

音高による音色変化に着目した 音源同定手法

北原 鉄朗[†] 後藤 真孝^{††} 奥乃 博^{†††}

[†]東京理科大学工学部情報科学科

^{††}科技団さきがけ21 / 産業技術総合研究所

^{†††}京都大学大学院情報学研究科知能情報学専攻

23 May 2001

発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音色変化に着目した音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音色変化に着目した音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

1. 導 入

音源同定の重要性

- 音源同定は , 人間にとって基本的な聴覚的情景分析の能力である .
- 自動採譜などにおいて , 有用である .

従来の研究事例は多くない .

本研究では , 音源同定を扱う .

将来は , 音全般に対象を拡張 .

1. 導 入

楽器音オントロジー構築にむけた アプローチ

楽器音オントロジー …

楽器音を一般的に表す枠組み

- 楽器音の理解は階層的なものである。
- 音源同定で扱う特徴量は、
特徴量によってさまざまな傾向を示す。
(音高との相関性など)

1. 導 入

楽器音オントロジー構築にむけた アプローチ

楽器音オントロジー …

楽器音を一般的に表す枠組み

- 楽器音の理解は階層的なものである。
階層的な知識表現
- 音源同定で扱う特徴量は、
特徴量によってさまざまな傾向を示す。
(音高との相関性など)
明示的に記述し、それに応じた処理。

発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

- 音色変化の問題
- 未知楽器の問題

2.音源同定における問題点 音色変化の問題

同一楽器であっても、音高・音の強さ・
楽器の個体差・演奏方法によって音色が
変化することをどう扱うか。

この問題を扱った従来研究は少ない。
たとえば、楽器の個体差に着目
「適応型混合テンプレート法」(柏野ら, '98)など。

音高による音色変化を扱ったものはない。

2.音源同定における問題点 音色変化の問題

同一楽器であっても、音高・音の強さ・
楽器の個体差・演奏方法によって音色が
変化することをどう扱うか。



音高による音色変化に着目

音色が音高によってどう変化するかを調べ、
変化の様子を基本周波数の関数として表現。

2.音源同定における問題点

未知楽器の問題

未知の楽器 (学習データに含まれていない楽器)
をどう扱うか。

従来の一般的な音源同定システムは、
既知の楽器 (学習データに含まれる楽器)の
どれかにむりやり同定。



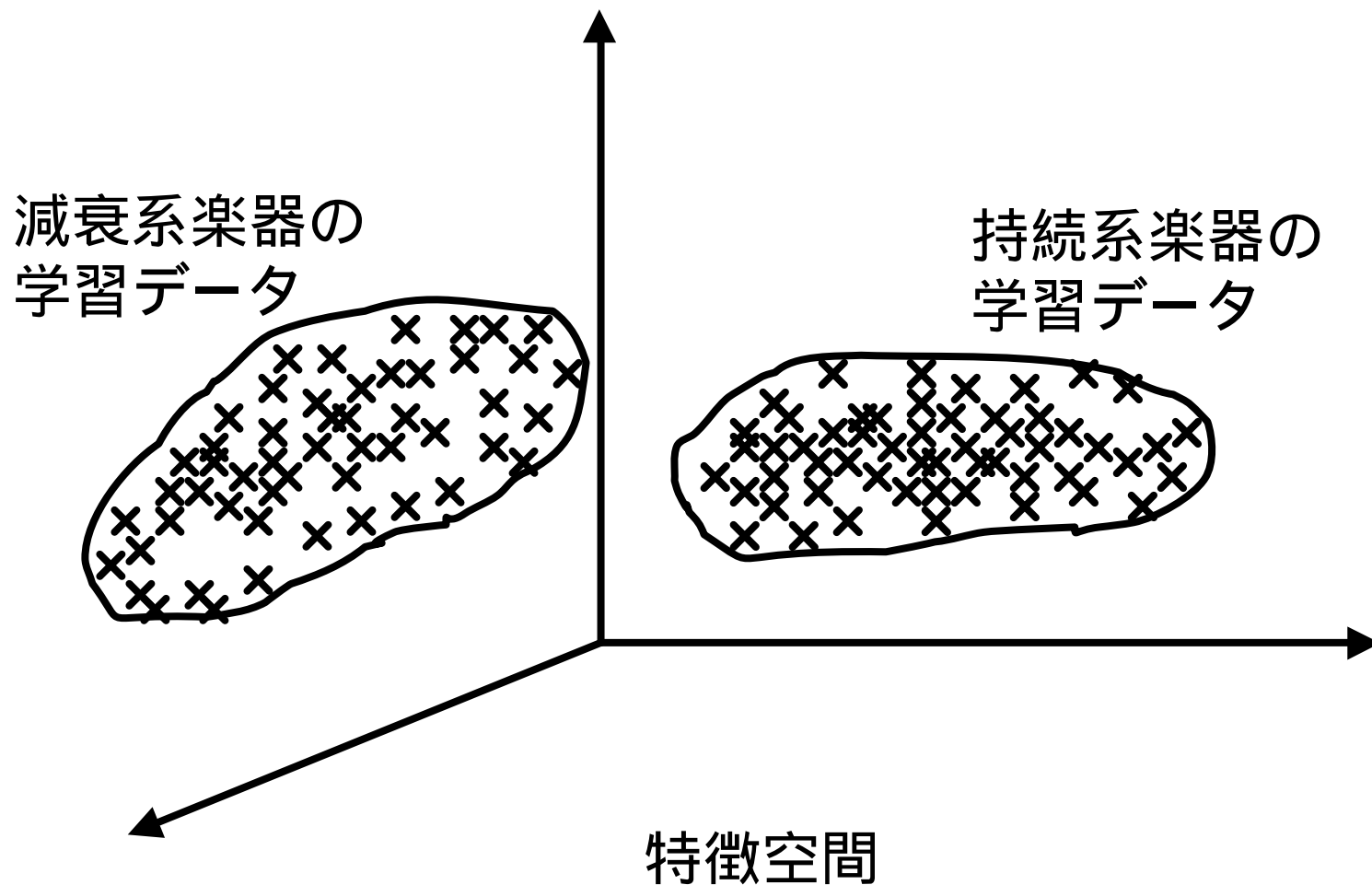
本音源同定システムでは、
楽器の階層的な分類に基づいて
一定の階層まで同定。

例 classical guitar 「減衰系楽器」

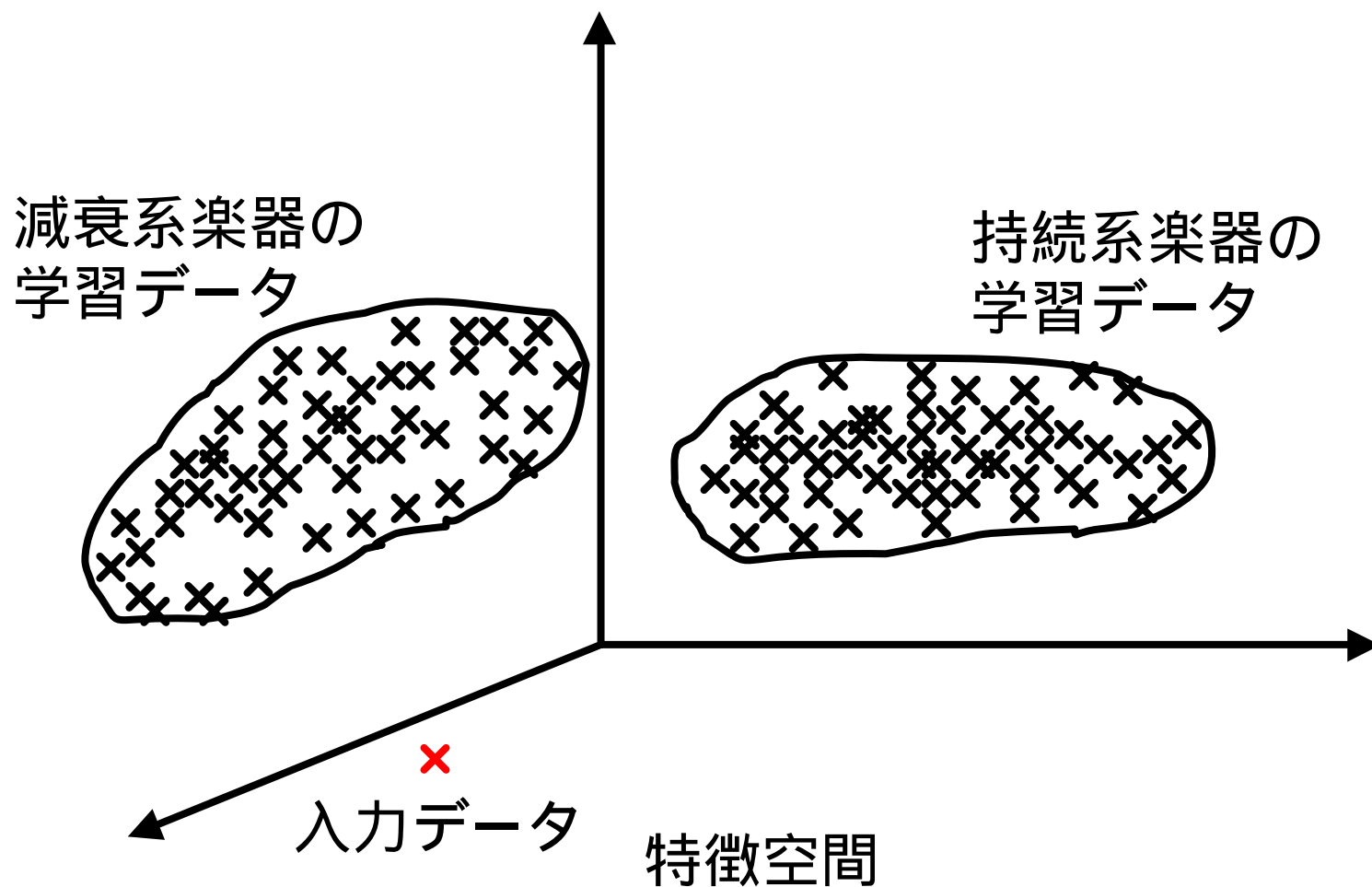
発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音色変化に着目した音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

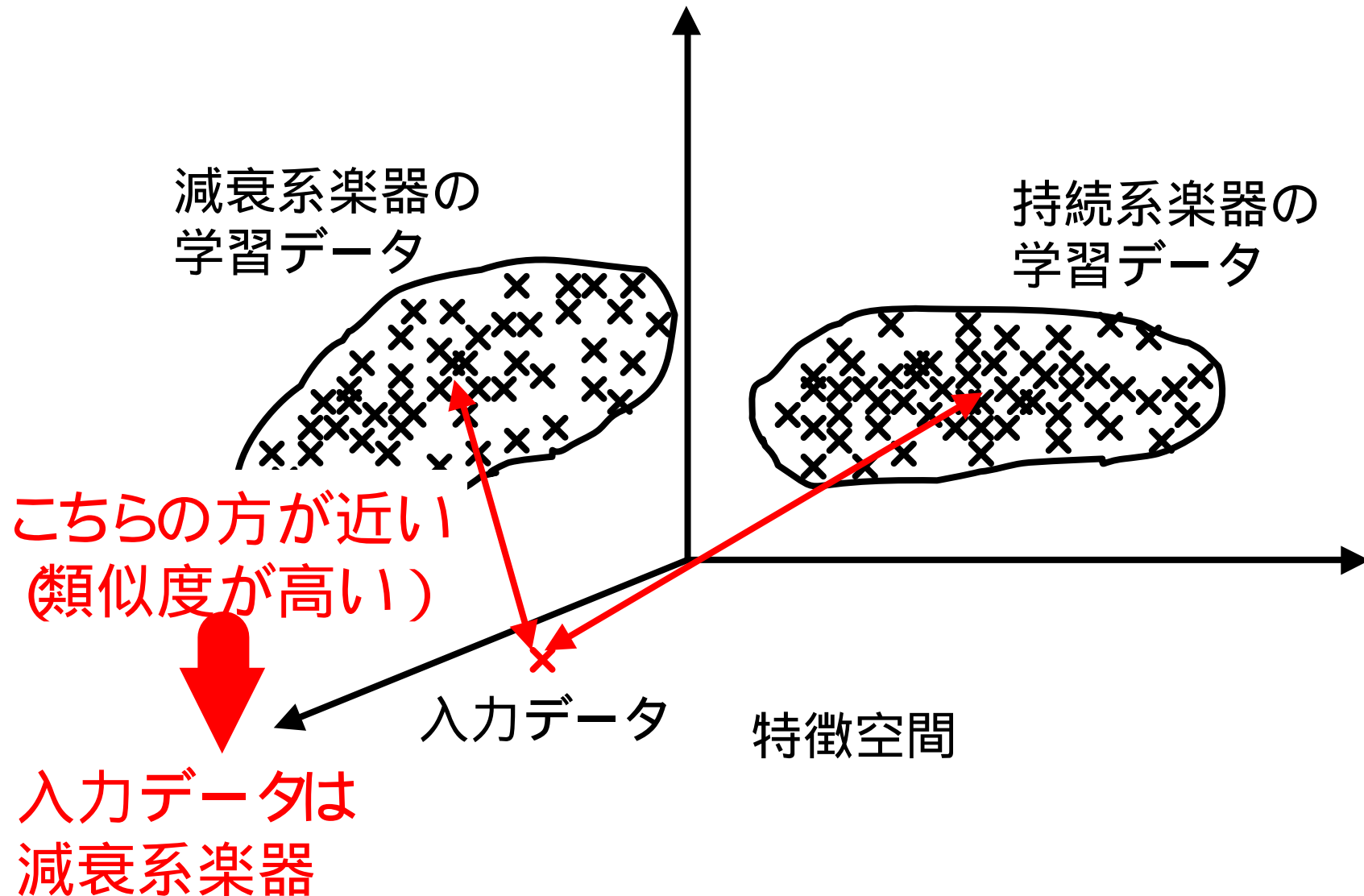
3. 本音源同定システムの概要



3. 本音源同定システムの概要



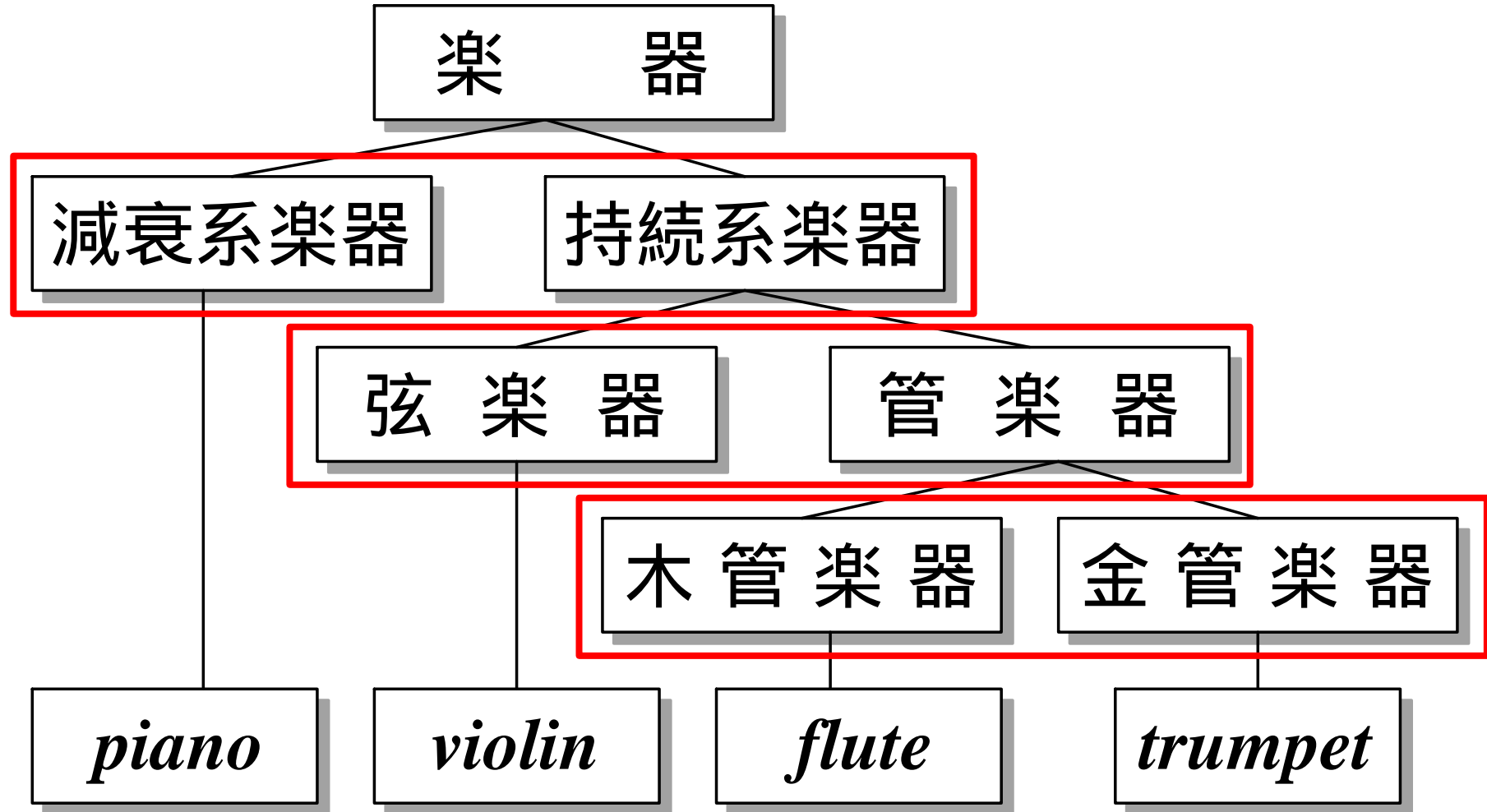
3. 本音源同定システムの概要



発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音色変化に着目した音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

