

K-024

音楽セッションを目的としたコミュニティサイトの提案

Proposing of Constructing a Music Session Site

鈴木 茂樹†
Shigeki Suzuki遠藤 守‡
Mamoru Endo山田 雅之‡
Masashi Yamada宮崎 慎也‡
Shinya Miyazaki

1. まえがき

音楽は古くからコミュニケーションの手段として我々の身近に存在している。総務省の調査[1]では、楽器演奏を趣味として楽しむ人は、日本国内だけで 1500 万人近く存在していることが分かっている。情報技術の進歩やインターネットの普及は、様々な分野に大きな影響を与えてきた。それは音楽という分野においても同様で、音楽の電子化、それに伴う共有、配信技術の向上や配信サービスの利用増加などからも見て取

ることができる。また、自作音楽の投稿サイトなどを利用すれば個人で音楽を作成、公開することは比較的容易である。SNS などネット上で音楽に関するコミュニティも形成され、物理的に遠く離れた相手との交流も盛んに行われるなど、音楽はネット上でコミュニケーションを図るための媒介となっている。一方で、実世界での「演奏」、特に独奏でなく複数のメンバーの合奏の形態においては、基本的に同一の時間・空間の共有が前提であるため、「予定が合わない」「居住地が離れている」「スタジオを借りなければならない」など、地理的、時間的な制約などにより達成が困難な場合が多い。そこで、本研究は、このような複数による音楽セッションをネットワーク上で実現し、新たなコミュニケーション手段として提案する。

1.1 関連研究

近年、ネットワーク上のコミュニケーションツールとして 3DCG アバタを用いた様々なサービスが注目されている。仮想空間内で 3DCG アバタを操作することにより、会話やスポーツなど実世界に近い形でコミュニケーションを行うことができる。

しかしながら、インタラクティブな演奏をコミュニケーションの手段としているものは見受けられない。これはネットワークを利用する場合の伝送遅延は不可避であり、地球規模の広域ネットワークでは音楽演奏に必要な同時性の確保は本質的に不可能だからである。

これまでもネットワークを利用した音楽演奏に関する研究は様々されており、特にネットワーク伝送による遅延をどう扱うかを大きな課題としているものが多い。[2]は、予め幾つかのパートを製作しておき、あるパートの演奏中に次のパートを決定し組み合わせていく、コラージュ的手法を用いて、ネットワーク上で楽曲制作を試みている。リアルタイムで曲の流れをある程度変えていけるものの、曲を作る段階では演奏を済ませており、自由さに欠ける。[3]は同一のコード進行による楽曲に限定し、遅延によって音楽セッションの同時性は無理という前提で演奏を行う、例えば 12 小節のブルース形式の 1 ループのような繰り返し構造を単位として、「相手の 1 単位前の演奏と自分の演奏との合奏」というセッションを相互に行う。またこれに

似たシステムで「GDS music」(Global Delayed Session music)があり、「1 小節遅れ音楽」、あるいはと名付けた新しい音楽のコンセプトを提案している[4]。これはインターネット時代に必須必然の「地球規模の遅れ」を前提とし、それを本質的・積極的なコンセプトとして置いた新しい音楽モデル、というものである。演奏できる音楽が遅れを前提とするものに限られるため、通常とは異なる感覚で演奏を行わなければならない。[5]は、音楽だけでなく演奏するキャラクタを画面に表示する点で本研究と似ているが、演奏相手となるデータを予め用意しておき、人とコンピュータで音楽のセッションを行おうという試みである。

1.2 本研究の目的

本研究では、音楽セッションを単なる音楽演奏の場としてではなく、コミュニケーションツールとして提供することを目的としている。実世界の音楽セッションでは演奏者の動作、表情など、音以外の部分もコミュニケーションを行うための重要な要素である。そこで、仮想スタジオや 3DCG アバタを使い、視覚的なアプローチも含めた音楽セッションシステムの作成を試みる。

音楽演奏に関してはネットワーク利用による伝送遅延が不可避であることは先述の通りだが、既に国内の通信においては遅延が致命的になるようなゲーム類もネットワーク上で商用として数多く利用されていることから、音楽もデータの形式によっては遅延を感じることなく通信ができるのではないかと考える。遅延を最小限に抑えるための通信データモデルを用いることで、できる限りリアルタイムに近い形でのセッションを実現する。

また、楽器を扱えない人向けに PC キーボードやジョイスティックデバイスなどを楽器として利用できるようにするための仕組みや、普段使っている楽器の音を再現するなど音の自由性を提供するためのサンプラーの実装も試みる。

