

## 聴取者違和感指標に基づく全自動 DJ ミックスシステム

石 先 広 海<sup>†1</sup> 帆 足 啓 一 郎<sup>†1</sup> 滝 嶋 康 弘<sup>†1</sup>

本稿では、連続する楽曲のテンポを合致させ、ビート時刻を重ね合わせることで、複数のテンポの異なる楽曲を連続再生することが可能な自動 DJ ミックス再生システムを提案する。聴取者に自然な自動 DJ ミックスを提供するための重要な課題として、(1) 楽曲再生順序の決定、(2) 聴取者違和感の軽減、(3) テンポ推定誤りへの対処があげられる。提案システムでは、自動 DJ ミックス再生時の聴取者違和感を測定する指標を定義し、本指標値を最小限に抑えることが可能な楽曲再生順序決定方法と、テンポ調整方式を導入する。さらに、自動テンポ推定におけるテンポ推定誤りを適応的に修正することが可能な自動 DJ ミックスシステムを構築する。被験者による主観評価実験の結果、提案システムが既存システムに比べて違和感が少なく、質が高い自動 DJ ミックスが可能であることが示された。

### Automatic DJ Mixing System Based on Measurement Function of User Discomfort on Music

HIROMI ISHIZAKI,<sup>†1</sup> KEIICHIRO HOASHI<sup>†1</sup>  
and YASUHIRO TAKISHIMA<sup>†1</sup>

This paper proposes a novel method to develop an automatic DJ mixing system, which enables smooth transition between individual songs by adjusting the tempi of the current and next songs, and aligning the position of beats. To develop such a system without disrupting listeners, there are three issues that must be solved: (1) decision of playing order, (2) minimization of user discomfort, and (3) octave errors of tempo estimation. In this research, we first define a measurement function of user discomfort, based on user experiments. This function is utilized in the proposed system to automatically calculate playing order and adjust the tempi of the songs to reduce user discomfort. Furthermore, the system implements a unique tempo adjustment method, which is robust to octave errors of tempo estimation. Subjective evaluation results showed that the proposed method is capable of high-quality automatic DJ mixing compared to conventional methods.

### 1. はじめに

近年の音楽圧縮技術の発展と音楽配信サービスの急速な普及とともに、電子データとして楽曲を購入し、個人が何千もの電子楽曲データを保持することが可能となっている。このような背景から、これまで主流であった自宅での音楽鑑賞に加え、携帯電話や携帯型音楽再生機で気軽に音楽を持ち運び、通勤時やジョギングなどの様々な状況の BGM として楽曲を利用することが一般的になりつつある。

個人が蓄積した楽曲データを聴取する方法としてはシャッフル(ランダム)再生やユーザが作成・編集したプレイリストを利用した再生が主流である。しかし、単なるシャッフル再生やプレイリストを利用した再生では、状況の BGM として不適切となることがある。たとえば、ジョギングやダンスなどの BGM として利用する場合に、選曲の好みや前後の楽曲の曲調の不一致、楽曲遷移における無音区間など、ユーザの没入感や雰囲気損なう要素が多く存在する。このように、状況に適した BGM を再生するためには、選曲や再生リストの構成などを考慮し、聴取者の雰囲気や没入感を維持させることが重要である。

一般的に、クラブハウスやディスコでは音楽を途切れさせないことが重要であるとされており、聴取者の没入感や場の雰囲気を維持するために、連続する楽曲のリズムを同期させたまま楽曲を移行させる。具体的には、楽曲どうしのテンポを合致させ、楽曲遷移部においてビート時刻を重ね合わせ、聴取者へ滑らかな楽曲移行を促している。このような技術を持つものは DJ (Disk Jockey) と呼ばれ、主に再生する楽曲の選択を行い、テンポを合致させ、DJ によっては効果音を挿入するなどして楽曲の連続再生 (DJ ミックス) を実現する。

しかし、個人が DJ ミックスを様々な状況で利用するためには専門知識や機材などが必要となり、個人が蓄積した楽曲を利用して楽しむことは困難であった。そこで本稿では、個人の楽曲に対しても DJ ミックスを容易に適用することができ、利用者が自然な楽曲再生を楽しむことが可能な自動 DJ ミックスシステムを提案する。さらに、聴取者主観評価実験により提案システムの有効性を検証する。

本稿では、2 章において本研究の関連技術を紹介し、3 章で自動 DJ ミックスを実現する際の従来技術の課題を提示する。4 章では自然な DJ ミックスを実現するための指標を被験者実験で検証し、5 章で提案システムの詳細を記述する。6 章で提案システムの有効性検証

<sup>†1</sup> KDDI 研究所  
KDDI R&D Laboratories Inc.

実験および考察を記述し、最後に 7 章でまとめを述べる。

## 2. 関連研究・関連方式

本章では、DJ ミックスの自動化において関連する研究および方式について紹介する。自動 DJ ミックスシステムを実現するための重要な要素技術として、主に「選曲技術」、「テンポ・ビート推定技術」、「再生技術」があげられ、これらの 3 項目を自動化するために有用な研究成果が多く報告されている。

「選曲技術」として、膨大な楽曲データベースから自動で聴取者の嗜好に適した楽曲や、音響的に類似する楽曲を抽出する技術の研究がさかに行われている。たとえば、協調フィルタリングを利用した音楽推薦技術<sup>1)–6)</sup>があげられ、聴取者の履歴情報を活用した音楽推薦、プレイリスト生成を実現している。また、楽曲信号の特徴に基づいて音楽検索を行うシステム<sup>7),8)</sup>や履歴情報と楽曲特徴の両者を考慮した楽曲推薦システム<sup>9)</sup>も提案されており、このような技術を利用して聴取者の嗜好や、音響的に類似した楽曲で構成されるプレイリストを自動で作成することが可能である。

「テンポ・ビート推定技術」においても多くの研究成果が報告されている。楽曲のビートを信号特徴量から推定するビートトラッキングシステムとして、Goto らは楽曲信号から推定したオンセットやコード遷移などから、ビートの階層構造を認識させるシステム<sup>10)</sup>を提案している。また、フィルタバンク分析に基づくシステム<sup>11),12)</sup>や、周波数解析を利用して信号の周期性を解析する手法<sup>13)–15)</sup>なども提案されている。

「再生技術」では、楽曲を連続して再生する際の無音区間を回避する方法として、クロスフェード再生があげられる。クロスフェード再生は楽曲遷移時に音量を調整することで前の楽曲をフェードアウト、後続の楽曲をフェードインさせる方法でラジオ放送などでも利用されている。さらに DJ ミックスにより近い技術として、クロスフェード再生に加えて自動で楽曲どうしのテンポを合致させることが可能なシステム<sup>16),17)</sup>が提案されている。これらのシステムでは、事前に CD データベースなどからテンポ情報を取得し、楽曲の音高を変化させずにテンポのみを変化させ、楽曲信号から自動もしくは半自動でビート情報を抽出し、ビート時刻を重ね合わせることで DJ ミックスの自動化を図っている(図 1)。

