

旋律特徴と楽器特徴を考慮した旋律クラスタリングによる 電子オルガン譜へのリダクションシステム Automatic Electronic Organ Reduction System Based on Melody Clustering Considering Melody and Instrument Characteristics

田中 大貴*

Daiki Tanaka

法政大学大学院 情報科学研究科 情報科学専攻

Email: 17t0016@cis.k.hosei.ac.jp

Abstract

To maintain the artistic side of music, it is important that many people play it. However, musical arrangement is difficult and time-consuming for amateurs, so even good music is not arranged for instruments that relatively few people play. In this study, we propose an arrangement system that can easily generate scores for various musical instruments, with the aim of helping to keep music alive. Here, we target electronic organ arrangements that can represent a full score. Since such scores include a large number of instrumental parts, we summarize the melodies that play the same role using four features (pitch, harmony, note length, and timbre of the instrument). Next, the system select a melody cluster for each part, considering the connections between the melodies and the characteristics of electronic organs. Finally, we adjust the score to make it easier to play on an electronic organ. We then conduct experiments to confirm whether the system can create arrangements that maintain the original piece's overall impression. Here, we compared our new arrangements of six pieces with the existing electronic organ scores. The average values of the right hand, left hand, and foot parts were 0.79, 0.76, and 0.73. The results showed that the proposed arrangement system can create an electronic organ score from the full score.

1 はじめに

現在、オーケストラなどで演奏される楽曲を別の楽器の楽譜へ編曲することが行われている。このような編曲が行われる目的として、楽曲の芸術的側面の維持、新たな創意工夫の機会、楽譜販売市場の拡大による経済的意義、作品の演奏用途の変更、レパートリーの拡大といった演奏実績の広がりなどがあげられる。

一方で課題もある。編曲は手間がかかり、素人には困難な作業である。そのため、どんなに良い曲でも人気がなかったり、おのおの楽器の演奏者が少なかったりすると、編曲されずに楽譜が販売されないことがある。しかし、芸術的側面を維持するためには、様々な楽器で多くの人に演奏してもらうことが重要である。

こういった状況がある中で、ピアノやギターなどの楽譜へ編曲は多々行われているが、電子オルガンを対象にした編曲はあまりやられていない。電子オルガンは他の楽器と違い、様々な楽器の音色が出せて、一台で多彩な楽曲を演奏できる素晴らしい楽器である。このように良い特徴を持つにもかかわらず、編

曲されていない。その理由として、電子オルガンの音色や演奏する旋律の数が多く、自由度が高いことから、重要な旋律を探して編曲することが難しいからだ。そこで本稿では、電子オルガンにおける編曲の意義が損なわれないようにするために、電子オルガン譜への編曲システムを提案する。これにより、編曲の負担が減らすこともできると考えられる。

オーケストラなどの演奏で使われる複数パートで構成される総譜(図1)を別の楽器の楽譜へ編曲する方法として、リダクションというものが存在する。リダクションとは、可能な限り元の楽譜に忠実になるよう、総譜から主要な旋律を残して、それ以外を削る編曲方法である。本研究では、リダクションに基づいた編曲システムを構築する。

まず、総譜の楽器パートが多いので、旋律の音高、和声、音価(音の長さ)、楽器の音色の特徴を用いて、同じ役割を持つ旋律をまとめる。次に、旋律のつながりと電子オルガンの特徴を考慮して、電子オルガンの右手パート、左手パート、足パートに対応する旋律クラスタを選択する。最後に、電子オルガンで演奏できるように楽譜を修正する。

図1. 総譜：オーケストラや吹奏楽の演奏で指揮者が利用する全パートの旋律が載っている楽譜

2 電子オルガンへのリダクション

2.1 電子オルガンの特徴

電子オルガンの演奏方法は、他の楽器と比べて独特である。ピアノは右手と左手を使って演奏するのに対して、電子オルガンは右手、左手、足を使って演奏する。また、ピアノは一つの鍵盤を使うのに対して、電子オルガンでは三つの鍵盤を使って演奏する。そのため、図2のように楽譜が三段組になっている。図2の上段は右手パート、中段は左手パート、下段は足パートに対応している。

電子オルガンの右手パートの旋律は、主旋律のような旋律を奏でる傾向がある。左手パートの旋律は、和音や和音を含む旋律を奏でることが多く、同時発音数が4音以下である特徴がある。足パートの旋律は、低音の旋律や和音の根音を主に奏で、必

* Supervisor: Prof. Katunobu Itou

ず単音である。右手パートと左手パートの同時に発音する時の音域は1オクターブ以内である。

電子オルガンは、約900種類の楽器の音色を搭載している。実際の楽器の音源を忠実に再現しているため、打弦楽器、撥弦楽器などの特徴を補って、電子オルガンの鍵盤で演奏できる。演奏時は、一つの鍵盤に複数の楽器の音色を重ねて演奏する。また、音色の自動切り替えが可能であるため、異なる楽器の旋律を続けて演奏できる。更に、電子オルガンでは打楽器などのリズムボックスが用意されている。楽曲のリズムを予め用意しておく、それに合わせて演奏するのが基本的な演奏スタイルである。



