

卒業論文 20xx 年度 (令和 xx 年)

リズム予測を用いたネットワーク音楽演奏における遅延の影響の軽減手法  
の提案

慶應義塾大学 環境情報学部  
松藤舜

卒業論文要旨 - 2023 年度 (令和 05 年度)

リズム予測を用いたネットワーク音楽演奏における遅延の影響の軽減手法の提案
--------------------------------------

キーワード:

1. 卒業論文, 2. 村井純研究室, 3. RG, 4. LaTeX

慶應義塾大学 環境情報学部  
松藤舜

Abstract of Bachelor's Thesis - Academic Year 2023

The Use of Rhythm Prediction to Reduce the Effects of Latency in Network Music Performance
--

I can't write English.

Keywords :

1. Thesis, 2. RG, 3. Jun Murai Lab., 4. L<sup>A</sup>T<sub>E</sub>X

Faculty of Environment and Information Studies, Keio University  
Shun Matsufuji

# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	はじめに . . . . .	1
1.2	本論文の構成 . . . . .	1
<b>第2章</b>	<b>背景</b>	<b>3</b>
2.1	ネットワーク音楽演奏 . . . . .	3
2.2	ネットワーク音楽演奏システムの構造 . . . . .	3
2.2.1	OpenSound Control . . . . .	3
2.3	Adaptive Metronome . . . . .	4
2.3.1	相互メトロノームの実験 . . . . .	5
2.3.2	共通テンポの算出 . . . . .	5
2.4	Tablanet . . . . .	5
2.5	Alexandraki . . . . .	5
<b>第3章</b>	<b>問題提起</b>	<b>6</b>
3.1	ネットワーク音楽演奏における遅延 . . . . .	6
3.1.1	遅延のモデル . . . . .	6
3.1.2	遅延の原因 . . . . .	6
3.2	Adaptive Metronome の問題点 . . . . .	7
<b>第4章</b>	<b>提案手法</b>	<b>8</b>
4.1	概要 . . . . .	8
4.2	本システムのアーキテクチャ . . . . .	8
4.3	仮説 . . . . .	8
<b>第5章</b>	<b>実装</b>	<b>10</b>
5.1	拍認識の実装 . . . . .	10
5.2	演奏予測の実装 . . . . .	10
5.3	位相修正 . . . . .	10
<b>第6章</b>	<b>実験</b>	<b>11</b>
6.1	共通テンポ算出の実験 . . . . .	11
6.1.1	目的 . . . . .	11
6.1.2	手法 . . . . .	11

6.1.3	結果	11
6.1.4	考察	11
6.2	予測システムの実験	11
6.2.1	目的	11
6.2.2	手法	11
6.2.3	結果	11
6.2.4	考察	11
<b>第7章</b>	<b>結論</b>	<b>12</b>
7.1	本研究のまとめ	12
7.2	本研究の課題	12
7.3	今後の展望	12
	<b>謝辞</b>	<b>13</b>

# 図 目 次

2.1	PTA フレームワーク. [10] より引用 . . . . .	4
2.2	Adaptive Metronome の実験 [5] . . . . .	4
3.1	連成振動のモデル. [12] より引用 . . . . .	7
4.1	本システムのアーキテクチャ . . . . .	8
5.1	本システムの予測システム . . . . .	10

# 表 目 次

# 第1章 序論

本章では本研究の背景，課題及び手法を提示し，本研究の概要を示す．

## 1.1 はじめに

世界的なインターネット，SNS の普及がもたらしたグローバルなコミュニケーションの発展により，世界中のアーティストをつなぐグローバルな音楽のコミュニティが形成している．それだけでなく，2019 年の新型コロナウイルス感染症の広がりによって多くの人々は自宅から行う遠隔のコミュニケーションを強いられ，Apple，Meta を含む多くの企業によるインターネットで人々をリアルタイムでつなげる仮想世界への関心が上がっている．これらの状況を踏まえて今まで以上にインターネットを介した複数の演奏者同士による遠隔にリアルタイムで行う音楽演奏のシステムへの需要が上がっている．こうした異なる場所にいるミュージシャンをネットワークを介して，同じ部屋にいるのと同じように演奏する音楽演奏を「ネットワーク音楽演奏 (Network Music Performance)[1]」と呼ぶ．

