

楽曲構造を考慮した音楽音響信号からの自動ピアノアレンジ

高森啓史^{†1} 中塚貴之^{†1} 深山覚^{†2} 後藤真孝^{†2} 森島繁生^{†3}

概要: 本研究はポピュラー楽曲を対象とした音楽音響信号からの自動ピアノアレンジを、楽曲構造を考慮しながら行う。従来手法では小節間の連続性を考慮することが困難であったため、出力譜面にピアノ譜の持つ構造的な特徴を反映させることが出来なかった。我々は、既存ピアノ楽譜の伴奏部分及び音響特徴量の自己類似度行列を生成し、ピアノ譜の持つ繰り返し構造を段階的に反映させ、1曲に対し伴奏の異なる複数のピアノ譜を出力した。また、どの音響特徴量がピアノ譜の持つ類似度行列を生成するために有効であるかを確認するため重回帰分析を行った。

キーワード: ピアノアレンジ, 重回帰分析, 楽曲構造, 信号処理, 音響特徴量

1. はじめに

実際に音楽を演奏することは、ただ音楽を聴くだけでは味わえない楽しさを楽しむことが出来る。あらゆる楽器の中でもピアノは広い音域を持ち、多くの音を同時に奏でることの出来る表現の幅が広い楽器である。そのため、様々な音が混在する楽曲でもピアノで表現することが可能である。特に、ポピュラー音楽はメロディ、和音、リズムといった要素が明確であり、ピアノはこれらの役割を同時に表現することが出来る。そこで、本研究はポピュラー音楽に焦点をあて、自動ピアノアレンジを行う。

ピアノアレンジ楽譜を作成する際は、演奏可能でありつつ原曲の雰囲気や反映させることが求められるため、多大な労力と専門知識が必要とされる。実際には、楽器店や書店などで編曲家により作成されたピアノアレンジ楽譜を手に入るが、その楽曲数は限られている。そのため、所望の楽曲のピアノアレンジ楽譜が入手できない場合には、自らピアノアレンジ楽譜の作成を行わなければならない。この問題に対処すべく、我々はこれまでポピュラー音楽の音楽音響信号から自動的にピアノアレンジ楽譜を出力する手法を提案してきた [1]。従来手法では、音楽音響信号より抽出したリズムと類似したピアノアレンジ楽譜の出力を行った。しかし、小節間の連続性やピアノ楽譜の持つ構造的な特徴が考慮されていなかった。そこで、本稿では、原曲の印象を維持しつつ、ピアノ楽譜の持つ構造を考慮した自動ピアノアレンジを行う手法について述べる。本手法の概要を図1に示す。本手法は3つの段階から構成される。(a)音楽音響信号から得られる自己類似度行列とピアノ楽譜の左手声部から得られる自己類似度行列について重回帰分析を行う。(b)テスト楽曲に対する自己類似度行列の推定を(a)で得られた偏回帰係数を基に行う。(c)音響データより得られる音楽要素より左右別々にピアノ楽譜を生成する。右手パートはメロディにコード構成音を付加することで生成

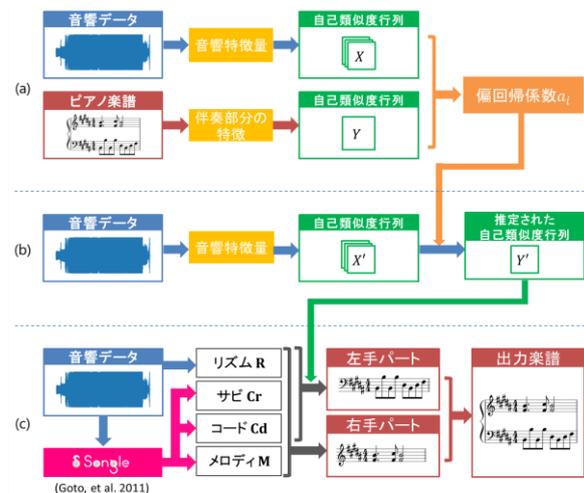


図1 本手法の概要図。(a)楽曲構造の解析。(b)音楽音響信号からの構造推定。(c)楽譜生成。

し、左手パートは事前に構築した伴奏データベースよりリズムの類似する伴奏を選択した後、(b)で得られた楽曲構造に近づくよう修正を行う。本研究の貢献は以下の2点が挙げられる。

- 原曲のリズム及びピアノ楽譜の持つ構造を反映させたピアノアレンジ楽譜を音楽音響信号から生成した。
- 自己類似度行列として表されたピアノ楽譜の持つ構造を、7つの音楽音響信号から推定した。

