

パーソナルロボットのための アクティブインタフェースの設計と実装

山崎 信行* 安西 祐一郎*

Design and Implementation of *Active Interface* for Personal Robots

Nobuyuki Yamasaki* and Yuichiro Anzai*

We believe that personal robots will be used in our office and home like current personal computers in the near future. In this paper, we discuss a user-friendly interface design for personal robots. Then we consider features of personal robots and their environments, and propose a new user interface concept called *Active Interface* for human-robot interaction. *Active Interface* keeps a system in an advantageous condition by using not only users' explicit inputs but also users' implicit inputs and the external information from various sensors. To show the effectiveness of *Active Interface* concept, we design and implement a speech dialogue system: *Chaser* based on *Active Interface*. From the experiments of *Chaser*, we ascertained the efficiency of *Active Interface*.

Key Words: *Active Interface*, Personal Robot, Human-Robot Interaction, Speech Dialogue System

1. はじめに

高性能ワークステーションとそれらを結合するネットワークの普及により、オフィスにおける計算機システムの環境は、1人が1台のマシンを占有する形態に発展しつつある。これに伴い、計算機も複数の人々が互いにコミュニケーションしながら共同で作業を行うために利用される機会が多くなってきた。このようなオフィスにおける協調作業支援の重要性は、すでに CSCW の研究によって広く認識されている [1][2]。

しかし、協調作業で交換される情報はネットワークで伝送可能な電子的な情報がすべてではない。例えば、書類や小荷物などの物理的な「物」も協調作業で交換されるべき情報であると考えられる。このように計算機による協調作業支援の分野では、取り扱う情報を物理的対象にも拡張することが望まれてきており、実験的なシステムも報告されている [3]。一方、一般家庭においても家電製品のインテリジェント化が進み、掃除ロボット [4] などのシステムが実用化されつつある。

このようにオフィスや家庭における今後の計算機システムでは、従来の計算機システムが扱っていた「電子世界」だけでなく、我々が活動する「物理世界」までもその処理の対象に含めることが重要になると考えられる。

そこで、我々は電子世界から物理世界への橋渡し役としてパーソナルロボットに注目している。パーソナルロボットとは、現在のパーソナルコンピュータと同様に、個人で所有し簡単に

機能の追加・変更やプログラミングなどが可能な、パーソナルユースのロボットのことである。現状では、オフィスや家庭内で簡単な仕事や趣味に使用できる小型の自律移動ロボットを想定している。用途としては、オフィスワーク支援ロボット、家事支援ロボット、アミューズメントロボット、福祉ロボットなどを考えている。

パーソナルロボットは電子世界と物理世界の両方を取り扱うことができる。つまり、コンピュータとして計算や通信などを行うことも可能であり、一方で物理的な物を運んだりすることも可能である。我々は、パーソナルロボットのこの能力を用いて、計算機科学が扱う対象を物理世界へ拡張することを試みている。現在、人間・ロボット・計算機が混在する環境で、従来の電子世界から我々の活動する物理世界までを取り扱う研究を進めている [5]。

そして近い将来、パーソナルロボットが人間社会に融けこんでいくことを想定している。例えば、オフィスにパーソナルロボットが数台いて、ルーチンワークなどの仕事は人間がパーソナルロボットに頼むという社会、さらにはもう一步進んでパーソナルロボットのほうから人間に仕事を依頼し、人間と協調作業を行うというような社会が来ると考えている。

このような状況でパーソナルロボットと人間がインタラクションを行う場合、従来のマン・マシン・インタフェースでは、パーソナルロボットに対してユーザがすべての情報を明示的に入力しなければならないので、ユーザの負担が大きくなってしまい、快適にパーソナルロボットを使用することが困難である。そこで本論文では、パーソナルロボットと人間ができるだけ自然に接することを可能とし、ユーザの負担を軽減することので

原稿受付 1995年4月24日

*慶應義塾大学大学院理工学研究科計算機科学専攻

*Department of Computer Science, Keio University

きるインタフェースとして「アクティブインタフェース」を提案する。

ユーザからの明示的な入力だけでなく、ユーザが無意識のうちに出す情報や外界の情報を積極的に収集・分析することにより、パーソナルロボットが自発的に行動を起こし、より有利な状況でユーザに接することを可能にする。これが本論文で提案するアクティブインタフェースの概念である。

アクティブインタフェースの有効性を検証するために、パーソナルロボット *ASPIRE-I* 上にアクティブインタフェースを応用した音声対話システム *Chaser* を設計・実装し評価を行う。

2. アクティブインタフェースの提案

2.1 アクティブインタフェースの定義

ユーザからの明示的 (explicit) 入力を持つだけでなく、ユーザの無意識 (implicit) の入力や外界 (external) の情報を積極的に取り入れ、それらの情報を基に自発的に行動を起こし、より有利な状況でユーザに接するインタフェースを「アクティブインタフェース (Active Interface)」と定義する (Fig.1 参照)。

