

連想型自己組織化マップを用いたリズム演奏支援システム

笠原 俊一[†] 三枝 亮^{††} 橋本周司[†]

本稿では、演奏者のリズム演奏に応じて共創的にリズム演奏を生成するシステムを提案する。一般の音楽演奏では、ある演奏者の演奏が他の演奏者の演奏を誘起し、これらが繰り返されることで、共創的に演奏が行われることが多い。本システムは、このような共創的な音楽演奏を機械システムとして実現する。これまでに自動伴奏システムなど、人間の演奏に対して自動的に協調して演奏を生成するシステムが提案されているが、システムの演奏想起にユーザの嗜好や感性が十分に考慮されているとはいえない。本研究では、人間のリズム知覚に基づいたリズムの類似尺度を導入し、複数のリズムマップとその連結機構を実際の演奏データを用いて構築することで、ユーザと動的な協調演奏が可能なリズム演奏支援システムを開発した。感性評価実験により、本稿で提案するリズム類似尺度の妥当性を統計的に検証し、この類似尺度に基づいて構築されたリズムマップを解析した。さらに、複数人の被験者によるシステム評価実験により、本システムが共創的なリズム演奏に有効であることを確認した。

Rhythm Generation with Associative Self Organizing Maps

SHUNICHI KASAHARA,[†] RYO SAEGUSA^{††} and SHUJI HASHIMOTO[†]

This paper describes an interactive rhythm generation system. Generally, musical sessions are often co-created through the interaction among the several players. The proposed system allows us to realize the co-creative musical performance. Even though many autonomous systems for musical accompaniment and musical session have been developed so far, musical preference and ideas of the system user are not enough considered in the performance association. We developed an interactive rhythm generation system with the novel rhythm similarity based on the human's perception, and the rhythm maps constructed by the learning scheme with real human's performance. In perceptual experiments, we statistically examined the validity of the introduced rhythm suitability, and analyzed the constructed rhythm maps by the similarity. Finally, we concluded that the proposed system was effective for the co-creative musical performance through many trials of performances with human players.

1. はじめに

近年、人間の音楽演奏に協調して演奏を行うインタラクティブな音楽演奏支援システムが多く提案されている^{1)~4)}。しかしながら、共創的な演奏生成を想定したシステムは少なく、ユーザがシステムと共演するためには音楽演奏の技術に加えてシステムに関する知識が必要とされることも多い。本研究では、人間のリズム知覚に基づいたリズムの類似尺度を導入し、リズムの連想機構を実際の演奏データを用いて構築することで、ユーザが直感的に音楽セッションを実現できるインタラクティブなリズム演奏支援システムの構築を目指す。

指す。

Raphaelの自動伴奏システム^{1),2)}は、実時間でテンポやリズムを推定し、伴奏を生成することができるが、システムが生成する伴奏演奏には、ユーザの嗜好や感性が十分に考慮されていない。ジェスチャや演奏を通して人間と機械が共演できる後藤らの仮想ジャズセッションシステム³⁾や、人間と機械が音楽演奏の主導権をインタラクティブに交換することができるTakiらの自動伴奏システム⁴⁾などでは、ユーザの感性を反映したインタラクションが実現されている。しかしながら、これらの手法では楽器を演奏する技術が必要であり、音楽演奏の経験のないユーザがシステムを使用することは困難である。

我々は、音楽演奏として最も基本的な演奏形態であるリズム演奏に着目し、操作が容易でインタラクティブな演奏支援システムを構築した。リズム演奏は音高の変化がないため、演奏の特徴抽出や再構成が容易で

[†] 早稲田大学大学院理工学研究科物理学及応用物理学専攻
Major in Pure and Applied Physics Graduated School
of Science and Engineering, Waseda University

^{††} Robotics, Brain and Cognitive Sciences Department,
The Italian Institute of Technology

あり、リズムの組合せを感性的に評価しやすいなどの特長がある。

リズム演奏システムの開発は、これまでに報告されている。Tokui らのリズム演奏システム⁵⁾では、対話型進化計算によりユーザの感性を反映したリズムを生成できる。しかしながら、人間による主観評価データを進化計算の評価値として用いるため、ユーザの負担が大きいという問題がある。直感的なリズム演奏の生成や再構成を実現するためには、リズムをテンポに依存しない特徴量として表現すると都合がよい。武田ら⁶⁾は、音符のまとまりをリズムベクトルとしてとらえ、テンポに依存しない特徴量を導入している。本研究でも、1小節におけるリズムを1つのベクトル(リズムベクトル)として表現しているため、処理はテンポには依存しない。

人間の感性に適合したリズム演奏を可能とするためには、リズムベクトルが何かしらの感性的な類似尺度で評価される必要がある。リズムのジャンル名やリズム特徴を表現する形容詞などの、インデックス情報を用いた類似度評価では、インデックス作成のための負担が大きく特徴の表現能力にも限界があり、数量的な計算処理も困難である。吉井ら⁷⁾による編集機能付きオーディオプレイヤーでは、リズムベクトルの類似度はベクトル間の距離に基づいて評価されているが、リズムの知覚的な類似性とそれを表現するベクトル間の距離は必ずしも一致しない。

