

音楽の要素構成構造に着目した曲断片のモーフィング

武藤 誠, 半田 伊吹, 日比 啓文, 坂井 修一, 田中 英彦

東京大学大学院工学系研究科
〒113-8656 東京都文京区本郷 7-3-1

{muto,handa,hibi,sakai,tanaka}@mtl.t.u-tokyo.ac.jp

あらまし：モーフィングは画像、音声、楽音を対象とする手法が提案されているが、楽曲を対象とした手法は提案されていない。本稿では音楽が音色、旋律、リズム、和声の各構成要素から成り立つことに着目して、旋律パート、伴奏パート、リズムパートの3パート構成の1小節の楽曲のモーフィング手法を提案する。各音楽構成要素について主観的な類似度に関するアンケートを行い、多次元尺度構成法とニューラルネットワークによって主観的類似度マップを得た。そして、それぞれの類似度マップ上でモーフィングを行った結果を再合成することによって、モーフィングされた楽曲を得た。実験によって得られた楽曲のモーフィング結果は、概ね満足のゆくものであった。

キーワード: 楽曲モーフィング, 主観的類似度, 多次元尺度構成法, ニューラルネットワーク, 楽曲合成

Musical Morphing by Parallel Processing of Musical Components

Makoto MUTO, Ibuki HANDA, Hirofumi HIBI, Shuichi SAKAI and Hidehiko TANAKA

The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656

{muto,handa,hibi,sakai,tanaka}@mtl.t.u-tokyo.ac.jp

abstract : Many morphing methods for image, speech and musical tone have been proposed. But no morphing method for musical song have been proposed. In this paper, we propose a morphing method for musical song (one measure, 3 tracks) by parallel processing of musical components - timbre, contour, rhythm and harmony. We analyzed questionnaire data about the similarity of components by multidimensional scaling and Neural Networks, and got psychological similarity map of each musical components. We got musical morphing by re-synthesizing morphing results on the similarity map of each musical components. By experiments, we succeeded to get satisfactory morphing results.

keywords : Musical Morphing, Psychological Similarity, Multidimensional Scaling, Neural Networks, Song Synthesis

1 はじめに

モーフィングとは一組の加工対象に対してそれらの間の滑らかな遷移を自動的に生成する技術である。モーフィングは主に画像、音声、楽音を対象とする研究が行われており、実際に映像における特殊効果などに応用されている。しかし楽曲を対象としたモーフィング手法は提案されていない。

本稿では、音楽が音色、旋律、リズム、和声の各構成要素から成り立つことに着目して、各構成要素の並列的モーフィング処理と再合成による手法を提案する。加工対象とする楽曲は、旋律パート、伴奏パート、リズムパートの3パート構成で、コード進行の無い1小節のものとする。モーフィング処理の手続きは、(1)加工対象曲の音楽構成要素への分解処理、(2)各音楽構成要素について主観的な類似度マップ上の布置の計算、(3)各類似度マップ上でのモーフィング処理、(4)各音楽構成要素のモーフィング処理結果の再合成処理の各処理から成る。ここで、主観的類似度マップは類似度に関するアンケートを多次元尺度構成法とニューラルネットワークで分析することによってあらかじめ知識ベースとして用意する。

評価実験ではポップスや演歌などのサンプル曲100曲を用いてモーフィング処理を行ない、楽曲のモーフィング結果を得た。その結果、遷移が滑らかで音楽性の面でも納得のゆく曲が得られた。

第2章では、楽曲のモーフィングについて目的と背景を述べる。第3章では、本稿が扱う音楽の構成要素について述べる。第4章では、本稿が提案するモーフィング手法について述べる。第5章では、本手法の評価実験について結果と考察などを述べる。最後に第6章で本稿のまとめを述べる。

2 楽曲のモーフィング

2.1 本研究の目的

図1のように、2つの曲断片を元に両者の間の遷移を主観的¹に滑らかなものとする曲断片群を自動的に生成する技術を楽曲のモーフィングとする。与えられた2つの曲断片に対して、図1のように生成された曲断片群を時間的に挿入して、連続して聴いた場合に主観的に滑らかな遷移の印象を受けるものとなることを目的とする。