## 3. 問題点とアプローチ

2 章で記述した選曲技術、テンポ・ビート推定技術、再生技術を用いて自動 DJ ミックスシステムを構築した場合に、各技術に対応する問題が存在する。本稿では各技術の問題点で

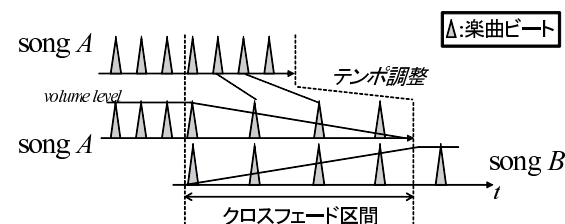


図 1 既存 DJ ミックス方式イメージ図

Fig. 1 Conceptual illustration of naive DJ mixing.

ある以下 3 項目について説明する。

- (1) 楽曲再生順序の決定
- (2) テンポ推定誤りへの対処
- (3) 連続再生時の違和感の低減

まず、選曲技術の問題点に関して述べる。自動 DJ ミックスシステムの再生対象楽曲として協調フィルタリングや楽曲信号の特徴に基づいて自動作成したプレイリストを利用することができる。しかし、自動で作成したプレイリストは再生順序について特に考慮されていないため、DJ ミックスを適用した場合に急激なテンポの変化が発生する可能性がある。そのため、プレイリスト内の楽曲の順序構成をシステム側で考慮して再生する必要がある。

次に、テンポ・ビート推定技術の問題点に関して述べる。自動 DJ ミックスを実現するためにテンポ・ビート推定技術を利用することが考えられる。しかし、テンポ・ビート推定技術では、推定結果が正解テンポ情報に対して、2 倍もしくは 2 分の 1 で得られる共通の問題(オクターブ誤り)が存在する。推定結果のオクターブ誤りはテンポ・ビート推定技術に共通かつ解決困難な問題であると Gouyon<sup>19)</sup>も報告している。テンポ推定結果がオクターブ誤りであった場合、DJ ミックスを適用する際に本来は不必要なテンポ変化が適用される可能性がある。文献 17) では実際の DJ がミックスした楽曲の分析に基づき、テンポの差が大きい場合には DJ ミックスを適用せず、無音区間が生じる通常の再生を適用している。したがって、テンポ推定結果がオクターブ誤りであった場合、本来は DJ ミックスが可能な楽曲どうしに対しても DJ ミックスが適用されない。そのため、自動 DJ ミックスシステムでは、テンポ・ビート推定結果のオクターブ誤りを考慮する必要がある。

最後に再生技術の問題点について述べる。従来の再生技術として、文献 16), 17) で提案されている技術があげられる。これら技術の根幹は DJ ミックスの適用対象となる楽曲どう

しのテンポを合致させて楽曲を遷移させることである。文献 18) によれば、聴取者に対して自然な DJ ミックスを提供するために、楽曲遷移の際にテンポの急激な変化を回避する必要があると記載されている。このことから、テンポの変化が聴取者に違和感を与える重要な要素であると推測できる。しかし、従来の再生技術では聴取者に自然な DJ ミックスを提供することに関して特に考慮していないため、聴取者に違和感を与える可能性がある。そのため、DJ ミックス時のテンポ変化が聴取者違和感に与える影響を詳細に調査し、聴取者の違和感を考慮した DJ ミックス方法を検討する必要がある。

#### 4. 予備実験

本章では被験者による主観評価実験に基づき、DJ ミックスにおけるテンポの変化の度合いと聴取者の違和感の関係について調査する。本稿では文献 17), 18) の知見に基づき、DJ ミックスにおけるテンポ変化の度合い(テンポ調整量)の増加が自然な聴取を妨げる一要因であると仮説をたて、予備実験で仮説の妥当性を検証する。具体的には実際に DJ ミックスが施された音源を生成し、ミックス音源に対する被験者主観評価値とテンポ調整量の相関分析により検証する。また、テンポ調整量の変化に対する聴取者の違和感の傾向を調査し、DJ ミックス時の聴取者違和感を測定するための指標を定義する。

##### 4.1 テンポ調整量と聴取者評価の相関調査

###### 4.1.1 実験データおよび実験方法

実験データは市販 CD から収集した 96 曲を利用し、専門家によってビート時刻、テンポ情報を事前に付与した。すべての楽曲からサビを含む 30 秒区間を抽出しミックス素材とした。なお、楽曲テンポ情報は楽曲速度の指標である BPM (beats per minute) を用いた。

本実験では、ミックス素材から 2 楽曲を選択し、図 1 に記載の DJ ミックスを適用してミックス音源を生成した。具体的には、ミックス対象楽曲を  $A$ ,  $B$  とすると、楽曲  $A$  の BPM ( $T_A$ ) を楽曲  $B$  の BPM ( $T_B$ ) に合致させるためのテンポを調整する係数(テンポ調整係数)  $f$  は以下の式で表せる。

$$f(T_A, T_B) = \frac{T_B}{T_A} \quad (1)$$

$f(T_A, T_B) > 1$  のとき、楽曲  $A$  のテンポは速くなり、 $f(T_A, T_B) < 1$  のとき、テンポは遅くなる。最終的に式 (1) に基づいて、一方の楽曲 BPM を他方の BPM と合致させることでミックス音源を作成した。以降、式 (1) により得られたテンポ調整係数を BPM に乗算し、楽曲テンポを乗算結果の BPM に変化させることをテンポ調整と記述する。また、一方の楽

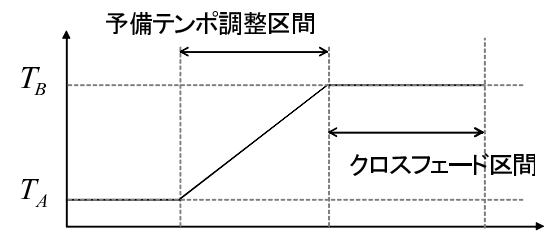


図 2 実験データにおける予備テンポ調整区間  
Fig. 2 The interval of gradual BPM change.

