

マルチモーダル情報を用いた仮想音楽環境

鈴木 健嗣 橋本 周司

早稲田大学 理工学部 応用物理学科
〒169-8555 東京都新宿区大久保 3-4-1
Tel. 03-5286-3233, Fax. 03-3202-7523
{kenji,shuji}@shalab.phys.waseda.ac.jp

あらまし 筆者らは、展覧会、劇場、芸術作品展示会場において、人間共存型ロボットと人間がインタラクションを行う仮想音楽環境の構築を試みている。人間と機械系が感性的に協調して芸術的な作業を行うための環境を構築するためには、聴覚への刺激である音響ばかりでなく、視覚、触覚への刺激も重要である。そこで、ロボットによる身体性のある「実体インタフェース」の可能性を検討している。本稿では、最近製作したロボットインタフェースを用いたパフォーマンスシステムを紹介する。

キーワード : 人間共存型ロボット、コンピュータ音楽、感性情報処理、マルチモーダルインタフェース、MIDI ネットワーク

Virtual Musical Environment using Multimodal Information

Kenji Suzuki Shuji Hasimoto

Dept. of Applied Physics, Waseda University
3-4-1 Okubo, Shinjuku-ku, 169-8555 Tokyo
Tel. 03-5286-3233, Fax. 03-3202-7523
{kenji,shuji}@shalab.phys.waseda.ac.jp

Abstract This paper introduces an interactive virtual musical environment with multimodal information for communication between human and human cooperative mobile robot. The developed system is a sort of active aids for the performance that allows users to get feedback for emotional activation in terms of not only sound and image but also motion using a mobile robot with multimodal input and output.

Keywords : Human Cooperative Robot, Computer Music, KANSEI Information Processing, Multimodal Interface, MIDI Network

1. はじめに

人間同士のコミュニケーションでは、言語による情報伝達に加え、表情やしぐさによる感性的な非言語情報の認識が重要である。このような観点から、最近のインタフェース研究では、マルチモーダルインタフェースが着目されている。これは人間とコンピュータ間で、感性情報など数値的に表現することが困難な情報伝達を、非言語的なインタフェースで補うものである[1]。また、ペットロボット、アミューズメントロボットの開発が広く行われているが[2]、これらは感性的な情報を伝える身体を持ったインタフェース[3]であると考えることができる。筆者らは、音楽演奏にこのような人間共存型ロボットを使用する研究を行っており、ロボットが、演奏者の動作によってパフォーマンスを行うシステムや、子供たち向けのエキシビションにおいてガイドやナビゲーションを行うロボット[4]について報告してきた。

音楽とは、楽器などを用いて音響により個人の感性や思想を表現する手段であるということができる。音楽の表現の自由度を上げるために、楽器や声帯という物理的制約を超えて、自由な体の動きから音楽を生成する試みがこれまでも発表されてきた[5][6]。楽器の形態を取らず、身体的位置あるいは動きから音を生成するシステムも見られる。さらに、近年、情報空間と実空間を結びつけるためにロボットを利用した試みも報告されている。例として、インターネットを利用して、展覧会場に設置されたロボットがリアルタイムで音楽を演奏し、複数のロボットとコミュニケーションするシステムがある[7]。しかしながら、仮想音楽環境における楽器としてロボットを取り扱う例は少ない。

我々の目的は、人間とロボットの感性的な相互作用を可能にするインタラクティブかつマルチモーダルな音楽環境の実現である。システムは、ユーザの動作や音響を観測し、ジェスチャーや移動、映像や音楽などを利用したマンマシンコミュニケーションを行う環境を提供する。この環境は、画像や音響あるいは音楽やダンスなどを介し、ユーザ

