

ブレスの合図を認識する伴奏システムの実装と評価

堀内 靖雄 増田 香織 西田 昌史 市川 薫

千葉大学

hory@faculty.chiba-u.jp

あらまし 従来の伴奏システムでは人間の独奏者の音響情報のみを用いて伴奏制御を行っていたが、実際の人間の合奏ではブレスが非常に重要な役割を演じている。本研究では伴奏システムが人間の演奏者のブレスによる合図を認識し、同時に演奏を開始できる手法を提案、実装し、評価実験を行なった。ブレス検出手法は先行研究で提案したオフラインの手法をオンラインアルゴリズムに改良することにより実現した。評価実験ではブレスによる合図に対するシステムの反応と演奏者の演奏のずれに関して、人間の演奏者が許容できる時間範囲を調べた結果、-60ミリ秒～97ミリ秒程度が許容範囲であることがわかった。この結果に基づいて、本提案システムの性能を評価したところ、94.6%が人間の演奏者の許容限界値内に含まれ、十分な精度でシステムがブレスの合図により演奏を開始できることがわかった。

Accompaniment System Using Cue by Breath: Implementation and Evaluation

Yasuo Horiuchi Kaori Masuda Masafumi Nishida Akira Ichikawa

Chiba University

hory@faculty.chiba-u.jp

Abstract Although accompaniment systems use only acoustic information for musical synchronization, human accompanists use breath of the soloist as musical cues in the actual performance. In this study, we will implement the accompaniment system which can use breath as musical cues and evaluate the system by human performers. In a previous study, the off-line method of detecting breath was introduced. In order to implement the breath detection algorithm into the accompaniment system, we introduce the on-line method. The system was implemented and evaluated by human performers. As a result, the performance of the implemented system for using breath as musical cues is quite well and synchronization was achieved in the tolerated range in musical situation.

1. はじめに

伴奏システム[1,2]とは人間の独奏者の伴奏をコンピュータが行なうもので、人間とコンピュータとのリアルタイム・インタラクティブ・システムの一つとして、音楽情報処理研究のみならず、人間とのリアルタイム・インタラクションを考える上で重要な研究テーマとなっている。伴奏システムでは楽譜情報はあらかじめ独奏パートと伴奏パートの両方がシステムに与えられており、独奏者の演奏と協調するように伴奏システムの演奏タイミングをリアルタイムで制御する。

従来の伴奏システムでは独奏者の音響情報（電子楽器の場合はMIDI情報）を用いて伴奏制御を行っていたが、曲の冒頭で独奏者と伴奏システムが同時に音を演奏するような楽曲に対しては対応が困難であり、独奏者がペダルスイッチを踏んで伴奏の演奏をスタートさせたり、独奏の音が演奏されてから即座に伴奏システムの演奏が開始される、などの方法がとられてきた。一方、人間同士の合奏の場合には、独奏者は体の動きやブレス（息継ぎ）を用いて伴奏者に合図を送り、両者の演奏を合わせている。

そこで本研究では伴奏システムが独奏者のブレスの合図を検出し、そのブレスから合奏開始のタイミングを推定し、実際に伴奏を開始するシステムの実現を目的とする。

2. 先行研究

著者らは先行研究[3]において、ブレスから楽音開始までの時間長が一拍の長さとはほぼ等しいことを示し、ブレスを人間の演奏者から伴奏システムへの演奏開始の合図として用いることが有効であることを示した。

- また、ブレスの音響的な特徴として、
- ・パワーがある程度の時間長持続する
 - ・調波構造をほとんど持たない

という特徴を利用して、ブレスの自動検出手法を提案し、実験の結果、100%の精度でブレスと楽音の分離に成功した（クローズド実験）。本研究ではこの手法に基づき、演奏開始の合図としてのブレスを利用可能な伴奏システムの実装と評価を行なう。

