

音響信号に対応した フットスイッチ式自動伴奏システムの開発及び評価

安部 綾太[†] 木村 敏幸[‡]

[†] 東北学院大学工学研究科電気工学専攻 〒985-8537 宮城県多賀城市中央 1-13-1

[‡] 東北学院大学工学部情報基盤工学科 〒985-8537 宮城県多賀城市中央 1-13-1

E-mail: [†] s1994201@g.tohoku-gakuin.ac.jp, [‡] t-kimura@m.ieice.org

あらまし 我々はコンピュータが人間のフットスイッチの操作に合わせて演奏の開始や停止を行う自動伴奏システムの実現を目指し、これまでにフットスイッチと MIDI を用いたシステムを開発してきた。本研究では音響信号を用いたシステムを開発するため、音響信号のループ再生条件とフットスイッチと組み合わせたシステムを主観評価実験によって評価した。その結果、実験参加者は曲の終わりにおいてタイミングがずれているのにも関わらず、タイミングが合っていると評価した。

キーワード 自動伴奏, 音響信号, ループ再生, フットスイッチ

Development and Evaluation of Automatic Accompaniment System Using Foot Switch for Acoustical Signals

Ryota ABE[†] and Toshiyuki KIMURA[‡]

[†] Graduate School of Engineering, Tohoku Gakuin University 1-13-1 Chuo, Tagajo-shi, Miyagi, 985-8537 Japan

[‡] Faculty of Engineering, Tohoku Gakuin University 1-13-1 Chuo, Tagajo-shi, Miyagi, 985-8537 Japan

E-mail: [†] s1994201@g.tohoku-gakuin.ac.jp, [‡] t-kimura@m.ieice.org

Abstract In order to realize the automatic accompaniment system in which a computer starts and stops a performance according to the operation of a performer's foot switch, we have developed the system using a foot switch and MIDI. In this study, in order to develop the system using acoustical signals, the loop-playing condition of acoustical signals and the system with a foot switch were evaluated by subjective experiments. As a result, participants evaluated that the timing was correct even though the timing of participants was varied widely at the end of songs.

Keywords Automatic Accompaniment, Acoustical Signal, Loop-playing, Foot Switch

1. はじめに

1.1. 研究背景

カラオケのような従来の伴奏システムでは人間が演奏のテンポやタイミングなどを正確に覚えていないとコンピュータの演奏とずれてしまい、満足のいく演奏ができないことがある。人間の演奏に追従するように伴奏を行う自動伴奏システムの研究は盛んに行われてきたが、テンポのない箇所（特に曲の終わり）に着目したシステムの研究は現状少ない。

そこで、我々はコンピュータが人間のフットスイッチの操作に合わせて演奏の開始や停止を行う自動伴奏システムの実現を目指し、これまでにフットスイッチと MIDI (Musical Instrument Digital Interface) のホールドオン機能を用いることで人が操作をするまで音が鳴り続ける自動伴奏システムを開発してきた[1,2]。しかしながら、MIDI は電子楽器の演奏データを機器間で伝

達するための規格で、音そのものではなく音の情報が入ったデータであるために、人間の演奏や合図に合わせて柔軟に音の速さや長さを変えることが容易である。一方、音響信号は初めから音が決まっていることから MIDI のように柔軟に対応させることが難しいため、音響信号を用いた自動伴奏システムの研究は十分にされていない。

1.2. 目的

本研究では、これまでに MIDI を用いて開発してきたシステムを、音響信号に対応させることを目指す[2,3]。特に MIDI のホールドオン機能と同様の動作を音響信号で再現することに重点を置き、音響信号のループ再生について深く検討する。また、開発したシステムの評価実験を行い、システムの有効性を示す。