2. 関連研究

これまで、複数のパートから構成される楽曲のピアノアレンジに関して研究されてきた。藤田ら [2]は原曲のアンサンブル楽譜からメロディとベースの推定を行い、コンデンススコアの作成を行った。さらに、ピアノで演奏できるよう同時打鍵における最大度数や演奏不可能な連打について制限をかけ、難易度別ピアノアレンジ楽譜を複数出力した。Chiuら [3]は楽曲の持つ5つの役割(lead, foundation, rhythm, pad, fill)を考慮し、ピアノアレンジを行った。これらの5つ役割は Owsinski [4]によって提唱されたものである。Chiuらは原曲スコアに含まれる各パートの重要度を上記の5つの役割を考慮して重みづけし、その重要度が

^{†1} 早稲田大学
 Waseda University.
^{†2} 産業技術総合研究所
 National Institute of Advanced Industrial Science and Technology (AIST)
^{†3} 早稲田大学理工学術院総合研究所
 Waseda Research Institute of Science and Technology

最大になるようパートの選択を行うことでピアノアレンジ楽譜を生成した。その際、原曲の印象維持に留意し、最大同時打鍵数や片手での度数制限といった演奏可能性も考慮した。中村ら [5]はアンサンブル楽譜からピアノアレンジを行う際に運指モデルを導入することで音の削減を行った。音の削減具合は難易度のパラメタを左右別々に調整することにより、複数段階の難易度を持つピアノアレンジ楽譜を出力した。

また、音楽音響信号から編曲を行う研究として、Percivalら [6]の Song2Quartet がある。Percivalらは、音楽音響信号から弦楽四重奏によるカバーソングの生成を可能とした。その目的においては、全てのパートの音高推定を行う必要はなく、原曲のメロディ、リズム、コード、音数といった音楽要素が抽出出来ればよい。本手法の一部は、それと類似した枠組みでありつつ、特にピアノアレンジを生成するための制約の考察に重点を置いた。

3. ポピュラー音楽におけるピアノアレンジ

メロディ、和音、リズム、音色、テクスチャといった要素は、音楽における代表的な要素として知られている [7]。これらの要素は音楽を表現する上で重要な役割を果たす。そのため、原曲の雰囲気を持する際にこれらの要素に注目することは重要であると言える。ただし、異なる楽器の音色をピアノで表現することは困難である。そのため、本研究ではメロディ、和音、リズム、テクスチャに注目し、原曲の雰囲気を保ったピアノアレンジを目的とする。各要素について以下4つの条件 [1]を満たすピアノアレンジ譜面を生成した。

- メロディが常に最高音であること
- 出力楽譜のコードが原曲と一致していること
- 出力楽譜が原曲のリズムを反映していること
- サビで原曲の盛り上がり表現していること

更に、本研究では上記の4点に加え、ポピュラー音楽における楽曲構造を考慮する。ポピュラー音楽はイントロ、Aメロ、Bメロ、サビ、アウトロ等といった区間が存在し、それらが繰り返し展開していくことにより楽曲が構成される [8]。本稿ではAメロやサビと言った区間を楽曲区間と呼ぶ。一般に、同一楽曲区間では似たようなリズムやメロディの展開がされ、新たな楽曲区間(特にサビ)に移る際はそれらの曲調が変化する。従って、ポピュラー音楽の持つ楽曲構造を考慮し、ピアノアレンジを行うことは重要であると考えられる。そこで、我々は新たに次の条件を加えピアノアレンジ楽譜を生成した。

- 出力楽譜の左手部分が同一楽曲区間で似たような伴奏であること

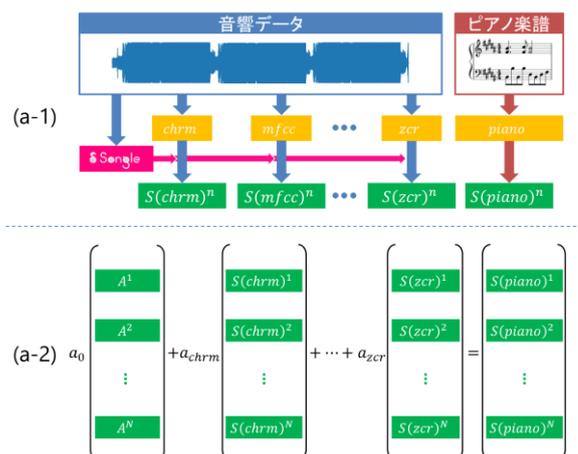


図 2 楽曲構造の解析の詳細。(a-1)音響データとピアノ楽譜から特徴抽出を行う。(a-2)全てのデータセットに対し行われる重回帰分析の図示。S()は音響特徴量及びピアノ楽譜の特徴の自己類似度行列を表し、aは偏回帰係数を表す。

