

複数の人間と協調する演奏システム

～ 人間とコンピュータとのアンサンブル ～

堀内 靖雄

藤井 敦

田中 穂積

(東京工業大学 工学部)

Abstract

In this paper we introduce a performance system which can play with several live (human) players. Normally, it is impossible for the system to adjust its tempo to all human players, as there are natural differences in timing for each human player. In order to create a natural-sounding ensemble, we introduce "the performance plan", "the independence rate" and "the time of the ensemble." "The performance plan" allows the system to perform with expression as a human player would. "The independence rate" is a weight corresponding to the musical importance of each player, including the system itself. When the system's "independence rate" is high, it can perform independently of the human players. "The time of the ensemble" is the average of the times of each player (including the system) weighted by "the independence rate." The system changes its tempo dynamically in order to match its time to "the time of the ensemble." Thus, the system is able to create a more natural-sounding ensemble with several human players.

1 はじめに

音楽におけるコンピュータの利用には長い歴史がある。ENIAC(1945)の開発の頃から、音楽への応用の試みは始まっており、コンピュータによる最初の自動作曲である「イリアック組曲」は1957年に発表されている。このような音楽とコンピュータとの関係の中で、最も重要な問題の一つは、ライブ音楽演奏におけるコンピュータの利用である。すなわち、コンピュータがライブ演奏家の役割を演じる訳である。しかし従来、コンピュータはその処理速度の問題のため、リアルタイムで用いることが困難であった。そこで、コンピュータによる演奏はテープに(多重)録音され、実際の演奏時にそのテープを流すという形態が取られていた。そのとき、もし人間の演奏家がこのライブ演奏に加わり、コンピュータと一緒に演奏するとしても、コンピュータはあらかじめプログラムされた演奏しか行なえないため、人間の演奏家がコンピュータの演奏に合わせて演奏しなければならなかった。その後、コンピュータの計算速度の飛躍的な向上により、演奏を続けながら、リアルタイムでコンピュータの演奏制御が可能になった。そこで、当然の要求として、人間のライブ演奏家に合わせてコンピュータが演奏するという考えが起こってきた。それがリアルタイム・インタラクティブ・システムとしての伴奏システムである。そこでまず、伴奏システムについて述べる。

2 伴奏システム

伴奏システムとは、コンピュータが一人の人間の独奏者にテンポなどを合わせて伴奏を行なうシステムであり、リアルタイム・インタラクティブ・システムの一つである。このような適応型の伴奏システムは、すでに様々な研究が行なわれている([Dannenberg 89] [Vercoe 85] [Baird 93] [Katayose 93] [Inoue 93] etc.)。これらのシステムは、独奏者の楽譜とコンピュータのパートの楽譜を与えられており、独奏者の演奏を楽譜を参照しながら追跡し、その演奏に合わせて伴奏パートの演奏を出力するものであった。しかし、これらの従来の伴奏システムの研究は独奏者の

テンポにできるだけ合わせることを目標としており、コンピュータによる伴奏が独奏者に及ぼす効果(コンピュータと人間とのインタラクション)についてはほとんど考慮していなかった。しかし自然なアンサンブルを実現するためには、独奏者と伴奏者の相互作用を考慮する必要がある。例えば、独奏者からの情報が少ないときや無いとき(独奏者が長い音符を演奏しているときや独奏者が休みのとき)には、従来の伴奏システムでは自主的な演奏を行なえないため、機械的で不自然な演奏になってしまうことが多かった。このような状況においては人間同士によるアンサンブルと同様、伴奏者(コンピュータ)がアンサンブルの主導権を握り、人間の独奏者をリードし、人間の独奏者が伴奏者(コンピュータ)に合わせるべきであると考えられる。良いアンサンブルを実現するためには、独奏者(人間)と伴奏者(コンピュータ)とが互いに協調して演奏することが必要であるため、従来の研究のように伴奏システムが独奏者に合わせるだけでは不十分である(このとき、双方向のインタラクションがなく、いわば一方通行による伴奏制御である)。そこで我々は伴奏システムの自主性を考慮することにより、自然なアンサンブルを実現できる伴奏システムを開発した[Horiuchi 93]。このシステムは、独奏者(人間)がアンサンブルの主導権を握るような音楽的状况では、従来の研究と同様、伴奏者(コンピュータ)が独奏者(人間)に良く合わせて追従し、逆に、伴奏者(コンピュータ)がアンサンブルの主導権を握るような状況では、伴奏者(コンピュータ)が自主的な演奏を行なうことにより、独奏者(人間)をリードすることができる。このシステムを評価するため、人間の独奏者と3通りの伴奏システム(人間の伴奏者、従来の研究に基づく伴奏システム、我々のシステム)による演奏を36人の被験者に聴き比べてもらう聴取実験を行なったところ、我々の伴奏システムは従来のシステムよりも良い評価を受け、また人間の伴奏者にも劣らないということが示された。