1.3. 自動伴奏システムの原理

Fig. 1 に自動伴奏システムの構成図を示す。あらか



図 1 自動伴奏システムの構成図
Fig. 1 Configuration of automatic accompaniment system

はじめ演奏者の足元にフットスイッチを配置する。本研究で用いたフットスイッチにはスイッチが3つ付いており、左から順に h, b, s のキーと対応している。スイッチを踏むことで、それぞれのキーに割り当てた動作（曲の再生や停止）が実行される。演奏者が演奏しながらフットスイッチを操作すると、コンピュータが演奏者のフットスイッチ操作に合わせて伴奏するので、これまでの自動伴奏システムでは実現できなかったライブ演奏（例えば、曲の終わりで人間の合図に合わせてコンピュータが同時に演奏を停止する）が可能になると期待される。

2. 音響信号のループ再生条件

音響信号は MIDI のようにファイルの最後や最初にホールドオン、ホールドオフの MIDI コントロールチェンジメッセージを入れることができないため、オーディオファイルの最後の音を鳴り続けさせることができない。従って、MIDI におけるホールドオンの状態を再現するためにはオーディオファイルの再生後にオー



図 2 自動伴奏システムの開発環境
Fig. 2 Development environment of automatic accompaniment system

ディオファイルの最後のごく短い時間をループ再生させることが必要となる。

しかしながら、ループ再生をする際には、ループの終わりと始まりが切り替わる際に急激に時間波形が変化することによってクリックノイズが発生するという問題点がある。そこで、本章では音響信号を用いた自動伴奏システムを開発する前段階として、クリックノイズが発生せずループ再生される条件を検討し、どの条件が自動伴奏システムに適用可能かを確認するために評価実験を実施する。

2.1. 開発環境

Fig. 2 に音響信号を用いた自動伴奏システムの開発環境を示す。

- ・ OS : Windows 10 Pro
- ・ ディスプレイ : iiyama X3272UHS-B1 (左)& MITSUBISHI (右)
- ・ ドライバインタフェース : ASIO4ALL
- ・ フットスイッチ : Route-R RI-FP3MG (左)
- ・ プログラミングソフトウェア : Max [4] Ver. 8.1.3

演奏音は PC 内蔵のサウンドカードを通して再生されるが、その際、サウンドカードを ASIO インターフェースとして認識させるために ASIO4ALL を使用した。

2.2. ループ再生条件

2.2.1. 従来法

まず、Max に装備されている groove オブジェクトを用いたループ再生条件を設定した。groove オブジェクトはバッファに格納されたオーディオファイルをループ再生するためのオブジェクトであり、開始時間と終了時間を指定するとその時間の間だけループ再生される。本研究では、ループの開始時間と終了時間は時間波形が急激に変化しない個所を手動で指定した。

2.2.2. Max を用いたクロスフェード法

しかしながら、従来法によってループの開始時間と終了時間を手動で指定する場合、時間の指定に非常に手間がかかる。そこで、次にクリックノイズが発生しないループ再生方法として、クロスフェードを用いたループ再生条件を設定した。

ループ個所の終わりの数 ms をフェードアウトすると同時にループ個所の始めの数 ms をフェードインすることにより、クリックノイズが発生しないループ再生が期待される。本研究では直線で変化する三角窓、曲線で変化する cos 窓の 2 種類の窓関数を用いて窓掛けを実施し、二つのオーディオファイル再生を順次切り替えることによりループ再生を実現した。

Max を用いた窓掛けでは、Max に装備されている line オブジェクトを用いて窓掛けを実装した。line オブジェクトは音量の目標値と到達時間を入力すると、現在の音量から目標値に向けて到達時間をかけて線形

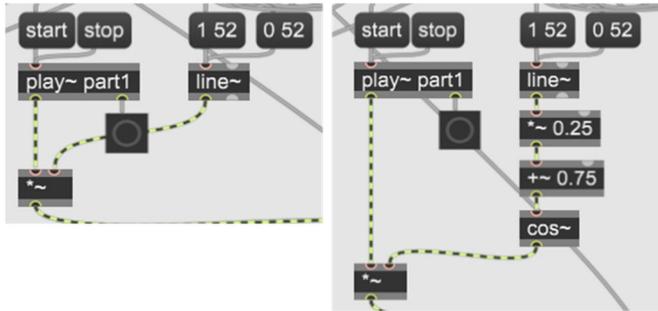


図3 Maxを用いたクロスフェード法
(左：三角窓，右：cos窓)