応用技術としては、作曲や編曲における特殊効果としての利用や、自動作曲や自動編曲の要素技術としての利用などが考えられる。

本研究で加工対象とする楽曲の形式は、図2のような表1の形式の楽曲とする。

¹本研究では主観的という言葉、物理的に対し心理的という意味と、客観的に対し主観的という意味の両方で用いる。

2.2 本研究のアプローチ

モーフィングは画像や音声を対象とした研究[1][2]が行われている。これらの研究では、加工対象とする一組の画像等が共通の物体等を含むことを想定し、これら共通物の間の対応付け処理を行い、物理的に対象間の補完を合成するという手法を用いている。しかし、音楽は明示的な意味を持たないため、加工対象間の対応付けは難しい。また、加工対象は任意であることが応用上の自由度が大きいといえる。

そこで本研究では、加工対象間の対応付けは行わず、主観的な類似度マップを用いて知識ベースから滑らかな遷移を実現する要素を選択するという手法を提案する。また、音楽を構成する要素の知覚上のモジュール性[3]を考慮して、各音楽構成要素に対して並列的なモーフィング処理と再合成による手法を提案する。

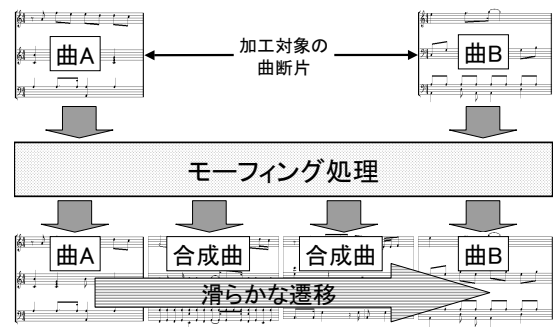


図1: 楽曲のモーフィング



図2: 本研究が加工対象とする楽曲の例 (リズムパートはドラム譜)

表1: 本研究が加工対象とする楽曲の形態

パート構成	旋律パート、伴奏パート、リズムパートの3パート構成
小節数	1小節
拍子	4分の4拍子
テンポ	120[四分音符 / 分]
リズム	8分音符の長さで量子化されたもの
ペロシティ	一定
コード	コードの変化はなし。伴奏パートの音符は全てコードの構成音の音名

3 音楽の構成要素

3.1 楽曲を構成するパート

音楽は、一般的に複数のパートで構成されている。それぞれのパートは、同一の演奏者によって同一の楽器によって演奏され一つの知覚上の流れを形作り、音楽の構成要素と考えることができる。

楽曲におけるそれぞれのパートの演奏には様々な形態があるが、本研究では旋律パート、伴奏パート、リズムパートの3種類に分類する。

3.1.1 旋律パート

楽曲における、メロディを担当するパートを旋律パートとする。主に単旋律²が用いられるので、本研究では単旋律のみを想定する。

3.1.2 伴奏パート

楽曲における、伴奏を担当するパートを伴奏パートとする。伴奏は一般的に時間的に一定の繰り返しのパターンとなっていることが多い。

3.1.3 リズムパート

楽曲における、打楽器によって演奏されるパートをリズムパートとする。現代の商業音楽の大多数の楽曲がリズムパートを持つ。

3.2 パートを構成する音色、旋律、リズム、和声

音楽には表 2 のような階層構造がある [4]。本研究で想定する曲は 1 小節の曲であるので、表 2 の C 「音楽構造」と D 「意味、内容」は考慮する必要はない。また、A 「音響信号」の属性の音高、長さ、大きさは楽曲の文脈で知覚される場合は、B 「知覚対象」の属性の旋律、リズムの一部として知覚されるので、旋律、リズムとしてまとめて扱うことができる。そこで、本研究では音楽の持つ属性として図 3 のように、音色、旋律、リズム、和声の 4 種類を扱う。

表 2: 音楽の階層構造 [4]

