

例による旋律の検索

牧口 智弥[†] 三浦 孝夫[†] 塩谷 勇^{††}

[†] 法政大学 工学部 情報電気電子工学科 〒184-8584 東京都小金井市梶野町 3-7-2

^{††} 産能大学 経営情報学部 〒259-1197 神奈川県伊勢原市上粕屋 1573

E-mail: [†]{c9843144,miurat}@k.hosei.ac.jp, ^{††}shioya@mi.sanno.ac.jp

あらまし 本稿では、キーボードやハミングで例示した単旋律から音楽データベース内の類似楽曲を検索する方式を提案し、実験により検証する。本研究では旋律は楽譜(シート音楽)で表現されているとし、小節毎にピッチスペクトルを比較することで、テンポ・表現の揺らぎ、移調などによらず効果的に検索できることを示す。

キーワード 音楽情報検索, 旋律類似度

Query-By-Melody to Music Information

Tomoya MAKIGUCHI[†], Takao MIURA[†], and Isamu SHIOYA^{††}

[†] Dept. of Elect. & Elect. Engr., HOSEI University 3-7-2, KajinoCho, Koganei, Tokyo, 184-8584 Japan

^{††} Department of Management and Information Science, SANNO University 1573, Kamikasuya, Isehara city, Kanagawa 259-1197 Japan

E-mail: [†]{c9843144,miurat}@k.hosei.ac.jp, ^{††}shioya@mi.sanno.ac.jp

Abstract We discuss how to look for *similar* melody in music databases by giving monophonic melody that human specify through MIDI-keyboard, microphone or any other devices. In this work, we utilize *text expression* (or sheet music) to describe music and examine *pitch spectrum* of melodies to obtain distinctive attributes. By this similarity definition, we concisely distinguish music from *tempo*, *swinging*, *transposition* or other arbitrary expressions. Also we show some experimental results.

Key words Music Information Retrieval, Melody Similarity

1. 前書き

本稿では、人間が与えた旋律の一部(単旋律)から楽曲を問い合わせる技法(音楽情報検索)を論じる。これらは、本来手書き文字認識や手書き絵による類似検索と同様に、内容検索を目的とし、曲名や作品番号といったメタ情報を仮定しない。この機能を用いた応用は広範囲に想定できる。例えば、音楽作業(作曲・編曲)に対して直接的な支援を行うことや、類似した曲を検索し著作権に抵触しそうな候補の絞込みを行うことができる。

他の情報検索分野と比較して、音楽分野では特に難しい点がある。まず、過去音楽情報をモデル化するのにデータモデリング機能が著しく欠けていることが挙げられる。例えば、順序概念(音符の並び、リズム)、長さ、速度(音符の時間軸配置)、音の高さ(絶対的あるいは移調による相対的)、表現(こぶし、ポルタメント)などを統一的に扱う試みはこれまで無い。第2に旋律質問の特定が難しい。実際、質問者の能力が十分ではないことから旋律が不正確であり、質問者の記憶があいまいであることから旋律が不明確であり、さらには質問者の思い込みにより旋律が作為的である。第3に、旋律の抽出方法が不明である

ことも重要である。通常、テーマ(主旋律)を検索することが中心になるが、和音、楽器・音色などが固定的でなく、またテーマであるといった記述表現もない。第4に、旋律の類似性の判定(何が答えか?)が明確に定義できず、解の表現もランキングによるのか、もっと直接的に楽譜や旋律演奏がよいのかが定まらない[13]。異なるキー(調性)に移調(ハ長調をト長調に)、音の補完・追加や除去(トリル、カデンツア)、調性の変更(転調、ハ長調をハ短調に)、単音を和音に変更あるいは音色を変更(編曲、異なる楽器で・複数の楽器で)、旋律を(同名調内で)変更(変奏曲など)、パートを除去あるいは追加(アカペラ化など)のどれが類似に相当するかは人間による判断が基本である。最後に、効率よいデータ管理・データ操作に関して、どのような技術が必要かが不明であることも重要である。新たなデータ構造化手法[8],[17]等を利用することなく、また情報検索(Nグラム)や探索エンジン(文字列データ処理)などの基盤を用いた報告がなされているが確定していない。

本研究では、例による旋律検索の精度を向上するため(楽曲および演奏の)楽譜から特徴量を抽出し、検索に利用することを提案する。即ち、小節ごとに音の高低(ピッチ)の分布状態を

スペクトル化し、特徴量として利用する。この方式の利点は、テンポ・表現の揺らぎ、移調などによらず効果的に検索できることにあり、以下で実験によりその有効性を検証する。