3. オンラインブレス検出手法の提案

先行研究[3]ではブレス検出・利用の可能性を調査することが目的であったため、その検出アルゴリズムはオフラインの手法であった。しかし、実際に伴奏システムに実装するためには、上記手法をオンラインのアルゴリズムに修正し、かつ、リアルタイムで動作可能であることが要求される。

先行研究[3]の提案手法はパワーの変動とケプストラムによる調波構造を分析することにより、ブレスを検出していた。しかしながら、ケプストラムの計算にはそれなりに計算コストがかかるため、伴奏システムのリアルタイム性を考慮すると、伴奏システムが独奏者の音高推定に用いている手法を利用できることが望ましい。そこで、ブレス検出における調波構造の推定には、伴奏システムが独奏者の音高推定に利用している自己相関[4]の値を用いることとした。

また、先行研究の手法では、ある一定時間以上、パワーが持続しているかどうかを判定することにより、有音区間（楽音またはブレス）と無音区間（瞬間的なノイズを含む）を分離していた。しかし、この手法ではパワーが減衰した時点でブレスか楽音かを判定していたため、そのまま実装すると判定結果を確定できるまでに時間がかかってしまい、適切な伴奏を行なえなくなってしまう。そこで本研究ではパワーが持続している場合でも、パワーが減衰する前に楽音／ブレス／無音の判別が可能となる手法を提案する。具体的には、有音区間とみなして良いかどうかを判定する時間長閾値 $Th_{duration}$ （この時間長以上にパワーが連続的に持続しているなら有音区間（楽音／ブレス）と判断する）に対し、新たな時間長閾値 $Th_{decision}$ ($> Th_{duration}$) を設け、この時間を経過した段階で自己相関による楽音／ブレスの分離判定を行なうこととした¹。具体的なアルゴリズムは以下の通りである。まず閾値 Th_{power} を越えるパワーが持続している状態で $Th_{decision}$ の時間が経過した段階で、その有音区間の自己相関を算出し、パワーで正規化を行なう。得られた自己相関において、独奏パートの対象としている楽器（本システムではフルート）の最低音と最高音の間の区間（音域）における最大値と同音域内の自己相関の平均値との差が閾値以下である場合には、調波構造をほとんど持たないといえるので、ブレスと判定し、それ以外の場合を楽音と判定する。なお、このとき、ブレス音、楽音それぞれの立ち上がり時刻は、パワーがパワー閾値 Th_{power} を越えた時刻とする。楽音の冒頭では調波構造は不安定であり、分離閾値 $Th_{decision}$ を小さくしすぎると楽音の区間をブレス音と誤検出してしまう可能性があるため、若干の遅延が生じてしまうのは否めないが、この手法によって、オフラインの手法に比べて遅

¹ 現在のシステムでは予備検討の結果、 $Th_{duration}$ を 80 ミリ秒、 $Th_{decision}$ を 120 ミリ秒としている。

延時間を大幅に短縮することができる。

また、パワーが存在するかどうかを判定する閾値 Th_{power} については、環境雑音を考慮し、システム起動時の周辺のノイズを計測することにより、自動的に決定される。

4. 提案システムの評価実験

提案手法を伴奏システムに実装し、提案システムの評価実験を行なった。

4.1. 実験の目的

人間の演奏者がプレスによる合図に関して、どの程度までのずれなら許容できるのか、という時間範囲を調べ、本提案システムの性能が人間の演奏者にどの程度、許容されるのかについて、その有効性を評価することを目的とする。

4.2. 被験者

本実験に参加した演奏者は音楽大学フルート専攻の卒業生二名（以下、演奏者A、演奏者Bとする）とアマチュアのフルート演奏家（フルート経験27年。以下、演奏者Cとする）の三名である。

