

# フレーズ間類似度に基づく楽曲構造提示機能を持つ 暗譜支援システム

伊藤 悠真<sup>1,a)</sup> 竹川 佳成<sup>2</sup> 寺田 努<sup>1,3</sup> 塚本 昌彦<sup>1</sup>

受付日 2015年6月4日, 採録日 2016年5月17日

**概要:** 楽器演奏者はコンサートやライブにおいて楽曲を暗譜して演奏することが一般的であるため、暗譜は重要な作業である。しかし、暗譜するためには楽曲を何度も演奏したり聴いたりする必要があり多大な労力を必要とする。そこで本研究では、楽曲のフレーズ間の類似度に着目した暗譜支援システムの構築を目的とする。提案システムは、楽曲に対してすべてのフレーズ間の類似度を算出し、それに基づいて楽曲の構造、類似しているフレーズ、類似しているフレーズ間の共通点および相違点を提示する。この機能により、学習者は直観的に楽譜の構造を理解できると同時に、共通点や相違点を意識することで暗譜のために記憶しなければならない量を削減でき、暗譜に必要な時間を短縮できる。提案する暗譜支援システムのプロトタイプを実装し、プロトタイプシステムと既存の楽譜を比較した評価実験により、提案手法の有用性を検証した。

**キーワード:** 音楽情報科学, 暗譜支援, フレーズ間類似度, 楽曲構造

## A System for Memorizing Songs by Presenting Musical Structures Based on Phrase Similarity

YUMA ITO<sup>1,a)</sup> YOSHINARI TAKEGAWA<sup>2</sup> TSUTOMU TERADA<sup>1,3</sup> MASAHIKO TSUKAMOTO<sup>1</sup>

Received: June 4, 2015, Accepted: May 17, 2016

**Abstract:** Players of musical instruments usually memorize musical scores for concerts and live performances. However, memorizing compositions requires much effort on the part of the players, as they have to play and listen to the musical compositions over and over again. The goal of our study is to construct a system for memorizing musical scores based on phrase similarity. The proposed system calculates the phrase similarity in the target musical composition, and based on this. It presents the musical structures of the piece, similar phrases, and the differences and similarities between these similar phrases. The learner can thus understand the musical structure intuitively, and can memorize the musical score in a short time because of the reduction of duplicated learning of the similar phrases. We evaluated the effectivity of the proposed system compared with conventional musical scores.

**Keywords:** musical information science, memorizing songs, phrase similarity, musical structure

### 1. はじめに

コンサートやライブなどでは楽曲を暗譜（譜面情報を暗記すること）して演奏することが一般的である。また、街角や路上での演奏では照明が暗かったり設置スペースの問題から楽譜を見られない場合がある。こういったことから、暗譜は演奏者にとって重要である。しかし、暗譜するためには楽曲を何度も演奏したり聴いたりしなければなら

<sup>1</sup> 神戸大学  
Kobe University, Kobe, Hyogo 657–8501, Japan

<sup>2</sup> 公立はこだて未来大学  
Future University Hakodate, Hakodate, Hokkaido 041–8655, Japan

<sup>3</sup> 科学技術振興機構さきがけ  
PRESTO, Japan Science and Technology Agency, Chiyoda, Tokyo 102–0076, Japan

a) yuma-ito@stu.kobe-u.ac.jp

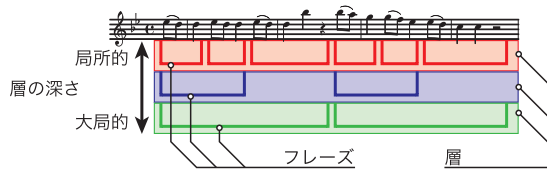


図 1 楽曲構造

Fig. 1 Musical structure.

ず多大な労力を必要とする。また、正確に暗譜することは難しく、本番の演奏では緊張などから同じ箇所を複数回演奏したり、一部の箇所を忘れてしまったりする場合も少なくない。

一方、効率の良い暗譜のためには、楽曲構造を分析し、意識することが重要であるとされている [1], [2]。図 1 に楽曲構造の例を示す。図 1 中の色分けされた部分は層を表し、上部の層は局所的な層、下部の層は大局的な層を表す。また、それぞれの括弧はそれぞれの層に属するフレーズ（音楽的なまとまりのある連続する音符群）を表している。これらの楽曲構造はフレーズ A やフレーズ A' といったように類似したフレーズで構成されている場合が多く、1 小節以下の短いフレーズ単位でみた場合、「1 音だけ音高が異なる」「左手の伴奏のリズムだけ異なる」「移調している」といったように同じ楽曲中にも多くの類似したフレーズが存在する。また、共通点を見つけることで記憶しなければならない量の削減を図れると同時に、相違点を意識することで正確にフレーズを記憶できる。しかし、この知見はプロフェッショナルな演奏家の経験則をもとにした暗譜法であり、科学的に検証されていない。さらに、新しく暗譜しようとする初見の楽譜に対して、手動で類似点などを探そうとした場合、楽曲の長さによっては探し出すだけでも多大な労力を要する。

そこで、本研究では、楽曲のフレーズ間の類似度に着目した暗譜支援システムの構築を目的とする。提案システムは、楽譜が存在する楽曲を対象とし、楽曲中のフレーズの類似度を算出する機能を持つ。本研究では、システムの利用対象者を「学習者」と称し、その学習者は練習中の楽曲をできるだけ短時間で正確に記憶する（以下、暗譜する）ことを目的としている。短時間で正確に楽譜を記憶することが求められる場合はよくある。たとえば、合唱などで急にピアノ伴奏での出演を求められた場合、ライブで演奏するために複数の楽曲を限られた時間の中でバンド練習しなければならない場合、次の 1 週間後のレッスンまでに練習曲を暗譜しないとイケないという課題が出された場合などがある。このような場合において、正確かつ短時間で暗譜できれば、細かな表現方法の練習などの演奏の質を上げるための練習に時間を割ける。ここで、本研究では五線譜形式の楽譜を利用するエレキギターの演奏者を対象とする。提案システムは、楽譜上に楽曲構造を視覚的に提示し、類似

したフレーズや類似フレーズ中の共通点および相違点にアノテーションを付与する。これにより、学習者は直観的に楽譜の構造を理解できると同時に、共通点や相違点を理解することで暗譜のために記憶しなければならない情報量を削減できる。さらに、システムは大局的な楽曲構造から局所的な構造まで多段の楽曲構造を保持している。大局的な層は楽曲のどのフレーズが似ているか似ていないかといったおおまかな構造の確認に適し、局所的な層は楽曲中のリズム、音高、演奏技法などの細かな音符の情報の確認に適している。このように各層において用途は異なるため、層ごとの用途を考慮した楽曲構造およびフレーズ間の類似度の提示方法について検討する。

以下、2 章では関連研究について述べ、3 章で提案システムの設計、4 章でシステムの予備実験について説明する。5 章でシステムの実装、6 章で提案システムの評価と考察について述べ、最後に、7 章で本論文のまとめを行う。

## 2. 関連研究

### 2.1 楽曲構造の提示

本研究のように楽曲構造の提示に着目した試みはいくつか事例があり、たとえば、ジャンルなど楽曲の特性を視覚化する研究 [3], [4], [5], [6], [7] や、音響信号から楽曲構造を抽出し提示する研究 [8], [9], [10] がある。これらの研究は、音楽を視聴するユーザを対象とする一方、本研究は効率的に暗譜することをめざした楽器演奏者を対象とし、本研究のように楽器演奏特有の指の動かし方、細かな演奏技法の違いなどに着目していない。

本研究と同じ演奏者および指揮者を対象とし、楽譜上に情報を可視化する研究もある [11], [12]。ScoreIlluminator [13] は、オーケストラスコア（楽曲のすべてのパートを記した楽譜）をメロディや伴奏など役割ごとに色分けすることで、指揮者や演奏者の読譜力を支援している。ScoreIlluminator では楽曲全体のフレーズに対してリズム、響き、メロディ、和声の 4 つの特徴量ごとにフレーズ間の距離を算出し、重みかけた和を最終的な距離としている。この距離を用いてフレーズをクラスタリングしている。この重みは学習者の経験則に基づき手動で決定される。本研究のように個々の音符を憶える行為である暗譜を目的とした場合については、この重みが具体的にどのような値になるか明らかにされていないという問題がある。ScoreIlluminator で用いられている特徴量は 4 つであるが、本研究では暗譜に適した特徴量を新たに採用し、さらに演奏者の特性をあらかじめ調べることで、暗譜に最適な重みづけを明らかにする。加えて、ScoreIlluminator で提示される情報は、楽曲全体の構造にとどまっておらず本研究のように楽曲の構造を階層化し各層の特性を考慮した情報の提示方法については検討していない。

さらに Songle [14] は演奏者や指揮者だけでなく、一般の

音楽視聴者を対象にしている web 上の楽曲鑑賞サービスである。Songle は音響信号解析により A メロ、B メロ、サビといった繰返し構造やボーカルの音高などを自動で解析および提示する。Songle を利用することで、楽曲の理解を深めることができる。しかし、Songle は五線譜ではなく、ピアノロール譜をベースにした情報を提示している。このため、大局的な楽曲の構造を学習者は理解できるが、音高やリズムなどの細かい情報は Songle 上で確認できないため、暗譜支援には不向きである。

## 2.2 暗譜支援

効率的な暗譜法に関して科学的に検証している研究は少なく、プロフェッショナルな演奏家の経験則をもとにした暗譜法がほとんどである。たとえば文献 [1] では、無意識の記憶は弱く、「次の音は何だったのだろうか」という簡単な疑問の答えを考えると崩壊してしまうとされ、意識的な記憶が必要であると述べている。具体的には、演奏者が図 1 に示すような楽曲構造を意識したり、転調のポイントを認識したりすることが意識的な記憶であるとしている。しかし、楽曲の規模が大きければ大きいほど、図 1 に示すような楽曲構造の規模も大きくなり、自力でこれらの構造を見つけ出すことは複雑な作業であると考えられる。したがって本研究では、学習者がこれらの構造を容易に理解することができ、かつ学習者の意識的な記憶に影響を与えられるように支援する。

文献 [2] では人が音楽を記憶するときのプロセスを認知心理学的な見地と情報理論的な見地から調査している。特に演奏者にとっては音楽構造を分析し、類似しているフレーズ間の違いを意識することが重要であると述べている。本研究では、楽譜上にフレーズ間の相違点を提示することで、曖昧に記憶してしまうことを防げるよう支援する。