本稿では情報検索技術の基本知識を仮定する[5],[20]。第2節で、必要な定義と音楽情報の検索処理における問題点を要約する。第3節で、本研究で提案する特徴量を定義したあと、この特徴や利点を述べる。第4節では実験結果を示し、本手法の有効性を示す。第5節は関連研究を要約する。

2. 旋律検索のための音楽表現

本稿では音楽表現に対する内容探索方式を提案する。初めに幾つかの用語を定義する。ピッチ (pitch) とは音の調子、即ち音の高さを表す。以下では A,B,...,G の記号で表現される(注1)。ピッチを音 (tone) とも呼ぶ。音程 (interval) はピッチ間の距離であり、1 オクターブ離れた音程を12等分した単位を半音 (semitone) という。全音階的 (diatonic) な音の並びとは、A,B,C,D,E,F,G,a または C,D,E,F,G,A,B,c のように、半音がそれぞれ 2,1,2,2,1,2,2 または 2,2,1,2,2,2,1 ずつ離れた配置を言う。前者を短調 (minor) 後者を長調 (major) という。各ピッチには長さ (duration) がある。旋律 (melody) とは (長さを持つ) ピッチの順次配置である。和音 (chord) は複数のピッチの並列配置であり、これらは同時に演奏されることを想定している。伴奏が無い音楽を単旋律 (monophony) という。

五線譜 (five horizontal lines あるいは score) は、音の高さと長さを記入した記号 (音符) を順次的あるいは並列的に記述した表現である。旋律は音符の並びで、和音は音符の重なりで記述される。音符の長さの合計を拍 (beat) という。

音楽データベースに蓄えられたデータに対して楽曲を検索する場合、指定される検索条件は不完全である場合が多く、その大半は部分的な旋律である。例えば、ハミングなどにより単旋律を質問として与える (旋律質問と呼ぶ) 場合、実際の楽曲から大きく異なる可能性があるものとして、リズム・調性・音色・発想・速度・奏法・装飾音・強弱・揺らぎがある。これに対して、拍子や長短調の差異は、極端に異なるとは思えない。このため、旋律質問に対しては演奏状況を記述する表現 (演奏表現) を配慮せねばならない。表象的にはピッチ、音程、長さ、強弱、音色等が、また抽象的には拍子、リズム、小節、繰り返し、調性、音階的などの項目がある深層的には表情、解釈・揺らぎといったものも含める必要がある。

これらを計算機で表現するには、これまで信号レベル (CD 等でのアナログ的变化を表現)、操作レベル (MIDI など、輝度・時間のゆれを表現) 及び楽譜レベル (表現・解釈の多様性を表現) でなされてきた。過去に提案された音楽情報の検索処理のための特徴値記述では、旋律の輪郭情報 (Melody Contour) が代表的である[1]。また、近年は楽譜化技法として、XML に基づくものや五線譜利用による抽象化記述がある。旋律の輪郭情報のうち、ピッチ輪郭 (pitch contour) 表現は多くの研究で検討されている。ピッチ輪郭表現とは、単旋律にだけ有効であり、

直前の音と比べてピッチが相対的に高い、低い、同じという状態をそれぞれ文字 U,D,S で表し、輪郭を文字列として記述する(注2)。この表現は、ピッチに対して相対的なので移調に強いが、雑音に反応して文字列が変わりやすく、旋律検索は不完全文字列一致処理の実行を意味する。

[例1] 以下では、「カエルの合唱」(または「カエルの歌」)を統一例として考察する。五線譜は図1で示されるが、2段譜の右手パートに主旋律が生じている。このピッチ輪郭は-UUU DDD UUUU DDD DSSS SSUUUUU DDD である。あるいは、長さを見捨てることはピッチが同じかぎり意味をなさないと考えたことに相当するので、記号 S を取り去り、-UUU DDD UUUU DDD D UU DDD と表すこともある。□



図1 カエルの合唱五線譜

しかし、同時にピッチ輪郭方式は数多くの問題を含む[2],[3],[12]。即ち、多声音楽から各声を旋律と見なした輪郭に対しては組み合わせ的探索になること、拍子・音符長を含まないため判別には多くの情報を必要とすること、不完全一致検索手続きは多項式時間でよいが複雑であること、(全走査を除けば)2次記憶域上での処理効率を向上させる方法は知られていないことなどである。加えて、輪郭抽出には雑音除去のための「しきい値」の設定が容易ではない。