4.3. 実験に用いた曲

データ収録には実際の曲だけでなく、単調なフレーズからなる練習曲も用いた。収録に用いた楽譜を譜例1～譜例6に示す。

譜例 1 M. Ravel
「亡き王女のためのパヴァーヌ」より

譜例 2 G. Fauré
「小舟にて」より

譜例 3 A. Dvořák
「ユモレスク」より

譜例 4 F. J. Gossec
「ガボット」より

譜例 5 J. S. Bach
「管弦楽組曲第2番よりパディネリ」より

譜例 6 P. Taffanel & Ph. Gaubert
「17のメカニズム日課大練習」より

各楽曲の設定テンポは譜例1から順に52, 63, 92, 116, 132（それぞれ1分あたりの拍数。単位はbeat/minute）とした。また、譜例6に関しては80, 100, 120の三種類のテンポを設定した。

これらの楽曲は演奏者に馴染みがあり、演奏が容易にでき、また、幅広いテンポを網羅するように選曲された。

4.4. 実験条件

本実験では人間の演奏者がプレスによる合図に関して、どの程度までのずれなら許容できるのか、という時間範囲を調べるため、上記の各楽曲に対し、プレスから推定されるタイミング通りに演奏を開始するものに加え、プレスから推測されるタイミングから故意にずらして演奏を開始するシステムを用意した。

具体的には、各楽曲に対し、プレスから伴奏の演奏開始までの時間として、下記の5種類を設定した。

- ・1拍より1/4拍短い
- ・1拍より1/8拍短い
- ・1拍と等しい
- ・1拍より1/8拍長い
- ・1拍より1/4拍長い

例外として、譜例2については楽曲が3連符系であるため、1/4のかわりに1/6、1/8のかわりに1/12とした。なお、これらの設定は、あらかじめ

め予備検討を行ない、伴奏の演奏開始が早すぎると感じられるところから遅すぎると感じられるところを含むように設定した。

4.5. 実験手続き

上述のように8種類の楽曲（譜例6は三種類のテンポがある）に対し、プレスに対する反応が異なる5種類のパターンを用意し、各パターンをそれぞれ二回実験に用いた。結果として、演奏者一人に対し、一曲あたり10回、合計80回の演奏を行なってもらった。各楽譜の順番は、直前の曲のテンポの影響を受けにくいように考慮して決定した。また曲中の10種類のパターンはランダムに提示された。

各演奏において、まず演奏直前に演奏者にメトロノームで指定のテンポを呈示する。メトロノームを停止してから、伴奏者に演奏開始の合図を示すように明確にプレスをした上で演奏を開始してもらった。演奏直後に、伴奏が独奏者に対して適切なタイミングで演奏を開始したと感じられたかどうかについて、5段階で主観的に評価してもらった。各評価は以下の5段階から選択してもらった。

- ・早すぎる(-2)
- ・早いが許容できる(-1)
- ・適切である(0)
- ・遅いが許容できる(1)
- ・遅すぎる(2)

これらは人間の演奏者がプレスによる合図に関して、どの程度までのずれなら許容できるのか、という時間範囲を調べることを目的に設定された。また、伴奏システムとしては、上記の評価で、-1~1の範囲内であればシステムの性能として十分であると考えられる。

4.6. 実験結果 1（人間の許容限界）

各演奏者が伴奏システムの演奏開始タイミングについて、どのように評価したのかについて分析するため、人間の演奏者が演奏を開始した時刻（冒頭の楽音の立ち上がり時刻）と実際に伴奏システムが演奏を開始した時刻（冒頭の楽音の立ち上がり時刻）の時間的なずれと人間の演奏者の評価結果との関係を図4に示す。横軸が時間的なずれで、演奏者の演奏よりも伴奏のタイミングが遅い場合を正としてある。グラフは上から下へ向かってテンポの遅い順に並べてある。

被験者ごと、楽曲のテンポごとに傾向が若干異なるが、おおむね右肩あがりの分布となっており、演奏のずれが評定値と高い相関関係にあることがわかる。また、テンポが遅い曲の方がその傾きがゆるやかであり、逆にテンポが早い