アクティブインタフェースでは、入出力を以下のように分類する。

| | |
|--------|----------------------|
| 明示的入力 | キーボード入力, 音声入力など |
| 無意識の入力 | 表情 [6], 体温, 音量, 抑揚など |
| 外界の情報 | 位置, 騒音, 気温など |
| 出力 | 動作, 画面表示, 発話など |

上記の分類は、インタフェースの設計段階における設計者から見た解釈であり、インタフェースの設計段階で、どのような入力をユーザに意識させ、またどのような入力をユーザに意識させないかを明確にするために必要となる。

従来型インタフェース (Fig.2 参照) とアクティブインタフェース (Fig.1 参照) を比較すると、従来型インタフェースでは明示的の入力しか取り扱わない。このため、ユーザが無意識に出す情報は情報量があるにもかかわらず喪失してしまっていた。また、外界の情報は従来のマン-マシン・インタフェース (特にヒューマン-コンピュータ・インタフェース) ではほとんど取り扱われていなかった。アクティブインタフェースでは、これらの従来までは捨ててしまっていた情報に注目し積極的に利用することにより、より有利な状況でユーザと接することを可能にしようとする。

2.2 アクティブインタフェースの設計方針

2.2.1 環境変化の取得・対応

パーソナルロボットは計算機が対象としている「電子世界」だけではなく、我々が活動する「物理世界」も対象にする必要がある。一般に物理世界は動的に変化する (人の位置, 物の配置, 騒音レベルなど), パーソナルロボットのインタフェースを設計する際には、環境変化に対応することを考慮しなければならない。そのためには、環境変化をセンサによって積極的に取得し、その情報をインタフェースに反映させればよい。つまり、環境変化にユーザが対応するのではなく、インタフェースが対応することでユーザの負担を軽減できると考える。

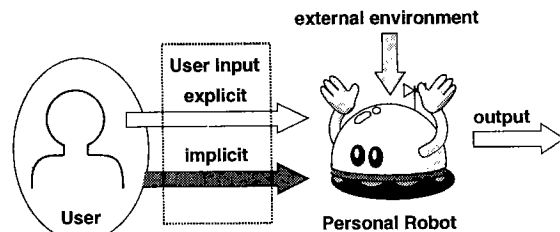


Fig. 1 A conceptual model of Active Interface

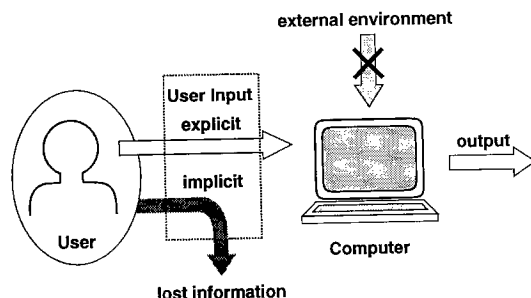


Fig. 2 A conceptual model of a conventional interface

2.2.2 パーソナルロボット固有の能力の活用

パーソナルロボットには物理世界に能動的に働きかける能力 (物を動かす能力, 移動能力など) がある。この能力を活かすことによって、パーソナルロボットがワークステーションなどの基地局を介すことなくユーザと直接インタラクションを行う場合に、より効率良く情報の授受をすることが可能になると考えられる。

例えば、パーソナルロボットには移動能力があるので、ロボット自身が情報を獲得しやすい場所まで移動することが可能である。アームを搭載したロボットであれば、ユーザに物理的接触で働きかけ注意を引き付けてから、何か重要な情報を伝えるようなインタフェースが考えられる。このようなアプローチは、データ入力の解析結果やタスクの目標に応じてデータ取得プロセスを制御し方策を変更するというアクティブセンシングの概念 [7] をパーソナルロボットのインタフェースに拡張したものとも考えることもできる。

2.2.3 無意識の入力と外界の情報の活用

パーソナルロボットには物理世界で活動するために多種多様なセンサが搭載され得る。センサはパーソナルロボットがユーザからの入力を受け取るためのデバイスでもあるが、パーソナルロボットに搭載されているセンサの機能のすべてをユーザが理解し活用することは困難である。したがって、ユーザに対してセンサをどのように見せる (意識させる) かが、パーソナルロボットのインタフェースを設計する上で重要な問題となる。

本研究では、可能な限りユーザにはセンサを意識させない方針を取る。まず、パーソナルロボットに対してユーザが明示的に入力すべきものを減少させる。その代わりとしてユーザの無意識の入力や外界の情報を取り込み、それら情報を積極的に利用する機構を設ける。この機構により、パーソナルロボットとユーザが直接向き合っているインタラクションする場合でも、ユー

ザがすべての入力を明示的に行う必要がなくなりユーザの負担を軽減できると考える。

以上をパーソナルロボット上の音声認識システムを例にして説明する。パーソナルロボットの取り扱うことのできるセンサ入力として、音声（言葉）、音量、抑揚、音源方向、熱源方向があるとするとする。

最初に、ユーザに対して「このパーソナルロボットは、あなたの命令（言葉）を理解します」と紹介する場合の設計とユーザの負担について考える。この場合、入力は次のように分類できる。

