

研究題目	アコースティック楽器の音響信号を利用した協調演奏システムの研究	報告作成者	青野 裕司
研究従事者	青野 裕司, 井口 征士		
研究目的	<p>ジャズやロックなど、軽音楽の演奏方法の一つにジャムセッションと呼ばれるものがある。ジャムセッションの特徴は、即興演奏を主体とする点にある。近年、音楽情報科学の分野において、このジャムセッションを人とコンピュータの間で行う、ジャムセッションシステム（以下セッションシステム）の研究が盛んである。それらの研究では、感性情報を利用したり分散協調環境で実装を行うなど、様々なアプローチのシステムが報告されているが、未着手の問題も残されている。我々はその中で、楽譜に代表される事前楽曲情報の問題と演奏入力楽器の問題に注目する。</p> <p>過去のセッションシステムでは、セッションをはじめる前に楽譜が与えられていたり、和音進行が決まっていたりするものがほとんどであった。しかし実際のジャムセッションはその即興的性質から、作曲や編曲の場として用いられることがある。そこでまず我々は、セッションシステムの持つ作・編曲環境としての可能性を引き出すために、楽譜を用いることなく、入力演奏のみで楽曲を構成することができるセッションシステムの開発をめざす。</p> <p>次に演奏入力楽器について考える。過去のセッションシステムは、MIDIと呼ばれる標準規格を持つ電子楽器を用いるか、そうでない場合は単音楽器を演奏入力に用いている。MIDI楽器はコンピュータと接続することを前提として開発された楽器であり、音の名前や音の強さといった演奏情報を、コンピュータに正確に伝達することができる。しかしそれらの音質や演奏の感触などを、アコースティック楽器（非電子楽器）と比較すると、確実に近づいてはいるが、いぜん差異があり、それらを別種の楽器と考えて使い分けしているミュージシャンも多い。本研究で開発するセッションシステムでは、ピアノやギターなどのアコースティック楽器を入力楽器として用いることとする。この試みはセッションシステム研究の新たな展開であると同時に、音響工学的にも興味もたれる。アコースティック楽器を用いるためには、それらの音響信号から音名などの音楽情報を抽出する必要がある。楽器の音名抽出の研究は、音響工学の分野を中心に行われてきたが、単旋律を対象としたものは実時間で処理が、採譜システムのように和音を含む場合ではオフラインでの処理が可能になっている。我々が目指す事前楽譜なしセッションシステムでは、コードに代表される複合音をリアルタイムに知る必要があるため、従来の方法を用いることはできない。本研究では、セッションに用いるという前提のもとに、その際に必要となる情報を中心に解析抽出するという制約を設け、複合音の実時間解析を行う。中心となる情報とは、コード名、ベースノートと思われる低いピッチの音の音名、リズムパターンなどである。一方、正確な発音時刻や音長など、アーティキュレーションに含まれる情報の解析は行わない。</p> <p>演奏生成に関しては、システム内に備える伴奏パターンを減らし、逆に演奏生成に関する基本的なルールを実装することによって、即興性を高めることを目指す。</p>		

## 研究内容

## [システム概要]

このシステムでは人が演奏において主導的な役割を果たす。一方コンピュータはドラムス、ベース、ギターを担当する。セッションを行うに当たってユーザはまず、コード進行とリズムパターンをシステムに教示する演奏を行う。システムは、演奏解析部と演奏生成部の二つの部分に分けることができる(図1参照)。

## [演奏解析部]

演奏解析部では、与えられた演奏からコード名を認識し、それらの最小反復単位(コード進行とリズムパターンを含む)を検出する。この処理は4つのプロセスで構成される。なお、本システムで扱うコードは3度の積み重ねによる三和音、もしくは四和音で、その入力においては、転回形や開離位置での演奏、さらに装飾音や少量のメロディーの付加を許すものとする。

## (1) FFTを利用した周波数スペクトル解析

ピアノなどの楽器音を対象として考えた場合、スペクトル解析には高い周波数分解能と、広い周波数帯域が同時に要求される。本システムでは、4つの異なるサンプリングレートを持つ1024点のFFTを並列で利用しており、0.5Hzの周波数分解能と10kHz強の周波数帯域を同時に獲得している(図2参照)。

