

自動 DJ ミックス再生における楽曲速度変化に伴う 聴取者違和感指標に関する検討

A study on measurement function of user comfortness according to song tempo change for DJ mixing.

石先 広海 † 帆足 啓一郎 † 滝嶋 康弘 †
Hiromi Ishizaki † Keiichiro Hoashi † Yasuhiro Takishima †

1 はじめに

近年の音楽圧縮技術の発展に伴い、現在多くの音楽配信サービスが普及している。これらサービスを利用して個人が容易に膨大な楽曲コレクションを構築することが可能になっている。一方で、個人が楽曲を聴取する方法としては、プレイリスト再生やランダム再生が主流である。前述の再生方法を用いて楽曲を連続再生させる場合、楽曲間に無音が生じるなど、個人が楽曲を聴取している際の“ノリ”が損なわれる事がある。本稿では、聴取者の“ノリ”を損なわずに楽曲を連続再生する DJ(ディスクジョッキー)に着目し、より自然な連続再生を実現するための聴取者違和感指標の定義について検討する。

2 関連研究

一般的に、聴取者の“ノリ”の損失を防ぐため、クラブハウスやディスコ等では音楽を途切れさせない事が重要であるとされており、実際に DJ は楽曲同士のリズムを同期させたまま複数の楽曲を連続させてなめらかな楽曲再生を実現している。

DJ のような楽曲再生を自動化するために井上らは再生対象楽曲におけるテンポを一致させ、ビート位置を調整したクロスフェード再生 [1] を用いることで、擬似的に DJ のような滑らかな連続再生 (以下, DJ ミックス) を実現している。また、筆者らは DJ ミックスの際の音質劣化を軽減するために、再生対象となる楽曲同士のテンポオクターブ関係に着目してテンポ調整率を算出する事でテンポ差が大きい場合の音質劣化を軽減している [2]。

3 課題

DJ ミックスにおける聴取者違和感の度合いを推し量ることは、聴取者により自然な連続再生を提供するための重要な要素であると考えられるが、これまでに楽曲のテンポ変化に対する聴取者違和感を推し量るための指標は存在していない。

文献 [1] では、ミックス対象となる楽曲のテンポ差が 30% を越える場合には、通常の連続再生を適用している。これは明らかに聴取者に不快を与えると思われるテンポ変化を制限するための閾値であり、聴取者のテンポ変化に対する許容範囲を反映させたものではないため、聴取者は違和感を覚える可能性がある。

文献 [2] では、テンポ調整率に基づいた評価指標を利用しており、評価実験によって DJ ミックス音源におけ

るテンポ調整率と被験者評価に相関がある事を示唆している。評価指標 V は、テンポ調整率 f に基づいて以下のように定義している。

$$V = \begin{cases} f - 1 & f > 1 \\ 0 & f = 1 \\ 1/f - 1 & f < 1 \end{cases} \quad (1)$$

V はオリジナルのテンポと、テンポ調整後のテンポの差異として理解することができる。

この様にテンポ調整率に対する聴取者違和感を推し量ることで、テンポ変化を柔軟に制御することで高精度な連続再生を実現する事ができると考えられる。

4 実験

本章では文献 [2] の結果に基づき、テンポ調整率と聴取者違和感に相関があるとして、被験者聴取実験結果から聴取者違和感とテンポ調整率との関係性を調査する。

第 3 章にて述べた様に、DJ ミックス再生時の聴取者違和感を回避するためには聴取者違和感を推し量るための評価指標を定義する必要がある。そこで本実験では特にテンポを上げる場合と下げる場合に対する聴取者違和感の傾向について調査を行った。

4.1 実験データ

本実験では、実験データとして RWC 研究用楽曲データベース [5] から 18 曲選択した。全ての楽曲に対し、徐々にテンポを上げた場合の音源と、徐々にテンポを下げた場合の 2 種類を用意した。

テンポ変化は *SoundTouch* ライブラリ, *SOLA* を用いて音の高さを変化させずに楽曲速度のみを変化させた。*SoundTouch* ライブラリは *SOLA* に比べ、高品質なテンポ変化を可能にするライブラリである。表 1 に 1 楽曲から作成した実験データの詳細を示す。最終的に合計 72 データを作成した。

表 1 実験データパラメータ詳細

テンポ変化方式	テンポ変化方向
<i>SoundTouch</i>	上昇
<i>SoundTouch</i>	下降
<i>SOLA</i>	上昇
<i>SOLA</i>	下降

対象となる音源のテンポ変化は以下の式で表せる。

$$T_{adj} = f \times T \quad (2)$$

ここで、 T_{adj} はテンポ変化後のテンポ、 f はテンポ調整率、 T は対象音源のテンポを表す。対象音源に対して、

† KDDI 研究所, KDDI R&D Laboratories, Inc.

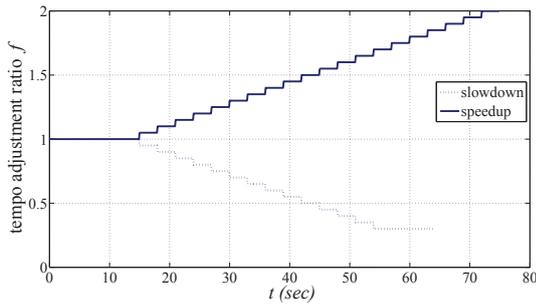


図1 実験音源時刻に対するテンポ変化率 f :テンポ上昇 (speedup), 下降 (slowdown) .

