

メロディラインによるジャンル判別

鈴木崇也[†] 長谷川智史[†] 穴田一^{††}

現在，世界中には様々なジャンルの楽曲が存在し，それらの楽曲の自動ジャンル判別に関する研究が盛んに行なわれている．その中で，旋律パートのみに着目した高精度のジャンル判別は未だに実現されていない．そこで，本稿では旋律のみを用いてジャンル判別を行なうための新たな特徴量を提案し，その有効性の検証を行なった．Rock, Pop, Blues, Country の 4 ジャンル，計 184 曲に対して実験を行なった結果，Pop の楽曲に対して約 64%，全体で約 46% の判別精度を示し，提案手法の有効性が示された．

Automatic Music Genre Classification Using Melody Lines

Takaya Suzuki[†] Satoshi Hasegawa[†]
Hajime Anada^{††}

Currently, there are musical pieces of various genres all over the world. Many people had studied about automatic music genre classification. But, accurate music genre classification with melodies has not yet realized. Therefore we propose new features in melodies and validate the utility of those. And we experimented on 4 genre classification (Rock, Pop, Blues, Country, total 184 musical pieces) by k-nearest neighbor algorithm. Experimental result by using features which we proposed show success rates of 64% for Pop, and 46% for 4 genres.

1. はじめに

現在，世界中で様々な楽曲が作曲，演奏され，人々の耳に届いている．また，インターネット環境の普及や，iPod に代表される大容量携帯オーディオプレイヤー，スマートフォンの普及により，数万曲の音楽をいつでもどこでも持ち歩けるのはもちろんのこと，聴きたい曲をその場でダウンロードして楽しむことも可能である．それに伴い，音楽情報処理の技術も，急速な発展が求められている．その中でも，音楽情報検索 (MIR: Music Information Retrieval) という分野では，大量に存在する楽曲の検索，分類，管理に対する取り組みに関する研究が盛んに行なわれてきている [1]．その中の 1 つに，楽曲の自動ジャンル判別に関する研究がある．

楽曲の自動ジャンル判別に関する研究は，国際会議 ISMIR (International Symposium on Music Information Retrieval) においても，主要なテーマとして扱われ，盛んに議論が行なわれている．現在まで，様々なジャンルの判別をする研究が行われているが，対象とするデータにより，大きく分けて 3 つに分類することが出来る．1 つ目は，音響データを用いる方法．2 つ目は，MIDI データを用いる方法．そして 3 つ目はそれらを組み合わせる方法である [2]．音響データを用いる方法として，土橋らの研究 [3] が挙げられる．土橋らは，楽曲のベースラインに着目し，ベースラインから抽出した特徴量と，音色，リズムに関する特徴量を用いてジャンル判別を行い，ベースラインから得られる特徴量の有効性を示している．MIDI データを用いる方法としては，McKay らの研究 [4] が挙げられる．McKay らは，楽曲の MIDI データからメロディやリズム，コードなどに関する様々な特徴量を抽出し，高い精度の判別を実現している．また，土橋らの研究のように，1 つのパートに着目したジャンル判別の研究として，Simsekli の研究 [5] がある．Simsekli はベースラインのデータのみに着目し，隣接する音符の音高差を用い，音楽理論を元に重み付けしたユークリッド距離を用いた新しい方法で，高い判別精度を示している．

ジャンル判別の分野では，楽曲には様々なパートが存在する中，ベースラインのみ注目した研究が多く行なわれている．ベースは，音程を持つ楽器であり，さらにドラムやパーカッションのようなリズムパートの役割も担っている．また，ほとんどの楽曲に存在し，無音時間も少ないという点で，ジャンル判別には有用であると言われている．

しかし，ほとんどの楽曲に存在すると言う点では，旋律パートも考えられる．また，

[†] 東京都市大学大学院 工学研究科 システム情報工学専攻
Systems Information Engineering,
Graduate School of Engineering, Tokyo City University

^{††} 東京都市大学 知識工学部 経営システム工学科
Department of Industrial and Systems Engineering,
Faculty of Knowledge Engineering, Tokyo City University