我々はこれまでに、リズム知覚の性質を考慮した階層的なリズム表現法を提案し、リズムの構造推定⁸⁾やリズム配置⁹⁾において、良好な結果を得てきた。本稿では、リズムの感性的な安定度を考慮したリズムの類似尺度を導入し、リズムの低次元配置(リズムマップ)を2つの打楽器ごとにそれぞれ構築する。さらに、これらのリズムマップを連結することで、ユーザの演奏に応じてリズムをインタラクティブに想起するリズムセッションを実現する。システムが人間のリズム演奏に対して適切なリズムを想起するためには、リズム間の知覚上の相性を評価する必要がある。本システムでは、同時に演奏される頻度が高いほどリズム間の相性が良いと仮定し、実際の協調リズム演奏を学習データに用いて、自然なリズム連想を試みた。

以下、本稿の2章では、本システムを構成するリズムマップとその連結、およびリズムの連想機構について述べる。3章では、リズム類似尺度を検証するための感性評価実験や、システムの評価実験などについて述べる。4章では、本稿のまとめと今後の展望について述べる。

2. リズム演奏支援システム

図1に、本システムと人間の演奏者による協調リズム演奏の模式図を示す。本システムは、2つのリズムマップとそれらの連結(マップリンク)により構成される。2つのリズムマップは、それぞれ異なる打楽器に対応する。2つのリズムマップとマップリンクは、2段階の学習により構築される。事前に採取した2名の演奏者による協調リズム演奏データから、まず2つのリズムマップを生成し、次にそれらを連結するマップリンクの結合強度を獲得する。ユーザはリズムマップをインタフェースとしてリズム演奏を生成する。それに対して、システムは得られたリズムマップと結合強度に基づいてリズムを想起し、演奏を行う。それぞれのリズム演奏は音響としてユーザに提示される。ユーザがその音響に触発されてリズムを変えると、システム側のリズムも変化する。このようにしてユーザとシステムによるダイナミックなリズムの共創が実現される。以降では、リズムマップ、マップリンク、リズムの連想機構についてそれぞれ述べる。

2.1 リズムマップ

本稿で定義されるリズムマップは、Kohonenの自己組織化マップ¹⁰⁾をリズム表現に適用したものである。図2にリズムマップを示す。2次元的に配置された各ユニットにリズムベクトルが格納される。近傍のユニットには類似したリズムが格納される。リズムベクトルは、1小節のリズムを16個の打点に標本化し、打点の強さ(Velocity)を128段階に量子化した16次元のベクトル $(v_1, v_2, \dots, v_{16})$ ($0 \leq v_i \leq 127$)とする。

2.1.1 リズムベクトルの類似度

リズム知覚の性質を考慮したリズムベクトルの類似度を導入する。以下では、まず、リズムの階層構造¹¹⁾に基づくリズムの安定性類似度 D_s と、ベクトル距離に基づくリズムの空間的類似度 D_v を定義し、節末にリズムベクトルの類似度を定義する。

本研究で仮定するリズム階層構造を図3に示す。標本数 2^p のリズムベクトルは、1周期の時間を上位階

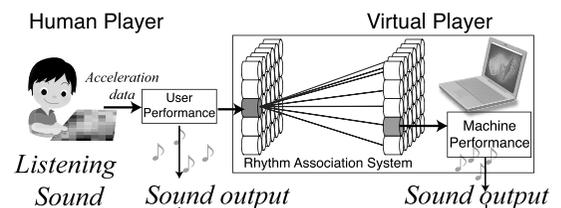


図1 本システムによる協調リズム演奏
Fig.1 Overview of interaction system.

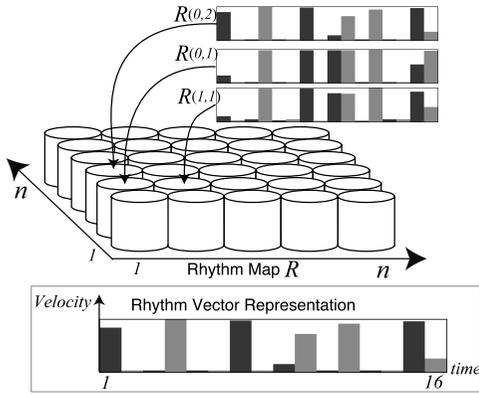


図2 リズムマップ
Fig. 2 Rhythm map.

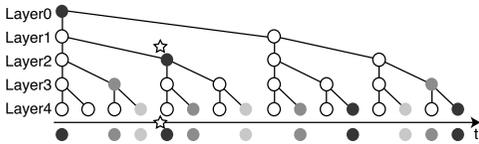


図3 リズムの階層構造．円形と円内の色の濃さは、それぞれ打点と打点の強さを表す
Fig. 3 Rhythm hierarchical structure.

層から p 回 2 分割した木構造として記述される．リズムベクトルの各成分の階層は、その標本点で最も高い階層とする．たとえば、図 3 中で 指示された 4 番目の打点の階層 L は 2 となる．リズムの階層的な知覚に関する研究は多くなされており、リズムの階層構造において上位階層の打点が下位階層の打点より強い場合により安定性が強く感覚されるという報告がある¹²⁾．これらの仮説に基づいて、以下にリズムの安定性類似度 D_s を定義する．

まず、リズムベクトルの各打点の安定度を表現するため、階層 L の各打点 v_i に対し、それよりも上位階層にある打点の中で直前の打点 v_i^b と直後の打点 v_i^f との強さの差 $v_i^b - v_i$ 、 $v_i^f - v_i$ を stb 、 stf と定義する．これらの値が大きいほどリズムに安定感があると仮定する．この仮定は、リズムの局所的な時間変化、特に、強拍と弱拍に関する知覚に対応する．