	次元	属性
A	音響信号	音色, 音高, 長さ, 大きさ
B	知覚対象	旋律, リズム, 和声
C	音楽構造	主題
D	意味, 内容	思想, 標題, 筋書

²同時に複数の音が演奏されない旋律

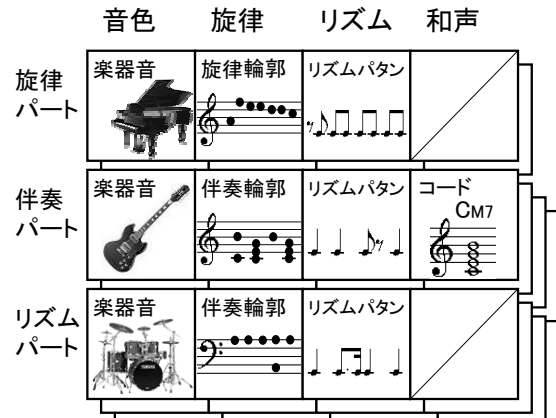


図 3: 本研究が想定する音楽構成要素

3.2.1 音色

各パート毎に、そのパートで用いられる楽器の楽器音を音色要素とする。

3.2.2 旋律

旋律パートにおける旋律要素として重要なものに、輪郭と音程の幅がある [5]。輪郭とは図 3 のような隣接する音符の上昇、下降の時系列のパターンである。この旋律輪郭を旋律要素として、旋律パートの構成要素とする。

伴奏パートにおける旋律属性として、伴奏輪郭が考えられる。伴奏輪郭とは旋律輪郭と同様に図 3 のような隣接する音符の上昇、下降の時系列のパターンであるが、伴奏パートでは同時に複数の音高が同時に生起する場合がある。

リズムパートの旋律属性は、ドラムセットにおける各打楽器の演奏の図 3 のような時系列のパターンを旋律輪郭とし、リズムパートの旋律要素とする。

3.2.3 リズム

音楽のリズム要素として重要なものに、音符の生起時刻と音長の時系列パターンがある。本研究では、それらをリズムパターンとし、各パートのリズム要素とする。

3.2.4 和声

旋律パートにおける和声要素は無しとする。

伴奏パートはコードに基づいて演奏され、その曲の和声を形成するので、コードは和声要素を表す。そこで、伴奏パートにおける和声要素として、コードを設定する。

リズムパートにおける和声要素は無しとする。

4 音楽構成要素の並列モーフィング処理

4.1 モーフィング処理の概要

モーフィング処理の手順を図 4に示す。加工対象として与えられた曲 A と曲 B の SMF ファイル³に音楽構成要素への分解処理を行い、各曲の音楽構成要素を得る。次に、各音楽構成要素に対して主観的類似度計算を行い、主観的類似度マップ上の布置を求める。次に、そのマップ上で滑らかな遷移となるような合成曲用の構成要素を知識ベースから選択する。最後に選択された音楽構成要素を元に楽曲を再合成して SMF ファイルに変換し、時間的に並べることによってモーフィングされた楽曲を得る。以下、各処理について述べる。



図 4: 楽曲のモーフィング処理の手順

4.2 譜面情報の音楽構成要素への分解処理

加工対象の楽曲の譜面情報を音楽構成要素に分解する。譜面情報としては、表 1の条件を満たす楽曲の SMF ファイルを想定する。また、SMF ファイルには各パートの音色番号の情報とコード名が含まれているとする。以下、分解処理について各音楽構成要素毎に述べる。

³Standard MIDI File, 音楽情報のデータフォーマット

4.2.1 音色要素

旋律パート、伴奏パート

譜面情報から音色情報を抽出し、この音色情報を音色要素とする。例として、図 2の曲の場合は、旋律パートの音色要素はピアノ音となる。

リズムパート

SMF ファイルの譜面情報では、打楽器の音色情報は一般的にドラムセットを一単位として扱われる。このようなドラムセットの情報を抽出し、音色要素とする。例として、図 2の曲のリズムパートの音色要素は「pop set」となる。

4.2.2 旋律要素

旋律パート

旋律パートの譜面情報から MIDI ノートナンバーで表した旋律輪郭

$$c_t = \{c_1, c_2, \dots, c_n\} \quad (1)$$

を抽出する。ここで、 $n = 1, 2, \dots$ は、それぞれ一つの音符に対応し、 $n = 1$ が時間的に最初の音符に相当し、添字の増加は時間の経過と対応する。 n は旋律パートの音符数である。例えば、図 2の曲の場合は、 $c_t = \{69, 77, 76, 76, 74, 74, 72\}$ となる。その時間差分 C_t を、

$$C_t = \begin{cases} c_{t+1} - c_t & (t=1, \dots, n-1) \\ -\sum_{k=0}^n c_k & (t=n) \end{cases} \quad (2)$$

と計算し、 C_t と c_1 を構成要素とする。例として、図 2の曲の場合は、 $C_t = \{8, -1, 0, -2, 0, -2, -3\}$ となる。