旋律パートは楽曲のメインとなるパートであり、音程やリズムに関して、その楽曲の特徴が大きく表れているパートであると考えられる。しかし、旋律パートのみに着目した高精度のジャンル判別は未だに実現されていない。また、旋律パートのデータのみでジャンル判別を実現できれば、少ない容量のデータで大規模な楽曲データの分類や、検索を可能にすることができる。そこで、本研究では、旋律パートであるメロディラインに着目したジャンル判別の手法について提案し、有効性の検証を行なっていく。

2. 特徴量の抽出

ジャンルを判別するためには、楽曲からそれぞれのジャンルの特徴を捉えた特徴量を抽出しなければならない。本研究では、電子楽器の演奏データを機器間でデジタル転送するための世界共通規格である MIDI のデータを楽曲データとして用い、特徴量の抽出を行う。本章ではまず、特徴量を抽出するための前段階である、MIDI データの加工について述べ、続いて楽曲データから抽出する特徴量について述べる。

2.1 MIDI データの加工

まず、MIDI データを編集するソフトを用いて旋律データのみを抽出を行なう。抽出する旋律データは最初の音符が発音されてから、最後の音符の発音が終わるまでとする。旋律のみを取り出した MIDI データより、音の高さと音の長さ、TimeBase 値の 3 種類を抽出する。TimeBase 値とは、通常その MIDI データの 4 分音符の長さ定めた値である。また、旋律パートにおいて同時刻に複数の音符が存在する楽曲は除外した。同時刻に発音されたわけではないが、音が鳴っている時間が重複している場合は、先に発音された音符の音の長さを重複している音符が発音するまでの長さとし、同時に発音されている部分は取り除いた。

次に、抽出した音の高さ、音の長さ、TimeBase 値を用いて、楽曲データを作成する。音の高さは、MIDI で元々割り当てられているノートナンバーをそのまま用い、これを「音高」とする。ノートナンバーはピアノ鍵盤上の中央 C の音を 60 とし、半音を 1 として表されている。続いて音の長さは、各音符の音の長さを TimeBase 値の 4 倍の値で割った値とし、これを「音長」とする。これにより、4 分の 4 拍子の楽曲において、音長が 1 で 1 小節の長さとなる。また、同じ楽曲を MIDI にした場合でも、MIDI の楽曲の制作者によって実際に音が鳴っている時間表すゲートタイムが異なる場合が多い。ゲートタイムとは、キーボードで例えるならば、鍵盤を押している長さのことで、通常は楽譜上で表されている音符の長さの 80~90% に相当し、残りの時間は無音となる。これを統一するため、各音符が発音されてから、次の音符が発音されるまでの長さに、各音符の音長を修正する。しかし、間奏部分等で、明らかに無音となると考えられる部分も多く存在する。そこで、音長が 1 より大きい、すなわち 1 小節より

大きい無音時間は休符とみなす。

以上の操作で得られた音高、音長、休符を楽曲データとし、分析を行っていく。

2.2 特徴量

本研究で提案するジャンルの特性を表すと考えられる特徴量は、大きく分けて 3 つに分類される。1 つ目は、音符同士の音高の差や、休符の割合など、旋律から計算される基本的な特徴量で、これを「基本特徴量」とする。2 つ目は、音高と音長の値から計算される「音符速度」という概念を取り入れた「音符速度特徴量」。3 つ目は、いくつかの音符を 1 つのかたまりとしてとらえ、このかたまりから計算される特徴量で、これを「フレーズ特徴量」とする。以下に、各特徴量について詳しく述べる。

(1) 基本特徴量

● 音高差頻度

隣接する音符同士の音高の差の頻度。各音符で隣接する音符との音高の差を求め、各音高差を全音高差数で割り、全音高差に対する割合にする。隣接する音符が休符の場合は除外する。本研究では、正の方向に 13 半音、負の方向に 13 半音、絶対値をとった 13 半音、そして音高差 0 の計 40 種類をそれぞれ特徴量とした。

● 音高差・音長差パターンの頻度

各音符の音高、音長の値がどのように推移しているかを、隣接する 3 つの音符の音高及び音長からなる 9 つのパターンに分類し、楽曲ごとに各 9 パターンの頻度を求めた。そのパターンについて以下に示す。

