

1 はじめに

筆者らは、音楽情景分析の処理モデル OPTIMA (Organized Processing Toward Intelligent Music-scene Analysis) を提案し、その実験システムを実装している [1]。OPTIMA では、一部の単音が取得できないという問題があったが、楽曲において旋律やリズムについて特徴的な部分を検出し、その情報に基づいて他の認識確度の低い部分を補うことが本研究の目的である。

OPTIMA では時間-周波数解析の結果から単音の抽出を行っているが、ここでは開始時刻と終了時刻が明確な単音から楽曲が構成されるものとする。つまり五線譜で表現するのに十分な情報があるとする。認識確度の低い部分を補うということの可能性の検証のため、単音のいくつかを取りこぼすというモデルも設けてみた。

2 楽曲の抽象化

まず、楽曲の解析のために抽象化を行う。楽曲を構成する音符の時間的長さの最小単位を決定する。対象とする曲は、この最小単位 n 個分の長さであるとする。すると、各最小単位ごとに音程が一意に決定される。また、最小単位の前で同じ音が切れずに継続するか、あるいはそうでないか、どちらかであるということが出来る。

音程を表す何らかの値 (例えば MIDI のノート番号) を導入すると、旋律は m 次元のベクトルで表現できる。例えば、図 1 の譜例は最小単位を 8 分音符とすると、

(64, 64, 76, 74, 78, 76, 74, 72, 71, 71, 71, 74, 67, 67, 67, 67, 64, 64, 76, 74, 78, 76, 74, 78, 79, 79, 78, 78, 76, 76, 76, 76)

と表現できる。

また、最小単位とその直後の最小単位とで音が途切れる場合を 1、継続する場合を 0 と表現することで、リズムは $m+1$ 次元のベクトルで表現できる。図 1 の譜例の区切りの情報は、

(1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1)

と表現できる。

* "Music scene analysis perceiving repetition in a musical piece"
Ibuki Handa, Tomoyoshi Kinoshita, Makoto Muto, Shuichi Sakai, Hidehiko Tanaka
Graduate School of Engineering, the University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan



図 1: 譜例

3 拍子の推定

楽曲が何拍子かを推定することを考える。最小単位 m 個で小節が構成されると仮定し (これを本稿では仮小節と呼ぶことにする)、その仮定が妥当か否かを評価することで最も適当と思われる拍子を決定するものとする。

仮小節が m 個の最小単位で構成されるとすると、楽曲全体は

$$l = \lceil n/m \rceil \quad (1)$$

個の仮小節に分解できる。任意の 2 つの仮小節 M_1, M_2 に対して、以下の相違度を定義する。

$$d_R(M_1, M_2) = \sum_{i=0}^m |R(M_1, i) - R(M_2, i)| \quad (2)$$

但し、 $R(M, i)$ は仮小節 M の i 番目と $i+1$ 番目の最小単位の区切りの情報 (1 または 0) とする。

一般に楽曲は小節を単位として同一な、あるいは酷似したリズムや旋律が現れる。従って l 個の仮小節のうち 2 つを取り出した相違度の、全ての組合せについての総和を評価量とすると、その仮小節の区切り方が妥当であれば小さな値となる。しかし、仮小節の区切り方が異なると組合せの総数や 1 つの仮小節を構成する最小単位の数も異なってしまう。そこで、これらに依存しないように正規化した (3) を評価量とする。

$$D = \frac{2}{l(l-1)(m+1)} \sum_{i=1}^{l-1} \sum_{j=i+1}^l d_R(M_i, M_j) \quad (3)$$

正解が 4/4 拍子で 36 小節からなる楽曲に対して解析を行った結果を図 2 に示す。最小単位を 8 分音符としたが、 D が小さくなる仮小節の設け方は最小単位が 8 の倍数のときで、正しい小節の区切り方と合致した結果を得た。

次に、単音の幾つかを取りこぼす場合をモデル化して、ランダムに休譜に代えてリズム解析を行ってみた。対象は先と同じ楽曲で、30% の確率で取りこぼすことにした。図 3 に示す結果から、取りこぼしがあっても拍子の推定は可能と判断してよい。