3 複数の人間と協調する演奏システム

一人の演奏者に合わせる伴奏システムの研究に対して、複数の人間の演奏者と演奏を行なうシステムの研究 [Baird 91] も行なわれているが、このシステムは、複数の演奏者の中で最も重要であると思われる演奏者一人を決定し、その演奏者のみに合わせていたため（本質的には伴奏システムと同じ処理による）、すべての人間と協調して演奏することはできなかった。もちろん、重要な演奏者に合わせることは必要であるが、他の演奏者をまったく無視してしまうのでは、自然なアンサンブルを行なうことはできない。何故なら、その重要な演奏者自身も、他の演奏者の演奏の影響を受け、自分の演奏を変化させるからである。そのため、良いアンサンブルを実現するためには、そのようなアンサンブルをリードしている演奏者に合わせるの言うまでもなく、それ以外の演奏者の演奏も考慮すべきである。先の Baird らのシステムも、この点で改良され [Baird 93]、マッチングの確からしさや、旋律の重要さなどを考慮に入れ、複数の人間の演奏情報を重み付けして演奏を制御しようとしているが、具体的な考えはほとんど示されていない。また、このシステムは人間の演奏者をリードするようなことができないが、音楽的な状況によっては、コンピュータがメロディを演奏するなど、他の演奏者をリードすべき状況も生じる。このような状況では、コンピュータが自主的で表情豊かな演奏を行ない、他の演奏者をリードすることが望ましい。そこで我々は先の伴奏システム [Horiuchi 93] を拡張し、複数の人間とのアンサンブルにおいて、すべての演奏者の演奏を状況に応じて考慮した自然なアンサンブルを可能し、また、状況によってはコンピュータが人間の演奏者をリードしながら、自主的で表情豊かな演奏を行なうことも可能となる [藤井 93]。

3.1 システムの構成

このシステムは人間と一緒にアンサンブルを行なうシステムなのでアンサンブル・システムと呼ぶ。アンサンブル・システムの構成を図 1 に示す。この図は n 人の人間と一緒に演奏を行なう場合であるが、 n は一人でも良い（そのときの処理は、伴奏システムと同じである）。すべての演奏者（人間とコンピュータ）の楽譜はあらかじめシステムに与えられている。実際に人間の演奏家と演奏を行なうとき、各演奏者の演奏はシステムに入力され、システムはこれらの演奏に合わせてリアルタイムでコンピュータの演奏を出力する。ここで、人間の演奏者の演奏（入力）とシステムの演奏（出力）はともに MIDI 信号である¹。アンサンブル・システムは処理の違いによって大きく二つのモジュール (Listener & Performer) に分けられ、各モジュールに Macintosh コンピュータが割り当てられている (図 1 参照)。以下、二つのモジュールについて順に説明する。