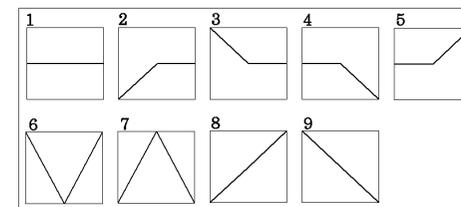


図 1 音高、音長パターン

3 つの音符の音高の変動で 9 パターン、音長の変動で 9 パターンに分類される。各パターンは、音高、または音長が上昇するか下降するか、変化しないかで表される。例えば音高パターンの場合、パターン 1 は、3 音の音高が全て同じであることを表し、パターン 2 は 1 音目よりも 2 音目の音高が高く、2 音目と 3 音目の音高が同じであることを表す。パターン 6 は、1 音目より 2 音目が低く、2 音目より 3 音目の音高が高くなることを表し、パターン 8 は、3 音の音高が順に高くなることを表す。音長についても同様に考える。

この図の 9 つのパターンの各楽曲の出現頻度を求め、それを全パターンの総出現数で割ることにより、それぞれ「音高差パターンの頻度」、「音長差パターンの

頻度」とする。

- 音長差平均

音高差と同様に、隣り合う音符同士の音長の差を計算し、その絶対値をとったものの平均である。

- 音域

旋律の音域である。楽曲で使用されている音符の最高音高と、最低音高の差で表される。

- 休符割合

旋律に存在する休符の割合を表す。休符の長さの合計を旋律全体の長さで割ることにより計算される。

(2) 音符速度特徴量

- 音符速度パターンの頻度

各音符の音高が、どの程度の時間で次の音高に変化しているかを表す「音符速度」という尺度を考案した。音符速度は、次の図のように隣接する音符同士の音高の差を、1つ目の音符の音長で割ることにより求められる。次にその計算方法の例を示す。

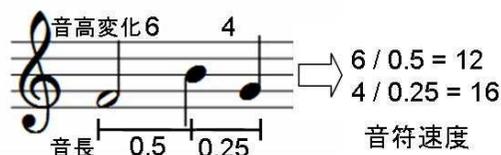


図 2 音符速度の計算例

音符速度の計算例を示す。左から1つ目の音符とその次の音符との音高の差の絶対値が6であり、これを音高変化と考え、1つ目の音符の音長で割ることにより計算された値が音符速度となる。すなわち、どの程度の時間で次の音高に変化しているかを表し、この場合12となる。2つ目の音符と3つ目の音符との音符速度も同様に計算され、16となる。

この図のように各音符間の音符速度を計算する。また、音長が TimeBase 値の8分の1、すなわち32分音符以下の音符は、装飾音符である可能性が高く、32分音符以下の音符は音符速度を計算する上では除外した。

各音符で音符速度を計算した後、音高差・音長差パターンと同様に、これらの値がどのように推移しているかを、隣接する3つの音符速度からなる9つのパターンに分類した。そして、楽曲ごとにそれぞれの頻度を求め、それらを全パターンの総出現数で割ることにより、「音符速度パターンの頻度」とする。

(3) フレーズ特徴量

- 音高フレーズ

旋律の楽譜上の音符を線で結んでみると、大小様々な山のような形状が繰り返されていることがわかる。本研究では、この山のような形状を1つのフレーズとして考え、次の図のように音高が上昇し、下降し、再び上昇する1つ前の音符までを「音高フレーズ」とした。

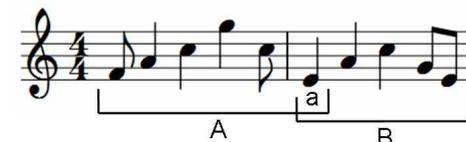


図 3 音高フレーズの定義

音高が下降から上昇に変化する音符をフレーズの開始音とし、再び下降から上昇に変化した音符の手前の音符をフレーズの最後の音とする。A、Bはそれぞれフレーズを表し、重複している音符aは両方のフレーズに含めることとする。

各楽曲を音高フレーズに分割した後、次の7つの特徴量を計算する。

- ・ フレーズに含まれる音符数の平均
- ・ フレーズ高さの平均
- ・ フレーズ幅の平均
- ・ フレーズ幅に対する高さの割合の平均
- ・ フレーズ頂点の位置の平均
- ・ 全音長に対するフレーズ幅の割合
- ・ 音域に対するフレーズ高さの割合