次に、リズムベクトルの各階層の安定度分布を表現するため、階層 L の stb と stf をそれぞれ降順に並べ替えた値の列を、階層 L の安定度分布 $STb_{(V,L)}$ 、 $STf_{(V,L)}$ と定義する．

$$STb_{(V,L)} = \text{sort}\{stb_1, stb_2, \dots, stb_{(2^{L-1})}\} \quad (1)$$

$$STf_{(V,L)} = \text{sort}\{stf_1, stf_2, \dots, stf_{(2^{L-1})}\} \quad (2)$$

リズムの安定度分布を図 4 に示す．値の並べ替えにより、各打点の安定度の時間的配置に依存せずに、リズムの安定度を評価できる．

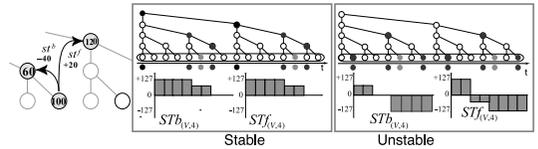


図4 リズムの安定度分布
Fig. 4 Rhythm stability distribution.

以上をふまえて、2 つのリズムベクトル $V^1 = (v_1^1, v_2^1, \dots, v_{16}^1)$ と $V^2 = (v_1^2, v_2^2, \dots, v_{16}^2)$ の安定性類似度 $D_s(V^1, V^2)$ を次式で定義する．

$$D_s(V^1, V^2) = \frac{1}{m(L^*)} \times \sum_i^{m(L^*)} \left\{ \frac{1}{2} \left(e^{-\frac{dSTb^2(i)}{2\alpha^2}} + e^{-\frac{dSTf^2(i)}{2\alpha^2}} \right) \right\} \quad (3)$$

$$dSTb(i) = STb_{(V^1, L^*)}(i) - STb_{(V^2, L^*)}(i) \quad (4)$$

$$dSTf(i) = STf_{(V^1, L^*)}(i) - STf_{(V^2, L^*)}(i) \quad (5)$$

ただし、2 つのリズムを表現するのに必要な最小の階層を L^* とする．なお、階層 L^* の打点数は $m(L^*) = 2^{L^*}$ となる．式 (3) より、 D_s はリズムの安定度分布 $ST(V^1)$ 、 $ST(V^2)$ の形状が類似しているほど大きく 1 に近い値となる．

リズムベクトルの空間的類似度 D_v を式 (6) で定義する．

$$D_v(V^1, V^2) = \frac{1}{16} \sum_{i=1}^{16} \exp \left\{ -\frac{(v_i^1 - v_i^2)^2}{2\beta^2} \right\} \quad (6)$$

空間的類似度は、リズムベクトルの時系列パターンとしての類似性を表す．このように、ここでは類似度の最大が 1 で、類似の度合いが下がるほど、0 に近づくようなガウス関数を用いている．

以上の安定性類似度と空間的類似度の重み付け和により、リズムベクトルの類似度 S_a を式 (7) で定義する．

$$S_a(V^1, V^2) = aD_v(V^1, V^2) + (1 - a)D_s(V^1, V^2) \quad (7)$$

ただし、 $0 \leq a \leq 1$ とする．上式よりリズムベクトルの類似度は $0 \leq S_a(V^1, V^2) \leq 1$ となる．なお、 a 、 α 、 β の値は、3.1 節の感性評価実験で人間のリズム感性尺度と最も近くなるように決定される．

2.1.2 リズムマップの学習

前項のリズムベクトル類似度を用いて、リズムマップの学習を行う．図 5 に学習時のリズムマップのデータフローを示す．リズムマップの座標 $X = (x, y)$ $0 \leq x, y < n$ のユニットに格納されたリズムベクトルを $R(X)$

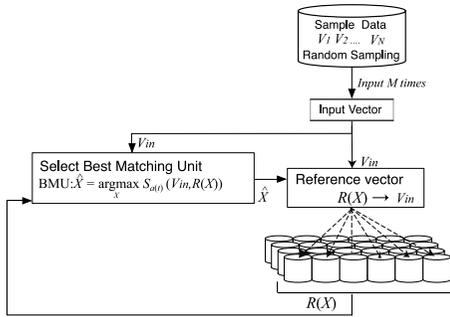


図5 学習時のリズムマップのデータフロー
Fig. 5 Procedures of rhythm map learning.

と表す。ただし、 n はリズムマップの1辺の長さ、 x, y は整数とし、各ユニットのリズムベクトルの初期値はランダムに決定した。

リズムマップの学習は、まず、入力リズムベクトル V_{in} に対する類似度が最大なリズムベクトルを持つユニット BMU (Best Matched Unit) を決定し、次に、すべてのユニットのリズムベクトルを式 (8) に従って更新する。

$$R(X)^{t+1} = R(X)^t + \eta(t)(V_{in} - R(X)^t) \quad (8)$$

ただし、上式は学習回数 t の更新を表す。学習係数 $\eta(t)$ は、式 (9) のように対象ユニットの位置 $X = (x, y)$ と BMU の位置 $\hat{X} = (\hat{x}, \hat{y})$ の距離によってガウス関数から決定され、その広がりを表すパラメータ $\sigma(t)$ は、式 (10) に従って更新される。

$$\eta(t) = \exp \left[-\frac{(\hat{x} - x)^2 + (\hat{y} - y)^2}{2\sigma^2(t)} \right] \quad (9)$$

$$\sigma(t+1) = 0.9999\sigma(t) \quad (10)$$

なお、BMU の決定には、本稿で導入したリズム類似度 S_α を用いているため、リズムベクトルの安定性と空間的距離が両方とも考慮されているが、各ユニットのリズムベクトルは空間的距離のみに基づいて更新されるため、本稿のリズム類似度に適合したリズムマップが得られない可能性がある。そこで学習段階に応じて、類似度定義の重み付け係数 a を式 (11) のように変化させる。

$$a(t) = \frac{1}{1 + A^{M-t}} \quad (11)$$

2.2 リズムマップの連結

複数人の演奏者による音楽演奏では、演奏者同士が互いに演奏を連想し合って共創的に演奏を生成することが多い。このような共創的な演奏を機械システムにより実現するためには、双方向かつ多対多対応な連想の仕組みが必要である。