## (2) スペクトル中にあらわれる倍音構造から、コード構成音の候補の選択

対象となる音が単音の場合、周波数スペクトル解析によって得られたパワースペクトルを倍音に添ってシフトさせ、その総和を取ると、基本周波数成分がエンハンスされピッチ抽出を行うことができる(図3参照)。この手法はSubharmonic Summationと呼ばれている。本研究ではこれを複合音のパワースペクトルに利用し、和音の構成音を同定する。ただしこの段階では、実際にはコードに含まれない音選ばれている可能性もある。また、それらの音によってどのような和音が形成されているかも不明である。

## (3) それら候補をもとにコード名の決定

和音の周波数スペクトルでは、倍音によるピークが多数生じたり、転回や開離の状態によってピークの大小関係や位置が変化するため、容易にパターン化できない。しかし3度重ねの三和音もしくは四和音を用いるという条件のもとでは、SHSスペクトルのピーク位置を音名で考えることによって、簡単に和音名を決定することができる。例えば入力された和音がC7であったとすると、C、E、G、B<sup>♭</sup>という4種類の音名の音が任意の転回形、および位置で弾かれていることになる。このときSHSスペクトルには、C、E、G、B<sup>♭</sup>の音名に相当する顕著なピークと、それ以外の相対的に小さなピークがあらわれる。ここで顕著であったピークの音名について、完全5度上(オクターブが変わる場合は完全4度下)の音名とペアを作ると、(1)C-G、(2)E-B、(3)G-D、(4)B<sup>♭</sup>-Fという7種類の音名を含む4つの組み合わせができる。このうちペアの両方が顕著なピークを示すのは(1)のみで、このペアは入力された和音の根音と5度音である。根音-5度音ペア検索は、上記の現象の逆問題を考え、ともに顕

著なピークをもつ完全5度離れたふたつの音名はその時の和音の根音と5度音であるとみなして和音名を決定する方法である。この方法を用いた場合、転回形や開離位置での演奏に対しても、1オクターブ内12種類の各音名について最大ピーク値が分かれば、原形や密集位置での和音と同様に扱える。我々が提案するこの手法は過去の採譜システムにはないもので、テンプレートマッチングなどのパターン認識手法に比べ、処理が速い上に認識率を向上させることが可能になる。

この方法は、単音によって励起される倍音のパワーが、次数が高くなるに従って減少するということを仮定しているが、実際にFFTのスペクトルを観察すると、ピアノやギターなどセッションで利用するアコースティック楽器の場合、この仮定はほぼ満たされている。

図4では、Cmajorの原形であるドミソ、C△7の第二転回形であるソシドミ、そして、C△7を開離の位置で弾いたドシミソの3入力について、SHSおよび根音-5度音ペア検索がどのようになるか示してある。図中便宜上、倍音を単独で聴いたときの音名も付記する。この図では演奏されたすべての音の基本周波数のパワーは一定で、倍音は一定の変化量で減衰する場合について示しているが、実際の演奏では常にこのようになるわけではない。打鍵のばらつきによって構成音のいずれかの音が極端に小さい場合や、構成音の一部を省略した場合などは、認識において誤る可能性がある。また、major7th系、minor7th系の和音が入力された場合、真の根音-5度音ペアのWeightの値を(長・短)3度上のペアのWeightの値が下回る場合がある。そこで最終的に和音名を決定する前に、根音に対して長・短6度上(オクターブが変わる場合は短・長3度下)の音名に相当するピークを探し、他の構成音に相当するピークと同程度以上の大きさを示している場合は、誤りとみなし(これを3度上誤りと呼ぶ)、和音名を再決定する。

#### (4) コード進行がくり返される最小単位の検出

本研究がモデルとして取り上げたセッションの形態では、一般的に4小節や8小節といった単位で繰り返されることが多く、システムにとっては、すみやかにそれらを発見する必要がある。

