

「機械学習」特別セッション

吉井 和佳 (産業技術総合研究所)

亀岡 弘和 (東京大学/NTT)

機械学習とは何か？

- 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より)
- 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

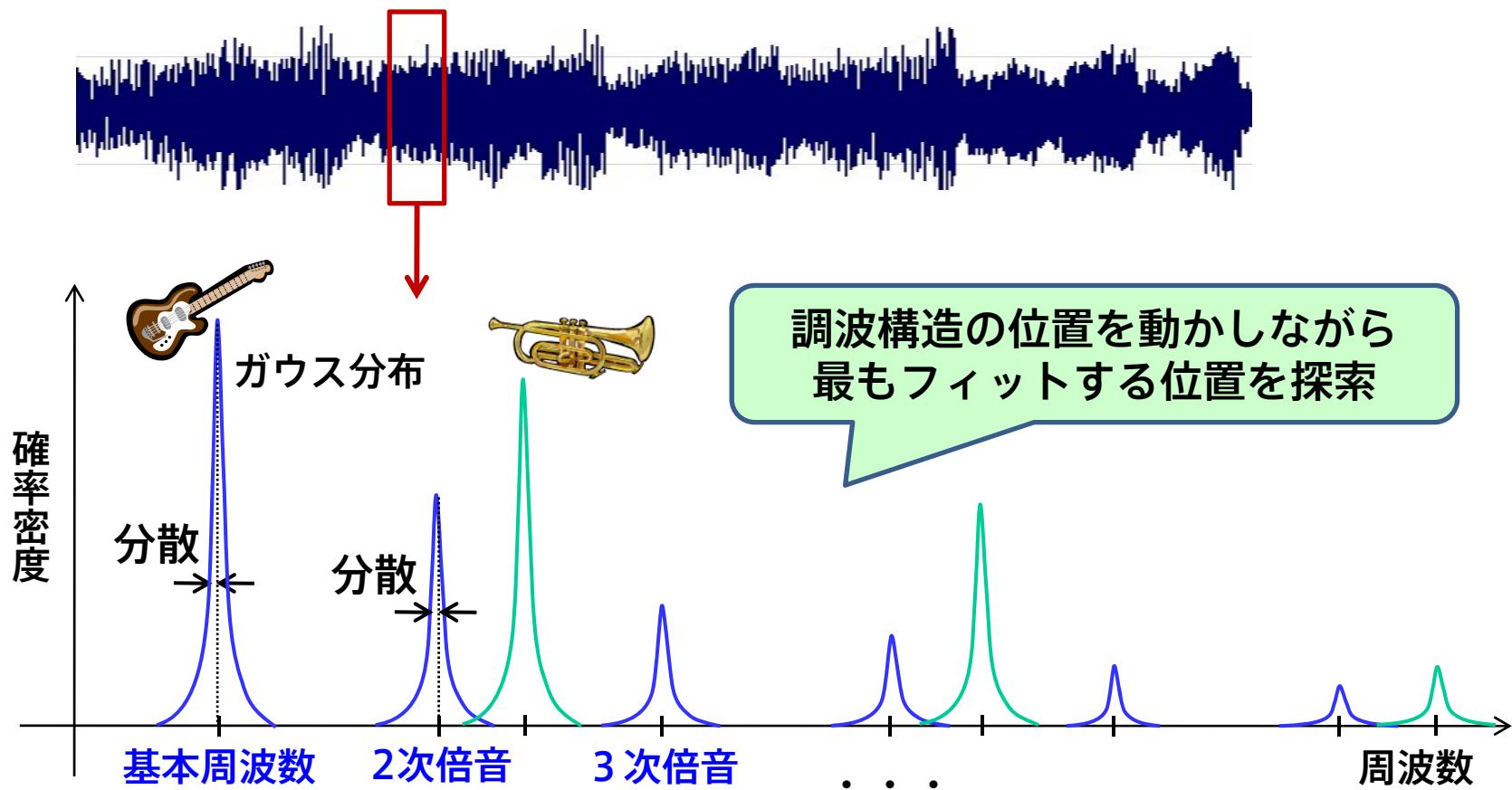
教師あり学習：楽器識別

- 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



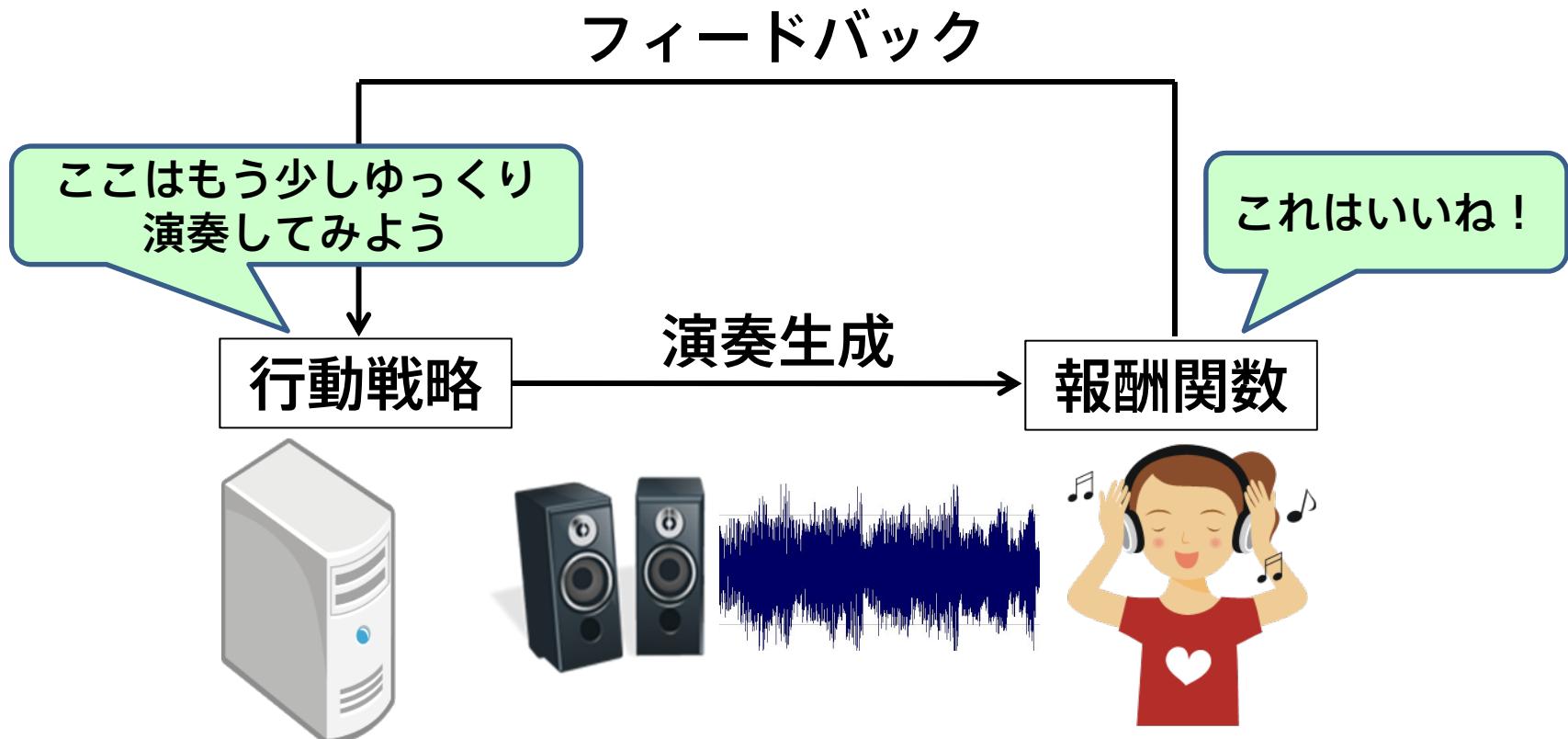
教師なし学習：音源分離

- 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



強化学習：演奏表情付け

- 音楽を表情豊かに演奏させたい
 - 計算機が生成した演奏が良かつたら褒める
 - cf. 良い演奏を直接教える → 教師あり学習
 - よりよい報酬が得られるように行動戦略を改善



音楽情報処理分野における機械学習

• 認識系の例

- 複数楽器混合モデルのパラメータ推定による楽器名同定
 - 糸山 克寿・後藤 真孝・駒谷 和範・尾形 哲也・奥乃 博(京大)
- 機能和声HMMによる音楽信号の和声推定
 - 上田 雄・小野 順貴・嵯峨山 茂樹(東大)
- 打楽器・低音旋律パターン特徴とクロマベクトルの線形識別によるムード・ジャンル分類
 - 角尾 衣未留・赤瀬 太一・小野 順貴・嵯峨山 茂樹(東大)
- 無限潜在的調波配分法による多重音解析
 - 吉井 和佳・後藤 真孝(産総研)
- 無限状態スペクトルモデルによる多重音解析 ★
 - 中野 允裕・ルルー ジョナトン・亀岡 弘和・小野 順貴・嵯峨山 茂樹(東大)
- MAHL: 演奏者間のインタラクション分析のためのスコアアライメント手法 ★
 - 前澤 陽・糸山 克寿・尾形 哲也・奥乃 博(京大)