3.2 Listener

各演奏者の演奏 (MIDI) は一つの MIDI 信号にマージされ Listener に入力される。Listener は各演奏者ごとに

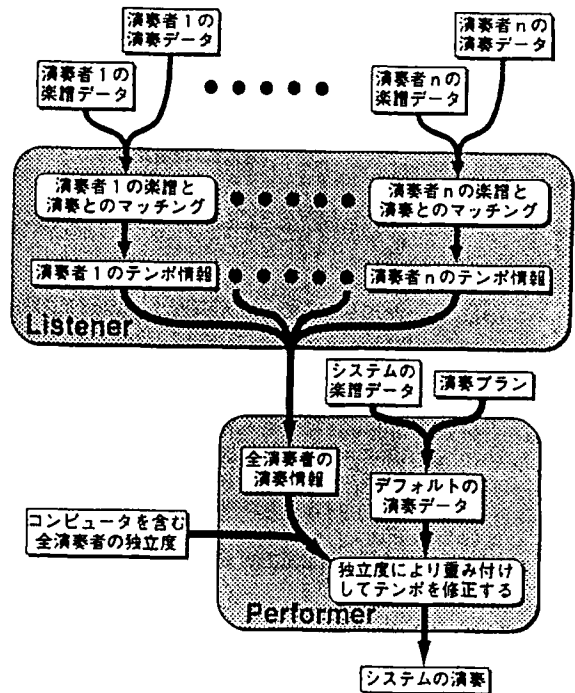


図 1: アンサンブル・システムの構成

独立して、その演奏を追跡し、各演奏者の演奏情報（現在のシステムではテンポ情報）を認識し、Performer に出力する。具体的には、各演奏者の Note On 信号（音符が演奏されたときに出力される MIDI 信号）、その演奏された時刻を用いて、その演奏者の楽譜情報とマッチングする事により、その演奏者が現在、楽譜上のどの位置を演奏しているのかを追跡する。しかし人間の演奏は、楽譜通りに一定のテンポで演奏されることはなく、また演奏ミスをすることもある。そこで、人間の演奏者のテンポの揺らぎに応じて、対応する楽譜の時間軸を伸縮したり、演奏ミスを許容して人間の演奏を追跡するため、音声認識などに用いられる DP マッチングをリアルタイムに適応する [Supawaree 92]。このとき、演奏ミスや大きなテンポ変化を考慮に入れるとマッチングの候補は通常複数個存在する。すなわち、各演奏者の現在の楽譜上における演奏位置が一意に決定できない場合である。その時は可能な候補を複数持ち、その時点で最も確からしいと思われる解釈を現時点におけるマッチング結果とする。その確からしさは各演奏イベントのマッチングに対してスコア付けをして、それらのスコアをもとに決定する。すなわち、楽譜通りの正しいマッチングに対しては高いスコアを与え、演奏ミスや、大きなテンポ変化に対してはその度合に応じてスコアを下げる。スコアがあらかじめ決められている閾値よりも小さくなった時、その解釈は破棄される²。Listener は、各演奏者ごとに、そのテンポ情報として、一拍ごとに楽譜上の位置（拍時間）、演奏時刻（実時間）、そのマッチングの確からしさを表わすスコアの3つを Performer へ出力する。

¹ MIDI とは電子楽器の通信などに用いられる規格であり、シンセサイザやコンピュータ間の通信は通常、この信号により行なわれる。

² 各演奏者の候補数は通常一つである。演奏ミスなどが発生したとき、複数になるが、正しい入力が続くと、再び一つの解釈に収束する。

3.3 Performer

Performer は Listener から送られる各演奏者の演奏情報をもとにコンピュータの演奏テンポを制御し、すべての演奏者と協調した自然な演奏を行なう。そのとき「演奏プラン」「独立度」「アンサンブル時間」の3つの考えを用いる。以下、それらについて詳細に述べる。

