

バンドライクな音楽アシスタントシステムについて

青野裕司* 片寄晴弘** 井口征士*

*大阪大学 基礎工学部 制御工学科

**イメージ情報科学研究所

音楽用語において、「セッション」とはミュージシャン同士がお互いの演奏を聴き、またそれに反応しながら進めてゆく即興的な作曲作業を意味する。その「セッション」においてより良い曲を作りまたより大きな満足を得るためには、ミュージシャンは注意深く、また迅速かつ正確に周囲の演奏を聴いて反応する必要がある。

我々はこの「セッション」を、人間とコンピュータの間で実現するシステム作成を試みた。このシステムにおいて、人間のプレイヤーは、コンピュータを他のメンバーとして構成されるバンドの一員となる。その人間のプレイヤーは様々な音楽的進行情報（例えば、コード進行・リズムパターン・ベースラインなど）をMIDIインターフェースを通じてコンピュータが担う他のバンドメンバーに送る。一方、コンピュータはそれらを聴き、そして次なる進行を予測し、正確にそしてエキサイティングに反応しなければならない。

Development of Band-like Musical Assistant System

Yushi Aono*, Haruhiro Katayose**, Seishi Inokuchi*

*Department of Engineering Science, Osaka University

**Laboratory of Image Information Science and Technology

In musical terms, 'session' means an improvised composition of tunes with interaction between musicians. To make a good tune, and to get more satisfaction, musicians should listen to and respond to other musicians' playing carefully, quickly, and correctly.

We tried to make a system which realizes this 'session' between a human and a computer. In this system, only one human player is an imaginary member of a band composed of other members performed by a personal computer. A human player gives many musical sequence information (for example chord sequence, rhythm pattern, bass line and so on) to computer-performed band members with MIDI interface. On the other hand, they listen to what, and must estimate next sequences, and make reactions accurately and excitingly.

1 はじめに

ロックやジャズなどの一般に軽音楽と呼ばれるジャンルの音楽は、バンドという形態で行われることが多い。このバンドという演奏形態においては、それを構成する人数にも依るが、メンバー間の意志疎通と相互作用が演奏に際して非常に重要になる。これは、メンバーの数が少なくなるほど顕著で、ドラム・ベースにキーボードもしくはギターなどといった3～4ピースのバンドにおいては、メンバー間のインタラクションは質の高い演奏をする上で欠かせないと言っても過言ではないだろう。さらにこういった小規模なバンドでは、メンバー間のインタラクション自体が楽しみの対象となることがある。つまり、即興的に演奏を進める中で互いの音を聴き、それに何らかのリアクションを加えてゆくことで、ひとつの楽曲を形成してゆく「セッション」と呼ばれるものがそれにあたる。本研究に於いてはこのインタラクションに注目し、人間のプレイヤーとコンピュータの間でセッションを実現するミュージックパートナーシステムの作成を行っている。

現在に至るまでも、ミュージックパートナーシステムにおけるインタラクティブ性というものは数多く研究されてきた。Robert RoweによるCypherは、音楽解析部と演奏部を持ったインタラクティブコンポーシングシステムといえる[1]。これは、人間や他のコンピュータからのMIDI演奏入力をリアルタイムに解析し、あらかじめ設定されたもしくはシステム内部に保存された音楽生成アルゴリズムに従って、即興的に作曲するものである。このシステムにおける主眼は作曲にあり、コンピュータと人間によって独自性の高い楽曲を生成することを目標としていると考えられる。また人間のソロプレイヤーに対してリアルタイムで伴奏するシステムもある。これらのシステムにおける重要な技術的基盤は、楽譜と演奏のマッチングによるテンポ・トラッキングである。Roger B. Dannenbergは楽譜と演奏との間のマッチングをとるために、整数行列演算を用いたシステムを生成し、伴奏システムの先駆をなした[2]。また、B.Vercoeは、同様の伴奏システムで、対象がアコースティック楽器であるシステムを提案している[3]。さらに大阪大学に於いてもインタラクティブ