ネットワーク音楽演奏においては，遅延は致命的な問題である．遅延があると演奏者同士のタイミングがずれ，リズムが崩れてしまう．相手が演奏している音をリアルタイムで聴き，それに合わせて演奏することが重要な音楽演奏においてはこのような遅延があると演奏は不可能になる．[2] などの研究によると 100ms 以上になると演奏が困難になる．Yamaha の Syncroom[3] や Zoom の S6 SessionTrak など，このような遅延を減らすシステムは多数開発されてきたが，これらには限界がある．パケットが速く届いてもジッターの関係で安定して情報を処理するまでにバッファは必要であるうえ，人それぞれのネット回線の混雑度，住む場所などで遅延が発生してしまう場合がある．それ以前に光の速さより速く伝達することはできないため，最低限光が地球を回る速度以上の遅延が発生してしまう．

本研究ではネットワーク音楽演奏における遅延を減らすのではなく，遅延を前提として演奏に影響を与えないシステムを提案する．その手法として本研究では演奏相手の演奏を予測し，その予測を再生することで遅延を補償することを提案する．[4] [5] [6] [7] [8] [9] これにより演奏者は相手の演奏の予測を聴き，それに合わせて演奏することで遅延を感じることなく演奏を行うことができる．

## 1.2 本論文の構成

本論文における以降の構成は次の通りである．



2 章では，背景を述べる． 3 章では，本研究における問題提起を述べる． 4 章では，本研究の提案手法を述べる． 5 章では， 4 章で述べたシステムの実装について述べる． 6 章では，本システムの実験評価を行い，考察する． 7 章では，本研究のまとめと今後の課題についてまとめる．

## 第2章 背景

本章では本研究の背景について述べる。

### 2.1 ネットワーク音楽演奏

ネットワークを介して複数の演奏者がリアルタイムで演奏を行うことをネットワーク音楽演奏と呼ぶ。一概にネットワーク音楽演奏と言っても同じローカルネットワークに繋がった小規模なものから、インターネットを介して世界中の演奏者が繋がった大規模なものまで様々なものがある。また演奏者の数も1対1のデュエットから、オーケストラ規模の多くの人を同時に繋げたシステムまで可能である。そのうえ近年では音声だけでなく視覚情報を含めた視聴覚体験を提供するネットワーク音楽演奏のシステムも研究されている。

上記の通りネットワーク音楽演奏には様々なものがあるが、いずれも演奏者の音声をリアルタイムで相手に届けることが重要であり、これには主に二種類の方法がある。音響情報をVoIPアプリケーションなどで送信する方法と、音声情報をMIDIやOSCなどの記号に圧縮し、抽象的なデータとして送信する方法である。

本研究では1対1のインターネットを介したネットワーク音楽演奏を想定し、後者の音声情報をOSCで送信する方法を用いる。

### 2.2 ネットワーク音楽演奏システムの構造

本研究で用いるネットワーク音楽演奏システムの構図はIorwerthらによる”Playing Together Apart Framework (PTA)”を用いる。PTAフレームワークについては図2.1に示す。

TODO: 日本語訳以上の通り片方の演奏者からもう一方の演奏者に届くまで5つの工程がある。

1. 演奏者の演奏
2. 演奏の発信
3. 情報の圧縮, ネットワークを介した送信
4. 情報の受信, 圧縮の解除
5. 演奏の再生

#### 2.2.1 OpenSound Control

OpenSound Control (OSC) は音響情報を記号に圧縮し、抽象的なデータとして送信するためのプロトコルである。本研究ではOSCを用いて音声情報を送信する。——OSCの詳細については後述する。

