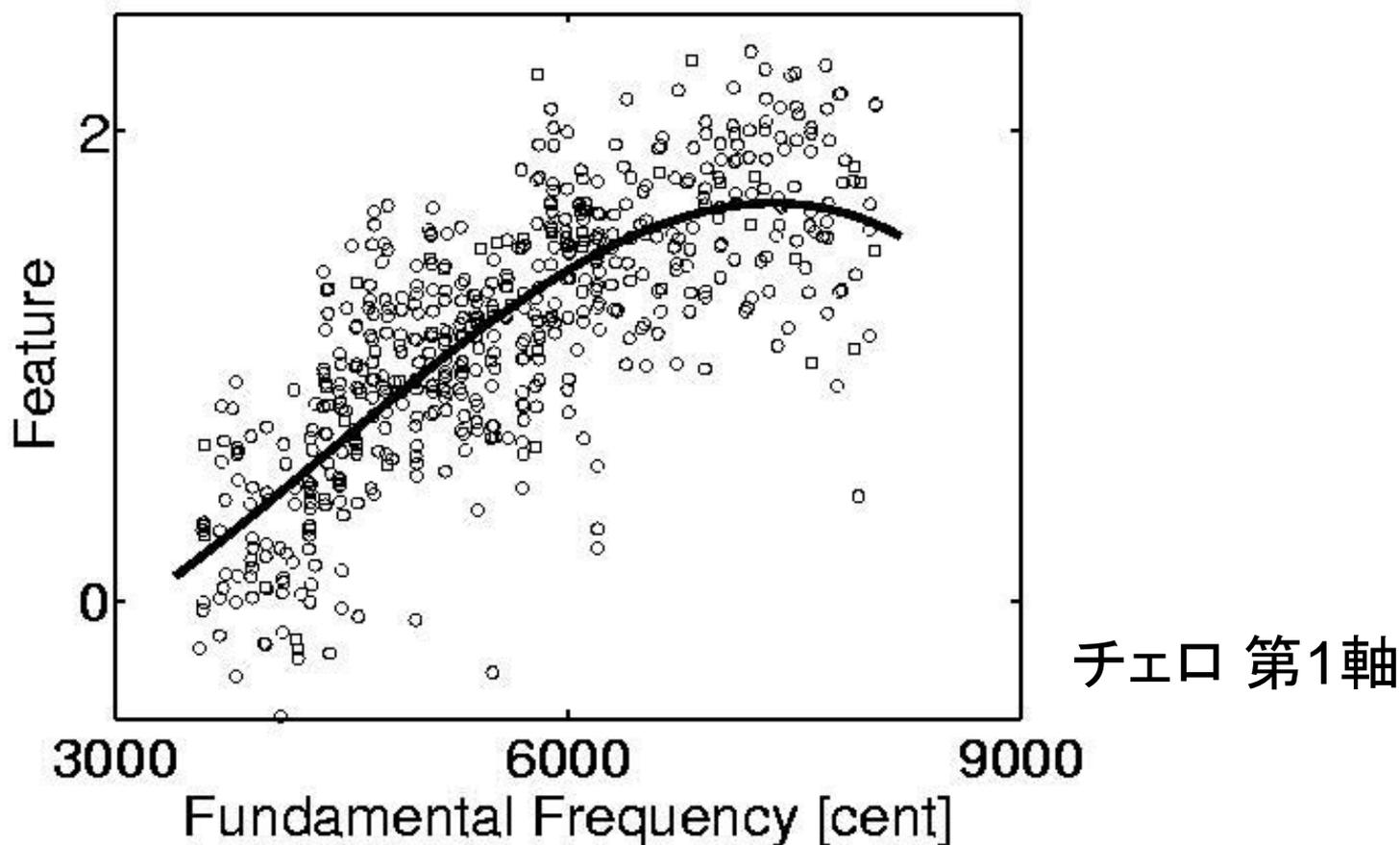
基本周波数によって変化する分布の平均を
関数近似(3次関数)により推定



ピアノ 第4軸

2. 音色空間の音高依存性を考慮した音源同定手法 代表値関数

基本周波数によって変化する分布の平均を