## 3. 設計

2 章で述べたように暗譜には楽曲構造と類似しているフレーズ間の相違点の意識が重要である。楽曲構造は図 1 に示すように大局的な層から局所的な層まで多段の層で構成されている。大局的な層は楽曲のおおまかな構造の確認に適しており、局所的な層は楽曲中のリズム、音高、演奏技法などの細かい情報の確認に適している。最適な層は学習者の理解度により異なるため、柔軟に学習者が利用する層を変更できる仕組みが求められる。

そこで、本研究では、図 2 に示すような楽曲のフレーズに着目した情報提示を行うことで、上記の要求を満たした学習環境を提供する。提案システムは各音符の音高、発音時間、指の配置といった楽譜データをもとにフレーズ間の類似度を算出し、暗譜支援に有用なアノテーション（類似したフレーズや相違点の強調など）が付与された楽譜を表示する。図 2 において、上部のメニューバーは現在表示しているモード、中部の表示はアノテーションが付与された譜面、下部のスライダは各パラメータを調整するものである。

### 3.1 フレーズ間類似度

本研究ではフレーズ単位で類似度を計算する。フレーズおよびフレーズの階層構造の生成には Lerdahl らが提案している GTTM [15] を用いる。

#### 3.1.1 類似度の算出方法

音符の個数など長さが異なるフレーズ間のマッチングが必要となるため、類似度算出には DP (Dynamic Programming) マッチング [16] に基づいた算出方法を用いる。また、GTTM により層ごとにフレーズが定義されているので、層ごとに類似度を計算する。楽曲中のある層に存在す

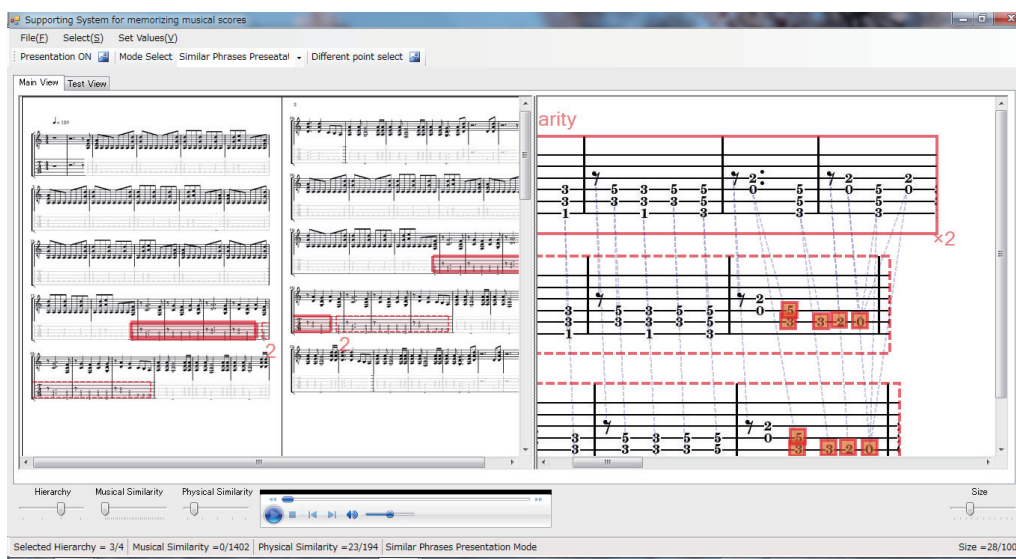


図 2 提案システムの提示例

Fig. 2 Presentation example of the proposed system.

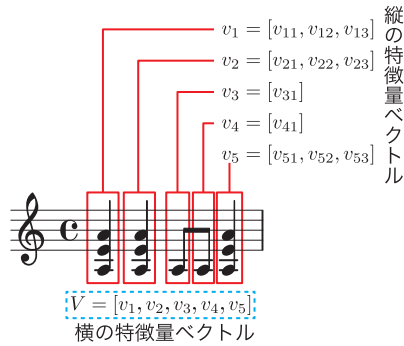


図 3 特徴量ベクトルの例  
Fig. 3 An example of feature value vector.

るフレーズの個数を  $N$  個とし、 $i$  番目と  $j$  番目のフレーズ間の距離を  $D(i, j)$  とする。  $D(i, j)$  は式 (1) で定義される。

$$\begin{aligned}
 D(i, j) = & w_{\text{timing}}d_{\text{timing}}(i, j) + w_{\text{pitch}}d_{\text{pitch}}(i, j) \\
 & + w_{\text{interval}}d_{\text{interval}} + w_{\text{art}}d_{\text{art}}(i, j) + w_{\text{abs}}d_{\text{abs}}(i, j) \\
 & + w_{\text{rel}}d_{\text{rel}}(i, j) + w_{\text{tech}}d_{\text{tech}}(i, j) + w_{\text{chord}}d_{\text{chord}}(i, j) \\
 & i = 1, 2, \dots, N, j = 1, 2, \dots, N \quad (1)
 \end{aligned}$$

ここで、 $d_{\text{timing}}(i, j)$ ,  $d_{\text{pitch}}(i, j)$ ,  $d_{\text{interval}}(i, j)$ ,  $d_{\text{art}}(i, j)$ ,  $d_{\text{abs}}(i, j)$ ,  $d_{\text{rel}}(i, j)$ ,  $d_{\text{tech}}(i, j)$ ,  $d_{\text{chord}}(i, j)$  はそれぞれ次項以降で説明する特徴量における距離を示しており、順にタイミング、音高、音程、アーティキュレーション、指の配置(絶対位置)、指の配置(相対位置)、特殊奏法、コードネームの特徴量における距離を表している。  $w_{\text{timing}}$ ,  $w_{\text{pitch}}$ ,  $w_{\text{interval}}$ ,  $w_{\text{art}}$ ,  $w_{\text{abs}}$ ,  $w_{\text{rel}}$ ,  $w_{\text{tech}}$ ,  $w_{\text{chord}}$  はそれぞれの特徴量に対応した重み係数であり、各々  $[0, 1]$  に正規化されている。重み係数の詳細については 4 章で説明する。 2 フレーズ間の距離  $D(i, j)$  の値は小さいほど 2 つのフレーズが類似していることを表し、類似度が高いといえる。  $D(i, j)$  が閾値以下になったものを類似しているフレーズとする。

次に本研究で想定する特徴量ベクトルの例を図 3 に示し、特徴量ベクトルの抽出方法および、2 つの特徴量ベクトル間の距離の計算方法を説明する。本研究では、図 3 のように和音も含むフレーズでのフレーズ間類似度の算出を想定しており、「縦の特徴量ベクトル」および「横の特徴量ベクトル」の 2 種類の特徴量ベクトルを抽出する。なお、フレーズからある基準に基づいて抽出されたベクトル状の数値の羅列を特徴量ベクトル、その基準を特徴量と呼ぶこととし、特徴量に関しては次節で詳細を説明する。縦の特徴量ベクトルとは「ある時間において同時に演奏される音のかたまりを抽出した」もので、図 3 の場合では、 $v_1$ ,  $v_2$ ,  $v_5$  は要素数が 3 の縦の特徴量ベクトルであり、 $v_3$ ,  $v_4$  は要素数が 1 の縦の特徴量ベクトルである。横の特徴量ベクトルとは「あるフレーズから抽出されるすべての縦の特徴量ベクトルを時系列順に並べてベクトルとした」もので、図 3 の場合では  $V$  が横の特徴量ベクトルにあたる。

$i$  番目と  $j$  番目のフレーズから抽出される横の特徴

量ベクトルをそれぞれ  $X(i) = [x_1, x_2, \dots, x_k, \dots, x_K]$ ,  $Y(j) = [y_1, y_2, \dots, y_l, \dots, y_L]$  とし、これらの間の距離を DP マッチング [16] に基づいて定義する。これら 2 つのベクトル間の距離を  $c(K, L)$  とすると、 $c(K, L)$  は次の手順で計算される。

(i) Initialization:

$$\begin{aligned}
 c(1, 1) &= DP(x_1, y_1) \\
 c(k, 1) &= DP(x_{k-1}, y_1) + 1 \text{ for } k = 2, \dots, K \\
 c(1, l) &= DP(x_1, y_{l-1}) + 1 \text{ for } l = 2, \dots, L
 \end{aligned}$$

(ii) for  $k = 2$  to  $K$  do  
for  $l = 2$  to  $L$  do

$$c(k, l) = DP(x_k, y_l) + \min \begin{pmatrix} c(k-1, l-1) \\ c(k-1, l) + 1 \\ c(k, l-1) + 1 \end{pmatrix}$$

end for

end for

(iii) Output:

Return  $c(K, L)$

ここで、 $DP(x_k, y_l)$  は 2 つのベクトル  $x_k$  および  $y_l$  間の距離を表す距離関数で、次項以降で各特徴量における距離関数について定義する。

上述の手順より  $d(i, j)$  をある特徴量における距離とすると、

$$d(i, j) = c(K, L)$$

と表される。次項以降で、特徴量や特徴量ごとの距離関数を定義し、その詳細を説明する。

本研究では使用する楽器に依存する類似度(演奏的類似度)と音高やリズムなどに関して楽器に依存しない類似度(音楽的類似度)、さらにクラシックやポップなどのジャンルに依存する類似度(ジャンルの類似度)の 3 つに類似度を分類して定義する。

### 3.1.2 音楽的類似度

各フレーズから抽出される音高やリズムなどの特徴量をもとにフレーズ間の音楽的類似度を算出する。これらの特徴量はどれも、楽器に依存せず、楽譜で最も基本的な特徴量であるため暗譜には重要である。また、図 4, 図 5, 図 6, 図 7 においてフレーズの下部の文字や数字は特徴量の具体的な値を示しており、外側の括弧は横の特徴量ベクトルを、内側の括弧は要素を複数持つ縦の特徴量ベクトルを表す。

#### タイミング

フレーズの最初の音を基準とし、各音符の 16 分音符を基準とした発音タイミングを特徴量とする。タイミングは和音内の各音符で値が変化することはないので、和音が



図 4 タイミングの類似例  
Fig. 4 An example of timing.

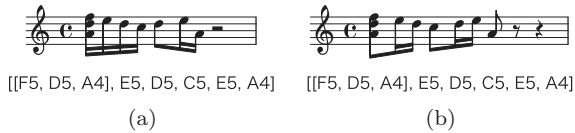


図 5 音高の類似例  
Fig. 5 An example of pitch.

