# 混合音からの特徴量テンプレート作成と音楽的文脈の利用による多重奏の音源同定

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

*ASJ Autumn Meeting 27-29 Sept. 2005*

# 発表の流れ

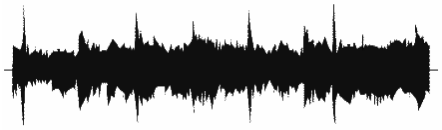
1. 研究の目的 --- 多重奏の音源同定
2. ベースライン手法
3. 性能改善のための3つの課題と解決法
   [課題1] 音の重なりによる特徴変動
   [課題2] 音高による音色変化
   [課題3] 音楽的文脈の利用
4. 評価実験
5. まとめ

# 研究の目的

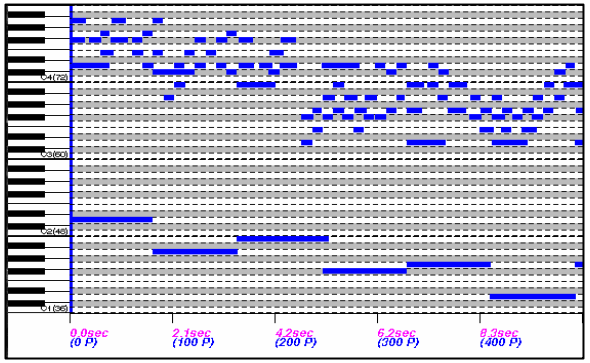
## 多重奏の音源同定

- 音楽の音響信号から演奏中の楽器を同定
- 自動採譜，音楽情報検索などに有用

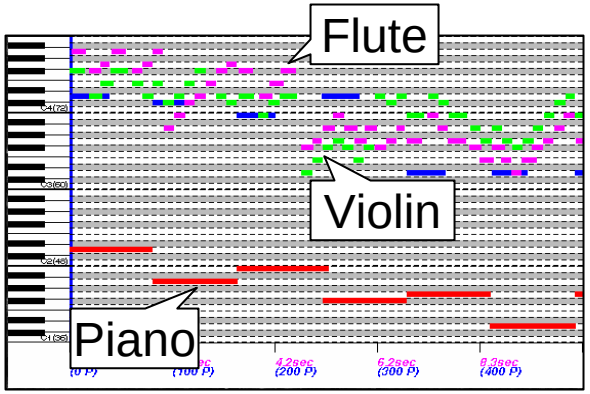
一般的な処理の流れ



音響信号



各単音の発音時刻・音高・音長を推定




各単音の楽器名を同定

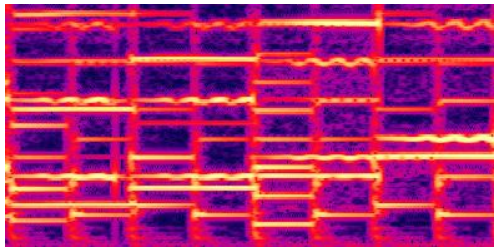
# ベースライン手法

Audio signal

STFT

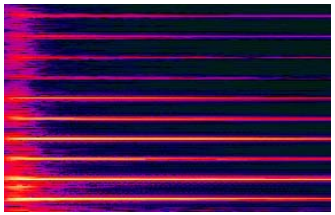
Musical notes

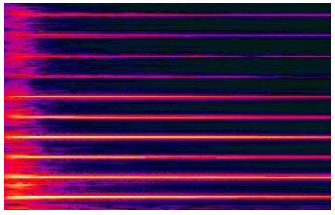
| | | |
|---|---|---|
| 1) | 0--1,000ms | C4 |
| 2) | 0-- 500ms | C2 |
| 3) | 500--1,000ms | G2 |
| 4) | 1,500--2,000ms | C4 |
| | …… | |

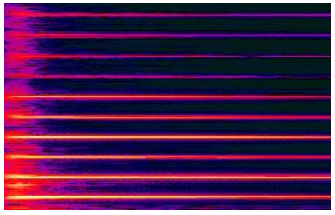


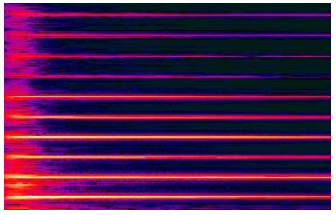
Spectrogram

Harmonic structure extraction

1) 

2) 

3) 

4) 

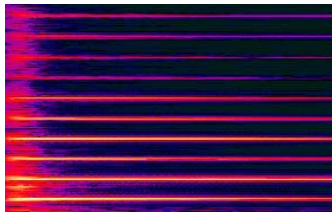
……

つづく

既存手法（e.g. HTC）で推定
今回は正解を付与

# ベースライン手法

各単音の調波構造抽出後，以下の処理を単音ごとに



Feat. extract.

Dim. reduct.