曲ではずれに対して評価が厳しくなっている（時間的なずれに対する許容度が低い）ことを示唆している。

また、ずれが負の部分、伴奏が人間の演奏者よりも先に演奏を開始している状況であるが、全体的なグラフの形状から、伴奏が先に出たことに演奏者が気付いて、予定していたよりも先に演奏を開始したため、ずれが短くなっていると考えられる。すなわち、人間の演奏者が予定していた発音開始時刻はもっと後であったにも関わらず、伴奏が先に演奏を開始してしまったため、人間の演奏者は反射的に演奏開始を早め、結果として、短いずれであるにも関わらず、低い評定値となっていると考えられる。

システム構築の観点からは-1~1の評定は許容可能であり、-2と2の評定値を避ける必要があると考えられる。そこで、今回の実験結果について、-2および2と評定されたケースに対し、各曲中でもっとも厳しく評定された時間的なずれを許容限界と考え、その許容限界とテンポとの関係を図1の実線に示す。また、参考として、同じく-1と1と評定されたケースに対して、各曲中でもっとも甘く評定された時間的なずれを図1の点線に示す。

点線がおおむね実線の外側に出ていることから、実線で示される許容範囲は演奏者によっては-1または1の評定を受けることもあるが、それ以上に厳しい評価基準での許容範囲を示していると考えられる。

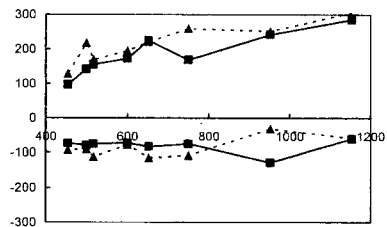


図1 縦軸は許容範囲と考えられる時間的なずれ（単位はミリ秒）を表わす。評価値2および-2に関して、各ケースの中で最も厳しく評価したものを実線、評価値1および-1に関して、各ケースの中で最も甘く評価したものを点線で示す。横軸はテンポから逆算した1拍の時間長（単位はミリ秒）。

図1から、テンポが早いほど許容限界が短くなることや伴奏が早く演奏を開始してしまった場合には人間の演奏者が予定よりも演奏開始を早めているが、そのタイミングがおおむね70~80ミリ秒で一定であることがわかる。逆に言えば、人間の演奏家は伴奏の演奏開始に非常に敏感に反応して、自分の演奏を制御している様子がうかがえる。

上記の理由から、ずれが負の場合、この許容限界は実際の限界値よりも小さく（厳しく）見積られていると考えられるが、人間の演奏家が実際に予定していたタイミングとのずれは不明であるため、以後の考察では、上記の限界値を用いて、システムの性能評価を行なう。

4.7. 実験結果 2 (システムの性能評価)

今回の実験では、システムは意図的にずれたタイミングで演奏を開始したケースがほとんどであるが、もし、システムが推定通りのタイミングで伴奏を開始した場合、人間の演奏家がどのような判断を下したと想定されるかについて検討を行なう。

そこでまず、システムがプレスを検出してから、計画した時刻に正しく演奏を開始できるかどうかについて評価を行なう。図2に今回の実験でシステムが実際に演奏を開始した時刻とプレスからシステムが計画した時刻（意図的に一拍後からずらした場合にはずらした時刻）とのずれの頻度を示す。

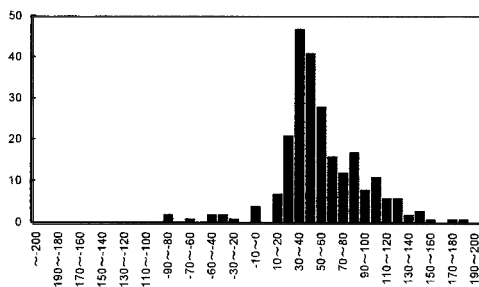


図2 システムの演奏のずれの頻度。横軸は予定していた時刻よりも遅い方が正（単位はミリ秒）。

図2からシステムの場合開始のタイミングが全体的に予定時刻よりも遅れていることがわかる。また、そのずれについてもばらついているのがわかる。これはOSの負荷や音源の発音遅延等の理由でシステムが予定した通りに発音がなされなかった結果であると考えられる。