本研究では以上の5つの点に留意し、ピアノアレンジを行った。

4. 楽曲構造の解析

本章では図 1(a)に示した楽曲構造の解析に関する説明を行う。その詳細を図 2に示す。本研究では27曲のポピュラー音楽について、音響データとピアノ楽譜の組をデータセットとして用意した。音響データはインターネット上で公開されているものを使用した。また、ピアノ楽譜は編曲家によって作成された楽譜 [9-11]を使用し、それらを正解楽譜とした。

4.1 特徴抽出

事前に用意したデータセットのうち、音響データからはクロマグラム、メル周波数ケプストラム係数(MFCCs)、オンセット、RMS値、スペクトル重心、スペクトルフラットネス、ゼロ交差率(ZCR)を音響特徴量として抽出した。特徴抽出の際、サンプリング周波数は44.1kHzに設定し、分析窓の窓幅は1024サンプルとし、256サンプルをオーバーラップさせて分析した。

クロマグラムは和音に関する特徴量であり、特定の周波数帯域のエネルギーを12のピッチクラス{C, C#, D, D#, E, F, F#, G, G#, A, A#, B}に投影することで得られる。クロマグラムを取得する際は、2ⁿ倍の周波数は同じピッチクラスとみなされる。メル周波数ケプストラム係数は人間の聴覚構造を考慮したケプストラム分析の結果を表すベクトルである。本手法ではメルフィルタバンクを20とし、低次12次元分を用いる。オンセットはスペクトルフラックスが大きく変化するフレームを取得することで得られる。我々はオンセット検出の手法としてBöckら [12]の手法を用いた。本手法で用いるオンセットはバイナリとして扱い、

スペクトルフラックスのピークが存在するフレームを 1, それ以外を 0 とし表した. RMS 値は音響データの音圧を測る指標, スペクトル重心は音の明るさを表す指標, スペクトルフラットネスはスペクトルの周波数分布を表す指標である. ゼロ交差率は音のノイズの程度を表す指標であり, 波形の符号が逆転する比率を表す. 本手法ではこれまで述べた 7 つの特徴量を小節ごとに求める. 音響データにおける各小節の開始時刻は能動的音楽鑑賞サービス Songle [13] より得る. これにより, 楽曲の小節位置と音響データの時間を対応付けることが出来る. m 小節目における各音響特徴量について, クロマグラムは $chr_m \in \mathbb{R}^{12 \times 16}$, メル周波数ケプストラム係数は $mfcc_m \in \mathbb{R}^{12 \times 16}$, オンセットは $onset_m \in [0,1]^{1 \times 16}$, RMS 値は $rms_m \in \mathbb{R}^{1 \times 16}$, スペクトル重心は $cent_m \in \mathbb{R}^{1 \times 16}$, スペクトルフラットネスは $flat_m \in \mathbb{R}^{1 \times 16}$, ゼロ交差率は $zcr_m \in \mathbb{R}^{1 \times 16}$ と表記する. 各特徴量の行列の行方向は時間を意味し, j 列目の要素は 1 小節を 16 ビートに分割した際の j 番目のビートを表す. オンセットを除く各音響特徴量は, 注目ビートと次のビートの間に含まれる値を平均し, 16 ビート毎に格納される. j 列目の $onset_m$ には, 注目ビートと次のビートの間にスペクトルフラックスのピークがあれば 1, なければ 0 を格納する.

事前に用意したデータセットのうち, ピアノ楽譜からは左手部分における音符の位置と数を用いて特徴抽出を行う. このピアノ楽譜の特徴も小節毎に表し, m 小節目を $piano_m \in \mathbb{Z}^{1 \times 16}$ と表記する. 行方向は時間を意味し, j 列目の要素は 1 小節を 16 ビートに分割した際の j 番目のビートに含まれる音符の数が格納される.