曲の BPM を他方の BPM に合致させることを既存 DJ ミックスと記述する。

生成した音源は計 40 音源で、各ミックス音源のクロスフェード区間は 7 秒とした。なお、テンポ調整対象となる楽曲に対し、クロスフェード区間の直前に予備テンポ調整区間として 5 秒区間を設定した。予備テンポ調整区間では、テンポ調整係数を 1 から  $f(T_A, T_B)$  まで直線的に徐々に変化させることで急激なテンポ変化を回避している(図 2)。なお、予備テンポ調整区間の時間設定は、急激なテンポ変化を回避するうえで十分な数値を事前実験で調査し、設定した。

楽曲のテンポ調整には文献 17) にも利用されている TSM (time scale modification) アルゴリズム, SOLA<sup>20)</sup> を利用した。SOLA により時系列信号に対して一定区間の間引きや補間を行うことで信号の伸縮操作を行い、音高を変化させずにテンポのみを変化させた音源を生成した。

本実験では被験者 22 名を対象として実施した。被験者は聴取したすべての音源に対してクロスフェード区間における自然さを 5 段階で評価した(5:違和感なし-1:違和感あり)。

###### 4.1.2 実験結果

テンポ調整量の増加が自然な聴取を妨げる一要因であるという仮説を検証するために、テンポ調整量を  $f > 1$  のとき  $f - 1$ ,  $f < 1$  のとき  $f^{-1} - 1$  として計算し、テンポ調整量と自然さ評価の相関を調査した。図 3 にテンポ調整量と自然さ評価における平均値の相関図を示す。テンポ調整量と自然さ評価の相関係数は  $-0.903$  となっており、図 3 から明らかとなり、テンポ調整量と自然さ評価値に強い負の相関が確認できる。この結果から DJ ミックス再生において、テンポ調整量の増加によって、聴取者の自然さに対する評価が下がる傾向にあることが確認された。

上記の結果より、テンポ調整量の増加が DJ ミックスにおいて自然な聴取を妨げる一要因

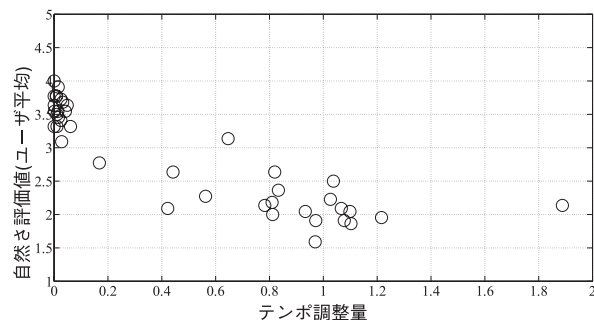


図 3 ユーザ評価における自然さ評価値平均とテンポ調整量の相関

Fig. 3 Correlation of average user ratings and amount of tempo adjustment.

であるという仮説は妥当であるといえる。つまり、DJ ミックスシステムにおいてテンポ調整量を低減させることが、聴取者に自然な DJ ミックスを提供することにつながるといえる。

#### 4.2 テンポ調整量に対する違和感の傾向調査

4.1.2 項で、テンポ調整量の増加が聴取者の違和感につながる事が示された。本節ではさらに詳細に違和感の傾向を調査するために、 $f > 1$  および  $f < 1$  区間において聴取者が違和感を覚えるテンポ調整係数を測定した。測定結果に基づいて聴取者のテンポ調整に対する違和感の傾向や許容範囲を明らかにすることで、DJ ミックス音源に対する聴取者の違和感を表現する指標として用いることが可能となる。

##### 4.2.1 実験方法・実験データ

被験者聴取用の音源として、RWC 研究用音楽データベース<sup>21)</sup> から 18 曲選択し、すべての楽曲に対して徐々に BPM を増加させた音源と、徐々に BPM を減少させた音源の 2 種類を用意した。音源作成の際にテンポ変化にともなう音質の影響を緩和するために、2 種類のテンポ変化ツール (*SOLA*, *SoundTouch*<sup>\*1</sup> ライブラリ) を用いて、音の高さを変化させずにテンポのみを変化させた。*SOLA* では時系列信号の単純な伸縮操作によりテンポを変化させているのに対し、*SoundTouch* ライブラリでは伸縮操作時の信号結合部において、信号の重複箇所の親和性が高くなる信号区間を探索し、該当区間に対して信号伸縮操作を行う。これにより、*SOLA* に比べて高品質なテンポ変化を実現している。これらのツールを用いて各 36 音源、合計 72 音源を作成した。

\*1 SoundTouch Audio Processing Library: <http://www.surina.net/soundtouch/>

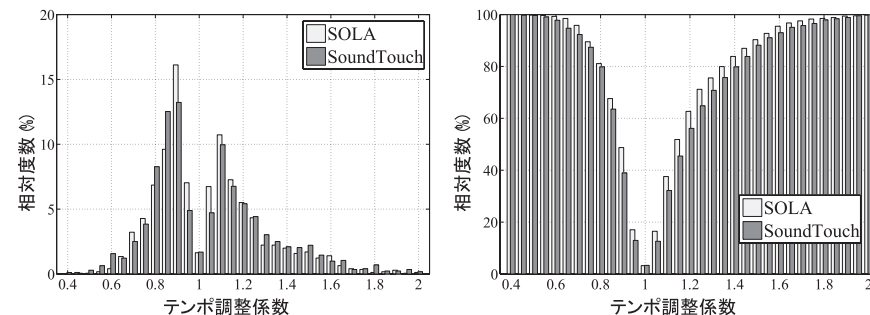
図 4 テンポ変化方式 (*SoundTouch*, *SOLA*) におけるテンポ調整係数  $f$  に対する聴取者違和感分布 (左) と累積度数分布 (右) (相対度数)

Fig. 4 Distribution of user discomfort to tempo adjustment factor (left: normal distribution, right: cumulative frequency distribution).

実験用音源では楽曲開始から 15 秒を起点として 3 秒ごとにテンポを変化させた。テンポを速くする場合にはテンポ調整係数  $f$  を 0.05 刻みで 1.00 から 2.00 まで変化させ、テンポを遅くする場合には 0.05 刻みで 1.00 から 0.30 まで変化させた。テンポ調整係数の変化量とテンポ変化の時間間隔は、急激なテンポ変化とならない数値を事前実験に基づいて設定した。なお、本実験は 96 名を対象とし、1 人あたり 36 音源を聴取してもらい、各音源に対して違和感を覚える時刻を収集した (有効データ数: 計 3,422 サンプル)。

##### 4.2.2 実験結果・考察

各音源について、被験者が入力した時刻から聴取者が違和感を覚えるテンポ調整係数  $f$  の分布を求めた。 $f$  の分布および累積度数分布を図 4 に示す。図 4 において、 $f > 1$ 、 $f < 1$  区間におけるピーク位置はそれぞれ 1.10, 0.90 となり、両テンポ調整方式において同じピーク位置を示した。

次に、 $f > 1$  および  $f < 1$  区間におけるテンポ調整係数の平均および分散を求めた (表 1)。表 1 におけるテンポ調整係数を用いて、テンポ調整量と  $f > 1$  および  $f < 1$  区間の聴取者違和感の傾向を調査した。*SoundTouch* および *SOLA* における  $f > 1$  区間の平均値のテンポ調整量はそれぞれ 0.227, 0.226 であり、 $f < 1$  区間の平均値のテンポ調整量は両者とも 0.174 となり、聴取者が違和感を覚えるテンポ調整量は、 $f > 1$  区間が  $f < 1$  区間を上回っていた。 $f > 1$  区間と  $f < 1$  区間における違和感の傾向の差異を検証するために検定を実施した。 $t$  検定 ( $p < 0.0001$ ,  $\alpha = 0.01$ ) および、 $F$  検定 ( $p < 0.0001$ ,  $\alpha = 0.01$ ) によって平均および分散の差は統計的に有意であることを確認した。このことより聴取者は BPM

表 1 テンポ調整係数平均および分散

Table 1 Average and variance of tempo adjustment factors.