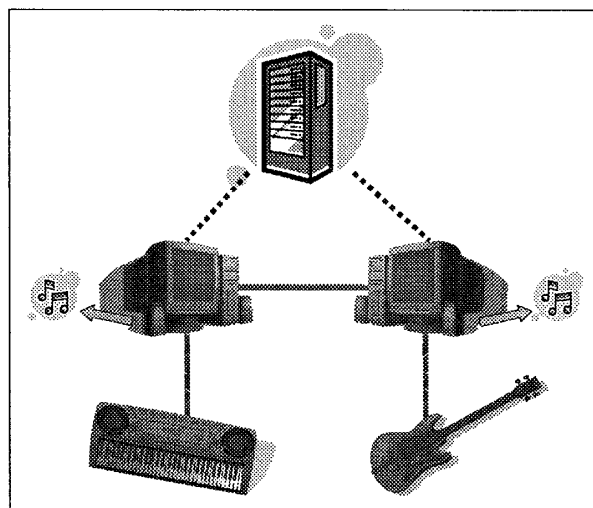


図1 音楽セッションシステムの概観

†中京大学大学院情報科学研究科

‡中京大学情報理工学部

2. システム概要

本システムは、セッションを行う相手を探すためのロビーサーバ、音楽セッションを行うためのクライアントソフトウェア、演奏のための楽器とキャラクタ操作のための入力機器で構成される(図1)。これにより、オンライン上で複数人による演奏を実現する。より多くのユーザが音楽によるコミュニケーションを取れるよう、サーバクライアント型とP2P型を複合し、手軽にセッションへ参加できる仕組みを提供する。音楽セッションを行う場合は、まずクライアントソフトウェアを起動する。ロビーサーバに接続し、相手を探す。相手が見つかったらP2P接続により音楽セッションを開始し、PCに接続した楽器などで演奏を行う。このとき、演奏に合わせてキャラクタを自動的に動かすことや、演奏以外の動きを操作することで演奏相手とのコミュニケーションを図る。

2.1 ロビーサーバ

ロビーサーバの役割は、ユーザ同士をマッチングさせることと、セッションの管理を行うことである。具体的には、接続ユーザの情報(ユーザ名、IPアドレス、ポート番号など)と現在演奏中の音楽セッションの情報(セッション名、開始したユーザの情報、参加者の情報)を保持・管理し、クライアントへ伝える役割を果たす。また、チャット機能を追加し手軽に相手を探せるようにした。

2.2 クライアントソフトウェア

クライアントソフトウェアの機能はネットワーク部、サウンド部、グラフィック部、入力制御部の4つに大別できる(図2)。

ネットワーク部はロビーサーバと通信し、接続者情報や、開始されているセッションの情報を受け取るためのロビークライアントの機能と他のPCとP2P接続により音楽セッションを行うためのピアの機能を有する。入力制御部は、演奏やキャラクタ操作のためのMIDI楽器または入力デバイスの制御を行う。サウンド部は、演奏音の保持や音の再生などサウンドデバイスの制御機能を有する。グラフィック部は、仮想スタジオ、キャラクタの表示やキャラクタ

