ヒューマンフレンドリーなマン・マシンインターフェースの構築の試み —人間のコミュニケーションの検討とその応用—

正 員 樹野 淳也 (東京農業大)
非会員 小久保 芳行 (法政大)
非会員 松村 勇 (法政大)
正 員 小林 尚登 (法政大)

An Attempt on Construction of Human Friendly Man-Machine Interface — Study and Apply about Human Communication —

Junya TATSUNO Member (Tokyo University of Agriculture)
Yoshiyuki KOKUBO Non Member (Hosei University)
Isamu MATSUMURA Non Member (Hosei University)
Hisato KOBAYASHI Member (Hosei University)

This paper describes an attempt on a construction way of human friendly man-machine interface. At first, we do a simple experiment to find out the characteristic of human verbal communication. From the experimental results, we get some rules in case in human verbal communication. We construct the man-machine interface which is based on these rules. Through teaching process, we examine our verbal communication interface comparing with conventional interfaces. From this comparison, we recognize that the verbal communication interface is valid to construct the user-friendly man-machine interface.

キーワード： ヒューマンフレンドリー、マンマシンインターフェース、バーバルコミュニケーション

1 まえがき

現在、ロボットは主に自動車産業や電気電子産業において利用されている。今後は、原子力産業や建設業、農業を始めとする第一次産業など非製造業向けのロボットの需要は確実に増えることが予想され、しかも、それらの多くは知能ロボットになると言われている[1]。上記に示したような非製造業分野では、ロボットは不均一で整備されていない環境の下で、多様な作業に従事しなければならない。また、ロボットの周囲には様々な人間が存在し、介在することになる。そのような場合、熟練した作業者が複雑な整備で作業の教示を行うのではなく、誰もが容易に教示作業を行えるようなヒューマン・フレンドリーなインターフェースが必要になる。

ロボットに望ましい作業を行わせるための基本的な方法であるティーチングは、オンラインプログラミングとオフラインプログラミングに分類できる[2]。オンラインプログラミングは、ロボットを実際の作業内容に従って動作させその履歴を記憶させる手法である。一方、オフラインプログラミングは、シミュレータを用いロボットや作業対象をモデル化し、作業行程におけるロボットの動作を生成する方法である。更に、テレオペレーションによりロボットを動作させる手法もティーチングの一手法であろう。このテレオペレーションは近ごろでは非産業分野で主に使われており、近い将来、テレ

オペレーションロボットは日常生活において、あらゆる所で使用されることになりそうである。

しかし、これらティーチングという作業はオペレータにとって容易な作業ではない。なぜなら、現段階ではティーチングに用いられているインターフェースのコミュニケーション形態が、人間の自然なコミュニケーションとは全く異なる形態を採っているからである。言い換えれば、人間の自然なコミュニケーションをマン・マシンインターフェースに取り入れることは、ヒューマンフレンドリーなインターフェースの構築の最初のステップになりうるということである。これまでに、ロボットの教示を目的とした人間の行動を画像認識プログラムを生成するような研究[3]や同様に画像認識よりスキルレベルの作業モデルを構築する研究[4]、視覚からの非言語情報とともに自然言語形式の命令をロボットに与える研究[5]、更に、ロボットに意を組み入れることの有効性が議論され[6]、人間の意図を人間の行動より能動的に理解する機能をもったインターフェースも研究されている[7]。筆者らもヒューマンフレンドリーなインターフェースの構築に関する研究を行ってきた[8][9]。

そこで、本論文では、このような研究の一環としてマン・マシンインターフェースの構築において人間のコミュニケーション形態を模倣することを提案する。まず、第2章においてコミュニケーションの形態について考察し、第3章にて人

Fig.1 Cannels of human - human communication in case of fase-to-fase

図1 対面した人間のコミュニケーションのチャネル

間のコミュニケーションの傾向を簡単な実験より検討する。更に、第4章にてその結果をもとにロボットと人間とのインターフェースを音声コミュニケーションを用いて構築し、従来の産業用ロボットに使用されているインターフェースとの比較を行うタスク教示実験を報告する。最後に、第5章にて結論を述べる。

2 コミュニケーション形態

まず最初に、コミュニケーションの形態について考察する。従来の産業用ロボットは、前述のように、オンラインティーチングやオフラインティーチングなどにより教示されている。ダイレクトティーチングに代表されるように、オンラインティーチングはコミュニケーションツールに大きく依存していないため、ここでは特に取り上げないことにする。一方、ロボット言語などを用いたオフラインティーチングの場合、そのインターフェースは人間が理解しやすいものになってきたが、人間が通常行っているコミュニケーションの形態とは大きな差が存在する。