図2. 電子オルガンの楽譜

2.2 関連研究

複数パートの旋律を圧縮して、他の楽器の楽譜へ編曲する方法に関する研究が複数ある。例えば、ポッターらの研究[1]ではギターアレンジの手法を提案している。これはギターの演奏上の制約を考慮して、重要な音符の集合を選んでギター譜へ編曲している。しかし、この方法は旋律のフレーズ単位ではなく、音符単位に処理される。したがって、選択された音符が異なる楽器から選択された場合、旋律のフレーズが失われる可能性がある。

ピアノ譜へのリダクションに関する研究として藤田らの研究[2]がある。これはピアノの右手は主旋律を主に奏で、左手は伴奏を奏でるといった観点から、平均音高、発音時間、音高・リズムパターンに着目して、総譜の旋律からメロディとベースラインを推定している。そして、それらの旋律をピアノの右手パートと左手パートに採用するという方法である。しかし、この方法ではポリフォニックのような曲の場合、メロディを一意に決めることができないので、不向きである。更に、ピアノが常にメロディとベースラインを引き続けているとは限らないので、この方法では不十分である。

他にも中村らの研究[3]がある。これは、可能な限り元の楽曲の旋律を使用して、鍵盤上の運指を考慮しながら総譜を圧縮してピアノ譜へ編曲する手法を述べている。しかし、音楽的な特徴が考慮されていないことから、演奏できない理由から楽曲の重要な旋律が消える可能性がある。

本研究では、電子オルガンを対象にした編曲をする。そのため、ピアノと異なり足パートの特徴を考慮した編曲が必要になる。また、電子オルガンの特徴の音色を考慮した編曲もしなければならない。したがって、上記の手法では技術では不十分である。そこで、本研究では楽曲の役割を一意に決めない。そして音符単位ではなく、小節単位の旋律として考える。更に、電子オルガンの特徴の音色を新たに考慮して編曲を行う。

まず、音楽的な特徴を考慮して、同じ役割を持つ旋律を小節単位にグループ化する。次に、既存の電子オルガン譜から右手パート、左手パート、足パートの特徴と各パートが演奏する旋律の特徴、更に小節間のつながりを考慮して、各パートに対応する旋律グループを選択する。最後に、電子オルガンで演奏できるように修正を行う。

3 電子オルガン譜への編曲システム

3.1 旋律クラスタリング

総譜の各々のパートは主旋律、ハーモニー、装飾などの役割を持っている。そのため、総譜が約30パートで構成されていたとしても、旋律の役割を考慮して分類することで、5パート程に集約できる。総譜のパートを集約するために、本稿ではクラスタリングを使用する。

3.1.1 特徴量

クラスタリングに使用する特徴量は、旋律の音高 (Melodic Pitch)、和声 (Consonance Activity)、音価 (音の長さ) (Sonic Richness)、楽器の音色 (Instrument Timbre) の五つである。これらの特徴を考慮して、総譜のパート数を削減する。

■**旋律の音高** 旋律の音高の特徴量は、旋律の音高の変化とリズムを考慮した特徴量である。旋律の多くは音高の上下があることが多いため、動きの少ない伴奏パートと旋律パートを別のフレーズとして扱うことが多い。また、複数の旋律は互いにハーモニーを奏でるため、同じ時刻 (リズム) で発音されたり、同じように音高を変化させたりすることが多い。例えば音程が異なる図3(a)(b)の旋律は、同じ音高の変化とリズムで奏でられているので、似た旋律とみなせる。それに対し、図3(c)の旋律は変化とリズムが異なるため、別のフレーズとみなせる。

一番短い音符の長さを基本単位として、単位時間ごとに音符が発音されたタイミングに音高に対応した数値、それ以外を0として、要素数 n のベクトル MP を生成する。この時、音高の数字はMIDI (Musical Instrument Digital Interface) のノートナンバーに1を足した1から128の値で表現する。

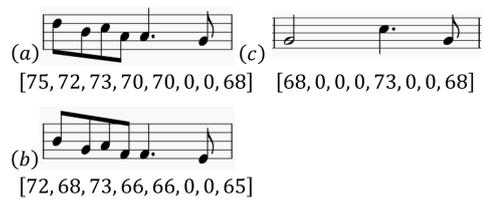


図3. 旋律の音高の特徴ベクトル

■**和声** 音高の動きが似ていても、旋律パートのように非和声音を多く含むパートと伴奏パートのように和音を多く含むパートを別のフレーズとして認識することが多い。そこで、単位時間ごとに全パートが鳴らしている上位3つの音を求め、図4(a)のような近似的な和声音を作成する。

単位時間ごとに和声音を含む場合を1、含まない場合を0として、要素数 n のベクトル CA を生成する。これにより、和声音をあまり含まない図4(b)の旋律と和音を多く含む図4(c)の旋律を区別できる。