関数近似(3次関数)により推定



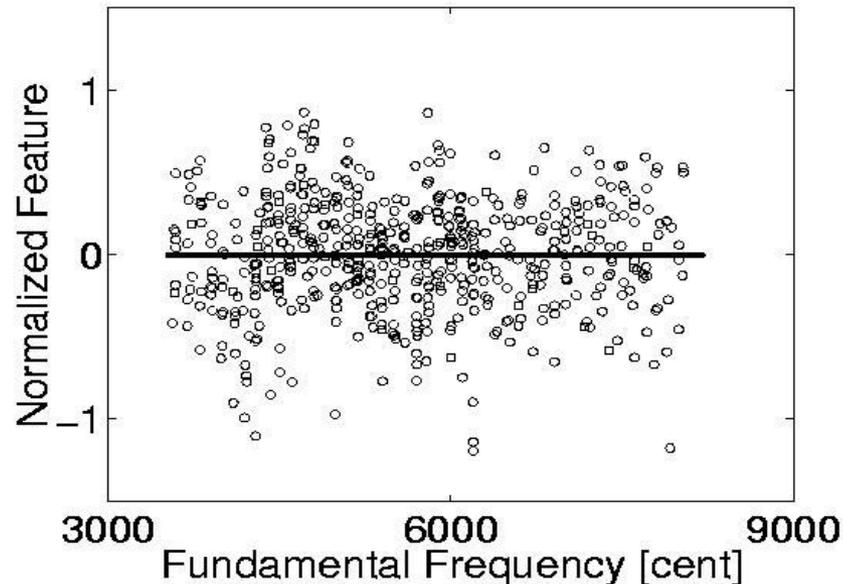
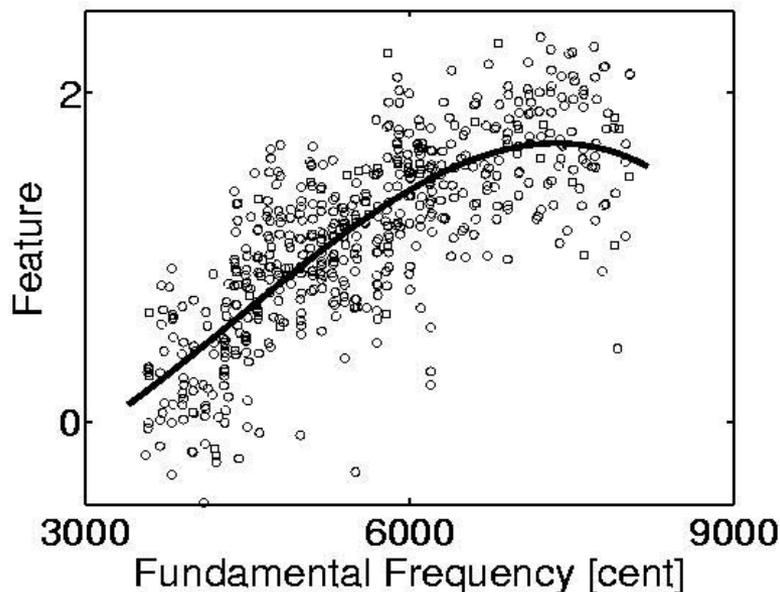
3. 音色空間の音高依存性を考慮した音源同定手法 FO正規化共分散行列