連想システムは、Hopfield 型ネットワーク¹³⁾ やアソ

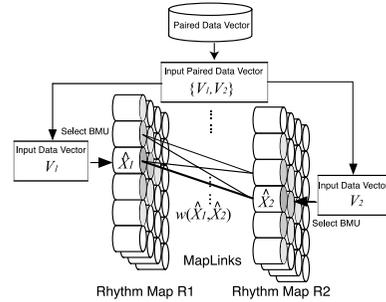


図6 学習時のマップリンクのデータフロー
Fig. 6 Procedures of map link learning.

シアトロン¹⁴⁾ などのニューラルネットワークモデルによって実現されることが多い。また、近年では Kosko による2層からなる双方向連想メモリ BAM^{15),16)} や Hagiwara による多層での多方向連想メモリ MAM¹⁷⁾ などが提案されているが、これらは音楽演奏支援には特化されていない。ここでは、前節で定義した2つのリズムマップを、実際のリズム演奏データに基づいて連結する手順について述べる。

2つのリズムマップの各ユニット間の連結する仕組みを、マップリンクと呼ぶことにする。連結する2つのリズムマップはそれぞれ2次元であるため、マップリンクは4次元のテーブルとなる。リズムマップ R_1 とリズムマップ R_2 のユニットの座標を X_1, X_2 、これらのユニット間の結合強度を $0 \leq w(X_1, X_2) \leq 1$ と表す。

マップリンクの学習は、同時に演奏された2つのリズムに対応するユニット間の結合強度を強化することで行われる。また、それらの近傍にあるユニット間の結合強度も同時に強化される。近傍ユニットの結合強化により2つのリズムマップは滑らかに結合される。図6に学習時のマップリンクのデータフローを示す。

リズムベクトル V_1, V_2 が同時に演奏された場合、2つのリズムマップ R_1, R_2 のマップリンクは次のように更新される。 V_1 の R_1 における BMU の座標を $\hat{X}_1 = (\hat{x}_1, \hat{y}_1)$ 、 V_2 の R_2 における BMU の座標を $\hat{X}_2 = (\hat{x}_2, \hat{y}_2)$ とするとき、 X_1 と X_2 のユニット間の結合強度 $w(X_1, X_2)$ を、式 (12), (13), (14) で更新する。ただし、 $w(X_1, X_2) = w(X_2, X_1)$ とする。

$$w_{t+1}(X_1, X_2) = 1 - (1 - w_t(X_1, X_2))(1 - K(X_1, X_2)) \quad (12)$$

$$K(X_1, X_2) = \exp \left\{ -\frac{|X_1 - \hat{X}_1|^2}{2\gamma_1^2(\hat{X}_1)} \right\} \exp \left\{ -\frac{|X_2 - \hat{X}_2|^2}{2\gamma_2^2(\hat{X}_2)} \right\} \quad (13)$$

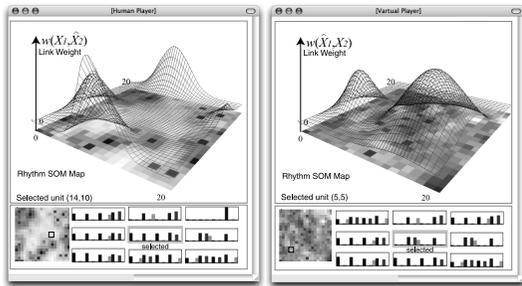


図7 リズム演奏支援システムの操作インタフェース。左ウィンドウはユーザのインタフェース、右ウィンドウはシステムのインタフェースに対応する

Fig. 7 Interaction interface.

$$\begin{aligned} \gamma_i(X_i) &= \frac{m}{8} \sum_{X_j \in D(X_i)} S(R_i(X_i), R_i(X_j)) \quad (14) \end{aligned}$$

なお、更新の影響範囲を表すパラメータ γ_i は各リズムマップ R_i における BMU \hat{X}_i の 8 近傍ユニットのリズムベクトルの類似度平均により決定される。 $D(X_i)$ は X_i の 8 近傍ユニットの集合を表す。したがって、近傍ユニットのリズムベクトルが類似しているほど、更新の影響範囲は大きくなる。

2.3 リズムの連想機構

ユーザのリズム演奏に適したリズムを連想し、インタラクティブに演奏を行う仕組みについて述べる。前節で述べた 2 つのリズムマップは、それぞれユーザとシステムに対応させ、ユーザとシステムが各々のリズムマップ上でユニットを選択することで、それらのユニットに格納されたリズムベクトルが音響として演奏される。システムのユニット選択は、ユーザの選択したユニットと 2 つのリズムマップの結合強度に基づいて決定される。以下に、システムの操作インタフェースとリズムの連想アルゴリズムについて述べる。