テンポ調整ツール	項目	$f > 1$	$f < 1$
SoundTouch	平均	1.227	0.852
SOLA	平均	1.226	0.852
SoundTouch	分散	0.0415	0.0092
SOLA	分散	0.0348	0.0074

の増加に比べ、減少に対して違和感をより覚えやすいことが判明した。

#### 4.3 聴取者違和感指標の定義

本節では、聴取者違和感を測定するための指標を前節での結果に基づいて定義する。聴取者違和感の指標は、 $f > 1$  および  $f < 1$  における聴取者違和感の傾向を考慮して、 $f > 1$  および  $f < 1$  のテンポ調整量に対して重み付けを行うことにより表現する。具体的にはテンポ調整係数  $f$  に値する違和感指標  $L_{dc}(f)$  を以下のように定義した。

$$L_{dc}(f) = \begin{cases} a(f-1) & f > 1 \\ 0 & f = 1 \\ b(f^{-1}-1) & f < 1 \end{cases} \quad (2)$$

式 (2) における重み係数  $a, b$  は、表 1 の  $f > 1, f < 1$  区間における平均テンポ調整係数から得られる  $L_{dc}$  が等しくなるように設定した。具体的には、 $a = 0.765, b = 1.000$  として重み係数を設定した。

また、図 4 において、 $f > 1, f < 1$  区間におけるピーク位置はそれぞれ 1.10, 0.90 となり、聴取者違和感を回避するためのテンポ調整の適用可能範囲（厳密適用範囲）として利用することとする。さらに、表 1 の聴取者が違和感を覚えるテンポ調整係数の平均値から、 $0.852 < f < 1.227$  の範囲を超えてテンポ調整を適用した場合、多くの聴取者が違和感を覚えると予測できるため、この範囲を平均的なテンポ調整の適用可能範囲（平均適用範囲）として利用する。

### 5. 提案システム

本章では、DJ ミックスの要素技術である「選曲」、「テンポ・ビート推定」、「再生」を自動化した自動 DJ ミックスシステムの詳細について説明する。本システムでは、DJ ミックス処理において聴取者違和感指標値を軽減させることで聴取者に対して自然な DJ ミックス

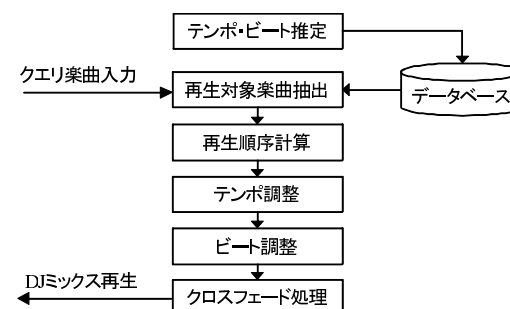


図 5 自動 DJ ミックス再生システム構成図

Fig. 5 Framework of full-automatic DJ mixing system.

を実現する。

はじめにシステム処理概要と流れを示す。本システムは主に、楽曲を格納するデータベースと、「テンポ・ビート推定」、「再生対象楽曲抽出」、「再生順序計算」、「テンポ調整」、「ビート調整」、「クロスフェード処理」の 6 つの処理によって構成されている。提案システム構成を図 5 に示す。システムは利用者から検索クエリとなる楽曲情報を入力してもらう。システムはクエリ楽曲情報に基づいてデータベースから曲調が類似する楽曲を再生対象楽曲として抽出し、DJ ミックス再生に適した順序を計算する。得られた順序に基づいて、楽曲遷移部における 2 楽曲の BPM を調整し、ビート時刻を重ね合わせる。最終的に、楽曲が重複する区間にクロスフェード処理を適用し、再生する。以降、各処理の詳細について記述する。

#### 5.1 テンポ・ビート推定

テンポ・ビート推定処理では、DJ ミックス対象となる楽曲群に対して、自動テンポ・ビート推定技術を用いてテンポおよびビート時刻を推定する。本システムでは、事前にシステムデータベースに格納されている楽曲群に対してテンポ・ビート推定処理を適用し、メタ情報として保持する。本システムでは、2 章で記述したテンポ・ビート推定技術の中から、事前実験により安定した精度が見込まれた BeatRoot<sup>14,\*1</sup>を用いてビート時刻を推定させる。なお、テンポ情報は推定されたビート情報に基づいて以下の式により求める。

$$BPM = \frac{60}{IBI} \quad (3)$$

\*1 <http://www.elec.qmul.ac.uk/people/simond/beatroot/>

ここで、 $IBI$  (InterBeat Interval) は推定ビート時刻から求められるビート間隔の平均 (sec) から算出する。

### 5.2 再生対象楽曲抽出

本処理では「選曲技術」の代替として、音楽推薦・検索技術を利用する。本システムでは、履歴データが不必要な音響的特徴に基づく音楽検索技術<sup>8)</sup>を適用し、データベースから再生対象楽曲を抽出する。具体的には、ユーザが指定した楽曲(クエリ)の音響的特徴を表現する特徴ベクトルと、検索対象楽曲の特徴ベクトルのベクトル類似度に基づいて楽曲を検索し、類似度上位  $N$  楽曲を再生対象楽曲として用いる。

### 5.3 テンポ調整

本処理では、テンポ推定結果がオクターブ誤りとなる場合にも違和感指標値を低減させることが可能なテンポ調整について記述する。既存 DJ ミックスではミックス対象楽曲に付与されたテンポ情報を基に、式(1)を適用することでテンポを調整する。したがって、ミックス対象楽曲の正解テンポ情報が近い場合には、テンポ推定結果のオクターブ誤りによって違和感指標が増大する。

本処理では、テンポ情報を楽曲組合せに応じて適応的に修正することで、テンポ推定結果がオクターブ誤りとなる場合にもロバストなテンポ調整を適用する。さらに、ミックス対象楽曲の両者に対してテンポ調整を適用することで楽曲単位でのテンポ調整量を軽減する。これにより、テンポ推定結果がオクターブ誤りとなる場合にもロバストかつ、違和感指標をより低減可能なテンポ調整が実現できる。以降、オクターブ誤りを適応的に修正したテンポ情報のことを最適テンポと記述する。