Fig. 3 Crossfade method using Max

(Left : Triangle window, Right : Cosine window)

に変化するオブジェクトである。従って、入力する音量の値を常に0もしくは1にすることで三角窓を、また、その入力値をcos関数で変換すればcos窓を実装することができる。Fig. 3にMaxを用いたクロスフェード法のプログラムを示す。クロスフェードに要する到達時間は手動で1msずつ調整していき、自分で聴いて最もクリックノイズを感じなかった52msとした。

2.2.3. Octaveを用いたクロスフェード法

Maxを用いたクロスフェード法では、二つのオーディオファイル再生をタイミング良く切り替えることができればクリックノイズが発生しないループ再生を実現できるが、実際にはフェードアウトと同時にフェードインが始まらない。そこで、Octave [5]を用いてオーディオファイルのクロスフェードを波形処理したうえで、grooveオブジェクトで再生するループ再生条件をさらに設定した。

Fig. 4にOctaveを用いたクロスフェード法のイメージ図を示す。Octaveでは1ループの長さにクロスフェ

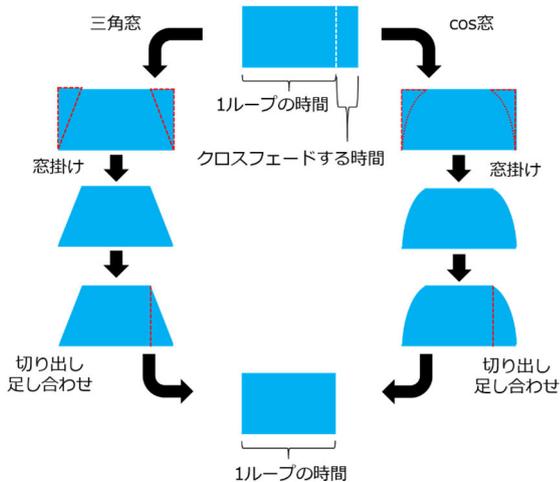


図4 Octaveを用いたクロスフェード法
(左：三角窓，右：cos窓)

Fig. 4 Crossfade method using Octave

(Left : Triangle window, Right : Cosine window)

表1 ループ再生条件の評価実験における
5段階評価

Table 1 Five scales in the evaluation experiment
of loop-playing condition

5	全く気にならない
4	気にならない
3	少し気になる
2	気になる
1	とても気になる

ードする時間を足したオーディオファイルの両端に窓掛けを行い、フェードアウトした部分を切り出してフェードインした部分に足し合わせることで本来の1ループの長さになるよう調整している。

2.3. 評価実験

2.3.1. 実験手順

実験はMaxで実験用のパッチを作成して行った。音源には500Hzのsin波と通常の曲の2種類を用いた。通常の曲としては、星野源のアイデア[6]の曲の最後尾をMIDIとして打ち込み、再生・録音してオーディオファイルとしたものを使用した。また、再生方法は以下の6種類を用いた。

- ・従来法
- ・Maxを用いたクロスフェード法（三角窓）
- ・Maxを用いたクロスフェード法（cos窓）
- ・Octaveを用いたクロスフェード法（三角窓）
- ・Octaveを用いたクロスフェード法（cos窓）
- ・ループなし

ループなしの場合では10秒間のファイルをそのまま再生し、その他の場合では1秒間のファイルをループ再生することにより10秒間再生した。

実験参加者は4名である。実験参加者にはまず実験手順を説明し、ループなしの音源を2回聴いてもらった。ここで実験手順と正確な音を理解できたら実験に入る。試行では、6種類の再生方法のうちのいずれかを用いた音が10秒間再生される。実験参加者は音源が再生された後の5秒の間にTable 1に示す5段階で評価する。1セットは6試行で、セット内には6種類の再生方法が必ず1回ずつ含まれる。再生方法の提示順番はセットごとにランダム化した。本実験では音源ごとに5セットずつ実施した。音源の提示順番は実験参加者ごとにランダム化した。