本稿では、旋律質問特性に対応するために、作曲者の意図の記述である「五線譜」あるいは演奏表現の情報を示す「楽譜記述」を用いる。特に音楽データベース情報は、特定されるべき楽曲であるから五線譜情報を有するほうが有利である。実際、これらは高度の抽象性を有し、厳格な指示 (五線譜上の記号による指示) に忠実に従っている。速度、装飾音、調性、長短調などは (ほぼ) 正確であり、音楽理論による知識を持つ (ソナタ形式の第2旋律は対立する属調で表現される(注3)等) ため、演奏表現上の情報は論理的な帰結と解釈できる (速度や強弱の変化)。一方、曖昧な指示 (文字列による指示) は五線譜上では表情記号で示され、音色・発想、揺らぎ等と共に、演奏者の自由な処理に委ねられている。無論、厳格な指示さえも曖昧 (ショパン夜想曲6番作品5-3では、革命的和声進行として4小節毎に半音階的転調を繰り返す[21]ため和声に比して演奏法は自明ではない) に扱われ、逆に指示の無い部分への指示に対する暗黙の仮定 (ウインナワルツの2拍目はやや長い) である経験知もある。信号レベル・操作指示レベルからこれらを抽出することは容易

(注2) : 開始文字は相対位置を定めることができないので - で表す。

(注3) : 第二主題は第一主題とは異なる調を持ち、主調が長調である場合は属調を、主調が短調である場合は平行長調を用いるのが通例である[19]。

(注1) : 以下では1オクターブ高いピッチを小文字で表す。

ではなく、楽曲の特性という観点からは、楽譜レベルでの”厳格な指示” 特性を活かしたほうが楽曲を特定しやすい。

単旋律による質問記述でも、楽譜による方法は有利である。実際、不完全な旋律 (不正確・欠落・旋律やリズムの近似、移調・長単調の変更、作為的な表情付け、こぶし、ずれ、ポルタメントといった揺らぎ) に対して、楽譜記述を用いれば、リズム・音色・発想・速度・奏法・強弱・拍子などの観点で問題を局所化する。反面、移調・装飾音・揺らぎという問題については、楽譜化では容易には解決できない。

本稿では、楽曲の五線譜記述・楽譜記述のため、abc 形式を利用する [16]。abc 楽譜ファイルはヘッダ部とボディ部からなる。例えば **M**: コマンド (拍子を示す), **L**: コマンド (音符記述の単位長), **K**: コマンド (調性) がある。ボディ部では実際の音楽記述を行う。記述は幾つかのフィールドが集まって小節をなし、最後は空白行で終わる。音符表現と休符表現が組み合わされて小節を構成する。ピッチは C, D, E, F, G, A, B, c, d, e, f, g, a, b で 2 オクターブ分のドレミで記述され、更に , と , を続けることでオクターブ上下を加え、都合 4 オクターブが記述される。臨時記号は ^ (sharp, 半音上), _ (flat, 半音下), =(natural, 原音) がある。休符は z, Z を用いる。ピッチや休符の長さは、単位長を 1 とした割合を整数倍または分数倍した値で表される。例えば z3 は 3 単位休符, Z4 は 4 小節休止, A/2B/4B/4 はラを 0.5, シを 0.25 ずつ 2 単位奏する指示を表す。

[例 2] 以下にカエルの合唱 (図 1) のふたつの abc 楽譜を示す。図 1 は 2 段五線譜であることから、V:1,V:2 の 2 部構成とし、2 声五線譜として表すことができる。

```
X:1
T:Kaeru
M:4/4
L:1/4
Q:1/4=120
K:C % 0 sharps
V:1
cdef|edcz|efga|gfez|zcz|zcz|
c/2c/2d/2d/2e/2e/2f/2f/2|edcz|z4||
V:2
z4|cdef|edcz|efga|gfez|zcz|zcz|
c/2c/2d/2d/2e/2e/2f/2f/2|edcz||
```

同じ曲を 和音を用いた 1 声五線譜で記述することもできる。

```
X:1
T:Kaeru
M:4/4
L:1/4
Q:1/4=120
K:C % 0 sharps
V:1
cdef|[ec]d[ce][zf]|e[fd][gc][az]|
[ge]f[eg][za]| [cg][zf][ce]z|zcz||
```

```
c/2c/2d/2d/2e/2c/2-|[e/2c/2]f/2f/2|
[e/2-c/2][e/2c/2]d/2d/2[c/2-e/2]
[c/2e/2][z/2f/2][z/2f/2]|edcz||
```

□

3. 音楽情報の検索方式

音楽情報を検索するためには、各楽曲の旋律と質問旋律の関連を評価する尺度が必要である。本研究では旋律から特徴量を抽出し、ベクタ空間モデル [5], [20] を用いて特徴項目ベクトル空間へ写像する。各次元は項目に対応するため、どの旋律も空間内の点で表せる。旋律質問の解はベクトル空間内の類似した点に対するランクつけで得られる。空間内では伝統的な余弦空間で定義する。即ち $\cos \theta = \frac{X \cdot Y}{|X| |Y|}$ を計算し、1.0 に近いほど類似していると判断する。ここで X, Y は空間ベクトルを、 $X \cdot Y$ は内積を、 $|X|$ はベクトル長を表す。

