

楽曲特徴量分析に基づく バッハ・コラールの構造的分析

福知山公立大学 情報学部情報学科

32345035 川人すばる

指導教員 橋田光代 准教授

提出日 2025 年 1 月 31 日

改訂日 2025 年 2 月 7 日

— 目次 —

1	はじめに	1
2	特徴量抽出	1
3	分析手法	1
4	分析 1 回目	1
4.1	考察	1
5	分析 2 回目	1
5.1	考察	2
6	分析 3 回目	2
6.1	考察	2
7	分析に対するまとめ	2
8	長さによる影響	2
9	長さに影響を及ぼす値を省いた分析	3
10	今後の課題	3

— 図目次 —

1	バッハのコラールを PCA で可視化した散布図	1
2	バッハのコラールの分析をヒートマップで可視化したもの	1
3	バッハのコラールクラスタ 1 を PCA で可視化した散布図（クラスタ 1 のみ）	1
4	バッハのコラールの分析をヒートマップで可視化したもの（クラスタ 1 のみ）	2
5	バッハのコラールクラスタ 1 を PCA で可視化した散布図（サブクラスタ 1 のみ）	2
6	バッハのコラールの分析をヒートマップで可視化したもの（サブクラスタ 1 のみ）	2

— 表目次 —

1	抽出した特徴量一覧	1
---	---------------------	---

1. はじめに

本研究では、楽曲の特徴量を数値的に抽出し、その構造を分析することを目的とする。本年度はバッハのコラールを対象とし、特徴量分析およびクラスタリングによる構造把握を試みた。研究目的としては、楽曲ごとの曲の違いをデータとして閲覧したかったというのと、今後の研究にも応用できる研究であると考えたためである。

2. 特徴量抽出

楽曲の解析には Python を用い、music21 ライブラリによって以下の特徴量を抽出した。

表 1: 抽出した特徴量一覧

特徴量	意味	どの性質を表すか
measures	小節数	曲の長さ
avg_chord_size	平均和音サイズ (同時発音数)	和声の厚み
sevenths_ratio	7th 和音の割合	和声の複雑さ
pitch_mean	平均音高	全体の音域の中心
pitch_std	音高の標準偏差	メロディの動きの大きさ
pitch_range	最高音 - 最低音	音域の広さ
leap_rate	跳躍進行の割合	メロディの飛び方の多さ
smoothness	音程差の逆数的な指標	滑らかさ
up_rate	上行音程の割合	メロディの方向性 (上昇か下降か)
dur_whole/half/quarter/eighth/sixteenth	各音価の出現比率	リズムの構成 (長音中心か短音中心か)
vel_mean	平均ベロシティ (強さ)	曲の平均的な音量
vel_std	ベロシティの変動	強弱表現の豊かさ
prog_unique	ユニークな和音進行数	ハーモニーの多様性
rn_seq_len	ローマ数字解析による進行列の長さ	和声進行の規模

3. 分析手法

抽出した特徴量に対して標準化を行い、主成分分析 (PCA) によって次元削減を行った。その後、クラスタリング手法を用いて楽曲の分類を試みた。

4. 分析 1 回目

PCA による散布図を図 1 に示す。

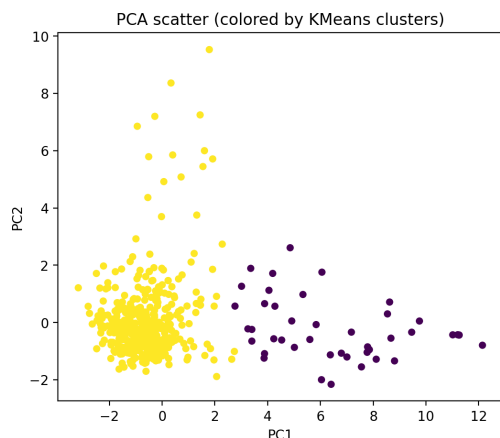
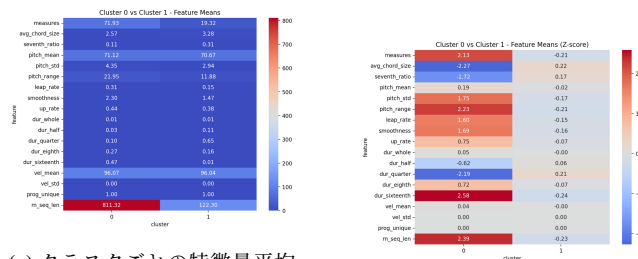