Fig.1に通常、我々人間が行っているコミュニケーションのモデル図を示す。我々人間のコミュニケーションは言語コミュニケーションと非言語コミュニケーションの二つからなる。更に、言語コミュニケーションの形態は音声と文字のチャネルに分けられ、非言語コミュニケーションは人体、動作、目、周辺言語、沈黙、身体接触、対人的空間、時間、色彩の9個のチャネルを持っている[10]。

人と人が対面してコミュニケーションする場合、その内容の35%が言語コミュニケーションで伝達され、残りの65%は非言語コミュニケーションを用いている[11]。しかし、この事実は人と人が対面している場面に限ってのことであり、常に、コミュニケーションにすべてのチャネルが使用できるとは限らない。たとえば、我々は日常よく電話を使用するが、その場合、当然、そこでのコミュニケーションは言語コミュニケーションで、さらに、音声チャネルのみが使用されている。し

かし、チャネルが音声に限られても、我々は、それほど不自由さを感じないし、仕事や作業の指示も十分に行うことができる。

更に、電話で誰かに物を扱ってもらう場合を想像してみる。このような場合、指示者はキーワードの組み合わせ(“腕を+伸ばす”, “腕を+縮める”, “手を+挙げる”, など)により動作指令を伝え、被指示者は指令を的確に遂行する事ができている。

従って、上記のことから、人間にとってフレンドリーなマン・マシンインターフェースを実現するには、人間のコミュニケーションの形態の模倣が有効であることが考えられる。次章において、その人間のコミュニケーションの傾向について議論する。

3 人間のコミュニケーション

ヒューマンフレンドリーなマン・マシンインターフェースを実現するには、人間のコミュニケーションの形態の模倣が重要であると考えているが、それらを応用するにはその特徴、傾向を熟知しておく必要がある。ここでは、人間のバーバルなコミュニケーション、特に音声コミュニケーションについてその特徴の抽出を行うための実験を報告する。

これまでに、ロボティクスの研究分野において、人間の五感(視覚、聴覚、触覚、味覚、臭覚)を再現するという試みがなされたきた。実用レベルに達しているものも数多くあるが、実際に産業界に導入されている知能ロボットは、知能という冠が付いていても、CCDカメラを付けた視覚機能を持つものや、手先に力・トルクセンサを付けた力覚フィードバック機能を持っている程度であろう。従って、その程度の知能を持つロボットとのコミュニケーションへの応用を仮定する。

以上のことをふまえ、まず、人間階層のコミュニケーションの特徴抽出のための実験を行った。

3.1 実験概要

Fig.2 Outline of experiment
図2 実験の概略

Fig.2は実験の概略図を示す。実験は被験者A、被験者Bの二人一組で行われる。被験者Aは被験者Bに指示を送り、被験者Bはその指示通りに動作をする。被験者Bの前に壁があり、作業空間を左右に分断している。3つの缶を壁の左側に置く。実験のタスクはこれらの缶を右側に移動させ、左側にあったように缶を再び積み直すことである。このタスクを、Aの指示に従いBが実行する。ここで、AからBへの指示は音声のみで送られる。Bの知的レベルを人為的に下げるために視界を遮り視覚情報なしで動作を行わせるようにした。実験は10組のペアについて行った。

3.2 結果

ここで我々が知りたいことは、人間のコミュニケーション形態の特徴抽出である。言い換えれば、被験者Aが被験者Bに対し自分の意志を伝えるためにどのような言葉を使用するかである。まず、例としてある被験者の組の実験の結果を示す。その被験者の組が行ったタスクのプロセスは次のようであった。また、Fig.3はそれぞれのプロセスにおける最終状態の姿勢である。