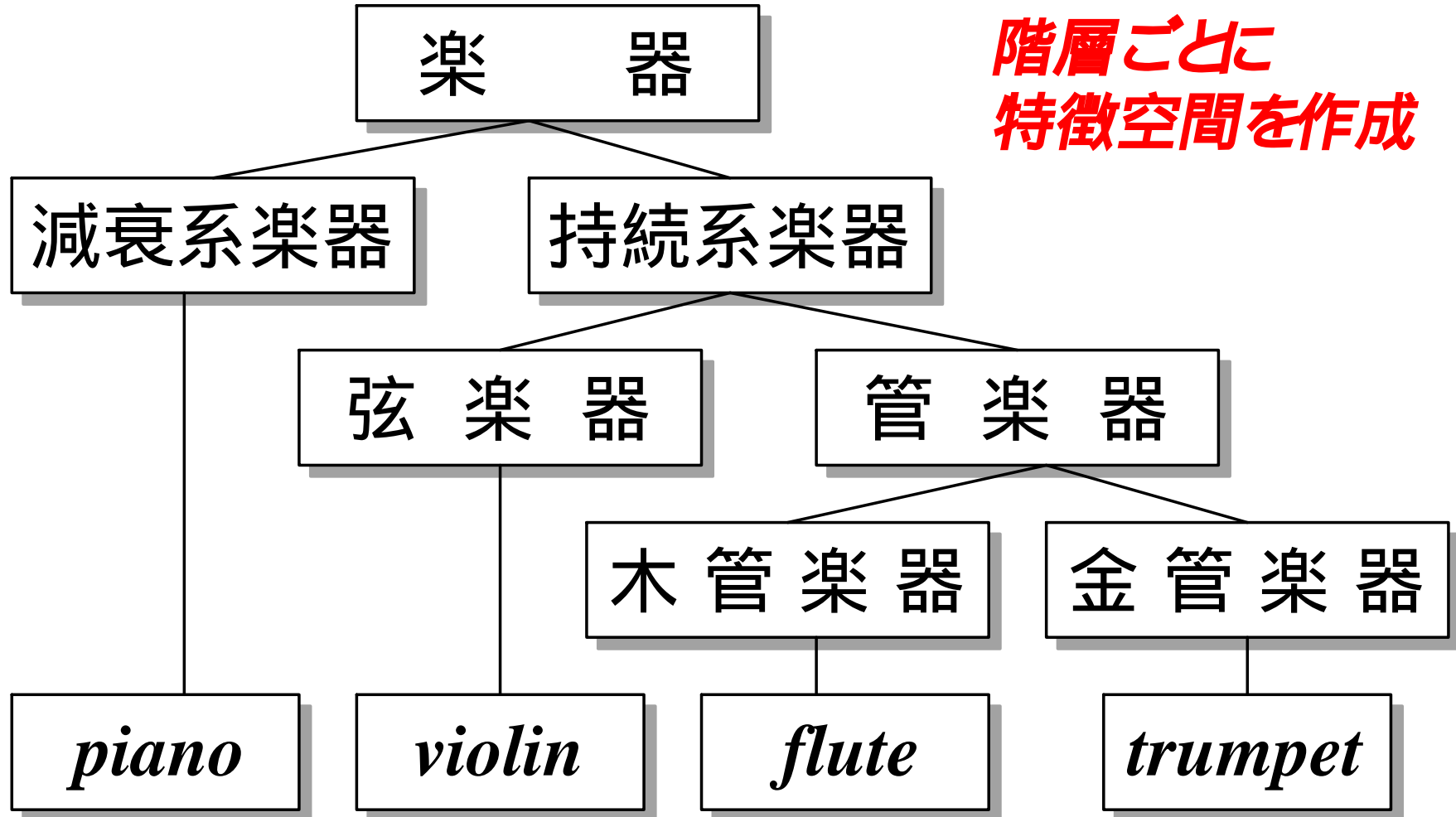
4.階層的な音源同定 楽器の階層的分類



今回は、音の長さが長いもののみを扱う。

4 .階層的な音源同定

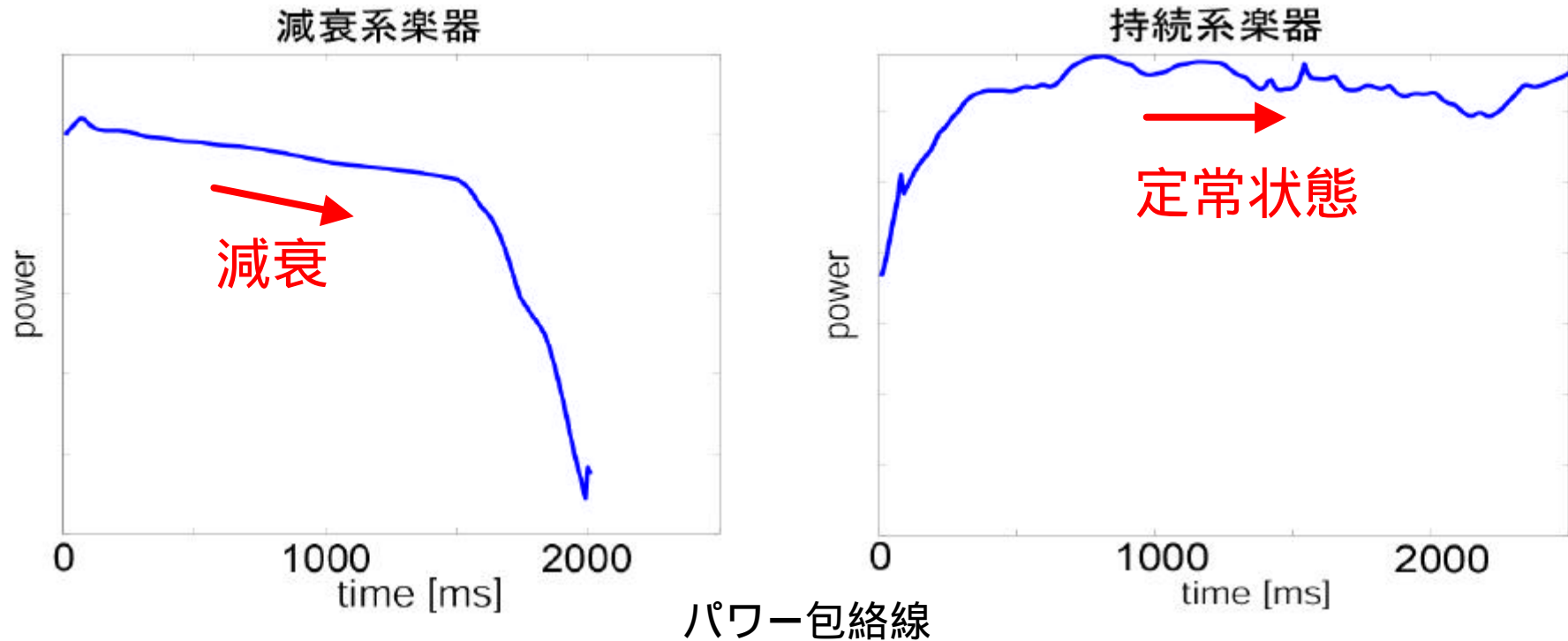
楽器の階層的分類



今回は ,音の長さが長いもののみを扱う.

4.階層的な音源同定

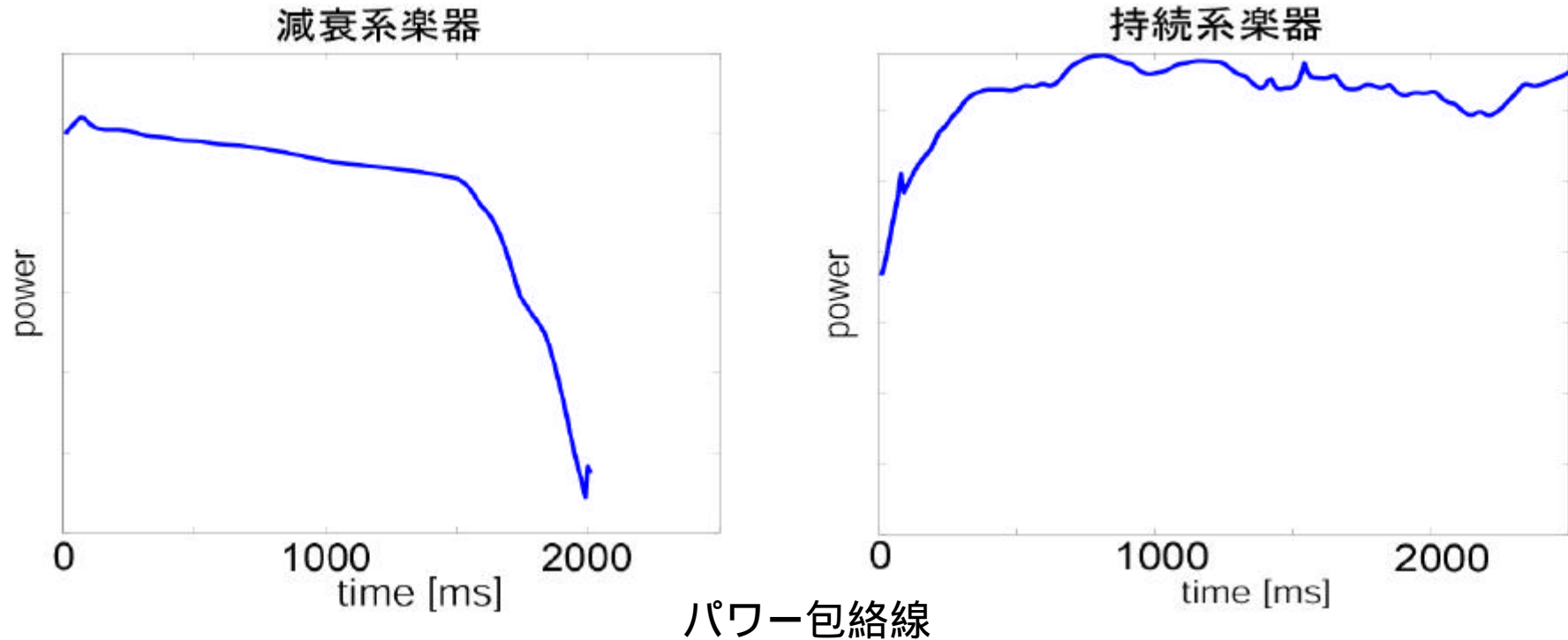
減衰系楽器---持続系楽器



減衰系楽器は ,立ち上がり後すぐに減衰 .

4 .階層的な音源同定

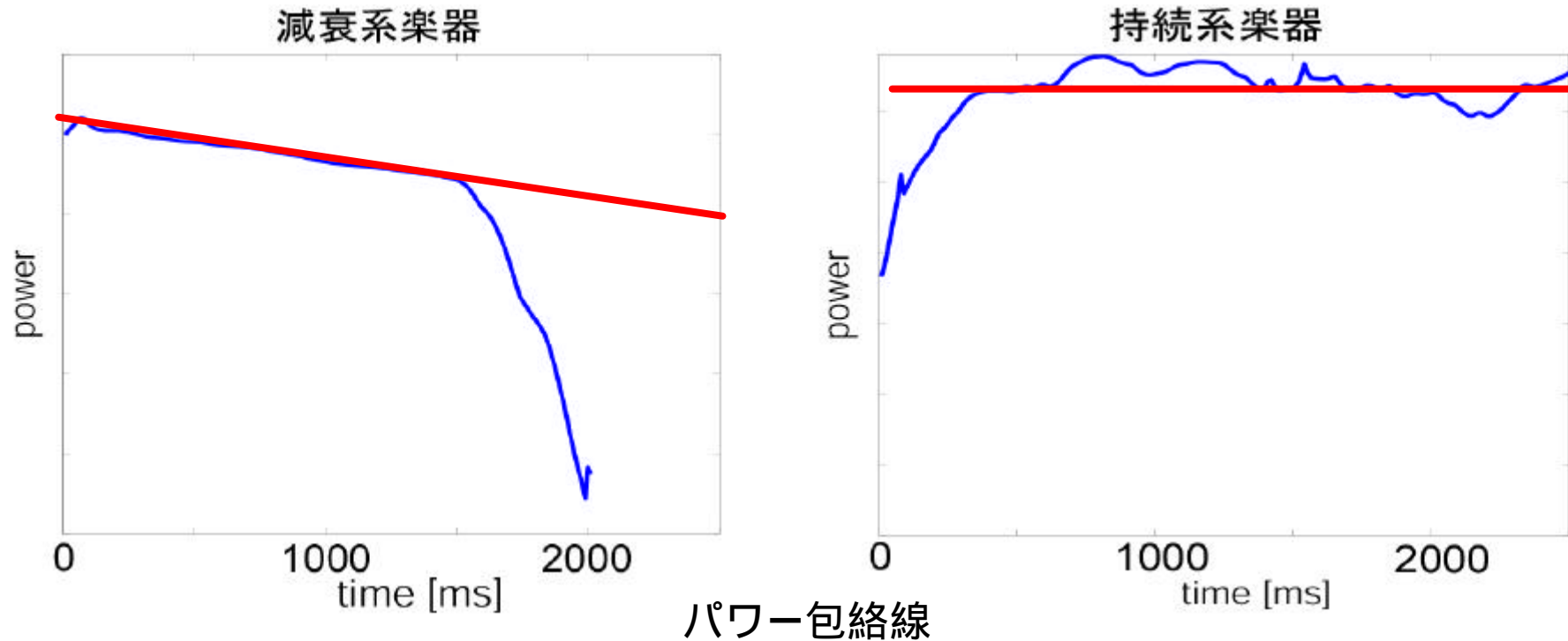
減衰系楽器---持続系楽器



- i) パワー包絡線の近似直線の傾き .
- ii) 最初から800msまでのパワー包絡線の微分係数の中央値 .
- iii) 最大パワー値と800msのときのパワー値との差 .