フレーズ高さの平均とは、各音高フレーズに含まれる音符の最高音高と最低音高の差の平均である。フレーズ幅の平均とは、各音高フレーズに含まれる音符の音長の合計の平均である。フレーズ幅に対する高さの割合の平均とは、各音高フレーズのフレーズ高さをそのフレーズ幅で割り、その平均を計算したものである。フレーズ頂点の位置の平均とは、各音高フレーズの最高音高が、フレーズ幅に対してどの位置にあるかを表すものであり、各フレーズの最高音高が出現する音符の1つ手前までの音符の音長の合計を、そのフレーズのフレーズ幅で割り、その平均を計算したものである。全音長に対するフレーズ幅の割合とは、フレーズ幅の平均を各楽曲の音長の合計で割ったものである。音域に対するフレーズの高さの割合とは、フレーズ高さの平均を、基本特徴量の1つである「音域」で割ったものである。

● 平均音高のフレーズ

音高フレーズにより、旋律をいくつかのかたまりに分割した。ここでは、そのかたまりの変動について考えてみる。各音高フレーズに含まれる音符の平均音高を計算してみる。その各フレーズの平均音高を1つの要素とし、旋律の楽譜のように、横軸時間（フレーズの出現する順番）、縦軸をフレーズの平均音高として配置していくと、楽譜上の音符と同様に山のような形状が繰り返されていることがわかる。そこで、音高フレーズと同様、音高フレーズの平均音高により形成される山のような形状を1つのフレーズとし、これを「平均音高のフレーズ」とする。平均音高のフレーズの定義に関しては、休符が存在しないこと以外は、音高フレーズと同様である。各音高フレーズの平均音高 p_{ave} は以下の式で計算される。

$$p_{ave} = \sum_{k=1}^n p_k l_k / \sum_{k=1}^n l_k \quad \dots$$

ここで、 n は音高フレーズに含まれる音符数。 p_k と l_k はそれぞれ、音高フレーズに含まれる k 番目の音符の音高と音長を表す。この値を1つの要素とし、特徴量の計算を行なう。平均音高のフレーズも音高フレーズとほぼ同様の特徴量を計算するが、各要素が音高の平均しか持っていないため、フレーズの幅や、全音長、音域等の概念が無い。そこで、平均音高のフレーズに関しては次の4つの特徴量を計算する。

- ・ フレーズに含まれる要素数の平均
- ・ フレーズ高さの平均
- ・ フレーズに含まれる要素数に対する高さの平均
- ・ フレーズ頂点位置の平均

フレーズ高さの平均とは、各平均音高フレーズに含まれる要素の最高値と最低値の差の平均である。フレーズ頂点位置の平均とは、各平均音高フレーズにおいて、最高値の要素が出現する1つ手前までの要素の個数を、フレーズに含まれる要素数で割り、平均を計算したものである。

以上の特徴量を楽曲データから抽出する。基本特徴量は、「音高差頻度」の40種類、「音高差・音長差パターンの頻度」の18種類、そして「音長差平均」と「音域」、「休符割合」の全61種類。音符速度特徴量は、「音符速度パターンの頻度」の9種類。フレーズ特徴量は、「音高フレーズ」の7種類と、「平均音高のフレーズ」4種類の11種類。計81種類の特徴量を用意した。

3. ジャンル及び対象楽曲

ジャンル判別を行なうにあたり、どのジャンルを判別の対象とするかを設定する必要がある。本研究では、土橋らの研究[7]で用いられている midi データが配布されているサイト[7]で分類されているジャンルから選定し、Rock, Pop, Blues, Country の4ジャンルとした。また、実験に用いる楽曲は、土橋らの研究でも用いられているものを含むフリーサイト[7][8][9]から収集した各ジャンル46曲の計184曲である。