3.1 ピッチスペクトル

旋律が五線譜・楽譜で記述される時、本稿では、小節単位にピッチスペクトル (pitch spectrum) を導入し特徴量とする。ピッチスペクトルとは、小節内に生じる各音符を長さで集積しそのヒストグラムをとるもので、以下では 4 オクターブ $12 \times 4 = 48$ 音を対象に、48 次元ベクトルを生成する^(注 4)。小節ごとにピッチスペクトルを算出し、これを対象楽曲の規定小節数だけ用意する。実験では 16 小節分 $8 \times 48 \times 16$ バイト、約 7KB で表現する。このアイデアは信号処理では提案されているが、楽譜レベルで特徴量化するというアプローチはない [18]。[例 3] カエルの合唱 (2 声) から抽出した単旋律 (第 1 声) に対する 16 小節分のピッチスペクトル (図 2) を示す。

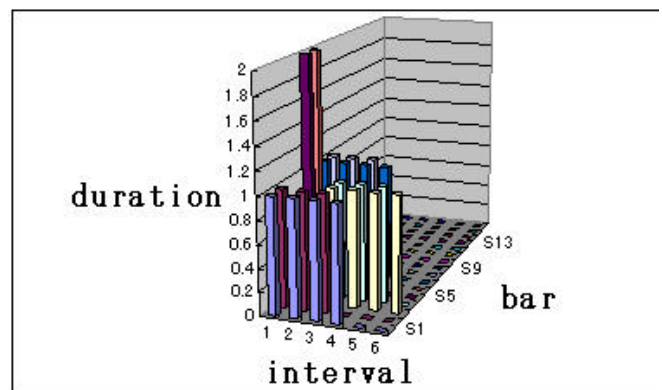


図 2 カエルの合唱ピッチスペクトル

同様に、和音を用いた 1 声楽譜から単旋律を 16 小節抽出した。ここでは、和音が生じる都度、第 1 音だけを旋律と見なし単旋律化した。次のようになる。

```
cdef|edcz|efga|gfez|zcz|zcz|
c/2c/2d/2d/2e/2e/2f/2f/2|
e/2-e/2d/2d/2c/2-c/2z/2z/2|edcz||
```

両者にはほとんど差が見られない^(注 5)。

(注 4) : 倍精度実数で表せば $8 \times 48 = 384$ バイト長となる。

(注 5) : 後述する類似度は 0.920 となる。

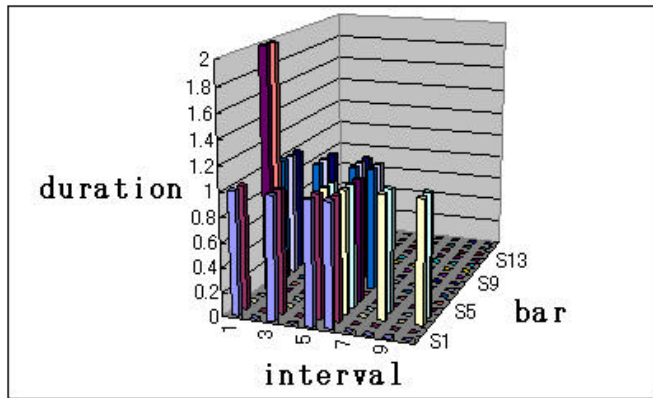


図3 和音カエルのピッチスペクトル

□

ピッチスペクトルは各ピッチを長さで合計したヒストグラムであり、(ギターコードのような)小節全体に和音を指定したものとは異なる。ピッチ輪郭とは異なって、小節単位の特徴を集約しているため、旋律の輪郭を表してはいない。むしろ、小節間の変化を記述しているため、単旋律の時系列表現とも異なる。代わりに、ピッチ輪郭は不完全な旋律に生じる問題を解消することができる。実際、装飾音・旋律・長短調・揺らぎがピッチスペクトルで局所化する。ただ、調性の違いについては難しく、本研究では移調処理を考察せざるを得ない。

3.2 楽曲データベースの作成

データベースに格納される楽曲ごとに、楽曲情報と同時にピッチスペクトルを特徴量(特徴ベクトル)として格納する。本稿では、 \langle 楽曲ID,ベクトル長,特徴ベクトル \rangle を単位として処理する。楽曲から旋律を抽出することは難しい問題であり、発見的に、1トラックにまとめ、最高音を *modulo 12*(オクターブ音を無視)する手法 *all-mono* が有効とされる[12]。