側の意欲を駆り立てる働きを持つフィードバックを可能にする。

本システムの技術的な特徴の一つは、音楽制御ばかりでなく、ロボット部の入出力信号としても、MIDI データを用いていることである。これまでも、各モジュール間の通信を MIDI で行い、Max によるリアルタイム制御を実現したインタラクティブシステムが報告され、有効性が示されているが[8]、ここでは、全システムを MIDI ネットワーク化したため、センサや音源ばかりでなくロボットの制御も Max などの音楽用ソフトウェアで可能になった。これより、音楽家、作曲家や振り付け師などを含むユーザが、音楽の生成だけでなく、ロボットの動作をも容易に取り扱うことが可能になった。

本稿では、最近製作したパフォーマンスシステムを紹介し、人間とインタラクションを行う身体を持ったロボットインタフェースを用いた音楽環境について報告する。

2. 実体インタフェース

従来のロボットは、主に人間の代わりに労働を行うエージェントとしての役割が中心であった。これに対して最近、我々の日常生活を支援する人間共存型ロボットについての研究が盛んになってきた。その応用としては、介護支援、リハビリテーション支援、家事支援、教育支援、アミューズメント、ペットなどがあげられる。

筆者らは、展覧会、劇場、芸術作品展示会場において、音楽を用いた人間共存型ロボットと人間がインタラクションを行う音楽環境の構築を試みている。このように人間と機械系が協調して芸術的な作業を行うための仮構の環境を構築するためには、聴覚への刺激である音楽だけでなく、視覚、触覚への刺激を可能にするロボットによる身体性のある「実体インタフェース」の導入が重要であると思われる。これを用いることで、エアロビクスやジムのためのインタラクティブ運動環境、さらには、動作や作業を伴う教育ツールなどへの応用が可能になる。

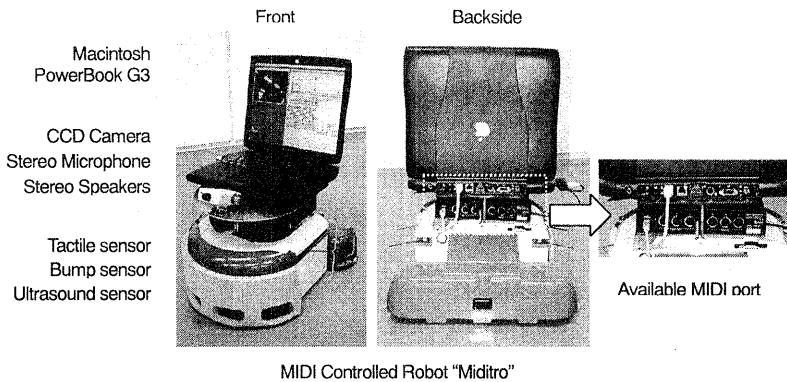


図1 “MIDItro” システム構成
Fig. 1 Overview of the Developed System “MIDItro”

これまで、人間の重心移動を取得するインタフェースを用いた全方向移動ロボットによるパフォーマンスシステム[9]を開発し、ダンサーと協力して実地実験を行ってきた。ここでは、画像、音響情報に加え、触角センサ、赤外線センサ、バンパセンサなどを用いることで、より直接的なインタラクションを可能にするシステムの構築を試みた結果を述べる。

3. システム概要

製作したシステム“MIDItro”の概観を図1に示す。システムのハードウェア部は、ロボット部、メインコントローラ、入出力デバイスの3つの部分よりなる。周囲の環境情報を得るため、CCDカメラとステレオ入力用に2つマイクロフォンをロボット上部に設置した。メインコントローラとして用いるPowerBook G3、およびスピーカを含めたすべての部品をロボット本体に搭載し自律移動を可能にしている。

ソフトウェアとして用いたMaxは、MIDIを扱うためのGUIを備えたオブジェクト指向プログラム環境である。作曲家、演奏家などに良く知られているだけでなく、オブジェクトをつなぎ合わせるというGUIは、提案するシステムのように複数の制御対象を持つ場合に有効であると考えられる。