ミュージックシステムが開発されている。和気らによるジャムセッションパートナーは、本研究と最も位置づけが近いと考えられるが、バンド形態を模したブルースセッションを人間とコンピュータの間で実現するシステムを発表している[4]。但しこの研究に於いては、コード進行があらかじめ限定されるなど、制約も多い。これらを含む今までのミュージックパートナーシステムと本システムが最も異なる点は、伴奏生成に於いてバンド的感触を主眼としている点である。このバンド的感触が何を意味するかについては、以下に順次述べるとして、本システムは楽譜を持たないリアルタイム伴奏システムといえる。

2 バンドにおけるセッション

普通セッションは、全く何も決めずに始めることは希であるにしても、何かの曲を演奏するときには比べると格段に混沌とした状況下でスタートする。例えば、「このキー（曲の主調）で始めよう。」とか、「このリズムパターンを使ってみよう。」とか、「このリフ（フレーズの繰り返し）を何かに発展させてみよう。」とかいった感じである。

（よく行われて有名なものにブルースのセッションがあるが、これは12小節分のコード進行のみ決められており、その繰り返しの中で即興的に演奏するものである。）さらにセッションにおいては、他のメンバーの音楽的解釈で許容されうる演奏なら何を演奏してもかまわないわけで、つまりスタートも混沌としているがその先どの様な曲になってゆくかについてもまたメンバー次第であり、様々な可能性が存在する。ここにセッションの有する二つの重要な特徴があるといえるだろう。

まずひとつは、メンバー間のインタラクションが非常に要求されるということである。これは楽器を通してのコミュニケーションが必要とされるとも言い換えられるだろう。演奏を発展させるためには誰かが何らかのアクションを起こさなければならない。そのためにはまず周囲の演奏をよく聴き、その流れから逸脱しないものを考え出さなければならない。そうやって起こされたアクションに対して、他のメンバーは同様な方法で応答してゆく（図1参照）。つまりミュージシャンは楽器でもって会話をするがごとくに、演奏を進めてゆくわけである。このことは、セッションに取っ

で必要不可欠な要素であると同時に、ミュージシャンには大きな楽しみとなることが多い。

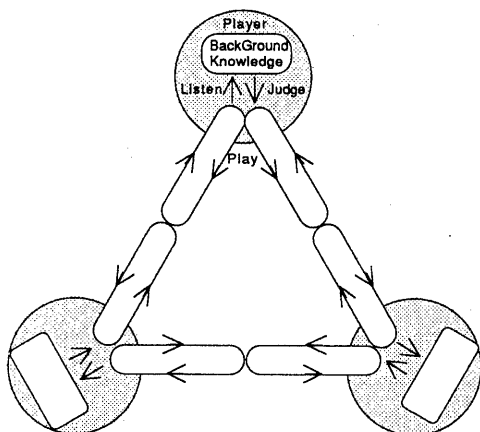


図1. バンドにおけるメンバー間のインタラクション

もう一つはこのセッションが作曲につながり得るということである。実際セッションを行っているうちに、新たな曲の原型やそれについてのアイデアが生まれることはよくある。また逆にセッションを作曲のための一手段として明確に位置付け、行う場合もある。セッションで作曲を行うと、大まかなアレンジまでリアルタイムに行えるので効率的である。一人では考えつかないようなアイデアが他のメンバーによって持ち込まれる可能性がある、などのメリットがある（図2参照）。

各メンバーの音楽性

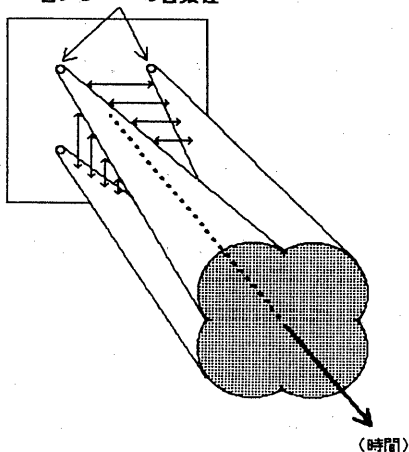


図2. セッションの経過に伴うメンバー間のインタラクションと音楽性の融合

この二つの特徴を持つバンドにおけるセッションの環境をコンピュータと人間の間で実現しようと試みたのが、この研究である。

3 バンドらしさを実現するために

3.0 システムの骨格となる機能と概念

本システムにおいてバンドらしさを実現するために考え出された概念及び機能について、その重要なものを列記すると、

- ・アクセントパターン抽出機能
- ・コード認識機能
- ・フレーズ抽出機能
- ・4小節単位のシーケンス抽出機能、及び保存機能
- ・シーケンスの変化に対する予測機能
- ・任意にシーケンスを組み立てることによる簡易作曲機能

