

演奏音全体の音響情報を用いた即興演奏システム

菅原芳晴[†] 山口友之[‡] 橋本周司[‡]早稲田大学 先進理工学研究科[†] 早稲田大学 理工学術院[‡]

1.はじめに

従来の即興演奏システムでは計算機が演奏を生成する際、人の演奏や計算機の演奏等を個別に入力データとして用いて、システムの演奏を決定していた[1][2]。そのため、人と計算機によって奏でられた全体の演奏は、明示的には考慮されていない。しかし、実際の即興演奏では各演奏者同士が互いの演奏を聴くだけではなく、演奏音全体を意識して、それぞれの演奏を行っていると考えられる。我々は、人と計算機の演奏を合わせた演奏音全体を計算機への入力として用い、演奏音全体の評価を計算機の演奏にフィードバックする手法を検討している[3]。本稿では全体演奏音の入力音響を用いて音量と音数等を評価することで、即興演奏システムの演奏モードを適宜切り替えて、場の発展に適合するシステムについて報告する。本システムは、全体音響を入力とするため、電子楽器以外や複数人による演奏にも適応可能である。

2.システム概要

提案システムではドラムとその他の楽器でのジャムセッションを想定し、計算機はドラムパートを担当し、人はギターやピアノなどを演奏する。ドラム演奏の生成には MIDI 音源を用い、マイクにより集音した全体演奏音をシステムへの入力とする。また、簡単のために、現在のところ演奏のテンポは一定としている。また、システムはマイクにより集音される計算機の演奏音の音量データを全演奏パターンをあらかじめ保持しておく。

2.1 演奏モード

ジャムセッションには主に二つの演奏形態があると考えられる。第 1 は、各演奏者が主導権を順次交換し、ソロ演奏の掛け合いをする対話型の演奏、第 2 は、全員が同時に参加して盛り上がる協調型の演奏である。ここでソロ演奏とは一人の演奏者が音量を大きくしたり音数を増やすことで演奏の主導権を執っている状態であ

A Session System based on Audio Signal of Whole Musical Performance Sounds.

[†] Yoshiharu Sugawara, Graduate School of Advanced Science and Engineering, Waseda University

[‡] Tomoyuki Yamaguchi, Shuji Hashimoto, Faculty of Science and Engineering, Waseda University

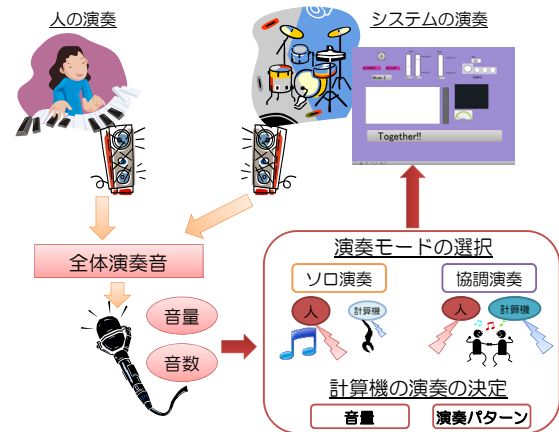


図1 システム概要図

る。この二つの演奏形態を演奏モードとして定義し、前者を **Mode1**、後者を **Mode2** とする。本システムでは、演奏形態を切り替える為のパラメータとして定義した同調率 $M(t)$ の値により演奏モードを切り替える。 $M(t)$ は次節で述べるように、計算機と他の演奏者が同調するように演奏しているかを表すものである。**Mode1** では人がソロ演奏を行うと、計算機は音量を下げ、人がソロ演奏を終えると、計算機は音量を上げてソロ演奏を行う。一方、**Mode2** では人が演奏音量を変化させるとシステムは音量を合わせるように演奏を行い、協調演奏を行う。したがって、二つのモードは全く反対の音量制御となる。

2.2 演奏モードの切り替え

人と人とのセッションでは演奏形態が切り替わるきっかけは、主として以下の場合がある。

- ・全体の演奏が盛り上がる協調演奏の時に急に演奏をやめる場合、又は全体が静かな協調演奏の時に急に音量を上げる場合に、セッションの演奏形態はソロ演奏に変わる。

- ・ソロ演奏の掛け合い中に、主導権を持っているソロ演奏者が音量を徐々に上げる場合に協調演奏に変化する。

本システムでは、全体の演奏の進行状況から、このような切り替えのキーに相当する特徴が認められた時点で演奏モードを切り替えることとした。

全体演奏音の音量は 1 小節毎に算出する。1 小節を 16 拍に分解し、 n 拍目の音響振幅のピーク

値を v_n として、1 小節の音量 $V(t)$ を次のように定義する。ただし、 t は小節数とする。

$$V(t) = \frac{\sum_{n=1}^{16} v_n}{16} \quad (1)$$

さらに、重み係数 α_m を用い音量変化値 $D(t)$ を次のように定義する。

$$D(t) = \sum_{m=1}^4 \alpha_m \{V(t) - V(t-m)\} \quad (2)$$

以上より、1 小節の音量 $V(t)$ と音量の変化値 $D(t)$ を用いて、同調率 $M(t)$ を算出する。ただし、ここでは $M(0)$ を 100 とする。