process1 初期状態から対象物aを掴む

process2 対象物aをx'地点へ移動

process3 対象物cを掴む

process4 対象物cをz地点へ移動

process5 対象物bを掴む

process6 対象物bをy地点へ移動

process7 再び対象物aを掴む

process8 対象物aをx地点へ移動

更にそれぞれのプロセスにおいてこの組の被験者Aが被験者Bに指令するために用いた言葉は以下の通りであった。

process1 左に → もう少し左に → 前に出す → もっと前に → 下げる → 行き過ぎ → 前 → もう少し前 → 掴む

process1

process2

process3

process4

process5

process6

process7

process8

Fig.3 Last posture of each process
図3 各プロセスの最終姿勢

process2 引く → 右に → もう少し右 → 下 → 少し前 → もう少し下 → 置く

process3 引く → 左 → 左 → 少し戻す → もう少し → 少し前 → 掴む

process4 引いて → 右 → もう少し → 少し前 → 離す

process5 引く → 左 → もう少し → 前 → 少し右 → もうちょっと右 → 少し前に → 掴む

process6 引いて → 左 → 少し左 → もう少し → 少し前 → 置いて

process7 引いて → ほんの少し右 → ほんの少し前 → 掴む

process8 上げて → 上げすぎ → 左 → 少し下げて → 少し前 → もうちょっと左 → 少し下げて → 離す

更に全被験者の実験結果より考察を進める。まず、用いられた指令のほとんどのものが上記の例のように単純な手先の動きやハンドの動きを表す原始的動作レベル[12]の指令であることが分かる。また、手先の動作を表す指令は、程度を表すキーワードと方向を表すキーワードの組み合わせによる構成である。方向を表すキーワードはFig.4に示すように直交座標成分を示すものがほとんどであった。それぞれの方向成

Fig.4 Model of orientation keywords
図4 方向を示すキーワードのモデル

分において用いられたキーワードをTable 1にまとめる。

Table 1 Result of orientation keywords	
表1 方向を示すキーワードの結果	
ORIENTATION	KEYWORDS
up	上に, 上, 上げる
down	下に, 下, 下げる, 下ろす
right	右に, 右
left	左に, 左
forward	前に, 前, 伸ばす
back	手前に, 手前, 引く
other	行き過ぎ, 戻す

どれもが単純なキーワードであるが、上下方向、前後方向に関しては副詞や名詞のみを用いたり、方向を示す動詞を用いている。なかでも、前後方向の「伸ばす」や「引く」といった言葉は腕の動作独特のものである。また、左右方向は副詞や名詞のみを用いた指令となっている。その他、特徴あるキーワードとして「行き過ぎ」や「戻す」といった表現が見受けられた。これらのキーワードにより全く逆方向の動作を期待する

ことができ、「行き過ぎ」というキーワードについてはそれ自体が移動距離の程度も含んでいるようである。

程度を表すキーワードは次のようなものが用いられた。尚、キーワードの後ろの数字は10組の被験者のうちそのキーワードを用いた組の数を示している。

Table 2 Result of adjective keywords
表2 副詞的なキーワードの結果

少し	(10/10)
もう少し	(10/10)
もっと	(7/10)
ちょっと	(5/10)
もうちょっと	(5/10)
大きく	(5/10)
ほんの少し	(3/10)
若干	(1/10)
ゆっくり	(1/10)

表2より、移動距離に関しての程度を表しているキーワードがほとんどで、移動速度を表すものは「ゆっくり」のみでその頻度はほとんどない。また、短い距離を表すキーワードが多様でかつ頻度が多い。長い距離に関しては、方向を表すキーワードを繰り返すことにより実現していることが確認できた。その他の特徴として、これらの程度を表すキーワードのみを用いた場合は直前のキーワードを同じ方向で移動距離程度を変化させることができる。

4 ロボットとのコミュニケーション

前章の実験の結果より、人間が指示のために用いるキーワードの数は限定され、また、文法形態にも規則性が見られる。更に、あるキーワードによる動作には特殊な性質を見出すことができた。以上のことより、人間と機械とのコミュニケーションにそれらの規則性を適用することによりヒューマンフレンドリーなマン・マシンインターフェースを構築できると思われる。そこで、本章では、実際にロボットとのコミュニケーションにキーワードによる命令規則を用いた方法とそれらを用いた実験について報告する。

4.1 概要

Fig.5に示すようなシステムを構築した。システムは大きく分けて音声認識を行うコンピュータ(CPU:Pentium 133MHz)、プログラムのエディット・コンパイルを行うワークステーション(Sun)、ロボットの制御を行うVMEコントローラ(CPU:Micro Sparc, OS:VxWorks)、オペレータからの命令を実行するロボットから構成される。また、音声認識処理を行うアプリケーションとしてIBM社製のVoice Type Ver. 3.0を用い、制御するロボットはFig.6に示すような人間の腕と同構造の7自由度のマニピュレータである。それぞれの端