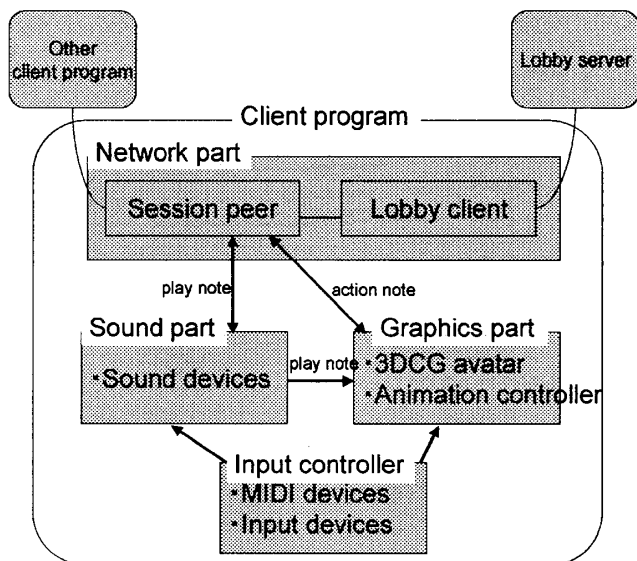


図2 クライアントプログラム構成図

演奏動作の生成などの機能を有する。それぞれの詳細については以降の章で述べる。

クライアントソフトウェアの主な処理内容は次の通りである。

- (1) 自分が演奏を行った場合、演奏情報がサウンド部に伝わり、音が再生される。演奏情報がグラフィック部へ渡り、演奏動作が合成、出力される。演奏情報がネットワーク部へ渡り、他の参加者に送信される。
- (2) 音楽セッションの参加者が演奏を行った場合、ネットワーク部が他の参加者から演奏情報を受け取る。演奏情報はサウンド部に渡り、音が再生される。演奏情報はグラフィック部へ渡り、演奏動作が合成、出力される。
- (3) 自分がキャラクタを操作した場合、動作情報がグラフィック部へ伝わり、キャラクタの動作を出力する。動作情報はネットワーク部へ渡り、他の参加者に送信される。

- (4) 音楽セッションの参加者がキャラクタを操作した場合、ネットワーク部が他の参加者から動作情報を受け取る。動作情報はグラフィック部へ渡り、キャラクタの動作が出力される。

2.3 開発環境

開発プラットフォームにはMicrosoft社のVisual Studio .Netを使用した。3Dグラフィックス、サウンド、ネットワーク、入力デバイス制御用のライブラリにはDirect X SDKを用い、C++言語での実装を行った。また3DCGのモデルはX形式のデータを用いた。

3. ネットワーク部

本システムの通信形態は、サーバクライアント型とP2P型の複合である。多くのユーザが利用する場合の負荷を考え、音楽セッション部分はP2P接続で実現する。また、手軽に相手を探せるようロビーサーバを用意した。ロビーは接続者などの情報の保持・管理を一括して行う必要がある。このことからロビーサーバとの接続はサーバクライアント

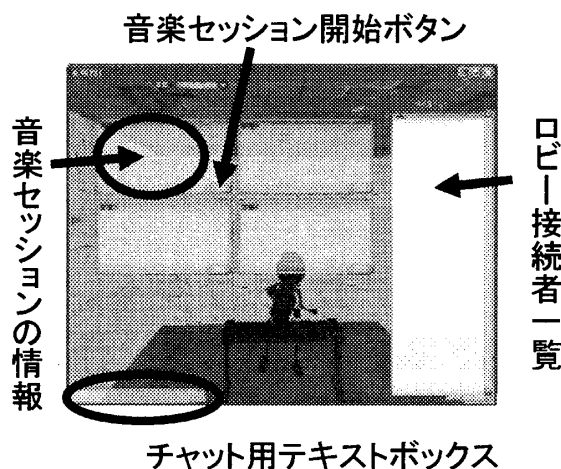


図3 ロビークライアントのインターフェース

ント型とした。

3.1 ロビークライアントの実装

ロビークライアントは、ロビーサーバからの情報を受け取り、画面に表示することで音楽セッションを行うためのインターフェースを提供する(図3)。

本システムで使用したネットワーク用ライブラリ DirectPlay の機能を使い、ロビーサーバとロビークライアントを実装した。

3.2 音楽セッション用ピアプログラム

本システムでは、音楽セッションを演奏だけでなく3DCG アバタの動きを交えて行うため、音データとキャラクター操作データの2つは特に重要な通信データである。これ以外にチャット用テキストデータも扱う。音データの詳細についてはサウンド部で述べる。

音楽セッション用ピアプログラムは、ユーザが演奏やキャラクター操作などの入力を行った場合、サウンド部から音データを、グラフィック部から動作データを受け取りセッション参加者へデータを送信する。演奏動作(例、ドラムスティックを振る、鍵盤を弾くなど)については、音データに同期させたため、送受信は行わない。ロビーサーバ・クライアントと同様に DirectPlay をライブラリとして使用し、実装を行った。

4. サウンド部

現実空間で実現される音楽セッションではユーザが用意できる音色は無数にあり、同じ空間で音楽を共有するため聞き手の環境によって音に変化することはない。

更に、ネットワーク上のリアルタイム音楽セッションでは、遅延を最小限に抑えるための通信データモデルを用いる必要がある。

もちろんコンピュータ同士を接続する仕組みは国内外を問わないため、遅延による非リアルタイム性をいくらか無視すれば、セッションそのものは可能である。

音(波形データ)を通信に利用することを考えた場合、通信を軽量化するためには波形データそのものを送らないことで実現が可能である。その仕組みとしてMIDIを用いる。本研究ではこれが最も軽量で、音源さえ用意できれば音質も悪くない状態で再現が可能であると考えた。