伴奏パート

伴奏パートの譜面情報から MIDI ノートナンバーで表した伴奏輪郭

$$c_t = \{c_1, c_2, \dots, c_n\} \quad (3)$$

を抽出する。各 c_t は、同時刻に発音される音符の MIDI ノートナンバーの集合を表す。例として、図 2の曲の場合は、 $c_t = \{(60, 71), (60, 64, 67), (71), (60, 64, 67)\}$ となる。

ここで、伴奏輪郭はコードに応じて様々な構成音から成るが、要素間の類似度計算のために CM7 のコードに対応した伴奏輪郭に変換したものを旋律要素とする。この変換処理は以下のように行なう。

まず、図 5のようにコード名の根音に応じて MIDI ノートナンバーに一定値を加算または減算することに

よって、根音がCのコードに対応する伴奏輪郭に変換する。次に、図5のように、コードタイプに応じて構成音の音名を変化させることによってCM7のコードに対応する伴奏輪郭を得る。これを旋律要素とする。

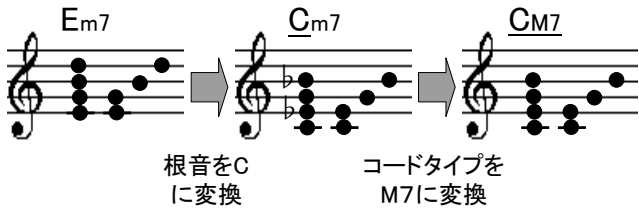


図5: 伴奏輪郭をCM7のコードに対応する伴奏輪郭に変換する処理

リズムパート

伴奏パートと同様の処理を行う。ただし、コードに関する変換処理は行わない。

4.2.3 リズム要素

それぞれのパートについて以下の処理を行う。

譜面情報から各音符の生起時刻と音長のMIDIクロック⁴表現をそれぞれ

$$a_t = \{a_1, a_2, \dots, a_n\} \quad (4)$$

$$l_t = \{l_1, l_2, \dots, l_n\} \quad (5)$$

とし、リズム要素とする。ただしMIDIクロックは、4分音符の長さが960クロックとする。ここで、各 a_t は同時刻に発音される音符の生起時刻を表す。 $t = 1$ が時間的に最初の音符群に相当し、添字の増加は時間の経過と対応する。 l_t は a_t の時刻に生起する音符群の音長を表す。同時刻に複数の音符が生起する場合は、 l_t の要素数は2以上となる。例えば、図2の伴奏パートの場合は、

$a_t = \{0, 960, 1920, 2880\}$,
 $l_t = \{(960, 960), (960, 960, 960), (480), (960)\}$
 などとなる。

4.2.4 和声要素

伴奏パート

譜面情報からコード名を抽出し、和声要素とする。ここで、コード名は和音の根音とコードタイプ⁵を表す。例として、図2の曲の場合は、CM7となる。

⁴MIDIで用いられる時間表現で、時間分解能の時間を一単位として数値表現される。

⁵和音における音の積み重ね方の種類

4.3 音楽構成要素の類似度計算

4.3.1 物理的類似度マップ

各音楽構成要素の物理的特徴量を表3のように設定した。各音楽構成要素の特徴量について、物理的非類似度を特徴量ベクトルの差のノルムとして計算する。この物理的非類似度に多次元尺度構成法[7]を適用し、図6のような物理的類似度マップを得る。

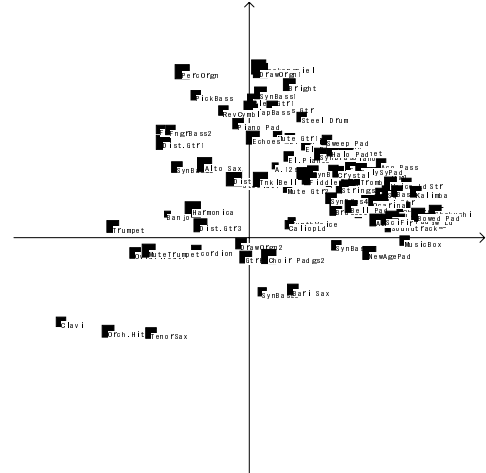


図6: 音色の3次元物理的類似度マップ(各四角の大きさは奥行き方向の軸の座標値の目安を表す。)