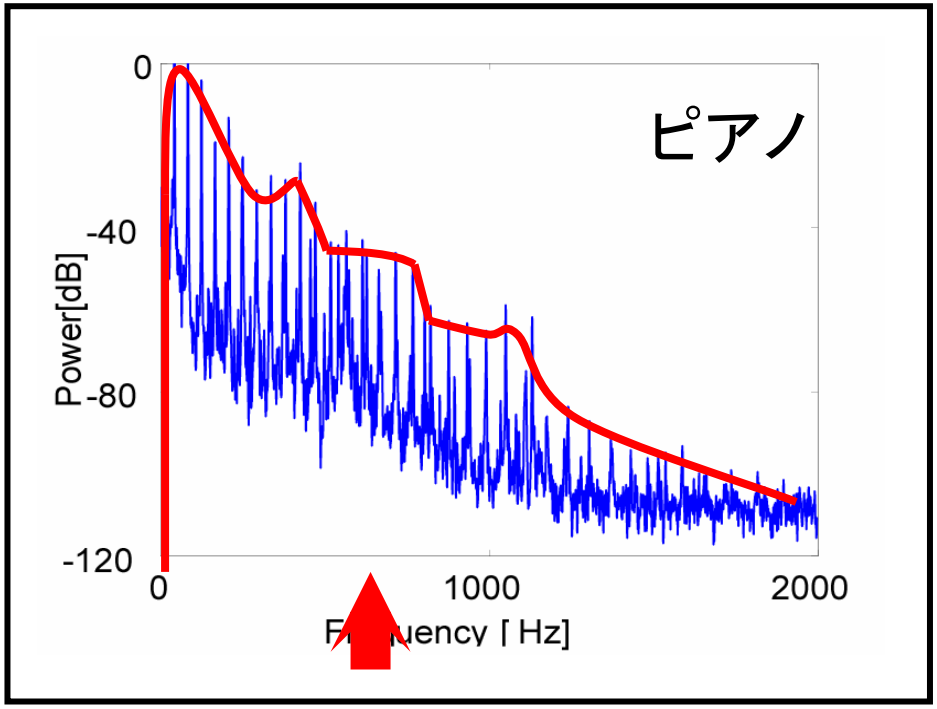
A post. prob. calculation

$p(\omega_i \mid \mathbf{x}_k)$

Instr. determ.

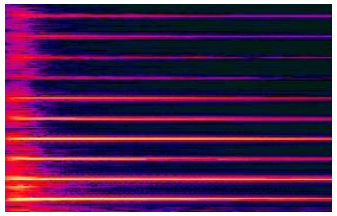
**Violin**

43個の特徴量を抽出
e.g. 周波数重心（時間軸方向の中央値）
パワー崩落の近似直線の傾き



# ベースライン手法

各単音の調波構造抽出後，以下の処理を単音ごとに



Feat. extract.

Dim. reduct.

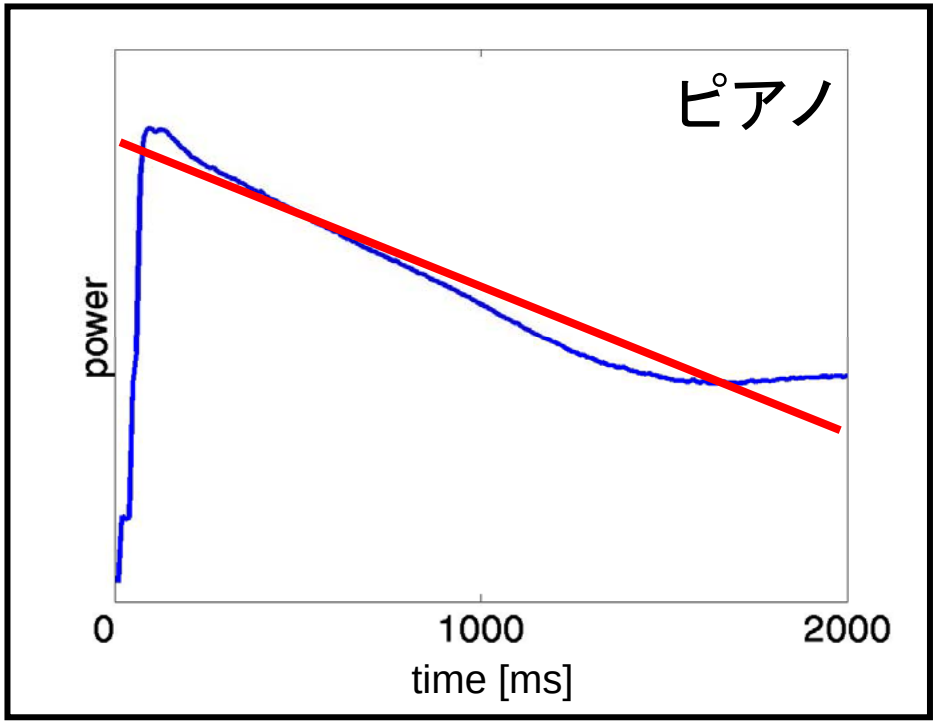
A post. prob. calculation

$p(\omega_i \mid \mathbf{x}_k)$

Instr. determ.

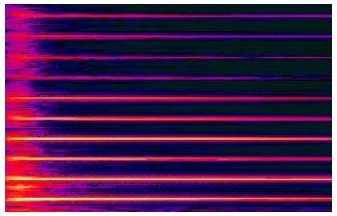
**Violin**

43個の特徴量を抽出
e.g. 周波数重心（時間軸方向の中央値）
　　パワー崩落の近似直線の傾き



# ベースライン手法

各単音の調波構造抽出後，以下の処理を単音ごとに



Feat. extract.

Dim. reduct.

A post. prob. calculation

$p(\omega_i \mid \mathbf{x}_k)$

Instr. determ.

**Violin**

43個の特徴量を抽出
e.g. 周波数重心（時間軸方向の中央値）
　　パワー崩落の近似直線の傾き

主成分分析（PCA）&線形判別分析（LDA）で
43→4次元に圧縮

楽器毎に $p(\omega_i \mid \mathbf{x}_k)$ を計算
（特徴ベクトル$\mathbf{x}_k$が楽器$\omega_i$のものである確率）
あらかじめ用意した学習データとの
特徴空間における距離に基づいて計算

