

相互引き込みモデルを用いたアンサンブルシステムの開発[†]

小林 洋平*・三宅 美博*

New Ensemble System Based on Mutual Entrainment Model[†]

Yohei KOBAYASHI* and Yoshihiro MIYAKE*

Mutual entrainment of rhythm is widely observed in the music performance between humans. However, one-sided action is dominant in the performance between human and musical machine. Thus, in this research, we try to realize human-like ensemble system by using mutual entrainment between human and machine, and constructed it by using coupled non-linear phase oscillator model. We also evaluated the system from objective side by using phase difference between human and machine, and from subjective side by using questionnaire to the performer. As the results, our system was estimated as more human-like than previous system, and the effectiveness of our model and the importance of mutual entrainment was shown.

Key Words : ensemble system, music ensemble, mutual entrainment, phase oscillator

1. はじめに

近年の音楽機器の発達により、人間の演奏が機械に置き換えられてきている。たとえば、カラオケは人間の伴奏を機械が担当することにより、どこでも手軽に歌を楽しめるようにしたものである。また、商用のレコーディングの現場などにおいても、機械演奏は多用されている¹⁾。しかし、実際の人間の演奏、特にアンサンブルなどと比較すると現状の機械は及ばない。

上記のカラオケなどの例で、歌い手が機械の演奏と人間の演奏に差を感じる理由として、機械の演奏が、人間の行なっているような協調動作が少ないということが報告されている²⁾。そこで、機械が人間の演奏に適應する方法が提案されている。特に、機械の演奏を人間の演奏に合わせるということを目的にしたものが多い³⁾。たとえば、人間の演奏ともっともマッチングの高い伴奏を楽譜データから検索し、それに合わせた演奏を行なう方法である^{4)~9)}。しかし、ここで提案されている方法は、一方向的に人間が機械を制御することを目的としている。そのため、人間の演奏で行なわれている双方向的な作用とは異なるため、これと比較すると差が生じてしまう。また、ジャムセッションシステムのように、一方向化された作用関係を交互に繰り返すことによって、双方向性を実現しようとするアプローチもある^{9),10)}。たとえば、西嶋らの掛け合いソロ演奏を実現するシステムの研究がよく知られている¹¹⁾。しかし、この方法も人間の演奏で行なわれている相互作用とは異なるため、違った印象を受けるという問題点がある。

このような背景から本研究では、より人間らしいアンサンブルシステムを実現するためには、演奏者同士の双方向的かつ同時的な作用が必要であるという作業仮説を導入する。そして、この相互作用をどのように構成するのか、構成した相互作用の人間らしさをどのように評価するのかという問題を解決する。

まず、前者の問題を解決する方法として、人間同士の演奏を上記の相互作用の観点から解析し、機械に組み込むことが考えられる^{12)~14)}。特にここでは、テンポやリズムなどの演奏者が演奏を行なうタイミングに着目して、相互作用を解析する必要がある。そして、その解析結果から得られたダイナミクスを機械に応用し演奏を行なわせることによって、再構成できると考えられる。また後者の問題を解決する方法として、人間同士の演奏をリファレンスとして評価することが考えられる。特に、本研究での人間らしさの定義と評価としては、チューリングが提唱するテストと同様のものを用いる¹⁵⁾。このテストでは、相手が機械か人間かわからない状況で、被験者が質問とその回答から相互作用を通じて相手がどちらであるかを判断するというものである。同様に、アンサンブルのパートナーが機械か人間かわからない状況を設定し、そのパートナーとの演奏による相互作用を通して、比較に基づいた評価で人間らしさとは何かの指標とする。

そこでわれわれは、まず第2章で、人間同士のアンサンブルの相互作用を解析する。そして第3章においてその相互作用モデルを作成し、第4章においてアンサンブルシステムを構築する。第5章でパラメータの設定を行ない、最後に第6章において人間同士のアンサンブルとの比較を行なう。特にここでは主観的評価を行ない、客観的評価と対比させながらシステムの有効性を論じる。

[†] 第41回計測自動制御学会学術講演会で発表(2002・8)

* 東京工業大学大学院総合理工学研究科 横浜市緑区長津田町4259

* Tokyo Institute of Technology, 4259 Nagatsuta, Midori, Yokohama

(Received August 11, 2003)

(Revised July 28, 2004)

2. 人間の演奏の相互作用解析

本章では人間同士の演奏の解析指標として、演奏のタイミングとその時間発展を用いる。すでに、人間の演奏機構を解析する研究として、人間の発音時刻のモデル推定をしたものがある¹⁰⁾。この研究では、機械の演奏のテンポ変動に対して、どのように人間の演奏が影響を受けて変化するかが解析されている。ここでは、同様の解析を二人の人間の共同演奏に拡張し、特にリズムの相互作用に重点をおいた実験を行なう。

リズムの相互作用について解析を行なうためには、演奏者の音楽的解釈の影響をできるだけ少なくし、作用関係の同定が容易になるような演奏を行なう必要がある。そこでこの実験では、二人の演奏者がFig.1に示したような単純な曲を2オクターブずらして同時に演奏し、打鍵周期や打鍵のずれの時間発展から作用関係の同定を試みる。この曲は打鍵周期が常に8分音符と一定周期になるように構成されており、演奏者同士の演奏内容が同じであるため解析が容易である。

それぞれの演奏者はFig.2のように、直接的な相互作用が遮断された状況で演奏を行なった。そして、演奏者は自分の演奏と相手の演奏をヘッドフォンを通して聞く。演奏実験はピアノ教育を10年以上受けた5人の20代学生を被験者として行なわれた。どの被験者もツェルニー100番練習曲程度の演奏をこなすことができる技術がある。また、実験は明るさ約2000ルクス、室温20℃程度の静かな部屋で行なわれた。5人の演奏者については組み合わせを変えて計5回の演奏を行なわせた。それぞれの演奏者はFig.1の曲について演奏を行なう十分な技術があり、また、ミスタッチなどのエラーがでないよう十分に練習を行なっている。ここでは両者のリズムの相互作用を調べるこ



Fig.1 Music score for ensemble

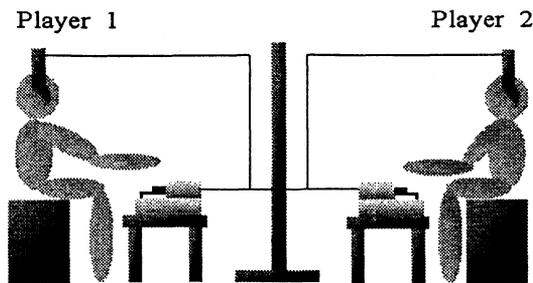


Fig.2 Ensemble between human and human

Table 1 Result of humans ensemble

Trial	1	2	3	4	5	All
Average of hitting cycle (msec)	214	194	191	195	191	197
Standard Deviation of hitting cycle (msec)	17.0	14.6	11.5	14.3	14.7	14.5
Average of hitting lag (msec)	22.1	8.2	1.7	8.2	0.0	8.1
Standard Deviation of hitting lag (msec)	20.0	25.3	22.4	25.2	22.9	23.2