2.3.1 リズム連想インタフェース

図 7 にリズム演奏支援システムの操作インタフェースを示す。左右のウィンドウは、それぞれユーザとシステムのリズムマップに対応し、これらはマップリンクにより連結されている。各ウィンドウの下部にはリズムマップが表示され、上部には 2 つのリズムマップ間の結合強度が表示される。

図 7 のリズムマップ部分の拡大図を図 8 に示す。ユーザは、マウスなどのポインティングデバイスを用いて、このインタフェース上でリズムマップのユニットを選択する。リズムマップでは近傍のユニットに似たリズムベクトルが配置されるため、選択するユニットの位置を少し変化させることで、類似した異なるリズムを生成できる。リズムマップの右横には、選択さ

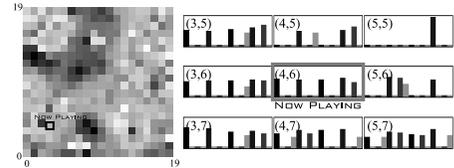


図8 リズムマップの拡大図

Fig. 8 Rhythm map display.

れたユニットとその 8 近傍ユニットのリズムベクトルが表示される。リズムマップの輝度値は、ユニットに格納されているリズムベクトルの成分の平均値を表し、平均値が大きいほど暗く表示される。すなわち、輝度値はリズムの打点強度に相当する。

2 つのリズムマップ間のマップリンクは 4 次元のテーブルであるが、一方のリズムマップの 2 次元座標を指定することで、図 7 のように 3 次元グラフとして可視化できる。左ウィンドウのユーザ側のインタフェースには、システムがリズムマップ R_2 で選択したユニット \hat{X}_2 と、ユーザ側のリズムマップ R_1 の全ユニット X_1 との結合強度 $w(X_1, \hat{X}_2)$ が 3 次元的に表示される。同様に、右ウィンドウでは、ユーザがリズムマップ R_1 で選択したユニット \hat{X}_1 とシステム側のリズムマップ R_2 との結合強度が表示される。なお、結合強度の表示には、Non Uniform Rational B-Spline¹⁸⁾ による曲面補完を用いた。リズムマップのサイズは計算コストを考慮し 20×20 程度としたが、結合強度の曲面は 50×50 程度の高分解能で描画した。

選択されたリズムの音響は、MIDI 音源により生成され、スピーカよりユーザに提示される。本システムでは、リズム音響のテンポと音色を任意に指定できるが、以降の実験では、テンポを BPM = 120 とし、ユーザ側の音色はスネアドラム、システム側の音色はバスドラムとした。これはシステムの学習時に使用した演奏音色と同一である。リズムの音響生成は 1 小節単位で行われ、各単位時間の 50 msec 前に選択されているリズムベクトルが次の単位時間で演奏される。

2.3.2 リズム連想のアルゴリズム

システムのリズム連想の手順について述べる。システムのリズム連想は、リズムマップ上を移動するエージェントの位置で表現される。時刻 t において、ユーザが選択したユニットの位置を $X'_1(t)$ 、エージェントの位置を $X'_2(t) = (x'_2(t), y'_2(t))$ とし、次の時刻 $t+1$ のエージェントの位置 $X'_2(t+1) = (x'_2(t+1), y'_2(t+1))$ を式 (15), (16) で定義する。ただし、 M は移動における慣性質量を表す。

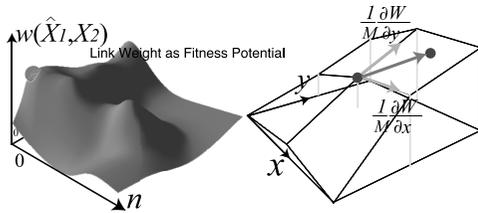


図 9 エージェントの動作
Fig. 9 Behavior of agent.

$$x'_2(t+1) = \frac{1}{M} \frac{\partial W}{\partial x_2}(X'_1(t), X'_2(t)) + x'_2(t) \quad (15)$$

$$y'_2(t+1) = \frac{1}{M} \frac{\partial W}{\partial y_2}(X'_1(t), X'_2(t)) + y'_2(t) \quad (16)$$

図 9 に示されるように、エージェントはユーザの選択したリズムに応じて、システム側のリズムマップ上を結合強度の最急勾配方向に移動する。なお、インタラクティブセッションの開始時において、ユーザは任意の初期地点を選択し、エージェントの初期値点はリズムマップの中心地点とする。エージェントが選択したリズムに対して、ユーザ側のリズムマップにも結合強度が表示されるので、ユーザはその時点でシステムが選択したリズムに対する相性を直感的に把握しながら、次のリズムを選択することができる。このようにユーザとシステムのリズム連想が繰り返されることで、インタラクティブなリズム演奏が実現される。

3. 実験

3.1 リズム類似度の感性評価実験

本稿で提案するリズム類似度が感性的尺度として妥当であることを統計的に検証するため、リズムの類似感覚に関する感性評価実験を行い、感性的尺度として最適なパラメータ α, β, a の値を調べた。

まず、リズムの類似感覚に関する感性評価実験を行った。本実験では、1 対のリズム音を被験者に提示し、それらの主観的な類似度を 5 段階評価させた。なお、提示するリズム音は、リズムマップ生成のために採取した 1 小節長のリズムベクトルサンプルから、無作為に抽出した。提示時のリズムのテンポは BPM=120 とし、被験者は 1 小節のリズムを繰り返し聞くことができる。リズムの音色には MIDI 音源のスネア音を使用した。リズム類似度の評価は、被験者 13 名に対してそれぞれ 100 サンプルずつ行った。被験者の内訳は、楽器経験者 9 名と楽器未経験者 4 名で、いずれも 20 代前半の男女とした。以降の実験では、この実験で得られた感性評価データを規準として用いた。