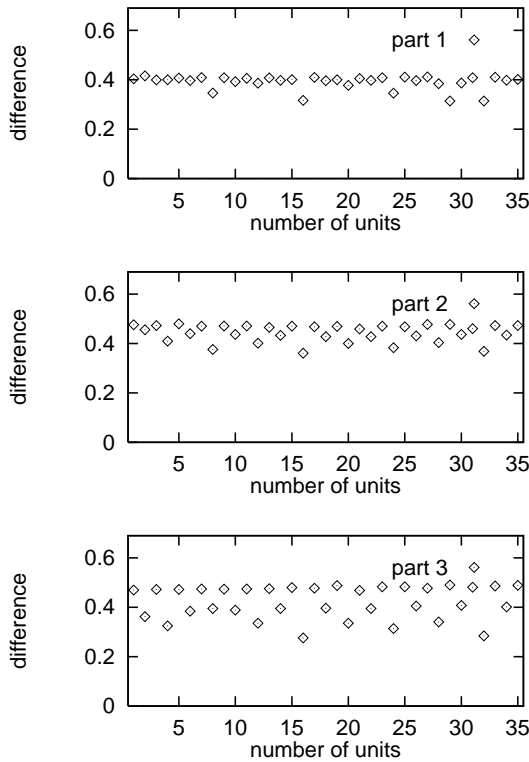


図 2: 4/4 拍子の楽曲のリズム解析の結果

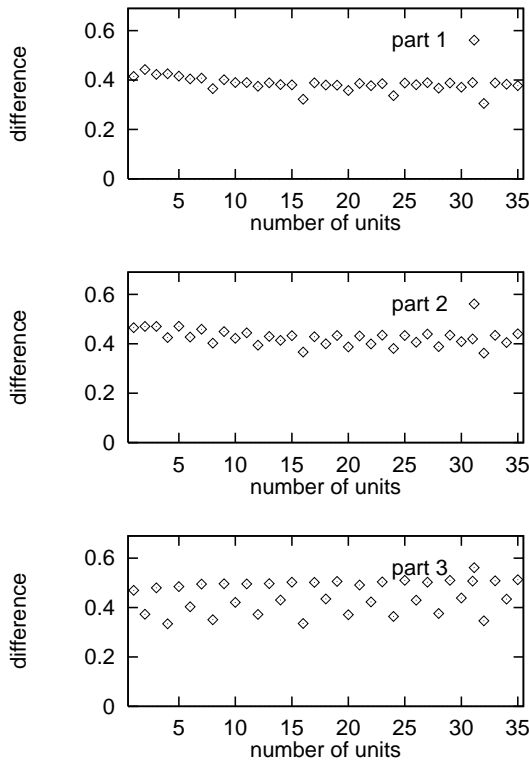


図 3: 30%のとりこぼしがある場合のリズム解析の結果

4 小節同士の旋律の類似性の評価

小節の区切り方が決定された上で、今度は似ている小節を見分ける方法を検討して見る。旋律の相違度の評価量として、(4) を設ける。

$$d_T(M_1, M_2) = \sum_{i=1}^m |T(M_1, i) - T(M_2, i) + \alpha| \quad (4)$$

但し、 $T(M, i)$ は小節 M の i 番目の最小単位の音程である。この式で α の値を変えると d_T の値も変わるが、ある 2 つの小節に対して最小の d_T を与える α の値とそのときの d_T の値の 2 つを組として評価量とする。平行移動することも考慮に入れているのである。図 1 の譜例に対する評価結果を表 1 に示す。第 1 小節と第 3 小節がよく似ていると認識されたことが結果に現われている。

表 1: 旋律の類似性の評価

比較する小節の組	平行移動度	相違度
(1,2)	5	41
(1,3)	0	6
(1,4)	2	120
(2,3)	5	47
(2,4)	8	95
(3,4)	2	126

5 欠落情報の補完

第 4 節での議論では、全ての単音が完全に把握されているものとしていたが、リズムの場合と同様、取りこぼしがあっても旋律の類似性を評価できる方法を作り、はっきり認識された小節から他の似ていると思われる小節を補完することを考えている。補完のためには、認識できた単音の他に、時間-周波数解析の結果および単音の遷移の確率の統計データなどが必要となる。

6 おわりに

一つの楽曲内のリズムと旋律の特徴的な部分を取り出す方法を提案し、その効果の評価を行った。

今後は時間-周波数解析の結果からのリズム、旋律の特徴抽出と、不確かな情報の補完を行う機構を構築したい。

参考文献

- [1] 木下智義, 村岡秀哉, 田中英彦: 単音の遷移に注目した単音認識処理, 日本音響学会誌, Vol.54, No. 3, pp. 190-198, March 1998