$$M(t+1) = M(t) + K(V(t), D(t), t) \quad (3)$$

ここで、 $K(V(t), D(t), t)$ は同調率の時間発展を表わす関数で、階層的な条件分岐のあるアルゴリズムで構成的に計算している。要約すれば、 $V(t)$ と $D(t)$ が共に低い値をとる継続小節数に応じて正の値となり、 $D(t)$ の絶対値が大きい場合は負の値となる性質がある。算出された $M(t)$ と閾値 T_p により演奏モードを以下のように切り替える。

$$\begin{cases} M(t) < T_p & \text{Model} \\ M(t) \geq T_p & \text{Mode2} \end{cases} \quad (4)$$

以上より、Model のときソロ演奏後の音量変化が小さい場合、協調演奏へと移行する。一方、Mode2 のときは急激な音量変化が起こった場合ソロ演奏へと移行することになる。

2.3 計算機の演奏音量

全体演奏の音量と計算機の演奏の音量を比較して、Model では全体演奏音の音量が小さい場合、計算機はソロ演奏のために自身の音量を大きくする。Mode2 では計算機の演奏の音量を全体の音量のバランスに揃えるように、つまり共に盛り上がるように制御する。

一対一のセッションでは問題ないが、全体演奏音から計算機の音量と複数人による演奏の音量を区別する際に、計算機の音量が既知だとしても、単に時間領域で全体の音量から計算機の音量を減算して、複数人による演奏の各演奏者の音量を正確に計測することはできない。そこで、まず全体演奏音を 7 つの周波数帯に分離する。分離した各周波数帯をそれぞれ逆フーリエ変換し、時間領域に戻す。同様の処理を、計算機の演奏音に適用する。各時間領域で、全体演奏音と計算機の演奏音の差分をとり、各時間領域でピークを求める。7 つのピークの中から最大のものを、複数人による演奏の中の代表的な音量 $V_h(t)$ とする。また、ソロ演奏の掛け合いを行う Model では人のソロ演奏時は計算機の音量は一定に、ドラムのソロ演奏時は $V_h(t)$ の平均値を

計算機の音量としている。協調型の演奏を行う Mode2 では、 $V_h(t)$ に合わせるように常に計算機の音量を調整する。

2.4 ドラムの演奏パターンの選択

ドラムの演奏パターンは $V(t)$ と人の演奏の 1 小節の音数 N_i に依存し、ドラムパターンデータベースより選択される。データベースは階層構造となっており、Model はまず $V(t)$ により、バックミュージックとソロ演奏に分けられ、さらに N_i の値によりパターンが選択される。Mode2 では $V(t)$ により三段階に分けられ、 N_i に合わせたパターンが選択される。

3. システム確認

本システムを用い、複数人による即興演奏を行った。人は電子ピアノ奏者 1 名とアコースティックギター奏者 1 名の計 2 名であり、システムはドラムを担当した。各楽器の中間地点にマイクを置き、演奏を行ったところ、協調演奏ではバランス良く音量を合わせ、音数にも対応できることが確認できた。また、演奏モードの切り替えも適切であることが確認でき、ソロ演奏の掛け合いも行えた。

4. まとめと今後の展望

入力信号として用いる全体演奏の音量と音数等を評価することで、演奏の場を意識した柔軟なシステムが構築できた。全体演奏音を用いるため、電子楽器以外の入力に対しても適応可能であり、3 つ以上の楽器によるジャムセッションにも適用できる。現在、音量変化だけでなくリズムの認識やテンポのずれを修正できるよう拡張を検討している。

謝辞

本研究の一部は、早稲田大学ヒューマノイド研究所、科学技術振興機構 CREST「人を引き込む身体的メディア場の生成・制御技術」、日本学術振興会グローバル COE プログラム「グローバルロボットアカデミア」の支援を受けて行われた。

参考文献

- [1] 日高, 後藤, 村岡, “すべてのプレーヤーが対等なジャズセッションシステム”, 情報処理学会研究報告, 1996
- [2] 後藤, 日高, 松本, 黒田, 村岡, “仮想ジャズセッションシステム: VirJa Session”, 情報処理学会論文誌, 1999
- [3] 菅原, 山口, 橋本, “演奏音全体からのフィードバックを有する即興演奏システム”, 情報処理学会全国大会論文集, 2010