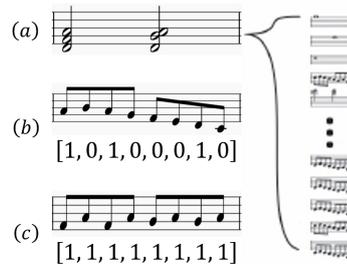


図4. 和声の特徴ベクトル

■**音価** 音価は楽器の響き方に関する特徴量である。これにより、休符を含まない図5(a)の旋律と休符を含む図5(b)の旋律を区別できる。

単位時間ごとに音が発音された拍を2、音を伸ばしている拍を1、鳴っていない拍を0として、要素数 n のベクトル SR を生成する。

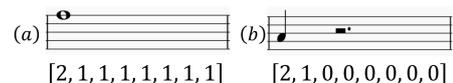


図5. 音価 (音の長さ) の特徴ベクトル

■**楽器の音色** 旋律の音高、和声、音価の特徴量は、視覚的に似ている旋律をまとめるための特徴量である。そのため、視覚的に旋律は似ていないが、同じ役割を持つような旋律をまとめることができない。音楽は音で表現されるものであるため、視覚情報以外にも音響的な情報が欠かせない。例えば、オーケスト

ラの楽曲では、木管楽器、金管楽器、弦楽器がそれぞれまとまって役割が同じ旋律を演奏する特徴がある。また、似た音色の楽器は同じような旋律を担当することで、音色のバランスが考慮されて楽曲が作られている。電子オルガンでも、この音色の特徴が反映されて編曲されている。そこで、視覚的な点に着目した特徴量に加えて音響的な特徴を考慮する。新たに音色に基づいた楽器間の距離の特徴量を追加する。

では、音色の違いはどこにあるのか。ローズの文献 [4] によると楽器の音色は、音の振幅エンベロープ、フォルマント構造、時間的な周波数成分の変動などであると述べられている。また、楽器が音を発した瞬間（アタック）に大きな特徴があるとも言われている。本稿では、音色の特徴が顕著に表れるアタックまで時間とフォルマント構造の時間変化、そして周波数成分の形状に着目して音色を表現する。

フォルマント構造を分析する手法として、LPC(Linear Predictive Coding)[5]、LPC ケプストラム [6]、ケプストラムやメル周波数ケプストラム係数(MFCC)[7]などが挙げられる。ここで対象とするのは楽器音であり、各楽器において発音可能な音域が異なる。したがって、同じ音域で比較することが難しい。また、楽器音は基本周波数の整数倍の周波数が強く表れる倍音構造の性質が見られる。そのため、周波数が強く表れている部分も含めたフォルマント構造を求める LPC は、ここでは適していない。そこで、音程の影響をあまり受けない分析方法として、本稿では MFCC を用いて音色の分析を行う。

周波数成分の形状を表現する特徴量として、スペクトル重心 [8] というものがある。これはスペクトルの強さをもとに重心を計算する。スペクトル重心の値が大きければ、より多くの高周波数成分を含むため、明るい音色を持っていると判断できる。

音色の特徴抽出をする際、オーケストラや吹奏楽の演奏で用いられる 28 種類の楽器音を使う。これらの楽器音は RWC データベース [9] の実音源を使用している。楽器の音色は音程によって変化するので、ここでは総譜で用いられる約 2 オクターブ分の音域の音程を用いる。本稿では、アタックまでの音声を用いて、フレーム長 64 ms、シフト長 8 ms とし、14 次元の MFCC でフォルマント構造を求め、それを楽器の音色の特徴量とする。また、アタックまでの時間を求め、それも特徴量として使用する。

求めた特徴量を用いて、音色のモデル化をする。実験的に調べた結果、アタックまでの時間と MFCC、スペクトル重心のすべての特徴量を用いた時に、一番良く音色を分類できた。よって、ここでは三つの特徴量すべてを用いてモデルを生成する。本研究では、AIC(赤池情報量規準)[10]に基づいて、混合数 2 の混合ガウスモデル (Gaussian Mixture Model) を使用してモデル化を行う。

各楽器の音色をモデル化した後、モデルの距離を測る。本稿ではバタチャリヤ距離を使用して、各モデルの分布間の距離を計算する。パート i, j の音色の確率分布を P, Q としたとき、バタチャリヤ距離は式 (1) のように表される。この時、 $p(x), q(x)$ は音色の確率密度関数である。バタチャリヤ距離は、値が 0 に近い時は二つのモデルの距離が近いということである (式 (2))。

$$D(i, j) = -\ln \left(\sum_{x \in X} \sqrt{p(x)q(x)} \right) \quad (1)$$

$$0 \leq D(i, j) \leq \infty \quad (2)$$

バタチャリヤ距離を GMM へ拡張する。ユーらの文献 [11] によると、GMM へ拡張したバタチャリヤ距離は式 (3) によって表される。 K は GMM の混合数である。また、混合数 K の確率分布 P において、 μ_k^p はの平均ベクトル、 $\det \Sigma_k^p$ は共分散行列、 w_k^p は混合比を表している。混合数 K の確率分布 Q においても同様である。