3.3.1 演奏プラン

人間は演奏を行なう際に、自分の意志に基づく独自の演奏プランに従って演奏しており [Sloboda 87]、また、アンサンブルの場合には、この独自の演奏プランを動的に修正し、他の人間と協調しながら演奏を行なっていると考えられる。そこでアンサンブル・システムも人間のように表情豊かに演奏しながら、自然なアンサンブルを実現するため、システム内部にあらかじめ演奏プランを与えておく。これはいわば、コンピュータの静的な演奏意志に相当するものである。具体的には、楽譜に書かれていない表情データとして、テンポの揺らぎ、音量変化、ペダル情報などを与える。コンピュータが自主的に演奏すべき状況では、この演奏プランに従うことにより、表情豊かで自主的な演奏を行なうことが可能となる。アンサンブル・システムは、この演奏プランに従ってコンピュータ自身のパートを表情豊かに演奏するが、すべての人間の演奏者と協調するように演奏プランを動的に修正しながら演奏を進める。その時、以下の「独立度」と「アンサンブル時間」の考えに基づいた演奏制御を行なう。

3.3.2 独立度

人間同士のアンサンブル演奏では、お互いに協調しながら演奏を行なっている。そのとき、すべての演奏者が対等に演奏を行なっている訳ではなく、その中には、アンサンブル全体をリードするような演奏者、その演奏者に合わせようとしている演奏者というように、各演奏者ごとにアンサンブルにおける重要性がある。また、この重要性は一曲を通じて常に一定ではなく、状況によって変化する。

このようなアンサンブルにコンピュータが入って、一緒に演奏する場合、コンピュータも各演奏者の重要性を考慮しながら、すべての演奏者に合わせて演奏すべきである。そこで各演奏者（コンピュータを含む）のアンサンブルにおける重要性を表わす尺度として「独立度」を導入する。これは各演奏者ごとに異なる値をとり、楽譜の位置ごとに変化する。独立度は基本的に、アンサンブルをリードする演奏パートほど高くし、他の演奏者に合わせるべきパートほど低くする。また、細かい音符（弦楽器のきざみのような）を演奏しているパートは若干高く、長い音符や休みのパートは低くする。独立度はこのような規則を用いて楽譜から決定し、あらかじめシステムに与えておく。システムはこの独立度に応じてすべての演奏者の演奏を考慮し、適当な演奏を行なう。すなわち、システムは独立度の高い演奏者のテンポには比較的良く合わせるが、他の演奏者の演奏もその独立度に応じて考慮するため、まったく無視することはない。またシステムの独立度が他の演奏者よりも高い状況では、システムは自主的で表情豊かな演奏を行なうことにより、他の演奏者をリードする。このような演奏

制御を実現するため、「アンサンブル時間」の考えを導入し、それを用いた伴奏制御について述べる。

3.3.3 アンサンブル時間

「アンサンブル時間」とは、全演奏者（コンピュータを含む）の演奏時刻を独立度により重み付けした平均（重心）の時刻であり、直感的には、アンサンブル全体の演奏時刻と考えられる。（ここで各演奏者の演奏時刻とは実時間を意味し、秒を単位とする。これに対して拍時刻とは楽譜中の時間を意味し、拍を単位とする。）具体的には、拍時刻 s において、各演奏者の s での演奏時刻を $t_i(s)$: ($i = 1..n$ ここで n は人間の演奏者の数)、システムの s での演奏時刻を $t_{n+1}(s)$ 、 s での各演奏者の独立度を $Indep_i(s)$: ($i = 1..n$)、システムの独立度を $Indep_{n+1}(s)$ とするとき、 s におけるアンサンブル時間 $t_{ensemble}(s)$ を以下のように定義する。

$$t_{ensemble}(s) = \frac{\sum_{i=1}^{n+1} \frac{Indep_i(s)}{\sum_{j=1}^{n+1} Indep_j(s)} t_i(s)}{\sum_{j=1}^{n+1} Indep_j(s)} \quad (1)$$