$p(\omega_i \mid \mathbf{x}_k)$ が最大となる楽器を出力

# 課　　題

［課題１］ 音の重なりによる特徴変動

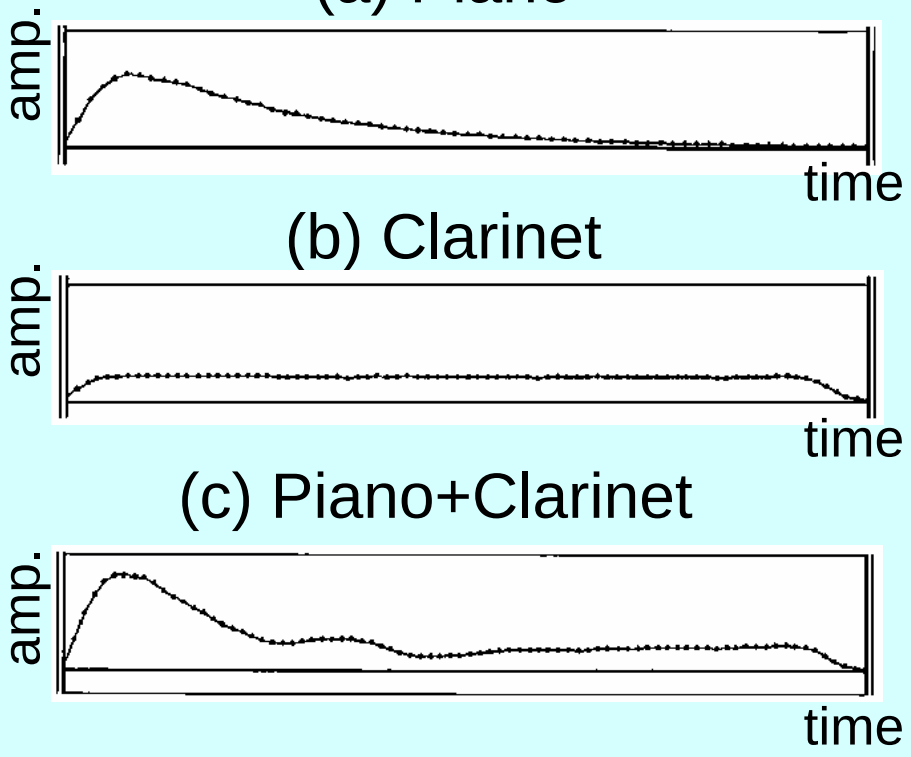
［課題２］ 音高による音色変化

［課題３］ 音楽的文脈

# [課題1] 音の重なりによる特徴変動

複数楽器が同時に鳴る
- ⇨周波数成分が重複
- ⇨周波数成分が干渉
- ⇨そこから抽出した
  特徴量が変化する

**周波数成分の重なりの例**




**ね ら い**

「**特徴量の重み付け**」

影響の小さな特徴量には大きな重みを

影響の大きな特徴量には小さな重みを

# [課題1] 音の重なりによる特徴変動

**課題**

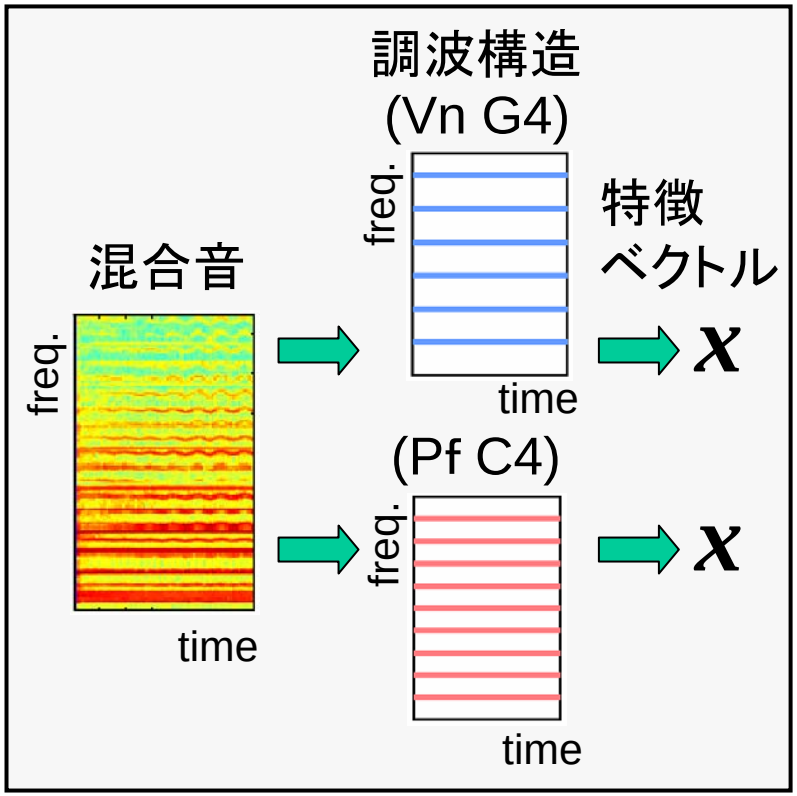
音の重なりの影響の大きさを
どのように測るか

**アイディア**

音の重なりによる特徴変動が
大きければ，「クラス内分散」
が大きくなるはず

**解決策**

**混合音から得た学習データ**に
**線形判別分析**（DAMS法）



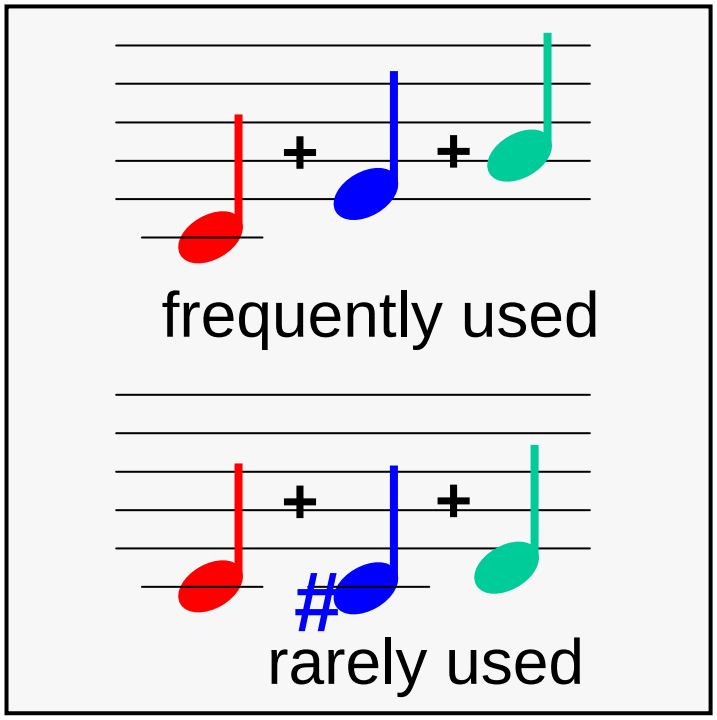