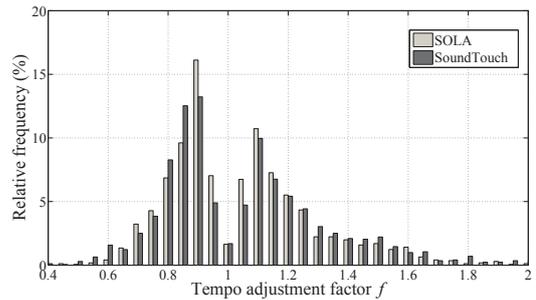


図2 テンポ変化方式 (SoundTouch, SOLA) におけるテンポ変化率 f に対する聴取者違和感収集結果 (相対度数) .

楽曲開始から 15 秒を起点として、テンポ調整率 f を 1.00 から 2.00, 1.00 から 0.30 として、0.05 刻みで変化させた。また、各テンポ調整率区間は 3 秒となるようにテンポ変化を適用している。図 1 に対象音源時刻に対するテンポ変化率の変化を示す。

被験者は一般から収集した 96 名で、一人当たり 36 音源を聴取してもらい、各実験データに対して違和感を覚える時刻を入力してもらった。(有効データ数: 計 3422 サンプル)

4.2 実験結果・考察

各音源において、被験者入力時刻からテンポ変化率 f を収集した結果を図 2 に示す。実験結果では $f > 1$, $f < 1$ 区間におけるピーク位置はそれぞれ 1.10, 0.90 となり、両テンポ変化方式において同じであった。

図 2 では、ピークを示すテンポ変化率は同じであるが、 $f < 1$ 区間におけるピークの度数が $f > 1$ 区間におけるピークよりも大きい事が分かる。各区間における平均および分散を表 2 に示す。

表 2 において、オリジナルテンポに対する差異に対して、 $f > 1$ 区間と $f < 1$ 区間における平均の違いを検証するために、式 1 を用いて V に変換し検定を実施した。 t 検定 ($p < 0.0001$, $\alpha = 0.01$) 及び、 F 検定 ($p < 0.0001$, $\alpha = 0.01$) によって平均及び分散の差は統計的に有意である事を確認した。

このことから、テンポを下げる場合の方がテンポを上げる場合よりも聴取者は違和感を覚えやすい事が分かる。

表 2 テンポ変化方式における違和感入力テンポ変化率平均と分散

テンポ変化方式	項目	$f > 1$	$f < 1$
SoundTouch SOLA	平均	1.227	0.852
		1.226	0.852
SoundTouch SOLA	分散	0.0415	0.0092
		0.0348	0.0074

4.3 定義

本節では、前述の実験結果に基づいて聴取者違和感を推し量る指標の定義を図る。本節では文献 [2] の結果に基づいて、テンポ変化対象楽曲のオリジナルテンポと変化後のテンポの差が聴取者違和感に関連していると仮定した。実験結果から明らかになった通り、聴取者は $f > 1$ 区間よりも $f < 1$ 区間で違和感を覚えやすいため、

式 1 に対して $f > 1$ の場合と $f < 1$ の場合で重みを付ける。具体的には重み係数をそれぞれ a, b として以下のように定義した。

$$V_{dc} = \begin{cases} a(f - 1) & f > 1 \\ 0 & f = 1 \\ b(1/f - 1) & f < 1 \end{cases} \quad (3)$$

ここで、 a, b は表 2 で与えられる平均テンポ変化率から求められる V_{dc} が $f > 1, f < 1$ において等しくなるように設定した。具体的には、 $a = 0.765, b = 1.000$ と設定した。さらに、表 2 から、 $0.852 < f < 1.227$ の範囲を越えてテンポ変化を適用した場合、多くの聴取者が違和感を覚える事が分かる。前述の平均的テンポ変化範囲を越える場合には、通常の再生を適用するなど違和感を回避する判断指標として利用する事が可能である。

5 まとめ

本稿では、DJ のような自然な連続再生を実現するために、聴取者のテンポ変化に対する違和感の程度を推し量る指標を定義する事を目標とし、実際にテンポ変化を適用した音源を被験者群に聴取してもらうことで、テンポを上げた場合とテンポを下げた場合における聴取者違和感の差異を調査した。実験結果からテンポを上げる場合に比べ、テンポを下げる場合の方が聴取者は違和感を覚え易い事が判明し、実験結果に基づいて違和感指標および、平均的なテンポ変化範囲を定義した。引き続き、本指標の妥当性及び、有効性検証を進めていく。

参考文献

- [1] 井上他, “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.
- [2] 石先他, “音質劣化を考慮した音楽自動リミックス手法”, 情報処理学会研究報告, 2008-EC-009, pp.43-50.
- [3] SoundTouch Library:
<http://www.surina.net/soundtouch/>
- [4] S. Roucos et al. “High quality time-scale modification for speech,” IEEE ICASSP, 1985, pp.493-496.
- [5] M. Goto, et al. : “RWC Music Database: Popular, Classical, and Jazz Music Databases,” Proc. of ISMIR 2002, pp.287-288.