アンサンブル時間は独立度の高い演奏者の s での演奏時刻に近づくが、他の演奏者も独立度に応じて考慮される。コンピュータの独立度が他の演奏者よりも高いときは、コンピュータ自身の演奏時刻に近づく。次節ではアンサンブル時間の考えを用いてコンピュータの演奏を制御する方法について述べる。

3.3.4 コンピュータの演奏制御

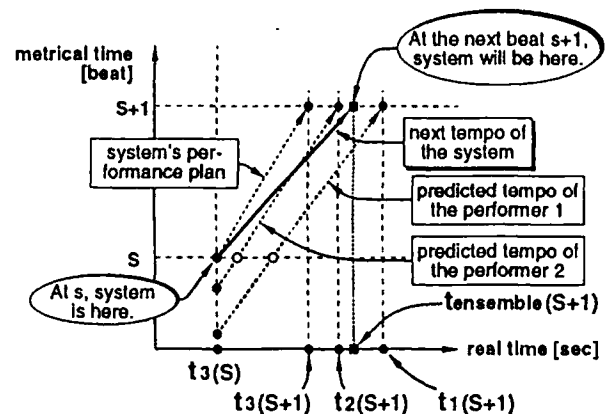


図 2: コンピュータのテンポ制御

今、実時刻 $t_{n+1}(s)$ にコンピュータが拍時刻 s を演奏したところであるとする（図 2 参照。この図は人間の演奏者が 2 人の場合）。コンピュータは演奏を続けるため、次のテンポ、すなわち拍時刻 $s+1$ を演奏すべき実時刻を計算する必要があるが、その時刻として、拍時刻 $s+1$ におけるアンサンブル時間 $t_{ensemble}(s+1)$ を適用する。

$$t_{ensemble}(s+1) = \frac{\sum_{i=1}^{n+1} \frac{Indep_i(s+1)}{\sum_{j=1}^{n+1} Indep_j(s+1)} t_i(s+1)}{\sum_{j=1}^{n+1} Indep_j(s+1)} \quad (2)$$

このアンサンブル時間を計算するために、全演奏者（コンピュータを含む）の $s+1$ での演奏時刻 $t_i(s+1)$: ($i = 1..n+1$) が必要となる。 $s+1$ での人間の演奏者の演奏

時刻 $t_i(s+1)$ ($i = 1..n$) は、過去数拍 (本システムでは 1 小節分) の演奏履歴を重み付けして予測する。コンピュータの $s+1$ での演奏時刻 $t_{n+1}(s+1)$ はコンピュータの s での演奏時刻 $t_{n+1}(s)$ と s での演奏プランのテンポ $V_{plan}(s)$ を用いて以下のように計算される。

$$t_{n+1}(s+1) = t_{n+1}(s) + \frac{1}{V_{plan}(s)} \quad (3)$$

これらから式 (2) に従って計算したアンサンプル時間 $t_{ensemble}(s+1)$ により、コンピュータの次の演奏テンポ $V_{n+1}(s)$ は以下のように計算される。

$$\tilde{V}_{n+1}(s) = \frac{1}{t_{ensemble}(s+1) - t_{n+1}(s)} \quad (4)$$

基本的には上記の考えでコンピュータは演奏を続けるが、いくつか例外がある。もし、式 (4) で $t_{ensemble}(s+1) < t_{n+1}(s)$ となり、計算されたテンポが負になるときは、コンピュータは各演奏者の演奏が追い付いてきて、そのテンポが正になるまで演奏を待つ。

またシステム以外の演奏者の独立度がすべて 0 (最小値) の時は、上記の式に従って演奏を行なうことにより、人間の演奏者の演奏から独立して、演奏プラン通りの自主的な演奏を行なうことができる。

式 (2) の例外的な処理として、拍時刻 s におけるコンピュータの独立度が 0 のとき (コンピュータが人間の演奏者に縦の線を合わせる必要があるとき) には、コンピュータは人間の演奏者の誰かが最初に s での音符を演奏した直後 (おそらく数ミリ秒以内) にその音符を演奏する。そのとき、コンピュータの演奏がその演奏者よりも遅れている場合、システムはその拍時間を s にジャンプし、逆に、コンピュータが他の演奏者よりも進んでいる場合、誰かが追い付くまでその音符を演奏するのを待つ。