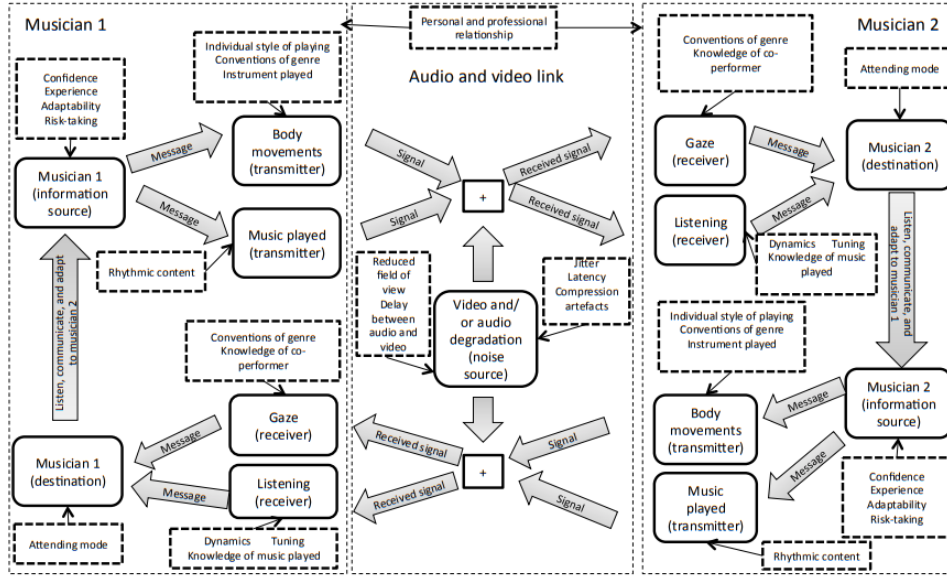


図 2.1: PTA フレームワーク. [10] より引用

## 2.3 Adaptive Metronome

Battello らによる Adaptive Metronome[5][6] は、演奏者の演奏をリアルタイムで分析し、演奏者の演奏に合わせてメトロノームのテンポを変化させるシステムである。

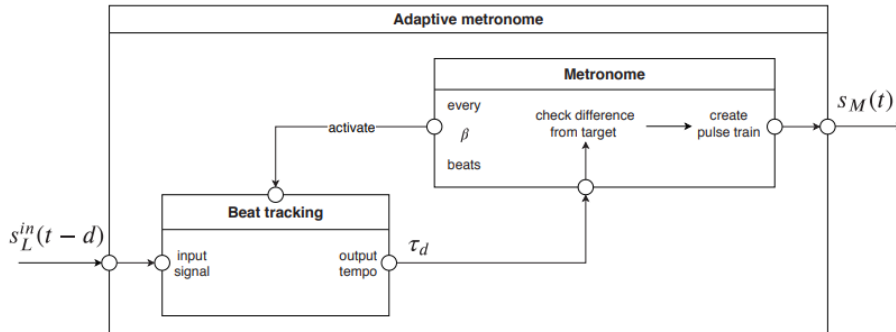


図 2.2: Adaptive Metronome の実験 [5]

Adaptive Metronome は親子構造を持っている。親の演奏者は演奏すると同時にシステムはリアルタイムでその演奏の拍を推定し、子演奏者に音声と拍情報を送信する。子演奏者のシステムはその情報を受取り、親演奏者の演奏に合わせてメトロノームのテンポを変化させる。またこのときメトロノームの音は親から子への遅延を考慮して位相をずれて再生される。

こうすることで小演奏者に聞こえるメトロノームは親演奏者の遅延された演奏に関わらず、常に親演奏者の無遅延の演奏に合わせたテンポで再生される。このシステムを用いた

実験では 120ms の遅延下での演奏を行ったうえでも、被験者は抵抗を感じることなく演奏することができたという結果が得られた。[5]