の6つが挙げられる。以下これらの項目について、説明を加える。

3.1 アクセントパターン抽出機能

まずアクセントパターンという言葉であるが、これは筆者がこのシステムを作成する際に新たに作り出した造語である。バンド経験のある方ならなんとなくは想像されるかもしれないが、バンドにはドラムスという楽器があり、これが楽曲のリズムパターンの重要な部分を担っている。さらにドラムスにはバスドラムとスネアドラムがあり、この二つが入るところがリズムパターン中でアクセントをなすことが多い。またドラムス以外のメンバーは、特にセッションの時などにおいては、このアクセントが小節中のどこに位置するかで、大まかなリズムパターンを認識する。またその他の細かなリズムパターンは、曲の進行に従って変化することも希ではない。つまりセッションの最中においては、リズムパターンはそのアクセントだけを抽出したアクセントパターンというより抽象的な形で、保たれるといえる。この考えに基づき本システムではまず演奏入力から、幾つかの決まりに従ってアクセントと思われる場所を推定し、それらを小節単位でまとめたアクセントパターンを作成している。決まりとは、

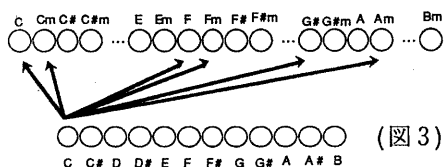
- ・コードの切り替わる場所
- ・ベースラインやソロフレーズの中でヴェロシティの強い場所
- ・2拍, 4拍

などである。そしてアクセントパターンが決定された後は、それを元に伴奏を逐次生成することによって、即興性に富んだ演奏を実現できるよう試みた。

3.2 コード認識機能

セッションを始めるにあたっては、まずメンバー間でコード進行についての意思疎通が図られる必要がある。そこで本システムでもセッションが始まるとまずコード進行をコンピュータ側に教える必要がある。この際コンピュータは3つ以上の音が同時に入力された場合のみコードとみなして、コードネームなどの判別を行う。つまりフレーズの構成音などからコードを推測するといった複雑な作業は行わない。これには二つの訳がある。一つはリアルタイムで伴奏を行うため複雑な処理は望ましくないということ、そして二つ目は実際のセッションにおいて「あるフレーズからコードを勝手に判断する」などということはまず行わないからである。つまり普通のセッションと同様に、演奏の始まった段階でコード進行を他のメンバーに知らせるようにコンピュータにも知らせるのである。

コンピュータ側が、演奏者がコードを入力してきていると判断すると、それらの音からコードネーム及びテンションノート・ベースノートを解析する。その処理はCypher[5]とほぼ同じ方法で、あらかじめ用意された48個の配列に(CからBまで12音に対しM, m, dim, augの48個)入力された音を構成音として含む場合ポイントを与え、全ての入力が終わった段階でもっともポイントが高いコードネームを結果とみなす。さらにそのコードネームのトライアドに含まれない音はテンションノートとみなし、また一つだけ突出して音程の低い音があればそれをベースノートとみなす(図3)。



また、コード認識機能と同時にアクセントパターン抽出機能も働いているので、コードを自分の意図したリズムパターンに乗っ取って演奏すれば、曲に関する重要な情報は一括してコンピュータに与えることができる。簡単な例を図4に示す。