4 .階層的な音源同定

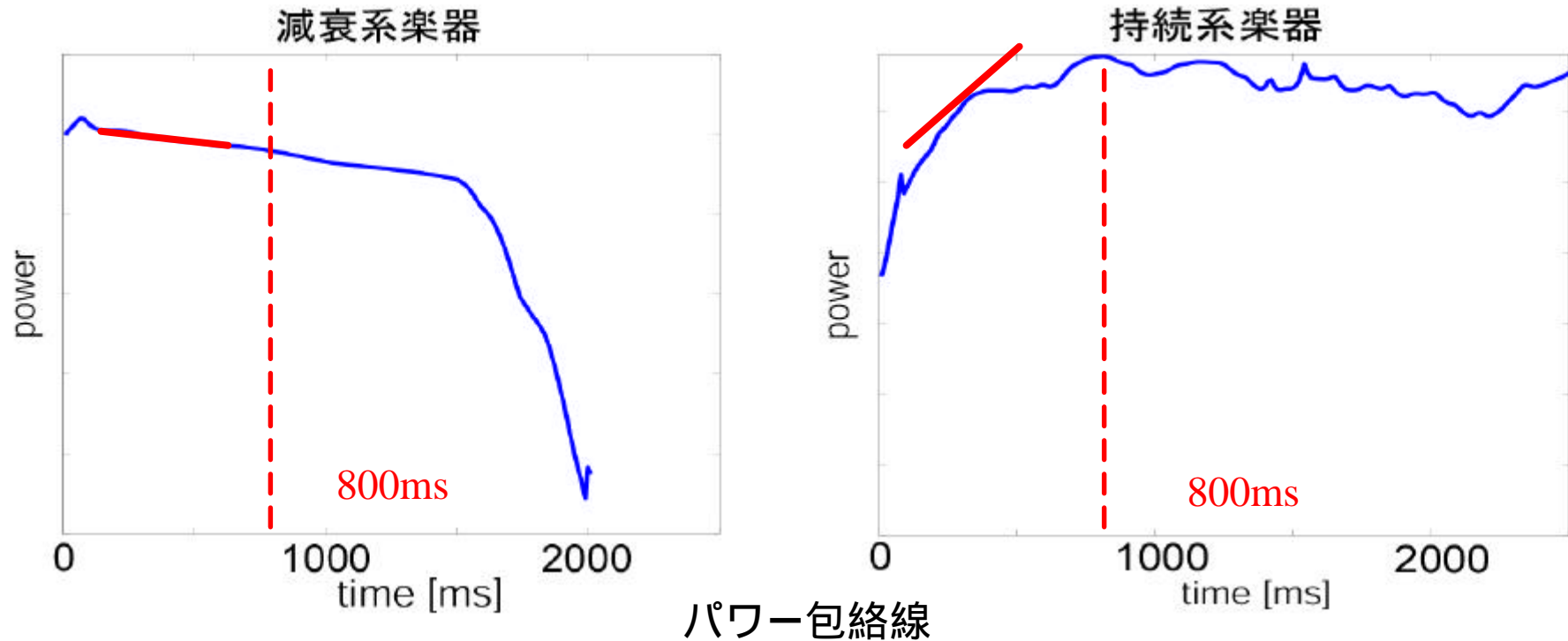
減衰系楽器---持続系楽器



- i) **パワー包絡線の近似直線の傾き .**
- ii) 最初から800msまでのパワー包絡線の微分係数の中央値 .
- iii) 最大パワー値と800msのときのパワー値との差 .

4 .階層的な音源同定

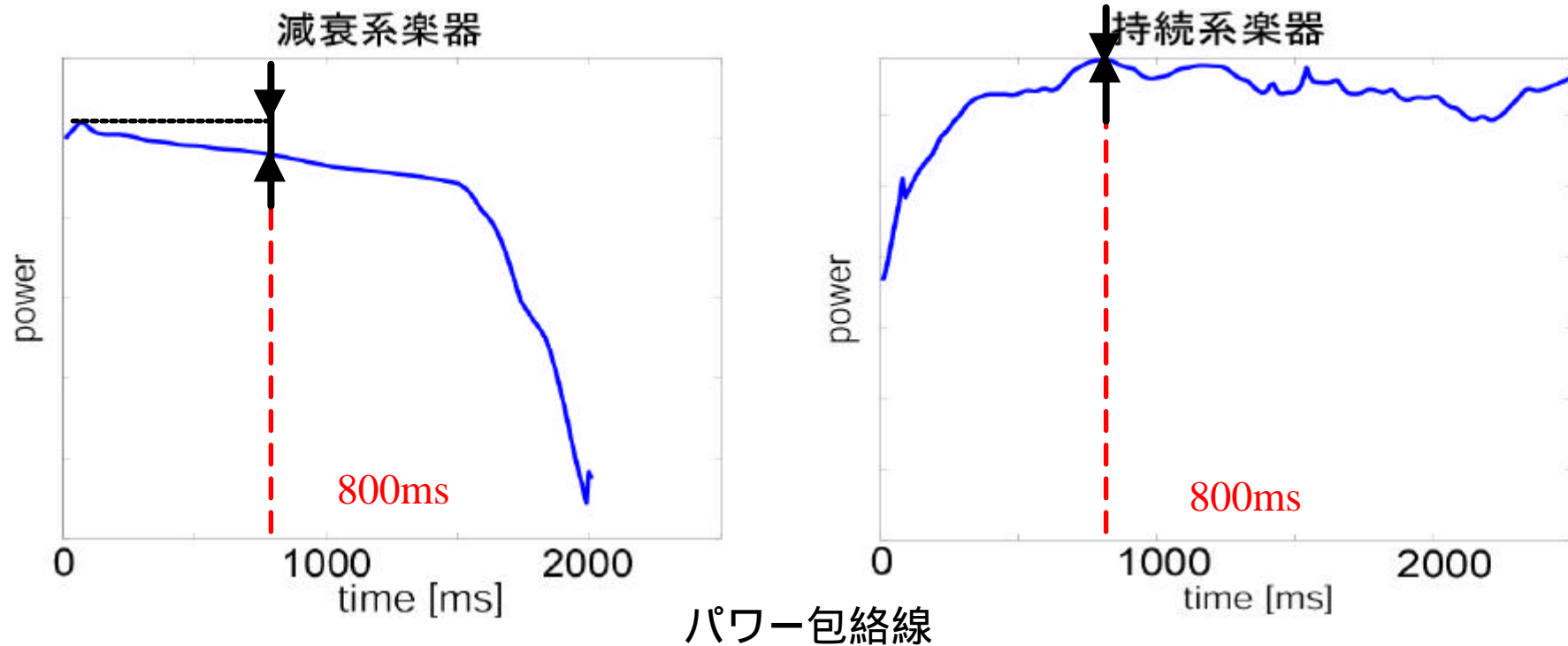
減衰系楽器---持続系楽器



- i) パワー包絡線の近似直線の傾き .
- ii) 最初から800msまでのパワー包絡線の微分係数の中央値 .
- iii) 最大パワー値と800msのときのパワー値との差 .

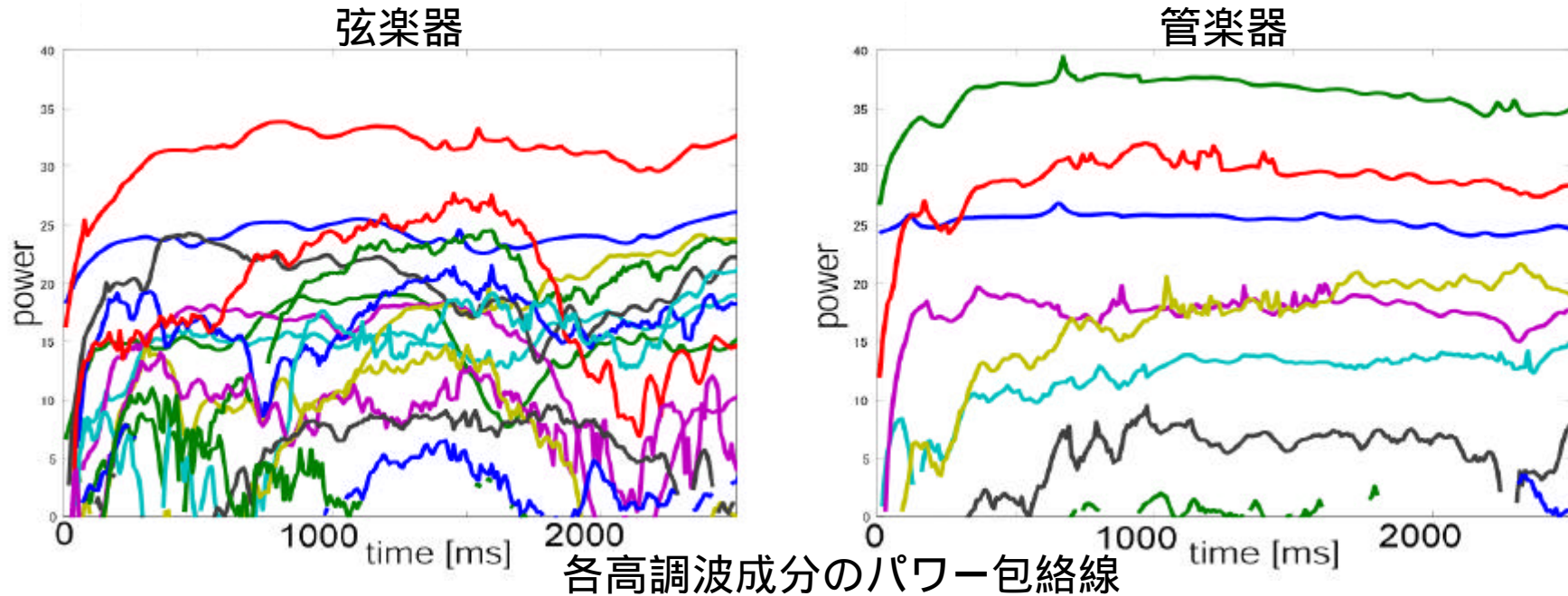
4 .階層的な音源同定

減衰系楽器---持続系楽器



- i) パワー包絡線の近似直線の傾き .
- ii) 最初から800msまでのパワー包絡線の微分係数の中央値 .
- iii) 最大パワー値と800msのときのパワー値との差 .

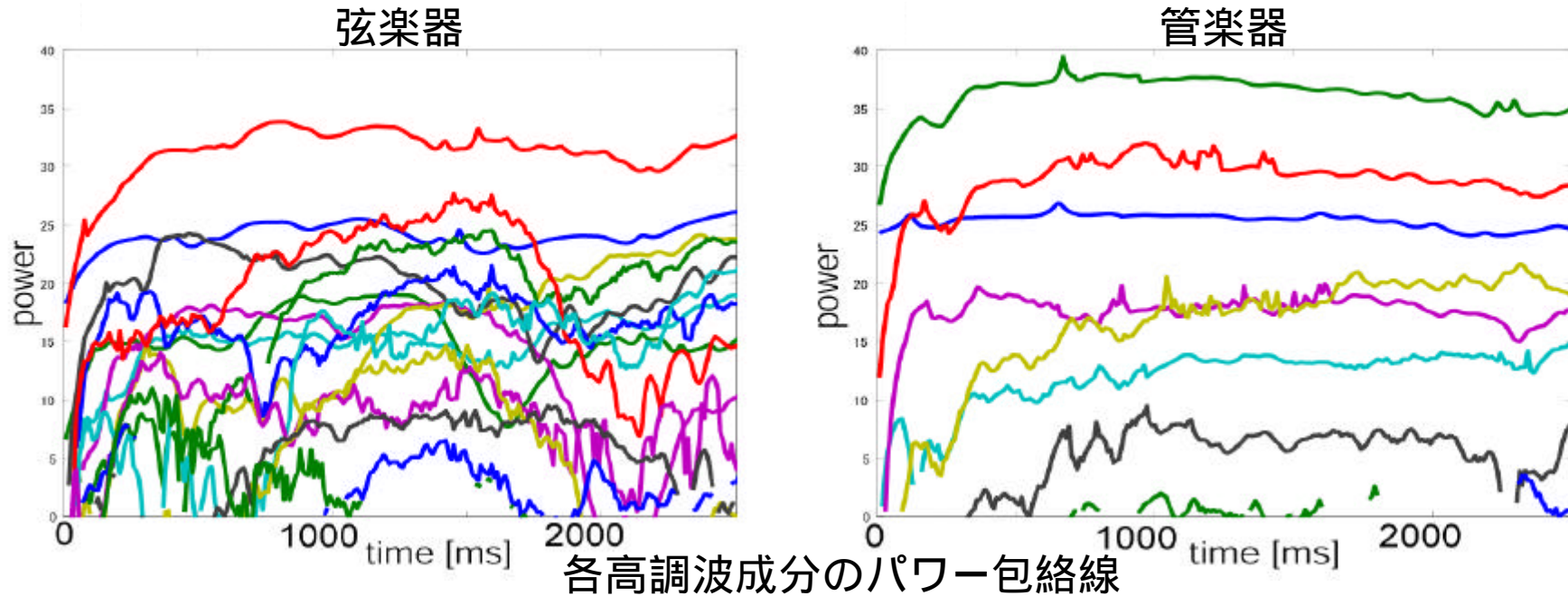
4.階層的な音源同定 弦楽器---管楽器



弦楽器は、

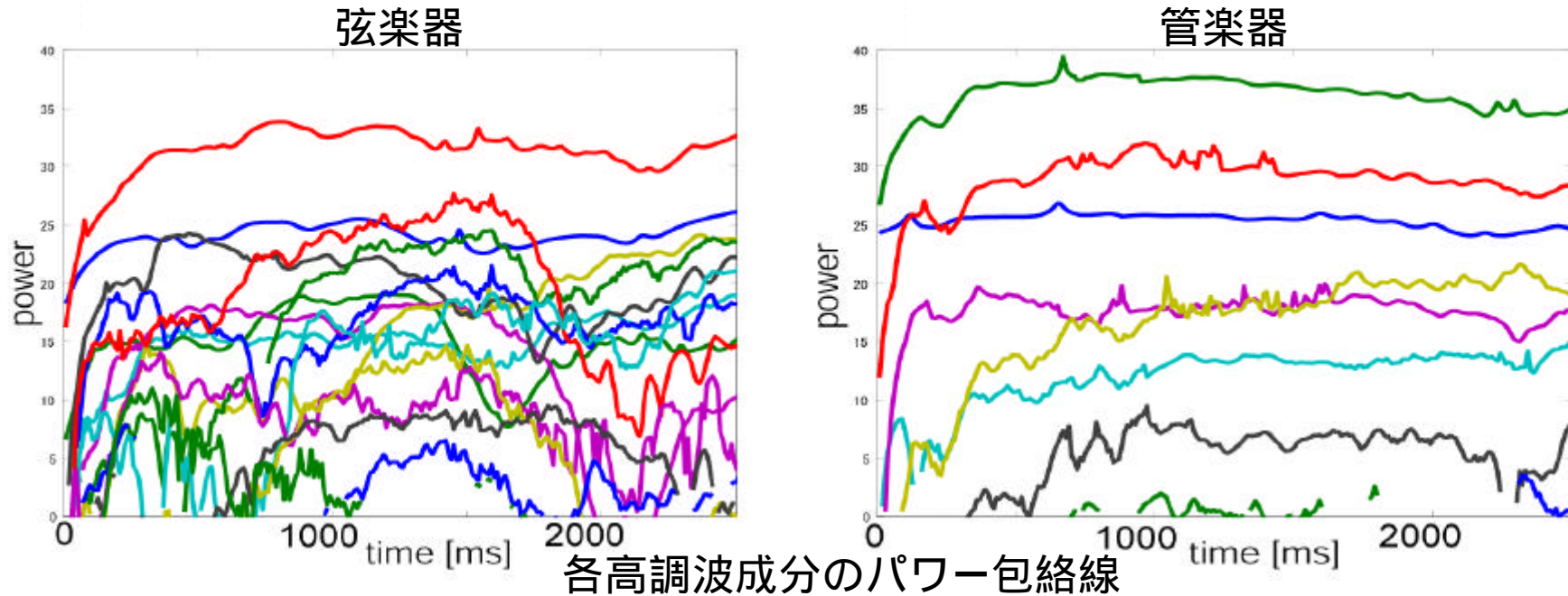
- (1) 高調波成分が豊富．
- (2) 各高調波成分のパワーの変動が激しい．

