

内合22-65

早稻田大学大学院理工学研究科

# 博士論文審査報告書

## 論 文 題 目

A Study on Multimodal Human-Machine  
Communication

- Realization of Physical, Intelligent and Kansel Interaction -

人間一機械系におけるマルチモーダルコミュニケーション  
に関する研究  
—物理的、論理的、感性的な相互作用の実現—

申 請 者  
鈴木 健嗣

Kenji Suzuki

物理学及応用物理学専攻・情報工学研究

2003年3月

従来、人間の論理的な活動としての知性を規範として研究されてきた情報処理が、最近は感性という人間の情緒的側面も対象にするようになってきた。人間の心の動きは、五感により検知された外界情報により大きな影響を受け、これが感性的な反応として現れる。感性的情報は、主観的かつ多義的であり個人や状況への依存性が強く、多くの点で従来の情報処理の対象であった論理的な知識情報とは異なる。しかしながら、我々人間の行動は、知識情報ばかりでなく感性情報によるところも多いのである。したがって、誰でも使える察しの良いシステム、いわゆる「人に優しいシステム」には、物理的あるいは論理的なインターフェースばかりでなく感性的な情報のやり取りも必要不可欠である。また、現在の GUI(Graphical User Interface)を超えて、人間と機械のより柔軟なコミュニケーションを実現するには、システムに身体性を付与することが重要であると言われている。情報処理技術、通信技術、機械技術の発展に伴い、PostGUIとも言うべき次世代インターフェースの研究が盛んになり、多くの試みがなされているが、未だ、決定的な解は見出されていない。

このような背景の下で、本論文の著者は、人間と機械の間に物理的、論理的、感性的な相互作用を実現するマルチモーダルコミュニケーションに関する研究を行ってきた。著者によれば、マルチモーダルコミュニケーションとは、視覚、聴覚、触覚、嗅覚、味覚に代表される人間の感覚を通じ、画像、音響、言語、身振り、表情などの様々なモダリティ(様式)を統合した情報の複合的なやりとりのことであり、本論文には、これに関わる一連の研究成果が纏められている。

本論文は英文で全9章からなっている。以下にその概要と評価を述べる。

第1章は序論で、本研究の位置づけと研究の背景を述べている。

第2章では、従来から現在までの代表的なマルチモーダルコミュニケーション研究を紹介した後、人間と機械間のコミュニケーションの階層型モデルを提案している。モデルは、物理的、論理的、感性的コミュニケーションの3階層からなるが、著者はコミュニケーションにおけるこれらの役割を述べて、階層化処理の特徴と優位性を明らかにしている。

第3章では、人間の感覚の各モダリティにおける物理的および論理的なコミュニケーションの具体的な内容を説明し、第4章で、これらを実現する人間型ロボットの設計・製作と実験について述べている。製作したロボットは26自由度の人間型ロボットであり、体内LANによってハードウェアおよびソフトウェアのモジュールを統合する並列分散型の新しいアーキテクチャをとっている。また、人間や環境との物理的な相互作用に関わる信号処理の経路とパターン認識を伴う知的な相互作用の処理経路をハードウェア的にも独立した階層としたところが、このロボットシステムの大きな特徴である。実験では、物理的コミュニケーションと論理的コミュニケーションの実例とし

て、力覚による制御・動作、把持行動（握手）による動作、物体追従／到達動作、環境音への反応動作および音声による動作を取り上げ、内部での情報処理、および外界からのセンサ情報とロボット動作の関連付けの方法が示されている。さらに、これらモダリティ毎に異なる動作の複合が引き起こす効果が実験結果と共に述べられている。これらの結果から、提案手法が人間と機械の間の柔軟なコミュニケーションを容易にするばかりでなく、対障性などシステムのロバストネスの向上にも有効であることが判る。