リズム類似度の構成要素である安定性類似度 D_s と

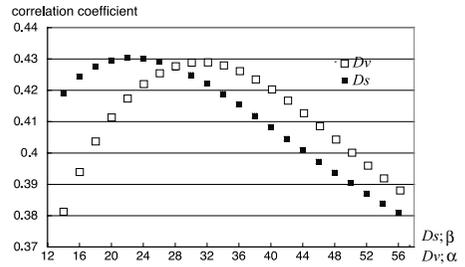


図 10 リズムの安定性類似度 (D_s) および空間的類似度 (D_v) と感性評価データとの相関係数

Fig. 10 Correlation coefficient between human evaluation date and rhythm stability similarity (D_s), vectorial similarity (D_v).

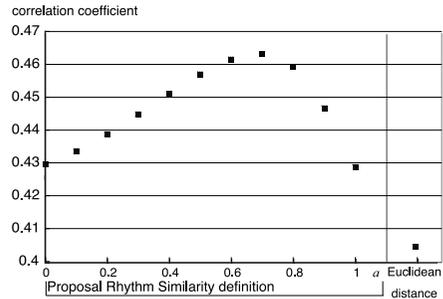


図 11 リズム類似度と感性評価データとの相関係数

Fig. 11 Correlation coefficient between human evaluation date and rhythm similarity definition.

感性評価データとの相関係数を、 α をパラメータとして求めた。また、空間的類似度 D_v と感性評価データとの相関係数を、 β をパラメータとして求めた。結果を図 10 に示す。これより、安定性類似度と空間的類似度は、ともに感性評価データと正の相関があり、リズム類似度の構成要素として妥当であることが分かった。また、パラメータが $\alpha = 22, \beta = 32$ であるとき、これらの類似度要素がリズムの感性的尺度として最適であることが分かった。

次に、リズム類似度と感性評価データの相関係数を、安定性類似度と空間的類似度の重み係数 a の値を変化させて求めた。ただし、これらの要素類似度のパラメータには最適値 $\alpha = 22, \beta = 32$ を用いた。結果を図 11 に示す。結果より、提案するリズム類似度 S_a は感性評価データと正の相関があり、重み係数が $a = 0.7$ であるときに相関係数は 0.463 で最大であった。このときのリズム類似度の相関係数は、構成要素の D_s と D_v をそれぞれ単独で用いた場合の最大相関係数より大きいことから、これらの構成要素を組み合わせることが有効であることも示された。これ以降、最適な重み係数 0.7 を \hat{a} と表す。

また、提案するリズム類似度と従来のユークリッド

距離 S_{Euclid} に基づいた類似度について、感性評価データとの相関係数を統計的に比較した。後者の類似度は次式で定義される。

$$S_{Euclid} = 1 - \sqrt{\frac{1}{16} \sum_{i=1}^{16} \left(\frac{v_i^1 - v_i^2}{127} \right)^2} \quad (17)$$

相関係数は、前述の $S_{\hat{a}}$ の場合 0.463 であったのに対して、上式の S_{Euclid} では 0.404 であった。サンプル数が十分に多いという仮定の下で、「これらの2つの相関係数に差がない」という帰無仮説を両側検定したところ、 p 値は 4.24% となり、棄却域 5% の有意水準において有意な差があることが確認された。

以上の実験により、本稿で提案するリズム類似度の妥当性および優位性が統計的に検証された。

さらに、リズムの音色による感性的尺度の違いを検討するために、リズムの音色にバズドラムを用いて、同様の実験を被験者 4 名に対して合計 400 回行ったところ、スネアの音色による結果と同じの傾向が見られたが、感性評価データに対する \hat{a} , α , β の最適値は、スネアの場合と異なる値であった。音色によるリズム知覚の違いについては、さらなる検証が必要であると考えられる。

3.2 リズムマップの近傍構造の評価

異なる学習条件で生成したリズムマップの近傍構造を調べるために、リズムマップ学習の実験を行った。リズムマップの近傍ユニットには、類似したリズムベクトルが配置される。

ここでは、リズムベクトルの近傍構造を評価する特徴量として、リズムマップの局所平均類似度 \bar{I} と分散 $dev(I)$ を用いる。 \bar{I} は、位置 X のユニットとその 8 近傍ユニットのリズムベクトルに関する平均類似度 $I(X)$ を、 X について平均した値とする。 $dev(I)$ は $I(X)$ の分散とする。リズムマップの歪みが小さいほど、 \bar{I} の値は 1 に近くなり、 $dev(I)$ は小さい値となる。 $I(X)$ の計算には、共通の値 \hat{a} を用いた。

実験では 20×20 の正方形形状のリズムマップを用い、学習データには MIDI 信号として演奏情報を出力可能な楽器で人間が演奏した 600 小節分のリズムサンプルを用いた。

リズム演奏には、YAMAHA 社製の電子打楽器 DTXEXTREME を使用し、打楽器部分にはゴム製パッド TP65 を使用し、演奏時の音色はスネアとバズドラムとした。