図4.コードネームとアクセントパターンを同時に読み取る例

3.3 フレーズ抽出機能

コードと同様にフレーズも曲の進行において重要な情報となる。現在のシステムでは、ひとまとまりのフレーズを切り出すのに

- ・一つ前のノートとの時間間隔
- ・数個前のノートとのヴェロシティの差
- ・一つ前の音との音程差

といった特徴を用いている[6]。これらの特徴からフレーズの切れ目らしさを計算しその値が一定値を越えた場合に新たなフレーズに切り替わったとコンピュータは認識する。ただしこの際各特徴からの寄与の度合いは、演奏者により前もって設定することが可能である。つまり演奏する曲に応じて、拍節に重きを置くか、音の間隔に重きを置くかなどが決められるわけである。実際に人間がフレーズのかたまりを判断する際には、これらの他にリズムやコード感からの寄与も用いているはずである。今後はそれらの要素も用いる予定である。

フレーズといった際に予想されるものがいくつか存在する。まず演奏者がメロディーラインを弾いている場合、ソロを弾いている場合、そしてベースラインをコンピュータに提示している場合である。この中でもっとも伴奏に影響を及ぼすものは、恐らくベースラインを弾いている場合であり、それ以外の場合は主に曲の雰囲気作りに寄与すると考えられる。しかし現在本システムでは採用し

ているフレーズ切り出しの特徴が少ないので、ソロやメロディーを人間の意図に忠実に切り出すには十分でない。この様な理由から現在ではできるだけ細かくフレーズを切り出すように設定して用いている。そしてそれがベースラインである場合（すなわちフレーズの構成音の大部分が低い音程で占められている場合）のみ保存し、それ以外の場合は音の密度やヴェロシティといった情報から曲の雰囲気来判断するために使い捨て的に用いて保存はしていない。

また同時にアクセントパターン抽出機能も働いており、フレーズからもアクセントパターンを推定することは可能である。

3.4 4小節単位のシーケンス抽出機能、

および保存機能

一般に軽音楽においては、4・8・16など偶数小節がひとかたまりとなっており意味を持つシーケンスを形成する。簡単に言うと、Aメロ・Bメロ・サビ・イントロなどの事である。つまりフレーズを切り出すのとはほぼ同じ考え方に基づき、加えて拍節構造の寄与を大きくとってやることで、シーケンスの始まりと終わりを見つけだすことが可能になる。さらにシーケンスで曲を管理するという事は、複雑な進行を持つ曲でもそれらの組み合わせで容易に理解・表現できることになる。本システムでは拍節構造が最も影響度が高いことを考え、演奏入力から抽出されたアクセントパターン・コード進行・フレーズの情報をリアルタイムで4小節単位のシーケンスにまとめあげ、曲を管理している。システム中にはこの4小節のシーケンスが最大20個まで保存できる。

演奏者はシーケンスを作成する際にとくに留意する必要はない。4小節入力が終わった段階で、現存するいずれかのシーケンスと同じと判断されればそのシーケンスをモデファイするにとどめ、また4小節の途中でも現存しないシーケンスであると判断されるとシステムは新たなシーケンスを展開する。

3.5 シーケンスの変化に対する予測機能

伴奏を連続的に行う場合は、いくつかのシーケンスをつないでゆかなければならない。しかし、現在の軽音楽の楽曲の様式は多岐にわたり、一様

のルールではシーケンスのつながりを予測することは不可能であると思われる。そこで現在は、過去の経緯を記録にとどめ、次に来るシーケンスとしてより確率の高いものを出力し、直後の演奏からその正誤を確認するという方法を取っている。また今後は、予め演奏者に曲の基本構造を指定してもらい（A→B→サビ、サビ→A→B、A→A→サビ、etc.）、その情報と現在行っている確率的予測とを加えた方法にしていきたいと考えている。

3.6 任意のシーケンスを組み立てることによる簡易作曲機能

上記の各機能によって抽出されたシーケンスはいわば曲の構成要素であり、これらを組み合わせることによって作曲が可能になると考えられる。まずこのシステムを用いいくつかのシーケンスをリアルタイムで作成し、それらをブロックを並べるようにディスプレイ上で組み合わせるだけで曲が完成するとしたら、これは非常に簡便な作曲ツールになる。つまりこのシステムは作曲支援ツールとして発展する可能性を持っているといえる。

4 ユーザーインターフェース

図5にこのシステムのユーザーインターフェースを示す。演奏者はこれらの画面からの情報をもとに、自分の演奏がコンピュータにどのように理解されているかを知ることができる。図5の番号にしたがってそれぞれの機能を示す。