本稿ではピアノ曲に限定して実験を行う。ピアノ曲では通常主旋律を奏するのは右手である。このため複数楽器を前提とした *all-mono* 手法にとられる必要はなく、右手に対応した第1声に限定し、和声があった場合は(高音とは小指や薬指になるので限定しにくいことから)和声記述の先頭音を旋律音と見なす発見的な方法を取る。

3.3 旋律質問の記述

旋律質問の記述は楽譜でなされ、単旋律で指示されるものとする。各小節をピッチスペクトルで記述してベクトル列で表現する。後述するように、この質問を移調(調性の自動検出と音程シフトの自動化)を行ったあと、質問ベクトルと各楽曲との類似性を余弦値計算により求める。類似度は、質問と楽曲の対応する小節ごとにピッチスペクトルの余弦を求める。質問者は一般的に、旋律の最初ほど正確に記憶しており、あとになるほど精度が悪くなると考えられることから、十分に小さい忘却率 $\rho \geq 0$ に対して、余弦値 c_k に $(1 - k \times \rho)$ 倍した値を k 番目の小節の類似度とする。全類似度の平均を、2つの旋律の類似度(similarity)と定義する:

$$\text{類似度} = \frac{\sum_{k=1..n} (1 - k \times \rho) c_k}{\sum_{k=1..n} (1 - k \times \rho)}$$

むろん、1.0に近いほど類似している(注6)。

3.4 移調判定処理

移調を行うには、調性を見出し、実際の移動量を計算するという2段階の操作が必要である。例えば、同じハ長調であってもオクターブ移動がある。また、ハ長調からヘ長調へは-7半音移動や+5半音移動、およびオクターブ移動との組み合わせがある。以下では、これらの手順を要約する。

旋律の調性(キー)を見出すことは容易ではない[12]。本研究では次のような前提の下で、移調の発見的方法を採用する。

- (1) 本来の楽曲は全音階的である
- (2) 曲の途中で調性は変わらない
- (3) 調性は臨時記号が7つまで、都合15種類の調性が検査される

先ず4オクターブの48の半音を(*modulo 12*により)12個に集約(同一視してその長さを合計)し、12音だけのピッチスペクトルを計算する。臨時記号が定期的に発生した音を調べ、調性と最も対応関係(キー移動量)のよいものを選ぶ。例えば、ヘ長調(F)ではBよりもBbの方が多く生じ、これ以外は定期的変化はない(注7)。以下の表は、+1(半音上昇, sharp)あるいは-1(半音下降, flat)の変化と調性の関係を示す。例えばヘ長調ではBにだけ-1が対応し、これ以外に変化がない(0で表現)。

臨時記号数	C	D	E	F	G	A	B	調性
0	0	0	0	0	0	0	0	C
+1	0	0	0	+1	0	0	0	G
+2	+1	0	0	+1	0	0	0	D
+3	+1	0	0	+1	+1	0	0	A
+4	+1	+1	0	+1	+1	0	0	E
+5	+1	+1	0	+1	+1	+1	0	B
+6	+1	+1	+1	+1	+1	+1	0	F
+7	+1	+1	+1	+1	+1	+1	+1	Csharp
-1	0	0	0	0	0	0	-1	F
-2	0	0	-1	0	0	0	-1	Bb
-3	0	0	-1	0	0	-1	-1	Eb
-4	0	-1	-1	0	0	-1	-1	Ab
-5	0	-1	-1	0	-1	-1	-1	Db
-6	-1	-1	-1	0	-1	-1	-1	Gb
-7	-1	-1	-1	-1	-1	-1	-1	Cb

質問旋律と候補旋律のそれぞれの実際の移動量を算出するため、移動方向および移動幅を算出する必要がある。方向については、A, A#, ..., G# の各ピッチに数値1, 2, ..., 12を対応させ、各小節の特徴ベクトルの代表ピッチを次のようにして求める:

$$\frac{\sum_{12 \text{ 音の各音}} \text{音の対応値} \times \text{音の長さ}}{\text{音の長さの総和}}$$

質問旋律と候補旋律のそれぞれに、全ての小節について代表ピッチを計算し平均値を求め両者の差を計算する。これが移動方向(差を小さくする方向に移動)を決める。移動幅の計算も、両者の差の値を小さくするように決める。即ち、音楽理論に従い、調性分(sharp記号あたり+7半音, flat記号あたり-7半音)の移動量のうちオクターブ移動量を同一視(つまり移動半音量を *modulo 12*)して, sharp記号あたり+7/-5半音(あるいはflat記号あたり-7/+5半音)の移動で差が小さくなる方を採用する[19]。