3.1 ロボット部

本システムで用いたロボットは左右2輪を独立駆動することで、直進、並進、カーブ動作、回転走行を行い、全方向移動を実現している。ロボットの外周部には、図2に示したように、前後左右に計9個のバンパセンサを備え、接触を検地することができる。また、ロボットの前方に2つ、および左右それぞれに1つずつ赤外線センサが取り付けられている。このセンサによって20cmから6m程度の範囲でロボットから障害物までの距離を測定できる。さらに、ロボットの左右側面にそれぞれ2個ずつ、計4個の超音波センサが取り付けられている。これは、壁などの障害物からの距離の測定に利用する。また、衝突時の衝撃を吸収するために、前方にはクッションを取り付けてある。

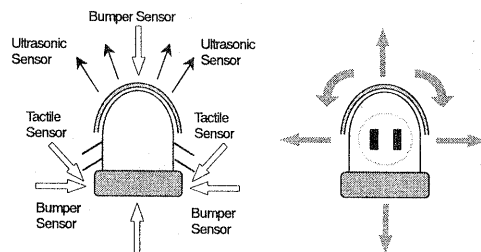


図2 ロボット部のセンサ(左)及び動作(右)
Fig. 2 Sensors (left) and motion (right) of mobile base

3.2 ロボット通信部

走行、停止などのロボットの制御、および各種センサからの情報は、MIDI コンバータを通じてメインコントローラに通信される。ロボットの制御信号およびセンサ情報は、RS232C 回線を経て入出力を行っている。これをマイクロチップを用いたコンバータを用いて MIDI データに変換し、メインコントローラへデータ送受信している。これにより、MIDI 楽器などの音階とロボットの移動動作を Max 上で容易に関連付けることが可能である。

これらにより、システムは、MIDI データをプロトコルとしてモジュール化された MIDI ネットワークを形成する。すべてのモジュールは、メインコントローラの Max によりリアルタイム制御される。図3に MIDI ネットワークの構成を示す。

3.3 システム入力部

ロボットが人間と感性的な音楽パフォーマンスを行うにあたり、システムは以下のような情報を環境情報として取得する。ここでは、ユーザの動作を制限しないよう、体に取り付けるセンサなどは用いなかった。

1) 音響入力モジュール

システムに装備された左右2つのマイクروفोनより、環境の音響情報を取得する。この2つのマイク入力からのボリュームを比較することで簡単な音源定位が可能となる。さらに、このモジュールはピッチ情報を取得することも可能である。基本周波数の推定にはケプストラム法を用いている[10]。また、音響のボリュームのピークを検出し、手拍子のテンポを取得するモジュールを作成した。本モジュールは、ユーザが手拍子やパーカッションでロボットとコミュニケーションすることを可能にする。(図4)さらに、ピッチ情報を用いることで、環境音や手拍子と同様に、歌声や口笛などもロボットに対する刺激として与えることができる。

2) 画像入力モジュール

ロボット周囲の環境をカメラからの動画像により抽出するモジュールを、外部の計算機なしで、Max のオブジェクトとして実現している。図5に示すように、ロボット前部に装備された小型 CCD カメラからより、320×240 (pixel)の画像データを100ms 毎に取得する。CCD カメラは SONY の