学習回数は 200 回、パラメータは $\sigma(0) = 10$, $\alpha = 22.0$, $\beta = 32.0$ とした。リズムベクトルの類似度のパラメータ a を、0, 0.7, 1.0 に固定した場合、

表 1 異なる条件で学習したリズムマップの近傍構造

Table 1 Neighborhood structure of rhythm map with different parameter.

a 値	$a = 0$	$a = 0.7$	$a = 1.0$	a 可変
\bar{I}	0.740	0.724	0.727	0.745
$dev(I)$	0.01573	0.00644	0.004448	0.005065

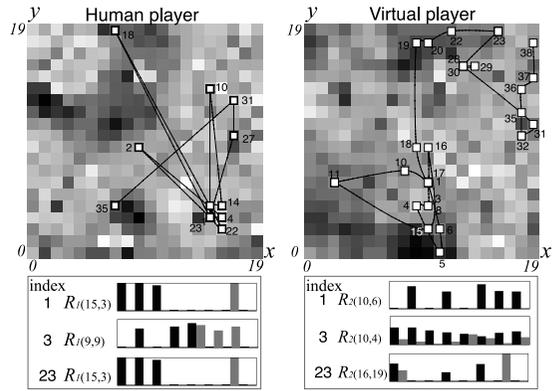


図 12 インタクションにおける選択リズムの軌跡

Fig. 12 Trajectory of performances during the interaction.

および、式 (11) に基づいて変化させた場合の、 \bar{I} と $dev(I)$ の値を表 1 に示す。ただし、0, 0.7, 1.0 の値はそれぞれ、安定性類似度、最適なリズム類似度、空間的類似度に対応する。式 (11) のパラメータは補足実験の結果より、 $A = 1.25$ と $M = 20 \times 600$ とした。

表 1 に示されるように、 a の値が固定値の場合、 a の値が小さいほど、リズムマップの局所的な歪みは小さいが、大局的なばらつきは大きくなる。一方で、 a が可変の場合には、リズムマップの局所的な歪みが小さく、大局的なばらつきも小さくなり、学習がより安定して行われていることが分かった。また、リズムマップを観察したところ、局所的には空間的類似度が大きいリズムが配置されていた。

3.3 ユーザとシステムのインタクション

本システムを総合的に評価するため、ユーザとシステムのインタラクティブセッションを行い、ユーザの行動とシステムの動作を観察した。

インタクションにおいてユーザとシステムが選択したリズムの軌跡を、図 12 に示す。図よりシステムのリズムがユーザのリズムに応じて動的に変化する様子が分かる。システムのリズム選択は直前のエージェントの位置にも依存するため、実際のインタクションでは、ユーザが同じリズムを演奏していてもシステムの演奏は一意的ではなく、異なるリズムを出力することがあった。

図 13 に示されるように、リズムマップの結合強度

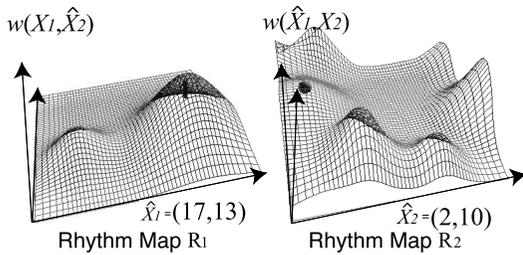


図 13 インタクションにおける結合強度
Fig. 13 Graph of map link.

の曲面形状は、一般に多峰性である。エージェントの直前の位置が異なれば、エージェントは異なる極大点に向かって結合強度の山を登ることもあり、システムの演奏リズムも異なる。リズム演奏は一方のリズム選択だけでは決定されず、本質的にインタクションにより生成される。

インタクションでのユーザの行動を観察したところ、ユーザは結合強度の表示を参考にリズムを選択していたが、「自分の意図したリズムがリズムマップ上のどの位置にあるのか分からない」といった意見もあり、リズムの可視化やラベリング処理などの検討が必要であることが分かった。また、インタクション中に、結合強度の弱いリズムをあえて選択し、演奏の流れに変化を付ける、といった行為も観察された。また、「システムのリズム演奏を変化させるように働かせたい」という意見もあった。

システム全体に関しては、「演奏経験がなくても演奏できることが楽しい」「表示グラフをもとに演奏を変えるとシステムの演奏も変化するので、協調して演奏をしている印象を持った」などの感想が得られた。

本システムのリズムの連想機構では、リズムマップ間の結合強度に基づいて連続的にリズムが変化するため、連続的なセッションアレンジには有効であるが、突発的な応答や 1 小節では完結しないリズム演奏は困難であり、ユニゾンやシンコペーションなどの音楽的意味を持ったリズムパターンの表現ができないという課題が残っている。これらについては、応答モデルの考案や、リズム演奏の時系列変化および楽曲の構造を考慮するなどの解決法を検討していきたい。

また、インタクションにおいて、学習データとは異なる音源を使用したところ、多少音色が異なっても、同じ楽器であれば違和感なくインタクションが行えたが、異なる楽器音の場合は演奏時に若干の違和感があった。今後は、音色まで含めたシステム構築を検討する必要がある。

4. ま と め

本稿では、演奏者のリズム演奏に応じて共創的にリズム演奏を生成するシステムを提案した。本システムでは、2つのリズムマップとそれらの連結機構を、実際の協調リズム演奏データを用いて構築することで、ユーザとシステムの共創的なリズムインタクションを実現した。