4 .階層的な音源同定 弦楽器---管楽器



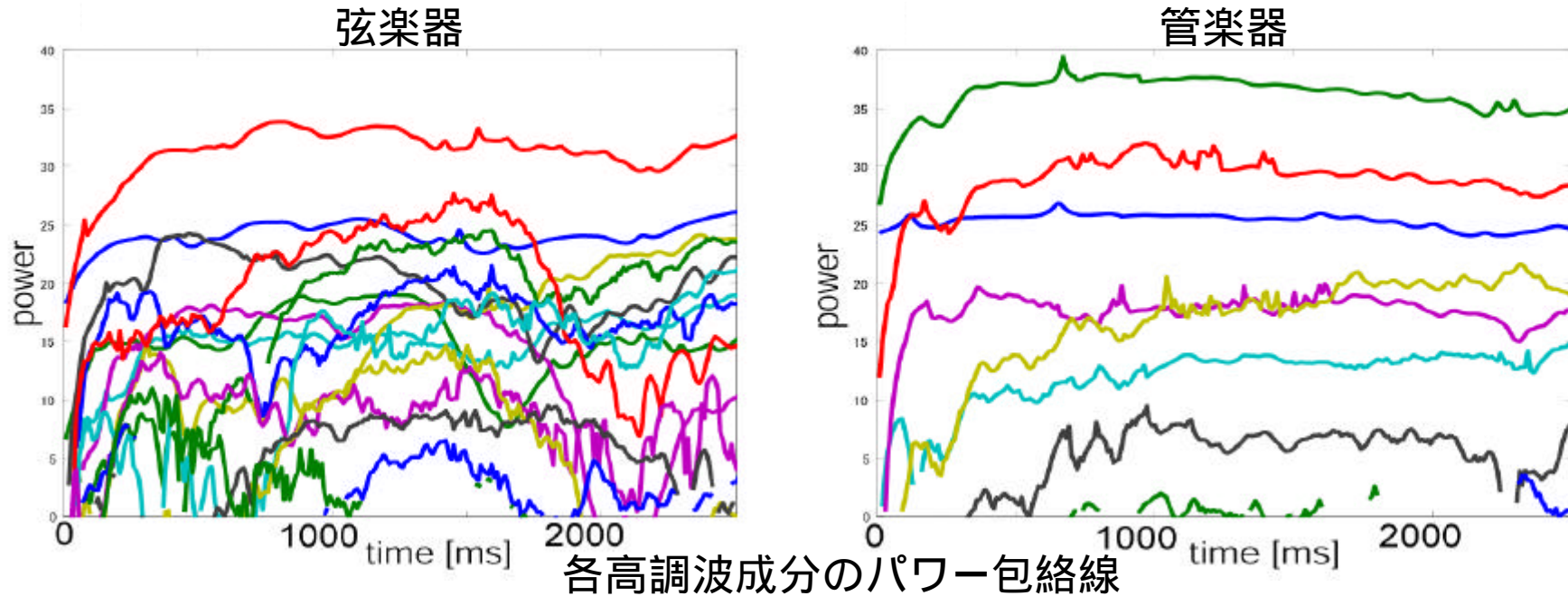
- i) 全持続時間の70%以上鳴り続けている
高調波成分の個数 .
- ii) 周波数重心 .
- iii) 各高調波成分のパワー値の時間変化の標準偏差を
全高調波成分にわたって平均した値 .

4. 階層的な音源同定 弦楽器---管楽器



- i) 全持続時間の70%以上鳴り続けている
高調波成分の個数 .
 - ii) 周波数重心 .
 - iii) 各高調波成分のパワー値の時間変化の標準偏差を
全高調波成分にわたって平均した値 .
- 弦楽器のほうが多い .

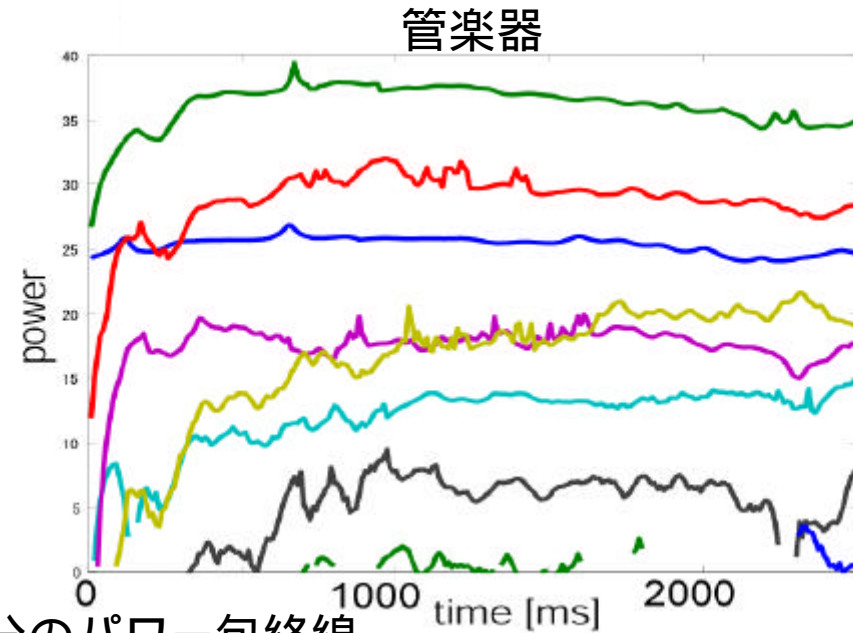
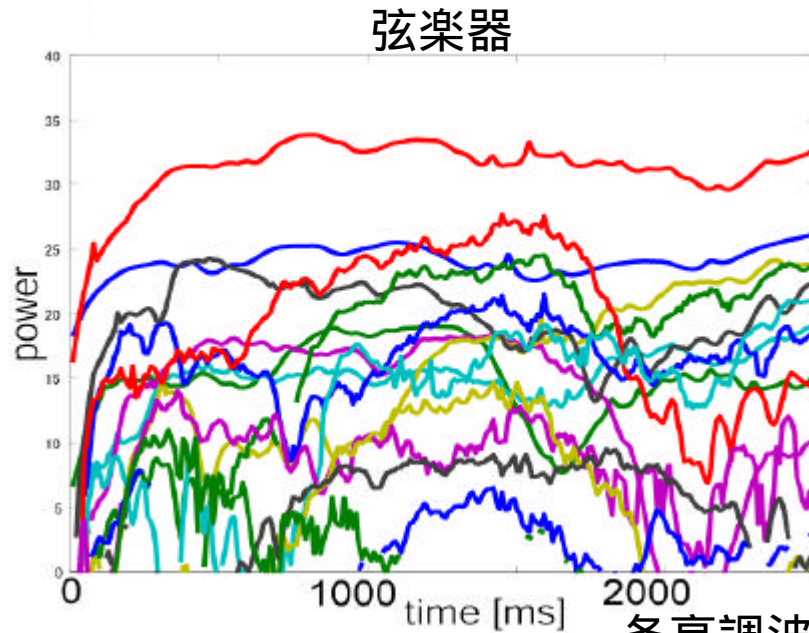
4. 階層的な音源同定 弦楽器---管楽器



- i) 全持続時間の70%以上鳴り続けている
高調波成分の個数 .
- ii) 周波数重心 .
- iii) 各高調波成分のパワー値の時間変化の標準偏差を
全高調波成分にわたって平均した値 .

弦楽器のほうが高い .

4. 階層的な音源同定 弦楽器---管楽器



各高調波成分のパワー包絡線

- i) 全持続時間の70%以上鳴り続けている
高調波成分の個数。
- ii) 周波数重心。
- iii) 各高調波成分のパワー値の時間変化の標準偏差を
全高調波成分にわたって平均した値。

パワー変動の程度を表す。
弦楽器の方が大きい。

4 .階層的な音源同定 弦楽器---管楽器

2つのカテゴリー間の音色が似ているため、
効果的な特徴量を直接抽出するのは難しい。

17個の特徴量を抽出し,PCAで次元を圧縮。

寄与率80% 6次元。

各主成分は,

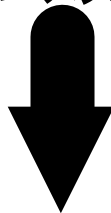
**「高調波成分の豊富さ」 周波数の変動の幅」
「ビブラート・トレモロ」などを総合的に表す。**

発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音色変化に着目した音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

5 .音高による音色変化に 着目した音源同定

本手法では ,特徴量テンプレート中の
学習データとの類似度に基づいて同定する .



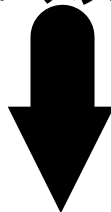
学習データは音高に依存

特徴量テンプレート 多量の音響信号から抽出された
特徴量からなるデータベース

学習データ 特徴量テンプレート中の個々のデータ

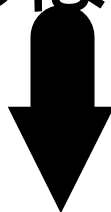
5. 音高による音色変化に 着目した音源同定

本手法では、特徴量テンプレート中の
学習データとの類似度に基づいて同定する。



学習データは音高に依存

あらゆる音高に対して学習データを
用意するのは困難。



有限の学習データから様々な音高に
対応できる手法が必要。

5 .音高による音色変化に着目した音源同定 課 題

課題 1 特徴量によって音高による変化の仕方が異なることをどのように考慮するか .

課題 2 有限の学習データから ,さまざまな音高の特徴量をどのように推定するか .

5 .音高による音色変化に着目した音源同定 課 題

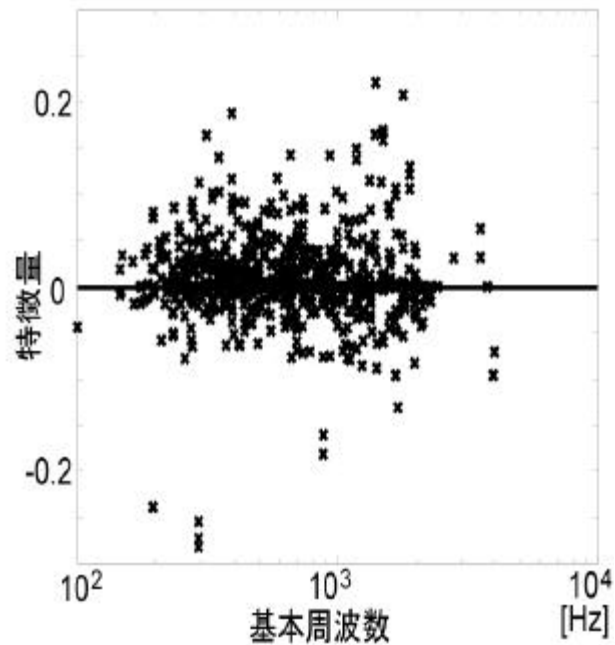
課題 1 特徴量によって音高による変化の仕方が異なることをどのように考慮するか .

音高による変化の仕方で特徴量を**3つに分類** .

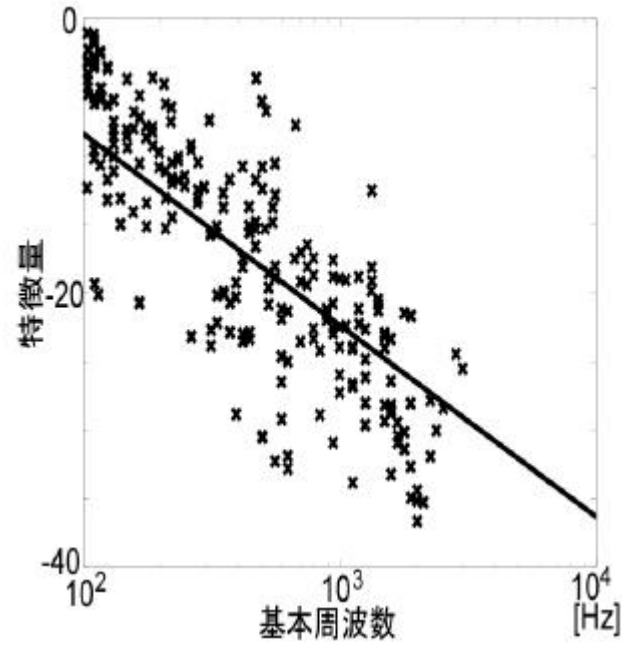
課題 2 有限の学習データから ,さまざまな音高の特徴量をどのように推定するか .

5. 音高による音色変化に着目した音源同定 特徴量の3つの分類

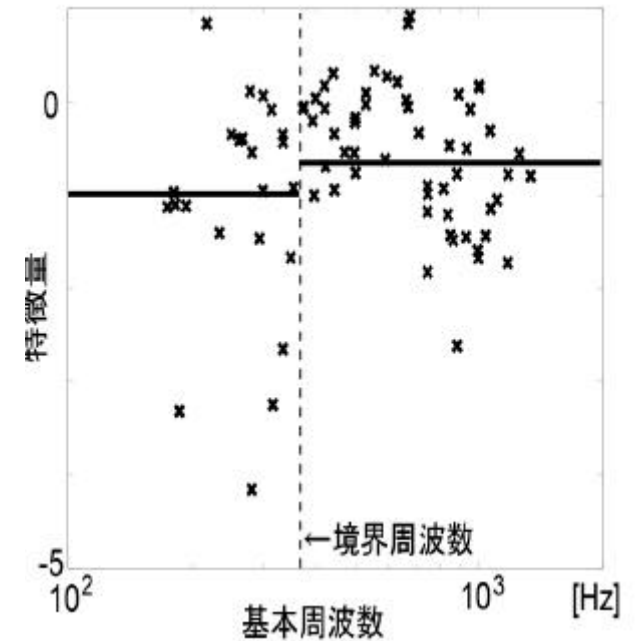
無相関型特徴量



連続変化型特徴量



離散変化型特徴量



5 .音高による音色変化に着目した音源同定 課 題

課題 1 特徴量によって音高による変化の仕方が異なることをどのように考慮するか .

音高による変化の仕方の特徴量を**3つに分類** .

課題 2 有限の学習データから ,さまざまな音高の特徴量をどのように推定するか .

代表値関数と変動値関数を導入 .

(**特徴量の分布を表現する基本周波数の関数**)

5 .音高による音色変化に着目した音源同定 代表値関数

各基本周波数で特徴量を代表する関数 .
(従来の平均値に相当)

(i) 無相関型特徴量の場合

各学習データの特徴量の中央値 .

(ii) 連続変化型特徴量の場合

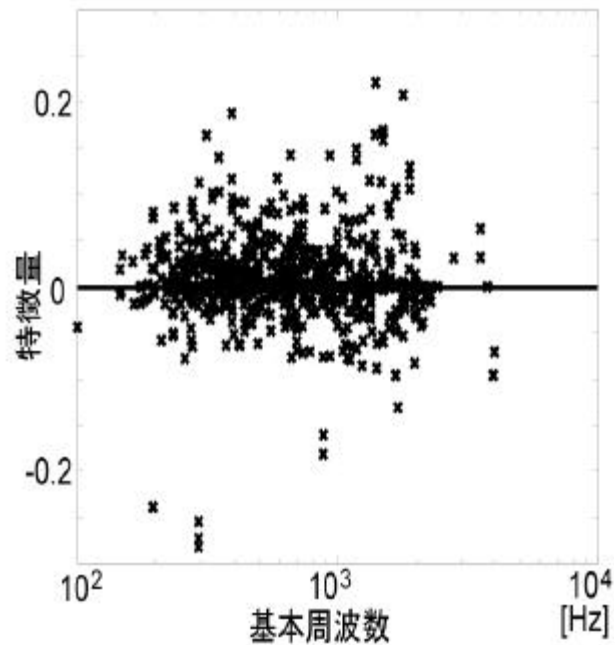
各学習データの特徴量の近似直線 .

(iii) 離散変化型特徴量の場合

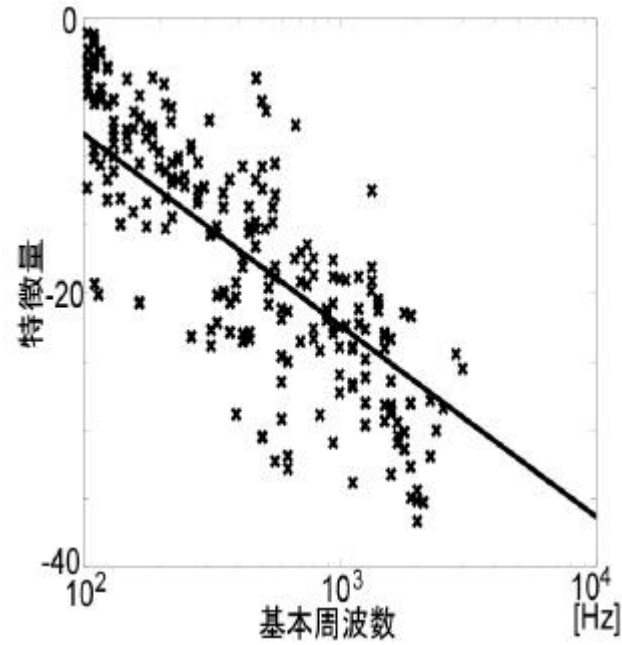
境界周波数で区切られた区間ごとの
各学習データの中央値 .