### 2.3.1 相互メトロノームの実験

当初の Adaptive Metronome の実験では、親演奏者の演奏を子演奏者が聴き、それに合わせて演奏するという形で実験が行われた。後に Battelo らは親子構造を持たず、相互的に Adaptive Metronome を聞きあう実験を行った。実験の結果は現状まだ不十分であるが、このような相互的なシステムでもある程度の効果が得られることが示された。

### 2.3.2 共通テンポの算出

相互的に Adaptive Metronome を聞きあう実験では、全体のテンポを決定させる親演奏者がいないため、演奏者同士のテンポと位相が異なる場合がある。このとき 2 人の演奏者の状況を踏まえたうえで両方の演奏が同期するような共通のテンポを算出する必要がある。

## 2.4 Tablanet

Tablanet[9] は、タブラ奏者の演奏をリアルタイムで分析し、演奏者の演奏に合わせてタブラのテンポを変化させるシステムである。

## 2.5 Alexandraki

Alexandraki らによる研究 [7][8] では、演奏者の演奏をリアルタイムで分析し、事前収録した演奏を実際の演奏に合わせて再生するシステムを提案している。

## 第3章 問題提起

本章では本研究における問題提起について述べる。

### 3.1 ネットワーク音楽演奏における遅延

遅延はネットワーク音楽演奏のしやすさに大きな影響を与える。[2]によると 100ms 以上の遅延があると演奏者は演奏が困難になる。これらの遅延の解決を考察するためには、まずネットワーク音楽演奏における遅延を理解する必要がある。

#### 3.1.1 遅延のモデル

この遅延を音楽演奏の文脈でどう捉えられるかを考えるためには、まずモデルを考える必要がある。[11], [12] では音楽演奏は一定の拍を基準として行われているため、音楽演奏者をオシレーターと捉えている。その場合演奏者による拍を基準とした誤差はオシレーターの周期の位相のずれとして表現できる。これはすなわち、「演奏の崩れ」と考えることができる。それを踏まえると、遅延がある中でのネットワーク音楽演奏においての2人の演奏者の関係は、次のような連成振動のモデルで理解することができる。

ここで PD は Phase Detector, すなわち演奏者の捉えた拍の感覚,  $\tau$  は遅延時間である。このモデルを踏まえて、遅延下の2人の演奏者同士が帰結するテンポは次の式であらわされる。

$$\Omega = \frac{\omega}{1 + K\tau'}$$

$\omega$  は両演奏者の平均テンポ,  $K$  は振動の状態更新を表す低数であり,  $\tau$  は遅延を表す。

#### 3.1.2 遅延の原因

遅延は2.2で述べたすべての工程において発生するが、特に情報の圧縮、ネットワークを介した送信において発生する。

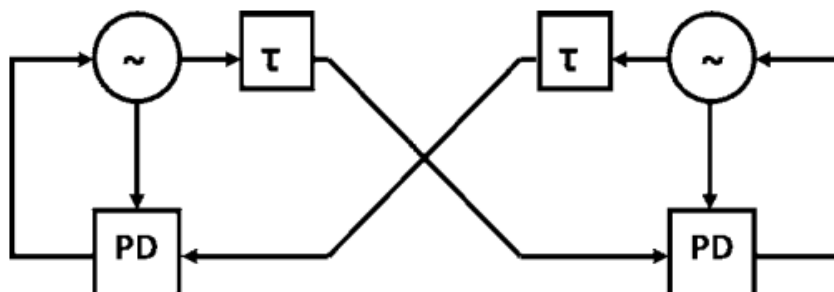


図 3.1: 連成振動のモデル. [12] より引用

## 3.2 Adaptive Metronome の問題点

Adaptive Metronome は演奏者の演奏をリアルタイムで追跡し、演奏者の演奏に合わせてメトロノームのテンポを変化させるシステムを作ることになった。しかし、このシステムではメトロノームに依存した演奏を行うことになり、特に遅延が大きい場合に各演奏者の演奏相手は相手の演奏者ではなく、メトロノームに必然的になってしまうと考えられる。