人間のリズム知覚に適合したリズム配置を生成するために、リズムの安定度を考慮した類似度を導入し、その妥当性を統計的に検証した。また、リズムマップ間の連結状態を表現する演奏インタフェースを作成し、インタラクティブなリズム想起を試みた。本システムによるデモ音源に関しては、<http://www.nicolb45.com/Research/>を参照されたい。

今後は、リズムの音色や演奏の応答モデルの考慮、複数人が演奏可能なインタクションシステムへの拡張などを検討する。また、演奏者のリズム嗜好を学習させることで仮想演奏者データベースの構築も検討している。

謝辞 本研究の一部は科学技術振興機構 CREST 研究「人を引き込む身体的メディア場の生成・制御技術」の研究助成を受けて行われた。

参 考 文 献

- 1) Raphael, C.: A bayesian network for real-time musical accompaniment, *Advances in Neural Information Processing Systems*, Dietterich, T.G., Becker, S. and Ghahramani, Z. (Eds.), NIPS 14, MIT Press (2002).
- 2) Raphael, C.: Synthesizing Musical Accompaniments with Bayesian Belief Networks, *Journal of New Music Research*, Vol.30, No.1, pp.59-67 (2001).
- 3) 後藤真孝, 日高伊佐夫, 松本英明, 黒田洋介, 村岡洋一: 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, Vol.40, No.4, pp.1910-1921 (1999).
- 4) Taki, Y., Suzuki, K., Konagaya, H., Hartono, P. and Hashimoto, S.: Machine Listening for Autonomous Musical Performance Systems, *Proc. 2002 International Computer Music Conference*, San Francisco, pp.61-64 (2002).
- 5) Tokui, N. and Iba, H.: Music Composition with Interactive Evolutionary Computation, *Proc. 3rd International Conference on Generative Art (GA2000)*, Milan, Italy (2000).
- 6) 武田晴登, 西本卓也, 嵯峨山茂樹: 確率モデルによる多音音楽演奏の MIDI 信号のリズム認識,

情報処理学会論文誌, Vol.45, No.3, pp.670-679 (2004).

- 7) 吉井和佳, 後藤真孝, 駒谷和範, 尾形哲也, 奥乃博: Drumix: ドラムパートのリアルタイム編集機能付きオーディオプレイヤー, インタラクシオン 2006 論文集, 3/2-3/3 (2006).
- 8) 笠原俊一, 橋本周司: 階層構造による音楽リズムの分析と生成, 情報処理学会全国大会講演論文集, Vol.68, No.2, pp.171-172 (2006).
- 9) 笠原俊一, 橋本周司: 階層構造を用いたリズムの構造分析—自己組織化マップによるリズムマッピング, 情報処理学会研究会報告, Vol.2006, No.90, pp.25-30 (2006).
- 10) Kohonen, T.: *Self-Organization and Associative Memory*, Springer-Verlag (1984).
- 11) Longuet-Higgins, H.C. and Lee, C.S.: *Perception of musical rhythms*, Perception (1982).
- 12) 末富大剛, 中島祥好: リズム知覚研究の動向, *Journal of Music Perception and Cognition* (音楽知覚認知研究), Vol.4, No.1, pp.26-42 (1998).
- 13) Hopfield, J.J.: Neural network and physical systems with emergent collective computational abilities, *Proc. National Academy of Sciences of the United States of America*, Vol.79, No.8, pp.2554-2558 (1982).
- 14) 中野 馨: アソシアトロン—連想記憶のモデルと知的情報処理, pp.1-56, 昭晃堂 (1979).
- 15) Kosko, B.: Adaptive bidirectional associative memories, *Applied Optics*, Vol.26, No.23, pp.4947-4960 (1987).
- 16) Kosko, B.: Bidirectional associative memories, *IEEE Trans. Syst., Man, & Cybern.*, Vol.18, No.1, pp.49-60 (1988).
- 17) Hagiwara, M.: Multidirectional associative memory, *Proc. IEEE and INNS International Joint Conference on Neural Networks*, Vol.1, pp.3-6 (1990).
- 18) Les Piegl and Tiller, W.: *The NURBS Book*, 2nd ed., Springer-Verlag (1997).

(平成 19 年 4 月 4 日受付)

(平成 19 年 9 月 3 日採録)



笠原 俊一 (学生会員)

2006 年早稲田大学理工学部応用物理学科卒業。現在, 同大学大学院理工学研究科物理学及応用物理学専攻に在学中。音楽リズムの構造分析, 音楽演奏によるインタラクシオン支援に関する研究に従事。



三枝 亮 (正会員)

1999 年早稲田大学理工学部応用物理学科卒業。早稲田大学助手を経て, 現在, Italian Institute of Technology 博士研究員。ニューラルネットワーク, 情報圧縮, 画像処理等の研究を通して, 機械学習やパターン認識, ロボティクスに興味を持つ。博士 (工学)。IEEE, 電子情報通信学会各会員。



橋本 周司 (正会員)

1970 年早稲田大学理工学部応用物理学科卒業。東邦大学講師, 助教授, 早稲田大学助教授を経て, 1993 年より早稲田大学理工学部応用物理学科教授。また, 2000 年 4 月より早稲田大学ヒューマノイド研究所所長。確率過程の応用, 画像処理, ロボティクス, 音楽情報処理, 等の研究を通して, 人間共存ロボット, メタアルゴリズム, 感性情報処理, ヒューマンインタフェースに興味を持つ。工学博士。主な著書 (共著), 『仮想音楽空間』, 『人間型ロボットのはなし』, 『岩波講座マルチメディア情報学第 1 巻』等。