$$\frac{\text{クラス間分散}}{\text{クラス内分散}}$$ 最大化

# [課題1] 音の重なりによる特徴変動

**混合音学習データ作成における問題**

**組合せ爆発**

音の混合パターンが無限にあり，
混合音の網羅的収集は不可能




**ね ら い**

音楽で使われる組み合わせは
もっと少ないはず
⇒音楽で使われるものだけ収集

**実楽曲の楽譜**から
混合音を作成

# [課題2] 音高による音色変化

## 音高による音色変化

楽器は他の音に比べて広音域

⇒音高による音色変化が顕著

## 解　決　策

**F0依存多次元正規分布**

特徴量の音高依存性を

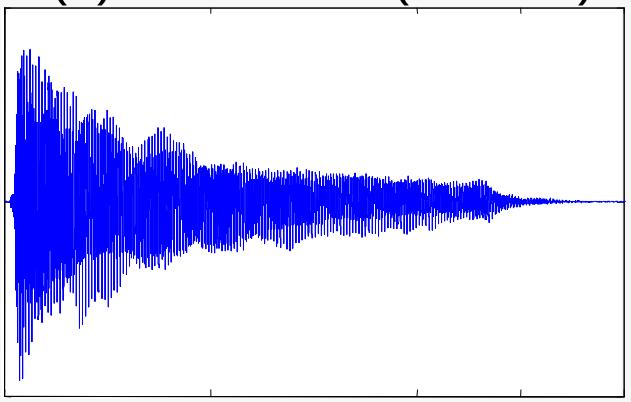
F0の関数で近似

情処学論'03，音講論'02秋で提案

（ただし単一音のみの実験）

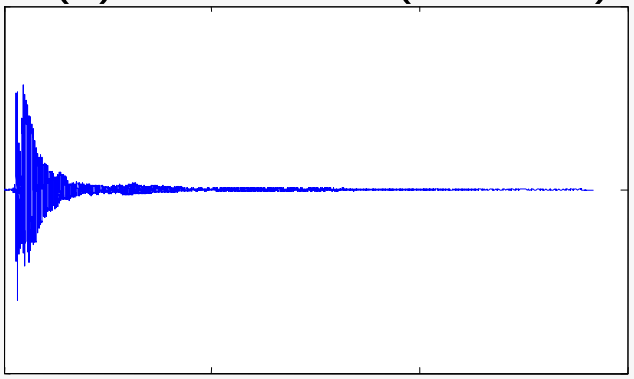
ピアノの場合

(a) Pitch = C2 (65.5Hz)



time

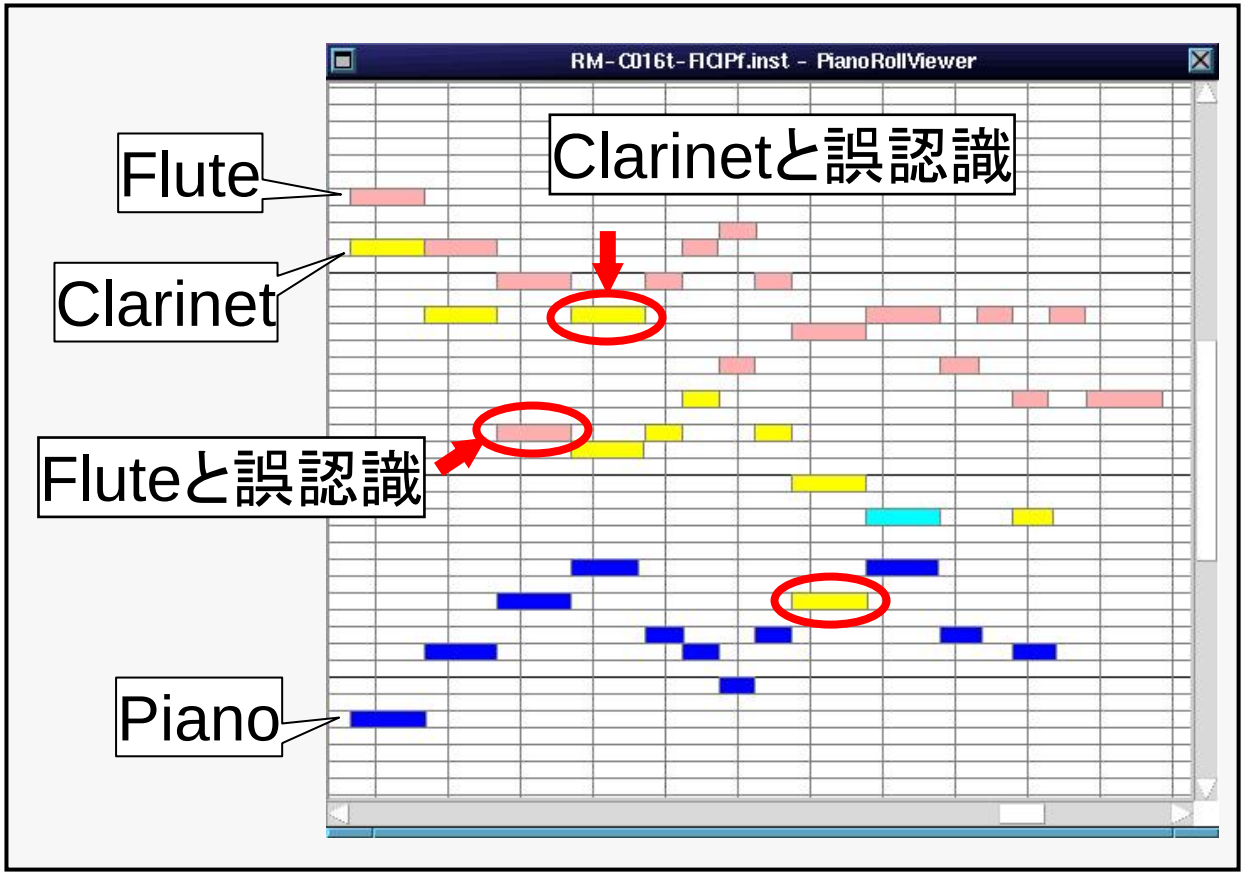
(b) Pitch = C6 (1048Hz)