4.2 自己類似度行列の重回帰分析

4.1 節で得られた音響特徴量と, ピアノ楽譜の特徴の自己類似度行列を算出する. 本手法で算出する自己類似度行列は, 1 曲に含まれるある 2 つの小節同士の類似度を全小節にわたって表したものである. 全 M 小節の楽曲における音響特徴量及びピアノ楽譜を $\mathbf{f} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_M\}$ と表記した際, 自己類似度行列 $S(\mathbf{f}) \in \mathbb{R}^{M \times M}$ は以下のように定義される.

$$S(\mathbf{f}) = [s_{ij}] = \begin{pmatrix} 1 & s_{12} & \dots & s_{1M} \\ s_{21} & 1 & \dots & s_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ s_{M1} & s_{M2} & \dots & 1 \end{pmatrix} \quad (1)$$

s_{ij} は i 小節目の特徴 \mathbf{f}_i と j 小節目の特徴 \mathbf{f}_j の類似度を示す. 類似度 s_{ij} は距離の逆数をとることで計算される.

$$s_{ij} = \frac{1}{1 + d_{ij}} \quad (2)$$

$$d_{ij} = \|\mathbf{f}_i - \mathbf{f}_j\| \quad (3)$$

d_{ij} は \mathbf{f}_i と \mathbf{f}_j のフロベニウスノルムとした. s_{ij} のとる値の範囲は $0 < s_{ij} \leq 1$ であり, s_{ij} が 1 であるとき \mathbf{f}_i と \mathbf{f}_j は完全に一致することを意味する.

式(1)で得られた各音響特徴量及びピアノ楽譜の特徴の

自己類似度行列を重回帰分析に用いる. 重回帰モデルは各音響特徴量からピアノ楽譜の特徴を推定する形をとり, 説明変数と目的変数はそれぞれ以下のように表す.

$$X_f = \{S(\mathbf{f})^1, S(\mathbf{f})^2, \dots, S(\mathbf{f})^N\}$$

$$Y_{piano} = \{S(\mathbf{piano})^1, S(\mathbf{piano})^2, \dots, S(\mathbf{piano})^N\}$$

N はポピュラー楽曲のデータセット数を表し, $S(\cdot)^n$ はデータセットの n 曲目における音響特徴量及びピアノ楽譜の特徴の自己類似度行列を表す. 重回帰分析は以下の式に従って行われる.

$$a_0 \mathbf{A} + \sum_Y a_{f_Y} X_{f_Y} = Y_{piano} \quad (4)$$

a_0 は切片を表し, \mathbf{A} はすべての要素が 1 の行列の集合を表す. また, 添え字 f_Y は各音響特徴量を表す. 式(4)を視覚的に表したものを図 2 の(a-2)に示す.

5. 構造推定

本章では図 1(b)に示したピアノ楽譜の持つ構造を音楽音響信号より推定する手法について説明する. 4 章で求めた重回帰係数を基に, ピアノ楽譜の自己類似度行列の推定を行う. 推定した自己類似度行列は, 変化点検出を用いて楽曲の分割を行う.

5.1 自己類似度行列の推定

4.1 節と同様にして, 音響データより音響特徴量の抽出を行い, 自己類似度行列を求める. その後, 4.2 節で求めた重回帰係数を基に, ピアノアレンジ楽譜のもつ自己類似度行列を推定する.

5.2 楽曲の分割

Jonathan ら [14] の変化点検出を用いて 5.1 節で推定された自己類似度行列を分割し, 楽曲区間の決定を行う. 変化点検出は, 自己類似度行列の変化度合いを表すノベルティスコアのピーク検出をすることで行われる. ノベルティスコアは, 自己類似度行列の対角線に沿って下記のチェッカーボード・カーネル \mathbf{C} をかけることで得られる.

$$\mathbf{C} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad (5)$$

本稿ではチェッカーボード・カーネルのサイズが (2×2) , (4×4) , (6×6) , (8×8) , (10×10) の 5 種類の行列を用いる. 本手法ではこれら 5 種のチェッカーボード・カーネルを用いて得られたノベルティスコアの平均をとり, 値の範囲が $[0,1]$ となるよう規格化したものを用いる. ノベルティスコアのピーク位置は, 1 階微分の値が正から負に変わり, 且つ 2 階微分の値が閾値 th よりも低い箇所をピーク位置と定める. 閾値 th は 0.00, -0.05, -0.07, -0.10, -0.15 の 5 段階に設定する. 以上の操作で得られたピーク位置を楽曲の境界と定め, その境界で挟まれた区間を同一楽曲区間とみなす.