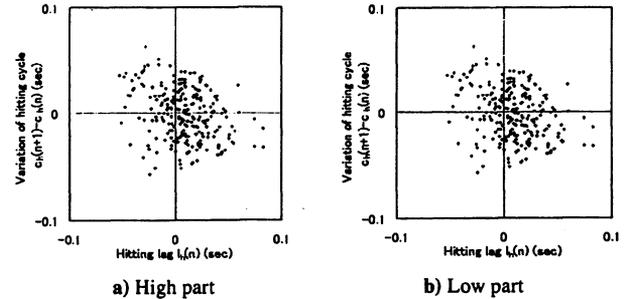


Fig.3 Relation between hitting lag $l(n)$ and variation of hitting cycle $c(n+1)-c(n)$ of each performer

Table 2 Correlation between hitting lag $l(n)$ and variation of hitting cycle $c(n+1)-c(n)$ of each performer

Trial	1	2	3	4	5	All
High	-0.29	-0.29	0.30	-0.29	-0.34	-0.30
Low	-0.42	-0.38	-0.33	-0.38	-0.39	-0.38

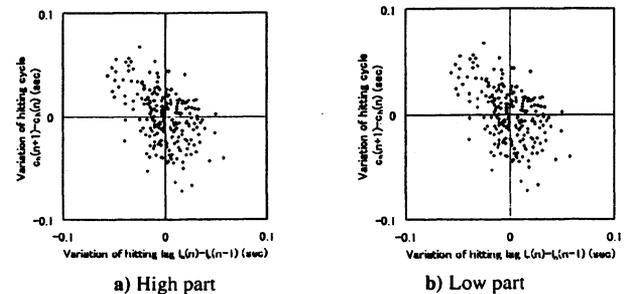


Fig.4 Relation between hitting lag variation $l(n)-l(n-1)$ and variation of hitting cycle $c(n+1)-c(n)$ of each performer

Table 3 Correlation between hitting lag variation $l(n)-l(n-1)$ and variation of hitting cycle $c(n+1)-c(n)$ of each performer

Trial	1	2	3	4	5	All
High	-0.40	-0.48	-0.60	-0.45	-0.50	-0.49
Low	-0.32	-0.45	-0.50	-0.48	-0.42	-0.43

とに重点を置き、両演奏者には演奏を同期させるように求め、楽曲の音楽的解釈などが入らない範囲で自由に演奏を行なわせた。計測は、両演奏者の n 番目の打鍵時刻 $t_h(n), t_l(n)$ を PC を用いて 1ms 単位で計測する。ここで、 $t_h(n)$ は高音側、 $t_l(n)$ は低音側の打鍵時刻である。そこから打鍵の時間間隔 $t_h(n)-t_h(n-1), t_l(n)-t_l(n-1)$ を計算し打鍵周期 $c_h(n), c_l(n)$ として、また両演奏者の打鍵の時間差 $t_l(n)-t_h(n), t_h(n)-t_l(n)$ を打鍵のずれ $l_h(n), l_l(n)$ として算出する。高音部側と低音部側の打鍵のずれは、 $l_h(n)=-l_l(n)$ であり符号が逆転するだけである。演奏は 40 小節以上行ない、演奏が安定するまでにある程度の時間が要することを考慮して、最初の 10 拍を削除し、そこから 128 拍を有効なデータとして使用

することとした。ここから、両者の打鍵周期 $c_h(n), c_m(n)$ 、打鍵のずれ $l(n)$ の平均と標準偏差を計算する。

まず試行ごとの打鍵周期、打鍵のずれの平均と標準偏差を Table 1 に示す。表中の ALL は 5 試行の平均である。このように打鍵のずれは周期と比較して非常に小さく、常に同期をとりながら演奏をしていることがわかる。この方法としてわれわれは 2 種類の指標を用いることを予想した。一つは相手の打鍵のずれ $l(n)$ を指標とする方法であり、もう一つはその変化量 $l(n) - l(n-1)$ を指標とする方法である。このどちらを主として用いているかを調べるために、打鍵周期変化との相関を計算する。

まず、打鍵のずれ $l(n)$ と自分のその直後の打鍵周期変化 $c(n+1) - c(n)$ の関係を Fig. 3 に示す。Fig. 3a が高音部の演奏者 $c_h(n+1) - c_h(n)$ への影響、Fig. 3b が低音部の演奏者 $c_l(n+1) - c_l(n)$ への影響を示しており、負の関係性があることがわかる。各試行の相関係数を両演奏者についてまとめたものが Table 2 である。この表では、High が高音部演奏者への影響、Low が低音部演奏者の影響を示している。この表より相関の値は $-0.3 \sim -0.4$ であることがわかる。

これに対して、打鍵のずれの変化量 $l(n) - l(n-1)$ とその直後の打鍵周期変化 $c(n+1) - c(n)$ の関係を Fig. 4 に示す。Fig. 4a が高音部の演奏者 $c_h(n+1) - c_h(n)$ への影響、Fig. 4b が低音部の演奏者 $c_l(n+1) - c_l(n)$ への影響を示しており、これも負の関係性があることが確認できる。各試行の相関係数を両演奏者についてまとめたものが Table 3 である。この表では、High は高音部演奏者への影響、Low は低音部演奏者への影響を示している。この場合では、 $-0.4 \sim -0.5$ 程度の相対的に大きい値を得ることができた。Table 2 と比較した場合、Table 3 の相関係数の平均値が有意に大きいことも確認された ($p < 0.01, t$ -test)。

なお、このずれの変化量 $l(n) - l(n-1)$ は展開すると $t_h(n) - t_l(n) - (t_h(n-1) - t_l(n-1))$ となり、両演奏者の打鍵周期の差 $c_h(n) - c_l(n)$ でもある。この作用がより強いということは先行研究においても示唆されており、本実験により、人間同士の 2 者間アンサンブルでも同様の現象が存在が確認された^{12), 16)}。このように、人間同士の 2 者間アンサンブルでは、両者がお互いの打鍵のずれとその変化量に基づいて、相互作用していることが示された。ただし、ここで解析したものは相互作用の定常的な側面であることに注意しなければならない。

3. 相互引き込みに基づいたアンサンブルシステムの提案

3.1 相互引き込みモデルの提案

前章の二つの作用関係のうち後者の方が強いことから、第一次近似として打鍵のずれの変化量によって、それぞれの打鍵周期が影響を受ける現象としてまとめると、以下ようになる。それぞれの打鍵周期を c 、打鍵のずれを l とすると、打鍵周期の変化量 \dot{c} と打鍵のずれの変化量 \dot{l} の関係はある負の定数 k を用いて、

$$\dot{c} = -kl \quad (1)$$

と表現することができる。つまり打鍵のずれの変化量 l が周

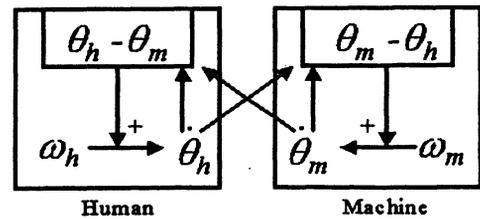


Fig. 5 Architecture of interactive performance system

期変化量 c にフィードバックされる作用関係である。これは積分すると、

$$c = kl + a \quad (2)$$

と表現することができ、ある定数 a を中心として打鍵のずれ l が打鍵周期 c にフィードバックされていることがわかる。打鍵のずれからの作用が無い場合 l は 0 で、 a と c は同じ値になるため固有周期であることがわかる。ここで、これを振動子が影響を受ける作用として、固有周期 a を固有振動数 ω を用いて $2\pi/\omega$ 、打鍵周期 c をみかけの振動数 θ を用いて $2\pi/\theta$ 、ずれ l を位相差 ϕ とみかけの振動数 θ を用いて $2\pi\phi/\theta$ で記述すると、

$$\dot{\theta} = \omega - k\omega\phi \quad (3)$$

となり、固有振動数を基準としてみかけの振動数が変化する現象であることがわかる。つまり、位相差がみかけの振動数にフィードバックされており、引き込み現象と同じ作用関係であることがわかる¹⁷⁾。引き込みは 2 つの振動子が相互に影響を及ぼし、固有振動数を基準としてみかけの振動数が一致する現象であり、非線形振動に特徴的なものである。