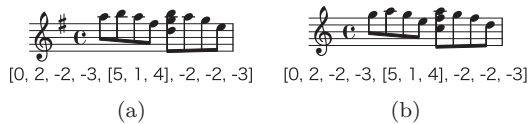


図 6 音程の類似例  
Fig. 6 An example of interval of pitch.



図 7 アーティキュレーションの類似例  
Fig. 7 An example of articulation.

存在するフレーズであっても、図 4 に示すように横の特徴量ベクトルはつねに 1 次元のベクトルとなる。さらに、図 4 に示すようにタイミングは整数値で与えられるので、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は式 (2) で定義される。

$$DP(\mathbf{x}_k, \mathbf{y}_l) = |x - y| \quad (2)$$

ただし、 $\mathbf{x}_k = [x]$ 、 $\mathbf{y}_l = [y]$ 、( $x, y$  は任意の整数値) である。

図 4 に示すフレーズにおいて個々の音符の長さは異なるが、発音タイミングが同じ音符 (図 4 の矩形で囲った音符) がある。図 4 の例ではタイミングの類似度は  $d_{\text{timing}}(i, j) = 2$  となる。

### 音高

各音符の絶対的な音の高さを音高の特徴量として定義する。図 5 に示すように音高は文字列で表されるタグを要素として持つ。また、音高は個々の音符に値が定義されるので、和音部分の縦の特徴量ベクトルは要素を複数保持している。音高の特徴量ベクトルはタグで与えられるので、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は式 (3) で定義される。

$$DP(\mathbf{x}_k, \mathbf{y}_l) = \begin{cases} 1 & x \neq y \text{ のとき} \\ 0 & x = y \text{ のとき} \end{cases} \quad (3)$$

ただし、 $\mathbf{x}_k = [x]$ 、 $\mathbf{y}_l = [y]$ 、( $x, y$  は任意の文字列) である。また、縦の特徴量ベクトルが複数の要素を持つ場合、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は式 (3) を距離関数として DP マッチング [16]

により計算される。

図 5 に示すように、フレーズ (a) とフレーズ (b) ではタイミングの類似度は  $d_{\text{timing}}(i, j) = 4$  となるが、音高の類似度  $d_{\text{pitch}}(i, j) = 0$  となる。

### 音程

フレーズの最初の音を基準とし直前の音符との音高の差を音程の特徴量として定義する。図 6 に示すように音高と同様に和音部分の縦の特徴量ベクトルは複数の値が保持される。また、図 6 のフレーズ (a) およびフレーズ (b) の最初から 5 つ目の音と 6 つ目の音のように前後の音符数が異なる場合は、すべての音符間で音高の差を計算し、その最小値を音程の特徴量とする。音程は図 6 に示すように、整数値で与えられるので、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  はタイミングと同様に式 (2) で定義される。また、音高と同様に縦の特徴量ベクトルが複数の要素を持つ場合、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は式 (2) を距離関数として DP マッチング [16] により計算される。

図 6 の場合、2 つのフレーズの音高の類似度は  $d_{\text{pitch}}(i, j) = 10$  となるが、音程の類似度は  $d_{\text{interval}}(i, j) = 0$  となる。

### アーティキュレーション

スラーやアクセントなどの各音符の表現に関わる記号をアーティキュレーションとして定義する。図 7 に示すように、アーティキュレーションはタグで与えられるので、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は音高と同様に式 (3) で与えられる。なお、図 7 ではスタッカートを「S」、アクセントを「A」、テヌートを「T」と表記している。

図 7 の場合、2 つのフレーズの音高の類似度およびタイミングの類似度はそれぞれ  $d_{\text{pitch}}(i, j) = 7$  および  $d_{\text{timing}}(i, j) = 9.4$  となるが、アーティキュレーションの類似度は  $d_{\text{art}}(i, j) = 0$  となる。

### 3.1.3 演奏的類似度

楽器固有の特性を対象とする演奏的類似度の特徴量として以下の 3 つがあげられる。

#### 指の配置 (絶対位置)

ある楽曲を演奏する場合、使用する指は楽器ごとに異なる。また、指の配置は、高速な演奏や音の強弱を制御するための重要な要素である。ギター演奏者はコードを指板上におく指の配置で記憶しており、指の配置の意識は重要である。

図 8 に指の配置 (絶対位置) の類似例と特徴量ベクトルの例を示す。図 8 下部に示している表は指の配置 (絶対位置) の特徴量ベクトルの具体的な値の例である。本研究では指の配置を各指が楽器上で指が置かれている位置もしくは指の状態で定義し、指がどの弦やフレットにも触れていない場合には「N.T. (not touch)」としている。指の配置は楽譜に記述されている情報ではなく、また楽譜から一意に決定できる特徴量ではない。しかし、一般的に指の配置は一貫性があり、楽器演奏者個人ではさらに強い一貫性が

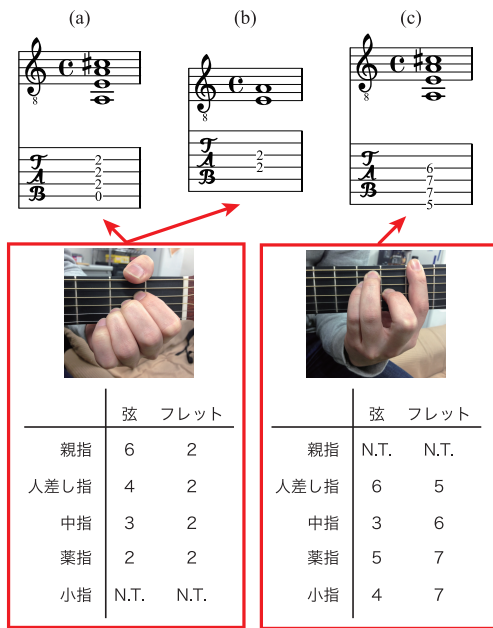


図 8 指の配置（絶対位置）の類似例

Fig. 8 An example of finger positions (absolute position).

ある。よって、楽譜制作者が指の配置情報を提供し、さらに楽器演奏者がカスタマイズすることで対応する。

指の配置（絶対位置）における縦の特徴量ベクトルは図 8 に示すように、和音であっても単音であってもつねに指 5 本分の値が存在するため 5 行 2 列の行列になる。したがって、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は式 (4) で定義される。

$$DP(\mathbf{x}_k, \mathbf{y}_l) = \sum_{k=1}^5 \sum_{l=1}^2 P_{kl} \quad (4)$$

ただし、

$$\mathbf{x}_k = \begin{pmatrix} x_{11} & x_{12} \\ \dots & \dots \\ x_{51} & x_{52} \end{pmatrix}, \mathbf{y}_l = \begin{pmatrix} y_{11} & y_{12} \\ \dots & \dots \\ y_{51} & y_{52} \end{pmatrix},$$

$$P_{kl} = \begin{cases} 1 & x_{kl} \neq y_{kl} \text{ のとき} \\ 0 & x_{kl} = y_{kl} \text{ のとき} \end{cases}$$

まず、図 8 において (a), (b), (c) の譜面はそれぞれ 1 つの全音符のみのフレーズとする。このときフレーズ (a) およびフレーズ (b) 間での音高の類似度は  $d_{pitch}(i, j) = 4$  となるが、指の配置（絶対位置）の類似度は  $d_{abs}(i, j) = 0$  となる。逆に、フレーズ (a) およびフレーズ (c) 間では音高の類似度  $d_{pitch}(i, j) = 0$  となるが、指の配置（絶対位置）の類似度は  $d_{abs}(i, j) = 10$  となる。

#### 指の配置（相対位置）

指の配置（絶対位置）では、2 フレーズ間で同じ位置で演奏されているかどうか類似度として現れるが、実際の楽器演奏においては、手の形は同じで演奏する位置は異なるという場合も存在する。この場合、「一方のフレーズを記憶すれば、もう一方はフレーズの開始位置のみを変えて

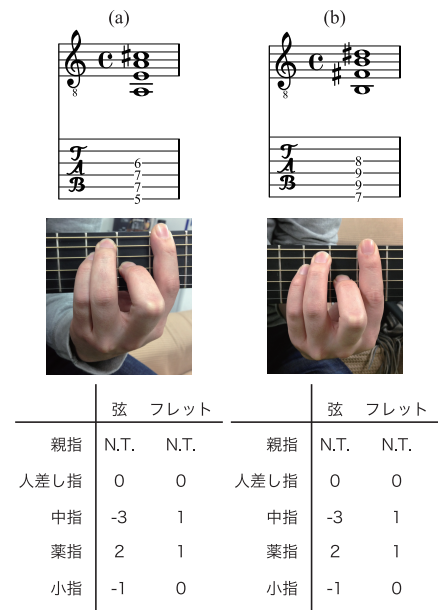


図 9 指の配置（相対位置）の類似例

Fig. 9 An example of finger positions (relative position).

残りは同様に演奏すればよい」と解釈することで暗譜に必要な情報の量を減らせる。

そこで、本研究では特徴量として指の配置（相対位置）を提案する。図 9 に指の配置（相対位置）の類似例と特徴量ベクトルの例を示す。ここで、図 9 のフレーズ (a) と図 8 のフレーズ (c) は同じフレーズである。指の配置（相対位置）は指の配置（絶対位置）の特徴量ベクトル中の各値を親指から小指にかけて順に直前の指の値との差をとることで定義している。ただし、直前の指の値が指の配置（絶対位置）において「N.T. (not touch)」の場合はもう 1 つ前の指、もしくは該当する指がない場合はその指の値を 0 として算出する。指の配置（相対位置）においても指の配置（絶対位置）と同様に、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は式 (4) で与えられる。

図 9 において、フレーズ (a) およびフレーズ (b) は音高の類似度および指の配置（絶対位置）の類似度がそれぞれ  $d_{pitch}(i, j) = 4$ ,  $d_{abs}(i, j) = 5$  となるのに対し、手の形や指の相対的な位置は同じであるため、指の配置（相対位置）の類似度については  $d_{rel}(i, j) = 0$  となる。

#### 特殊奏法

特殊奏法は楽器特有の奏法で、この奏法の有無によって演奏の際の手の動き方などが大きく異なるため、暗譜するうえで重要である。図 10 に特殊奏法の例としてカッティング奏法で演奏する 2 つのフレーズを示す。図 10 の中部には特徴量ベクトルの例を示している。ただし、図 10 中の「C」は「カッティング」、「N.T.」は特殊奏法が割り当てられていないことを示す「no technique」の略である。カッティング奏法とは、弦を押さえこまず軽く触れることで音が途切れるのを利用し、歯切れ良い音色を出すギターの奏法である。また、 $DP(\mathbf{x}_k, \mathbf{y}_l)$  は式 (3) で与えられる。

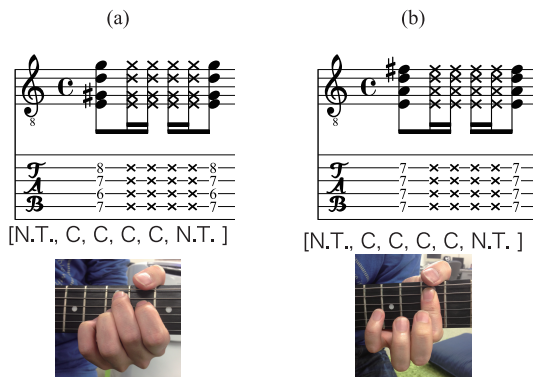


図 10 ギターにおける特殊奏法の類似例

Fig. 10 An example of musical performance technique of guitar.