4. ジャンル判別実験

(1) 特徴選択に用いる指標

ジャンル判別実験に用いる特徴量として、本研究では81種類の特徴量を用意した。しかし、全特徴量がジャンル判別に有効であるとは限らず、判別精度を低下させる特徴量が存在する可能性もある。特徴量の全組み合わせに対して実験を行なうことが出来れば問題ないのだが、膨大な組み合わせ数となってしまう計算が困難なため、有効な特徴選択を行なう方法を考える必要がある。そこで本研究では、各特徴量のジャンル判別に有効であるかどうかの指標として、クラス分離尺度[3][6]を用いることにした。クラス分離尺度は、クラス内分散をクラス間分散で割ることにより求められ、質的変数と量的変数の相関とも言われている。本研究の場合、クラスはジャンルを指す。このクラス分離尺度の値が高い特徴量から1つずつ順に特徴量を増やしていく方法で実験を行なう。

(2) 交差検証法を用いた k-近傍法

ジャンル判別に用いる手法として、ユークリッド距離を用いた k-近傍法を採用した。近傍数は1から30までとし、それぞれの近傍数で交差検証法を行った。交差検証法は、全楽曲を1曲ずつ抜き出しテストデータとして実験を行い、誤判別率の平均をその近傍数での結果とした。そして、各近傍数に対して10試行ずつ行い、誤判別率の平均が最小となる近傍数での結果を出力とした。

5. 実験結果

実験に用いるデータとして、特徴量81種類から、予め他の特徴量と0.9以上の相関係数を持つ特徴量については、クラス分離尺度が高いものを採用し、75種類の特徴量とした。クラス分離尺度の高い特徴量から順に1つずつ特徴量を追加し、全75種類のデータパターンに対して1から30までの近傍数で交差検証法を行い、誤判別率の平均が最小となる近傍数の結果を出力した。また、各特徴量は予め平均0、分散1に正規化したものを用いた。その結果について次に示す。

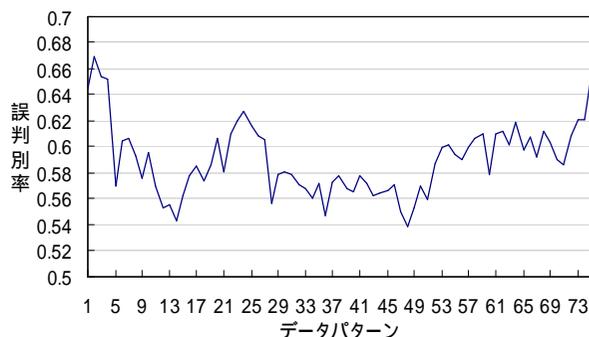


図 4 75 データパターンでの k-近傍法の結果

縦軸は誤判別率、横軸はデータパターンを表す。誤判別率は、1 から 30 までの近傍数で交差検証法を行い、その平均が最小となる近傍数での結果である。データパターンは 1 から 75 までクラス分離尺度の高い特徴量を順に 1 つずつ追加したものである。48 種類の特徴量を用いたデータパターンでの誤判別率が約 0.54 と最も低いことがわかる。

この図より、48 種類の特徴量を用いたデータパターンでの誤判別率が約 0.54 と最も低く、ジャンルの判別に有効な特徴量の組み合わせであると考えられる。ここで、ジャンル毎の判別率を見るために、48 種類の特徴量を用いたデータパターンにおいて、再び k-近傍法で実験を行なった。試行回数は 100 回とし、近傍数は最も誤判別率の低かった 10 近傍とする。各ジャンルに対する判別率及び正答率を以下の表に示す。

表 1 各ジャンルにおける判別率

	Rock	Pop	Blues	Country
Rock	0.46	0.39	0.09	0.06
Pop	0.15	0.64	0.16	0.04
Blues	0.08	0.24	0.50	0.18
Country	0.20	0.27	0.31	0.22

各行名はテストデータのジャンル、各列名は k-近傍法により出力されたジャンルを表す。各数値は、その行名のジャンルのデータが列名のジャンルに判定された割合を表す。例えば Rock の行に着目すると、Rock のデータが Rock と判別された割合が 0.46、Rock のデータが Pop と判定された割合が 0.39 となる。

この表より、Pop の正答率が約 64% と最も高く、次いで Blues、Rock とある程度高い正答率を示していることがわかる。しかし、Country は 22% と正答率が低く、正し