このような引き込み現象は人間の身体的な相互作用にも広く観察されている。たとえば母親のかけた声と同調して新生児が体を動かすという報告がある¹⁸⁾。また、成人の対面コミュニケーションにおいて、うなずき動作が発話に引き込まれたり、発話の間の取りかたが引き込まれるという報告もなされている^{19), 20)}。そして、これらを利用したコミュニケーション支援の方法も提案されている²¹⁾。また、歩行リズムの引き込み現象も観察されている²²⁾。さらに音楽においても、呼吸が音楽の小節周期に引き込まれるという現象が報告されている²³⁾。

引き込みからアンサンブルシステムを捉えると、従来までの装置と、われわれが目指す装置を切り分けることが可能となる。それは強制引き込みと相互引き込みである。強制引き込みは一方の振動数が固定であり、固定された振動数に他方のみかけの振動数が同調していく現象である。このとき実現される作用関係は振動数が固定された系から変化する系への一方向のみとなる。これは音楽演奏に対応させるとカラオケにあわせて歌うような一方向的な状況であったり、機械が一方向的に人間に適応するような状況である。これに対して相互引き込みは双方のみかけの振動数が変化し同調していくことである。これは、人間同士のアンサンブルで実現されている双方向的かつ同時的な相互作用に近く、われわれが目指すところである。

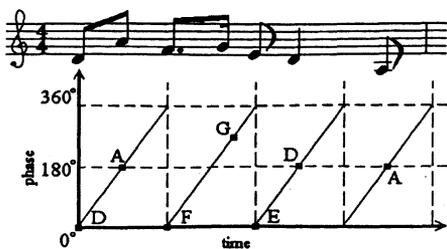


Fig.6 Relation between phase and notes

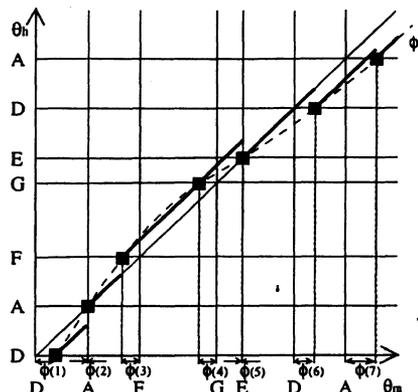


Fig.7 How to simulate phase of human and calculate phase difference

3.2 アーキテクチャ

人間と機械が協調して演奏を行なえるような装置を実現するための基本的なアーキテクチャをFig.5に示す。この図で左側が人間、右側が機械である。ここでは人間側および機械側の演奏の基本のテンポが、おのおの固有振動数 ω_h と ω_m で表わされている。この固有振動数は前節における(3)式の定数 ω に相当する。また、そのような基本のテンポに基づいて演奏されるテンポは、それぞれ人間側と機械側で位相 θ_h と θ_m という演奏位置の時間発展として表わされている。この速度 $\dot{\theta}_h$ と $\dot{\theta}_m$ は、(3)式の θ に相当する。さらに、両者の間での相互引き込みは、それぞれのリズムの位相差に基づき規定されるから、(3)式の ϕ に相当する $\theta_h - \theta_m$ 、 $\theta_m - \theta_h$ がフィードバック信号として、人間側と機械側の演奏過程に再入力される。これらの構造を仮定することにより、両者の演奏はおのおのの基本のテンポだけでなく、両者の位相関係に基づいても規定されることとなり、それぞれのみかけの振動数が時々刻々と変化し演奏されるテンポが相互に引き込まれる。

3.3 位相振動子を用いたモデル化

上記のようなアーキテクチャを実現するために、(3)式をもとにして、人間側と機械側の演奏の時間発展を以下のようにモデル化する。(3)式を上記のアーキテクチャに沿うように書き直すことにより、以下のような非線形な位相振動子モデルとして表現できる¹⁷⁾。

$$\dot{\theta}_m = \omega_m + K_m \cdot \sin(\theta_h - \theta_m) \quad (4)$$

$$\dot{\theta}_h = \omega_h + K_h \cdot \sin(\theta_m - \theta_h) \quad (5)$$

ただし、 ω_m と ω_h は、それぞれ機械側と人間側の固有振動数であり(3)式の ω に相当する。 θ_m と θ_h は、それぞれ機械側と人間側の位相であり(3)式の θ に相当する。また、 K_m と K_h は、それぞれ機械側と人間側における、相手側の振動子との結合係数を表わしており(3)式の $-k\omega$ に相当する。また、 $\sin(\theta_h - \theta_m)$ は(3)式の ϕ に相当する。ここで相互作用に \sin を使用する理由としては、この相互作用が周期的なりズムの引き込みであり、その周期を1拍単位とするためである。こうすることにより、曲の基本単位となる拍単位での相互作用が可能となる。このとき、固有振動数 ω が大きいほど演奏されるテンポは速くなり、小さいほどテンポは遅くなる。位相 θ は、楽譜上の演奏位置に対応し、 θ の値が増えると演奏が進行する。ただし前述したように、 \sin を使用したモデルであるため、扱える範囲は、 θ_h と θ_m の差が1拍以内であることを注意しなければならない。さらに、結合係数 K は共同演奏における位相差からの影響のされやすさに対応している。

カラオケのような一方的な演奏機器では $K_m=0$ であり、常に固有振動数と同じ一定の基本テンポで演奏が行なわれる。しかし、双方向的システムでは $K_m>0$ を用いることにより、時々刻々と相互引き込みしつつ、みかけの演奏テンポが時間発展することに特徴がある。ただし、実際にわれわれがアンサンブルシステムに組み込むのは、機械側のモデルだけであることに注意しなければならない。また、この系では単純に人間に追従するというだけでなく、その背景には $K_h>0$ という前提があり、両者が適応しようという相互適応的な関係が構築されることが、従来までの適応型の演奏機械とは異なる点である。

3.4 振動子モデルと楽譜情報の対応付け

上記の位相振動子モデルから実際に演奏される楽譜情報への対応付けについては、以下のように定めた。まず、固有振動数 ω に対応する基本のテンポ、BPM(Beat Per Minute)は、そのまま振動子の固有振動数と対応させた。つまり、通常の演奏では1拍となる演奏量である4分音符を、振動子の1周期と対応させるのである。4分の4拍子であれば1小節で振動子は4周するのである。4分の4拍子であれば1小節で振動子は4周する。たとえばBPMが120に設定された曲の場合であれば、固有振動数の回転数も120rpmとなるので ω の値は2Hzとなる。