(注6): 実験では $\rho = 0.005$ とした。

(注7): 長短調の判断は行わない。

ここで示した移調処理は、音楽理論に基づく体系化に従うため、当該半音変化が相当教生じ、しかも長短調の変化がない場合には極めて有効である。しかし、(真の) 臨時変化が調性決定に影響することを阻止できない。長短調の差がある場合でも、平行調 (Relative Key ハ長調とイ単調のように同じ調性の長短調) ならば本方式で対応できる。しかし同主調 (Parallel Key, ハ長調とハ短調のように主音が同じ長短調) で 2 つの旋律の比較が生じるとうまくいかない。その旋律は音程レベルでは E で半音だけ異なるにもかかわらず、ハ長調とハ短調 (flat 3 つで示される調性) は、本方式を適用する結果、ハ長調と変ホ長調 (flat 3 つの長調) の関係に移調され、全てのピッチが +3 半音移動することになり、類似度を低下させる。

4. 実験と評価

この節では、本稿で提案する方式の有効性を示すため、演奏情報をもとに類似旋律の検索を実験する。

精密に考えると”類似検索”には少なくとも 3 種類あると考えられる。類似旋律の検索、編曲・変奏の検索および演奏情報による旋律の検索である。類似旋律の検索とは、与えた旋律に類似した旋律を音楽データベースから検索することを意味する。本来独立した楽曲・旋律であるはずにもかかわらず、このような類似性が得られる論拠として、作曲者の性質・傾向のほか、時代的背景・流行、あるいは違法コピーなどがある。このような検索は、カラオケ楽曲検索や著作権検証などを目的としている。ただ、評価方法を定めることは容易ではない。

編曲・変奏の検索とは、各楽曲が与えた旋律にどの程度独立しているかを検証する狙いがある。作者が同一モチーフを多様に変化させたいとき、同一の旋律を意図的に独立のものに見せたい場合に、その度合いを検証することができる。

演奏情報による旋律検索とは、先と同様に旋律検索を行う。しかしこの場合、質問は演奏情報で与えられるため、良く似た旋律を解とはするものの厳格に一致することはない。言い換えると、旋律はもとの曲と異なるがはるかに離れたものではない。

ここでは、多くの候補曲を与え、解候補となる曲がこれらとどの程度分離できるかを調べる。本稿の狙いは、このような機能要求にピッチスペクトル特徴量が対応できるかを評価することにある。

以下の実験では、まず音楽データベースを作成する。本実験では、ピアノ演奏情報 (五線譜ではない) を表現する MIDI データを利用する^(注 8)。このファイルを abc 記述に変換したあとで、先頭から単旋律 16 小節を抽出し特徴ベクトルを計算する。既に述べたように、第 1 声、和音の第 1 音にのみ着目する得られた単旋律から、特徴量 16 小節を計算し、<楽曲名、ベクトル長、特徴ベクトル> のような項目を対応させる。同様に、質問旋律の計算についても特徴量 8 小節を計算しておく。質問旋律に対して、音楽データベースの各楽曲毎に類似度計算を行う。類似性計算は、移調処理・精度悪化の想定を加えた計算を行い、最終的にランキングする。

これらの処理は Windows 上で開発・実施する。実験に当たっては、”abc” 関連フリーソフト (abc2abc など) を利用・修正したほか、特徴ベクトルの生成と類似度計算のために、新たに C++ で 1000 行程度のソフトウェア (abc2vec, vecmatch) を開発した。

演奏された旋律をデータベースから検索するために、以下では 2 種類のピアノ曲を用意する。J.S.Bach (1685-1750) は、同じ考え方のもとに 2 種類の曲集を作曲した。これらはそれぞれ 2 声のインベンション (*Two Voice Invention*, BWV 772-786) と 3 声のインベンション (あるいはシンフォニア, BWV 787-801) と呼ばれる。共に 15 曲の小曲 (それぞれは数分程度) からなる。質問する演奏曲はこのうち 2 声のインベンション第 1 番 (BWV 772) であり、図 4 にその五線譜を示す。一方、図 5 はこの曲の演奏表現 (を五線譜表現) を BachQ の例を含む。



図 4 BWV 772: Two Voice Invention No.1



図 5 BWV 772: Performance Information

ここでは BWV772 の 3 つの演奏例 (BachQ, BachP1, BachP2 と 30 曲の対象曲 (Inv2-xx, Inv3-xx)) に対して、提案する検索を実験した。以下の表でその結果を示す。また、図 6 はこの結果のトップ 10 である。検索結果の平均は 0.2293 であり、先頭 4 曲は検索曲とほぼ一致して BWV 772 であることを示している^(注 9)。また、ふたつの作品 Inv 2-13, Inv 3-08 だけが 0.3 以上の類似度であることがわかった。