$D(i, j)$ はペアワイズな値である。それに対して他の特徴量(旋律の音高, 和声, 音価)は、楽器ごとに求めているため表現が

異なっている。本稿では、多次元尺度構成法 (MDS)[12] を用いて、 $D(i, j)$ の値を楽器ごとの値へと変換する。ここでは 6 次元の特徴量へと変換し、正規化したベクトルを IT とする。

$$D(i, j) = \frac{1}{8} \sum_k^K (\mu_k^p - \mu_k^q)^T \Sigma^{-1} (\mu_k^p - \mu_k^q) + \frac{1}{2} \sum_k^K \ln \frac{\det \Sigma}{\sqrt{\det \Sigma_k^p \det \Sigma_k^q}} - \frac{1}{2} \sum_k^K \ln (w_k^p w_k^q) \quad (3)$$

$$\Sigma = \frac{\Sigma_k^p + \Sigma_k^q}{2} \quad (4)$$

図 6 は、各楽器の音色の距離をわかりやすく可視化するために、Isomap[13] で IT ベクトルを 3 次元の楽器空間で示した図である。例えば、トランペットとホルネットは似たような楽器であり、音色も似ている。そのため、近いところに位置していることが確認できる。その他、チューバとユーフォニウム、トロンボーンとテナートロンボーンも近いところに位置している。以上より、作成した音色のモデルは適切なモデルである。

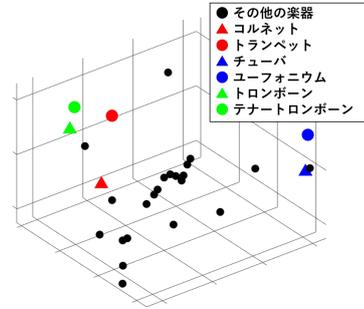


図 6. 3 次元の楽器空間で音色の距離を可視化した結果

3.1.2 尺度

本研究ではファジィ・クラスタリング [14] を使用する。ファジィ・クラスタリングとは、k-means 法のようにクラスを 0,1 で分類するのではなく、あるクラスに 0.8 の割合で属し、もう一方のクラスには 0.2 で属す、というように曖昧に分類できるクラスタリングである。したがって、旋律が重複したクラスを生成できるようになる。電子オルガンの場合、右手パートと左手パートが同じ旋律を演奏することもあるため、曖昧に分類して重複を許すファジィ・クラスタリングを採用した。

次にクラスタリングに用いる尺度を定義する。まず、求めた四つの特徴量ベクトルに重み $w_k (k = 1, 2, 3, 4)$ を掛け合わせ、一つの特徴量ベクトルとして表現する。式 (5) はパート i の特徴量ベクトルを示している。重みは足して合わせて 1 になる値である。

$$\mathbf{X}_i = [w_1 \mathbf{M} \mathbf{P}_i \ w_2 \mathbf{C} \mathbf{A}_i \ w_3 \mathbf{S} \mathbf{R}_i \ w_4 \mathbf{I} \mathbf{T}_i] \quad (5)$$

式 (5) で定義した特徴量ベクトルを用いてクラスタリングを行う。式 (6) は目的関数を示している。 N は楽器パート数。 C はクラスターの重心数。 m は曖昧性を決めるパラメータであり、1 より大きな値である。 u_{ic} はパート i のクラスター c に対する帰属度。 μ_c はクラスター c の重心である。

$$J = \sum_{i=1}^N \sum_{c=1}^C (u_{ic})^m \|\mathbf{X}_i - \mu_c\|^2 \quad (6)$$

帰属度 u_{ic} は式 (7) を用いて更新する。

$$u_{ic} = \left(\sum_{l=1}^C \left(\frac{\|\mathbf{X}_i - \mu_l\|^2}{\|\mathbf{X}_i - \mu_c\|^2} \right)^{\frac{1}{m-1}} \right)^{-1}, \quad (\mathbf{X}_i \neq \mu_c) \quad (7)$$

また、更新した u_{ic} から新しいクラスタの重心 c を求めるために、式 (8) を用いる。

$$\mu_c = \frac{\sum_{i=1}^N u_{ic}^m \mathbf{X}_i}{\sum_{i=1}^N u_{ic}^m} \quad (8)$$

3.1.3 旋律のクラスタリング

小節単位にクラスタリングを行う。実験的に調べた結果、重心の数が5である時、適切なクラスタリングができた。よって、ここでは初期の重心の数を5とする。そして、各重心に近い旋律を4つ取ることで最終的なクラスタを生成する。クラスタリングは次の手順で行う。

- I 重心の数 $C = 5$, $m = 1.5$, パート数 N , 一小節分の各パートの音符データを取得。
- II 取り出した各パートの音符データから 3.1.1 で定義した特徴量を用意して、重心数 K のファジィ・クラスタリングを実行。
 - (a) 初期クラスタ重心 C を与える。
 - (b) 式 (7) を用いて、 C から帰属度 u_{ic} を更新。
 - (c) 式 (8) を用いて重心 C を更新。
 - (d) 収束条件を満たしたら終了。満たさなかったら b) へ戻る。
- III クラスタリングをした結果、各重心から近い旋律を4つ取得。一小節ずらしてIに戻る。もし最後の小節であれば、クラスタリングを終了。

3.2 旋律クラスタの選択

小節単位の選択では、右手パート、左手パートが入れ替わる問題が起きるので、旋律のつながりを考慮した隠れマルコフモデル (HMM)[15] を利用してクラスタを選択する。