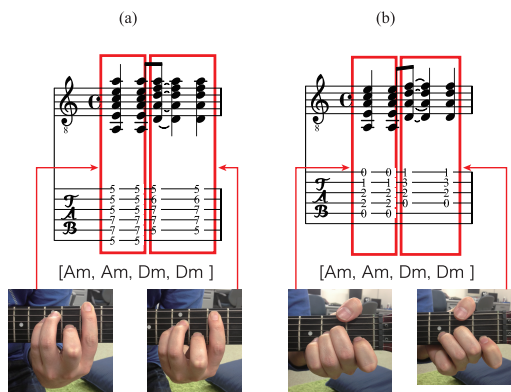


図 11 コードネームの類似例

Fig. 11 An example of chord symbol.

図 10 において指の配置（相対位置）や音高の類似度はそれぞれ  $d_{rel}(i, j) = 36$ ,  $d_{pitch}(i, j) = 12$  となるが、どちらも同じカッティング奏法で演奏しているので、特殊奏法の類似度は  $d_{tech}(i, j) = 0$  となる。

本研究で使用するギターの特殊奏法はほかに、ハンマリング・オン、プリング・オフ、チョーキング、チョーク・ダウン、ピブラート、スライド、グリッサンド、ミュート、ブリッジ・ミュート、アルペジオ、ハーモニクスがあり、これらも上で説明したカッティング同様の処理を行う。

### 3.1.4 ジャンルの類似度

#### コードネーム

ジャンルの類似度の特徴量としてコードネームを定義する。コードネームは、ジャズやポップスといった音楽ジャンルに多用される一方、コードネームが生まれる以前のクラシック音楽では基本的に使用されない。特にジャズやポップスの伴奏においては、コード進行という一定のパターンに従ってコードネームが進行することで楽曲を構成している場合が多く、暗譜するうえで重要な要素であるといえる。

図 11 にコードネームの類似例を示し、さらに図 11 の中部には特徴量ベクトルの例を示している。図 11 に示すように、コードネームの特徴量はタグで与えられるので、

$DP(x_k, y_l)$  は式 (3) で与えられる。

図 11 において、フレーズ (a)、フレーズ (b) 間の指の配置（絶対位置）の類似度は  $d_{abs}(i, j) = 40$  となるが、コードネームは同じであるためコードネームの類似度は  $d_{chord}(i, j) = 0$  となる。

### 3.2 提示方法

提案システムは図 12 に示す「全フレーズ表示モード」、図 13 に示す「類似フレーズ表示モード」および特にアノテーションのない白紙の楽譜のみを表示する「表示なし」の 3 つのモードを持つ。「全フレーズ表示モード」は楽曲の構造を理解したり、類似フレーズ表示モードで基準となるフレーズを選択したりするとき用い、「類似フレーズ表示モード」は選択したフレーズと他のフレーズとの類似度や相違点を見るとき用いる。「表示なし」は全フレーズ表示モードや類似フレーズ表示モードで表示した内容を楽譜を見ながら復習するときなど、アノテーションなしで楽譜を見たいときに用いる。以下、「全フレーズ表示モード」および「類似フレーズ表示モード」について説明する。

楽曲には図 1 に示すようなフレーズの層がいくつかあり、提案システムではこの層を学習者が自由に切り換えられるようになっている。図 12(a) は最も局所的な層を示しており、図 12(b) は最も大局的な層を示している。このモードでは学習者が選択した層に存在するすべてのフレーズを矩形で囲んで提示している。ただし、矩形の色に関しては隣り合う矩形どうしが同じにならないようになっているだけであり、特別な意味はない。学習者はこのモードで、練習する楽曲がどのようなフレーズに分かれているか、あるいはそれらのフレーズはどのような階層構造をなしているかを理解する。そのうえで学習者は、類似フレーズ表示モードにおいてどのフレーズの類似フレーズを見たいかを選択する。

#### 類似フレーズ表示モード

類似フレーズ表示モードでは、学習者が全フレーズ表示モードで選択したフレーズ（以下、基準フレーズ）およびそのフレーズと類似しているフレーズ（以下、類似フレーズ）を表示する。ここで、3.1.1 項で述べた提案システムでは、類似フレーズの閾値のデフォルト値は 0 に設定されており、そのうえで学習者が閾値を自由に変更できるようになっている。これにより、学習者個人の判断で類似フレーズを設定できる。図 13 に示すように類似フレーズを視覚的に表すために 2 種類の提示方法を提案している。1 つは、図 13 上部に示すように既存の楽譜の上に色付きの矩形を重畳することで類似フレーズを示し、もう 1 つは図 13 下部に示すように基準フレーズに類似するフレーズをまとめて表示している。実際には、図 2 に示すように図 13 上部と図 13 下部をそれぞれ左右に配置し表示する。

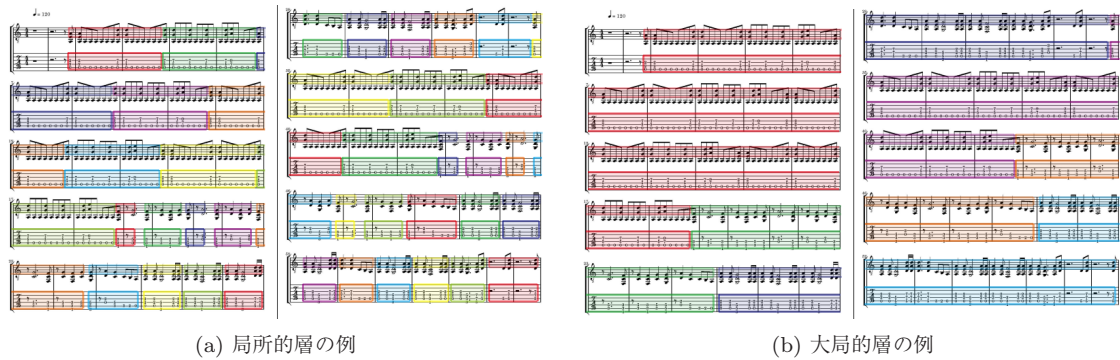


図 12 全フレーズ表示モードの提示例  
 Fig. 12 An example of all phrases presentation mode.

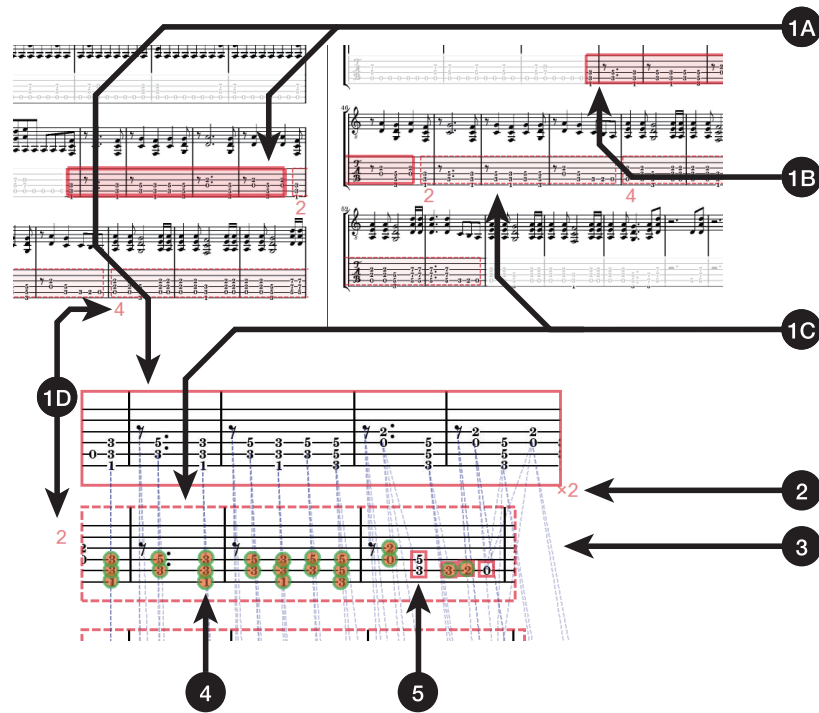


図 13 類似フレーズ表示モードの提示例  
 Fig. 13 An example of similar phrases presentation mode.