さらに、位相 θ と楽譜との対応関係はFig.6のように定めた。今回のシステムでは振動子の位相が 0° となる位置を拍の最初および小節の最初の位置と対応させた。たとえばFig.6のような楽譜を演奏する場合、1拍目のD音を1周目の位相 0° の位置で演奏し、その拍のつぎのA音を1周目の位相 180° の位置で、2周目は2拍目のF音を 0° で、G音を 270° の位置で、3周目は3拍目のE音を 0° で、D音を 180° の位置で、4周目には 0° の位置では演奏せず 180° で4拍目のA音を演奏する形になる。4拍目を過ぎるとつぎの小節に入り再び第1拍目がその拍における振動子の1周目と対応する。以上のような対応関係を定めることによって、位相振動子における位相の時間発展と楽譜の演奏内容の進行を対応させて演奏することが可能となる。

3.5 楽譜情報から位相差の推定

前節のように位相と楽譜を対応付けることによって、人間

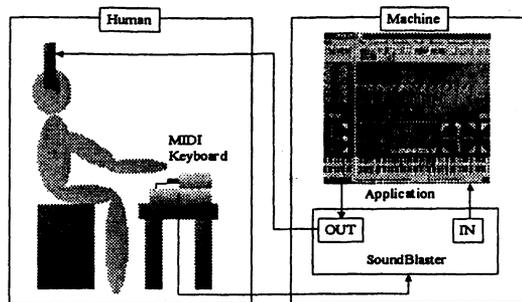


Fig.8 MIDI data flow

の演奏情報から位相の推定が可能となる。機械の位相と人間の位相の関係と、人間の位相の推定方法は Fig.7 になる。この図において、横軸は機械の位相、縦軸は人間の位相である。すでに楽譜と位相の関係で定義されたように、位相の値にしたがってそれぞれの演奏イベント (DAFGED と推移) が決定されている。仮に機械と人間が完全に同期したとすると、機械と人間の位相差は一定値で推移するので、人間の位相と機械の位相の関係は図中の 45° の傾きをもった直線となる。しかしながら、実際は人間と機械の演奏にはずれがあるため、図中の破線のように 45° の中心線の付近でゆらいている。この破線は、人間の演奏イベント (DAFGED と推移) の位相と、その人間の演奏イベントが行なわれた瞬間の機械の位相の交点を通るような曲線である。つまり、図中のプロットを通る破線のようなになるはずである。それぞれのプロットにおいては $f(1) \sim f(7)$ のように位相差が決定される。しかしながら、これ以外の場所においては推定が必要である。そこで、機械と人間が同期しているという前提のもとで、プロット以外の場所では、位相差が一定になるように人間の位相が推移すると仮定した。つまり、図中の太い 45° の傾きをもった線分によって、人間の位相を推定することにした。このような方法をとった場合、計算される位相差がイベントごとに不連続に変化してしまう。この問題は人間の位相が離散的にしか計測できないため不可避であるが、この推定方法はシンプルであるため機械の挙動が安定しやすいという利点がある。

4. システムの実装

4.1 アルゴリズム

実際に人間の演奏と機械の演奏を同期させるためにわれわれは以下のようなアルゴリズムを考案した。まず振動子に対応させて音楽を演奏させる方法として以下の手順を繰り返す。

- 1: 実時間と同期した時計を使用し、設定された時間刻みごとに 2 に進む
 - 2: (4)式にしたがって設定された時間刻みだけ位相を進める
 - 3: 進められた位相の値から楽譜上の演奏位置を計算
 - 4: 楽譜上の演奏位置において、まだ行なわれていない演奏イベントが存在すればそれを実行し、1 に戻る
- ここまでは位相を基準に行なう一般的な演奏アルゴリズム

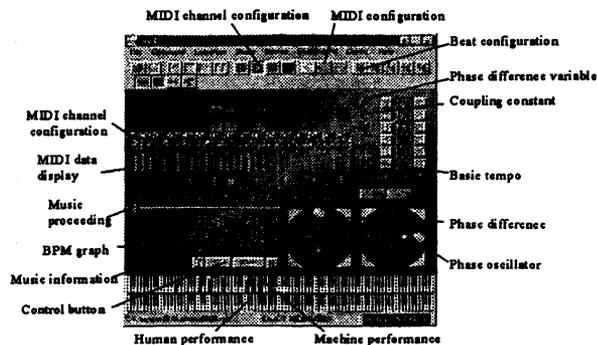


Fig.9 Application

である。これに加えて、設定された時間刻みごとに人間の演奏イベントを計測し、それに対応して位相差を計算する方法として以下の手順を繰り返す。

- 1: 人間の演奏イベントを検知する
- 2: 演奏イベントを解析し音符に変換
- 3: 設定された探索窓内で、演奏された音符と同じ音符を人間が演奏する楽譜から探索。
- 4: 複数対応の場合は位相差の絶対値が最小のものを選ぶ
- 5: 位相差を計算し(4)式に用い、1に戻る

この位相差計算アルゴリズムでは人間の演奏位置を推定するために、人間が演奏するはずの楽譜情報を使用する。機械側と人間側の演奏が引き込んでいる状態であれば、この位相差計算アルゴリズムによって機械の位相を基準として離散的ではあるが人間の位相を決定することができ、(4)式に代入することによって演奏テンポを変化させることが可能となる。ただし、演奏に大きなずれがともなった場合や、同じ音程の音符が密集してある場合では演奏者が意図した音符とは違った音符に対応付けてしまう可能性がある。よって本研究での実験では、そのような対応付けの間違いが起りにくい楽譜を使用する。これらの計算はすべて 10ms 単位で行ない、実時間を基準として演奏位置を決定し累積誤差がでないようにした。また、探索窓は振動子の一周を一拍としたため前後半拍とした。

4.2 入出力インタフェース

演奏をあつかうインタフェースの作成には MIDI 規格を用いることとした²⁴⁾。Fig.8 に示すのが人間と機械の間におけるデータの流れである。MIDI 信号の出力はヘッドホンやスピーカー等の一般的な発音機器によって瞬時に人間の演奏者に提示される。一方、人間の演奏者は MIDI キーボードを用いて演奏を行ない、その内容はサウンドブラスターを通して MIDI 信号として機械に送信される。われわれは、演奏アルゴリズムを用いて演奏を行なうため、Fig.9 に示すような Windows 上で動作するアプリケーションを作成した。このアプリケーションは、SMF フォーマットの MIDI ファイルを読み込み、制御することが可能となっている^{25),26)}。アプリケーションはインタフェースからの信号を音程と入力されたタイミングについて解析し、4.1 節で説明したような位相差計算アルゴリズムに利用する。そして位相差計算アルゴリズムの結果は演奏アルゴリズ

ムに反映され機械側の演奏に変化をおよぼすこととなる。このアプリケーションはGUIを考慮して作成されており、振動子のパラメータやMIDI機器のセッティングはすべてこのアプリケーションを通じて行なうことができる。

5. パラメータ設定

5.1 実験設定

システムの挙動はモデルの結合係数 K に強く依存する。よって最も人間同士の演奏に近い演奏が行なえるように K を設定しなければならない。そこで種々の結合係数のもとで人間と機械が演奏を行ない、人間同士の演奏と比較することとした。

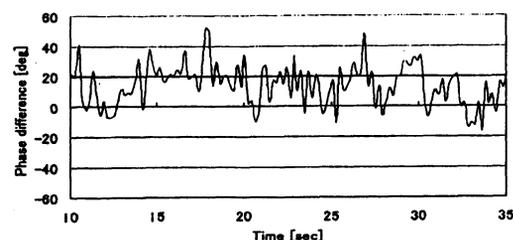
第2章で用いた被験者と同レベルの技術を有する5人の演奏者がこのシステムと演奏を行なった。結合係数 K を基本BPMの0%~50%までの5%きざみで変化させ、それぞれを試行ごとにランダムに並べて演奏実験を行なった。また人間同士の演奏として、5人の被験者すべての組み合わせについて計測し、結合係数を設定するためのリファレンスとした。