第5章では、この論文で取り上げる感性的コミュニケーションの意味を明確にし、第6章で、人間と機械の感性的コミュニケーションを音楽によって行うロボットに関する研究成果が述べられている。まず、ロボットの自律性と人間を含む環境への従属性、および、人間との相互作用の強さに着目し、音楽を通じたコミュニケーションに関する人間と機械の関係をモデル化し、それに基づいて視覚、聴覚および自動作曲機能を持った3機のロボットを製作している。第1は、ニューラルネットワークによる人工感情モジュールを内蔵し、環境からの画像情報に基づいて音楽を生成しながら動き回るロボットであり、第2は、音楽を生成するばかりでなく人間を乗せることができるロボットである。乗った人間の重心移動によりロボットの運動が制御される。第3は、電子楽器用インターフェースであるMIDIにより制御され、人間のダンサーと即興的な共演を行うロボットである。これらはいずれも、人間の感性的な表現の補助あるいは活性化に有効なものであり、身体性を持った一種の楽器ということもできる。このような感性的なシステムの評価を定量的に行うこととは困難であるが、美術館、音楽関係の学会などの実演によりその効果は確認されており、メディアとしてのロボットの新しい応用の可能性を拓くものとして評価できる。

第7章は、マルチモーダルコミュニケーションのためのシステム製作に関する前章までの研究と並行して行なわれた感性的な情報処理の理論的な研究成果を述べた章である。ここでは、感性情報を定量化するためのニューラルネットワークを提案し、提案アルゴリズムの構造および動作特性について数学的な解析が示されている。さらに、感性的な情報を教師信号とする実データでの実験を通じて、提案モデルが感性情報の定量化に有効であることが示されている。提案されている学習法は、2つの入力パターン間の距離を教師信号として与えるもので、入力パターンの空間から出力パターンの空間へこの距離関係を満たすように移す非線形写像を得ることができる。これにより、物理的特徴で記述されたパターンを心理的な位相関係を表現する空間への写像を得ることができ、感性情報の新しい数量化が可能となった。著者は、色彩、音響、顔表情に関する一対比較の主観評価尺度と計測可能な物理特徴の関連付けに提案手法を適用し、従来の代表的な数量化法である多次元尺度構成法（Multidimensional Scaling, MDS）および林の数量化理論（Hayashi's Quantization Theory, HQT）と比較検討した。提案手法は、心理評価に基づ

くデータの関係が低次元の空間で可視化できるばかりでなく、解析時に用いなかつたデータの評価が可能であるという従来手法にはない利点も有している。また、この手法を個人の主観評価データに適用することで、感性の個人差の視覚化が可能になるということも示されている。これらの結果は、前章までの具体的なシステムの設計においても有効に活用できるものである。

第8章では、全体としての考察と評価を行い、第9章は、研究成果のまとめと将来展望を述べている。

以上のように、本論文は、人間の五感に対応したマルチモダリティを有するシステムによって、人間と機械が物理的、論理的、感性的にコミュニケーションを行うためのシステム設計の原理として、新しい階層モデルを提案し、これを人間型ロボットの設計、製作に適用して有効性を示している。また、論理化が困難な感性情報に関わる人間と機械の間のコミュニケーションを、音楽の協調演奏として実現するロボットを製作して、多くの実演によって实用性を検証している。さらに、感性情報の理論的な扱いについても検討し、パターン間の距離を規範とするニューラルネットワークの新しい学習法によって、パターンの物理的特徴空間から感性的心理空間への非線形写像を生成することを可能にした。

これらの成果は、感性情報処理およびヒューマンインターフェースのみならず、計測工学、情報工学およびロボット工学の発展に資すること大である。よって、本論文は博士(工学)の学位論文として相応しいものであると認める。

2003年2月

#### 審査員

(主査) 早稲田大学教授	工学博士 (早稲田大学)	橋本周司
早稲田大学教授	工学博士 (東京工業大学)	久村富持
早稲田大学教授	博士(工学) (早稲田大学)	中島啓幾
早稲田大学教授	博士(理学) (大阪大学)	竹内 淳
早稲田大学教授	工学博士 (早稲田大学)	高西淳夫