6. ピアノアレンジ楽譜生成

本章では図 1(c)に示した音楽音響信号からのピアノアレンジ楽譜生成について説明する。まず事前に、伴奏データベースの構築を行う。次に、音響データより音楽要素の抽出を行う。サビ位置、コード、メロディは Songle [13]より取得し、リズムはスペクトルフラックスのオンセット検出をすることで得る。最後に、抽出された音楽要素に基づき左右のピアノ楽譜を出力する。

6.1 伴奏データベース

伴奏データベース DB は既存ピアノ楽譜の左手部分を 1 小節ごとに抜き出すことによって構築される [1]。抜き出された左手部分は 88×16 のサイズを持つ行列によって表され、列方向は音高方向、行方向は時間方向を表し、これを伴奏行列と呼ぶ。88 はピアノの鍵盤数を表し、16 は 16 分音符の時間単位まで考慮することを示す。また、伴奏行列として既存ピアノ楽譜の伴奏部分を格納する際は、根音が C となるよう平行移動させてから格納する。音高が異なる伴奏であったとしても、相対的な音の推移が全く同じであれば、それらを同一な伴奏とみなすことが出来る。伴奏行列の要素には、音符のある場所には音価が格納され、無い場所には 0 が格納される。これにより、既存ピアノ楽譜の伴奏部分の相対的な音高推移とリズムを伴奏行列で表すことが出来る。本稿では、 n 番目の伴奏行列を $DB_n \in \mathbb{Z}^{88 \times 16}$ と表す。また、それぞれの伴奏行列にリズムを表す $DBR_n \in [0,1]^{1 \times 16}$ を定める。伴奏行列 DB_n の j 列目に 1 以上の数字、つまり 1 小節を 16 ビートに分割したうちの j ビート目に音符があれば、 DBR_n の j 列目に 1 が格納され、それ以外は 0 が格納される。

6.2 音楽要素の抽出

メロディ M 、コード Cd 、サビ位置 Cr 、リズム R を音楽音響信号より抽出する。メロディ M 、コード Cd 、サビ位置 Cr は Songle [13]より取得する。リズム R は Böck [12]らの手法を用いたスペクトルフラックスのオンセット検出により求める。Songle [13]はインターネット上にアップロードされている約 120 万曲の楽曲について、メロディ、コード、ビート、楽曲構造の解析を行った Web サービスである。我々は以下に示される楽曲解析の結果を Songle Widget [15]より取得した。

- ビート : インデックス, 時刻, ビート位置
- コード : インデックス, 時刻, 長さ, コードネーム
- メロディ : インデックス, 時刻, 長さ, ノートナンバー
- サビ : インデックス, 時刻, 長さ

指数は楽曲の始めから何番目のビート、コード、メロディの音、サビなのかを表す。時刻はそれぞれのイベントが起こる開始時刻を表し、ビート位置は各小節における何拍目の位置であるかを表す。長さはそのイベントが続く長さを表す。コードネームは、根音とコードタイプによって表さ

れる。ノートナンバーは MIDI フォーマットでの音高を表す数値である。以上の情報より、各小節におけるメロディ、コードを取得でき、サビとなる小節番号を取得することが出来る。ここで、 m 小節目のメロディ、コード、サビ位置、リズムをそれぞれ M_m , Cd_m , Cr_m , R_m と表す。 M_m は 88×16 のサイズの行列で、88 はピアノの鍵盤の数に基づいて決め、16 は 16 分音符の時間単位まで表現できるように設定した。列方向と行方向はメロディの音高方向と時間方向を表し、対応する要素に音価が格納される。コードはコード構成音のノートナンバーの集合として Cd_m と表される。 Cr_m は m 小節目がサビであるならば 1、そうでなければ 0 が格納される。リズム R は 4.1 節で述べたオンセット $onset_m$ の抽出方法と同様にして得られ、1 小節を 16 分割した際の j 番目にスペクトルフラックスのピークがあれば $R_m \in [0,1]^{1 \times 16}$ の j 列目の要素に 1 が格納され、それ以外は 0 が格納される。

6.3 ピアノ楽譜の右手パート生成

右手パート RH はメロディ M を以下の式に従い割り当てることにより生成する。

$$RH_m = \begin{cases} Add(M_m, R_m, Cd_m) & Cr_m = 1 \\ M_m & otherwise \end{cases} \quad (6)$$

RH_m は m 小節目の RH を表す 88×16 の行列である。 m 小節目がサビである場合、 Add は R_m の要素が 1 であるビート位置にメロディの音高とともに和音をなすようにコード構成音を付加する。コード構成音は Cd_m より得られ、そのうちメロディよりも音高が低く、かつメロディと 3 度以上離れた最も近い音が付加される。