基本的な実験環境は2章で述べたのと同じであり、20代5人の演奏者がこの実験を行なった。使用する曲については、リズムの相互作用を扱うため2章と同じものを使用した。アプリケーションはPC(Pentium3 1GHz WindowsXP)上で実行され音源(Roland SC-8850)を通してヘッドフォンから発音される。人間側は電子ピアノ(Roland RD-600)を用いて演奏を行なった。PCは演奏者の後ろ側に配置し、演奏者からは見えないようにした。なお、演奏の発音はすべてSC-8850のみから行ない、人間の演奏の計測についてはすべてMIDI情報によって行なった。音色にはプログラムナンバー0のピアノの音色を用い、この音色のSC-8850の演奏の遅延は3ms以下であり10ms単位の計測精度には十分であると考えられる²⁷⁾。

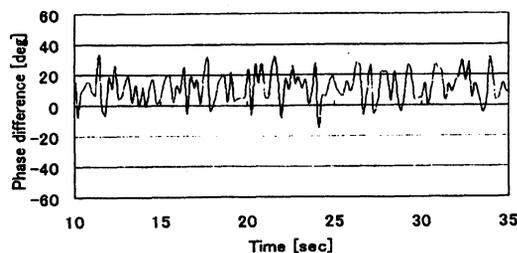
5.2 実験結果とパラメータの決定

Fig.10に演奏者間の位相差の時間発展を、人間同士のものとわれわれのシステム(結合係数20%のもの)について、それぞれ一例を示す。人間同士の位相差については位相差計算アルゴリズムが使えないため、打鍵のずれを各演奏者の演奏位置での一拍の周期で割った値を用い、人間と機械の位相差については位相差計算アルゴリズムによって計算された値を用いた。これらの時間発展はある一定の値を中心としたゆらぎとして観察されることから、われわれはこのゆらぎの中心である一次統計量としての平均値と、ゆらぎの大きさである二次統計量としての標準偏差を計算し比較することとした。位相差は正負の値ができるが、平均値を算出する時はその値を絶対値化せずに計算した。位相差の平均値は曲全体を通しての演奏者同士の演奏のずれを表わしている。また位相差の標準偏差は、この平均から位相差がどの程度ゆらいでいたかを示している。

すべての演奏者について結合係数に対応させてまとめたものがFig.11, Fig.12である。それぞれFig.11に平均値, Fig.12に標準偏差を示す。位相差の平均値は試行ごとに正負の値ができるが、試行全体の値を算出するときは試行ごとの値を絶対値化



a) Human-Human



b) Our new ensemble system

Fig.10 Temporal development of phase difference

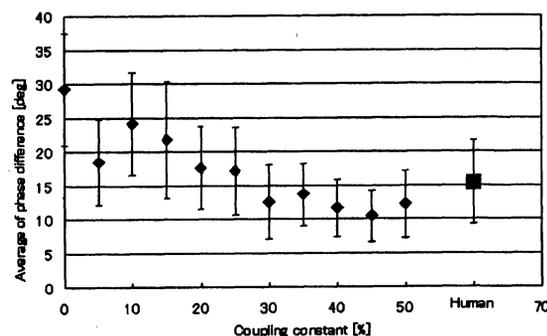


Fig.11 Average of phase difference

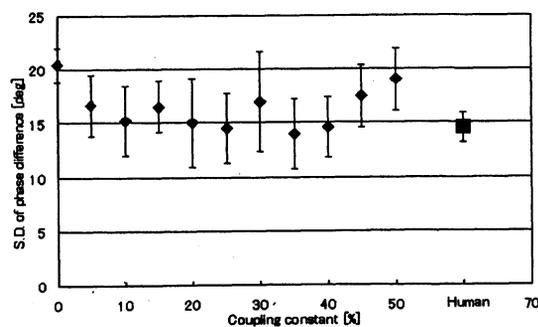


Fig.12 S.D. of phase difference

した。平均値については、結合係数が増加すると共に値が減少していくことが観察される。人間同士の値と比較すると結合係数20%~30%で最も接近することがわかる。これに対して標準偏差については、結合係数が増加するといったん減少した増加するというU字型の傾向が観察される。人間同士の値と比較すると、20%~40%で最も接近することがわかる。これらの結果から一般的な人間の演奏者に対して、比較的アンサンブルがうまくできる結合係数は20%~30%の範囲と考えられ、ここでは K 値として20%を採用した。

6. 評価実験

6.1 実験設定

本章では前章で得られた結合係数 K を用い、われわれのシステムの評価を行なう。この実験では3種類の状況を、客観的な統計指標に加え、演奏者の主観でも比較する。それは人間同士の共同演奏、われわれのシステムと人間との共同演奏、そして機械演奏に一方的に人間があわせるものである。これらと比較する方法として、一種のチューリングテストを行ない、人間らしい演奏の評価とした¹⁹⁾。チューリングテストは、人間か機械かわからない相手に質問を行ない、その相互作用を通じて相手がどちらであるかを比較に基づいて評価する、一種のブラックボックステストである。同様にこの実験でも、行なわれた演奏による相互作用を通して、人間らしさを比較に基づいて評価することを行なった。

二人の被験者を視覚的にも聴覚的にも遮断される条件に置き、Fig.13のような3種類の実験状況を設定した。一つはお互いの演奏が直接相手に送られる状況である(状況A)。もう一つはそれぞれの演奏者が個別にわれわれのシステムと演奏を行なうものである(状況B)。最後にそれぞれの演奏者が相手の単独演奏を録音したものと個別に演奏するものである(状況C)。たとえば、演奏者1と演奏者2とが各状況で演奏を行なった場合、状況Aでは演奏者1の演奏はそのまま演奏者2に送られ、演奏者2の演奏も演奏者1に送られる。状況Bではそれぞれの演奏