time

# [課題3] 音楽的文脈の利用

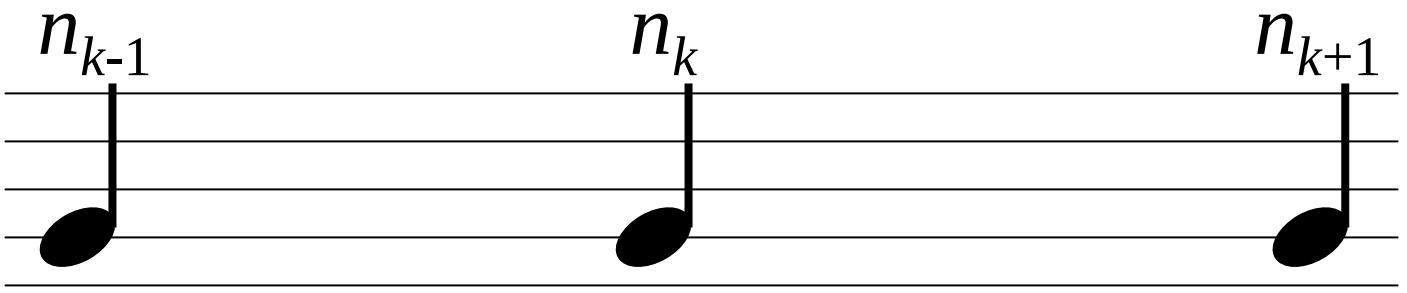
単音ごとに独立に同定処理をすると，
音楽の流れに合わない認識結果



# [課題3] 音楽的文脈の利用

**基本的アイディア**

事前確率に文脈（前後関係）を反映させる




事後確率 $p(\omega_i \mid \boldsymbol{x}_{k-1})$

事後確率 $p(\omega_i \mid \boldsymbol{x}_k) = \dfrac{p(\boldsymbol{x}_k \mid \omega_i)p(\omega_i)}{p(\boldsymbol{x}_k)}$