3.3.5 演奏ミスへの対応

このようにしてテンポ修正を行なうが、人間の演奏者の演奏ミスなどにより、その演奏者の演奏のトラッキングに失敗している可能性もある。そのとき間違ったマッチングを考慮したテンポ修正を行なってしまうと、その演奏者が意図していないテンポになってしまう可能性がある。この問題点を解決するため、Listener から受け取るマッチングのスコアが低い場合は、それに応じてその演奏者の独立度を低くして、その演奏者に影響されすぎないようにする。

4 おわりに

本論文では複数の人間の演奏者と協調して演奏を行なうインタラクティブ・コンピュータ・システム (アンサンプル・システム) について述べた。アンサンプル・システムは演奏プラン (コンピュータの演奏意志) を持ち、音楽的状况に従って変化するすべての演奏者 (コンピュータを含む) の独立度によって、すべての人間の演奏を考慮しながら、それと協調するように、コンピュータ自身の演奏を制御する (アンサンプル時間の考えに基づく) ことにより、人間との自然なアンサンプルを実現する。

今後の研究課題として、本論文ではテンポ制御についてしか述べなかったが、ダイナミクス (音量) などの演奏

ニュアンスも状況に応じて変化させることが考えられる。その時にも独立度の考えにより、独奏者と伴奏システムの音量のバランスを考慮した制御を行なうことを考えている。また、現在は人間が与えている演奏プランや独立度を自動/半自動で与えたり、あるいは、リハーサルを行なうことにより学習したりすることも考えたい。

References

[Baird 91] Baird, B.: The Artificially Intelligent Computer Performer and Parallel Processing, in *Proceedings of International Computer Music Conference*, pp. 340-343, 1991.

[Baird 93] Baird, B., D. Blevins, and N. Zahler: Artificial Intelligence and Music: Implementing an Interactive Computer Performer, *Computer Music Journal*, Vol. 17, No. 2, pp. 73-79, 1993, (Summer 1993).

[Dannenberg 89] Dannenberg, R. B.: Real-Time Scheduling and Computer Accompaniment, in Mathews, M. V. and J. R. Pierce eds., *Current Directions in Computer Music Research*, System Development Foundation Benchmark Series, chapter 18, pp. 225-262, MIT Press, 1989.

[Horiuchi 93] Horiuchi, Y. and H. Tanaka: A Computer Accompaniment System With Independence, in *Proceedings of International Computer Music Conference*, pp. 418-420, 1993.

[Inoue 93] Inoue, W., S. Hashimoto, and S. Ohteru: A COMPUTER MUSIC SYSTEM FOR HUMAN SINGING, in *Proceedings of International Computer Music Conference*, pp. 150-153, 1993.

[Katayose 93] Katayose, H., T. Kanamori, K. Kamci, Y. Nagashima, K. Sato, S. Inokuchi, and S. Simura: Virtual Performer, in *Proceedings of International Computer Music Conference*, pp. 138-145, 1993.

[Sloboda 87] Sloboda, J. A.: 「演奏」, Deutsch, D. ed., 音楽の心理学, 第 16 章, pp. 587-609, 西村書店, 1987.

[Supawaree 92] Supawaree, K.: リアルタイム DP マッチングを用いた演奏者情報の抽出, 東京工業大学工学部卒業論文, 1992.

[Vercoe 85] Vercoe, B. and M. Puckette: Synthetic Rehearsal: Training the Synthetic Performer, in *Proceedings of International Computer Music Conference*, pp. 275-278, 1985.

[藤井 93] 藤井, 堀内, 田中: 複数の人間と協調して演奏するシステム, 人工知能学会 第 7 回全国大会 論文集, pp. 567-570, 1993.