• 生成系の例

- CRFに基づく伴奏の演奏表現の予測モデルと協調演奏システム ★
 - 山本 龍一・酒向 慎司・北村 正(名工大)
- 楽譜コンテキストのベイジアンマイニングに基づく自動伴奏付与システム ★
 - 多田 圭吾・岡村 亮吾・山西 良典・加藤 昇平(名工大)
- ガウス過程による名演奏の学習
 - 寺村 佳子・前田 新一(京大)
- Polyhymnia: 多重音演奏の統計モデルと演奏記号解釈による自動ピアノ演奏表情づけ
 - 金 泰憲・深山 覚・西本 卓也・嵯峨山 茂樹(東大)

音楽情報処理分野の2大トレンド

- ・ ノンパラメトリックベイズ
 - 機械学習法の性能を向上させる汎用的枠組み
 - ・ 教師あり学習：未知の入力に対する汎化能力を向上
 - ・ 教師なし学習：データの自己組織化（構造学習）
 - ・ 強化学習：行動戦略の柔軟性を向上
- ・ 非負値行列因子分解
 - スパースネスに基づく教師なし学習法
 - ・ データが少数の「パート」から構成されていると仮定
 - （）多くの実世界の事象で仮定が成立（応用範囲が広い）
 - （）データ行列が行列積で記述可能（とてもシンプル）

ICASSP2011での動向

音声信号処理のトップカンファレンス

- ・ ノンパラメトリックベイズ
 - 機械学習の第一人者M. I. Jordan氏が招かれてのプレナリートーク
 - Bayesian Nonparametrics for Speech and Signal Processing



・ 非負値行列因子分解

- 下記21セッション中に33件もの関連発表が！

1. Acoustic Source Separation I
2. Music Signal Processing I
3. Acoustic Source Separation II
4. Acoustic Source Separation and Noise Reduction
5. Image Coding
6. Restoration, Enhancements I
7. Non-negative Tensor Factorization and Blind Separation
8. Machine Learning for Speech and Audio Applications
9. Signal Separation
10. Detection and Estimation

11. Signal and System Modeling and Estimation
12. Adaptation for ASR
13. Robust ASR II
14. Speech Coding and Analysis
15. Speech Analysis III
16. Innovative Representations
17. Learning Theory and M
18. Adaptation for ASR
19. Audio/Visual L
20. Vocal Outbursts
21. Systems Biology
22. Music Signal Pro
23. Acoustic
24. Speech
25. Signal Processing
26. Machine Learning
27. Non-negative
28. Tensor Factorization
29. Blind Separation
30. Restoration
31. Enhancements
32. Image Coding
33. Non-negative

知っておいて
損はしない！

ML特別セッション: 第1部

- チュートリアル1
 - ノンパラメトリックベイズ
 - 吉井 和佳 (産総研)
 - ベイズ推論の枠組み・ディリクレ過程の解説
 - 典型的なノンパラメトリックベイズモデルの紹介
- 記号処理
 - 和音進行解析のための語彙フリー無限グラムモデル
 - 吉井 和佳, 後藤 真孝 (産総研)
 - 和音系列に対する可変長n-gramモデルの学習
 - あらゆる音符の組合せを和音として許容可能
 - 階層Pitman-Yor言語モデルを用いたメロディー生成手法の提案
 - 白井 亨, 谷口 忠大 (立命館大学)
 - メロディー系列に対する可変長n-gramモデルの学習
 - コードや歌詞を考慮したメロディ生成

ML特別セッション: 第2部

- チュートリアル2
 - 非負値行列因子分解
 - 亀岡 弘和 (東京大学/NTT)
 - 非負値行列因子分解の基本的な考え方・性質・求め方の解説
 - 音楽情報処理への応用の紹介
- 信号処理
 - 時間周波数分解能の異なるスペクトログラムの並列NMFによる多重音解析
 - 落合 和樹, 中野 允裕, 小野 順貴, 嵐峨山 茂樹 (東京大学)
 - さまざまな時間周波数分解能での並列NMFの結果の統合
 - スペクトログラムのベイジアンノンパラメトリックモデリングに基づく音楽信号の解析
 - 中野 允裕 (東大), ルルー・ジョナトン (NTT)
亀岡 弘和 (東大/NTT), 中村 友彦, 小野 順貴, 嵐峨山 茂樹 (東大)
 - ノンパラメトリックベイズ+NMFの拡張
 - 混合音は無限個の基底の足し合わせ・各基底の取りうる状態数は無限

Just relax!

Enjoy the ML day!