MIDI のように、波形データを含まない情報を利用して

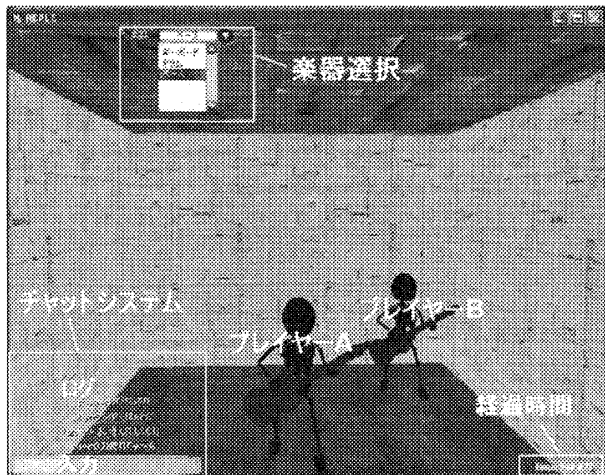


図4 仮想スタジオに配置した3DCGキャラクター

通信を行えば、非常に軽量のデータでオンラインセッションを行うことができる。音源はサンプラー機能を使うことで自由な個性のある音を表現することができる。

サウンド部の具体的な機能は、ユーザが演奏を行った場合、楽器に用意された音を再生し、グラフィック部に演奏データを伝える。更に、ネットワーク部にも演奏データを伝える。ネットワーク部から演奏データを受け取った場合、対応する楽器の音を再生する。

実装はサウンド用ライブラリである DirectSound を使用し行った。

4.1 サンプラーの実装

演奏データにはMIDIを用いたが、MIDIを音源として利用する場合、環境により再生される音が異なることがある。最近のWindows PCには「Microsoft GS Wavetable SW Synth」というソフトウェアMIDI音源が標準でインストールされるため、これを標準の音源としてもよいがソフトウェア音源のため発音まである程度時間がかかる。また、MIDI規格に存在しない楽器を利用することも不可能である。

実際の音楽制作では楽器から出た音をエフェクターを通すなどして変化させる場合がある。音作りも個性を表す重要な要素となるのである。

そこで我々は、ユーザ(または開発者)が自由に音源を用意し、且つ全てのユーザが同じ音で利用できる仕組み、サンプラーの機能を実装した。サンプラーは再生時に音の高さを変更して出力することにより、一つの音源データから、複数の音階を生成する。

本研究ではサウンド用ライブラリとして使用した DirectSound の機能を利用し、再生速度を変化させることでサンプラーを実現している。

5. グラフィック部

これまでのネットワークを利用した音楽セッションの研究では、単純な音だけのやりとりであった。実世界では音だけでなく、演奏者の動作、表情など、音以外の要素も演奏に影響する。そこで、本研究では音だけでなく、視覚的なアプローチによるコミュニケーションも試みる。仮想空間に音楽スタジオを設置し、さらに演奏中のユーザを表す3DCGのキャラクターを表示させる(図4)。キャラクターには演奏データと同期するようなアニメーションを追加し、実際にそのキャラクターが演奏しているように表現する。また、演奏以外の動作も追加し制御することで、ユーザがお互いに画面の向こう側に相手がいることをイメージでき、視覚的にもコミュニケーションが取れるのではないかと考えた。

具体的な機能は、入力制御部からキャラクター操作データを受け取った場合はキャラクター動作のアニメーションを出力し、さらにネットワーク部へキャラクター操作データを伝える。ネットワーク部からキャラクター操作データを受け取った場合はキャラクター動作のアニメーションを出力する。サウンド部から演奏データを受け取った場合は演奏アニメーションを生成して出力する。

5.1 演奏にあわせたアニメーションの生成

本研究で3Dグラフィックスライブラリとして使用したDirectXには複数のアニメーションを合成して出力するAPIが用意されている。本システムではこの機能を利用して3DCGアバタが演奏に合わせて演奏するアニメーション