この図2のデータと上述の人間の演奏家の許容限界値（図1の最も小さい値として、-60ミリ秒から97ミリ秒の間とする）とを比較してみると、全体の85.0%が許容限界値内に含まれており、比較的高い精度でシステムが演奏を開始できることがわかる。

さらに、図2の結果から、全体的に伴奏の演奏開始が遅延していることがわかるため、これらのデータの中央値（47.9ミリ秒）の分だけ早く出力するようにシステムを変更することにより、精度の向上が期待できる。図3に上記補正を行なった結果の頻度分布を示す。

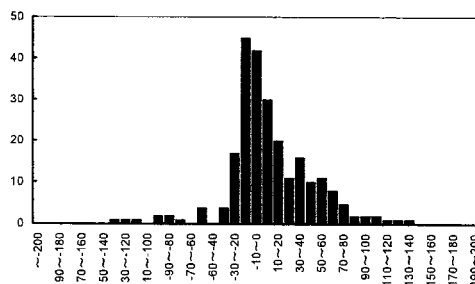


図3 中央値を用いて補正したシステムの演奏のずれの頻度。横軸は予定していた時刻よりも遅い方が正（単位はミリ秒）。

このようにシステムを修正したと仮定した場合、全体の94.6%が許容限界値内に含まれ、推定ではあるが、非常に高い精度でシステムが演奏を開始できることがわかる。

図1から人間の演奏者はテンポが早い曲ほど許容限界が短いことがわかるが、今回実験した範囲内での最も早い曲であっても高い精度で伴奏開始が可能であり、逆にテンポの遅い曲であれば、ほとんど問題なく伴奏を開始できると考えられる。

5. おわりに

本研究では伴奏システムが人間の演奏者のプレスによる合図を認識し、同時に演奏を開始できる手法を提案・実装した。評価実験の結果、本システムのプレスにタイミングを合わせる性能は人間の演奏者のずれの許容限界に対し、十分な精度を有していることがわかった。

今後の検討課題として、上述の補正をシステムに実装すること、また、演奏のずれの原因を追求し、可能であれば、より正確に演奏を出力できるようにシステムを改良したいと考えている。また、冒頭以外の箇所においてもプレスによる合図は重要であるため、そのような利用方法も検討していきたい。

参考文献

- [1] 堀内靖雄, 橋本周司: 「伴奏システム」情報処理 Vol.35, No.9, pp.815~821 (1994)
- [2] 堀内靖雄: 「自動伴奏」, 共立出版「コンピュータと音楽の世界」長嶋洋一, 橋本周司, 平賀譲, 平田圭二 (編), pp.252~269 (1998)
- [3] 堀内靖雄, 飯塚涼子, 西田昌史, 市川薫: 「伴奏システムでのプレス情報利用に関する検討」情報処理学会研究報告 Vol.2005, No.45, pp.13~18 (2005)
- [4] Curtis Roads: 「コンピュータ音楽—歴史・テクノロジー・アート—」青柳龍也・小坂直敏・平田圭二・堀内靖雄 (訳・監修), 東京電機大学出版 (2001)