2.3.2. sin波の実験結果

500Hzのsin波を音源として用いた時の平均評定値をFig. 5に示す。エラーバーは95%信頼区間を表す。Anova4 on the Web [7]で6水準の一要因分散分析を実施したところ、主効果に有意差が見られた。さらに多重比較を実施したところ、Max使用クロスフェード法

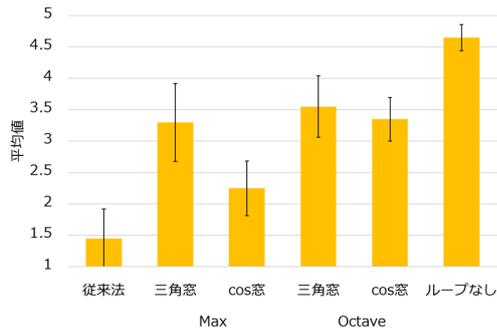


図5 sin波の実験結果

Fig. 5 Experimental result of sin wave

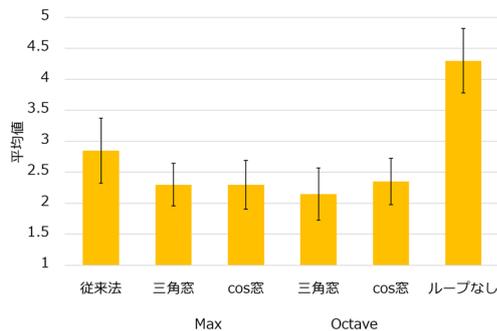


図6 曲の実験結果

Fig. 6 Experimental result of song

(三角窓), Octave 使用クロスフェード法 (三角窓), Octave 使用クロスフェード法 (cos 窓) の 3 種類の再生方法以外の全ての組み合わせにおいて 5%水準で有意差が見られた。

分散分析の結果から, sin 波をループ再生する場合, Octave を用いたクロスフェード法を適用するのが良いと考えられる。但し, Max を用いたクロスフェード法の場合でも三角窓のときは評定値が下がっていないことから, クロスフェードの設定の手間を考慮すると, Max を用いたクロスフェード法において三角窓を掛ける方法が最も妥当であると考えられる。

2.3.3. 曲の実験結果

通常の曲を音源として用いた時の平均評定値を Fig. 6 に示す。エラーバーは 95%信頼区間を表す。Anova4 on the Web で 6 水準の一要因分散分析を実施したところ, 主効果に有意差が見られた。さらに, 多重比較を実施したところ, ループなしとそれ以外のいずれかの組み合わせにおいてのみ 5%水準で有意差が見られ, ループ再生方法同士には有意差が見られなかった。

曲の最後尾を用いた音源では, 10 秒間集中して聴くことでいずれのループ時間でもループ部分に違和感を覚え, 評価としては 1~3 としたためにいずれの方法でも有意差が見られなかったと考えられる。しかしながら, 実際に自動伴奏システムを用いて演奏を行う際には演奏自体に集中するため, ここまでの違和感を覚

えない可能性がある。また, groove オブジェクトにおけるループの開始時間と終了時間は手動で調整する必要があるため, 用いる音源のループにより適した点を設定できれば, MIDI におけるホールドオンの状態により近づけることができると考えられる。

3. 音響信号を用いた自動伴奏システム

第 2 章では, 通常の曲をループ再生する時は手動でループ時間を設定したうえで groove オブジェクトを用いた方が良い可能性が示された。本章では, 第 2 章で得られた結果をもとに音響信号を用いた自動伴奏システムを開発し, 開発したシステムの評価実験を実施する。開発環境は第 2 章と同じである。

3.1. システム概要

従来法では左のスイッチを踏むと 1 曲全てのオーディオファイルが再生される。中央と右のスイッチを踏んだ時には何も起こらない。

提案法では曲が 2 つのオーディオファイルに分割されており, まず左のスイッチを踏むと 1 つ目のオーディオファイルが再生される。次に中央のスイッチを踏むと 2 つ目のオーディオファイルが再生される。最後に右のスイッチを踏むと停止する。その際, 2 つ目のオーディオファイルの最後尾のごく短い時間をループ再生させることにより, 右のスイッチを踏むまで最後の音が鳴り続けているように聴こえる。一方, 提案法にループ再生を導入しない場合は, 2 つ目のオーディオファイルの最後の音を演奏者がフットスイッチを踏むと想定される時間よりも長く演奏しているため, 右のスイッチを踏むまで最後の音が鳴り続けている。