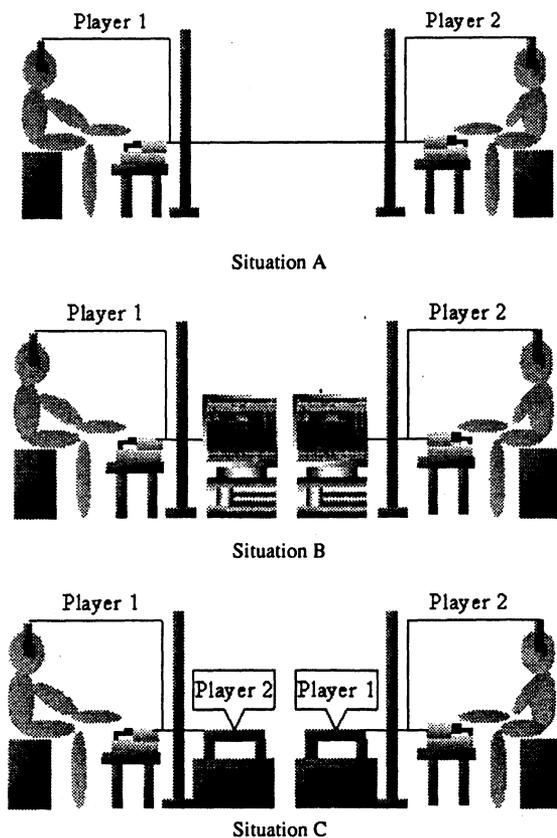


Fig.13 Experimental situations

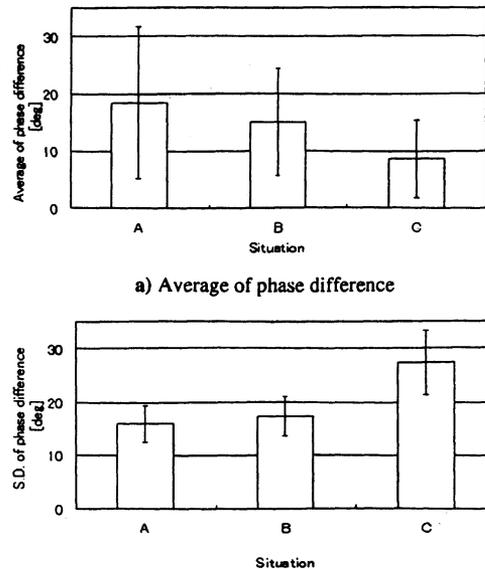


Fig.14 Phase difference of each situation

者が個別にわれわれのシステムと演奏を行なうため、演奏者同士に相互作用はない。状況Cではあらかじめ演奏者1と演奏者2の単独演奏を記録しておき、演奏者1は記録された演奏者2の演奏と演奏を行ない、演奏者2は演奏者1の記録された演奏と演奏を行なう。

被験者は演奏者の1人となり、上記の3種類の状況に対応する3つの試行から構成される実験セッションを行なう。実験各セッションでの状況の提示順番はランダムとし、1セッション終わるごとに演奏中についての評価を行なってもらった。演奏者の主観的評価指標としては、以下のような質問を用いた²⁰⁾。

まず演奏がどれだけそろっていたか、またどれだけ変化したかを調べるために、それぞれに対応するものとして以下の二つの質問を用意した。

a): 相手と同期して演奏できたか?

これは、相手と演奏をそろえることができたかという、同期の容易さを比較することを期待した。

b): 即興性、創出性はあったか?

これは、単独演奏と比較してどれほど自分の演奏が変化したかという、非定常性の有無を比較することを期待した。

また、これらの要素を包括する意味で、演奏の出来具合や人間らしさといった総合的内容の以下の二つの質問を用意した。

c): よい演奏ができたと思うか?

これは、自分がどれほどよく演奏することができたかどうかを比較することを期待した。

d): 相手の演奏が人間らしかったか?

これは、自分のパートナーが機械か人間か判断することを期待し、これによって人間らしさの評価とすることとした。

これらの指標にしたがって各セッションごとに最もよかったもの、悪かったもの、中間として3種類の順位付けを行なわせた。差があまり感じられず順位付けが難しい試行については中間の評価をさせ、差がつけやすかった試行のみよいか悪いか

で評価させた。よって、3試行すべてについて差がつけにくい場合などでは中間のみの評価となる。このような比較に基づいた方法を用いることによって、項目d)のような通常では困難な評価についても対応することが可能と考えた。

基本的な実験環境は2章で述べたものと同じであり、20代4人の演奏者がこの実験を行なった。使用する楽曲については、リズムの相互作用の影響を調べるため2章と同じものを使用し、より一般的な楽曲については今後の課題とした。演奏は、すべての組み合わせ12のパターンについて行なわれた。なお、この実験では音楽の情報として演奏のタイミングだけを用いた。つまり、鍵盤を叩く強さや、押した長さなどは常に一定のものに変換して両演奏者に送信し、ミスタッチで楽譜に記載された以外の鍵盤が叩かれた場合には送信されないようにした。

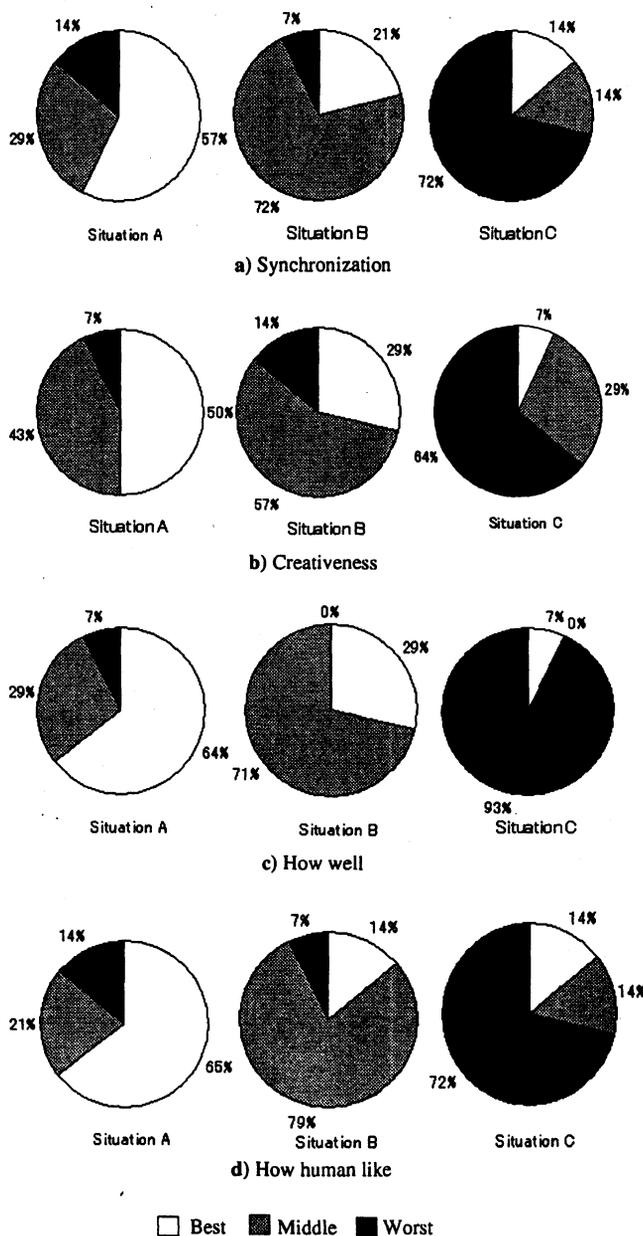


Fig.15 Ratio of evaluation

6.2 結果

Fig. 14はそれぞれの状況における位相差の平均値と標準偏差をまとめたものである。位相差の平均値は、実験セッションごとに計算された平均値の、絶対値によって算出し比較している。グラフ中のA, B, Cはそれぞれ状況A, 状況B, 状況Cに対応している。位相差の平均値についてはこれらの状況の間に有意な差は見られなかった($0.05 < p < 0.1$, ANOVA)。しかし、位相差の標準偏差については、これらの状況の間に有意な差が認められた($p < 0.05$, ANOVA)。各状況間では状況Cとほかの状況間について有意な差が見られ($p < 0.05$, LSD), 状況Aと状況Bの間には有意な差は見られなかった($p > 0.5$)。これらのことは、位相差のゆらぎについて、中心値の差は状況間で大きくはないものの、変動幅は状況Cのみ大きいことを示している。位相差の変動が大きいということは、演奏の同期がとれていないことを意味している。