以下、図 13 の提示例について詳細に説明するが、簡条書きの番号は図 13 中の黒丸の番号と対応している。ただし、(1)のみは 1A, 1B, 1C, 1D の 4 種類に分かれている。

(1) 実線の太い矩形 (1A) で囲まれているフレーズは基準フレーズ、実線の細い矩形 (1B) で囲まれているフレーズは基準フレーズと同一のフレーズ、さらに破線の矩形 (1C) で囲まれているフレーズは類似フレーズを表す。また、類似フレーズの矩形近くの矩形と同色で表された数字 (1D) は類似度の高さのランキングを表している。図 13 下部では基準フレーズを最上段に、以下、類似度が高い順に類似フレーズを並べて表示する。これにより、学習者は容易に類似フレーズを確認できる。

(2) 同一フレーズが複数ある場合、その個数を「×2」のように表示することにより同一フレーズが何回繰り返さ

れているかを確認できる。(1) および (2) により散在する類似フレーズが整理されるため、訓練中に同一フレーズに差しかかった際に再度フレーズを覚えなおすことをしなくてすみ記憶に要する量を削減できる。

(3) 類似フレーズが存在する場合、音楽の類似度においては発音タイミングが基準フレーズと対応している音符どうしを、演奏的類似度においては指の配置（絶対位置）が基準フレーズと対応している音符どうしを DP マッチングの結果から決定し点線で結ぶ。これにより、発音タイミング、または指の配置（絶対位置）について類似性を確認できる。

(4) 類似フレーズ内で基準フレーズと対応する音符が基準フレーズと同じ特徴量を持つ場合、その音符を丸で囲む。囲む色や塗りつぶす色は特徴量ごとに色分けし、どの特徴量を表示させるかは学習者が選択できる。こ



れにより各特徴量における共通点を確認できる。

(5) 類似フレーズ内で基準フレーズと対応する音符が基準フレーズと異なる特徴量を持つ場合、その音符を矩形で囲む。(4)と同様に特徴量ごとに色分けでき、学習者が表示する特徴量を選択できる。これにより各相違点を理解し、誤った状態で記憶してしまうことを防ぐ。

また、提案システムは類似度の閾値および式(1)における重み係数のデフォルト値を持っており、学習者はいつでもこれらの値を自由に設定できる。これによって学習者は学習者自身の感覚で類似フレーズを自由に選べる。

#### 4. 予備実験

3.1 節でフレーズ間類似度とその算出に用いる特徴量について述べた。しかし、たとえば、指の配置が直観的に理解しやすい tab 譜を使って演奏を学習するギター演奏者は、音高情報が記載されている五線譜を利用するピアノ演奏者と比べて、指の配置に関して強い意識を持っていると思われる。また、ポピュラー音楽やジャズを演奏する演奏者はコードネームを重視するが、コードネームが生まれる以前のクラシック音楽を主に演奏する演奏者はコードネームを意識しないなど、楽器や音楽ジャンルによっても楽譜の読み方が異なる。したがって、演奏者が利用する楽器やジャンルを考慮して暗譜に必要な情報を適応的に提示すれば、効率的に暗譜することができるとと思われる。そこで、ピアノ演奏者とギター演奏者を対象とし、各楽器の演奏者が類似フレーズと判断するとき注目する特徴量について調査する。これにより、学習者が使用する楽器に対して最適な類似フレーズを提示できる。

##### 4.1 実験の手続き

実験は web 上のアンケート形式で行った。図 14 に実際に実験で利用した web ページの例を示す。ページ最上部には「手本フレーズ」が提示されており、「手本フレーズ」の下には「実験フレーズ」が 5 つ並んで提示されている。これらを 1 つの問題セットとし、4 つの問題セットに回答してもらった。被験者にはまず、手本フレーズを覚えたのち、手本フレーズと比べて似ている順に 3 つのフレーズを選んでもらった。また、図 14 に示すように、フレーズの提示だけでなく、フレーズの模範演奏を再生するボタンも用意し、被験者には手本フレーズや実験フレーズを必要に応じて実際に楽器で演奏しながら回答してもよいと指示した。手本フレーズは実験フレーズの右側に表示することで、手本フレーズをいつでも確認できるようにした。さらに、実験フレーズや問題の配置順は被験者ごとにランダムに提示した。

##### 4.2 課題曲

4 つの問題セットは 2 小節と 8 小節のピアノ用のフレーズと、伴奏用ギターフレーズ 8 小節およびソロ用ギターフ

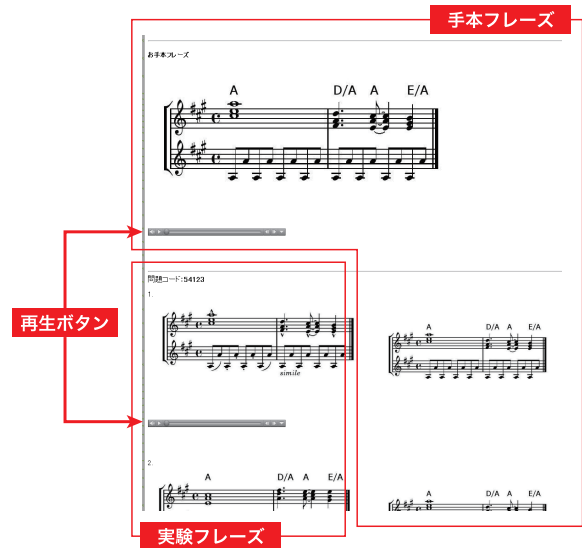


図 14 実験画面例

Fig. 14 An example of a test screen.

ズと、伴奏用ギターフレーズ 8 小節およびソロ用ギターフレーズ 8 小節をそれぞれ手本フレーズとした。表 1 に使用した楽曲の詳細を示す。表 1 において、抜粋箇所は課題曲のために抜粋した箇所の小節番号を示している。

本実験では、5 つの特徴量 (音高, タイミング, アーティキュレーション, 指の配置 (相対位置), コードネーム) に注目した。これら 5 つの特徴量から 1 つを選択し、以下に示すアルゴリズムで実験フレーズを生成した。アルゴリズムは大きく分けて 2 種類あり、ある特徴量のみを手本フレーズから変化させた実験フレーズ (以下、変化フレーズと呼ぶ) と、ある特徴量だけを変化させず、それ以外の特徴量を変化させた実験フレーズ (以下、非変化フレーズと呼ぶ) である。

なお、被験者には各実験フレーズにおいて、どの特徴量を適用させたかを知らせずに実験した。また、変化フレーズおよび非変化フレーズは手本フレーズと大きく異なることがないように設計した。ここで、図 15 (a) は手本フレーズを表し、図 15 (b), (c), (d), (e), (f) は手本フレーズに対する各特徴量の実験フレーズを表す。

##### 音高

変化フレーズであり、すべての音符に対して音高のみを 1 音以上遷移させ、ほかは変化させない。図 15 (b) に示すように手本フレーズに対して、すべての音を長 3 度下げている。

##### タイミング

変化フレーズであり、図 15 (c) に示すように、フレーズの随所に発音タイミングの異なる音符を挿入する。他の特徴量に関しては何も変化させない。

##### アーティキュレーション

変化フレーズであり、図 15 (d) に示すように、スラー、スタッカートなどの演奏記号を随所に挿入し、ほかは

表 1 予備実験の課題曲

Table 1 Trial songs of preliminary examination.

曲名 (作曲者)	抜粋箇所	ピアノ用, ギター用
君の知らない物語 (ryo) [17]	180-181	ピアノ用
Too Much Love Will Kill You (Brian May, Frank Musker, Elizabeth Lamers) [18]	21-28	ピアノ用
こっち向いて Baby (ryo) [19]	40-47	ギター用
バリバリ最強 No.1 (DARIA, FUYUKI 作曲) [20]	69-76	ギター用

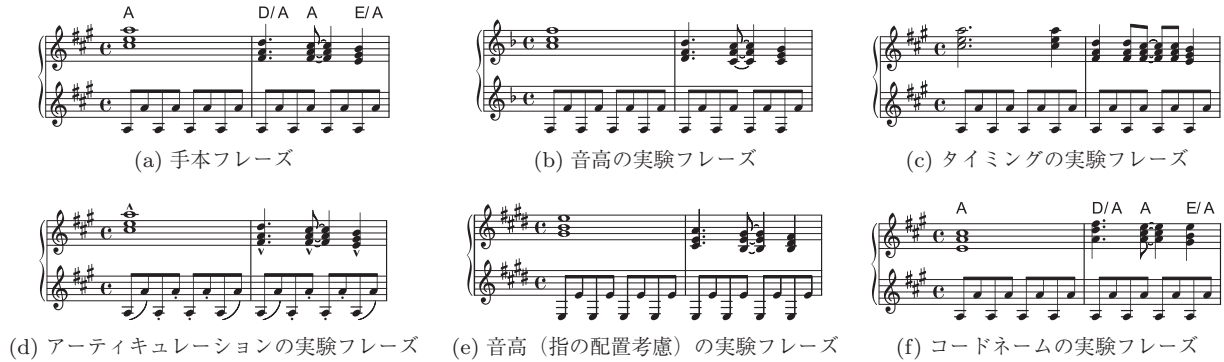


図 15 実験に使用したフレーズ例

Fig. 15 Examples of a standard phrase and similar phrases.

何も変化させない.

音高 (指の配置考慮)

非変化フレーズであり, 図 15 (e) に示すように, すべての音に対して指の相対的な位置が同じになるように音高を遷移させ, ほかは何も変化させない. ここでいう指の相対的な位置が同じとは, ピアノの場合は各指が押さえる鍵盤のパターン (白鍵か黒鍵か) が同じことで, ギターに関しては各指が押さえる位置が弦またはフレットに沿ってシフトしただけということを表す.

コードネーム

非変化フレーズであり, 図 15 (f) に示すようにコード構成音は同じであるが, すべてのコードに対して転回 (コードの構成音を並べ替えること) する. 他の特徴量に関しては何も変化させない.

4.3 被験者

被験者はギター歴半年から 6 年のギター演奏者 4 名で 19 歳から 24 歳, およびピアノ歴 3 年から 27 年のピアノ演奏者 6 名で 20 歳から 30 歳とした. ギター演奏者の被験者は 4 名ともポピュラー音楽を主に演奏する演奏者で, ピアノ演奏者は 6 名中 3 名がクラシック音楽, 2 名がポピュラー音楽, 1 名がジャズを主に演奏する演奏者である. 被験者には, 各フレーズにおいてまず手本フレーズを暗記してもらった後, 5 つの実験フレーズのうち手本フレーズと似ていると感じる実験フレーズを似ている順に 3 つ選ぶよう指示した.

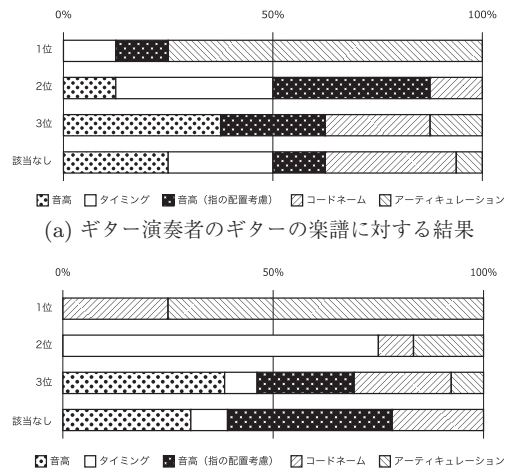


図 16 フレーズの類似度に影響を与える要素

Fig. 16 Effective factors on phrase similarity.