4.3.2 主観的類似度マップ

主観的な類似度関係を表すマップを得るために、多次元尺度構成法で得られた図7右のような主観的類似度マップ[8]と、物理的特徴量空間との間の対応関係をニューラルネットワークによって分析する。ニューラルネットワークの入力は各音楽構成要素の特徴量であり、出力は3次元主観的類似度マップ上の座標値である。学習データに対して主観的類似度関係の一対比較法によるアンケートを行い、その結果に多次元尺度構成法を適用して得られた主観的類似度マップの座標値を、学習時の教師信号とする。そして、学習後に得られたネットワークを用いて、任意の構成要素の主観的類似度マップ上の座標値の推測値を得る。

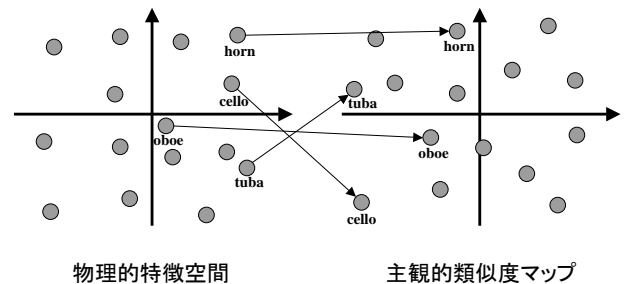


図7: 物理的特徴空間から主観的類似度マップへ写像

表 3: 各音楽構成要素の特徴量 (括弧内は特徴量の数を表す)

	音色	旋律	リズム	和声
旋律パート	楽器音の時間周波数解析結果のパワー値 (36)	音符列の音高の輪郭 (16)	音符の時系列パターン (16)	-
伴奏パート	同上 (16)	C,E,G,B の各音の音符生起の時系列パターン (64)	同上 (16)	コードの根音の半音音度, 五度圏 [6] の座標表現値 (4)
リズムパート	-	各楽器の演奏パターン (112)	同上 (16)	-

4.4 音楽構成要素選択処理

音楽構成要素選択処理では、各音楽構成要素の主観的類似度マップ上で曲 A と曲 B の構成要素に対応する点を滑らかに補完する構成要素を知識ベースから選択する。曲 A と曲 B の構成要素の主観的類似度マップ上の座標値をそれぞれ a, b とする。また、モーフィング処理で生成する合成曲の数 (曲 A, 曲 B を除く) を n とする。点 x_i を、

$$x_i \equiv \frac{(n-i+1)}{n+1}a + \frac{i}{n+1}b \quad (i = 1, 2, \dots, n) \quad (6)$$

とし、 i 番目の合成曲用に選択する構成要素は、点 x_i とのユークリッド距離が最小の構成要素とする。例として、音色の主観的類似度マップが図 8 のようであった場合、合成曲用の音楽構成要素として「オーボエ」、「トランペット」、「ハーモニカ」が選択される。

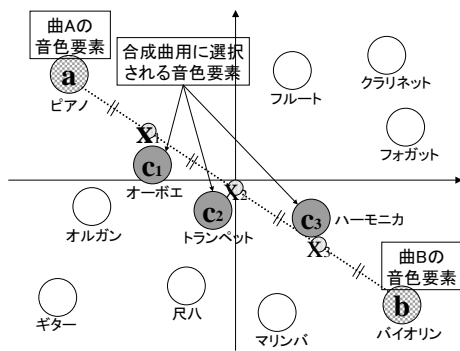


図 8: 音色の主観的類似度マップ上での音楽構成要素選択処理の例 (合成曲数 $n = 3$ の場合)

4.5 楽曲の再合成処理

音楽構成要素選択処理で選択された構成要素を元に楽曲を再合成する。また、主観的類似度のアンケートでは特定の要素の類似度を比較する必要があるので、ある単一の音楽構成要素についてそれ以外の構成要素に表 4 の標準設定要素を用いて再合成した 1 パート構成の曲を単一構成要素曲として、主観的類似度のアンケートに用いる。

4.5.1 旋律パートの再合成処理

再合成曲の旋律輪郭を c_t とする。 c_t の要素数はリズム要素 a_t の要素数 n_r とする。 n_C を C_t の要素数とし、 c_t を旋律要素の C_t と c_1 を用いて

$$c_{t+1} = c_t + C_t \% n_C \quad (t = 1, 2, \dots, n_r) \quad (7)$$

の漸化式を計算して求める⁶。この旋律輪郭 c_t にリズム要素の a_t, l_t を各 t について順次対応付けて、音長と発音時刻を定める。そして、得られた譜面情報を音色要素を用いて演奏する。例として、表 4 の旋律パートの標準設定要素を用いて旋律パートを再合成すると、図 10 の旋律パートとなる。