## 第4章 提案手法

本章では提案手法について述べる。

### 4.1 概要

本研究では従来の手法の「遅延を減らす」との違い、「遅延を補償する」システムを提案する。その手法として演奏者の演奏を予測し、その予測を再生することを提案する。

### 4.2 本システムのアーキテクチャ

本システムは従来のネットワーク音楽演奏のシステムに加え、拍認識部分、演奏予測、位相修正の3つの部分を追加している。

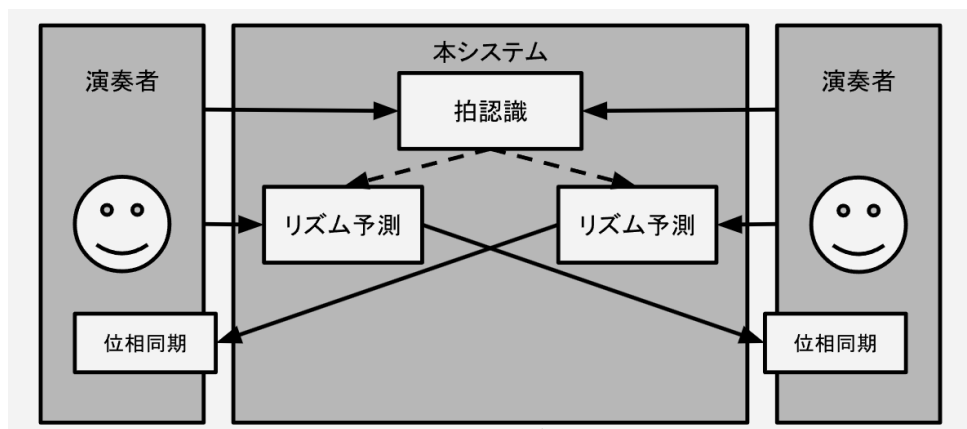


図 4.1: 本システムのアーキテクチャ

### 4.3 仮説

遅延のあるネットワーク下でも、演奏相手の演奏を予測しながら演奏を行うことができれば、遅延の量に関係なくまるで同じ部屋にいるかのように演奏できると仮説を立てる。

従来の Adaptive Metronome と違い、メトロノームの音ではなく予測の音を再生することで演奏者はまるで人間の演奏相手と演奏しているかと同等の体験を得られることがで

きると考える．本システムを用いると音楽における音のニュアンス，表現を保ちつつ，遅延の補償を行うことができる．なお音楽的な表現を保つことが目的であるため，表現を伝達するのに十分な予測精度を得られればよいと考える．



## 第5章 実装

本章では提案手法の実装について述べる.

### 5.1 拍認識の実装

拍認識には様々な手法が存在する.

### 5.2 演奏予測の実装

演奏予測には様々な手法が存在する. — ここで過去の手法について紹介今回は音楽的な表現の伝達に十分な精度しか必要ないため, /citetablenet で用いられた手法を参考にした簡易的な予測手法を用いる.

