

画像符号化シンポジウム・映像メディア処理 シンポジウム講演予稿の書き方

How to Prepare a Camera-Ready Paper for Picture Coding Symposium of Japan and Image Media Processing Symposium

藤沢 貴典[†]
Takanori Fujisawa[†]

池原 雅章[†]
Masaaki Ikehara[†]

[†] 慶應義塾大学理工学部電子工学科

[†] EEE Dept., Keio University

Abstract: ここに日本語概要を 300 字程度で記述します。本稿では非負値行列因子分解 (NMF) をベースにした楽曲の音高指定の手法を提案する。NMF ベースの音高指定は短時間フーリエ変換によって得られた楽曲データの振幅スペクトログラムを、その構成要素である楽器音の基底スペクトルおよび、その出現の線形結合の形に分解することで実現される。この分離段階において、スペクトル方向には係数のスパース性が満たされ、時間軸方向には滑らかに係数が出現しているのが望ましい。著者は行列ノルム拘束を用いて、係数のスパース性と時間軸の連続性を同時に満足する手法を導入し、それに基づく NMF 更新式を提案する。この手法を用いたピアノ演奏データの音高推定を評価して、新しい NMF 手法が入力スペクトル形状にロバストであり、またスペクトルの時間変化に強いことを示す。

1 分量、原稿サイズ

- 1 件 2 ページ以内 (図表を含む)
- A4 サイズ
- 幅 180mm
- 高さ 252mm
- 段間の幅 8mm 程度 (二段組の場合)

としてください。

2 はじめに

楽曲の音声信号は、楽器音、声、およびその他の信号の混ざりあわせとして表現される。多声音高推定はこの音声信号を入力として、これら構成要素の基本周波数と、その時間変化を推定するものである。この技術は楽曲の自動採譜やコード推定、特徴抽出といったアプリケーションへの応用が可能である。従来の音高推定の手法としては *correlogram* を用いるもの [1]、スペクトルピーククラスタリングを用いるもの [2]、サポートベクターマシンを用いるもの [3] 等がある。

NMF (Non-negative Matrix Factorization: 非負値行列因子分解) はスペクトル分解の手法としてよく用いられているものの一つである。これは音声信号のスペクトログラムを辞書行列と、その時間変化行列の積に分解する手法である。NMF はその解が非負の値となるために、楽器ごとのパワースペクトルの加法的に近似的に成立する音声信号処理との親和性が高い。また乗算更新則に基づく単純な計算で最適な解が得られる利点がある。

本稿では NMF をベースにした多声音高推定のアルゴリ

ズムを提案する。本手法は NMF によるスペクトル分離に対し実音楽に即した 2 つの特徴を組み込んでいる。1 つ目は、出現するアクティベーションのピッチ方向のスパース性の確保であり、もうひとつはアクティベーションの時間方向の連続性である。両特徴を満たすために、本研究では行列ノルムを用いた結合スパース性拘束をアクティベーションに対して導入している。この拘束条件の導入により、入力スペクトルが辞書行列中の基底に合わない部分での誤推定を減らす効果が期待される。それに加え時間方向の連続性拘束は、入力スペクトル要素の時間的な変化に対して追従できる効果が期待される。

2.1 NMF を用いた楽曲の音高推定

50 ~ 100 回の更新式適用によって得られたアクティベーション行列 A を元に楽曲のピアノロール表現を作成する。この際に、辞書 S の各インデックスが示す音高を推定する必要がある。 S の初期値に楽器音データベースから取得した基底スペクトルを用いた場合は、該当する基底スペクトルの音高をそのまま用いることが出来る。

2.1.1 スパース性と時間軸連続性拘束を付した NMF ベースの音高推定

実音楽のデータに対し (??) および (??) 式の更新式で音高推定を行った場合、アクティベーション行列 A において、実音楽の 2 つの特徴が無視される問題がある。問題の 1 つめが、出現する基底インデックスのスパース性であり、もう 1 つが、隣接するフレームのアクティベーション同士の連続性である。

表 1 符号化器の設定

設定ア	aa	aaa
設定イ	bb	bbb
設定ウ	cc	ccc

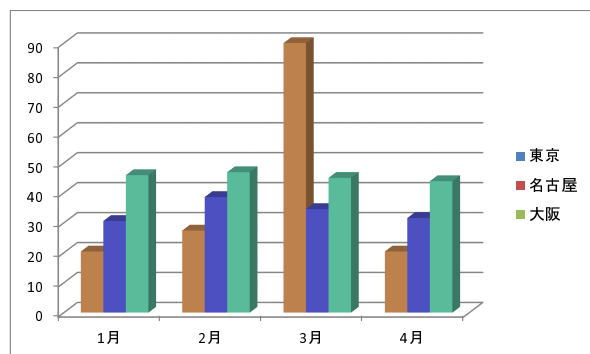


図 1 符号化実験結果

のようにしてください。

参考文献

- [1] S. Nagashima and T. Aoki, “Improved performance for subpixel image matching based on phase-only correlation,” 2004.

3 配置

- 表題、著者名、勤務先は、本サンプルに従ってそれぞれ記入して下さい。
- アブストラクト (和文の場合 300 字程度、英文の場合 100 語程度) をお入れ下さい。
- 本文は一段または左右二段に書いても差し支えありません。
- 用紙最終頁右下に、発表者連絡先をお入れ下さい (本サンプルのマクロ `\simplefootnotetext` を使われると便利です)。

4 文字サイズ

文字サイズは、本サンプルを目安にしてください。

表題 16 ポイント (`\LARGE`)

和文著者名 14 ポイント (`\Large`)

英文著者名 12 ポイント (`\large`)

所属および本文 10 ポイント (`\normalsize`)

程度です。

5 図表

予稿集は白黒・網がけ印刷になります。見にくくならないよう、作成の際にはご注意ください。

図表の出力位置を指定するオプションは、`[h]` は使わず `[t]`、`[b]` などを指定してページの天か地に置くことを基本にします。

図表内は英文でも和文でも結構です。参照は図 1、表 1

慶應義塾大学 電子工学科 池原研究室
〒223-8522 神奈川県横浜市港北区日吉 4-5-33 Kohoku-ku, Yokohama, Kanagawa, 223-8522, Japan
E-mail: ikehara@tkhm.elec.keio.ac.jp