3.2. 評価実験

3.2.1. 音源

曲は星野源のアイデアの第 2 章の評価実験で用いた箇所とは異なる部分を Domino で打ち込み, MIDI で作成したものを再生・録音してオーディオファイルとしたものを用いた。

3.2.2. 再生方法

従来法は 1 曲を 1 つのオーディオファイルで再生する方法である。本実験では, 曲の最後尾が普通の長さのファイルを再生する方法 (以下「従来法 (普通)」) に加えて, 従来法 (普通) のタイミングが分からない状況を模擬するために, 1s 短いファイルを再生する方法 (以下「従来法 (短い)」) と 1s 長いファイルを再生する方法 (以下「従来法 (長い)」) を設定した。

提案法は 1 曲を 2 つのオーディオファイルに分けて再生する方法である。本実験では, 第 2 章における実験の結果に基づき, Max に装備されている groove オブジェクトを用いて曲の最後尾の音を手動で設定した時間だけループ再生する方法 (以下「提案法 (ループ)」) と, あらかじめ曲の最後尾の音を長く演奏したファイ

表 2 自動伴奏システムの評価実験における
5段階評価

Table 2 Five scales in the evaluation experiment
of automatic accompaniment system

5	タイミングが合っていた
4	タイミングが大体合っていた
3	どちらとも言えない
2	タイミングが少しずれていた
1	タイミングがずれていた

ルを再生する方法（以下「提案法（ループなし）」）を設定した。従って、再生方法は従来法3種類、提案法2種類の計5種類である。

用いたオーディオファイルの長さは、従来法（普通）は15172.41 ms、従来法（短い）は14172.41 ms、従来法（長い）は16172.41 ms、提案法の1つ目のファイルは12069.05 ms、2つ目のファイルは3103.36 msである。

3.2.3. 実験手順

まず、従来法（普通）の音源を用いて3つのフットスイッチを踏むタイミングを説明する。実験参加者が理解でき次第演奏に入る。

演奏はフットスイッチを指定したタイミングで踏むことによって行う。左のスイッチを踏むと曲が始まる。中央のスイッチは曲中のテンポがない箇所、右のスイッチは曲の終わりで踏む。

実験参加者は演奏を1回行う毎に自分がフットスイッチを踏んだタイミングが指定されたタイミングに合っていたかどうか Table 2 に示す5段階で自己評価する。1つの手法2回ずつ、計10回演奏を行う。演奏する再生方法は全てランダム化した。なお、実験参加者は10人である。

3.2.4. 実験結果(時間)

Max に装備されている timer オブジェクトを用いて左のスイッチを踏んでから中央のスイッチを踏むまで（以下「曲中」）の時間と中央のスイッチを踏んでから右のスイッチを踏むまで（以下「曲の終わり」）の時間を測定し、その時間と提案法のオーディオファイルの長さとの差を取ることで、実験参加者がフットスイッチを踏んだタイミングが本来のタイミングからどれだけずれていたか算出した。

平均値を Fig. 7 に示す。エラーバーは95%信頼区間を表す。曲中の結果に Anova4 on the Web で5水準の一要因分散分析を実施したところ、主効果に5%水準で有意差が見られなかった。曲の終わりの結果に Anova4 on the Web で5水準の一要因分散分析を実施したところ、主効果に5%水準で有意差が見られた。さらに多重比較を実施したところ、従来法と提案法（ループなし）の組み合わせと従来法（短い）と提案法（ループなし）の組み合わせにおいて5%水準で有意差が見られた。