Fig.5 Structure of the System

図5 システムの構成

末はイーサネットで結合されており、WindowsとVxWorksとの通信プロトコルにはTCP/IP、インターフェースにはソケットを採用した。

Fig.6 View of 7 DOF redundant manipulator

図6 7自由度マニピュレータの外観

処理の流れは次のようになる。オペレータの指示音声はPC/AT互換機の音声認識アプリケーションにより認識される。指示データはソケットを介しVMEコントローラのロボット制御を行うモジュールに送られ、ロボットはそれらの指示データをもとに動作を行う。

4.2 実験

このシステムを用い以下の様な実験を行った。

Fig.7に示すようにロボットの全面に2つの対象物を設置し、ロボットにその対象物を順に叩かせるというタスクを行う。このタスク教示を、構築した音声入力によるインターフェースと通常用いられるティーチングペンダントを用いて行い、その結果を比較する。

音声入力インターフェースに用いたキーワードを以下に示す。

動作方向 : 上, 下, 右, 左, 前, 手前

Fig.7 Outline of experiment

図7 実験の概略

動作程度 : 大きく, もっと, 少し, ちょっと

3章で述べた人間同士のコミュニケーションの実験の結果を基に、入力指令は程度を表すキーワードと方向を表すキーワードの組み合わせにとし、動作方向は直行座標方向のみとした。また、それぞれのキーワードは3章の実験結果にて頻度の高い以下のキーワードを用いた。程度を表すキーワードは動作距離を示すもののみとし、“大きく”、“もっと”、“少し”、“ちょっと”はそれぞれ40cm, 40cm, 10cm, 5cmの動作距離としデフォルトは25cmとした。

ティーチングペンダントも同様に直交座標指令のみを用い、動作速度はいずれのインターフェースもほぼ同様の速度に設定した。

被験者はロボット教示に携わっていない学生10名とし、それぞれのインターフェースを用い各被験者はタスク教示を5回ずつ実行した。

Fig.8 Experimental result (time)

図8 実験結果 (時間)

Fig.8, Fig.9に結果を示す。Fig.8は試行回数とタスク教示の所要時間の関係を示し、Fig.9は試行回数とタスク教示中のロボットの手首位置の総移動距離を示している。それぞれ

Fig.9 Experimental result (length)
図9 実験結果 (距離)

のグラフにおいて、塗りつぶし丸及び実線はティーチングペンダントを用いた場合の全被験者の平均データ、その近似曲線を表している。三角印及び波線は音声入力の場合の平均データとその近似曲線である。これらのグラフより、ティーチングペンダントを用いた場合は試行回数が増えるにつれて所要時間、移動距離ともに減少することが見受けられる。一方、音声入力の場合は、試行回数によらず各データはほぼ一定値を示している。ティーチングペンダントを用いた場合のデータが左下がりな結果を示す理由として、被験者がティーチングペンダントの操作の熟練や経験からの教示手順の学習効果等が挙げられる。音声指令のデータがほぼ一定値を示すのは、このコミュニケーション手法が熟練を必要としないことを裏付けている。

両方式における結果を比較すると、試行回数が低い場合は音声入力による教示のデータはティーチングペンダントより低い値をとり、試行回数が増えるとティーチングペンダントのデータが低い値を示している。これは、試行回数が増えた場合にデータが逆転する原因として、実験に用いた音声認識ソフトの認識率によるものであると考えられる。十分な認識率を得ることができないため、入力を繰り返すことがしばしば起こったことや、さらに認識処理が重い処理のため実行に時間が掛かることが見受けられた。

しかし、試行回数が低い場合のデータを比較した場合のデータを見る限り、音声入力によるインターフェースは熟練を多く必要としないため容易にタスク教示を行うことができる。このことから音声入力を用いたマン・マシンインターフェースを用いることは、ユーザーフレンドリーなインターフェースが構築する一手法であることがわかる。

5 むすび

本稿では、ユーザーフレンドリーなマン・マシンインターフェースの構築について論じた。まず、人間機械系のコミュニ

ケーション形態に人間が通常行っているコミュニケーション形態を模倣することを提案し、人間のバーバルコミュニケーション、特に音声コミュニケーションの傾向を見いだした。その傾向に基づきキーワードにより構成されるロボット指令体系を構築し、実際にロボット教示の実験を行った。実験は非常にシンプルなものであったが、バーバルコミュニケーションを用いたインターフェースは熟練を必要としないため、比較的容易に教示が行えるという結果を得ることができ、このインターフェースの方向性が有効であることを確認した。しかし、我々の目標とするユーザーフレンドリーなインターフェースの完成には解決しなければならない多くの問題がある。まず、音声入力に関する問題である。今回の実験では市販の音声認識アプリケーションを用いたため、高い認識率は得ることができなかった。認識率の高い音声認識システムの導入は不可欠である。また、今回の実験に用いた指令キーワードは原始的動作レベルの指令のみであった。多様な動作を実現するにはタスク動作レベルの指令も用意する必要がある。そのためには画像情報や力情報を用いた制御手法を取り入れなければならない。