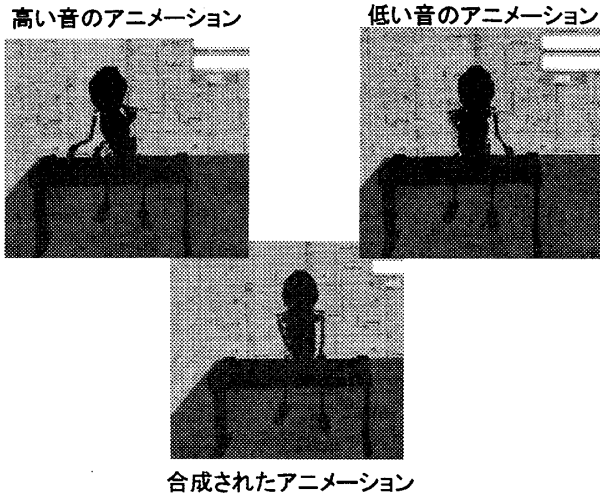


図5 アニメーション合成の様子

を生成している。具体的には、一番低い音と一番高い音を弾いている2つのアニメーションを用意しておき、実際に弾いた音の高さにより合成の割合を決定し出力する(図5)。楽器の演奏部分が直線状に配置されているのであればこの方法を適用し、アニメーションに合わせた演奏動作を生成できる。この方法を適用できない楽器、例えばドラムなどの演奏動作については現在のシステムでは生成することが出来ない。この点は今後早急に改良すべき課題と考えている。

6. 入力装置

本システムで扱うことのできる入力装置は、MIDI規格に準じたMIDI楽器、PCが認識するジョイスティックデバイス、PCキーボード、マウスである。MIDI楽器はキーボード、ドラム、ギターなど様々なものが市販されており、これらを使うことで通常に近い感覚で音楽演奏をすることができる。

また、これら楽器がない場合を想定し、仮想楽器を定義することで、様々な楽器を使って自由に音楽演奏できるようにした。

6.1 仮想コントローラ・仮想楽器

ユーザによる入力設定の違いや今後の対応楽器の拡大などを考えると接続されるであろう入力機器は無数に存在する。概念的な仮想コントローラや仮想楽器を定義することで入力情報を適切な情報へ変換する。

入力デバイスからの入力情報は仮想コントローラ・仮想楽器を通して演奏情報、動作情報に変換され、それぞれサウンド部、グラフィック部に伝えられる。

PCキーボードや、ジョイスティックデバイスなどの入力を仮想楽器に割り当てることにより、MIDI楽器を持たないユーザも演奏に参加することが可能である。本システムではグラフィカルな設定画面を採用し、直感的に設定できるように考慮した(図6)。

入力装置の制御はライブラリにDirectInputを使用して実装を行った。

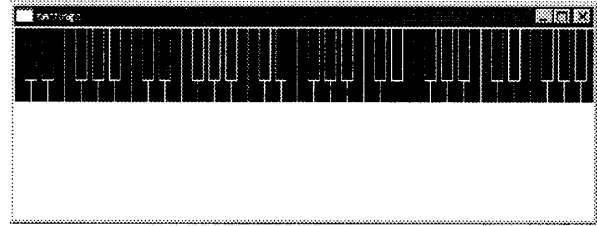


図6 仮想キーボードへの入力割り当て画面

7. むすび

本研究では、ネットワークを利用したコミュニティサイトに関する研究として、音楽セッションのためのクライアントプログラム、ロビーサーバの実装を行った。3DCGアバタを用い、動作を交えながら演奏することにより実世界に近い感覚でのコミュニケーションを行うための工夫をした。通信データとしてMIDIを用い、サンプラーの仕組みを組み込むことで環境によらない再生音を用意することができ、音源選択の自由度も増した。

今回、音楽セッション以外のコミュニケーション機能はロビーチャットの実装程度にとどまった。今後の展望として3DCGアバタのパーツ変更機能などコミュニケーションツールの拡充や、SNSなどの導入によりコミュニティ形成促進を図ることが重要であると考えている。

謝辞

本研究の一部は文部科学省科学研究費補助金、文部科学省私立大学ハイテク・リサーチ・センター補助金による。

文献

- [1] 『社会生活基本調査』総務省(平成13年度), <http://www.stat.go.jp/data/shakai/2001/kodo/pdf/hobby.pdf>
- [2] 小嶋克徳, 近藤育雄, 上島紳一, “コンテンツ集約による音楽セッション方式の検討,” DEWS2004 I-5-02, July 2004
- [3] 後藤真孝, 根山亮, “後藤真孝, 根山亮, “Open RemoteGIG: 遅延を考慮した不特定多数による遠隔セッションシステム,” 情報処理学会論文誌, vol.43, no.2, pp.299-309
- [4] <http://nagasm.suac.net/ASL/GDSM/>
- [5] 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, Apr.1999