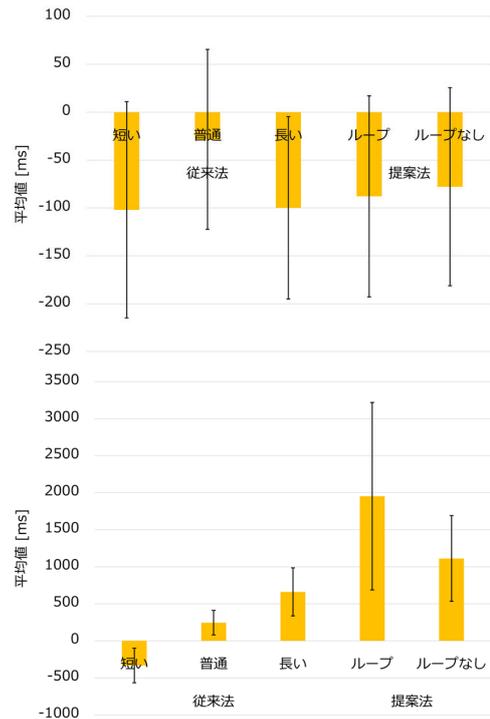


図 7 時間の平均値
(上：曲中，下：曲の終わり)

Fig. 7 Average of time

(Up : Middle of song, Down : End of song)

プ)の組み合わせと従来法（短い）と提案法（ループなし）の組み合わせにおいて5%水準で有意差が見られた。

標準偏差を Fig. 8 に示す。エラーバーは95%信頼区間を表す。曲中の結果に等分散検定を実施したところ、どの組み合わせでも0.5%水準で有意差は見られなかった。曲の終わりの結果に等分散検定を実施したところ、全ての従来法とループ、普通とループなし、短いとループなし、ループとループなしの組み合わせにおいて0.5%水準で有意差が見られた。

曲中において有意差が見られなかった理由は、フットスイッチを踏むタイミングではテンポがないとはいえその直前までテンポがあり、タイミングが取りやすいためであると考えられる。一方、曲の終わりににおいて有意差が見られた理由は、テンポがなく、提案法では曲が終わる手掛かりがないためであると考えられる。

3.2.5. 実験結果(5段階評価)

5段階評価の平均評定値を Fig. 9 に示す。エラーバーは95%信頼区間を表す。曲中の結果に Anova4 on the Web で5水準の一要因分散分析を実施したところ、主効果に有意差が見られた。さらに多重比較を実施したところ、従来法（短い）と提案法の組み合わせにおいて5%水準で有意差が見られた。曲の終わりの結果に Anova4 on the Web で5水準の一要因分散分析を実施し

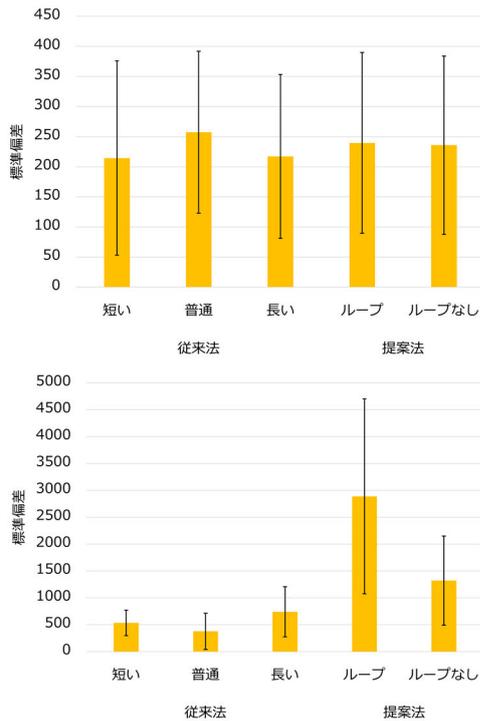


図 8 時間の標準偏差
(上：曲中，下：曲の終わり)

Fig. 8 Standard Deviation of time
(Up : Middle of song, Down : End of song)

たところ、主効果に有意差が見られた。さらに多重比較を実施したところ、従来法（普通）や従来法（長い）は従来法（短い）に比べて 5%水準で有意に評定値が高く、さらに提案法は従来法（普通）や従来法（長い）に比べて 5%水準で有意に評定値が高かった。