#### 5.3.1 最適テンポ計算

ミックス対象となる連続した2楽曲を  $m_i, m_{i+1}$ 、それぞれの BPM を  $T_{m_i}, T_{m_{i+1}}$  とし、楽曲  $m_i$  における最適テンポ  $T_{opt|m_i}$  を求める。 $T_{opt|m_i}$  は楽曲  $m_i$  におけるテンポ調整候補群  $T'_{m_i} = 2^C \times T_{m_i}$  ( $C = -2, -1, 0, 1, 2$ ) から抽出され、DJ ミックスにおけるテンポ調整量が最小となるようなテンポオクターブ関係にあるテンポを表すこととし、以下の式で求めることができる。

$$T_{opt|m_i} = 2^{C_{opt}} \times T_{m_i} \quad (4)$$

ここで、 $C_{opt}$  は  $C_{opt} = \operatorname{argmin}_C (|T'_{m_i} - T_{m_{i+1}}|)$  を満たす値を表している。 $T_{opt|m_i}$  は任意の楽曲組合せ  $m_i, m_{i+1}$  が与えられた際に複数の解が得られることがあるが、その場合には  $|C_{opt}|$  がより小さいものを利用する。たとえば、 $(T_{m_i}, T_{m_{i+1}}) = (50, 75)$  とし与えられたときに、 $T_{opt|m_i} = 50, 100$  が得られるが、 $T_{opt|m_i} = 50$  を最適テンポとして利用する。

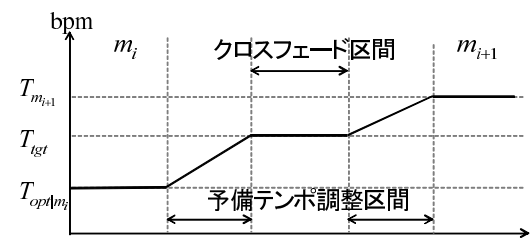


図6  $T_{tgt}$  を用いたテンポ遷移イメージ図

Fig. 6 Conceptual illustration of tempo adjustment with  $T_{tgt}$ .

#### 5.3.2 デュアルテンポ調整

次に、最適テンポ  $T_{opt|m_i}$  を用いてミックス対象楽曲  $m_i, m_{i+1}$  の両者をテンポ調整し、楽曲単位でのテンポ調整量を軽減することが可能なデュアルテンポ調整について説明する。デュアルテンポ調整では、楽曲  $m_i$  および  $m_{i+1}$  の BPM から求められる違和感指標の和がクロスフェード区間において最小となるテンポ(遷移テンポ)を用いてテンポを合致させる。その際に、遷移テンポに向けて徐々にテンポを調整し、遷移テンポからオリジナルのテンポ情報に徐々に戻すための予備テンポ調整区間を設定することで急激なテンポ変化を回避する。遷移テンポ  $T_{tgt}$  を用いたデュアルテンポ調整のテンポ遷移イメージを図6に示す。

遷移テンポ  $T_{tgt}$  は、 $Ldc(f(T_{opt|m_i}, T_{tgt})) = Ldc(f(T_{m_{i+1}}, T_{tgt}))$  を満たすテンポを表しており、以下の式で表せる。

$$T_{tgt} = \frac{(a-b)T_{low} + \sqrt{(a-b)^2 T_{low}^2 + 4abT_{high}T_{low}}}{2a} \quad (5)$$

ここで、 $T_{low}$  は、 $T_{m_i}, T_{m_{i+1}}$  において値の小さい方の BPM を表し、 $T_{high}$  は値の大きい方の BPM を表している。楽曲  $m_i, m_{i+1}$  に対して違和感指標を最小化することが可能な最適テンポ調整係数  $f_{opt}$  を以下の式に基づいて求める。

$$f_{opt}(T_{opt|m_i}, T_{tgt}) = \frac{T_{tgt}}{T_{opt|m_i}}, \quad f_{opt}(T_{m_{i+1}}, T_{tgt}) = \frac{T_{tgt}}{T_{m_{i+1}}} \quad (6)$$

最終的に楽曲  $m_i, m_{i+1}$  それぞれの最適テンポ調整係数  $f_{opt}(T_{opt|m_i}, T_{tgt})$  および  $f_{opt}(T_{m_{i+1}}, T_{tgt})$  を用いてテンポを調整する。

### 5.4 再生順序計算

本処理では、再生対象楽曲群を DJ ミックスに適した順序に最適化する方法について述べる。再生対象楽曲群 ( $N$  曲) に対し、ある順序  $P = \{m_1, m_2, \dots, m_N\}$  が与えられたとき、

$P$  における違和感指標の総計  $S_{dc|P}$  は各楽曲に対する最適テンポ調整係数を用いることで、以下のように表すことができる。

$$S_{dc|P} = \sum_{i=1}^{N-1} (L_{dc}(f_{opt}(T_{opt|m_i}, T_{tgt})) + L_{dc}(f_{opt}(T_{m_{i+1}}, T_{tgt}))) \quad (7)$$

$S_{dc|P}$  は順序  $P$  における楽曲遷移部分の違和感指標の総計を表現しており、順序  $P$  において  $S_{dc|P}$  を最小化することで再生楽曲全体に対する違和感指標を軽減させる。

最終的に、 $S_{dc|P}$  が最小となる順列  $P_{min}$  を再生順序として用いる。 $S_{dc|P_{min}}$  が複数の解を保持する場合には、 $P_{min}$  を構成する楽曲間に対する違和感指標の和 ( $L_{dc}(f_{opt}(T_{opt|m_i}, T_{tgt})) + L_{dc}(f_{opt}(T_{m_{i+1}}, T_{tgt})))$  から得られる分散値が最小となる  $P_{min}$  を利用する。なお、本処理の計算時間は、再生対象楽曲数  $N$  に対して指数関数的に増加する。したがって、 $N$  が大きい場合は、順序計算対象となる楽曲を絞り込む必要がある。たとえば、5.2 節で得られた類似度順に楽曲をソートし、上位から順番に順序計算対象楽曲を所定の数だけ選択する方法などが適用できる。

### 5.5 ビート調整・クロスフェード

本節ではミックス対象楽曲どうしのビート調整およびクロスフェード区間の決定について説明する。DJ ミックスにおいて強ビートどうしを重ね合わせることは重要であると考えられるため、本処理ではテンポ・ビート推定処理によって得られたビート時刻の周波数強度 (ビート強度) の相互相関を利用して、強ビートどうしまたは、弱ビートどうしを重ね合わせる。