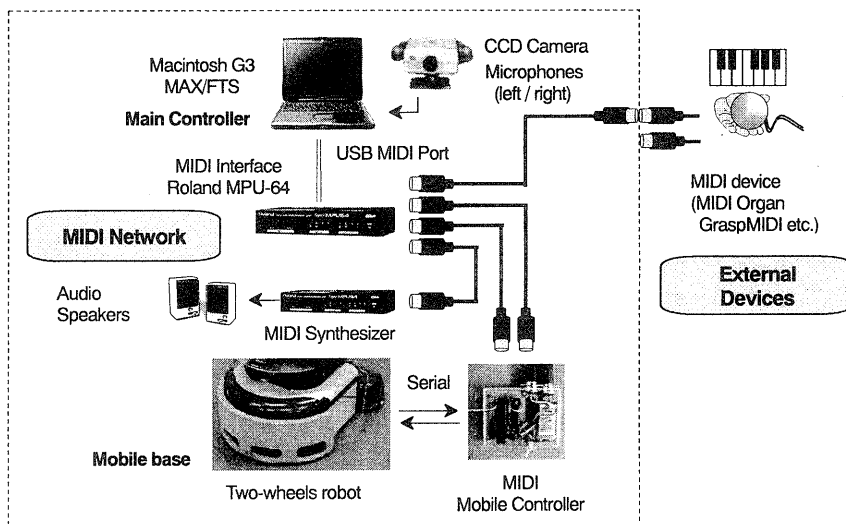


図3 MIDI ネットワーク
Fig. 3 Diagram of MIDI Network

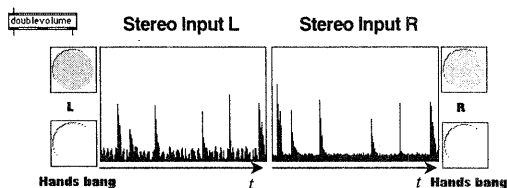


図4 Maxによるステレオマイク入力
および手拍子テンポトラッキング
Fig. 4 Tempo Tracking Object on MAX



図5 CCDカメラと画像入力モジュール
Fig. 5 Camera-based Sensor System and
Image Analyzer Module

CCD-Z1 を用いた。得られた各ピクセル単位でのRGB 情報より、色相、彩度、明度、エッジ密度を計算する。これにより、ロボットは環境の時間変化を色情報の変化として取得する。この時間差分を計算し、その値が一定以上の場合を、環境の顕著な変化の場面と考え、シーンチェンジと定義した。

このモジュールにより、ユーザはロボットに対する意図として、色のついた旗や、LED ライトなどを用いることで、ロボットに刺激を与えることが可能になる。

3) センサ入力部

ロボットに備えられた各センサの情報は、1sec 毎に1パケットとして、MIDI データに変換されてメインコントローラへ送信される。

4) 外部 MIDI 入力部

システムは、後部に外部 MIDI 入力ポートを備えている。MIDI ネットワークによりすべての入力は MIDI データとしてメインコントローラへ送られ

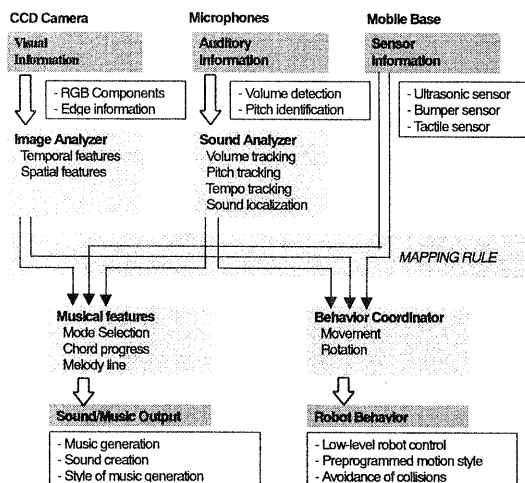


図6 システムによる音楽生成
Fig. 6 System Configuration

るため、MIDI キーボード、MIDI フルートなどの楽器をコントローラとして使用できる。また、握り動作を MIDI データに変換する GraspMIDI[11]なども利用できる。

4. 実地実験

本システムの有効性を確かめるため、実地実験を行った。取得可能な入力データによる音楽生成ルールは、作曲家あるいは演奏家によって決定される。システムの出力は、音響音楽出力、ロボット動作の2つの部分からなる。ここでは、実験に用いた音楽生成ルールの一例をあげる。

1) 音響音楽出力

ロボットは、ユーザがパフォーマンスを開始した後、演奏を始める。本システムはステレオスピーカを搭載するため、音響音楽出力はすべてロボット本体から出力される。音楽はセンサからのイベント情報により動的に変化させることができ、Max プログラミングにより多様な作曲が可能であるが、実験では、あらかじめ準備されたノートセットから、カメラからの色情報により確率的に音響を生成し、ロ