4.4 結果と考察

選ばれた特徴量の順位とその割合を図 16 に示す. 4 位以下のものは「該当なし」とし, 図 16 (a) はギター演奏者がギターの楽譜に対して回答した結果で, 図 16 (b) はピアノ演奏者がピアノの楽譜に対して回答した結果である.

どちらの楽器においても, アーティキュレーションが 1 位を占める割合は高い. したがって, たとえばアーティキュレーションのみ異なる 2 つのフレーズがあったとしても, 両楽器の演奏者ともに強い類似を感じるといえる. また, ギターにおいては音高 (指の配置考慮) とタイミングが, ピアノにおいてはタイミングが高い順位を占めている. これは各楽器の特性の違いが原因の 1 つとして考えられる.

まず、ギターおよびピアノにおいて共通して順位が高かったアーティキュレーションやタイミングについて考察する。アーティキュレーションが音符に付与されていても、鍵盤上や指板上での手の動かし方やリズムは変わらないため、類似しているフレーズとして認識されやすいと考えられる。タイミングに関しては、タイミングが異なる音符が挿入されたとしてもフレーズ中のリズムが大きく異ならなかったため、結果として類似しているフレーズと認識されたと考えられる。よって、これら2つの特徴量についてはフレーズのリズムが大きく異ならなかったことが、順位を上げた原因になっていると考えられる。

次にギターとピアノにおいて異なる結果となった音高（指の配置考慮）について考察する。ギター演奏においては五線譜よりも、tab 譜を重要視する演奏者が多い。tab 譜は弾く弦とフレットの番号が記されており、ギター演奏者は個々の音符の音高を意識しなくても tab 譜に表示された場所を弾けば正しい音高で演奏できる。さらに、指の配置が類似している場合、フレーズの最初の音符の弾く場所を把握しておけば、残りの音符は手本フレーズと同様に動かすだけなので、結果として指の配置の順位が上がったと考えられる。一方、ピアノ演奏者は五線譜を一般的に利用する。したがって、ピアノ演奏者にとって五線譜に記されている情報である音高とタイミングが重要な特徴量であるといえる。タイミングに関しては、上述の理由により順位が高くなったと考えられるが、音高に関しては順位が3位以下になっている。ピアノ演奏者にとって、音高は演奏する音を決定する重要な特徴量であり、音高に対して敏感になっていると考えられる。音高の少しの違いだけで類似していないと判断されたため音高の順位は低くなったと考えられる。

#### 4.5 提案システムへの適用

このように、訓練してきた楽器によって類似していると判断する特徴量が異なるという結果が得られた。これらの結果を利用することで、提案システムにおいて提示方法の最適化を図れる。具体的には、類似度算出の際に式(1)における重み係数の値を、学習者が使用する楽器ごとに変えるようにする。また、3.2節で述べたように重み係数は学習者が自由に設定できる値としているため、本予備実験の結果は暗譜支援を開始する際のデフォルトの値として設定する。これにより、学習者は学習の最初から学習者の演奏楽器に適した類似フレーズ情報を得ることができ、かつ学習者自身が類似度の基準を自由にカスタマイズできる。

4.2節で説明したように、本予備実験では変化フレーズと非変化フレーズの2種類があり、それぞれによって解釈が異なる。変化フレーズは上位であるほど、特徴量の変化に対して類似度が変化しないと認識されるため重み係数を0にする。一方、非変化フレーズでは上位であるほど、他

の特徴量を差し置いてその特徴量の類似度が高いことがフレーズ自体の特徴量に直結するといえるので、重み係数を1にする。ここで、4.2節より、変化フレーズの特徴量は「音高」「タイミング」「アーティキュレーション」が含まれ、不変化フレーズの特徴量には「音高（指の配置考慮）」および「コードネーム」が含まれる。

ギター奏者において重みを0にする特徴量は変化フレーズかつ上位である「アーティキュレーション」「タイミング」および、非変化フレーズかつ下位である「コードネーム」「音高」の4つとする。非変化フレーズかつ上位である音高（指の配置考慮）に関してはデフォルトではその重み係数を1としておく。

同様に、ピアノ奏者に対して重みを0にする特徴量は変化フレーズかつ上位である「アーティキュレーション」「タイミング」および、非変化フレーズかつ下位である「音高（指の配置考慮）」「コードネーム」とし、変化フレーズかつ下位である音高の重み係数は1とする。

## 5. 実装

3章で述べた暗譜支援システムのプロトタイプを実装した。プロトタイプシステムでは、楽譜の画像ファイルとXML形式の音高、リズム、指の配置といった楽譜データをもとにフレーズ間の類似度を算出し、3.2節で述べた情報を提示する。表示モードの切替えは画面上のボタンを押すことで切り替えられるが、ショートカットキーとして「Enter」キーを押すことで「表示なし」「全フレーズ表示モード」「類似フレーズ表示モード」の順に表示が切り替わるようにした。楽曲のフレーズへの分割は、文献[15]をもとに手動で分割することとした。なお、ソフトウェア開発にはWindows7 PC上のMicrosoft Visual C# 2010を使用した。

## 6. 評価

提案システムの暗譜時間の削減に関する有効性を評価するために、実装したプロトタイプシステムを用いて既存の楽譜を比較対象とした評価実験を行った。

### 6.1 実験方法

被験者には従来手法として既存の楽譜を用いた場合と提案手法としてプロトタイプシステムが生成した楽譜を用いた場合とで楽曲を暗譜できるまで練習してもらい、暗譜できるまでに要した時間を計測した。本実験では、課題曲を3曲用意しており、2曲もしくは1曲を提案手法で、残りの曲を従来手法で暗譜してもらった。

以下に従来手法および提案手法の詳細を示す。

**従来手法** 紙に印刷された楽譜を用いて暗譜する手法。楽譜は1ページがA4の紙の片面に印刷されおり、被験者は譜面台上に広げたり並べ替えたりなど自由に扱っ

表 2 課題曲

Table 2 Trial songs.

	曲名 (作曲者)	抜粋箇所	演奏時間	テンポ
曲 A	Brave Song (Gldemo ver.) (麻枝准) [21]	1-94	3 分 3 秒	124 BPM
曲 B	私は私の道を行く (Tom-H@ck) [22]	2-96	3 分 10 秒	120 BPM
曲 C	Hammer to Fall (Brian May) [23]	1-112	3 分 26 秒	130 BPM

表 3 暗譜時間

Table 3 Memorization time.

楽曲	被験者番号	暗譜時間	平均	S.D.	被験者番号	暗譜時間	平均	S.D.
A	2	52	43	9	1	41	37	4
	3	34			4	33		
B	3	44	40	4	1	40	40	0
	4	36			2	40		
C	1	46	44.5	1.5	2	28	29	1
	4	43			3	30		
計			42.5	6.05			35.3	5.22

てよい。ただし、ペンなどで紙上に書き込むことは禁止した。

**提案手法** プロトタイプシステムを用いて暗譜する手法。被験者には実験中に必ず、全フレーズ表示モードあるいは類似フレーズ表示モードのどちらかを少なくとも 1 度は表示するように指示し、それ以外は自由に暗譜してよいとした。ただし、従来手法と同様に紙に書き込むことや、テキストエディタなどにメモを残すことも禁止した。

また、被験者にはエレキギターと楽曲の音源を渡し、これらは自由に使ってよいと指示した。さらに、提案手法で練習する前に被験者には、プロトタイプシステムの使い方およびシステムを用いた学習方法を説明したデモムービー\*1を見せ、使い方や学習方法を理解してもらった。その後、練習用の楽曲を用いて実際にシステムを操作してもらい、使い方に慣れてもらった。加えて、提案手法の曲と従来手法の曲をそれぞれ練習する順番はランダムであり、偏りがでないようにした。暗譜ができたかどうかは被験者の自己申告で、実際に暗譜できているかどうかを確認するために、楽譜を見ない状態で課題曲をギターで演奏してもらった。このときに以下の条件をすべて満たしていれば、暗譜できているとした。

- 課題曲の最初から最後まで楽譜を見ずに正しく演奏できている。なお、演奏の途中で間違えてもすぐに修正できる場合は正しく弾けていると見なした。
- テンポは原曲と異なってもよく、被験者が自由に設定してよい。
- 譜面に記されているとおりに演奏することが被験者にとって難しいなどの理由で楽器での演奏が困難な場

合、その箇所を歌って補完してもよい。

さらに、プロトタイプシステムを用いて暗譜するときは、システム上で使用している表示モードの使用履歴のログをとった。

## 6.2 被験者

被験者は楽譜を読めて、エレキギターを演奏できる電気電子工学を専攻する大学院生および大学生の 6 名である。被験者のギター歴は 3 年から 10 年である。また、被験者はいずれの課題曲も練習したり聞いたりした経験はない。ただし、うち 2 名はblankが長いなどの理由で読譜能力や演奏能力が他の被験者に比べて著しく乏しく、分析の対象から外した。

## 6.3 課題曲

表 2 に使用した楽曲の情報を示す。表 2 中の抜粋箇所は課題曲のために抜粋した箇所を小節番号で表している。使用した楽曲は 3 曲で、ロックまたはポピュラー音楽の楽曲である。3 曲とも、それぞれ類似したフレーズや、層の数が同程度になるような箇所を演奏時間で 3 分から 4 分程度抜粋し課題曲とした。なお、これらの曲をそれぞれ「曲 A」、「曲 B」、「曲 C」と呼ぶ。

## 6.4 結果

実験結果を表 3 に示す。表において暗譜時間とは暗譜に要した時間で分単位で示している。平均、S.D. の列はそれぞれの手法における楽曲中での平均および標準偏差を表しており、最下段はそれぞれの手法での平均と標準偏差を表している。分散分析を行った結果、手法間 ( $F_{(1,6)} = 3.98$ ,  $p < .10$ ,  $\eta^2 = .29$ ) には有意な傾向がみられ、楽曲間 ( $F_{(2,6)} = 0.36$ ,  $\eta^2 = .052$ ) および手法と楽曲の交互作用

\*1 提案システムのデモムービーを下記 URL にアップロードする。  
<http://youtu.be/DAtJQQDsH60>

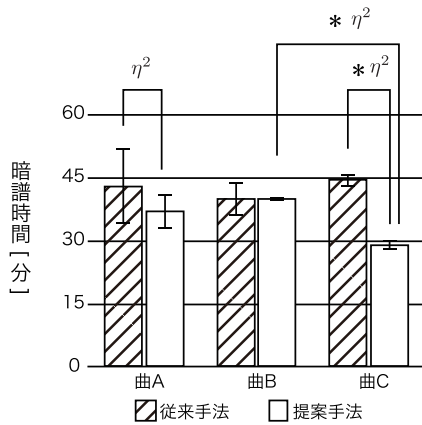


図 17 単純主効果の検定結果

Fig. 17 Result of analysis of simple main effect.