5. 音高による音色変化に着目した音源同定 特徴量の3つの分類

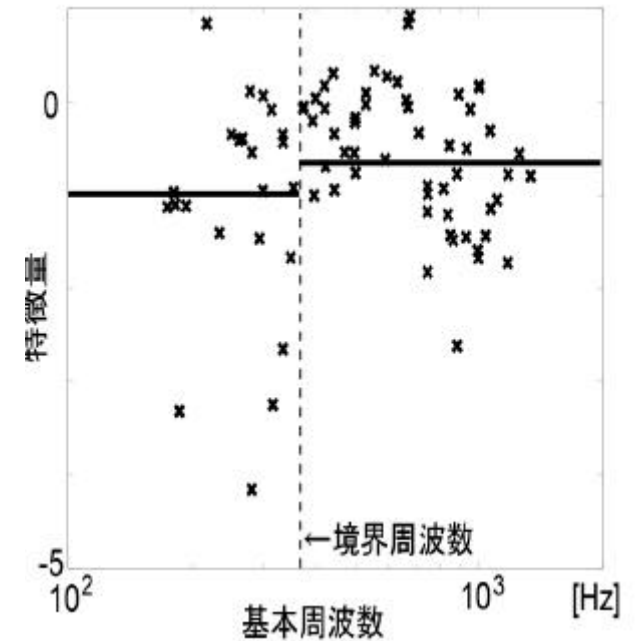
無相関型特徴量



連続変化型特徴量



離散変化型特徴量

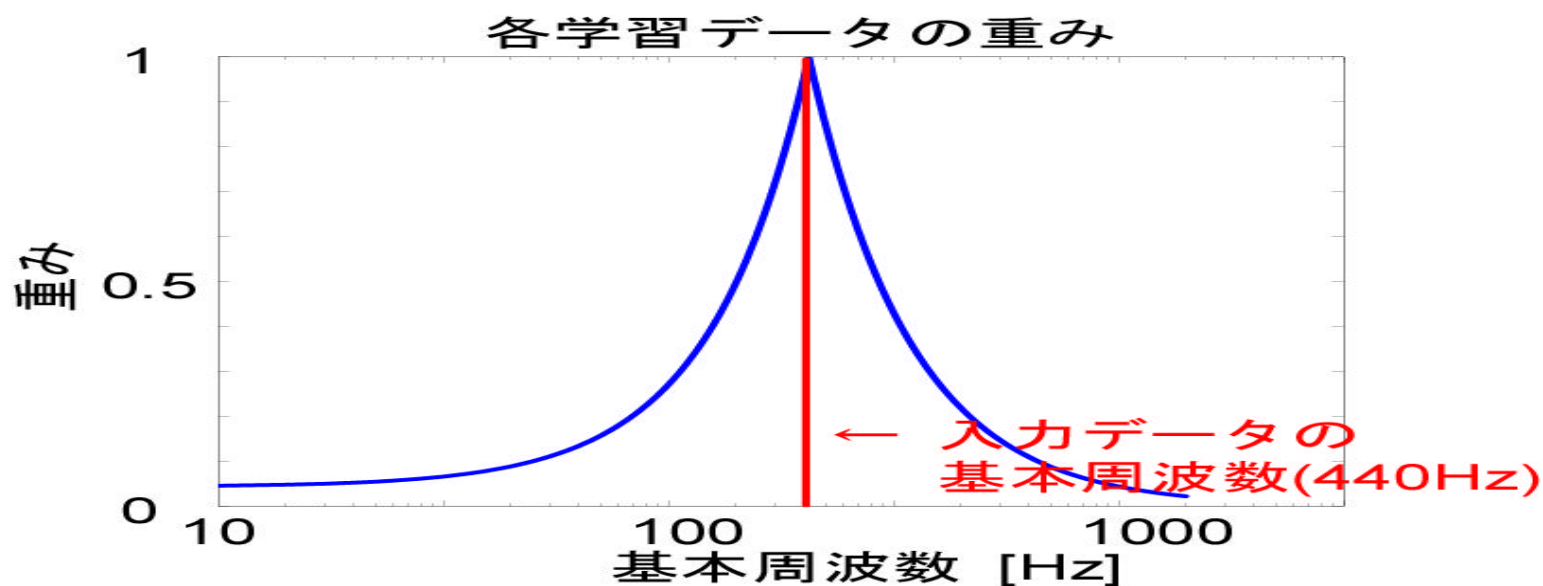


5 .音高による音色変化に着目した音源同定 変動値関数

各基本周波数での

代表値関数からの散らばりの様子を表す関数。
(従来の標準偏差に相当)

各学習データの特徴量と代表値関数との
差の重み付き二乗平均の平方根。



5 .音高による音色変化に着目した音源同定 変動値関数

各基本周波数での

代表値関数からの散らばりの様子を表す関数。
(従来の標準偏差に相当)

変動値関数:

$$\sigma_s^i(f) = \sqrt{\frac{\sum_k w_k(f) \cdot (x_k^i - \mu_s^i(f))^2}{\sum_k w_k(f)}}$$

重み: $w_k(f) = \exp(-|F_k - f|)$

x_k^i : 特徴量 F_k : 学習データの基本周波数 i : 特徴量番号
 $\mu_s^i(f)$: 代表値関数 f : 入力データの基本周波数 s : カテゴリー
 k : 学習データ番号

発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音色変化に着目した音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

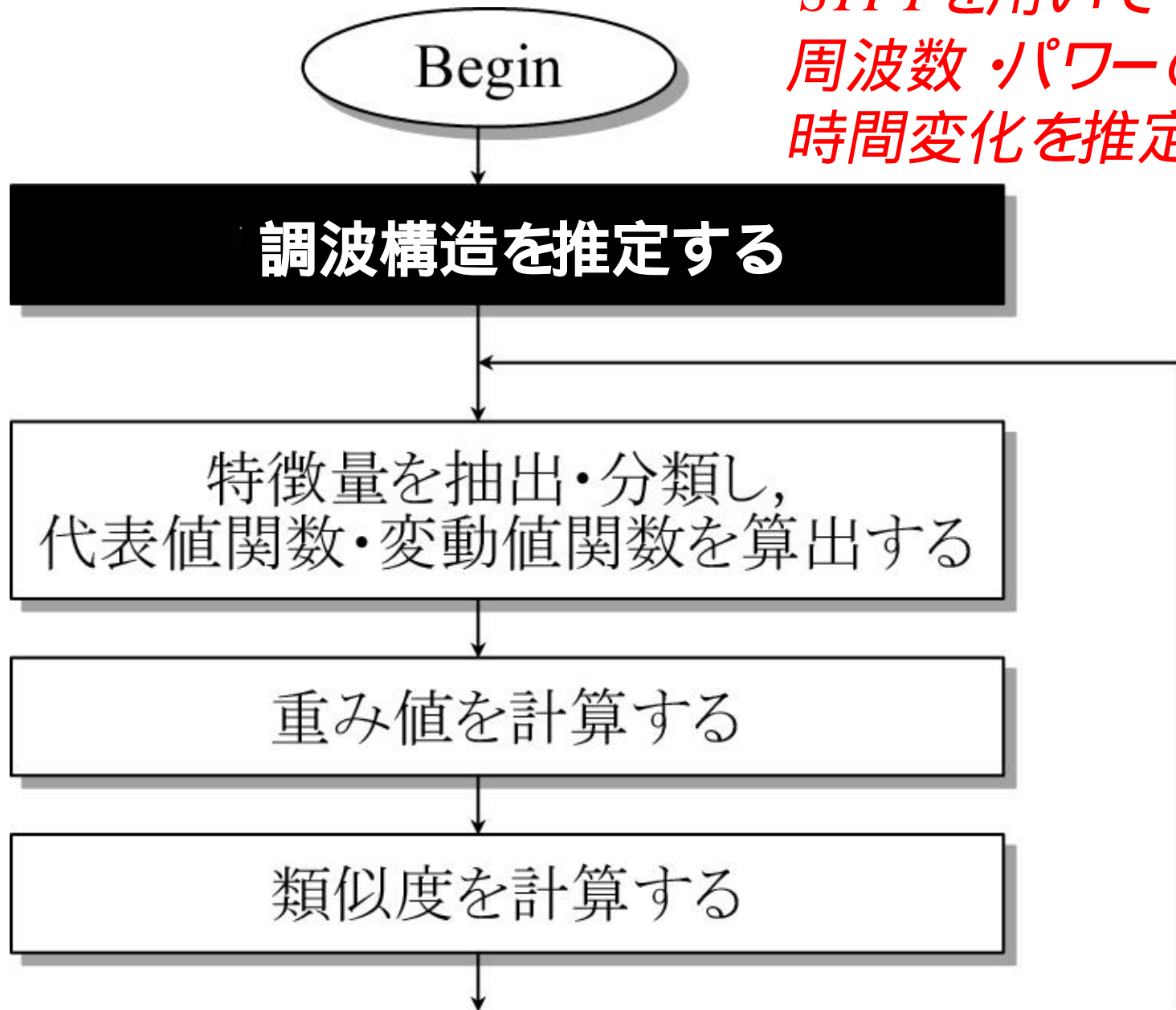
6 .処理の流れ

概 要

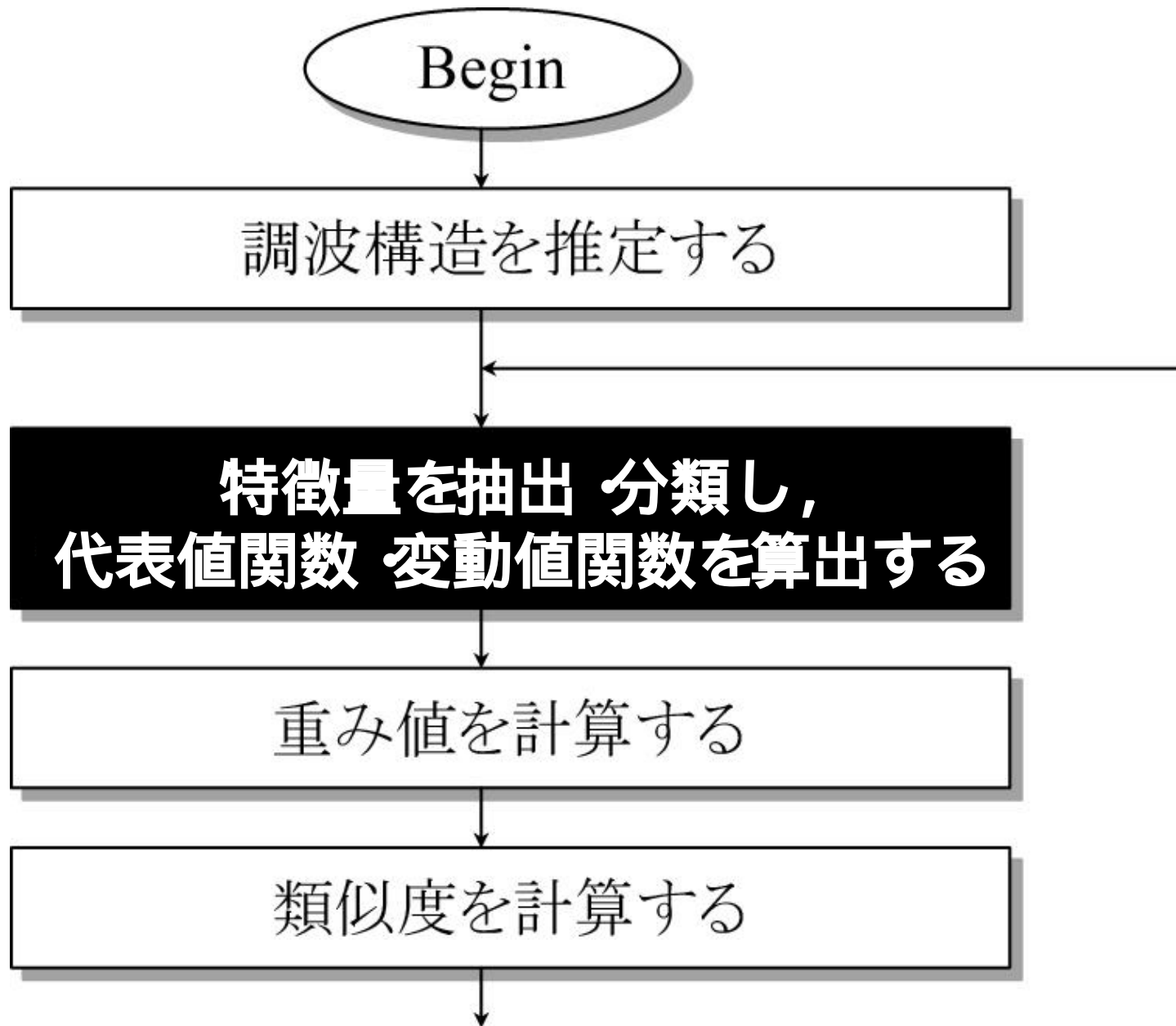
- 入力は単音の音響信号
入力信号に対応する楽器名を出力 .
- 最も抽象度の高い階層からより低い階層へ
反復的に処理 .
- 1回の反復処理で
同一レベルのカテゴリーを同定 .