4.5.2 伴奏パートの再合成処理

旋律要素の伴奏輪郭 c_t は CM7 のコードに対応するものに交換されている。そこで、 c_t を和声要素のコードに対応する伴奏輪郭に変換する以下の処理を行う。

まず、図 9 のように和声要素のコードタイプに応じて旋律要素の構成音の音高の変換を行う。次に、コードの根音に応じて旋律要素に一定値を加算または減算を行う。これによって、和声要素のコードに対応する伴奏輪郭となる。

次に、リズム要素と音色要素を用いて、前項の旋律パートの再合成処理と同様の処理を行う。例として、表 4 の伴奏パートの標準設定要素を用いて伴奏パートを再合成すると、図 10 の伴奏パートとなる。



図 9: 伴奏輪郭の和声要素のコードに応じた変換

⁶(7) 式の % はモジュロ演算子である。

表 4: 各音楽構成要素の標準設定要素

	音色	旋律	リズム	和声
旋律パート	Grand Piano	$C_t = \{0\}, c_1 = 60$	$a_t = \{240t t = 1, 2, \dots, 16\},$ $l_t = \{216 t = 1, 2, \dots, 16\}$	-
伴奏パート	同上	$c_t = \{60, 64, 67, 71\}$	同上	CM7
リズムパート	Standard Set	$c_t = \{38\}$	同上	-

4.5.3 リズムパートの再合成処理

前項の伴奏パートの再合成処理と同様の処理を行なうが、コードに関する処理は行わない。例として表 4 のリズムパートの標準設定要素を用いてリズムパートを再合成すると、図 10 のリズムパートとなる。



図 10: 標準設定要素による再合成曲

4.5.4 再合成処理の例外

音色要素と和声要素の単一要素構成曲は、前途の手法を用いずに、それぞれ中央のドの音、和音を二分音符で演奏した音とする。また、リズムパートの音色要素については、2小節の単純なドラムパターン⁷を演奏した音とする。

5 評価実験

5.1 方法

5.1.1 曲素材の準備

評価実験に用いるサンプル曲を 100 曲用意した。旋律パートとコードは、歌謡曲、ポップス、演歌など幅広いジャンルの音楽の楽譜から引用した。伴奏パートとリズムパートは、ヤマハのミュージックシーケンサー QY20⁸に内蔵されているものを用いた。

この 100 曲に対して、音楽構成要素への分解処理を行った。そして、各構成要素について単一構成要素曲の再合成処理を行なった⁹。

⁷典型的な 8 ビートのドラムの演奏パターンで、最後に簡単なフィルインを含むものを設定した。

⁸幅広いジャンルの音楽の典型的な伴奏パターンを内蔵する

⁹リズム要素の単一構成要素の曲では、拍子を明確にするため

5.1.2 主観的類似度のアンケート

各構成要素に対して、100 曲の単一構成要素曲から主観的類似度計算の学習用に、各 10 曲の単一構成要素曲を適当に選択した。これらの曲を、2 の順列で並べた ${}_{10}P_2 = 90$ 曲を作成した。そして、アンケート用のホームページを作成し、それぞれの構成要素の主観的非類似度を、一対比較法によって 0 から 10 の 11 段階の尺度評価で被験者 1 名¹⁰に解答させた。

5.1.3 主観的類似度マップの作成

前項で得られたアンケートデータを元に多次元尺度法によって学習データ用の主観的類似度マップを得た。そして表 5 の仕様の階層型フィードフォワード型ニューラルネットワークによって主観的類似度計算のバックプロパゲーション学習を行い、学習データ以外の音楽構成要素を含む主観的類似度マップを得た¹¹。また、評価実験の参考用に物理的類似度マップも作成した。

表 5: 本研究で用いたニューラルネットワークの仕様

入力層	16 ~ 112 ユニット (特徴量数)
中間層	30 ~ 120 ユニット (1 層)
出力層	3 ユニット (3 次元座標値)
入出力関数	入力層: 線形関数 中間層, 出力層: シグモイド関数
初期加重	$[-0.05, 0.05]$ の一様乱数
学習係数	0.9