楽譜とは離散記号である音符の集合で表現され、音響信号などで存在している。これらの離散記号に対する確率モデルとして、ここでは離散分布 HMM を使用する。本研究での HMM は、図 7 のようなトポロジーになっていて、8 分音符を一つの状態とし、開始状態は 1,2,3,4、終了状態は 5,6,7,8 となっている。また、各状態では $[C, C\#, \dots, B, \text{休符}]$ の 13 次元の音高情報のベクトルを一次元に圧縮したシンボルを出力する。

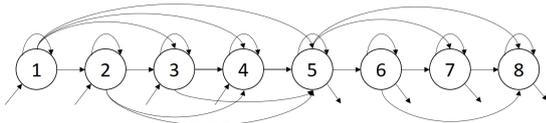


図 7. 旋律のつながりを考慮した HMM のトポロジー

13 次元の音高情報のベクトルを圧縮するために、ベクトル量子化 (VQ)[16] を使用する。VQ とは、ベクトルで表現されたデータの集合を代表パターン (クラスタの重心) に置き換える処理である。代表パターンのリストはコードブック (クラスタの重心のリスト)、クラスタ番号はコードと呼ばれている。各ベクトルデータは、代表パターンとの距離が一番近いコードへと置き換えられる。コードブックの学習には、k-means クラスタリングのアルゴリズムを用いる。本研究では、音高情報のベクトルが 13 次元なので、13 次元のコードブックを 32 個生成する。すなわち、音高情報のベクトルが 1~32 の数字のシンボルに表現しなおされるということである。実験の結果、コードブックの数は 32 個が適切であった。VQ されたベクトルを図 7 の HMM に対して使用する。

HMM の学習では、既存の電子オルガン譜を利用して旋律のつながりを学習させる。既存の電子オルガン譜を利用することで、各パートが演奏する旋律の特徴を学習できると考えられる。学習には図 8 にあるように、小節の区切りの前後二拍 (計四拍) の旋律を用いる。この旋律を VQ よりシンボル列に変換して HMM の学習を行う。右手パートの旋律を使用して右手の HMM、左手パートの旋律で左手パートの HMM、足パートの旋

律で足パートの HMM といったように、それぞれのパートに対応する HMM を生成する。

旋律クラスタを選択する時は、各クラスタ間の旋律を右手パート、左手パート、足パートの HMM に対して尤度を計算して、値が大きい旋律クラスタを選択する。この時、右手パート、左手パート、足パートの順に選択し、同じクラスタが選ばれないものとする。

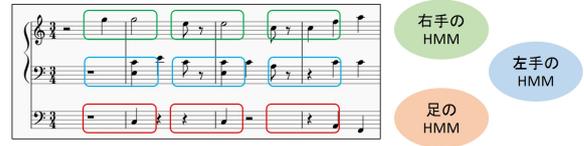


図 8. 既存の電子オルガン譜を用いた HMM の学習

3.3 楽譜の修正

まずは、クラスタリングで生成されたクラスタの修正を行う。ファジィ・クラスタリングは、重複した旋律を含むクラスタが生成できる。一方で、旋律を含みすぎたり含まずに足りなかったりする問題が起きる。これは前後の小節の関係を見ていないことが原因である。そこで、前後 2 小節 (計 5 小節) のクラスタリングの結果を用いてクラスタを修正する。例えば、左手パートに前後のクラスタに 3rd, 4th ヴァイオリン、ヴィオラ、チェロが採用されているのに対して、間的小節では 1st, 2nd, 3rd, 4th ヴァイオリンが採用されているとする。このように、前後のクラスタのパートが全て一致しているのに対して、間的小節のクラスタが異なる場合、間的小節のクラスタを修正する。

次に、電子オルガンでの演奏上の制約を考慮した修正を行う。クラスタ内の音符は各楽器に対応した音高のままであるため、実際に鍵盤で演奏することができないことがある。そこで、電子オルガンで演奏できるように修正する。楽譜は小節単位に同じ時刻の右手パート、左手パート、足パートを見て修正する。

まず、足パート旋律の修正する。足パートは二音同時に演奏することがないので、二音同時に音が存在してしまう箇所を修正する。重なっている二音が同じ音名の音である時、一つの音へまとめる。この時、発音楽器数が多い音を基準にしてまとめる。例えば、図 9(a) の上の旋律の発音楽器数が多いならば、図 9(b) のように上の旋律へと集約される。二音が異なる音名の音である時、和音の根音に対応する音を残して他の音を消す。また、和音の根音に対応する音がない場合は、発音楽器数が多い旋律を残す。以上の修正の他に、足鍵盤の音域は約 1.5 オクターブ (C2 から G3) であるため、その音域内になるように修正する。



図 9. 音域の修正 (a) 修正前 (b) 修正後

次に右手パートと左手パートの修正である。右手パートと左手パートは、同時に鳴っている最小の音高と最大の音高の差が 1 オクターブ以上ある時に修正をする。なぜなら、指が届かず演奏できないことがあるからだ。足パートと同様に、同じ音名の音の場合は発音楽器数が多い音高を基準に一つの音にまとめる。これも、図 9 のように修正される。異なる音名の音の時は、1 オクターブ未満になるように音の集約を行う。左手パートは、右手と交差して演奏しにくいことがあるので、C3 から C4 の間の音域を基準にして集約を行う。例えば、1 オクターブ以上離れている図 10(a) は、同じ音名の音が使用されている。これを一つの音にまとめると図 10(b) になる。更に、この状態では右手と重なってしまうので、全体の音域を変更すると図 10(c) のように修正される。