本システムでは、相補的な2つの方法で反復単位の認識を行っている。まず1つは、曲の主調に基づく認識である。演奏解析部では、和音名の認識と同時に曲の主調の簡単な推定を行う。ここでは、周波数スペクトル解析によって得られたパワースペクトルをすべて積算し、さらにオクターブを無視し12個の音名のパワーに集約する。この集約されたパワー列と、12種類の長音階との間で相関をとり、最も相関の高いものをその時のスケールとし、主調を推定する。その主調における主和音もしくは属和音をもって反復単位の終止とするこの方法は、認識における時間遅れがないという特徴をもつ。逆に、反復単位中に起こった転調により、正確に主調が推定できなかった場合は、終止位置を誤認識しやすいという欠点も存在する。もう1つは、現在のコード進行と過去のコード進行との間で、数小節にわたって相関をみる方法である。この方法は、認識に必ず2回以上の反復単位を必要とするが、より正確である。このときコード誤認識の悪影響を抑えるために、図5に示すような2次元平面を複数用いる。それぞれの平面が4, 8, 12, 16小節分の長さを持ち、縦軸は和音名に対応している。このとき、3度上誤りの起こりやすい和音名が隣接するように配置してある。例えば、Aminorの両端はFmajorとCmajor、Cmajorの両端はAminorとEminorになる。そして、認識された和音名のセルには3を、その両端のセルには2を、さらにその外側のセルに1を代入する。このようにすると、2値的に値を代入した場合などに比べ、誤認識がおこった時の相関係数が向上し、反復単位の認識が容易になる。図5では数字の代わりに明るさで表している。4種類の平面について並行して計算を行い、ある

長さの平面の組の相関係数が閾値を越えた段階で、繰り返し単位がその平面のもつ小節の長さであると確定される。4, 8, 12小節の繰り返しパターンで構成される曲について、この方法で相関係数を計算したときの時間推移を図6に示す。例えば4小節パターンの曲の場合(図6最上段)、演奏の入力が始まってから8小節経過した時点で、4小節の長さをもつ類似和音平面を用いた相関係数が(図中実線)ほぼ1に近い値となる。この結果、システムは反復単位の長さが4小節であることを認識し、そのために要した時間は反復単位2回分の8小節である。

#### 【演奏生成部】

演奏の生成は、セッション中の人の思考をモデルとした「ミュージカルジェスチャーの再構築」という手法で行う。人が即興演奏、特にインタラクティブな演奏を行う場合、過去の演奏情報を記憶し、再現もしくは変形させながら演奏を進める。このとき過去の演奏情報を楽曲として完全な形で記憶しているとは考えにくく、なるべく覚えやすいように抽象化もしくは情報の圧縮・切り捨てを行う。ミュージカルジェスチャーとは、例えばリズムパターンやコード進行のように、それらの時系列データが単一の意味・性質を持つように適度に分離された音楽要素ということができる。そしてそれらを再度組み合わせることにより新たな演奏を生成する。この再構成の過程において生じる微妙な変化が演奏に多様性をもたらし、即興演奏の基本となる。本システムが対象としているミュージカルジェスチャーは、コード列(コード進行)とリズムパターンである。このうちコード列は、演奏解析部の結果をもとにコードの切り替わりを追跡することで形成される。リズムパターンは、バンド形態の演奏ではドラムスなどの打楽器によって提示される場合が多い。しかし本システムの入力楽器はピアノやギターなどの非打楽器であるので、その演奏からリズムパターンを推定する必要がある。そこで、(1)コードの切り替わる場所 (2)相対的に音量の大きい場所に聴感上のアクセントがあると仮定し、それらを時系列的に並べたアクセントパターンを、リズムパターンとして用いる。得られたミュージカルジェスチャーは個別のパターンとしてシステム内に保存され、数小節ごとに(現行システムでは1小節ごとに)再度それらを組み合わせることによって基本演奏パターンを生成する(図7参照)。再構成を行う際、音の選択や発音時間(リズム)の選択に自由度を持たせることにより、演奏は作り出されるたびに微妙に異なってくる。しかし基本演奏パターン中の音数は、アクセントパターン中のアクセントの数に依存するので、音の隙間(休符)が多いことがある。このような隙間を埋めるために装飾音が付加され、最終的な演奏となる。装飾音は、そのときの和声的制約を受けながら、システムに実装された音楽的知識を用いて自動生成される。音楽的知識とは、あるコードのもとで使用できる音の、使い方や組み合わせ方に関する妥当性と頻度を記述したものである。以上の処理の結果、コンピュータの出力する演奏は、シーケンサ的な画一的なものとは異なり即興性豊かなものとなる。