曲中において全ての従来法と提案法の組み合わせに有意差が見られなかった理由は、提案法では中央のスイッチを踏んだ瞬間に音が鳴り始めるものの、時間の結果から従来法と提案法でスイッチを踏んだタイミングに差がなく、評価に影響しなかったためと考えられる。一方、時間にばらつきがあったのにも関わらず、曲の終わりにおいて従来法と提案法に有意差が見られた理由は、提案法では音がどんな長さでもフットスイッチを踏んだ瞬間に止まり、実験参加者は自分のフットスイッチを踏んだタイミングが合っていると感じたためであると考えられる。

4. おわりに

本研究では、人間のフットスイッチ操作に合わせて伴奏を行う音響信号を用いた自動伴奏システムの実現を目指した。音響信号をループ再生する際の条件を検討したところ、通常の曲をループ再生する時は groove オブジェクトを用いた方が良い可能性が示された。音響信号を用いた自動伴奏システムを開発し、評価実験を行った結果、曲の終わりでは実験参加者は実際には時

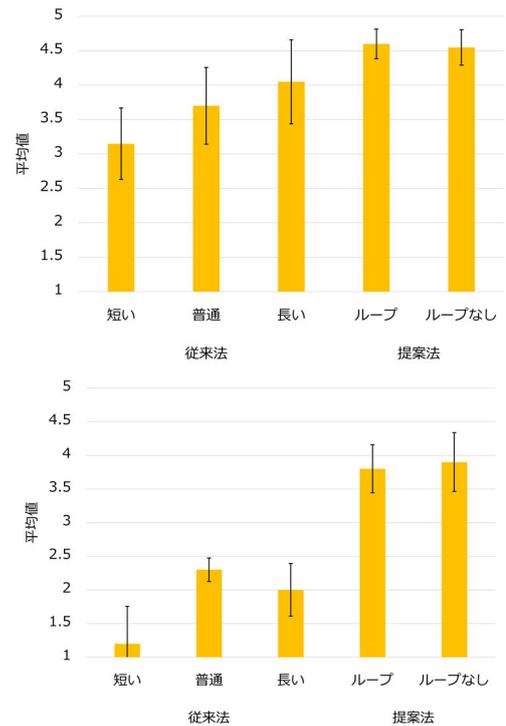


図 9 5 段階評価の平均値
(上：曲中，下：曲の終わり)

Fig. 9 Average of five-scale evaluation
(Up : Middle of song, Down : End of song)

間がずれているのにも関わらず、自分がフットスイッチを踏んだタイミングが合っていると評価した。

以上の結果から、音響信号を用いた自動伴奏システムでは、曲の終わりのテンポがない個所にてクリックノイズが鳴らないようにループ再生すると演奏者は満足いく演奏ができると考えられる。

本研究で用いた Max の groove オブジェクトでループ再生する方法は 1 曲ごとに手動で調整する必要があるため、非常に手間が掛かる。そこで、今後はループ処理を自動化する手法を開発する必要がある。

文 献

- [1] 安部綾太, 木村敏幸, "ペダル操作による自動伴奏システムの評価," 平成 31 年東北地区若手研究者研究発表会講演資料, no. YS-17-P30, pp. 141-142, Mar. 2019.
- [2] 安部綾太, 木村敏幸, "フットスイッチを用いた自動伴奏システムの検討," 第 2 回東北地区音響学研究会発表資料, no. 2-7, pp. 1-4, Nov. 2019.
- [3] 安部綾太, 木村敏幸, "音響信号に対応したフットスイッチ式自動伴奏システムの評価," 令和 2 年東北地区若手研究者研究発表会講演資料, no. R2-A-12, pp. 23-24, Feb. 2020.
- [4] Max, <https://www.mi7.co.jp/products/cycling74/>.
- [5] Octave, <https://www.gnu.org/software/octave/>.
- [6] 青木紀, バンドスコア「アイデア／星野源」, 株式会社フェアリー, 2018.
- [7] Anova4 on the Web, <https://www.hju.ac.jp/~kiriki/anova4/>