楽曲  $m_i, m_{i+1}$  における  $k$  番目のビート強度を  $Pw_{m_i}(k), Pw_{m_{i+1}}(k)$  と表現すると、楽曲  $m_i, m_{i+1}$  間のビート強度の相互相関は以下のように表現できる。

$$Cr(\tau) = \frac{\sum_{k=1}^{\tau} (Pw_{m_i}(BN_{m_i} - k + 1)Pw_{m_{i+1}}(k))}{\tau} \quad (8)$$

ここで、 $\tau$  は重ね合わされたビートの数を表しており、 $BN_{m_i}$  は楽曲  $m_i$  のビート数を表している。ビート強度は各ビート時刻における近隣区間 ( $\pm 50$  ms) において、ローパスフィルタ適用後の信号に対して FFT を適用することで求める。ローパスフィルタはカットオフ周波数 1,500 Hz とし、20 次の FIR フィルタを適用する。最終的に式 (8) において、

$$\tau_{max} = \operatorname{argmax}(Cr(\tau)) \quad (9)$$

を満たすように楽曲  $m_i, m_{i+1}$  のビート時刻を重ね合わせ、重複する区間に対しクロスフェード処理を適用する。

## 6. システム評価実験

本章では提案システムの有効性を検証するため、以下 2 項目の評価実験について記述する。

- 提案システムにおける聴取者違和感指標の軽減量評価実験
- 提案システムにおける被験者主観評価実験

実験 1 では、最適テンポを用いたデュアルテンポ調整処理と、再生順序計算処理の有効性を、既存 DJ ミックスと再生順序計算処理を適用しない場合の違和感指標値の比較により検証する。実験 2 では、既存 DJ ミックスを利用した既存システムと、5 章に記載した提案システムの両システムにより生成された DJ ミックス音源に対する被験者評価に基づいて提案システムの有効性を検証する。

### 6.1 実験 1

#### 6.1.1 実験方法

本節では文献 8) に記載の音楽検索システムによって得られた検索結果楽曲群に対し、検索結果をそのまま利用した再生リスト (検索結果リスト) と、5.4 節に記載の順序計算処理により作成した再生リスト (最適化リスト) と、既存および提案 DJ ミックスのすべての組合せを利用して違和感指標の総計を比較する。組合せ情報および以降での名称を表 2 に示す。

DJ ミックス方式に提案 DJ ミックスを用いたリスト (提案検索リスト・提案最適リスト) では式 (7) により違和感指標の総計  $S_{dc|P_{min}}$  を計算した。既存 DJ ミックスを用いたリスト (既存検索リスト・既存最適リスト) では、4.1.1 項に記載の既存 DJ ミックスを適用し、前楽曲の BPM を後続する楽曲の BPM に合致させることでテンポ調整係数を求め、式 (7) に基づいて違和感指標の総計を計算した。既存検索リストでは、BPM の小さい楽曲を他方の BPM に合致させるテンポ調整係数を計算し、式 (2) に基づいて計算した違和感指標の総和となる。既存最適リストでは、既存検索リストにおける違和感指標の総和が最小となる順

表 2 再生リストと DJ ミックス方式組合せ表

Table 2 Combination of playlist and DJ mixing method.

リスト名称	再生リスト	DJ ミックス方式
既存検索リスト	検索結果リスト	既存 DJ ミックス
提案検索リスト	検索結果リスト	提案 DJ ミックス
既存最適リスト	最適化リスト	既存 DJ ミックス
提案最適リスト	最適化リスト	提案 DJ ミックス

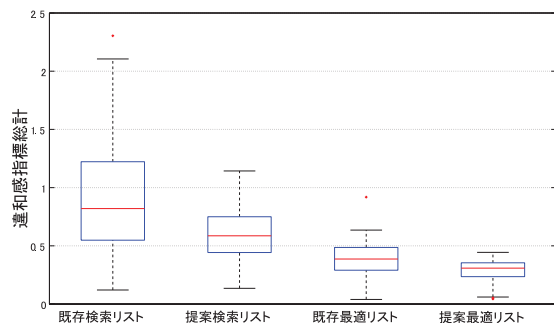


図 7 再生リストに対する違和感指標の総計のボックスプロット (平均値, 200 サンプル), 左から既存検索リスト, 提案検索リスト, 既存最適リスト, 提案最適リスト

Fig. 7 Boxplot of amount of  $L_{dc}$  in naive and proposed DJ mixing in each playlist.

序を計算した。これは再生対象楽曲群の BPM の昇順となる。

### 6.1.2 実験データ

利用した楽曲データベースは, Jamendo<sup>\*1</sup>より収集した 1,434 曲を用いた。実際に DJ ミックスを適用する際には, サビ区間を含む 30 秒間をすべてのデータから抽出し, *BeatRoot* を用いてビート時刻および BPM 情報をあらかじめ抽出した。検索要求として前記楽曲データ群からランダムに 1 曲選択し, 検索結果上位 5 楽曲をプレイリストとして計 200 プレイリストを作成した。すべてのプレイリストに対して表 2 に記載のすべての組合せを用いて違和感指標の総計を計算した。

### 6.2 実験結果・検証

図 7 に表 2 を適用したプレイリストの違和感指標の総計をボックスプロットとして示す。図 7 では, 各項目について最小値, 第 1 四分位数, 中央値, 第 3 四分位数, 最大値を示している。図からも明らかとなっており, 検索結果をそのまま利用した場合に比べ, 順序計算処理を適用することでリストの違和感指標の総計を大きく軽減可能であることが分かる。各リストにおける違和感指標の相加平均は, 図 7 の左から 0.88, 0.58, 0.38, 0.29 となっており, 既存検索リストの違和感指標総計の最大値は 2.30, 既存最適リストの総計の最大値は 0.91 であった。また, 各リストで使用された最大のテンポ調整係数は左から 2.07, 1.21, 1.94, 1.20 となっており, 再生順序計算処理を適用することで, 少ないテンポ調整で DJ ミックス

を適用することが可能である。

また, DJ ミックス方式に着目して比較すると, 提案 DJ ミックスを適用することで, さらに違和感指標の総計を軽減できることが分かる。この結果から音楽検索結果として得られたリストに対して最適テンポを用いたデュアルテンポ調整と, 順序計算処理を適用することで, 検索結果リストおよび既存 DJ ミックスを利用する場合に比べて聴取者の違和感を軽減することが期待でき, より自然な DJ ミックスを提供可能であるといえる。

## 6.3 実験 2

### 6.3.1 実験方法

本実験では, 既存および提案 DJ ミックスを適用した音源を被験者に聴取してもらい主観評価を付与する。初めにクエリ楽曲をランダムに選択し, 楽曲データベースからクエリ楽曲を含む 5 曲を再生対象楽曲として抽出した。提案システムでは 5 章に記載の処理を適用し, 既存システムでは, 式 (1) に基づく DJ ミックスを適用した。システムごとに 6 音源を作成し被験者聴取用の音源とした。楽曲データベースは 6.1.2 項と同様のものを利用した。

