

# スペシャルセッション 「音楽情報処理と機械学習」

オーガナイザ  
吉井 和佳 (産総研)

# 機械学習とは

- 人間が自然に行っている学習能力と同様の機能をコンピュータで実現させるための技術・手法
  - ある程度の大きさのサンプルデータ集合を解析し、**有用な規則・知識表現・判断基準などを抽出** (Wikipedia)
- 極めて多様な分野で顕著な成果
  - 画像処理・信号処理・自然言語処理・音声認識・ロボット・ゲーム  
文字認識・遺伝子解析・金融・通信・医療・気象・異常検知 …
- 主な3つの形態
  - **教師あり学習 (Supervised Learning)**
    - 入力に対して正しい出力ができるように訓練する
  - **教師なし学習 (Unsupervised Learning)**
    - データに内在している構造や規則を明らかにする
  - **強化学習 (Reinforcement Learning)**
    - 過去の行動結果を参考に報酬が最大になるような行動を選択する

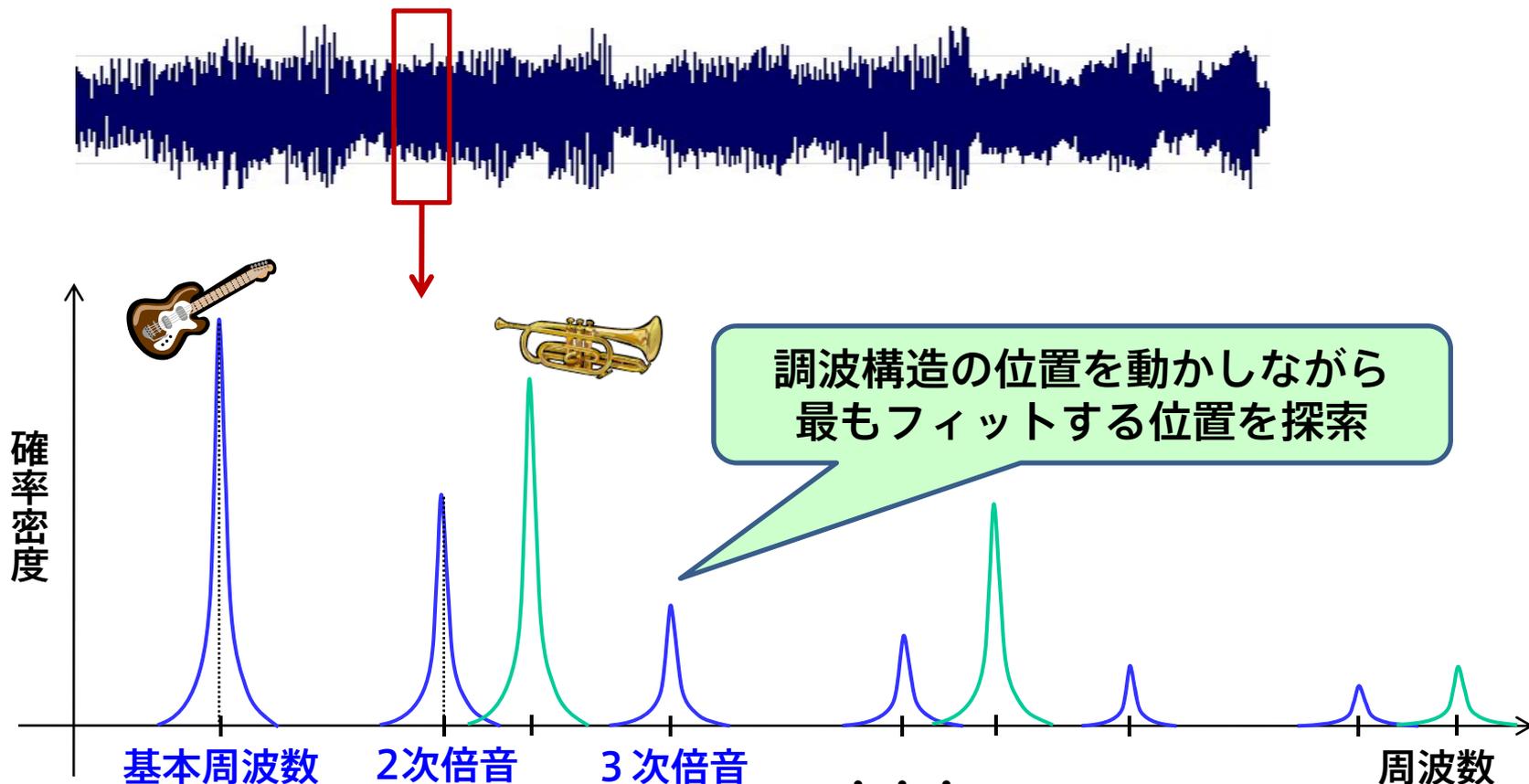
# 教師あり学習：楽器識別

- 混合音に含まれる楽器の種類を判別したい
  - 「これは～の楽器の音」という識別ができる
  - 「聴いたことがない音」という判断もできる



# 教師なし学習：音源分離

- 混合音を楽器音ごとに分離したい
  - 調波構造モデル (データの構造) と個数を与える
  - スペクトルにフィッティング



# 機械学習に関する研究発表

## • 歌声・歌詞関連

- 統計的言語モデルを用いた作詞補助システムのための歌詞特徴に関する検討
  - 阿部ちひろ, 伊藤彰則 (東北大学)
- VocaListenerによる学習データ生成を利用した多対多固有声変換に基づく歌声声質変換
  - 土井啓成, 戸田智基 (NAIST), 中野倫靖, 後藤 真孝 (産総研), 中村哲 (NAIST)
- 声質空間上での変換に基づく歌声らしさの転写に関する検討
  - 齋藤大輔, 石原達馬, 橘秀幸 (東京大学), 亀岡弘和 (東京大学/NTT), 嵯峨山茂樹 (東京大学)