- (1)現在の小節数及び拍数を表示するデジタルカウンター。
- (2)録音、停止、再生ボタン。再生ボタンで演奏を開始すると、入力演奏は保存された曲の情報に影響を与えない。
- (3)アクセントインジケーター。演奏された音のヴェロシティを表示する。
- (4)フレーズインジケーター。演奏された音列において、フレーズの切れ目らしさの度合いを表示する。このインジケーターが振り切れレッドゾーンに入ると、フレーズの切れ目と見なす。
- (5)フレーズシグナル。フレーズの切れ目が発見されたときに点灯する。

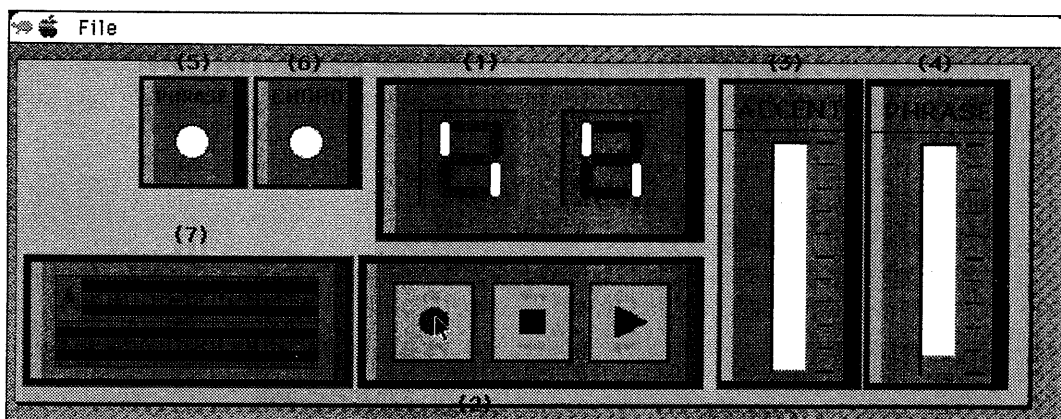


図5.ユーザーインターフェース

- (6)コードシグナル. コードが入力されたとき、システムが見なしたときに点灯する。
- (7)シーケンスボード. 演奏が始まってから、現在に至るまでにシステムが認識したシーケンスの数と名前を表示する。

5 むすび

本稿では、今迄に提案された様々な伴奏システムとまた一つ異なる、バンドライクという性格を持つ新たな試みについて、その基本となる考えと機能の概要及びその動作を述べた。

現在このシステムは、発展途上にあると言える。開発に当たっては、リアルタイム処理に伴って生じる様々な問題をクリアすることが先決であり、今まで述べてきた機能のそれぞれにおいて改良を加える余地が多分に残されているのが実状である。その中でも特に早急に行いたいのは、フレーズ切り出しにおいて判断基準に用いるパラメーターを増やすことである。現段階ではこれらパラメーターが少なすぎるということもあり、人間の意図に忠実であるとは言にくい。この部分が改善されると、フレーズ終端予測もよりの確になり、オカズを挿入するときなどはとくに効果が増すと考えられる。つまり一層バンドらしい伴奏が可能になる。さらに最終的には、作曲支援ツールとしての機能を盛り込んでゆく予定である。

参考文献

- [1] Rowe R. : "Machine listening and composing with Cypher", CMJ, Vol.16, No.1, pp.43-63 (spring 1992).
- [2] Dannenberg R. : "An Online Algorithm for Real-Time Accompaniment" Proc. ICMC, pp.193-198 (1984).
- [3] Vercoe B. : "The Synthetic Performer in the context of live performance" Proc. ICMC, pp.199-200 (1984).
- [4] 和気他. : "演奏者の感情を考慮した協調型演奏システム-JASPER-", 音楽情報科学研究会夏のシンポジウム予稿集, pp.43-46 (1989).
- [5] Rowe R. : "Interactive Music Systems", MIT Press.
- [6] 村尾忠廣 : "クロージャーの客観的測定に基づく構造音の抽出について", 音楽情報科学研究会夏のシンポジウム予稿集, pp.67-72 (1992).