研究のポイント	<p>○セッションシステムの持つ作・編曲環境としての可能性を引き出すために、楽譜を用いることなく、入力演奏のみで楽曲を構成することができるセッションシステムの開発を行った。そのために、入力演奏の繰り返しのパターンを自動的に認識する手法を提案した。</p> <p>○従来の音楽アプリケーションはMIDI楽器を対象としたものが多数を占めていたが、本システムは一般的なアコースティック楽器を対象とした。そのため楽器の発する複合音について認識を行う必要が生じるが、本研究で提案する演奏解析のプロセスは、実時間で和音名の解析を行うことが可能である。</p>
研究結果	<p>従来MIDI楽器を用いるのが一般的であったセッションシステムに、アコースティック楽器を導入する一つの試みを示した。我々が提案する手法により、和音構成音のピッチや正確な発音時刻など、採譜システム等で必要となる詳細な音楽情報は解析の対象としないという制約のもと、実時間で和音名の抽出を行うことが可能となった。軽音楽では一般的に用いられる3度重ねの四和音を中心に、さまざまな条件でこの手法の認識率について実験を行った結果、90%を越える認識率を示し、セッションを行うにあたって十分な性能があることを確認した。また、曲の反復単位を自動的に認識することにより、事前の情報としてはテンポを与えるだけで、自由度の高い楽譜無しセッションを実現した。本研究では反復単位の認識に、和音名の誤認識の影響を緩和する仕組みを取り入れ、最短で2回の反復単位の提示でそれを認識することが可能になった。</p>
今後の課題	<p>和音名認識の時間分解能については、MIDIシステムと比べると現段階では不十分といわざるを得ない。しかしこの点については、試作システムを用いた予備実験で、向上の余地があることが分かっており、今後改善していく予定である。</p> <p>楽譜無しセッションでは複数の反復単位を演奏中に切り替えることがある。また、今後作曲支援システムとして発展させるためにも、反復単位の切り替えは重要な機能となる。しかし、実際のセッションでは切り替えに際して、ジェスチャーやアイコンタクトなど、音以外の情報を用いている場合が多い。これをどのように実現するか、また音情報のみで切り替えを行うことができるかについて、検討を行う予定である。</p> <p>現在は、入力楽器としてピアノ、エレクトリックピアノを中心に開発を行っている。エレクトリックギターも、多少コードの認識率が低下するものの、セッションが可能であることを確認している。ギター系の楽器はMIDI化が難しいとされており、その普及も鍵盤楽器に比べて極めて低い。引き続き、これらの楽器を視野に入れながらシステムの改良を行っていく。</p>