図 1: バッハのコラールを PCA で可視化した散布図

黄色で示した部分がクラスタ 0、青色で示した部分がクラスタ 1 である。また、クラスタごとの特徴量平均の比較を図 2a に、クラスタごとの Z-score (平均からの相対距離) を図 2b に示す。



(a) クラスタごとの特徴量平均

の比較

(b) クラスタごとの Z-score

図 2: バッハのコラールの分析をヒートマップで可視化したもの

4.1 考察

分析の結果、図 2b より、クラスタ 0 がクラスタ 1 に対して曲の長さ、音域の広さ、16 分音符の割合、和声進行の規模が大きくなっており、クラスタ 0 は曲が長く、多数の音域と細かいリズムで複雑な曲が含まれていると推測できる。一方、クラスタ 1 は曲が短く、音域が狭く、長い音符が多い単純な曲が含まれていると考えられる。

5. 分析 2 回目

クラスタ 1 を見た所、まだ約 400 曲含まれていたため、さらに分類ができるのではないかと思います、クラスタ 1 のみを使い再びクラスタリングを行った。PCA による散布図を図 3 に示す。

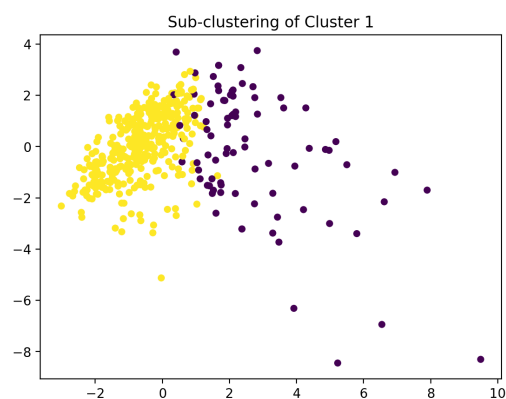
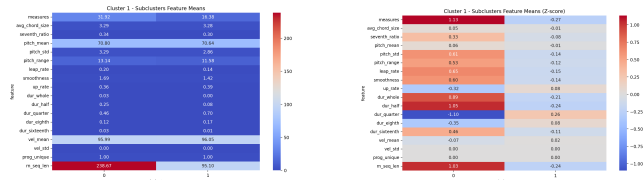


図 3: バッハのコラールクラスタ 1 を PCA で可視化した散布図 (クラスタ 1 のみ)

黄色で示した部分がサブクラスタ 0、青色で示した部分がサブクラスタ 1 である。また、クラスタごとの特徴量平均の比較を図 4a に、クラスタごとの Z-score (平均からの相対距離) を図 4b に示す。



(a) クラスタごとの特徴量平均の比較 (クラスタ 1 のみ) (b) クラスタごとの Z-score (クラスタ 1 のみ)

図 4: バッハのコラールの分析をヒートマップで可視化したもの (クラスタ 1 のみ)

5.1 考察

分析の結果、図 4b より、サブクラスタ 0 は半音符の割合が高く、四分音符の割合が低くなっています。曲の長さもサブクラスタ 0 とサブクラスタ 1 では相違が見られます。このことから、サブクラスタ 0 はリズムがゆったりとした曲が多く含まれていると推測できます。

6. 分析 3 回目

サブクラスタ 1 を見た所、まだ約 300 曲含まれていたため、さらに分類ができるのではないかと思います。クラスタ 1 のみを使い再びクラスターリングを行った。PCA による散布図を図 5 に示す。