(注 8) : <http://www.classicalarchives.com/>

(注 9) : 実は BachP2 と Inv2-01 は同一であることをこの実験で知った。

Description	Similarity	Trans.	Judgment
BachQ.abc	0.9825	0	C
BachP1.abc	0.921172	0	C
BachP2.abc	0.893104	0	C
Inv2-01.abc	0.893104	0	C
Inv2-13.abc	0.39747	0	C
Inv3-08.abc	0.328203	-2	Bb
Inv3-10.abc	0.290064	-5	G
Inv2-04.abc	0.274021	0	C
Inv3-01.abc	0.263477	0	C
Inv3-13.abc	0.239543	-5	G
Inv3-05.abc	0.212071	-2	Bb
Inv2-10.abc	0.210963	-5	G
Inv3-04.abc	0.162852	-12	C
Inv3-06.abc	0.162085	4	E
Inv2-06.abc	0.147679	4	E
Inv2-09.abc	0.141909	-4	Ab
Inv3-03.abc	0.140305	-5	G
Inv2-07.abc	0.130268	2	D
Inv3-07.abc	0.122077	-10	D
Inv3-14.abc	0.104519	-2	Bb
Inv3-11.abc	0.09103	-2	Bb
Inv2-15.abc	0.090888	2	D
Inv3-12.abc	0.083129	-8	E
Inv2-03.abc	0.078591	-3	A
Inv3-15.abc	0.071857	2	D
Inv2-02.abc	0.057013	-9	Eb
Inv2-12.abc	0.049853	-3	A
Inv2-05.abc	0.028571	3	Eb
Inv2-08.abc	0	12	C
Inv2-11.abc	0	12	C
Inv2-14.abc	0	12	C
Inv3-02.abc	0	-9	Eb
Inv3-09.abc	0	12	C

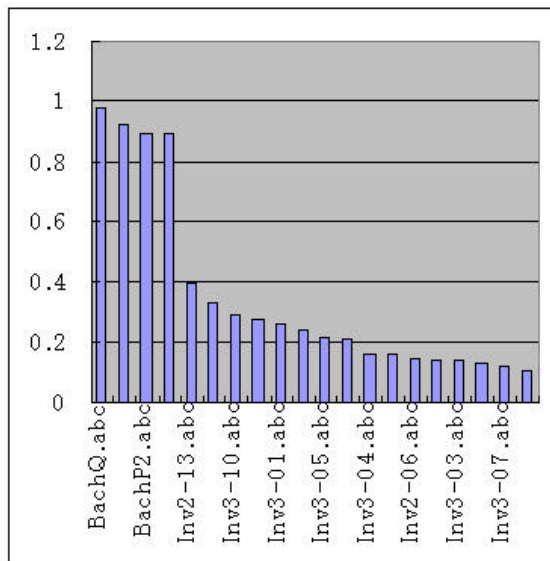


図6 Similarity: Results

この結果からピッチスペクトルがどれほど有効に作用しているかを知ることができる。即ち演奏情報から抽出された旋律をもとに原曲を特定するには十分である。実際、類似度トップ4曲は高い類似性を有した正しい結果であり、類似した曲があってもこれに比べれば小さく、類似でしかないことがわかる。ピッチスペクトルは、実際、効果的な手法であり旋律を他と区別するに十分な機能を有することがわかった。

5. 関連研究

当該分野の研究は60年代から開始され、マルチメディア処理技術(信号処理からメタ情報による検索まで)や情報検索理論との関連(Nグラム、ベクトル空間モデル、検索エンジン等)が

論じられてきた。現在、研究の流れは音楽情報のモデル化、旋律の自動抽出、音楽情報検索の3点に要約できる[12]。音楽情報のモデル化の狙いは、これまでの単旋律表現に加え、多声音楽やリズム記述を目的としたものであり今後大きな発展が予想される[2],[10]。旋律の自動抽出は、デジタル音楽には欠かせない技術である。インターネットに数多く存在する楽曲ファイルは、多重トラック・多重チャンネル(複数楽器)、和声、リズム(MIDIを想定)など多様な表現を有しているため、旋律(特に主旋律)の抽出は重要な技術を提供する。これまで発見的に、1トラックにまとめ、最高音を *modulo 12* (オクターブ音を同一視)する手法が知られている[12]。あるいは、“重要な旋律(テーマ)は繰り返し現われる”という経験的規則からテーマを抽出する手法もある[7]。音楽情報検索分野では、ハミング検索(Query By Humming)が注目されている[4],[6]。すなわち、人間が与えた不正確なハミング情報、移調、部分的旋律などだけで検索する技法が、どのような特徴値表現を採用することで可能になるかがここでの焦点である。信号処理[11]、旋律の輪郭情報[1]による文字列変換などの他、確率論的アプローチも報告されている[9],[10]。