( $F_{(2,6)} = 1.58, \eta^2 = .23$ ) に有意差は認められなかった。ここで、 $\eta^2$  は効果量と呼ばれる値で、各要因において母集団の違いがどれだけ結果に影響しているかを示す値である。分散分析の結果について述べている文献 [24] によると、効果量はサンプルの数に影響されない標準化された値であり、分散分析の結果は  $F$  値と効果量の双方から考察する必要があると述べられている。文献 [24] に基づいて本研究で得られた各効果量を分析すると、手法間および手法と楽曲の交互作用には大きな効果、楽曲間には中ぐらゐの効果があるといえる。これらのことから、提案手法は楽曲によって効果が異なる可能性があることが示唆された。

次に手法および楽曲それぞれの単純主効果について検定を行った結果を以下の箇条書きおよび図 17 に示す。図 17 は楽曲、手法それぞれの値の平均および標準偏差を表している。「\*」がついた括弧は 2 つの要素間に有意差 (有意水準 5%) があることを示し、「 $\eta^2$ 」がついた括弧は 2 つの要素間に効果量大があることを示す。

- 曲 A における手法間の単純主効果：有意差なし，効果量大 ( $F_{(1,2)} = 0.37, \eta^2 = .16$ )
- 曲 B における手法間の単純主効果：有意差なし，効果量なし ( $F_{(1,2)} = 0.0, \eta^2 = .00$ )
- 曲 C における手法間の単純主効果：有意差あり，効果量大 ( $F_{(1,2)} = 73.92, p < .05, \eta^2 = .98$ )
- 従来手法における曲間の単純主効果：有意差なし，効果量中 ( $F_{(2,3)} = 0.16, \eta^2 = .10$ )
- 提案手法における曲間の単純主効果：有意傾向あり，効果量大 ( $F_{(2,3)} = 5.71, p < .10, \eta^2 = .79$ )

ここで、提案手法における曲間の単純主効果について、LSD 法を用いた多重比較により、曲 C が他の 2 曲に比べて有意に時間が短縮されている結果となった ( $MSe = 11.33, p < .05$ )。

上述の箇条書きおよび図 17 より、今回の実験では曲 C、曲 A の順で提案手法が効果的にはたらき、曲 B は従来手法と同程度の結果となった。次節以降で詳細に考察を行う。

## 6.5 考察

6.4 節で得られた結果は概念実証には十分な結果であると考えているが、一方で提案手法の制限や課題は残っている。

まず曲ごとの効果の違いについて考察する。提案手法が最も効果的にはたらいた曲は曲 C で、その次に効果的にはたらいたのは曲 A である。実験中の各モードの使用履歴より被験者 1 から被験者 4 全員が、全フレーズ表示モードで音源を聞き楽曲の概要を把握し、フレーズを選択して類似フレーズを確認していた。このように被験者 4 名とも提案システムが提示する情報を活用していた。また、曲 A および曲 C の特性としては、フレーズの切れ目で大きくリズムが変わることはなく、一見するとフレーズの切れ目が見にくい楽譜となっていた。このため、フレーズの切れ目が視覚的に分かりやすく提示される提案手法が効果的にはたらいたと考えられる。

曲 B は提案手法と従来手法で効果がほぼ同じだった。これは曲 B の特性が影響していると考えられる。曲 B は「単音のみで構成されているフレーズ」、「和音のみで構成されているフレーズ」、「高音域のみで構成されているフレーズ」などがはっきり分かれており、提案手法を適用しない状態の楽譜でもある程度楽曲の大まかな構造を視認できることが要因になったと考えられる。加えて類似フレーズより完全一致のフレーズの割合が高く、新たに記憶する情報が少なかったため、提案手法の有用性を下げた要因であるといえる。

また、楽曲と被験者の組合せについて考察する。従来手法では、曲 B が平均 40 分要しているのに対して、曲 A および曲 C はほぼ 44 分要している。同様に提案手法では、曲 A および曲 B が平均 37 分と 40 分要しているのに対して、曲 C は平均 29 分で暗譜されている。被験者からは、「普段演奏している楽曲に近いフレーズが多く含まれている楽曲は早く暗譜できたが、そうでない楽曲は暗譜に時間を要した」という意見があり、被験者の楽曲に対する得意不得意具合が影響した結果であると考えられる。今後、被験者が普段演奏する楽曲や得意楽曲などと課題曲の類似性に基づいた評価も行う必要がある。

次に各被験者が提案手法で用いた表示モードの推移を図 18 に示し、被験者ごとの考察を行う。図 18 の各図において、横軸は実験開始からの経過時間を秒単位で示しており、上側の図はそれぞれの時間でどの表示モードが選択されたかを表示している。下側の図はそれぞれの時間でどの層が選択されていたかを表示している。層は番号で表示されており、番号が大きくなるほど大局的な層であることを示す。ただし、「表示なし」モードが選択されているときは、どの層も選択されていないため、下側の図では空白として表している。取得した各モードの使用履歴のログの結果から、同一の被験者では各モードの使用頻度や時間推移

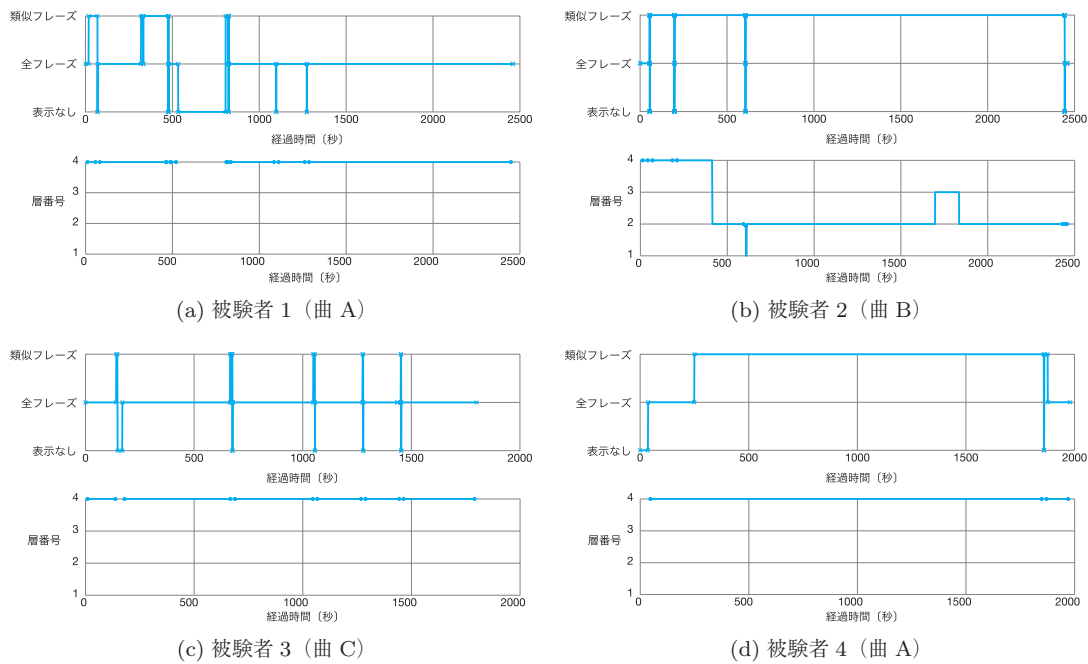


図 18 使用された表示モードの時間推移

Fig. 18 Transition of the used mode.

に同様の傾向がみられた。よって各被験者が提案手法を用いて暗譜した2曲のうち、どちらか1曲の結果のみを示す。

図 18 より、被験者ごとにモードの利用傾向は異なるものの、実験開始直後はどの被験者も全フレーズ表示モードを使用している。実験中の観測より、実験開始直後はどの被験者も音源を聞きつつ楽譜を眺めていた。ただし、被験者 4 だけは普段の暗譜時の癖から1度白紙状態の楽譜で音源を再生していたが、全フレーズ表示モードに移行してから、再度音源を最初から最後まで再生するという方法をとっていた。よって暗譜のプロセスの初期段階では楽曲の全体像をつかむことが重要であり、提案システムの全フレーズ表示モードは、全体像をつかむことに関してうまくはたっていたといえる。すべての被験者が最初は最も大局的な層を選択しており、このことを裏付ける結果となっている。また、被験者全員に共通した結果として、利用された特徴量の重み係数が初期状態から変化しなかったことである。さらに、被験者 2 はひととおり重み係数を切り替えたのち、初期状態に戻して暗譜していた。このことから、予備実験の結果を反映させた特徴量は適切であったといえる。具体的な値は以下のようにになっている。

$$(w_{\text{timing}}, w_{\text{pitch}}, w_{\text{interval}}, w_{\text{art}}, w_{\text{abs}}, w_{\text{rel}}, w_{\text{tech}}, w_{\text{chord}}) = (0, 0, 1, 0, 1, 1, 1, 0)$$

まず、図 18(a) より、被験者 1 は実験時間の前半に頻繁に表示モードを切り替え、後半は全フレーズ表示モードのみを表示している。被験者 1 は主に全フレーズ表示モードで楽曲の大まかな構造を把握しつつ、気になる箇所があれば類似フレーズ表示モードに切り替え、類似フレーズを探

していた。しかし、相違点や共通点の表示といった細かい情報の表示機能は使用しておらず、あくまで類似フレーズはどれかという情報のみを提示していた。同様の使い方は被験者 3 (図 18(c)) にも見られた。さらに、図 18 より、これらの被験者は、最も大局的な層を選択していた。この層は A メロ、B メロといったポピュラー音楽によく見られるフレーズで構成されており、ほぼ個々のフレーズは8小節単位からなっている場合が多い。また類似フレーズ間の相違点は集中していることが多いという特性がある。この2名の被験者は最初から最後まで最も大局的な層しか選択していない。2名とも「大まかな構造と、どのフレーズが類似フレーズかという情報が分かれば、細かな相違点や共通点は自力で見つけ出せる」とコメントしており、ポピュラー音楽の特性を活かしながら暗譜をしていたと考えられる。