代表値関数からのちらばりの程度を表す

⇒ **音高以外の要因による音色変化**を表す

⇒ 音色空間を代表値関数で正規化してから、
共分散行列を求める

↓
音高による音色変化を除去



3. 処理の流れ

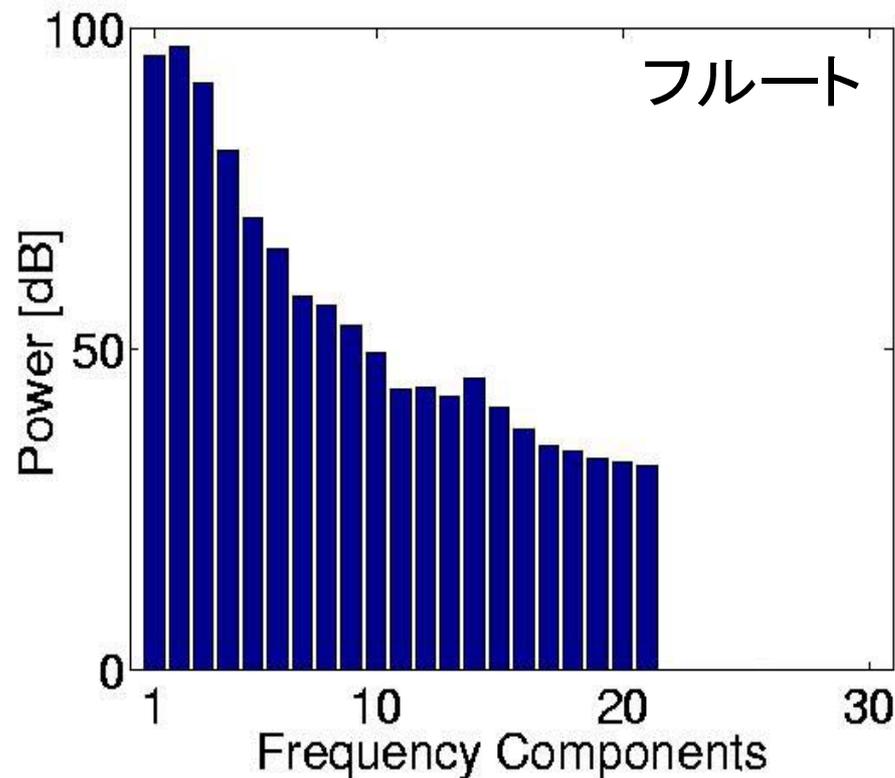
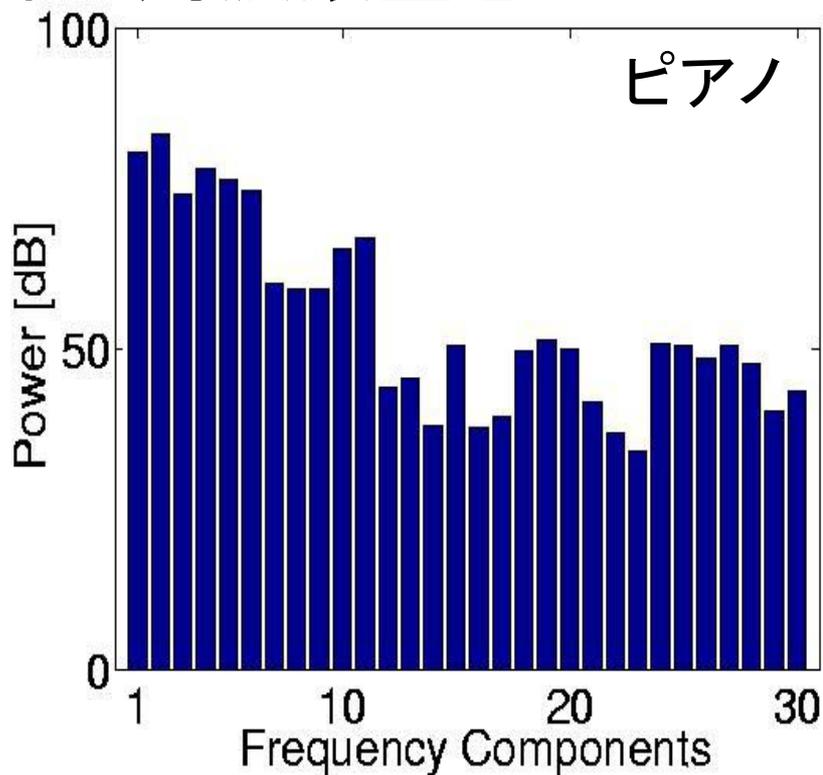
1. 特徴抽出 (129個)
2. 主成分分析で次元圧縮
(累積寄与率99%で79次元に圧縮)
3. 線形判別分析でさらに次元圧縮
(19楽器なので18次元に圧縮)
4. F0依存多次元正規分布のパラメータ推定
5. ベイズ決定規則に基づいて楽器名を同定
(事後確率が最大になる楽器名を見つける)

$$g_i(x; f) = \log p(x | \omega_i; f) + \log p(\omega_i; f)$$

3. 処理の流れ

1. 特徴抽出(129個)