く判別することが出来ていない。

以上の結果より、Rock、Pop、Blues の 3 ジャンルにおいてはある程度高い正答率が得られ、提案手法の有効性が示された。また、楽曲のメロディラインにもジャンルの特性が表れることが示唆された。しかし、旋律以外のパートから得られる特徴量も用いた手法に比べると、判別精度が高いとは言えない。また、Country に関しては全く判別できていないことから、改良すべき点があると考えられる。

図 4 を見てみると、各データパターンにおける誤判別率のばらつきが大きく、クラス分離尺度が高いからといって、必ずしもジャンル判別に有効な特徴量であると断定するのは難しいと考えられる。このような特徴選択の方法に関してだけでも様々な研究が行われており、今後更なる議論が必要であると考えられる。

表 1 を見てみると、Rock、Blues とともに Pop と誤判別されている割合が比較的大きいことがわかる。また、Country に関しては、どのジャンルに関しても同じ程度の誤判別率を示している。このような結果となった原因を探るため、主成分分析を行い、関係性の高い特徴量を少数の次元へ縮約し、ジャンルと特徴量との関係性を見定める。主成分分析には、k-近傍法で正答率の最も高かったデータパターンである 48 種類の特徴量を用いた。

主成分分析を行った結果、第 15 主成分までで累積寄与率が約 80% となった。ここでは、第 1 主成分と第 2 主成分に着目し、分析を行っていく。まず、第 1 主成分の主成分負荷量を見定める。主成分負荷量の大きい上位 3 つの特徴量を次に示す。

表 2 第 1 主成分負荷量上位 3 つ

音高フレーズ幅の平均	0.75
音高差頻度 (正の方向に 3 半音)	0.73
音高フレーズ高さの平均	-0.69

第 1 主成分の主成分負荷量の絶対値が大きい特徴量上位 3 つを示す。1 列目は特徴量、2 列目は主成分負荷量を示す。

この表より、音高フレーズに関する特徴量が正の方向、負の方向とも高いことがわかる。しかし、音高差・音長差パターンの頻度や、音符速度特徴量等、上位 3 つの特徴量以外に関しても、主成分負荷量が高いものも多く、単純に音高フレーズに関する主成分であると解釈することは難しい。第 2 主成分に関しても同様であり、主成分自体の解釈は困難である。ここで、第 1 主成分と第 2 主成分の主成分得点について見てみる。横軸を第 1 主成分得点、縦軸を第 2 主成分得点とし、各楽曲のジャンルをプロットした図を次に示す。

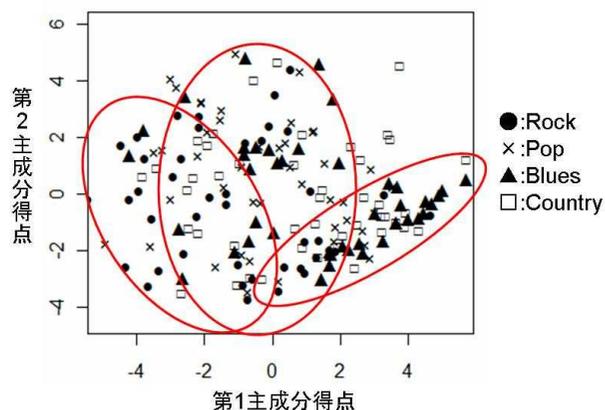


図 5 主成分得点のプロット

横軸は第 1 主成分得点、縦軸は第 2 主成分得点を表し、各楽曲の主成分得点に対応した座標に、ジャンルがプロットされている。それぞれ ●:Rock, ×:Pop, ▲:Blues, □:Country を表す。赤の楕円で囲っている部分は、左から Rock, Pop, Blues の楽曲が比較的密集していることを表している。

この図を見ると、各ジャンルの楽曲が第 1, 第 2 主成分得点の軸において、明確に分類されてはいない。しかし、比較的左下の楕円には Rock, 中央の楕円には Pop, 右下の楕円には Blues の楽曲が密集しているように見える。また、Rock と Blues の楽曲に着目し見ると、第 1 主成分の軸に対して、Blues は正の相関、Rock は負の相関が高い楽曲が多いことがわかる。Country の楽曲に関しては、全体に満遍なく分布しており、第 1, 第 2 主成分において特徴的な傾向は見られないことがわかった。