一方、図 18(b) および図 18(d) より、被験者 2 および被験者 4 は類似フレーズ表示モードを主に使っている。特に被験者 2 はほぼ最初から最後まで類似フレーズ表示モードを使っており、相違点の表示も頻繁に利用していた。被験者 2 が利用した相違点はタイミングのみであり、被験者 2 が類似フレーズを見る際にタイミングの相違点を重要視していることが分かる。これら2名の被験者も上述のようなポピュラー音楽の特性は知っていたが、そういったことを意識しなくても、システムが類似フレーズや相違点の位置を視覚的に提示してくれるので、自分の経験則と一致させることでより確実に暗譜することができたとコメントしている。また、被験者 2 は基本フレーズを切り替えるたびに、類似度の閾値と層を切り替えていた。閾値に関しては

フレーズによって異なる値に設定されていたが、いずれの場合においても類似フレーズが数個表示される値に設定されていた。この被験者は「閾値を設定後類似フレーズが表示されるが、最終的に似ているか似ていないかの判断は自分の主観で判断している」とコメントしており、あくまで1つの情報として利用していたといえる。使用した層については図 18 より、被験者 4 は被験者 1 および被験者 3 と同じく最も大局的な層を用い、被験者 2 は最初に最も大局的な層を選択したのち、局所的な層を中心に選択していた。いずれの被験者においても用いられた層は A メロ、B メロといったフレーズの基準になるフレーズまたは、それらフレーズより 1 段階細かいフレーズで構成されているものである。すなわち、特にポピュラー音楽の暗譜においては A メロや B メロといったフレーズが基準になると考えられる。

上述の A メロや B メロといった構造があるという楽曲の特性は、クラシック楽曲でもよく見られる特性であり、同様の効果が期待できる。また、ピアノなどギター以外の楽器の場合、4 章で述べた実験により重み係数を調整して適用できると考えられる。しかし、ピアノでは左手で主旋律、右手で伴奏を演奏するという通常のギターにはない演奏上の特性がある。これにより、ピアノではある 2 つのフレーズ間で左手は同じメロディだが、右手は異なるメロディといったことがありうる。この場合の類似フレーズの重み係数のデフォルト値や新しい特徴量の導入に関してさらに調査する必要がある。

以上より、提案手法は適用する楽曲について、楽譜を眺めただけでは構造が視認しにくい楽曲に対して効果的にはたらくことが分かった。さらに、主に全体を見て暗譜するタイプの学習者に対しては全フレーズ表示モードが、細かく情報を見て暗譜したいと考える学習者に対しては類似フレーズ表示モードが効果的に機能したといえ、提案手法が有効にはたらく条件が明らかになった。

## 7. まとめ

本研究ではフレーズ間類似度に基づく楽曲構造提示機能を持つ暗譜支援システムを構築した。提案システムは楽曲の構造や、類似したフレーズおよびその相違点を提示することで学習者が直観的に暗譜学習を進められるように支援している。提案システムの有効性を評価するために、既存の楽譜を比較対象とした評価実験の実験結果から、提案手法は暗譜時間の楽曲や学習者の特性によって有効にはたらくものとそうでないものがあることが分かった。

今後の課題として、より多くの被験者で評価実験を行うことがあげられる。このとき、被験者が普段演奏する楽曲と課題曲との類似性を考慮して実験を行うことも必要である。また、演奏レベルやジャンルの異なる被験者を対象とした実験や、ギター以外の楽器への応用もあげられる。さ

らに、提案手法による暗譜が時間的にどれだけ持続するかといった効果についても検証する必要がある。

謝辞 本研究の一部は、科学技術振興機構戦略的創造研究推進事業（さきがけ）および文部科学省科学研究費補助金基盤研究（A）(23240010) によるものである。ここに記して謝意を表す。

## 参考文献

- [1] Bernstein, S.: *With Your Own Two Hands*, G. Schirmer, Inc. (1981).
- [2] Snyder, R.: *Music and Memory*, The MIT Press (2001).
- [3] Pampalk, E., Dixon, S. and Widmer, G.: Exploring Music Collections by Browsing Different Views, *Computer Music Journal*, Vol.28, No.2, pp.49-62 (2004).
- [4] Torrens, M., Hertzog, P. and Arcos, J.L.: Visualizing and Exploring Personal Music Libraries, *Proc. 5th International Conference on Music Information Retrieval (ISMIR2004)*, pp.421-424 (2004).
- [5] Mayer, R., Lidy, T. and Rauber, A.: The Map of Mozart, *Proc. 7th International Conference on Music Information Retrieval (ISMIR2006)*, pp.351-352 (2006).
- [6] Lamere, P. and Eck, D.: Using 3D Visualizations to Explore and Discover Music, *Proc. 8th International Conference on Musical Information Retrieval (ISMIR2007)*, pp.173-174 (2007).
- [7] Yoshii, K. and Goto, M.: Music Thumbnailer: Visualizing Musical Pieces in Thumbnail Images Based on Acoustic Features, *Proc. 9th International Conference on Musical Information Retrieval (ISMIR2008)*, pp.211-216 (2008).
- [8] Cooper, M.L. and Foote, J.: Automatic Music Summarization via Similarity Analysis, *Proc. 3rd International Conference on Musical Information Retrieval (ISMIR2002)*, pp.81-85 (2002).
- [9] 後藤真孝: SmartMusicKIOSK: サビ出し機能付き音楽聴機, 情報処理学会論文誌, Vol.44, No.11, pp.1737-2747 (2003).
- [10] 北原鉄朗, 後藤真孝, 奥乃 博, 片寄晴弘: 楽器音認識技術を用いた音楽の可視化, エンタテインメントコンピューティング 2007, pp.145-148 (2007).
- [11] 渡邊ふみ子, 藤代一成, 平賀瑠美: デジタルスコアによる楽曲学習支援インタフェース, 情報処理学会論文誌, Vol.45, No.3, pp.710-718 (2004).
- [12] Hayashi, A., Itoh, T. and Matsubara, M.: Colorscore: Visualization and Condensation of Structure of Classical Music, *Knowledge Visualization Currents*, pp.113-128 (2013).
- [13] 松原正樹, 岡本紘幸, 佐野智久, 鈴木宏哉, 延澤志保, 斎藤博昭: ScoreIlluminator: スコア色付けによるオーケストラスコアリーディング支援システム, 情報処理学会論文誌, Vol.50, No.12, pp.2937-2948 (2009).
- [14] 後藤真孝, 吉井和佳, 藤原弘将, Matthias, M., 中野倫靖: Songle: 音楽音響信号理解技術とユーザによる誤り訂正に基づく能動的音楽鑑賞サービス, 情報処理学会論文誌, Vol.54, No.4, pp.1363-1372 (2013).
- [15] Lerdahl, F.A. and Jackendoff, R.S.: *A Generative Theory of Tonal Music*, The MIT Press (1983).
- [16] Bellman, R.E.: *Dynamic Programming*, Princeton University Press (1957).
- [17] 君の知らない物語, フェアリー (2009).
- [18] バンドスコア クイーン/ジュエルズ ヴェリーベストオブクイーン, オクト出版社 (2004).

- [19] こっち向いて Baby, フェアリー (2010).
- [20] アニソン・グレート・コレクション (バンドスコア), ケイ・エム・ビー (2008).
- [21] バンドスコア Angel Beats!/Girls Dead Monster BEST SCORE, ケイ・エム・ビー (2011).
- [22] TV アニメ「けいおん!」オフィシャル バンドやろーよ!! PART2 OFFICIAL BAND SCORE, 株式会社ポニーキャニオン (2010).
- [23] BEST OF QUEEN, Hal Leonard Corp. (2005).
- [24] 水本 篤, 竹内 理: 研究論文における効果量の報告のために—基本的概念と注意点, 関西英語教育学会紀要『英語教育研究』, Vol.31, pp.57-66 (2008).



伊藤 悠真

1989 年生。2011 年神戸大学工学部電気電子工学卒業。2013 年同大学院工学研究科電気電子工学専攻修士課程後期課程修了。現在、同大学院工学研究科電気電子工学専攻博士課程後期課程に在籍。ウェアラブル・ユビキタス・

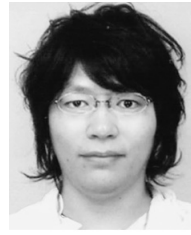
コンピューティング, 音楽情報科学に興味を持つ。



竹川 佳成 (正会員)

2007 年大阪大学大学院情報科学研究科博士課程修了。同年より神戸大学自然科学系先端融合研究環重点研究部助教。2012 年より公立ほこだて未来大学システム情報科学部助教。2014 年より公立ほこだて未来大学システム

情報科学部准教授, 現在に至る。2011 年には MIT Media Lab. にて Assistant Visiting Professor を兼務。博士 (情報科学)。音楽情報科学, 教育工学, エンタテインメントコンピューティングの研究に従事。



寺田 努 (正会員)

1974 年生。1997 年大阪大学工学部情報システム工学科卒業。1999 年同大学院工学研究科博士前期課程修了。2000 年同大学院工学研究科博士後期課程退学。同年より大阪大学サイバーメディアセンター助手。2005 年より同講師。

2007 年神戸大学大学院工学研究科准教授, 現在に至る。2004 年より特定非営利活動法人ウェアラブルコンピュータ研究開発機構理事, 2005 年には同機構事務局長を兼務。工学博士。アクティブデータベース, ウェアラブルコンピューティング, ユビキタスコンピューティングの研究に従事。IEEE, 電子情報通信学会, 日本データベース学会, ヒューマンインタフェース学会の各会員。



塚本 昌彦 (正会員)

1964 年生。1987 年京都大学工学部数理工学科卒業。1989 年同大学院工学研究科修士課程修了。同年シャープ (株) 入社。1995 年大阪大学大学院工学研究科情報システム工学専攻講師。

1996 年同専攻助教授。2002 年同大学院情報科学研究科マルチメディア工学専攻助教授。2004 年神戸大学電気電子工学科教授, 現在に至る。2004 年より特定非営利活動法人ウェアラブルコンピュータ研究開発機構理事長を兼務。工学博士。ウェアラブルコンピューティングとユビキタスコンピューティングの研究に従事。ACM, IEEE 等, 8 学会の会員。