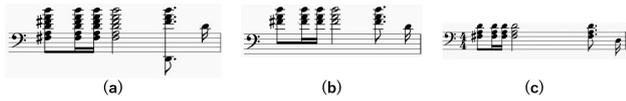


図 10. 左手の修正 (a) 修正前 (b) 音符の集約後 (c) 音域の修正後

4 評価実験

4.1 実験概要

原曲の特徴が保たれ、電子オルガンの特徴が考慮された適切な楽譜ができていのかを評価する。従来の評価は、編曲した楽曲を聴き比べたり、編曲した楽譜が演奏できるか楽器演奏の上級者に演奏してもらったりして評価している。しかし、前者の評価だと演奏可能であるか評価されていない。また、後者の評価だと十分な評価ができる程の楽器演奏者を集めることが困難である。そこで、本研究では元の楽曲の特徴を保って編曲され、演奏可能な楽譜になっているのか、既存の電子オルガン譜との比較で評価した。

評価の一致率の計算は式 (9) で行う。不足とは、模範の楽譜において音符がある箇所に、編曲後の楽譜では音符が何もないということである。挿入は、音符がない箇所に音符が入ってしまうこと。置換は、本来音符がある箇所に音符があるのだけれども異なる音になってしまっているということである。

右手パート、左手パート、足パートごとに一致率を計算し、その結果を確認する。また、ここでは各楽曲においてクラスタリングで使用する特徴量の最適な重みを実験的に求め、それらの特徴量を全楽曲に対して網羅するように適用して一致率を計算する。

$$\text{一致率} = \frac{\text{全体の音符数} - (\text{不足数} + \text{挿入数} + \text{誤り数})}{\text{全体の音符数}} \quad (9)$$

一致率と編曲後の楽譜を聴き比べ、主観的に判断基準を設けたところ、一致率が 0.8~1.0 の時は演奏可能な楽譜ができていて、元の曲の特徴を保った編曲であることが確認できた。0.5~0.8 の時は、演奏しにくい楽譜であるが、聴いたところ曲の特徴を保った編曲がされている。0.5 以下の場合には、演奏しにくい楽譜であり、演奏も違和感があるような編曲になっていた。一方で、編曲者によって細かな編曲が異なるため、楽譜も多種多様に存在する。そのため、一致率を 1 にすることは極めて難しいと言える。模範にする楽譜によって、一致率が左右されることに注意しなければならない。

編曲する総譜は、既存の電子オルガン譜と同じようなアレンジになっているものを使用する。また、パート数が異なり、曲全体の小節数も異なる。更に、曲の途中で拍子や調が変わる楽曲を使用している。評価対象の楽曲は、『1. モテット「アヴェ・ヴァルム・コルプス」K.618』『2. 組曲「展覧会の絵」よりプロムナード』『3. 愛のあいさつ』『4. 交響曲第 36 番「リンツ」K.425 より第 3 楽章メヌエット』『5. 「クーブランの墓」より前奏曲』『6. 「ソナチネ」よりメヌエット』の 6 曲である。

4.2 実験結果

編曲結果の例として『モテット「アヴェ・ヴァルム・コルプス」K.618』を示す。図 11 は、曲中の 4 小節分を示した楽譜である。図 11(a) は編曲の模範となる既存の電子オルガン譜で、図 11(b) は本システムで編曲した楽譜である。緑色が不足している音符で、赤色が誤って挿入された音符を示している。この曲の一致率は、表 1 に示されていて、右手が 0.84、左手が 0.80、足が 0.75 となっている。このことから、楽曲『モテット「アヴェ・ヴァルム・コルプス」K.618』は、右手、左手、足パートの一致率がすべて 0.7 以上であるので、演奏しにくい箇所もあるが、曲の特徴を保った編曲ができた。



図 11. 編曲結果 (a) 既存の電子オルガン譜 (模範となる編曲例)、(b) システムでの編曲結果

表 1. 各楽曲の評価結果

曲名	パート数	小節数	右手	左手	足
アヴェ・ヴァルム・コルプス	13	46	0.84	0.80	0.75
展覧会の絵	20	24	0.80	0.77	0.71
愛のあいさつ	6	99	0.80	0.82	0.73
交響曲第 36 番「リンツ」K.425	14	57	0.82	0.79	0.74
「クーブランの墓」より前奏曲	6	98	0.75	0.69	0.73
「ソナチネ」よりメヌエット	10	82	0.77	0.73	0.72
平均	—	—	0.79	0.76	0.73

5 考察

5.1 不適切な編曲

右手パート、左手パート、足パートの挿入誤りの平均は 0.13、不足誤りの平均は 0.08、置換誤りの平均は 0.03 であった。これより、挿入と不足の誤りによる不適切な編曲が多かった。これは、クラスタの選択を改善する必要があると考えている。誤った旋律クラスタを選択した時、次の二つのことが原因で挿入と不足による誤りが起こる。