6 .処理の流れ

*STFTを用いて
周波数・パワーの
時間変化を推定*

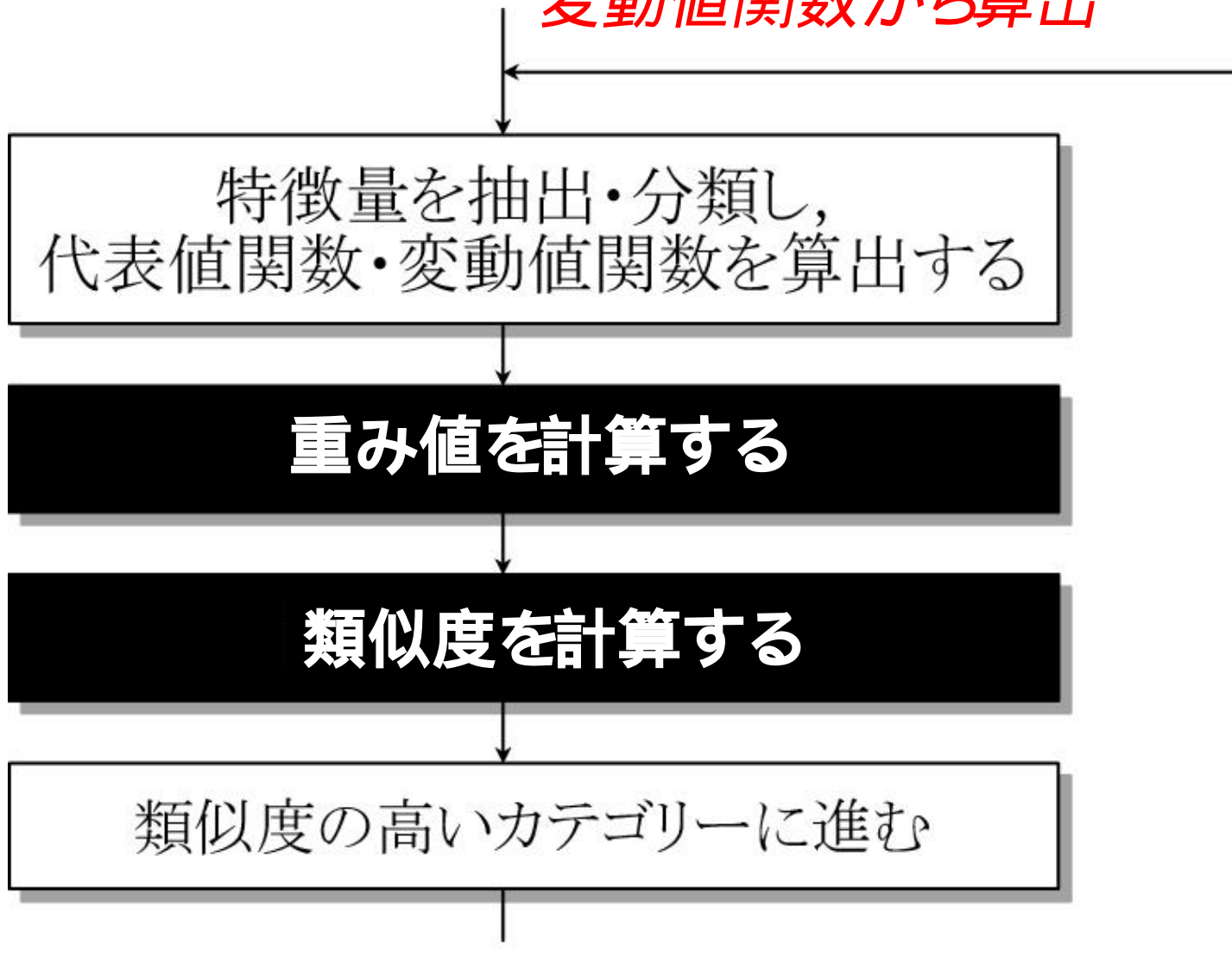


6 .処理の流れ



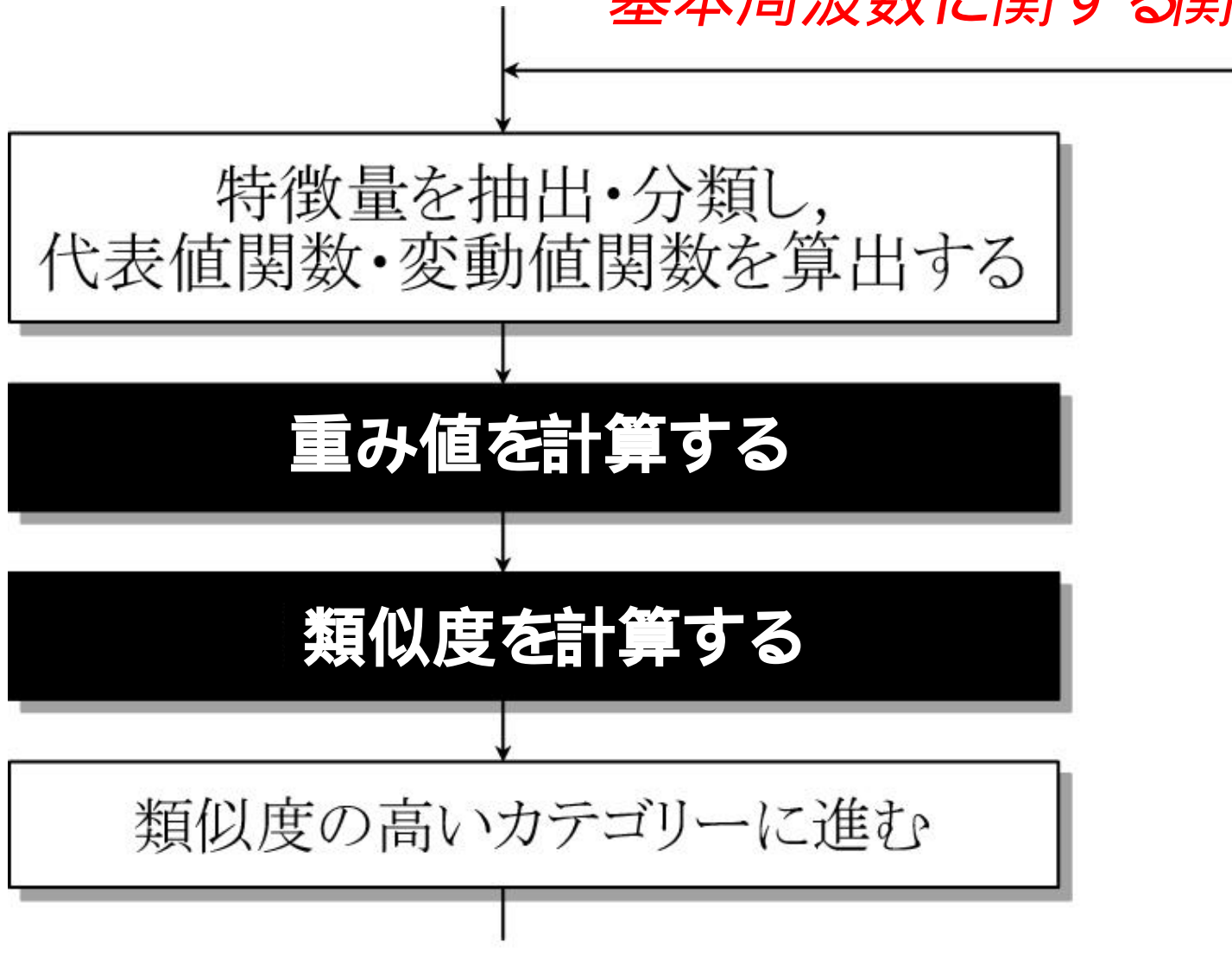
6 .処理の流れ

重み値・類似度を代表値関数
変動値関数から算出

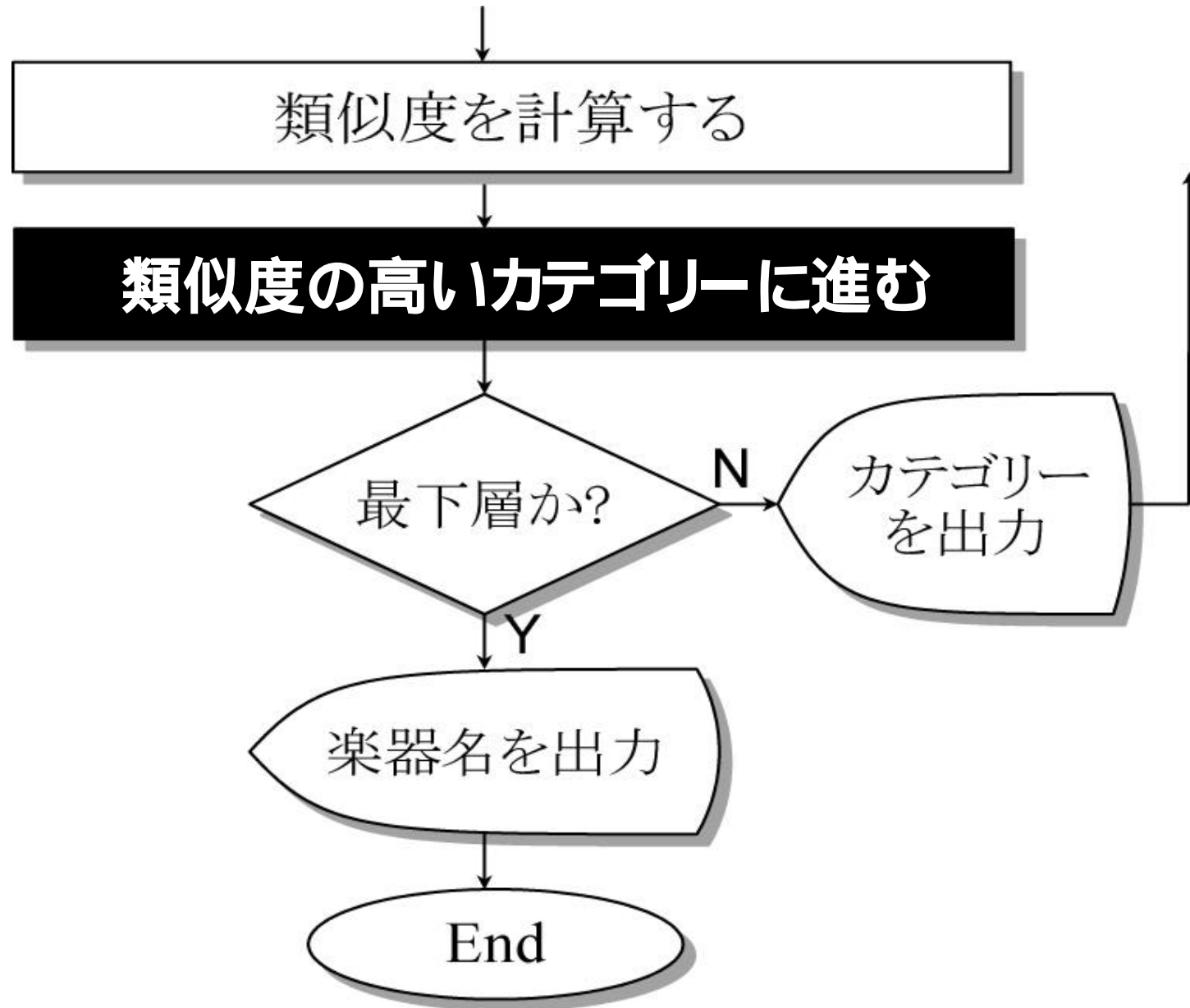


6 .処理の流れ

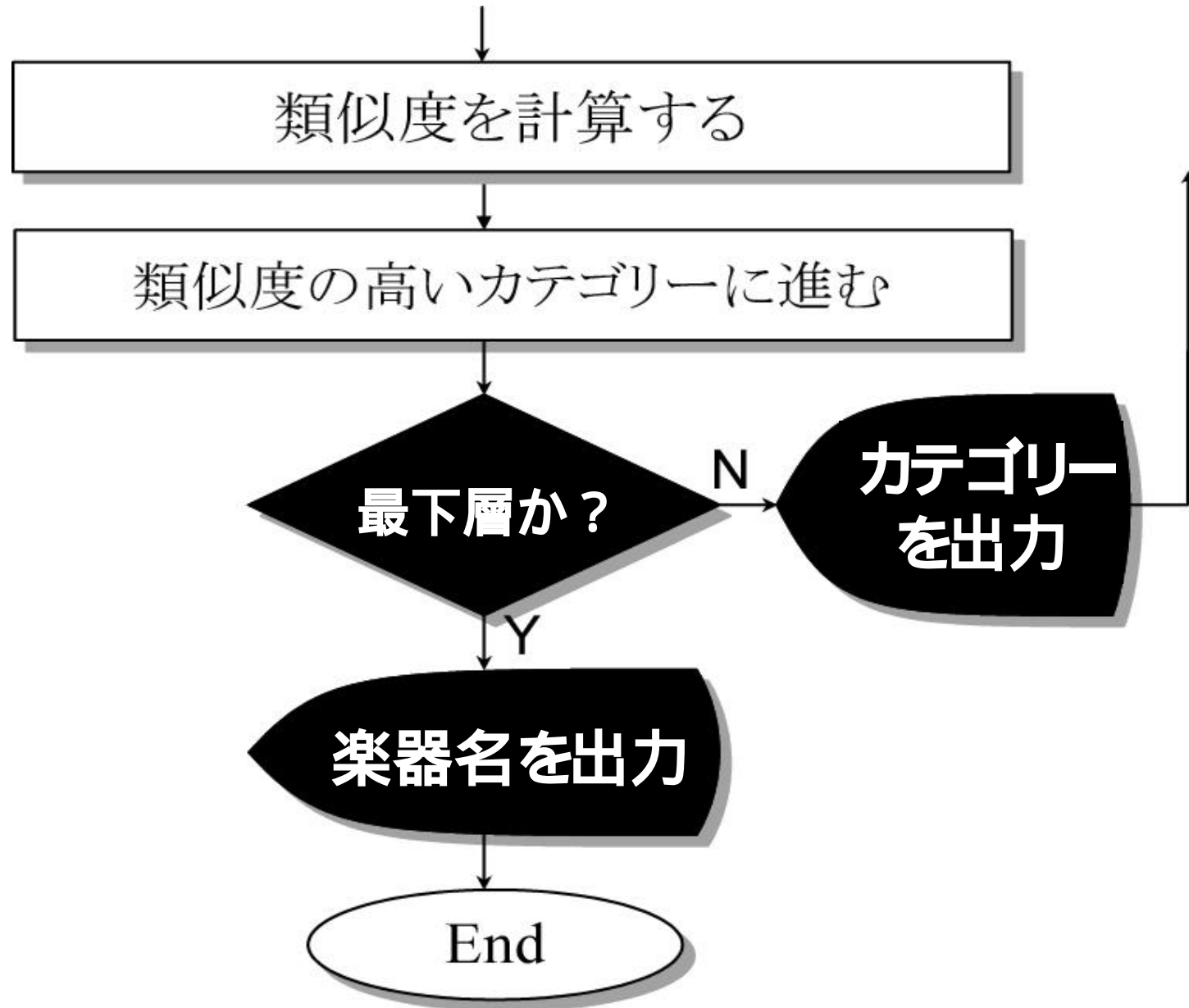
重み値 類似度も
基本周波数に関する関数



6 .処理の流れ



6 .処理の流れ



発表の流れ

1. 導入
2. 音源同
3. 本音源
4. 階層的
5. 音高
6. 処理の
7. 評価実験
8. まとめ

- 評価実験 1
(提案手法の有効性の検証)
 - 1) 実楽器データ
 - 2) MIDI音源
- 評価実験 2
(未知楽器の問題)
 - 1) MIDI音源

7 .評価実験 1- 1

実楽器の単音データベース NTTMSA-P1

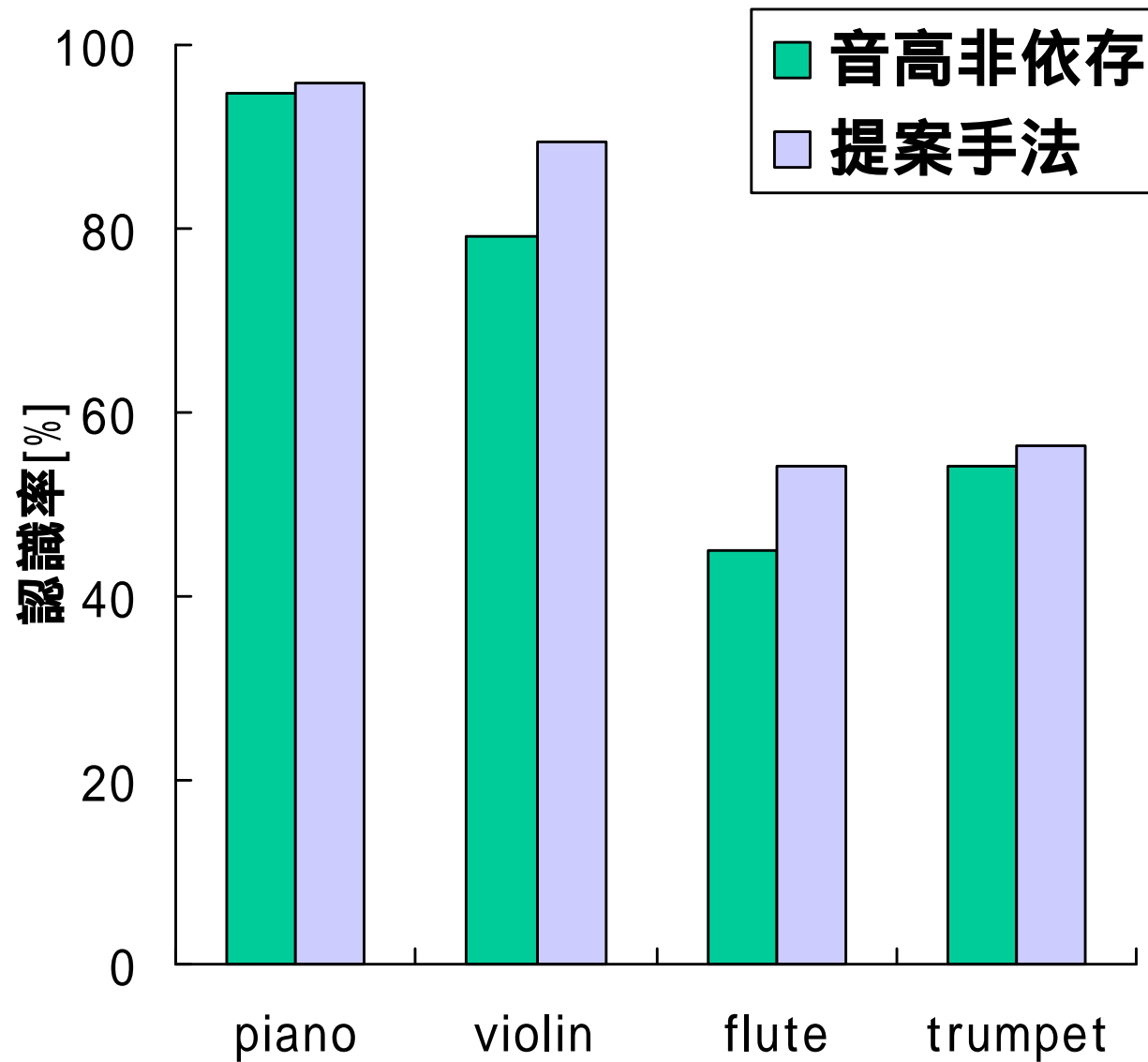
- 実楽器の単独発音を半音ごとに収録
- 各楽器に対して2種類の楽器個体
- 3種類の音の強さ
- 通常の奏法の他にビブラート奏法
(piano, trumpetを除く)
- データ総数: 967個