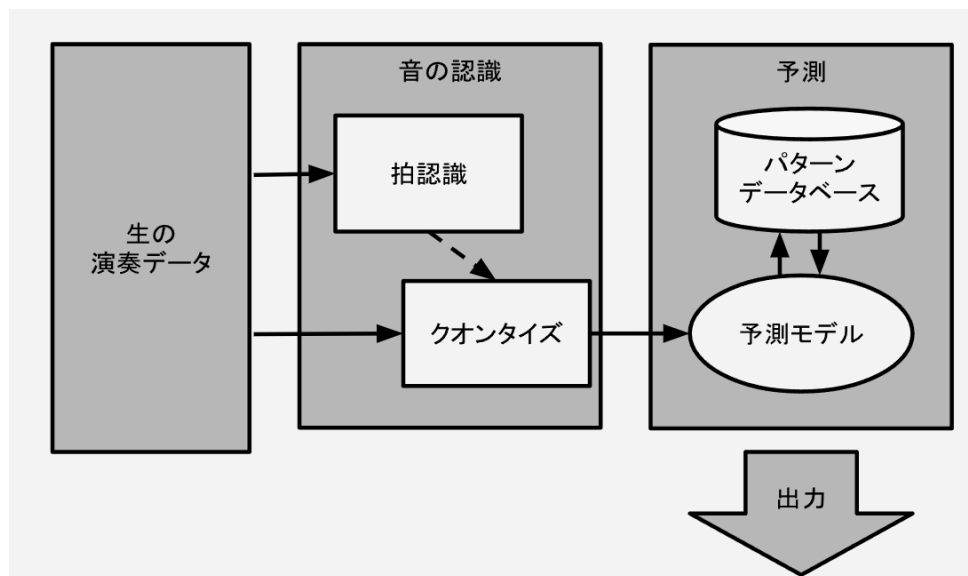


図 5.1: 本システムの予測システム

### 5.3 位相修正

受信部分において演奏相手の演奏予測を受け取った際, 遅延を考慮して位相をずらして再生する必要がある.

## 第6章 実験

本章では，提案システムの実験評価を行い，考察する．

### 6.1 共通テンポ算出の実験

#### 6.1.1 目的

#### 6.1.2 手法

#### 6.1.3 結果

#### 6.1.4 考察

### 6.2 予測システムの実験

#### 6.2.1 目的

#### 6.2.2 手法

#### 6.2.3 結果

#### 6.2.4 考察

## 第7章 結論

本章では，本研究のまとめと今後の課題を示す．

### 7.1 本研究のまとめ

### 7.2 本研究の課題

### 7.3 今後の展望

# 謝辞

俺に関わった全てに感謝

## 参考文献

- [1] John Lazzaro and John Wawrzynek. A case for network musical performance. In *Proceedings of the 11th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, NOSSDAV '01, page 157–166, New York, NY, USA, 2001. Association for Computing Machinery.
- [2] Christopher Bartlette and Mark Bocko. Effect of network latency on interactive musical performance. *Music Perception - MUSIC PERCEPT*, 24:49–62, 09 2006.
- [3] Yamaha. Syncroom について.
- [4] Iorwerth Miriam. *Networked Music Performance: Theory and Applications (English Edition)*. Focal Press, 2023.
- [5] Riccardo Battello, Luca Comanducci, Fabio Antonacci, Augusto Sarti, Stefano Delle Monache, Giovanni Cospito, Enrico Pietrocola, and Filippo Berbenni. An adaptive metronome technique for mitigating the impact of latency in networked music performances. In *2020 27th Conference of Open Innovations Association (FRUCT)*, pages 10–17, 2020.
- [6] Riccardo Battello, Luca Comanducci, Fabio Antonacci, Giovanni Cospito, and Augusto Sarti. Experimenting with adaptive metronomes in networked music performances. 69:737–747, 10 2021.
- [7] Chrisoula Alexandraki and Rolf Bader. Real-time concatenative synthesis for networked musical interactions. *The Journal of the Acoustical Society of America*, 133:3367, 05 2013.
- [8] Chrisoula Alexandraki and Rolf Bader. Using computer accompaniment to assist networked music performance. 01 2014.
- [9] Mihir Sarkar and Barry Vercoe. Recognition and prediction in a network music performance system for indian percussion. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*, pages 317–320. Association for Computing Machinery, 2007.

- [10] Miriam Iorwerth and Don Knox. The playing together, apart framework: A framework for communication in networked music performance. *Journal of Music Technology and Education*, 01 2021.
- [11] Cristina Rottondi, Chris Chafe, Claudio Allocchio, and Augusto Sarti. An overview on networked music performance technologies. *IEEE Access*, 4:8823–8843, 2016.
- [12] Peter Driessen, Thomas Darcie, and Bipin Pillay. The effects of network delay on tempo in musical performance. *Computer Music Journal*, 35:76–89, 03 2011.