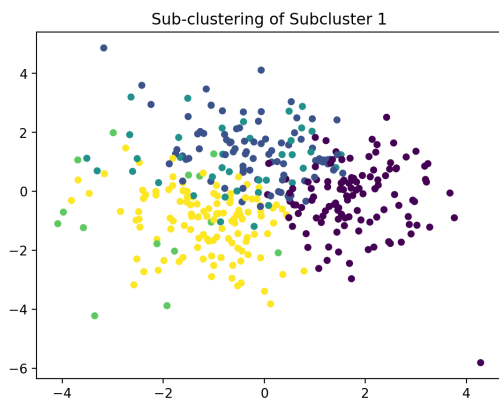
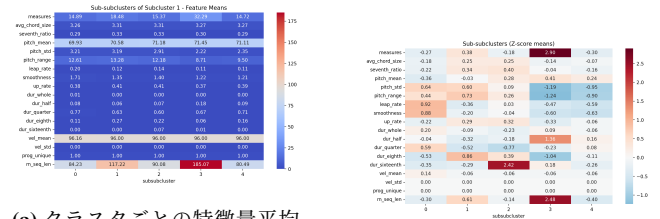


図 5: バッハのコラールクラスタ 1 を PCA で可視化した散布図 (サブクラスタ 1 のみ)

また、クラスタごとの特徴量平均の比較を図 6a に、クラスタごとの Z-score (平均からの相対距離) を図 6b に示す。



(a) クラスタごとの特徴量平均の比較 (サブクラスタ 1 のみ) (b) クラスタごとの Z-score (サブクラスタ 1 のみ)

図 6: バッハのコラールの分析をヒートマップで可視化したもの (サブクラスタ 1 のみ)

6.1 考察

分析の結果、図 6b より、サブクラスタ 2 とサブクラスタ 3 では大きな相違が見られている。しかし、図 6a を見た所、例で言うところの `dur.sixteenth` の割合が他のクラスタが 0 に近いのに対し、サブクラスタ 2 は 0.07 であり、全体的に数値が小さくなっていることが分かる。また、図 5 を見ると、各クラスタに大きな境がなく、クラスターリングがうまくいっていない可能性があると考えられる。このことから、クラスターリングを行い過ぎていると考えられる。また、サブクラスタ 1 に含まれている曲がバッハが作成するコラール曲の特徴を多く含んでいるものと推測できる。

7. 分析に対するまとめ

本研究では、バッハのコラール曲を特徴量に基づいてクラスターリングし、それぞれのクラスタの特徴を分析した。最初のクラスターリングでは、曲の長さや音域、リズムなどの特徴に基づいて 2 つのクラスタに分類された。さらに、クラスタ 1 を再びクラスターリングすることで、より細かい分類が可能となった。その結果、サブクラスタ 1 に含まれる曲がバッハが作成するコラール曲の特徴を多く含んでいることが示された。しかし、過度なクラスターリングにより、クラスタ間の明確な境界が失われる可能性があることも示された。また、今回はバッハが作成するコラール曲のみで分析を行ったため、他の作曲家やジャンルのと比較しての分析は行っていない。他の楽曲に対しても同様の分析を行うことで、より、バッハが作成するコラール曲の特徴が理解できるだろう。

8. 長さによる影響

先ほどのまとめで、楽曲の長さがクラスターリングに影響を与えている可能性があると言った。そのため、楽曲の長さが各特徴量にどのような影響を及ぼしているのかを相関係数を用いて考えてみる。

ここから先は研究途中であるため省略する。

9. 長さに影響を及ぼす値を省いた分析

楽曲の長さに影響を及ぼすと考えられる特徴量を省き、再度クラスタリングを行った。

ここから先は研究途中であるため省略する。

10. 今後の課題

今後の研究では、特徴量の選択やクラスタリング手法の改善を検討し、より精度の高い楽曲分類を目指していく。その際、特徴量に対しては、楽曲分析を行っている先行研究を参考にし、より音楽的な意味合いを持つ特徴量を多数選択、抽出し、分析に必要な特徴量の取捨選択をすることが重要だと考えている。また、バッハが作成するコラール曲以外の楽曲も含め同様の分析を行い、比較研究を進めることで、楽曲の特徴や構造に関する理解を深めていきたいと考えている。橋田准教授より、データセット (PEDB) への運用を提案されたため、今後の研究では PEDB を用いた分析も検討していく。