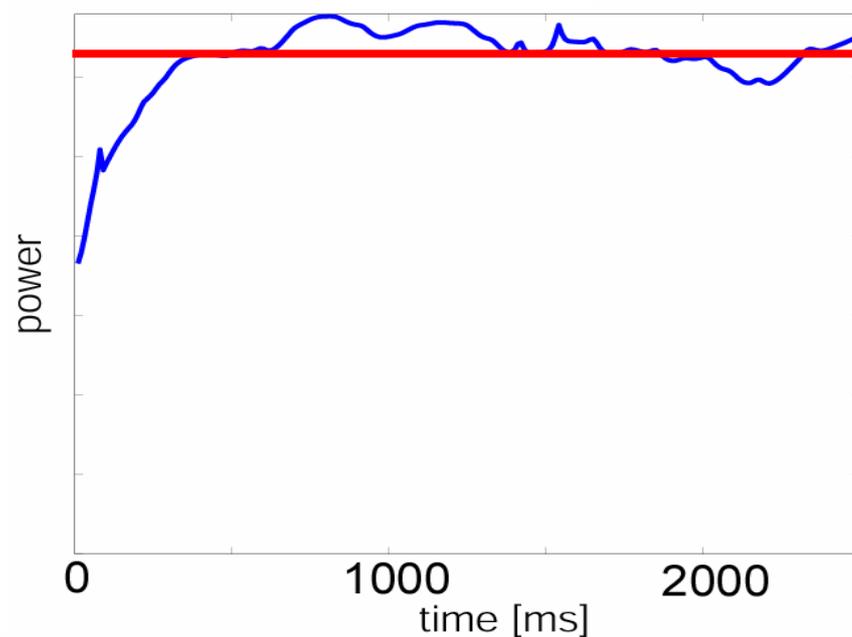
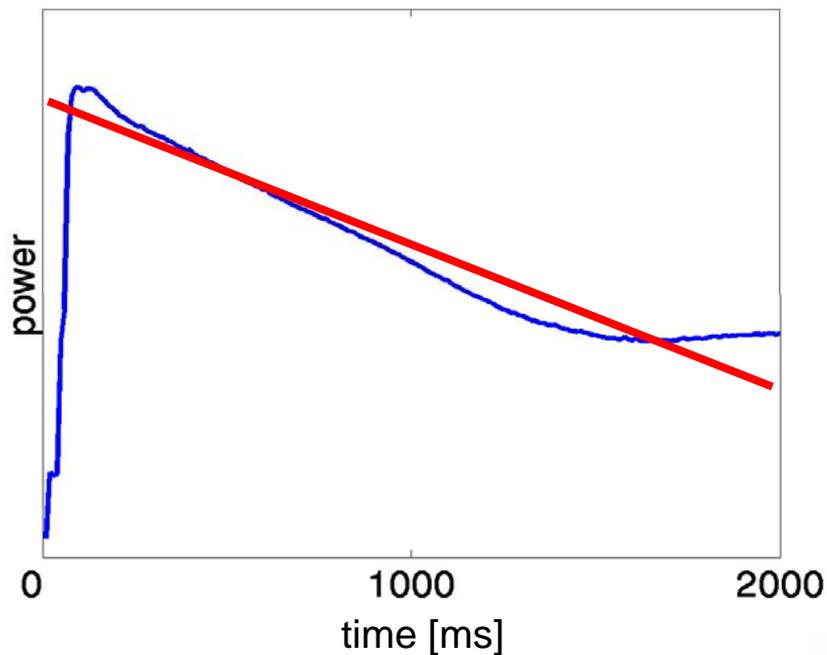
例: 周波数重心



3. 処理の流れ

1. 特徴抽出(129個)