事後確率 $p(\omega_i \mid \boldsymbol{x}_{k+1})$

事前確率

$$p(\omega_i) = \alpha \prod_j p(\omega_i \mid \mathbf{x}_j)$$

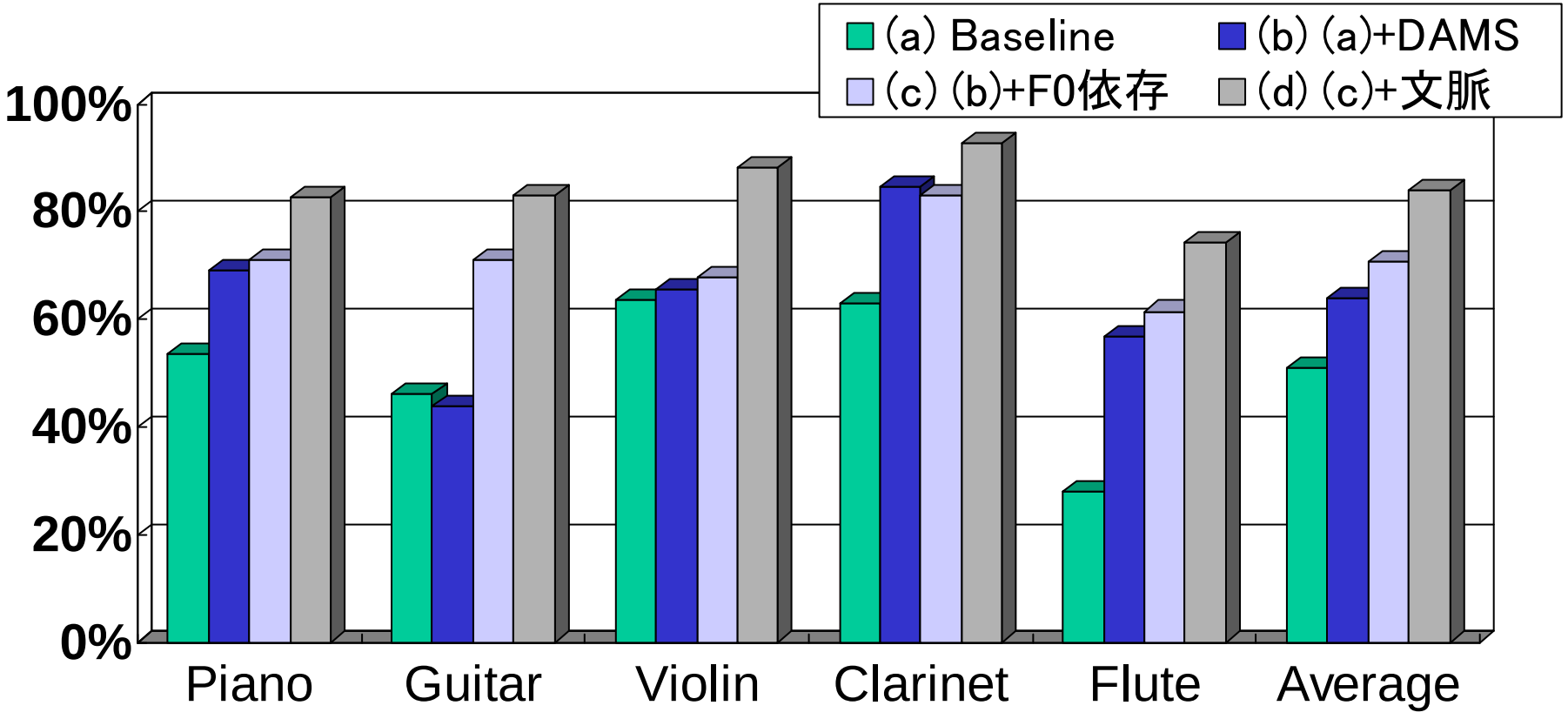
前後の単音の事後確率に基づいて計算する

# 評価実験／データ作成

- 単一音の楽器音データベースの音響信号を
  SMFに基づいて切り貼りして多重奏を作成
- 使用SMF：RWC-MDB-C-2001 No.13, 16, 17
  上記から2～4パートを抽出した二～四重奏
- 元となる音響信号：RWC-MDB-I-2001
- 使用楽器：
  Piano, Classical Guitar, Violin, Clarinet, Flute
- 楽曲単位でLeave-one-out法
  （1曲抜いて学習して残りの1曲を評価）