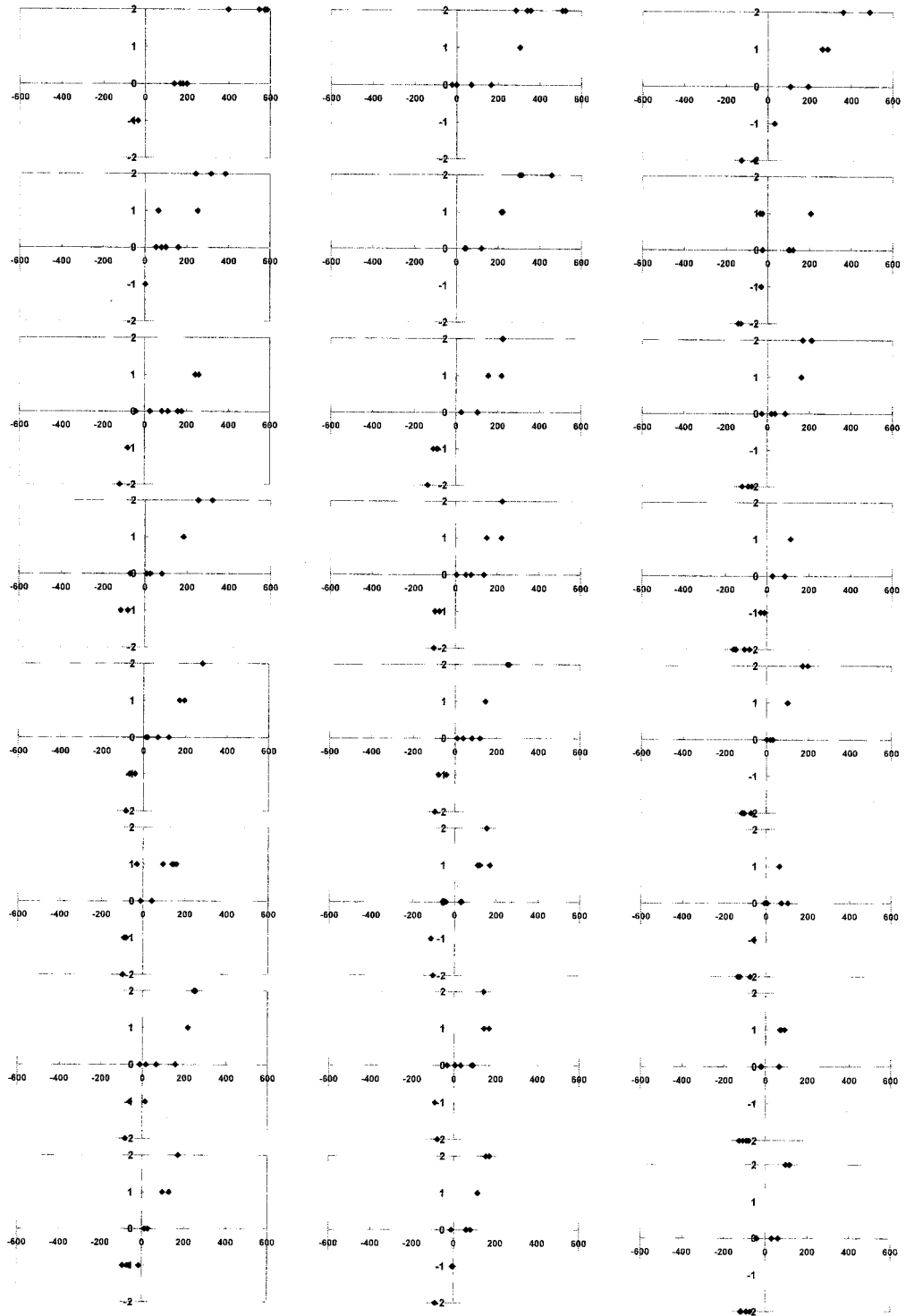


図4 ずれと評定値の関係。横軸がずれ（単位はミリ秒で伴奏が独奏よりも遅い場合を正とする）、縦軸が評定値。各行は上から譜例1(52)、譜例2(63)、譜例6(80)、譜例3(92)、譜例6(100)、譜例4(116)、譜例6(120)、譜例5(132)（括弧内はテンポ[beat/minute]）の順に並べてある。各列は左から演奏者A、演奏者B、演奏者Cである。