例: パワー包絡線の最小二乗法による近似直線の傾き
ピアノ フルート



3. 処理の流れ

1. 特徴抽出 (129個)
2. 主成分分析で次元圧縮
(累積寄与率99%で79次元に圧縮)
3. 線形判別分析でさらに次元圧縮
(19楽器なので18次元に圧縮)
4. F0依存多次元正規分布のパラメータ推定
5. ベイズ決定規則に基づいて楽器名を同定
(事後確率が最大になる楽器名を見つける)

$$g_i(x; f) = \log p(x | \omega_i; f) + \log p(\omega_i; f)$$

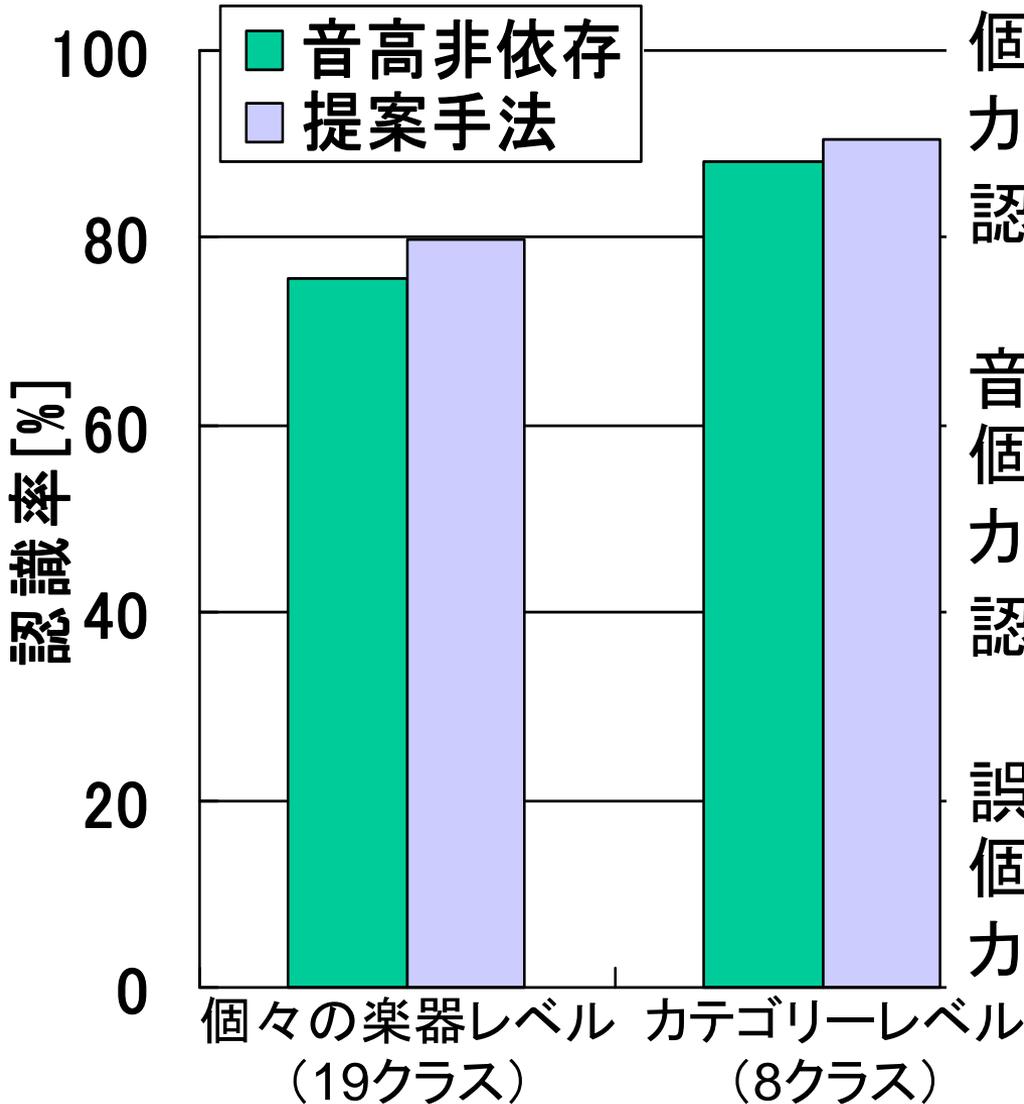
4. 評価実験

実験方法

- 使用データベース: RWC-MDB-I-2001
- 実楽器の単独発音を半音ごとに収録
- 今回は19種類の楽器を使用
- 各楽器に, 3楽器個体, 3種類の音の強さ
- 今回は, 通常の奏法のみ使用
- 使用したデータ総数: 6247個
- 上記のデータを無作為に10等分し,
クロスバリデーション
- 音高は既知
- カテゴリーレベルの認識率も算出