6.4 ピアノ楽譜の左手パート生成

左手パート LH は 6.1 節で構築した伴奏データベース DB より選択することで生成する。まず初めに、音響データより得られたリズム R と似たリズムを持つ伴奏を伴奏データベースより選択する。この過程で得られた全小節にわたる伴奏の並びを LH' と表す。また、 LH' に含まれる m 小節目の伴奏を LH'_m と表す。この伴奏列 LH' はリズムのみ考慮して選ばれた伴奏列であることを意味する。次に、5 章で述べた楽曲構造を伴奏に反映させる。楽曲区間を意識した伴奏選択を行うため、同一楽曲区間内で出現する伴奏の種類に制限をかけ、類似した伴奏が選ばれるようにする。伴奏の種類に制限をかけた後の伴奏列を LH と表す。1 つの楽曲区間に含まれる伴奏の種類に制限をかけるパラメータとして λ を導入し、一つの楽曲区間に含まれる伴奏の種類数の上限を λ によって設ける ($1 \leq \lambda$)。上記で説明した過程を以下の式に示す。

$$LH = FuncS(LH', \lambda, th) \quad (7)$$

$$LH'_m = \operatorname{argmin} CostR(DB, R_m) \quad (8)$$

$$CostR(DB, R_m) = \sum_n \|DBR_n - R_m\| \quad (9)$$

$FuncS(\cdot)$ はリズムのみ考慮されて選択された伴奏列 LH' から各楽曲区間内で出現する伴奏の種類を減らす関数である。具体的には、一つの楽曲区間内において出現頻度の低い伴奏を出現頻度の高い伴奏に差し替えることで伴奏の種類を減らす。この作業を各楽曲区間に含まれる伴奏の種類が λ 種類以下になるまで行う。 $\lambda = \infty$ の場合は伴奏の種類に制限がないことを示し、 LH' がそのまま LH となる。 $\lambda = 1$ の場合には各楽曲区間において1種類のみ伴奏が選ばれることになる。閾値 th は5章で述べた楽曲区間の分割に用いられる。 $CostR(\cdot)$ は音響データより抽出されたリズムと似たようなリズムを持つ伴奏を選択するために導入した。上記の過程で選ばれた伴奏 LH は、すべて根音がCとなるよう平行移動されたものであるため、 Cd よりコードを取得し、目的のコードとなるよう移調する。また、移調した際のコード構成音でない音については、最も近いコード構成音に移動させる。

7. 結果と評価

4章で述べた重回帰分析の結果を表1に示す。 a_0 は切片、 a_f は各音響特徴量の偏重回帰係数を表す。 t 値は重回帰分析において各説明変数が目的変数に対してもたらす影響力の大きさを示す値として知られ、絶対値が大きいほど推定に大きな影響があることを示す。 t 値は偏重回帰係数を標準誤差で割ることで得られる。 R_{adj}^2 は自由度調節済決定係数であり、以下のように定義される。

$$R_{adj}^2 \equiv 1 - \frac{\sum_i (y_i - y'_i)^2 / (N - p - 1)}{\sum_i (y_i - \bar{y})^2 / (N - 1)} \quad (10)$$

y は実際のデータ、 \bar{y} は y の平均、 y' は予測データ、 N はサンプルデータの数、 p は説明変数の数を表す。 R_{adj}^2 は回帰分析によって求められた目的変数の予測値が実際のデータとどのくらい一致しているかを表す指標である。予測データが実際のデータと完全に一致しているとき $R_{adj}^2 = 1$ となり、誤差が大きくなると R_{adj}^2 は低くなる。また、回帰分析における各説明変数間の有意差を検証するためにp値を算出した。p値が有意水準よりも低くなる時に有意であることを示せ、有意水準は一般に設定される0.05とした。本研究で行った重回帰分析において算出したp値はすべて 10^{-18} 以下のオーダーであった。

本研究では、データセットに含まれるポピュラー楽曲27曲についてピアノアレンジ楽譜の出力を行った。本手法の有効性を確かめるために1つ抜き交差検証を行った。初めに、データセットの中から1曲テストデータとして選ぶ。続いて、残りの26曲について重回帰分析を行い、その結果得られた偏重回帰係数を用いてテストデータのピアノ楽譜を生成する。また、ピアノ楽譜を生成する際は、楽曲分割の閾値 th と楽曲区間に含まれる伴奏の種類数 λ を変え、1曲につき th は5通り、 λ は7通りの計35通りの楽譜を生成する。最後に、生成されたピアノ楽譜の左手部分と正解データの

表1 重回帰分析の結果

	a_0	a_{chr}	a_{mfcc}	a_{onset}
係数	-0.0373	0.4439	-0.0329	0.1403
t 値	-8.974	74.50	-8.863	74.88

a_{cent}	a_{flat}	a_{rms}	a_{zcr}	R_{adj}^2
0.2224	0.35153	0.0494	-0.1175	0.1590
111.3	87.47	27.87	-34.95	