なお, 既存システムにおいても楽曲順序計算およびビート調整処理を適用することとし, 楽曲順序計算処理では式 (7) を利用する。また, 最適テンポ調整係数の代替として式 (1) から得られるテンポ調整係数に基づいて聴取者違和感の楽曲順序を計算した。ビート調整処理は 5.5 節に記載の処理を適用した。予備テンポ調整区間は提案, 既存システムともに 5 秒と設定し, テンポ調整は *SoundTouch* ライブラリを用いて音程を変えさせずに速度のみを変化させた。

### 6.3.2 被験者および評価指標

被験者はふだん音楽に親しんでいる 20 代から 50 代までの 27 名を対象とし, すべての音源に対する 5 段階評価を収集した。評価項目として, 自然さ (5: 違和感なし, 1: 違和感あり), リズム保持性 (5: リズムを維持できる, 1: リズムを維持できない), 面白さ (5: 面白い, 1: 面白くない) の 3 項目を設定した。表 3 に被験者に提示した評価基準を示す。また, 各音源に対する評価の理由を自由回答で収集した。

### 6.3.3 実験結果・考察

被験者によって付与された評価データを基に提案および既存システムの項目ごとに被験者評価の平均値を求めた。図 8 に結果を示す。図 8 から明らかとなっており, 既存システムを用いた音源に比べ, 提案システムを用いた音源はすべての項目において高い評価を得た。 $t$  検定により各評価項目の評価値の差は統計的に有意であることが確認できた ( $p < 0.001$ ,  $\alpha = 0.01$ )。

\*1 <http://jamendo.com/>



表 3 被験者主観評価基準

Table 3 Subjective evaluation measures for subjects.

評価値	違和感の有無 (自然さ)	リズム感保持の可否 (リズム感)	面白さ
5	まったくない	完全にできる	非常に面白い
4	ほぼない	ほぼできる	やや面白い
3	問題ない	問題ない	普通
2	ややある	ほぼできない	やや面白くない
1	とてもある	まったくできない	面白くない

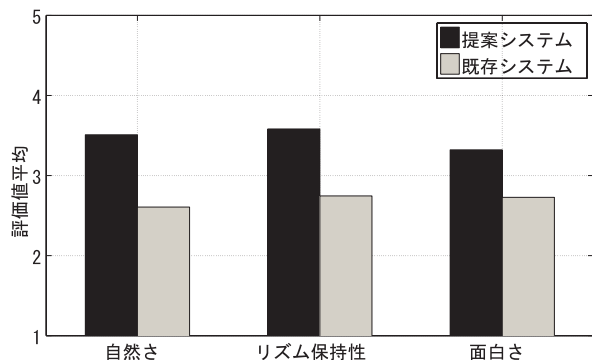


図 8 提案および既存システムにおける被験者主観評価実験結果 (平均)

Fig. 8 Average of user ratings in proposed and naive DJ mixing systems.

次に、各システムにおいて作成された実験用音源で適用されたテンポ調整係数を調査した。図 9 に提案および既存システムで用いられたテンポ調整係数の分布を示す。図 9 から明らかとなり、提案システムを用いた場合のテンポ調整係数は、既存システムにおけるテンポ調整係数に比べて、オリジナルにより近いテンポを利用して DJ ミックスを適用することができる。4.3 節で記載した平均適用範囲 (点線) と、厳密適用範囲 (実線) を図 9 に表記したところ、既存システムでは平均適用範囲内に収まるテンポ調整係数は 46%、厳密適用範囲内に収まるテンポ調整係数は 21%であった。一方、提案システムにおけるテンポ調整係数は、すべて平均適用範囲内に収まっている。また、厳密適用範囲においても 50%のテンポ調整係数が収まっており、自然さ評価結果の向上につながったと考えられる。

さらなる分析をするために、各項目・各音源に対してユーザ評価値の分析を行った。相関分析の結果、「自然さ」と「リズム保持性」については強い正の相関 (相関値: 0.934) が確認できた。提案システムではテンポ調整係数を平均適用範囲内でテンポを調整することが

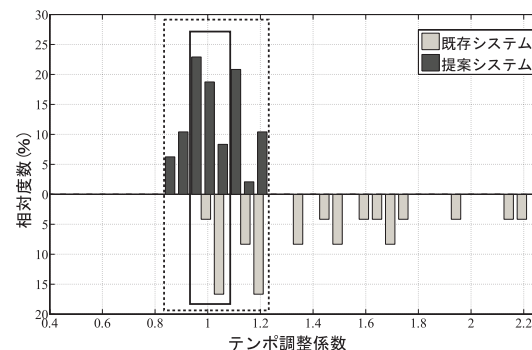


図 9 提案および既存システムにおけるテンポ調整係数分布 (相対度数)

Fig. 9 Distribution of relative frequency of tempo adjustment factors in song-to-song transitions of proposed and naive DJ mixing system.

きるため、既存システムに比べてテンポの急激な変化を抑えることができる。ゆえに、クロスフェード区間においてリズム保持性評価が向上し、それとともに自然さ評価が向上したと考えられる。実験結果より、提案システムが既存システムに比べ聴取者の違和感を軽減可能であるといえ、提案システムの有効性が示された。

一方で、「自然さ」の項目において提案システムを用いた音源に対する評価値は全体的に上回っていたが、既存システムの評価平均を下回る音源が 1 音源存在していた。このときの提案システムの違和感指標は既存システムの値よりも低い値を保持していたが、音源内に存在する 2 カ所の楽曲遷移部で楽曲ビート調整処理がうまく動作せず弱ビートと強ビートが重ね合わせられていた。これにより該当区間のリズムを保持することができず評価が下がったと考えられる。また、「面白さ」の項目において、既存システム評価平均が提案システム評価平均を上回る被験者が 27 名中 6 名存在した。これらの被験者によるアンケート回答を分析した結果、テンポ調整量の増加に由来した急激なテンポ変化が顕著になる際にエフェクト処理が適用されたような音源となること新鮮に感じられたことがアンケート記述より推測できる。本稿ではテンポ調整量に着目して違和感指標を定義したが、テンポ調整量以外にも聴取者違和感の要因や考慮すべき項目が実験で示唆された。今後はテンポ調整量以外の聴取者違和感の要因を調査するとともに、個人的嗜好や意外性などを考慮することが可能な DJ ミックスシステムを検討することで、評価値が改善可能であることが示唆された。