ピアノ	ピアノ	
ギター	クラシックギター ウクレレ	アコースティック ギター
弦楽器	バイオリン ビオラ	チェロ
金管楽器	トランペット	トロンボーン
サククス	ソプラノサククス アルトサククス	テナーサククス バリトンサククス
複簧楽器	オーボエ	ファゴット
クラリネット	クラリネット	
無簧楽器	ピッコロ フルート	リコーダー

4. 評価実験 実験結果



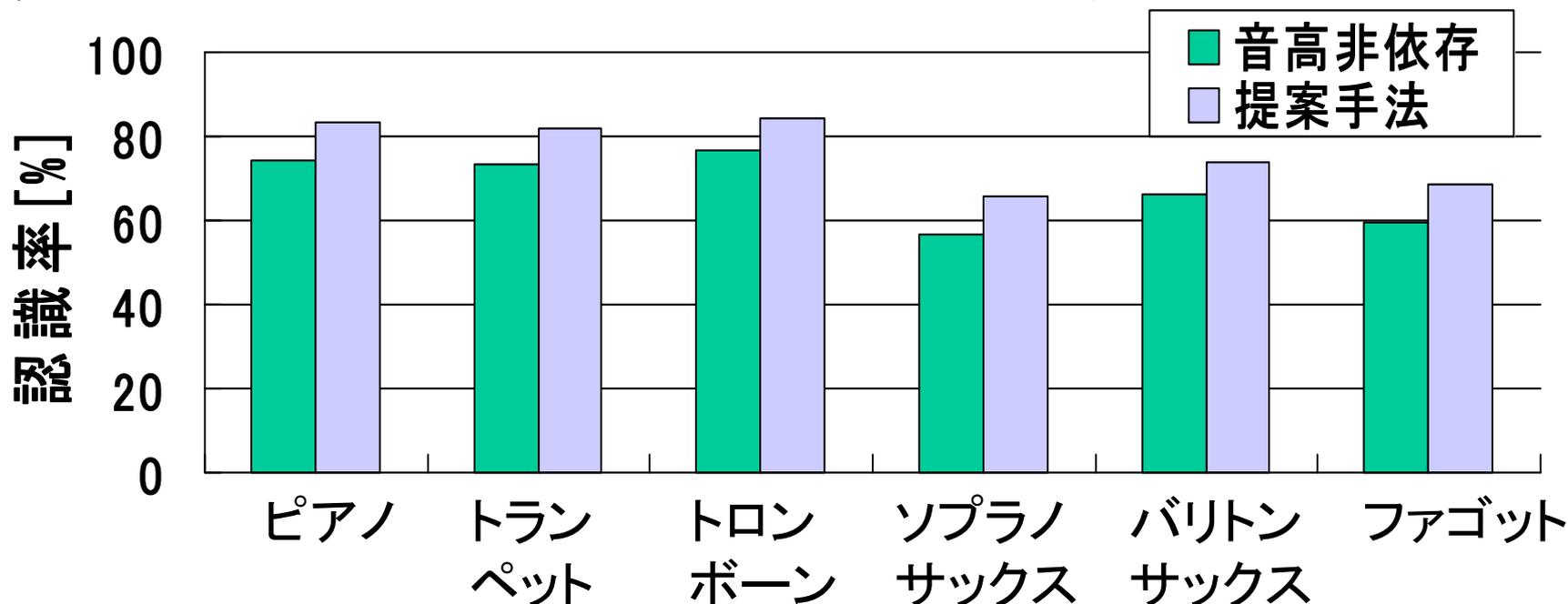
個々の楽器レベルで**約80%**,
カテゴリーレベルで**約90%**の
認識率を実現