表2 一つ抜き交差検証の結果 ($\times 10^{-3}$)

閾値 th	伴奏の種類数 λ						
	1	2	3	4	5	6	∞
0.00	3.550	3.630	3.674	3.685	3.714	3.730	3.760
-0.05	3.535	3.612	3.684	3.708	3.738	3.742	3.760
-0.07	3.277	3.357	3.452	3.522	3.591	3.647	3.760
-0.10	3.257	3.451	3.413	3.430	3.460	3.506	3.760
-0.15	3.311	3.503	3.538	3.535	3.564	3.560	3.760

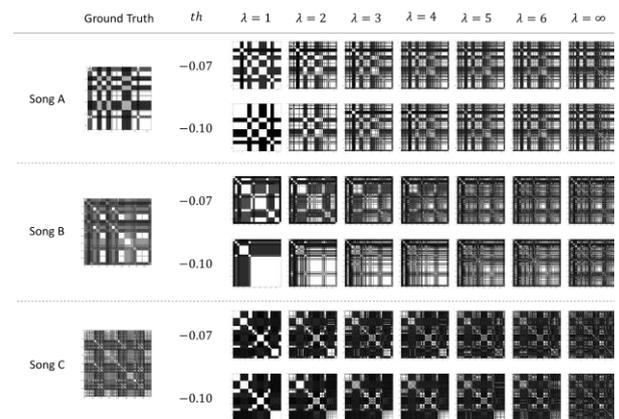


図4 推定ピアノ楽譜と正解楽譜の自己類似度行列

ピアノ楽譜の左手部分について自己類似度行列を求め、2つの行列のフロベニウスノルムを算出することにより誤差を求める。誤差は行列の要素数で割ることで正規化される。この操作を全27曲について行い、求めたそれらの誤差の平均を表2に示す。また、データセット中の3曲について、 $th = -0.07, -0.10$ において λ を変化させた際の自己類似度行列の推移と正解楽譜の自己類似度行列を図3に示す。本手法の結果として、RWC研究用音楽データベース [16]に収録されている楽曲 (RWC-MDB-P-2001 No.7) についてピアノアレンジを行い、その生成結果を <https://youtu.be/lep84y9dpGE> にアップロードした。

8. 考察

表1よりスペクトル重心、スペクトルフラットネス、オンセット、クロマグラムのt値が比較的高い値であることが読み取れる。この結果から、これらの音響特徴量はピアノアレンジ楽譜における構造推定に有効であると言える。

スペクトル重心, スペクトルフラットネス, クロマグラムは音の高さやその組み合わせに関する特徴量である. 従って, ピアノ楽譜の持つ構造推定において, 音域やメロディライン, コードに注目することは重要であると言える. また, オンセットはリズムに関する特徴量であるため, ビートを刻むパーカッション等の音に注目することもまた重要であると言える.

一方, メル周波数ケプストラム係数, RMS 値, ゼロ交差率の t 値は比較的低い値となった. この結果から, 音色やテクスチャに関する特徴は, 前述の音の高さやリズムに関する特徴ほど構造推定において大きな影響を持たないと言える. これらの音色やテクスチャに関する特徴量から得られる繰り返し構造は, ピアノ楽譜の持つ繰り返し構造と対応したものではないと考えられる. 本手法では, ピアノ楽譜のもつ繰り返し構造として左手のリズムの要素のみから自己類似度行列を求めた. しかし, 原曲の音色やテクスチャに関する特徴は, ピアノ楽譜において演奏記号や音数でも表現が可能である. そのため, 音色やテクスチャに関する特徴を考慮したピアノ楽譜の自己類似度行列が作成できず, t 値が低くなったのではないかと考えられる.

自由度調節済決定係数 R_{adj}^2 は最大値 1 に比べ低い値となった. この原因として, ある 1 つの楽曲に対して定まった一つの正解ピアノアレンジ楽譜があることはなく, 何通りものピアノアレンジが存在することが考えられる. 本研究では 3 章で述べた通り, 一般にピアノアレンジにおいて意識される大事な点に注目して生成したが, 編曲家によって表現の捉え方が異なる場合もある. そのため, R_{adj}^2 が比較的低い値となったと考えられる. しかしながら, R_{adj}^2 は正の値であり, 音響特徴量よりピアノ楽譜の持つ構造を推定することはある程度有効であると言える.