5.2 結果

適当な 2 曲を選び、中間曲数を 6 として提案手法のモーフィング処理を行ったところ、図 11 のような結果が得られた。また比較のため、物理的類似度マップ上で構成要素選択処理を行った場合のモーフィング結果も得た (図 11(b))。

に、曲の始めに四分音符 4 回によるカウントをハイハット音で挿入した。

¹⁰音楽の専門教育を受けていない東京大学四年生の男子

¹¹音色要素に関しては、旋律パートと伴奏パートとで共通の形式であるので主観的類似度マップも共通とする。同様に、リズム要素に関しても各パートで共通の主観的類似度マップを用いる。

図 11: 楽曲のモーフィング結果 ((a),(c):1 小節目と 8 小節目を元にモーフィング処理を行った結果。(b): (a) と同じ 2 曲 (曲 A, 曲 B) に対する物理的類似度マップ上での構成要素選択処理によるモーフィング処理結果)

5.3 考察

実験で得られた楽曲のモーフィング結果を MIDI 音源 QY20 で演奏したものを聴き [9]、その印象を元に以下の考察を行う。

5.3.1 構成要素毎のモーフィング処理結果の考察

得られたモーフィング結果を各パート毎に単独で聴いたところ、それぞれのパートで滑らかな遷移の楽曲が得られているという印象である。ただし、主観的類似度マップを用いた場合と物理的類似度マップを用いた場合の差が明確に感じられなかった。このことから、主観的類似度マップの改善の余地があると思われる。

今後の課題としては、構成要素毎のモーフィングをより滑らかなものとするために、知識ベースの構成要素数を増加させることが有効であると思われる。

5.3.2 モーフィング処理結果の音楽性に関する考察

モーフィング結果の 2～7 小節の合成曲は、音楽として認識するのに十分な規則性や統一感を有しているという印象である。ただし、本稿の手法では各パート間の相互関係を考慮していないために、パート間の和声的な機能などの点においてやや不自然と思われる部分がある。しかし、全体的な聴取では概ね納得感のあるモーフィング結果が得られたという印象である。

今後の課題としては、合成曲をより音楽的に自然な印象を与えるものとするために、構成要素選択処理において音楽性を示す何らかの評価関数を設定し、より音楽として自然な構成要素の組み合わせを選択する機構が必要であると考えられる。

6 まとめ

本稿では音楽が音色、旋律、リズム、和声の各構成要素から成ることに着目して、各構成要素の並列モーフィング処理と再合成処理による楽曲のモーフィング手法を提案した。評価実験によって得られた楽曲のモーフィング結果は音楽として納得のゆく合成曲による滑らかな遷移が実現されており、概ね満足のゆくものであった。ただし、合成曲の音楽としての自然さなどの点には改善の余地があるという印象である。

参考文献

- [1] 佐藤洋一, 佐藤いまり, 池内克史: 「3 次元物体形状および表面反射特性を考慮したモーフィング手法」, 電子情報通信学会論文誌, Vol.J80-D-II, No.9, pp.2287-2297, 1997.
- [2] 坂野秀樹, 武田一哉, 鹿野清宏, 板倉文忠: 「包絡と音源の独立操作による音声モーフィング」, 電子情報通信学会論文誌, Vol.J81-A, No.2, pp.261-268, 1998.
- [3] 岡田顕宏, 阿部純一: 「音楽知覚におけるモジュール性」, 音楽知覚認知研究, Vol.5, No.1, pp.22-43, 1999.
- [4] 梅本堯夫 編: 「音楽心理学の研究」, ナカニシヤ出版, 1996.
- [5] Aiello, R.: Musical Perceptions, Oxford Univ. Press, 1994.
- [6] 武藤誠, 半田伊吹, 日比啓文, 坂井修一, 田中英彦: 「音楽構成要素の主観的類似度の計算」, 情報処理学会第 61 回全国大会講演論文集 (2), pp.3-4, 2000.
- [7] 斎藤堯幸: 「多次元尺度構成法」, 朝倉書店, 1980.
- [8] 鈴木健嗣, 橋本周司: 「ニューラルネットワークを用いた感性情報の数量化」, 電子情報通信学会論文誌, Vol.J82-D-II, No.4, pp.677-684, 1999.
- [9] <http://www.mtl.t.u-tokyo.ac.jp/~muto/>