音高非依存に比べて,
個々の楽器レベルで**4.00%**,
カテゴリーレベルで**2.45%**,
認識率向上

誤り削減率は,
個々の楽器レベルで**16.48%**,
カテゴリーレベルで**20.67%**

4. 評価実験

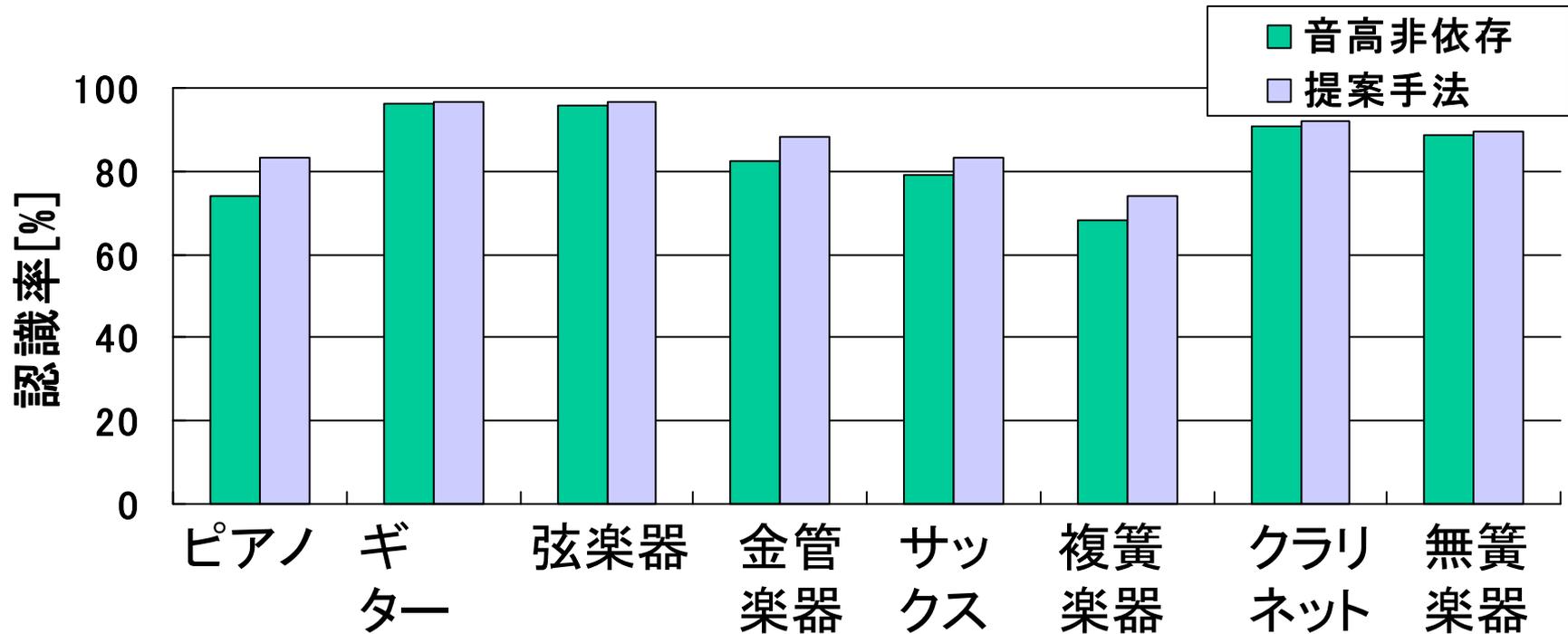
実験結果／認識率が7%以上改善された楽器



- ・ **ピアノ**: 最も性能改善
(認識率 **9.06%** 改善, 誤り削減 **35.13%**)
 - ∵ 音域が広く音高による音色変化が顕著
- ・ PF, TR, TB で **約33~35%** の認識誤りを削減
- ・ SS, BS, FG でも **20%以上** の認識誤りを削減

4. 評価実験

実験結果／カテゴリーレベルの認識率



誤り削減 35% 8% 23% 33% 20% 13% 15% 8%

- すべてのカテゴリーで認識率改善
- ギター, 弦楽器の認識率(提案手法): 96.7%
- 最も低いカテゴリーでも72%の認識率(提案手法)

5. まとめ

- 音高による音色変化を考慮するため、**F0依存多次元正規分布**を提案
- F0依存多次元正規分布のための識別関数をベイズ決定規則から定式化
⇒ **音源同定の性能向上に貢献**
(個々の楽器で16.48%,
カテゴリーレベルで20.67%認識誤りを削減)
- 今後の課題
 - ベイズ決定規則以外への応用
 - 混合音への適用など