一つ目は、正しいクラスタを選ばず、他パートのクラスタやそれ以外のクラスタを選択した時だ。例えば、既存の電子オルガン譜の右手パートの旋律と、システムが右手パートに選択したクラスタの旋律とで異なる時、誤った音が挿入されたり不足したりするため一致率が下がってしまう。

二つ目が、誤ったクラスタの選択により、3.3 節で述べたクラスタの修正処理が機能しない時である。この修正では、前後の小節のクラスタ内のパートを確認し、間的小節を修正するという方法である。例えば、前後の小節が誤ったクラスタを選択していた場合、間的小節のパートは修正されない、もしくは異なるパートに修正されてしまう。そのため、既存の電子オルガン譜の旋律と異なり、挿入や不足が起きる。

以上、2 点の問題は誤りの約 4 割を占めている。よって、旋律の選択処理を改善する必要があると考えている。

現在のクラスタの選択では、音高は考えずに旋律のつながりのみを考慮した処理を行っている。そのため、旋律の途中で音高が異なっても、つながりが良ければ正しくない旋律クラスタでも選択されることになる。

この問題を解決するために、旋律のつながりに加えて電子オルガンの各パートの音域を考慮した選択を行うことが良いと考えている。電子オルガンの鍵盤や楽譜は、右手、左手、足の順に、演奏する旋律の音域が低くなる性質がある。この特徴を反映させるために、既存の電子オルガン譜から右手パート、左手パート、足パートの各パートに対して使用される音高のヒストグラムを生成する。そして、あるクラスタにおいて、クラスタ内に含まれる音高からどのパートに属するのが尤もらしいか判断、つまりクラスタ内の音高と各パートの音高のヒストグラムとの尤度を計算して選択することを考えている。これにより、電子オルガンの性質がさらに考慮された選択ができるため、正しい旋律クラスタを選びやすくなると考えられる。そして、上記 2 点の問題の解決につながり、精度が向上すると考えている。

5.2 音色の特徴量の追加

本稿では、新たに音色の特徴量を加えた。表 2 は音色の特徴量を加える前のシステムの 4 曲分の評価結果である。音色の特

微量を加えた後(表 1)と比べて、左手パートは 0.10、足パートは 0.11 上昇した。

音色の特徴量を加える前、左手パートでは旋律の不足による誤りが多かった。従来の特徴量は、旋律のリズムや動きが同じ旋律をまとめていた。そのため、旋律の動きが異なるけれど、同じ役割を持つ旋律をまとめてられなかった。これが、明示的に似ていないけれど、役割が同じ旋律をまとめてられない問題を引き起こす原因である。音色が似ている楽器は、旋律の動きが異なっても同じ役割を果たし演奏する特徴がある。よって、音色の特徴量の追加により、異なる動きをしている旋律でもまとめられるようになった。したがって、精度の向上につながった。

右手パートの精度は 0.02 上昇していたが、他のパートと比べるとそれほど変わっていない。これは鍵盤楽器の特徴が現れていると考えられる。鍵盤楽器の場合、右手パートの旋律は他のパートと比べて動きが多くて早い。実際の電子オルガンの右手パートは、異なる動きをする旋律を演奏するより、リズムが揃い、同じ動きをする旋律が、1 曲の中で 8 割程占めている。したがって、従来の明示的に似ている旋律をまとめる特徴量だけでも十分であった。以上のことより、音色の特徴量を加えても不足していた旋律が少し増えただけであって、精度はそれほど変わらなかった。

足パートの精度も 0.11 ポイント上昇した。音色の特徴量を追加する前は、追加後よりも約 1.6 倍の音の挿入があった。電子オルガンの足パートは、チューバやユーフォニウムなどの低音楽器の旋律を奏でることが、1 曲の中の約 8 割と多い。しかし、従来の特徴量を用いたクラスタリングでは、音域が高いフルートの旋律とまとまるがあった。そのため、高音楽器の旋律が挿入されたり置換されたりして精度が落ちてしまった。今回の音色の特徴量を追加したことで、低音楽器のみの旋律クラスタが生成されやすくなり、低音楽器の旋律どうしがまとまりやすくなった。これより、高音の旋律が除外され、従来よりも挿入数が減った。

表 2. 音色の特徴量を除いた編曲結果

曲名	右手	左手	足
アヴェ・ヴァルム・コルプス	0.79	0.73	0.56
展覧会の絵	0.75	0.53	0.60
愛のあいさつ	0.82	0.75	0.71
交響曲第 36 番「リッツ」K.425	0.74	0.66	0.63
平均	0.77	0.66	0.62

5.3 システムの応用

実験に使用した各楽曲において、クラスタリングに使用した特徴量の重みは、表 3 のようになっている。表 3 より、旋律の音高の特徴量 MP が良く効いていることがわかる。今回は鍵盤楽器の電子オルガンが対象であることから、これらの特徴量が効いていると考えられる。鍵盤楽器の場合、複数ある旋律を一人で演奏すると、同じ動きをしたり同時に発音できたりする旋律でないと、演奏するのが難しい。従って、鍵盤楽器の特徴がうまく反映された編曲ができた。