一方、Fig. 15はアンケートの各項目に対する演奏者から得られた解答を各項目ごとにまとめたものであり、状況ごとに受けた評価の割合が示されている。それぞれの評価項目において、中間の割合が多い理由は、差がつけにくかった試行については中間として評価させたからである。まず項目a)では、状況Aが最も同期がとりやすく、逆に状況Cは同期がとりにくいという主観的感覚を演奏者に与える結果となった。また、項目b)の即興性については、状況Aがより即興性があり、状況Cは即興性が少なく、状況Bはその中間に位置した。このように相互作用の有無によってアンサンブルの非正常性に差が生じている。ここで、同期がとりやすかった状況において、演奏の変化も大きいという結果がでており、演奏が変化する中でも両者がそろおうという興味深い結果となっている。また、項目c)の演奏のよさについても状況Aはよい評価を受け、状況Cは悪い評価を受け、項目d)の人間らしさという点においても、状況Aが人間らしいという評価を受け、状況Cはよい評価を受けなかった。

いずれの項目においても、状況Aと状況Cの間に有意な差が見られ($p < 0.05$, Friedman), 状況Aと状況Bの間では有意な差がなかった($p > 0.5$)。また、状況Bと状況Cの間では項目c)のみにおいて有意な差が見られたものの($p < 0.05$)、ほかにおいては有意な差は確認できなかった($0.1 < p < 0.2$)。

7. 考察

本研究では人間らしいアンサンブルを実現するには、同時的かつ双方向な相互作用が必要であるという作業仮説のもとに、位相振動子を用いた相互引き込みモデルによってアンサンブルシステムを再構成した。再構成したシステムの有効性を評価するために、客観的な指標としては位相差の統計量、演奏者の主観的な評価としてはチューリングテストを用いた^{19),20)}。その結果、客観的指標については、位相差の統計量が状況Bが状況Cよりも状況Aに近く、われわれのシステムが同期のとれた演奏が行なわれていたことが明らかになった。また主観的

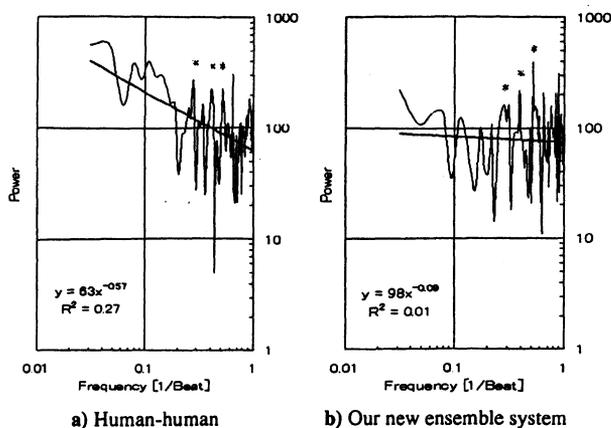


Fig.16 Power spectrum of phase difference

指標についても、状況Bが状況Cよりも状況Aに近くわれわれのシステムがより人間同士の演奏に近いという感覚を、演奏者に与えていたことがわかった。特に、いかに人間らしいかという、評価が複雑で困難な項目についても差がでることが確認された。

これらの主観的評価項目のうち、項目a)の同期については、演奏者の主観と客観的な同期度合いと対応がつく。Fig.14に示したように、各実験状況の位相差の標準偏差を比較すると、状況Bが状況Cよりも状況Aに近く、これの同期度合いの傾向は項目a)の傾向と同じである。そして、この結果は演奏者同士の相互作用の有無が、主観的にも客観的にも強く影響していることを示している。

これに対して項目b)の創出性については項目a)と対照的である。特に客観的なゆらぎの大きさを考えた場合、位相差の標準偏差が最も大きい状況Cがゆらぎも大きいと考えられる。しかし、主観的結果からは単純にゆらぎが大きいことが、創出性には必ずしも結びつかないということが示唆されている。つまり、相互作用によってゆらぎが生じていることが演奏者の感じる創出性に重要であると考えられる。このように対照的な音楽的要素の両方について相互作用の重要性が示されており、また、これらを包括する内容である音楽的な演奏のよさについても、項目c)の評価結果のように、相互作用の有無が演奏者にとっての音楽性に大きな影響を与えていることが示されている。

このように相互作用が演奏のよさに大きな影響を与えていることがわかったが、これらのことは、最終的に項目d)の結果のように人間らしさについても関係していると考えられる。これまで人間らしい演奏を実現するために、データベース的に人間の演奏パターンを利用することが試みられる傾向にあった。しかし、今回の結果は人間機械間の相互作用を構成する方法が、演奏者が感じるアンサンブルの人間らしさに強く関わっており、適切な相互作用系の構築が必要であることを示している²⁰⁾。また、今回のような評価方法を用いることによってその妥当性を調べることが可能となるであろう。今後はより一般的なアンサンブルについても、相互作用の構成方法や評価方法を応用したいと考えている。

ここまでで述べてきたように、アンサンブルシステムの構築において人間機械間の相互作用が必要不可欠であること、そしてその方法として相互引き込みが有効であることが示された^{21,30)}。しかし、人間同士の演奏と比較した場合、われわれの提案する方法との間にはまだ差が観察される。特に、客観的指標とでの状況Aと状況Bの差と比較して、主観的な状況Aと状況Bの差は大きい。この理由として、今回のモデルでは相互作用の第一次近似として相互引き込みのみを利用したことが考えられ、今後は相互作用モデルを改良する必要がある。

また今回、客観的指標として用いたものは位相差の平均値と標準偏差という一次と二次の統計量であった。しかし位相差の時間発展は相互作用と強い関係性があることから、このような静的な指標ばかりでなく動的な指標も用いて解析することで、より詳細なメカニズムを調べることが可能になるであろう。その一つの候補が、Fig.16に示す位相差のゆらぎのスペクトルである³¹⁾。このグラフは人間同士の演奏とわれわれのシステムを用いた演奏、それぞれ一試行についての位相差ゆらぎの時間発展を、拍単位で離散フーリエ解析したものである。人間同士の演奏では位相差ゆらぎのスペクトルは、高周波成分になるほど減衰するフラクタル性を示し、これは人間の演奏に特徴的といわれている³²⁾。しかしながら、われわれのシステムでは、このような特性は観察されず、ホワイトノイズに近づいている。これらは統計的指標では現れなかった部分であり、これが主観的な違いに対応している可能性がある。つまり、これらは時系列的な差に起因していると考えられ、確率過程的な観点からは、人間同士のものは弱定常過程に近く、われわれのシステムは強定常過程に近いことが考えられる³³⁾。特にわれわれの実装したモデルは直前の位相差のみに従うため、一種のマルコフ過程となり定常性が強くなっている可能性がある。今後はこのような確率過程的な観点からも、システムを比較評価する必要がある。

特にわれわれのシステムでは、Fig.16中に*印で示したような、周波数0.25,0.4,0.5付近のピークが観察され、特に高周波成分のピークが大きい。一方、人間同士のスペクトルでも同様のピークが観察され、このようなそれぞれ独立したピークと比較すると、われわれのシステムは人間同士の相互作用のダイナミクスを再構成できていると考えられる。これは周期に変換すると2, 2.5, 4ビートに相当し、拍単位の短い音楽的な構造に依存したゆらぎが独立に現れていることを示している。一方、全体的なスペクトルの勾配で両者を比較した場合、人間同士とわれわれのシステムでは差が観察される。人間同士のスペクトルは周波数相関が高くフラクタル的な広い帯域での相互作用の存在の可能性が示唆されているが、われわれのシステムのスペクトルは周波数相関が低く、楽譜構造に依存した相互作用しか再現できないことが示されている。