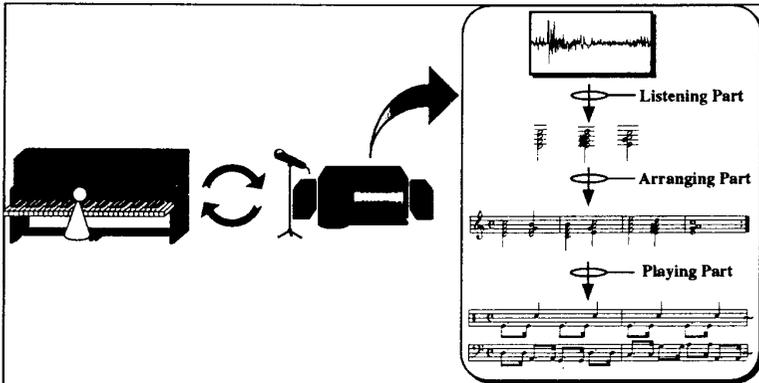


図1 システム概要

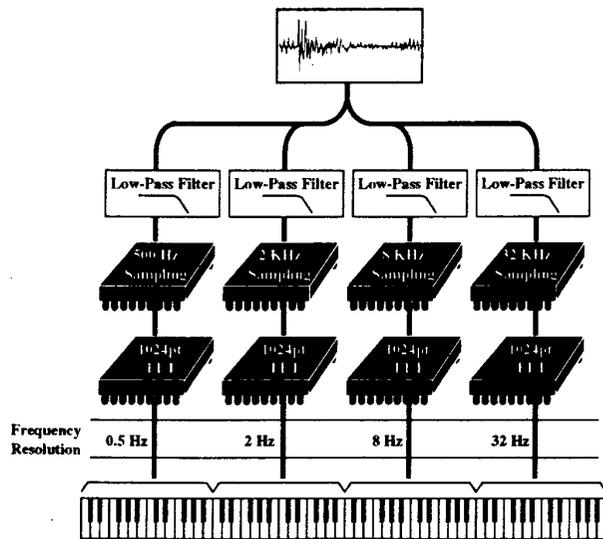


図2 スペクトル解析

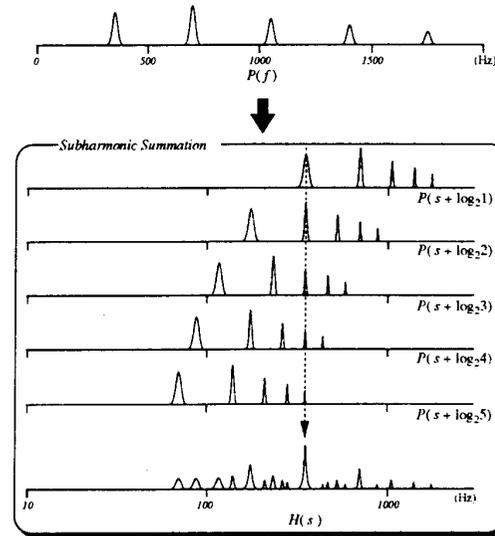


図3 Subharmonic Summation

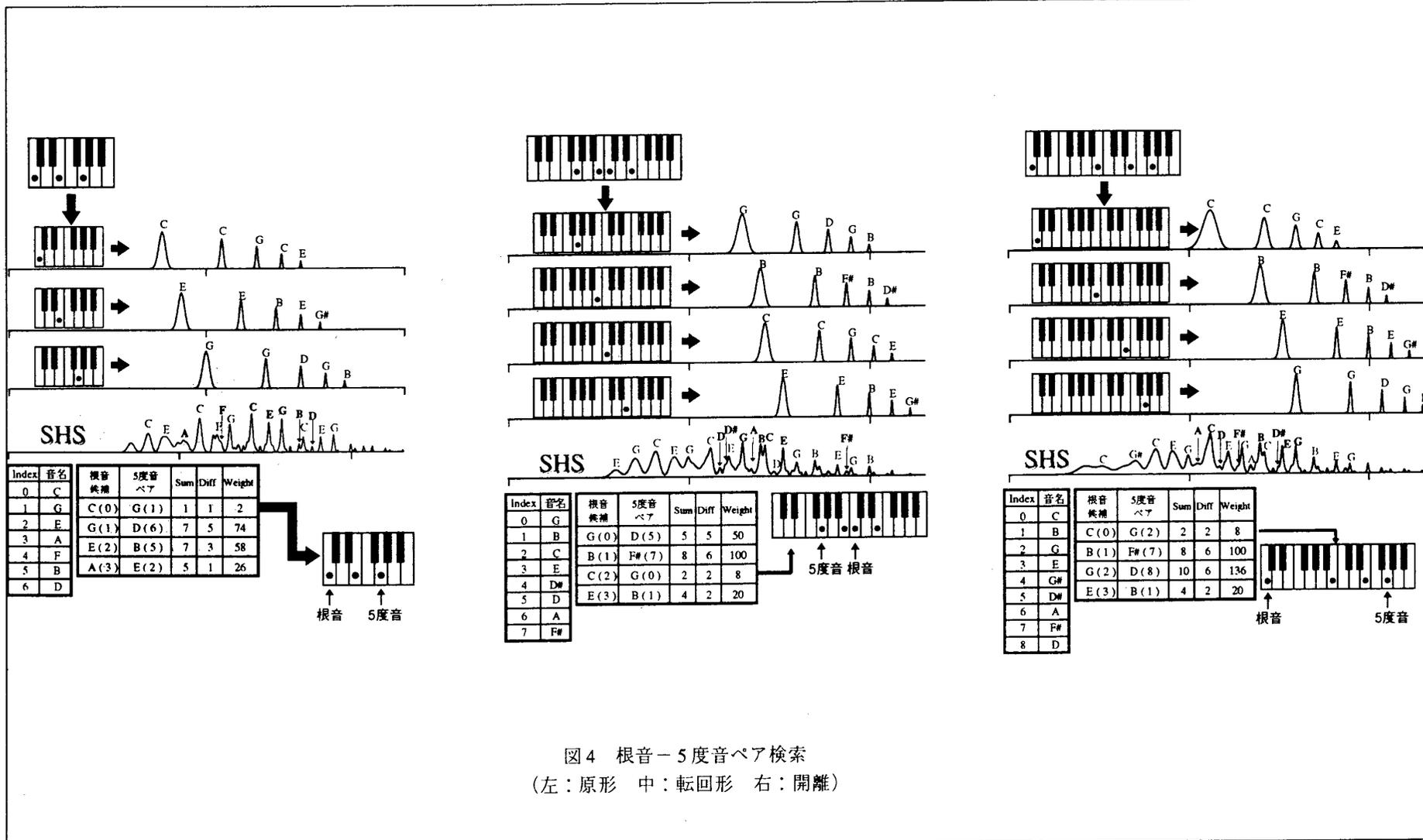