## • 採譜・分離・同期 (明日午前&午後をお見逃しなく！)

- 多重音基本周波数解析のための無限複合自己回帰モデル
  - 吉井和佳, 後藤真孝 (産総研)
- ベイジアン非負値調波因子分解と多重基本周波数推定への応用
  - 阪上大地, 大塚琢馬, 糸山克寿, 奥乃博 (京都大学)
- 楽譜の文脈自由2次元木構造表現に基づく多重音スペクトログラム生成モデルによる音響信号からの自動採譜
  - 土屋政人, 落合和樹 (東京大学), 亀岡弘和 (東京大学/NTT), 嵯峨山茂樹 (東京大学)
- 押弦制約と運指制約を用いたタブ譜自動生成システム
  - 矢澤一樹, 阪上大地, 糸山克寿, 奥乃博 (京都大学)
- 補助関数法によるMIDI信号の音楽音響信号への詳細なフィッティング
  - 高宗典玄 (東京大学), 亀岡弘和 (東京大学/NTT CS研), 土屋政人, 嵯峨山茂樹 (東京大学)
- 隠れセミマルコフモデルと線形動的システムを組み合わせた音楽音響信号と楽譜の実時間アライメント手法
  - 山本龍一, 酒向慎司, 北村正 (名古屋工業大学)
- 多声MIDI演奏の楽譜追跡における演奏の不確定性のモデル化と自動伴奏への応用
  - 中村栄太 (東京大学), 山本龍一, 酒向慎司 (名古屋工業大学), 齋藤康之 (木更津高専), 嵯峨山茂樹 (東京大学)

# 音楽情報処理の新潮流

## • 確率モデルに基づく手法が著しい発展!

– 解析したい対象・現象の「生成過程」を考える

• 「ゆらぎ」を確率的にとらえる

– あるサイコロ → 実際にふってみて出る目

– 文字列「愛してる」 → 歌声「あ～いしてる～♪」

– 楽譜に指定されている音高 → 実際に演奏される音高

– ベイズモデル

• 「信念の強さ」も確率で表現できる!

– あるサイコロを何度かふってみた。サイコロが精密である確率は?

– 彼女は怒っている。その理由が遅刻したせいである確率は?

真実は既に確定済み (ゆらぎがない)



従来の頻度主義的な確率では  
説明できない

真実が分からないことに  
起因する不確実性



確率の公理を満たす!

# スペシャルセッション

- 世界的なトップランナーが集結！
  - 戸田 智基 氏 (奈良先端科学技術大学院大学)
    - 歌声の声質変換の確率モデル
  - 大石 康智 氏 (NTT)
    - 歌声の音高軌跡の確率モデル
  - 亀岡 弘和 氏 (東京大学/NTT)
    - 音響信号の生成過程の確率モデル
  - 中野 允裕 氏 (NTT)
    - 音楽に対するノンパラメトリックベイズモデル
- 異なる視点から3つのテーマを議論
  - 「機械学習を用いた研究の進展」
  - 「機械学習の効果と限界」
  - 「今後の展開および実用化に向けて」

世界最先端の研究が  
一挙に外観可能！

# 音楽情報処理と機械学習

- 「機械学習を用いた研究の進展」
  - 戸田 智基 氏
    - 歌声合成・声質変換・時系列データ生成
  - 大石 康智 氏
    - 話声と歌声の違い・歌声のF0の時系列変化
  - 亀岡 弘和 氏
    - 多重音解析・打楽器音分離・音声イントネーション解析
  - 中野 允裕 氏
    - ノンパラベイズモデルによる多重音解析・コード進行解析
- 「機械学習の効果と限界」
  - 解析対象をしっかりと観察することが何より重要！
- 「今後の展開および実用化に向けて」
  - 人間の知の解明のための数理的アプローチ
    - 大規模データの扱い・人間との協調・創作支援
    - 音楽情報処理から他分野に影響を与えられるか？

# 講演概要

- 戸田さん

- 人間の振る舞いを反映した音声合成

- 最尤系列＝平均的→不明瞭になる・もごもご発音
    - 静的＋動的＋系列内変動のモデル

- 声質変換

- 固有声変換
    - 声質コントローラ (性別・音高・年齢・かすれなど)

- 大石さん

- 歌の「うまさ」「らしさ」の数理モデル

- 文字で書き起こせない情報(パラ言語)の大切さ
    - 歌声と話声の違い
    - 音高軌跡のモデル

# 講演概要

- 亀岡さん

- 生成問題をたてて逆問題を解くアプローチ

- 尤度＋事前分布
    - 物理的な制約を考えて尤度関数を決める！

- 研究トピック

- 多重音解析
    - 調波・非調波音分離
    - 音声イントネーションモデル

- 中野さん

- ノンパラメトリックベイズによる構造学習

- 音楽の構造を自己組織的に学習
    - 「抽象度」という軸で音楽要素を考える
      - 楽器音・コード・セグメント etc.
    - 階層ノンパラメトリックベイズモデル