このように、人間同士のコミュニケーションでは、短い時間スケールの相互作用と長い時間スケールの相互作用の両方が重要であることがわかる。従来までの適応型伴奏システムでは人間に伴奏を正確にあわせるために、応答速度を向上させるよう

な、いわば短時間相関のみに着目していた。本研究でも結果的には、人間のような長時間相関の再構成は不十分であったが、われわれの研究グループはこの部分の重要性についても着目し、そのメカニズムについて明らかにしようとしている^{34),35)}。特に、二つの時間スケールの関係性が人間の創出的活動を行なう上で重要であることを明らかにしつつあり、このような二重性を考慮したコミュニケーション支援システムの実現が期待されている^{22),36)}。

8. まとめ

本研究では、人間同士の演奏を人間機械系で再構成するため、同時的かつ双方向な相互作用の必要性を示し、相互引き込みモデルを用いた。構築したシステムが、どれだけ人間らしいかを評価するためにチューリングテストを用い、客観的指標との対応関係も調べた。その結果、われわれの提案するシステムは従来のものよりも有意に人間同士の演奏に近いことが示され、アンサンブルにおける相互作用の重要性、相互引き込みモデルの有効性が示された。しかし、人間同士の演奏と比較すると、われわれの提案するモデルにはまだ多くの改善の余地が残されている。今後は、位相差ゆらぎの周波数解析などの動的な指標を導入することや、短周期の楽譜情報に限定されない広い帯域のゆらぎについても再構成することなどが必要になるであろう。

参考文献

- 1) 矢口木三一:マッキントッシュミュージック入門, 西東社(1993)
- 2) 井上, 橋本, 大照:適応型歌声自動伴奏システム, 情報処理学会論文集, 37-1, 51/56 (2002)
- 3) 長島, 橋本, 平賀, 平田:コンピューターと音楽の世界, 共立出版(1998)
- 4) R.B.Dannenberg:An On-Line Algorithm for Real-Time Accompaniment,Proc. of ICMC,275/283 (1984)
- 5) J.J.Bloch and R.B.Dannenberg:Real-Time Computer Accompaniment of Keyboard Performances,Proc. of ICMC,279/289 (1985)
- 6) B.Veroce:The synthetic performer in the context of live performance,Proc. of ICMC,199/200 (1984)
- 7) P.Desain,H.Honing and H.Heijink:Robust Score-Performance Matching:Taking Advantage of Structural Information,Proc. of ICMC,337/340 (1997)
- 8) L.Grubb and R.B.Dannenberg:A Stochastic Method of Tracking a Vocal Performer,Proc. of ICMC,301/308 (1997)
- 9) R.Rowe:Interactive Music Systems Machine Listening and Composing,The MIT Press (1993)
- 10) 和気, 加藤, 才郎, 井口:テンションパラメータを用いた協調型演奏システム -JASPER-, 情報処理学会論文集, 35-7,106/107 (1995)
- 11) M.Nishijima and K.Watanabe:Interactive music composer based on neural networks,Proc. of ICMC,53/56 (1992)
- 12) 井川, 直井, 大, 橋本:相互作用モデルによる実時間適応自動伴奏とその動作解析, 電子通信情報学会春季全国大会講演論文集, 7, 216/216 (1990)
- 13) Y.Kobayashi and Y.Miyake:New Ensemble System Based on Mutual Entrainment,IEEE 9th International Workshop on Robot and Human Interactive Communication (ROMAN2000), Osaka,Japan, 288/291 (2000)
- 14) Y.Kobayashi and Y.Miyake:New Ensemble System based on Mutual Entrainment, Proc. of 2002 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2002), Hammamet, Tunisia, 1045/1048 (2002)
- 15) A.M.Turing:計算する機械と知性について, Mind, 59, 433/460 (1950)
- 16) 堀内, 坂本, 市川:合奏における人間の発音時刻制御モデルの推定, 情報処理学会論文集, 43-2, 260/267 (2002)
- 17) 蔵本, 川崎, 山田, 甲斐, 篠本:パターン形成, 朝倉書房(1991)
- 18) W.S.Condon and L.W.Sander:Neonate Movement is Synchronized with Adult Speech,Science, 183, 99/101 (1974)
- 19) 渡辺, 大久保, 小川, 檀原:InterRobot/InterActer 音声に基づく身体引き込みコミュニケーションシステム, ヒューマンインタフェースシンポジウム 2001 論文集, 397/400 (2001)
- 20) 長岡, 小森, 中村:対話における交替潜時の2者間相互影響, 人間工学, 38-6, 316/323 (2002)
- 21) H.Ogawa and T.Watanabe:InterRobot Speech-driven Embodied Interaction Robot, Advanced Robotics, 15-3, 371/377 (2001)
- 22) 武藤, 三宅:歩行介助を目的とする人間-ロボット協調系における共創出過程の解析, 計測自動制御学会論文集, 38-3, 316/323 (2002)
- 23) 山本, 三宅:音楽の生演奏時における演奏者と聴取者の相互作用の解析, 計測自動制御学会論文集, 38-9, 800/805 (2002)
- 24) 大蔵康義:音と音楽の基礎知識, 国書刊行会(1999)
- 25) R.J.Simon, T.Davis, J.Eaton and R.M.Goertz:Windows95APIバイブル3, 翔泳社(1998)
- 26) 新井純:SMFリファレンスブック, リットーミュージック(1996)
- 27) 長嶋洋一:MIDI音源の発音遅延と音源アルゴリズムの検討, 情報処理学会研究報告, 99-68, 99-SIG-31 (1999)
- 28) 堀内, 田中:自主性を持つ伴奏システム, 人工知能学会誌, 10-1, 72/79 (1993)
- 29) 瀧陽, 橋本:一定の主従関係を持たないインタラクティブな自動伴奏システム, 情報処理学会第60回全国大会講演論文集, 243/244 (2000)
- 30) 東, 橋本:音声認識とピッチ検出を併用した歌声の自動伴奏, 音楽情報科学研究会例会, MUS-22-1 (1997)
- 31) 井上, 宮本, 大橋, 上田, 斎藤 共著:ゆらぎの科学1, 森北出版(1991)
- 32) 土井, 秋山, 松田:楽器演奏時の1/f ゆらぎに関する考察, 94年音楽音響研究会予稿(1994)
- 33) 中川, 真壁:確率過程, 培風館, 75/113 (2002)
- 34) 三宅, 宮川, 田村:共創出コミュニケーションとしての人間-機械系, 計測自動制御学会論文集, 37-11, 1087/1096 (2001)
- 35) 清水, 久米, 三輪, 三宅:場と共創, NTT出版, 339/397 (2000)
- 36) 高梨, 三宅:共創型介助ロボット "Walk-Mate" の歩行障害への適用, 計測自動制御学会論文集, 39-1, 74/81 (2003)

「著者紹介」

小林 洋平



2001年, 東京工業大学大学院総合理工学研究科知能システム科学専攻博士前期課程修了。現在, 同大学院博士後期課程在学中。ヒューマンインタフェース, ヒューマンコミュニケーションの研究に従事。

三宅 美博 (正会員)



1989年東京大学大学院薬学系研究科博士課程修了(薬学博士)。金沢工業大学情報工学科助手, 講師, 助教授を経て, 96年より東京工業大学大学院総合理工学研究科知能システム科学専攻助教授。99年よりミュンヘン大学客員教授併任。主として生命的自律性の研究に従事, 生物物理学会, IEEEなどの会員。