図4 根音-5度音ペア検索  
(左:原形 中:転回形 右:分離)

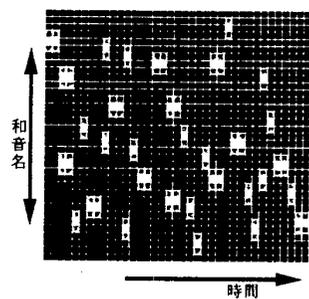


図5 類似和音平面

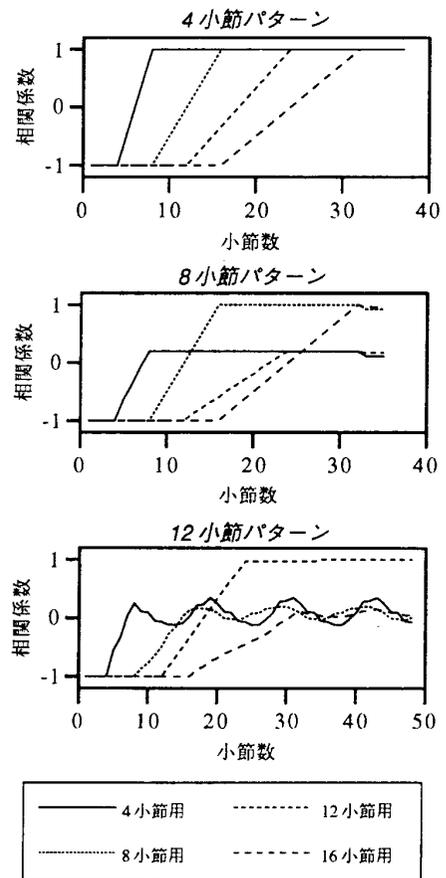


図6 3種類の長さの反復単位に対する相関係数の推移

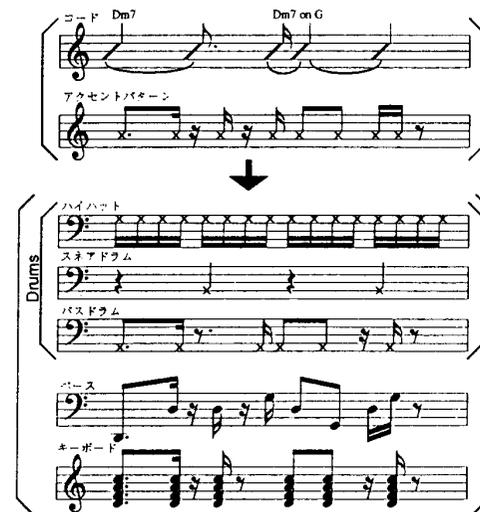


図7 基本伴奏パターンの生成例