謝辞 本研究において、貴重なご助言を賜りました東京農業大学農業工学科玉木浩二教授、田島淳講師に感謝の意を表します。

(平成9年7月23日受付)

文 献

- [1] 例えば、梅谷: "ロボットに輝かしい未来はあるか?", 日本機械学会論文集 (C編), Vol. 62, No. 599, pp. 2517-2519, 1996
- [2] 例えば、日本ロボット学会誌 (教示特集), Vol. 13, No. 5, 1995
- [3] 國吉, 井上, 稲葉: "人間が実演して見せる作業の実時間視覚認識とそのロボット教示への応用", 日本ロボット学会誌, Vol. 9, No. 3, pp. 295-303, 1991
- [4] 池内, 末広: "視覚による組立作業理解のための作業モデルとそれに基づく動作生成", 日本ロボット学会誌, Vol. 11, No. 2, pp. 281-290, 1993
- [5] 藤沢, 荒牧, 吉村, 首藤, 阿久根: "視覚情報と言語情報を用いたヒューマンロボットインターフェースの実現法", 第14回日本ロボット学会学術講演会予稿集, pp. 337-338, 1996
- [6] 佐藤, 溝口: "心 (知・情・意) と「意」の情報処理 - 「意」をもちいるロボットの実現 -", 日本機械学会誌, Vol. 98, No. 914, pp. 38-42, 1995

- [7] 佐藤, 西田, 市川, 畑村, 溝口: "ロボットによる人間の意図の能動的理解", 日本ロボット学会誌, Vol. 13, No. 4, pp. 545-552, 1995
- [8] J.Tatsuno, S.Matsuyama, Y.Kokubo, H.Kobayashi: "Human Friendly Teaching for Industrial Robot", Proc. of 5th IEEE International Workshop on Robot and Human Communicaiton, pp.456-460, 1996
- [9] 小久保, 松村, 小林, 樹野: "バーバルコミュニケーションを用いたマン・マシンインターフェースの一構成法", 日本機械学会ロボティクス・メカトロニクス講演会講演論文集, Vol. B, pp. 981-982, 1997
- [10] Marjorie fink Vargas, 石丸正訳: "非言語コミュニケーション", 新潮選書, 1987
- [11] 立川敬二ら: "コミュニケーションの構造", NTT 出版, 1993
- [12] 井上: "ロボット言語の研究課題", 日本ロボット学会誌, Vol. 2, No. 2, pp. 87-90, 1984

松村 勇 (非会員) 1974年8月29日生まれ。平成9年法政大学工学部電気工学科電気電子専攻卒業。同年同大学大学院工学研究科電気工学専攻修士課程入学。現在に至る。主として、遠隔操作およびヒューマンインターフェースの研究に従事。

小林 尚登 (正員) 1951年11月27日生まれ。昭和51年理工学研究科電気工学専攻博士前期課程修了。現在、法政大学工学部電気電子工学科教授。1988年～1989年 Stuttgart 大学 (ドイツ) 客員研究員 (フンボルト財団招聘)。計測自動制御学会論文賞等受賞。主として、ロボティクス、制御理論及びその応用に関する研究に従事。IEEE, 計測自動制御学会, 日本ロボット学会, 日本機械学会などの会員。工学博士 (早稲田大学)。

樹野 淳也 (正員) 1971年2月8日生まれ。平成5年法政大学工学部電気工学科電気電子専攻卒業。平成7年同大学大学院工学研究科電気工学専攻修士課程修了。同年同博士後期課程入学, 平成8年退学, 同年東京農業大学農学部農業工学科助手現在に至る。主として、遠隔操作およびヒューマンインターフェースの研究に従事。日本ロボット学会, 日本機械学会, 日本人間工学会などの会員。

小久保 芳行 (非会員) 1974年1月8日生まれ。平成8年法政大学工学部電気工学科電気電子専攻卒業。同年同大学大学院工学研究科電気工学専攻修士課程入学。現在に至る。主として、遠隔操作およびヒューマンインターフェースの研究に従事。日本機械学会の会員。