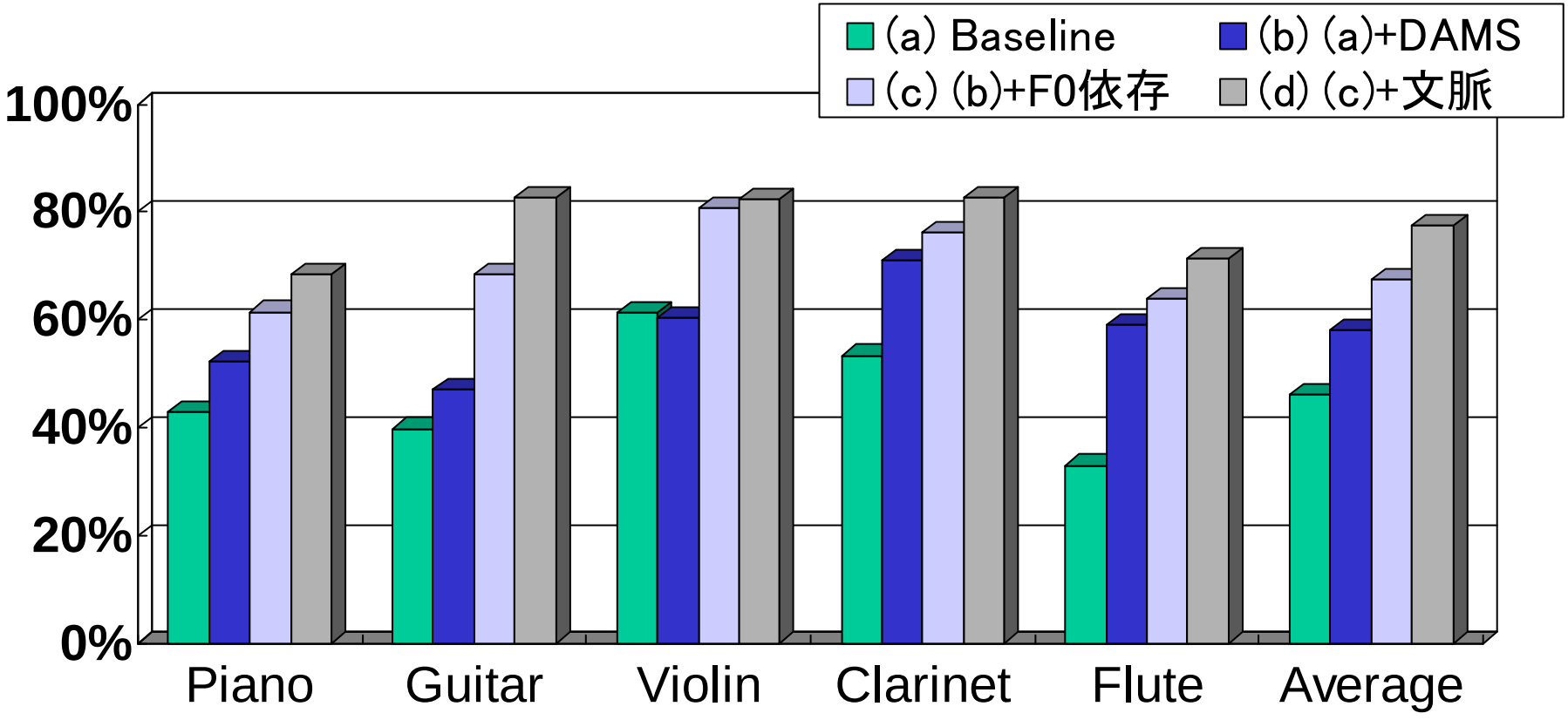
# 評価実験／結果

認識：二重奏，学習：単旋律＋二重奏＋三重奏



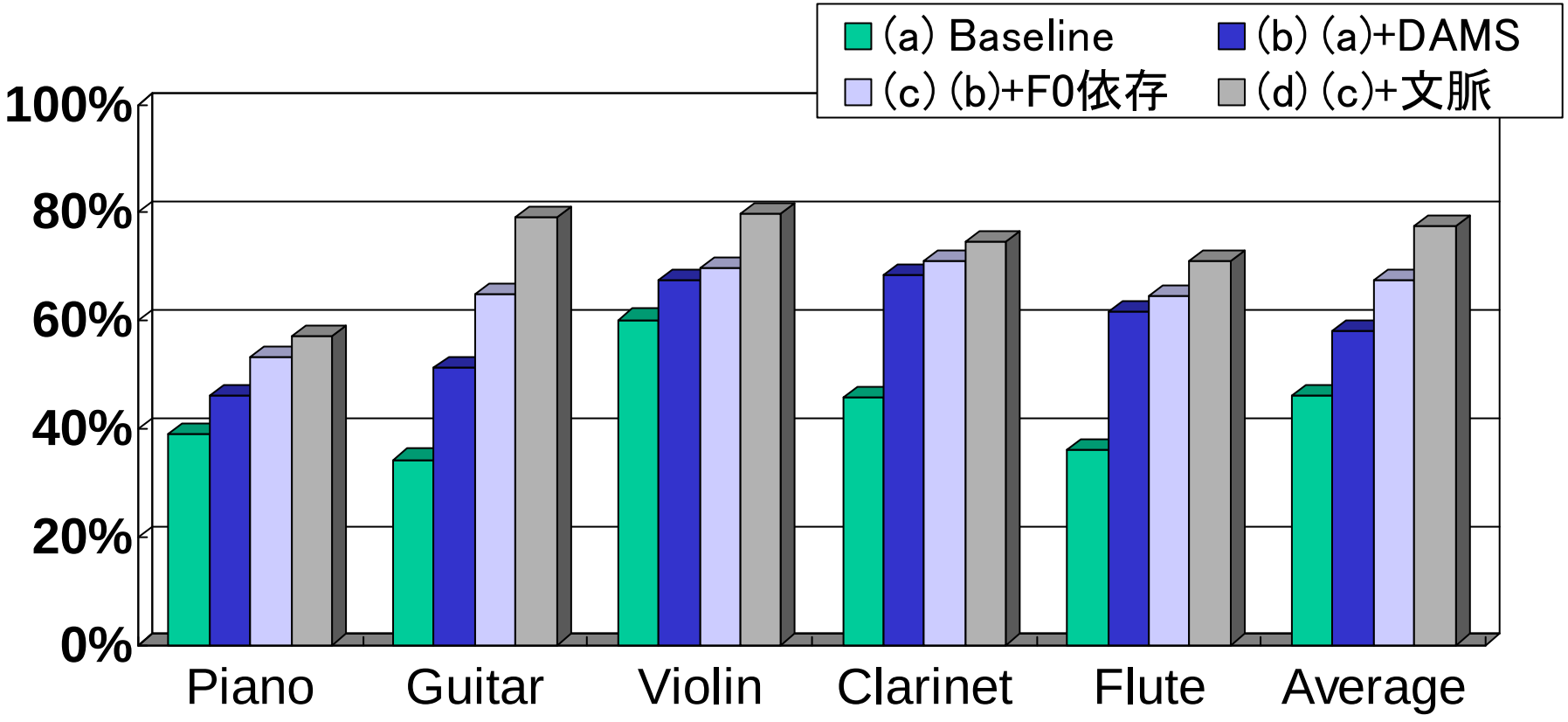
# 評価実験／結果

認識：三重奏，学習：単旋律＋二重奏＋三重奏



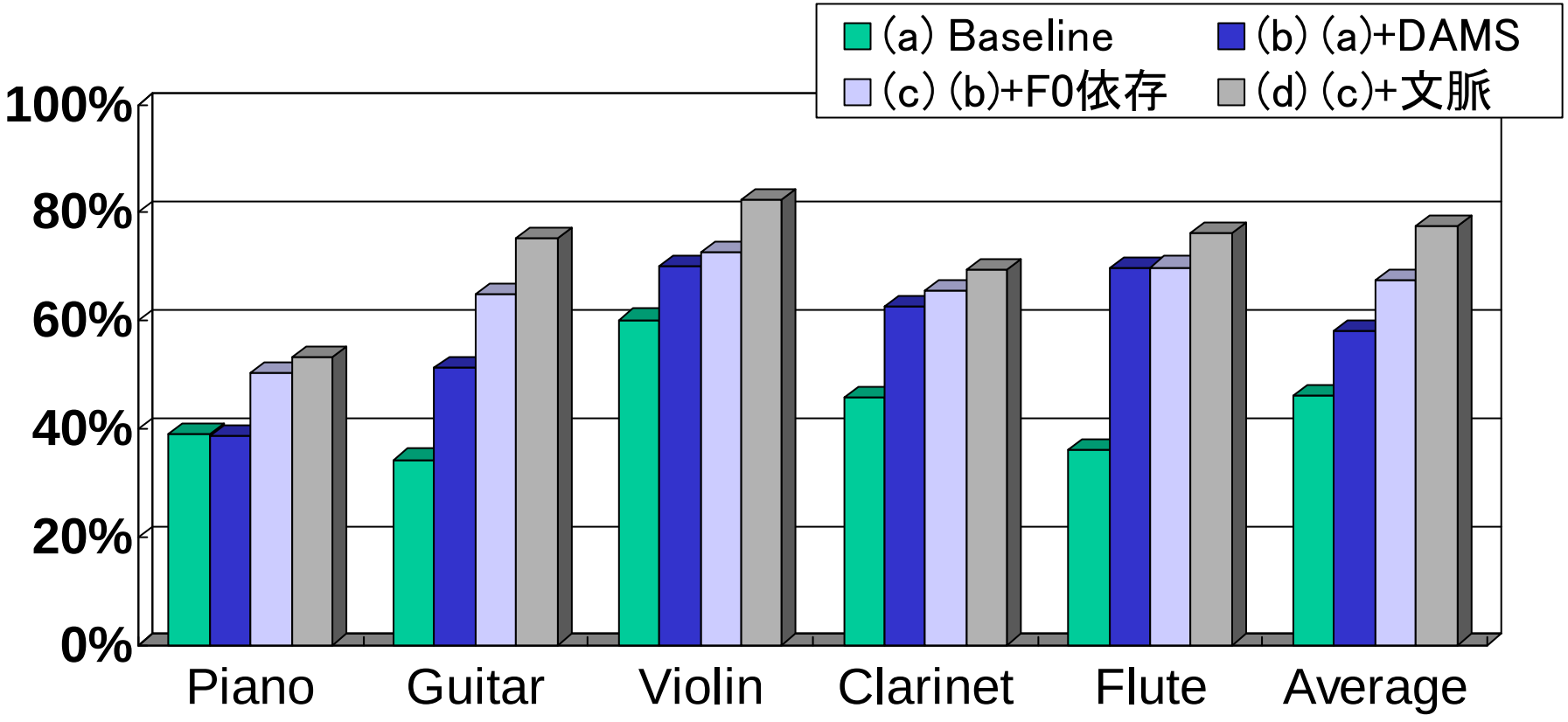
# 評価実験／結果

認識：四重奏，学習：単旋律＋二重奏＋三重奏

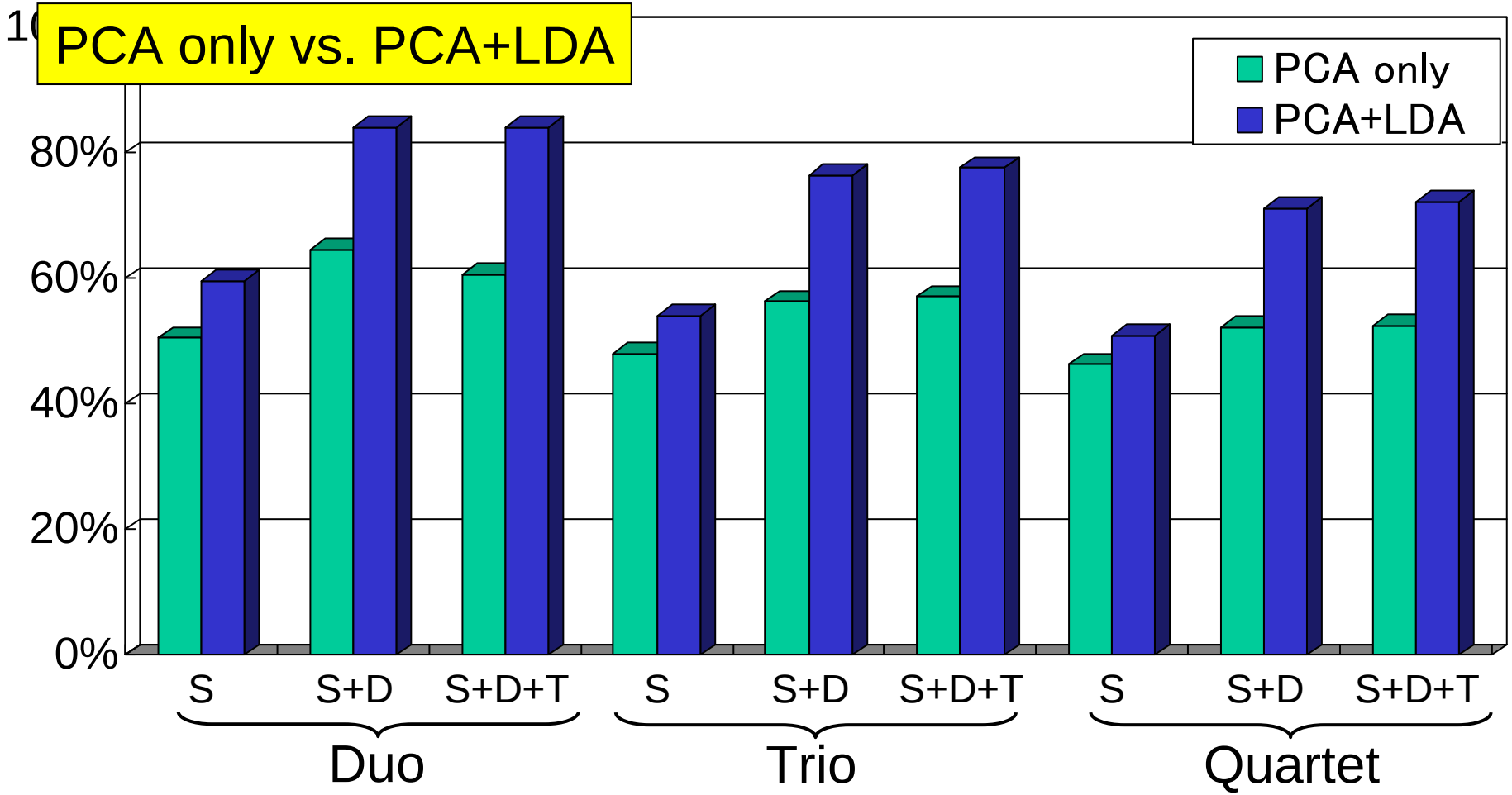


# 評価実験／結果

認識：四重奏，学習：単旋律＋二重奏



# 評価実験／結果




- PCA onlyでは，混合音学習データの効果小
  ⇒「混合音学習データ＋LDA」（DAMS法）が有効

# ま と め

- 目　的: 多重奏に対する音源同定
- 課　題: ①音の重なりによる特徴変動
  ②音高による音色変化
  ③音楽的文脈
- 解決策: ①混合音学習データ+LDA
  ②F0依存多次元正規分布
  ③事前確率への文脈情報の適用
- 実　験: 5楽器二～四重奏で性能向上確認
  （少ない混合音学習データでも）