## 7. ま と め

本稿では、連続する楽曲のテンポを合致させ、ビート時刻を重ね合わせることで複数の楽曲を連続的に再生することが可能な自動 DJ ミックスシステムを提案した。提案システムでは既存技術が考慮していなかった、楽曲再生順序、自動テンポ推定誤りへの対処、連続再生時の違和感の軽減の 3 項目について改善を図った。具体的には、DJ ミックス再生時のテンポ変化が聴取者に与える影響を予備実験によって検証し、予備実験の結果に基づいて聴取者違和感を測定する指標を定義した。前記指標値を最小限に抑えることが可能なテンポ調整方式と楽曲順序計算方法をシステムに導入することで、聴取者違和感を軽減させる自動 DJ ミックスシステムを構築した。さらに、連続する楽曲間のテンポ調整において、テンポ推定結果がオクターブ誤りであった場合にもロバストな最適テンポ調整係数の計算方法を導入した。

評価実験では、楽曲順序計算時の聴取者違和感指標の総計を測定し、提案システムの楽曲順序計算方法が既存技術に比べて聴取者違和感指標の総計を軽減可能であることを示した。さらに主観評価実験において、既存システムと提案システムによって作成した DJ ミックス音源の評価を実施した。評価実験により提案システムが既存システムに比べて、聴取者に自然な DJ ミックスを提供可能であることを示した。

謝辞 本稿を作成するにあたり、日頃ご指導いただく KDDI 研究所伊藤泰彦代表取締役会長、秋葉重幸代表取締役所長、中島康之副所長および菅谷史昭執行役員に深く感謝いたします。

## 参 考 文 献

- 1) Anderson, M., Ball, M., Boley, H., Greene, S., Howse, N., Lemire, D. and Mcgrath, S.: RACOPI: A rule-applying collaborative filtering system, *Proc. COLA '03* (2003).
- 2) Shardanand, U. and Maes, P.: Social information filtering: Algorithms for automating “word of mouth”, *Proc. CHI '95*, pp.210–217 (1995).
- 3) Cohen, W.W. and Fan, W.: Web-collaborative filtering: Recommending music by crawling the web, *Proc. WWW 2000* (2000).
- 4) Uitdenbogerd, A. and van Schyndel, R.: A review of factors affecting music recommender success, *Proc. ISMIR*, pp.204–208 (2002).
- 5) Alghoniemy, M. and Tewfik, A.H.: A network flow model for playlist generation, *Proc. ICME 2001* (2001).
- 6) Pauws, S. and Eggen, B.: PATS: Realization and user evaluation of an automatic playlist generator, *Proc. ISMIR 2002*, pp.222–230 (2002).
- 7) Foote, J.: Content-based retrieval of music and audio, *Proc. SPIE*, Vol.3229, pp.138–147 (1997).
- 8) Hoashi, K., Matsumoto, K. and Inoue, N.: Personalization of user profiles for content-based music retrieval based on relevance feedback, *Proc. ACM Multimedia 2003*, pp.110–119 (2003).
- 9) Yoshii, K., Goto, M., Komatani, K., Ogata, T. and Okuno, H.: Improving efficiency and scalability of model-based music recommender system based on incremental training, *Proc. ISMIR*, Vienna, pp.89–94 (Sep. 2007).
- 10) Goto, M.: An audio-based real-time beat tracking system for music with or without drum-sounds, *J. New Music Res.*, Vol.30, No.2, pp.159–171 (2001).
- 11) Scheirer, E.: Tempo and beat analysis of acoustic musical signals, *J. Acoust. Soc. Amer.*, Vol.103, No.1, pp.588–601 (1998).
- 12) Klapuri, A., Eronen, A. and Astola, J.: Analysis of the meter of acoustic musical signals, *IEEE Trans. Speech Audio Process.*, Vol.14, pp.342–355 (2006).
- 13) Alonso, M., David, B. and Richard, G.: Tempo and beat estimation of musical signals, *Proc. Int. Conf. Music Information Retrieval*, pp.158–163 (2004).
- 14) Dixon, S.: Automatic extraction of tempo and beat from expressive performances, *J. New Music Res.*, Vol.30, No.1, pp.39–58 (2001).
- 15) Ishizaki, H., Hoashi, K. and Takishima, Y.: Autocorrelation-based beat estimation adaptive to drastic tempo change in a song, *Proc. ICME 2009*, pp.478–481 (2009).
- 16) 宮島 靖: Music Mosaic Generator: 高精度時系列メタデータを利用した音楽リミックスシステム, *Proc. WISS2007 (オンライン)*, 入手先 <http://www.wiss.org/WISS2007Proceedings/> (2007).
- 17) 井上亮文, 天野玲奈, 備瀬翔平, 市村 哲, 松下 温: 携帯型音楽プレイヤーのためのコンテンツ再生・配信方法, 情報処理学会研究報告, 2006-DBS-138, pp.133–138 (2006).
- 18) 沖野修也: DJ 選曲術, 東京書籍 (2005).
- 19) Gouyon, F., Klapuri, A., Dixon, S., Alonso, M., Tzanetakis, G., Uhle, C. and Cano, P.: An experimental comparison of audio tempo induction algorithms, *IEEE Trans. Audio, Speech and Lang. Process.*, pp.1832–1844 (Sep. 2006).
- 20) Roucos, S. and Wilgus, A.M.: High quality time-scale modification for speech, *IEEE Int. Conf. Acoust., Speech, Signal Processing*, Tampa, FL, pp.493–496 (Mar. 1985).
- 21) Goto, M., Hashiguchi, H., Nishimura, T. and Oka, R.: RWC Music Database: Popular, Classical, and Jazz Music Databases, *Proc. ISMIR 2002*, pp.287–288 (Oct. 2002).

(平成 22 年 3 月 18 日受付)

(平成 22 年 11 月 5 日採録)



石先 広海 (正会員)

2004 年早稲田大学理工学部電子情報通信学科卒業。2006 年同大学大学院修士課程修了。同年 KDDI 株式会社入社。現在、株式会社 KDDI 研究所知能メディアグループ研究員。この間、音楽情報処理の研究に従事。FIT2009 ヤングリサーチャー賞受賞。電子情報通信学会会員。



帆足啓一郎 (正会員)

1995 年早稲田大学理工学部情報学科卒業。1997 年同大学大学院修士課程修了。同年国際電信電話株式会社 (現 KDDI 株式会社) 入社。現在、株式会社 KDDI 研究所知能メディアグループ主任研究員。この間、マルチメディア情報検索等の研究に従事。2001~2005 年早稲田大学メディアネットワークセンター非常勤講師。工学博士。FIT2004 ヤングリサーチャー賞受賞。電子情報通信学会、ACM 各会員。



滝嶋 康弘

1986 年東京大学工学部電気工学科卒業。1988 年同大学大学院修士課程修了。同年、国際電信電話株式会社 (現 KDDI 株式会社) 入社。現在、株式会社 KDDI 研究所執行役員。この間、動画像符号化、画像処理、情報理論の研究開発に従事。工学博士。電子情報通信学会、映像情報メディア学会、画像電子学会各会員。