Robot movement	MIDI	Primary Factor
Basic movement		Bumper Sensor
Forward	B-01	Contact at backside
Backward	B-02	Contact at front
Sliding		Tactile Sensor
left direction	B-03	Contact at right part
right direction	B-04	Contact at left part
Rotate		Sound Localization
clockwise	B-05	Source located at right
counterclockwise	B-06	Source located at left
Zigzag motion	B-07	Randomly occurred
Circle motion	B-08	Scene Change

表1 ロボット動作とセンサ入力対応関係の例
Table 1. An example set of primary factors for robot movement

ロボットのセンサ情報より音色の変更、強弱を決定することにした。

2) ロボット動作

メインコントローラにより生成されるロボット動作は、表1に示すような対応付けを行った。今回、ロボットの動作は、前進・後進、左右方向への並進、時計回り・反時計回りの回転の6種類に限定した。最大移動速度は30cm/sである。

5. まとめ

本稿では、音楽を用いた人間共存型ロボットと人間がインタラクションを行う音楽環境について述べ、移動ロボットを用いたパフォーマンスシステムの構築例を報告した。実地実験では、ダンサーは音楽的な構造よりも、ロボットの動作、反応などを注目していた。ロボットを仮想楽器として用いる試みはいまだ発展段階ではあるが、ダンサー・演奏者から好意的な意見を得られた。このようなパフォーマンスシステムに対する定性的な評価は一般に困難であるが、身体性のあるインタフェースであるロボットに対し、ユーザが動作で自身の意図を伝えることによって音楽が生成されるだけでなく、ロボットの動作もフィードバックされることにより、仮想音楽環境における演奏者の創造性を拡張できると思われる。また、提案したMIDIネットワークを用いることで、作曲家や舞踏家などが容易にロ

ボットを用いた作曲、演出活動を行うことが可能になるため、新しい音楽芸術の可能性が広がると考えられる。

現在、作曲家と協力し、提案システムを用いたマルチメディアコンサートの計画を立てている。

参考文献

- [1] 黒川 隆夫, ノンバーバルインタフェース, 電子情報通信学会編ヒューマンコミュニケーション工学シリーズ, オーム社, 1994
- [2] 藤田 雅春: "ペット型ロボットの感性表現", 日本ロボット学会誌, Vol. 17, No. 7, pp.947-951, 1999
- [3] 中津 良平, "人間の感性とロボット, エージェント", 日本ロボット学会誌, Vol. 17, No. 7, pp. 920-027, 1999
- [4] K. Suzuki, A. Camurri, P. Ferrentino and S. Hashimoto: "Intelligent Agent System for Human-Robot Interaction through Artificial Emotion", in Proc. of IEEE Intl. conf. on System, Man and Cybernetics, pp. 1055-1060, 1998
- [5] H. Katayose, T. Kanamori, K. Kamei, Y. Nagashima, K. Sato, S. Inokuchi and S. Shimura, "Virtual Performer", in Proc. of ICMC93, pp. 138-0145, 1993
- [6] Joseph Paradiso, Kai-Yuh Hsiao and Eric Hu, "Interactive Music for Instrumented Dancing Shoes" in Proceedings of ICMC99, pp. 453-456, Beijing, 1999
- [7] Kouichirou Eto, SoundCreatures, 1998
<http://www.canon.co.jp/cast/artlab/scweb/>
- [8] 長嶋 洋一, 光岡 正樹, 井口 征士: "生体センサによる音楽表現の拡大と演奏表現の支援について", 情報処理学会研究報告, 音楽情報科学, 25-11, pp. 75-81, 1998
- [9] Kenji Suzuki, Takeshi Ohashi and Shuji Hashimoto: "Interactive Multimodal Mobile Robot for Musical Performance", in Proc. of ICMC99, pp. 407-410, 1999
- [10] 東 英司, 橋本 周司: "母音認識とピッチ検出を用いた歌声のテンポ抽出 第3報", 情報処理学会研究報告, 音楽情報科学, 26-3, pp.17-22, 1999
- [11] 澤田 秀之, 橋本 周司: "ハンドジェスチャ入力デバイスによる音響生成", 電子情報通信学会論文誌 D-II, Vol.J81-D-II, No.5, pp. 795-803, 1998