7 .評価実験 1-1

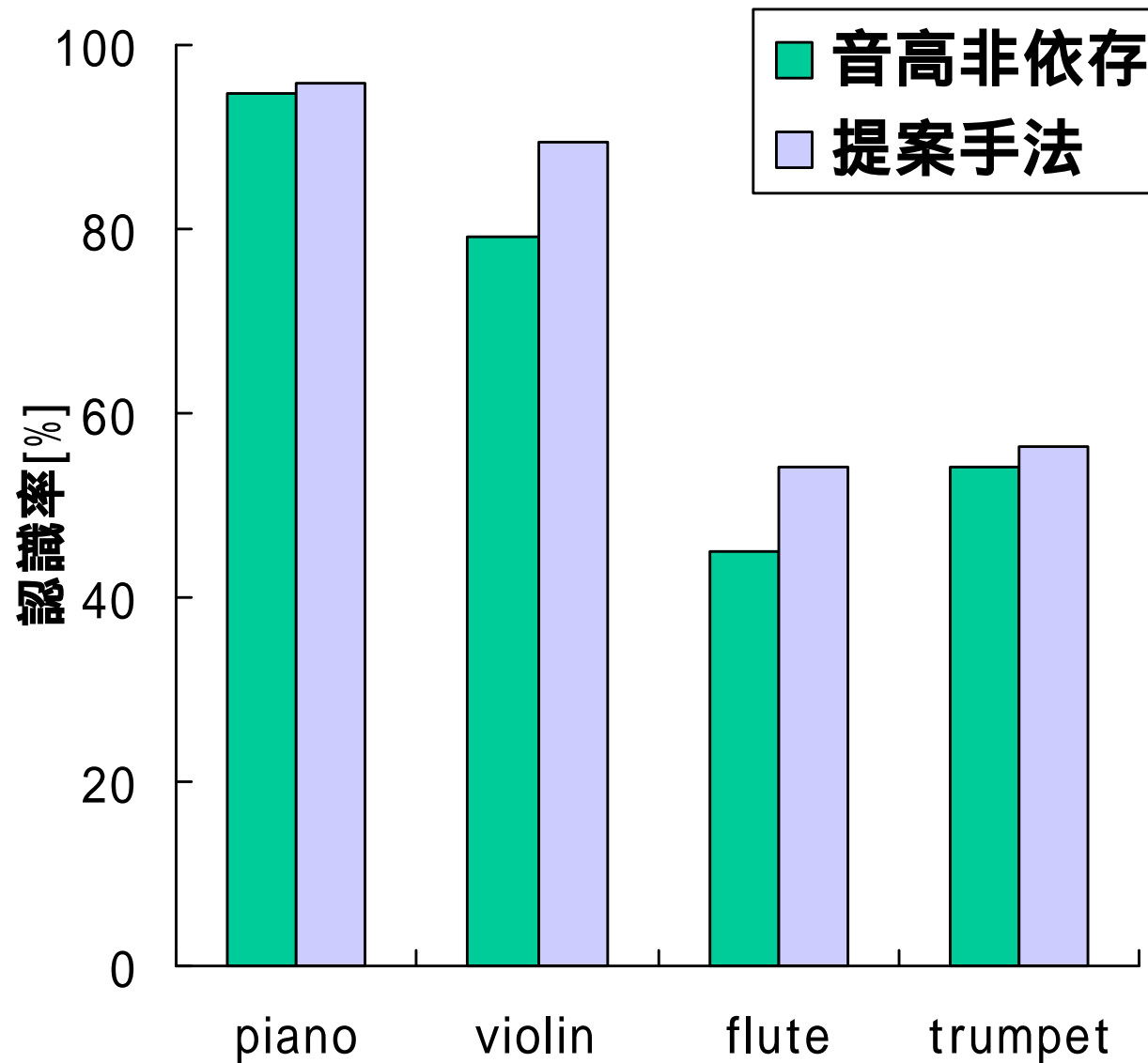
実験方法

- Piano, violin, flute, trumpet のデータを使用 .
- ランダムに2割のデータを学習データに残りを入力データに割り当てる .
- 実験は比較実験とする .
 - ・音高による変化を考慮しない場合 (音高非依存)
 - ・提案手法
- 実験を20回繰り返す .
 - ・学習データにどれを割り当てるかが毎回変化 .

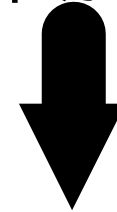
7. 評価実験 1-1 実験結果



7. 評価実験 1-1 実験結果

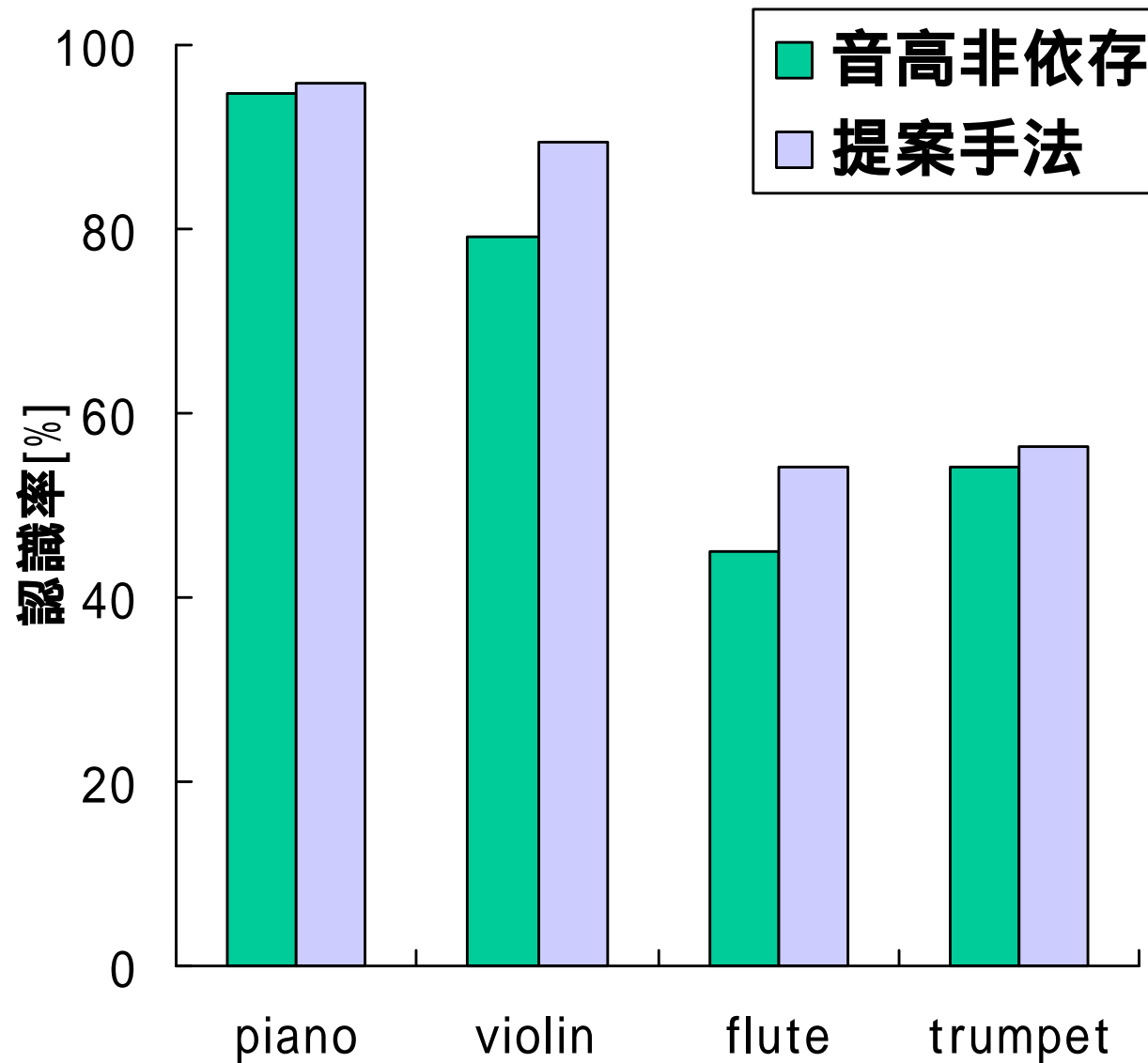


すべての楽器で
提案手法によって
認識率が改善。



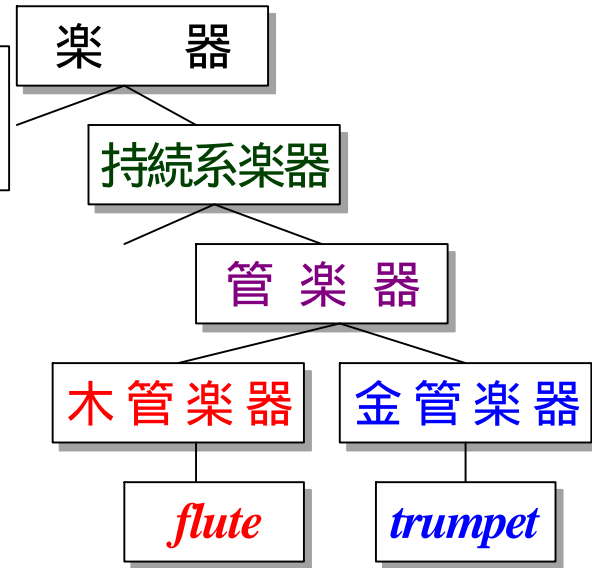
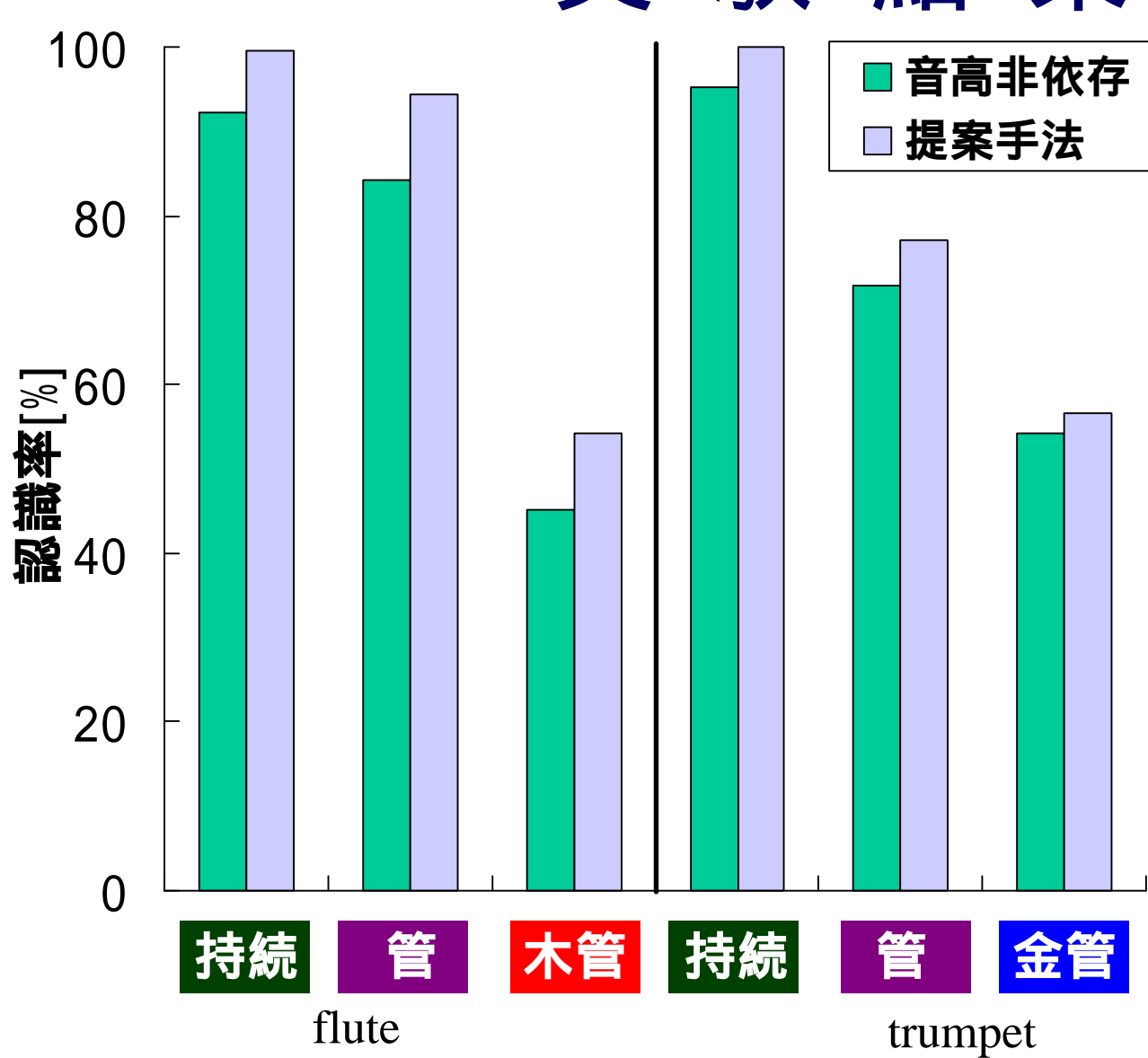
音高による変化の
分類と関数表現が
有効。

7. 評価実験 1-1 実験結果



Flute, trumpet
では ,60%程度 .