主成分分析の結果と表 1 の結果を照らし合わせてみると、Rock の楽曲が Blues と誤判別される割合と、Blues の楽曲が Rock と誤判別される割合が比較的低いのは、第 1 主成分得点の値が、それぞれ正と負の方向に離れているからであると考えることが出来そうである。第 1 主成分の解釈を踏まえて考えてみると、Blues の楽曲は、音高フレーズの幅が大きく、音高の上昇と下降をゆっくり繰り返している楽曲が多いと考えられる。また、Rock に関しては、音高フレーズの高さが高く、音高の変動に対する振幅が比較的大きいことを示す。Blues は、スローテンポで落ち着いた雰囲気のある楽曲が多いのに対し、Rock は、Heavy Metal や Hard Rock などのように、アップテンポで激しいサウンドの楽曲が比較的多く、それに伴い旋律の音高も大きく変動する楽曲が多いと考えられる。このことから、提案した特徴量が Rock や Blues の持つ特性を捉えられていることが示唆されたと言える。また、Rock と Blues の楽曲が Pop と誤判別される割合が高いのは、第 1 主成分軸において、Pop が Rock と Blues 楽曲に挟まれる形

で分布していることが原因であると考えられる。

そして、Country の正答率が低いのは、第 1 主成分と第 2 主成分を見る限りでも、Country 独自の傾向を捉えることが出来ておらず、提案した特徴量が Country の特性を捉えられていないことが原因であると考えられる。

6. おわりに

本稿では、旋律パートから得られる新しい特徴量を提案し、ジャンル判別に対する有効性の検証を行なった。クラス分離尺度を基準に選択した 48 種類の特徴量を用いた、k-近傍法によるジャンル判別実験では、Rock, Pop, Blues, Country に対して、約 46% の判別精度を得た。また、Pop の楽曲に対しては 64% と、高い判別精度を得ることが出来た。

ジャンル判別は様々なアプローチから研究が行われているが、旋律パートであるメロディラインにもそれぞれのジャンルの特性が表れていることがわかり、音楽情報検索に有用な一手法としての可能性を示せたのではないだろうか。しかし、本研究で対象とした 4 ジャンルすべてにおいては、高い精度で判別することは実現できなかった。このことから、今後は、旋律とジャンルの関係性を更に調査することはもちろんのこと、旋律以外のパートから得られるジャンル判別に有効な特徴量と旋律との関係性についても考慮し、新しい特徴量の考案を行なう必要がある。また、特徴選択の手法や、サポートベクターマシンやニューラルネットワークなど他の判別器に関しても、検討を進めていきたい。

参考文献

- 1) 後藤真孝, 平田圭二: 音楽情報処理の最近の研究(<小特集>音楽音響における最近の話題), 日本音響学会誌 60(11), pp.675-681 (2004)
- 2) Z. Cataltepe, Y. Yaslan, A. Sonmez: Music Genre Classification Using MIDI and Audio Features, Eurasip Journal on Advances in Signal Processing, Vol.2007, pp.1-9 (2007)
- 3) 土橋佑亮, 北原鉄朗, 片寄晴弘: 音響信号を対象としたベースラインからの音楽ジャンル解析, 情報処理学会研究報告, 2008-SLP-70, Vol.2008, No.12, pp.217-224 (2008)
- 4) C. McKay, I. Fujinaga: Automatic Genre Classification Using Large High-Level Musical Feature Sets, ISMIR, pp. 525-530 (2004)
- 5) U. Simsekli: Automatic Music Genre Classification Using Bass Lines, 2010 20th International Conference on Pattern Recognition, Vol.2, No.1, pp.4137-4140 (2010)
- 6) 土橋佑亮, 片寄晴弘: SOM を用いたベースラインからの音楽ジャンル解析, 情報処理学会研究報告, 2006-MUS-66, Vol.2006, No.90, pp.31-36 (2006)
- 7) freemidi.org, <http://www.free-midi.org/>
- 8) MIDI DataBase, <http://mididatabase.com/>
- 9) homemusician.net, <http://www.homemusician.net/>