6. 結 び

本稿では、新しい楽曲の検索方式を提案した。本アプローチでは楽曲を楽譜で記述する。このことにより、音楽理論的領域知識の利用や旋律質問のもつ不完全・不正確な表現から厳格な指示を抽出しやすくなる。楽譜から単旋律を抽出してピッチスペクトルによる特徴量化を行い、旋律質問時に移調処理を加えることにより、テンポや表情などの表現や奏法による差異を吸収しやすくなる。本稿では、これらを類似度ランキングという方式で評価した。

実験による評価で、本アプローチが極めて効果的であることを示した。演奏情報から抽出した旋律をもとに検索を行い、本来同じはずの曲が、十分な類似性を持って判定できることを示した。

本稿では実験できなかったが、実は奏法、移調、テンポ、奏法などが異なっても評価でき、高精度で分離することができる。

我々の実験では先に述べた対応が問題として残る。さらに、多声音楽への展開についても同様の手法を適用できるものと考えている。

謝辞

本研究の一部は文部科学省科学研究費補助金(課題番号14580392)の支援による。

文 献

- [1] W.J.Dowling: Scale and Contour – two components of a theory of memory for melodies, *Psychological Reviews* 85-4, 1978, pp.341-354
- [2] M.Droettboom et al.: An Approach Towards A Polyphonic Music Retrieval System, *Intn'l Symp. on Music Information Retrieval (ISMIR)*, 2001
- [3] M.Droettboom et al.: Expressive and Efficient Retrieval of Symbolic Musical Data, *Intn'l Symp. on Music Information Retrieval (ISMIR)*, 2002

- [4] A.Ghais et al.: Query By Humming, *ACM Multimedia Conf.*, 1995
- [5] D.Grossman, O.Frieder: Information Retrieval – Algorithms and Heuristics, Kluwer Academic Press, 1998
- [6] T.Kageyama et al.: Melody Retrieval With Humming, *ACM Intn'l Computer Music Conf. (ICMC)*, 1993
- [7] C.Meek et al.: Thematic Extractor, *Intn'l Symp. on Music Information Retrieval (ISMIR)*, 2001
- [8] T.Miura, W.Matsumoto, I.Shioya, and Y.Wada: Extensible Perfect Hashing, *ACM Conf. on Information and Knowledge Management(CIKM)*, 2000
- [9] J.Pickens: A Comparison of Language Modelling and Probabilistic Text Information Retrieval Approaches to Monophonic Music Retrieval, *Intn'l Symp. on Music Information Retrieval (ISMIR)*, 2000
- [10] J.Pickens, T.Crawford: Harmonic Models for Polyphonic Music Retrieval, *ACM Conf. on Information and Knowledge Management(CIKM)*, 2002
- [11] T. Sonoda, et al.: A WWW-based Melody Retrieval System, *Intn'l Computer Music Conf. (ICMC)*, 1998
- [12] A.L.Uitdenbogerd et al.: Manipulation of Music For Melody Matching, *ACM MultiMedia Conf.*, 1998
- [13] A.L.Uitdenbogerd et al.: Matching Techniques for Large Music Databases, *ACM Multimedia Conf.*, 1999
- [14] A.L.Uitdenbogerd et al.: MUSIC IR: Past, Present and Future, *Intn'l Symp. on Music Information Retrieval (ISMIR)*, 2000
- [15] A.L.Uitdenbogerd et al.: Music Ranking Techniques Evaluated, *Intn'l Symp. on Music Information Retrieval (ISMIR)*, 2000
- [16] C.Walshaw: abc Version 1.6,
<http://www.gre.ac.uk/~c.walshaw/abc2mtex/abc.txt>
- [17] S.Watanabe, and T.Miura: Reordering B-tree, *ACM Symp. on Applied Computing(SAC)*, 2002
- [18] C.Yang: Music Database Retrieval Based on Spectral Similarity, *Intn'l Symp. on Music Information Retrieval (ISMIR)*, 2001
- [19] 石桁他：新装楽典－理論と実習，音楽之友社，2001
- [20] 北研二他：情報検索のアルゴリズム，共立出版，2002
- [21] 吉田秀和：シヨパンに関する覚書，吉田秀和全集 13，白水社，1979