7. 評価実験 1-1 実験結果



第2階層までは、
80%程度。

7. 評価実験 1-2

MIDI音源から作成した 単音データベース

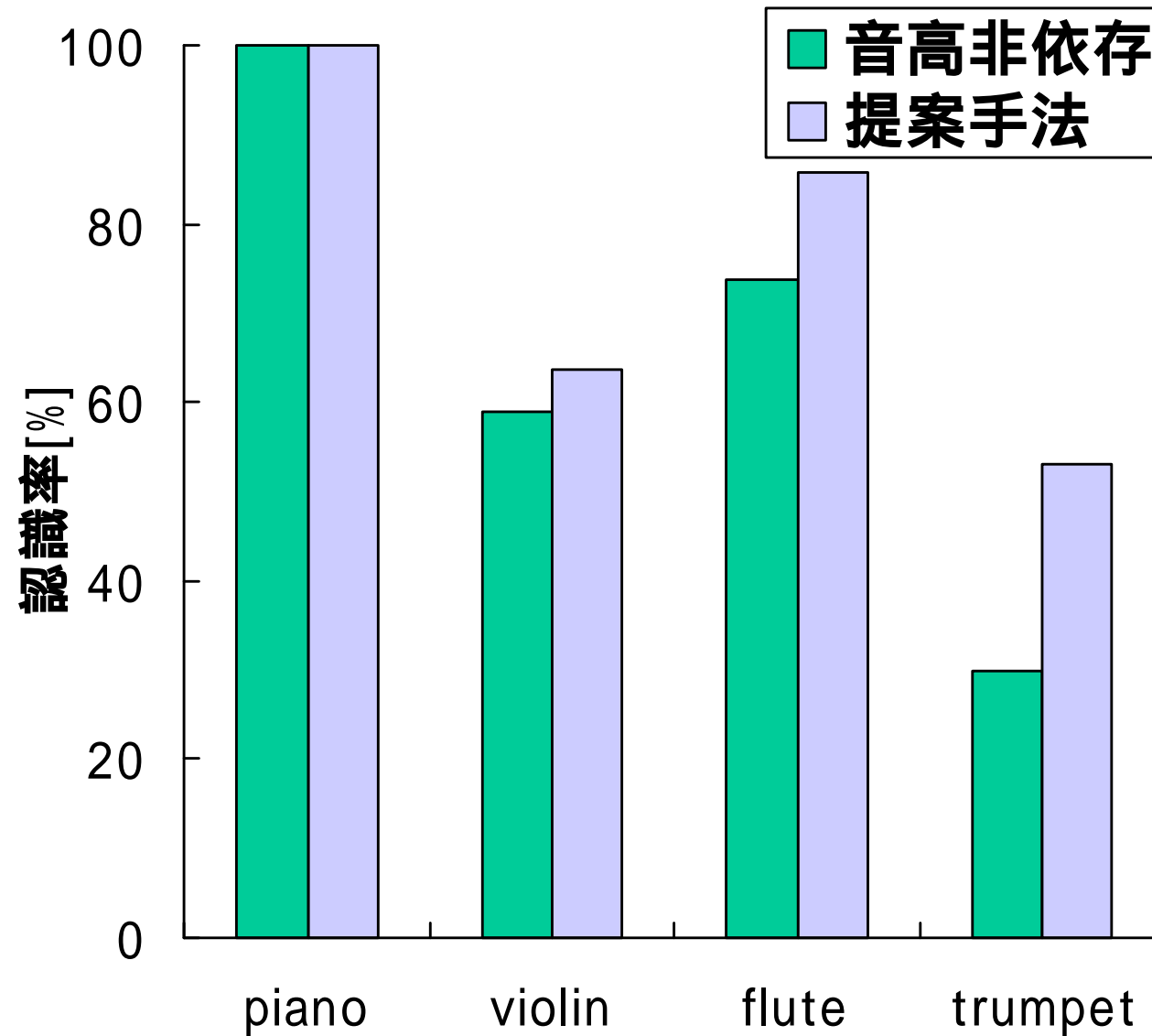
- ヤマハ「MU2000」の単独発音を半音ごとに収録
- 各楽器に対して2種類の楽器個体
- 3種類の音の強さ
- 通常の奏法のみ
- データ総数: 637個

7 .評価実験 1-2

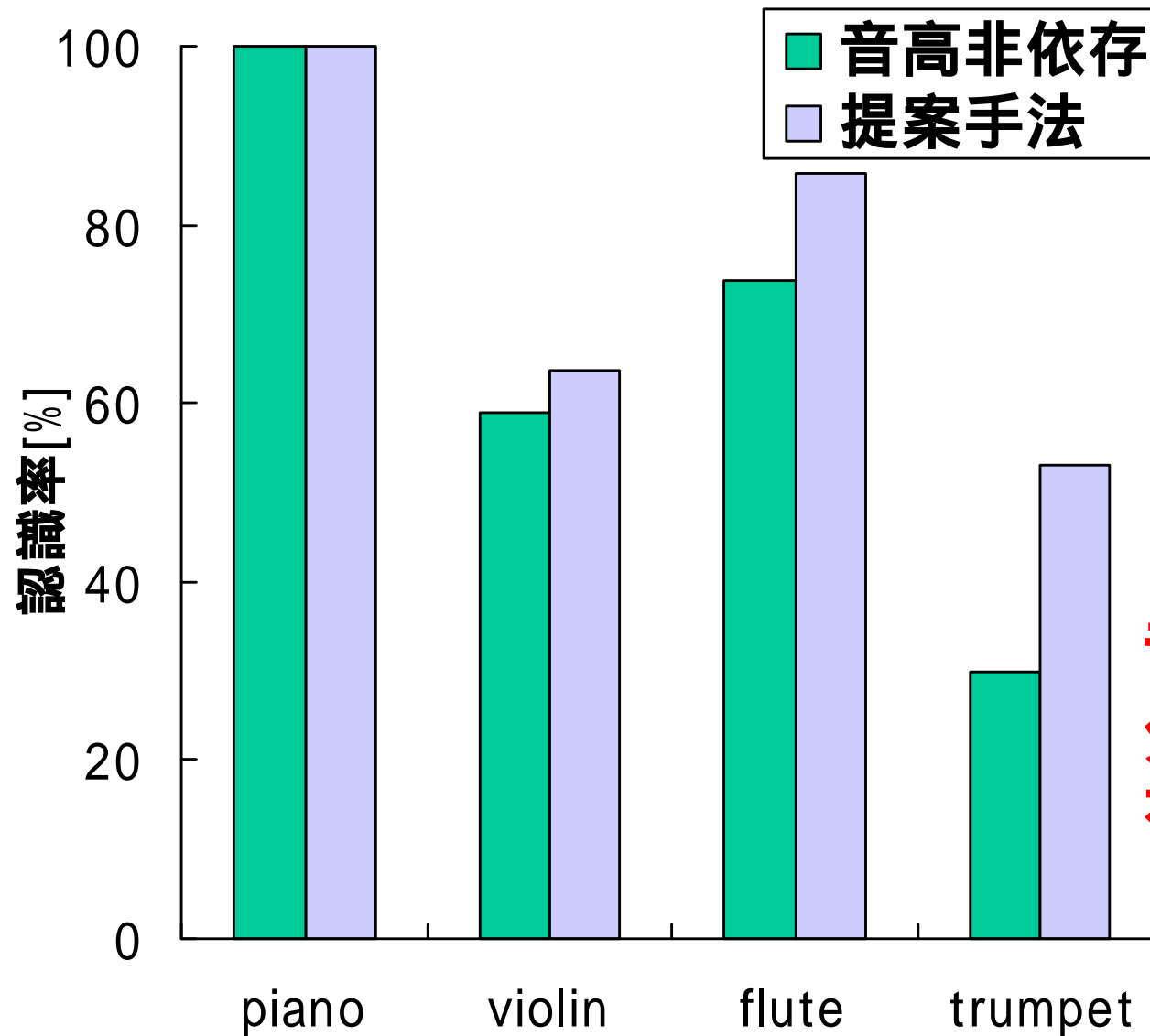
実験方法

- Piano, violin, flute, trumpet のデータを使用 .
- ランダムに2割のデータを学習データに残りを入力データに割り当てる .
- 実験は比較実験とする .
 - ・音高による変化を考慮しない場合 (音高非依存)
 - ・提案手法
- 実験を20回繰り返す .
 - ・学習データにどれを割り当てるかが毎回変化 .

7. 評価実験 1-2 実験結果



7. 評価実験 1-2 実験結果



Piano以外の
すべての楽器で
提案手法によって
認識率が改善。



音高による変化の
分類と関数表現が
有効。

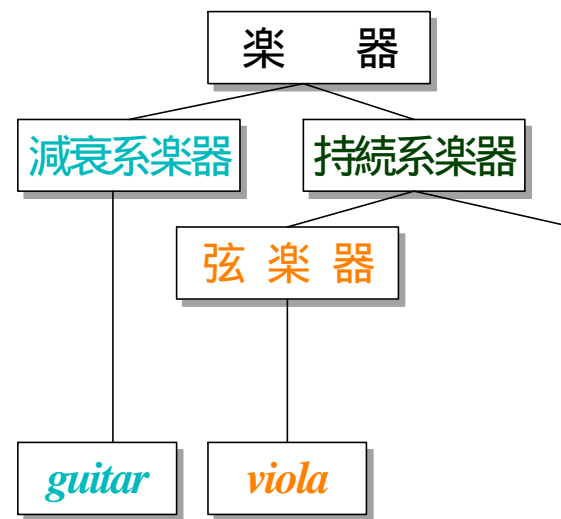
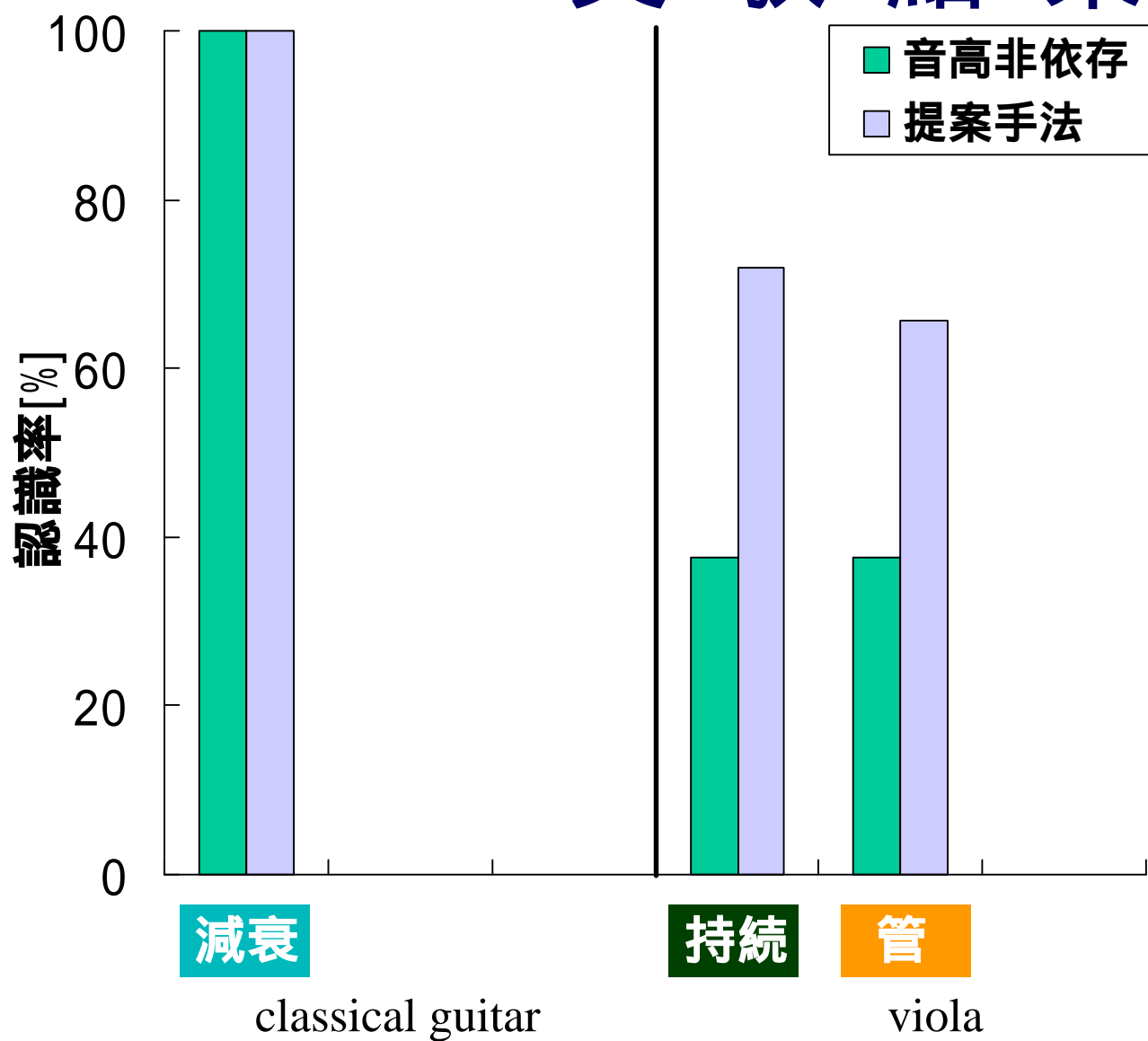
7 .評価実験 2---未知楽器の問題に対する実験---

実験方法

- 学習データには ,MIDI音源から作成した単音データベースを使用 .
- 同音源中の
 - classical guitar (減衰系) ,
 - viola (持続系 | 弦楽器) ,
 - piccolo (持続系 | 管楽器 | 木管楽器) ,
 - clarinet (持続系 | 管楽器 | 木管楽器)の音を入力 .
- 比較実験を行う .
 - ・音高による変化を考慮しない場合 (音高非依存)
 - ・提案手法

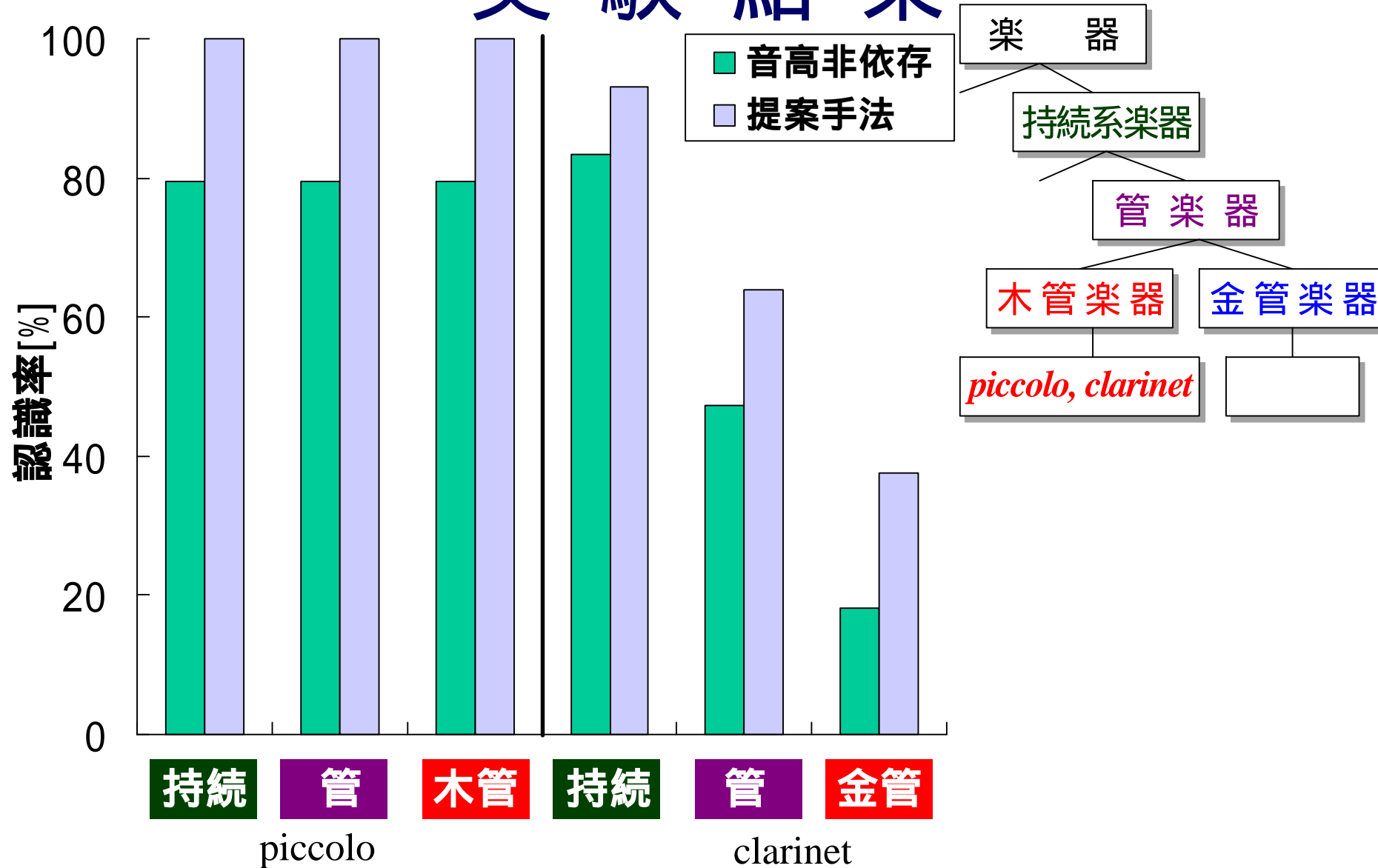
7. 評価実験 2---未知楽器の問題に対する実験---

実験結果



7. 評価実験 2---未知楽器の問題に対する実験---

実験結果



発表の流れ

1. 導入
2. 音源同定における問題点
3. 本音源同定システムの概要
4. 階層的な音源同定
5. 音高による音色変化に着目した音源同定
6. 処理の流れ
7. 評価実験
8. まとめ

8.まとめ

- 各特徴量を音高による変化の仕方
3つに分類。
- 各特徴量の分布を表現する
基本周波数に関する関数を導入。

性能の向上に貢献

- 階層的に処理することで
未知の楽器でも一定の階層まで
同定可能であることを確認。

8.まとめ

今度の課題

- 3つの分類の自動化
- より多くの楽器・奏法への対応
短い音 (打楽器音を含む) への対応
階層の詳細化 など
- 混合音への適用