表 2 より, 推定された自己類似度行列は, 一つの楽曲区間に含まれる伴奏の種類数 λ が小さくなるほど正解楽譜との誤差が小さくなり, 特に閾値 th が -0.07 と -0.10 の時に誤差が小さくなった. この結果より, 推定された自己類似度行列を楽曲区間に分割する際は程よい粗さで分割を行い, 各楽曲区間内でなるべく統一された伴奏を選択することで正解楽譜の持つ繰り返し構造に近づくと言える. しかし, 実際の楽譜には同じ楽曲区間において異なる 2 種類の伴奏が繰り返し出現することや, 次の楽曲区間を導く役割として楽曲区間の最終小節で他とは違った伴奏が出現することがある. 従って, より実際のピアノ楽譜に近づけるためには, 同一楽曲区間内での繰り返し構造や楽曲の転換にも焦点を当てることが重要であると考えられる.

9. まとめと今後の課題

本研究は音楽音響信号より楽曲構造を考慮したピアノアレンジ手法を提案した. 本手法を用いて 1 つの楽曲に対し, 構造の反映具合が異なる複数のピアノ楽譜を出力した. ま

た, 既存ピアノ楽譜と音響データの持つ構造の関係を調べるため, それぞれの自己類似度行列を算出し, 重回帰分析を行った. その結果, 音の高さやリズムに関する特徴が構造推定に有効であることが確認された. また, 楽曲構造を考慮することで, より既存ピアノ楽譜に近い特徴を持つことが確認された. 今後は, より編曲家により作成されたピアノアレンジ楽譜に近づけるために, 同一楽曲区間内の繰り返し構造や曲の転換に焦点を当て, ピアノアレンジ楽譜の生成を行いたい.

謝辞 本研究の一部は JST ACCEL (JPMJAC1602) の支援を受けた.

参考文献

- [1] 高森啓史, 深山覚, 後藤真孝ほか: 音楽音響信号から得られる音楽要素に基づく自動ピアノアレンジ, 情報処理学会研究報告音楽情報科学, Vol. 2017(116), No. 13, pp. 1-4.
- [2] 藤田顕次, 大野博之, 稲積宏誠: 習熟度を考慮した複数楽譜からのピアノ譜生成手法の提案, 情報処理学会研究報告音楽情報科学, Vol. 2008(77), No. 10, pp. 47-52.
- [3] Chiu, S., Shan, M. and Huang, J.: Automatic system for the arrangement of piano reductions. *Proc. IEEE International Symposium on Multimedia*. 2009, pp. 459-464.
- [4] Owsinski, B.: *The Mixing Engineer's Handbook*. Thomson Course Technology, 1999.
- [5] Nakamura, E. and Sagayama, S.: Automatic Piano Reduction from Ensemble Scores Based on Merged-Output Hidden Markov Model. *Proc. ICMC*, 2015, pp. 298-305.
- [6] Percival, G., Fukayama, S., and Goto, M.: Song2Quartet: A System for Generating String Quartet Cover Songs from Polyphonic Audio of Popular Music, *Proc. ISMIR*, 2015, pp. 114-120.
- [7] Schmidt-Jones, C.: *The Basic Elements of Music*. Lulu.com, 2014.
- [8] Doll, C.: Rockin' Out: Expressive Modulation in Verse-Chorus Form. *Society for Music Theory*, 2011, Vol. 17, No. 3, pp. 1-10.
- [9] Depuro MP: ワンランク上のピアノソロ ポカロ神曲大集合ベスト 30, 菊倍版, 2016.
- [10] Depuro MP: 上級ピアノ・グレード ポカロ名曲 ピアノ・ソロ・コンサート, 菊倍版, 2015.
- [11] ヤマハミュージックエンターテインメントホールディングス, Print Score, 入手先 (<https://www.print-gakufu.com/>) (参照 2018-07-27)
- [12] Böck, S. Krebs, F. and Schedl, M: Evaluating the Online Capabilities of Onset Detection Methods, *Proc. ISMIR*, 2012, pp.49-54.
- [13] 後藤真孝, 吉井和佳, 藤原弘将ほか: Songle: 音楽音響信号理解技術とユーザによる誤り訂正に基づく能動的音楽鑑賞サービス, 情報処理学会論文誌, 2013, Vol. 54, No. 4, pp. 1363-1372.
- [14] Jonathan, F.: Automatic audio segmentation using a measure of audio novelty, *Proc. IEEE International Conference on Multimedia and Expo*, 2000, Vol. 1, pp.452-455.
- [15] 国立研究開発法人産業技術総合研究所, Songle Widget, 入手先 (<http://widget.songle.jp/>) (参照 2018-07-27)
- [16] 後藤真孝, 橋口博樹, 西村拓一ほか: RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース, 情報処理学会論文誌, 2004, Vol. 45, No. 3, pp. 728-738.