今回の結果から、各楽器の特徴に合わせた特徴量の重みを設定することで、他の楽器にも対応できると考えられる。例えばギターを対象とした編曲を行う場合を考える。ギターは主にコード弾きをする楽器であるため、和声の特徴量 CA の重みを大きくするのがいいだろう。和音が考慮されたクラスタができるので、ギターらしい旋律を作れる。また、クラスタの選択の時、和音を含む旋律と低音部分の旋律を選ぶことで、ギターの楽譜が作成できる。更に、ギターの演奏しやすさ、運指を考慮することで、より完成度が高い編曲を行えるだろう。

予備実験として、キーボード、ギター、ベース、サクソ、ヴァイオリンによるアンサンブル譜への編曲を試みた。手順としては、クラスタリングで使用する特徴量の重みを変更し、各楽器に対応しそうな旋律クラスタを選択した。楽譜の修正は、手作業で行ってしまったが、各楽器の特徴に合わせた修正アルゴリズムを考案すれば自動で楽譜が作れるだろう。編曲の結果、それ

なりに良いクラスタが生成され、各楽器で演奏できるような楽譜になった。ここで作成した楽譜を用いて、大学内の懇親会でアンサンブルを披露した。

表 3. クラスタリングに使用した重み

曲名	MP	CA	SR	IT
アヴェ・ヴァルム・コルプス	0.4	0.2	0.2	0.2
展覧会の絵	0.4	0.2	0.2	0.2
愛のあいさつ	0.4	0.2	0.3	0.1
第 3 楽章メヌエット	0.4	0.1	0.3	0.2

6 おわりに

本研究では、総譜から電子オルガン譜への編曲システムを提案した。旋律のクラスタリングでは、音色の特徴量を取り入れたことで、音響的に似た旋律をまとめることができた。また、ファジィ・クラスタリングのアルゴリズムを用いたことで、各々のクラスタに重複した旋律を含めるようになった。これにより、右手、左手パートが同じ旋律を演奏することがあるという電子オルガンの特徴に沿ったクラスタを生成できるようになった。旋律クラスタの選択では、HMM を使用することで小節間の旋律のつながりを考慮した選択を行えるようになった。実験の結果、右手パート、左手パート、足パート共に約 70%、模範となる既存の電子オルガン譜と一致した編曲ができた。模範の電子オルガン譜と異なった原因は、旋律クラスタの選択による、旋律の挿入と不足であった。既存の電子オルガン譜と異なった旋律クラスタが採用されたため、一致率が低下した。つながりのみを考慮している選択に加え、電子オルガンの各パートの音高も考慮した選択を行うことで、改善できると考えている。

参考文献

- [1] R. Daniel and W. D. Potter: "GA-based Music Arranging for Guitar," Proceedings of International Congress on Evolutionary Computation, 2006.
- [2] K. Fujita, H. Oono, and H. Inazumi, "A proposal for piano score generation that considers proficiency from multiple part", IPSJ Special Interest Group on Music and Computer, 2008, pp. 47-52.
- [3] E. Nakamura and S. Sagayama, "Automatic Piano Reduction from Ensemble Scores Based on Merged-Output Hidden Markov Model," in Proc. ICMC, 2015, pp.298-305.
- [4] C. Roads, Computer music—History · Technology · Art, Tokyo Denki University Press, 2001.
- [5] J. Makhoul, "Linear prediction: A tutorial review", Proceedings of the IEEE, Volume:63, Issue:4, April 1975, pp. 561-580.
- [6] S. Yoshii, Digital audio processing: Tokai University Press, 1998.
- [7] T. Kobayashi, "Cepstral and Mel-Cepstral Analysis of Speech", The Institute of Electronics, Information and Communication Engineers, 1998, pp. 33-40.
- [8] G. Tzanetakis, P. Cook, "Musical genre classification of audio signals", IEEE Trans. on Speech and Audio Processing, 10(5):293-302, 2002.
- [9] S. Goto et al., "RWC Music Database: Music Genre Database and Musical Instrument Sound Database", Information Processing Society of Japan, IPSJ SIG Technical Reports 2002-MUS-45-4, Vol.2002, No.40, pp.19-26, May 2002.
- [10] Akaike, H., "Information theory and an extension of the maximum likelihood principle", Proceedings of the 2nd International Symposium on Information Theory, Petrov, B. N., and Caski, F. (eds.), Akadimiai Kiado, Budapest: 267-281 (1973).
- [11] C.H. You, K.A. Lee, and H. Li, "Gmm-svm kernel with a bhattacharyya-based distance for speaker recognition," Audio, Speech, and Language Processing, IEEE Transactionson, vol.18, no.6, pp.1300-1312, 2010.
- [12] J. Edwards and P. Oman, Dimensional Reduction for Data Mapping-A practical guide using R, R News, Vol. 3/3, 2003, pp. 2-7.
- [13] J.B. Tenenbaum, V. de Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. Science, 290(5500):2319-2323, 2000.
- [14] H. Ichihashi, Introduction to Cluster Analysis, Morikita Publishing, 1999.
- [15] T. Kathiresan, "Automatic melody generation", Master's thesis, KTH Royal Institute of Technology, 2015, pp. 25-43.
- [16] R. M. Gray, "Vector quantization", IEEE ASSP Mag., vol. 1, pp. 4-29, 1984.