- (1) 明示的入力：音声（言葉）
- (2) 無意識の入力：音量、抑揚
- (3) 外界の情報：音源方向、熱源方向

ここで、ユーザにとっての入力手段は(1)だけなので、ユーザは音声（言葉）による入力に対してだけ注意を払えばよいと感じる。（ロボットは(1)だけでなく、(2)や(3)についても情報を読み取り考慮しているが、ユーザはそれらの情報を意識していないのでユーザがそれらの入力を負担であると感じることはない。）

次に、ユーザに対して「このパーソナルロボットは、あなたの命令（言葉）以外に、あなたの体温、声のする方向、声の大きさ、抑揚を理解します」と紹介する場合の設計とユーザの負担について考える。この場合、入力は次のように分類できる。

- (1) 明示的入力：音声（言葉）、音量、抑揚、音源方向、熱源方向

このような紹介をされると、ユーザは音声だけでなく、その他の入力手段についても意識的に注意を払う必要があると感じる。これはユーザにとっての負担が最初の例に比べ増加していることになる。

つまり、すべての入力手段をユーザに見せることは、ユーザに対してシステムの入力デバイス（マイクの入力ゲインやセンサの種類・感度など）を常に把握・意識させることになり、物理的および精神的負担になると考えられる。

2.3 アクティブインタフェースにおける能動性

アクティブインタフェースの能動性は、ユーザに入力を促すというような能動性ではない。もしこのような能動性であったならば、逆にユーザにとって厚かましいものになってしまう。本論文では以下の2点をアクティブインタフェースの能動性とする。

- (1) 明示的入力を待つだけでなく、無意識の入力や外界の情報を獲得する能動性
- (2) より有利な状況でユーザに接するために、自発的に行動を起こす能動性

第一の能動性は、ユーザの無意識の入力や外界の情報を積極的に獲得することにある。従来の多くのインタフェースは受動的にユーザからの入力を待つだけであったが、アクティブインタフェースではユーザがすべての入力を明示的に与える必要がないため、ユーザの入力に関する負担を軽減できる。また、無意識の入力や外界の情報などを参照し併用することで、ユーザからの明示的入力をより正確に解釈できる。

第二の能動性は、無意識の入力や外界の情報を基にロボット固有の能力（物を動かす能力、移動能力など）を活用して自発的に行動を起こし、できるだけ有利な状況でユーザに接することにある。これにより、ユーザの入力が容易になり、同時にパーソナルロボットが入力を獲得しやすくなる。さらに、出力としてもより正確な情報をユーザに伝達できる。

3. アクティブインタフェースを用いた実装

3.1 アクティブインタフェースを応用した

音声対話システム：Chaser

アクティブインタフェースの概念は、人間とパーソナルロボットが直接向かい合ってインタラクションを行う場合の様々なインタフェースに適用することができると考えられる。本論文では、アクティブインタフェースを応用した一実装例として音声対話システムを取り上げる。アクティブインタフェースの有効性を検証するために、移動能力を持ったパーソナルロボット *ASPIRE-I* にアクティブインタフェースを応用した新しい音声対話システム *Chaser* を設計・実装し評価を行う。ユーザに対して「このパーソナルロボットは、あなたと音声による簡単な対話ができます。」と紹介する場合の設計と実装をアクティブインタフェースに基づいて行う。

3.2 音声対話システムにおける問題点

音声による入出力はユーザにとって親しみやすい入出力方法の一つである。しかしながら、パーソナルロボットのインタフェースとして音声対話システムを利用することを考えると、ロボット周辺の環境はユーザも含めて動的に変化するので、十分な認識率での音声認識、および適切な音量での音声合成を実現することは難しいと考えられる。

この問題の解決策として、ユーザにヘッドセットなどを利用させることが考えられる。しかし、この方法はユーザに入出力デバイスを強く意識させると共に束縛感を与えることになり、好ましくないと考えられる。

そこで、本論文ではユーザ側にヘッドセットを利用させず、入力部に指向性マイクロホン、出力部に通常のスピーカを用いる。これによってヘッドセットを利用する際に生じる煩わしさを回避することができる。しかし、インタラクションを困難にする以下のような要因が生じる。

- ユーザの発声音量の変化
- 周囲の騒音レベルの変化
- ユーザの位置変化（方向と距離）

そのため、上記の要因により以下のような問題が生じる。

音声認識 認識率の低下

音声合成 騒がしい環境ではユーザに十分な情報を伝達できない。静かな環境では大きな音量でユーザに不快感を与える。

近年、ノイズに強い音声認識システムが報告されている [8] [9] が、これらの技術に加えアクティブインタフェースを応用することで、さらに認識率の低下を防ぐことが可能であると考えられる。同時に、適切な音量でユーザに発話することを可能にする。

3.3 アクティブインタフェースに基づいた設計

アクティブインタフェースに基づいて、音声対話システム *Chaser* の設計を行う。そのために、まず、音声対話システムの入出力をアクティブインタフェースの入出力 (2.1節参照) に従い以下のように分類する。

| | |
|--------|---|
| 明示的入力 | ユーザからの音声入力 |
| 無意識の入力 | ユーザの発話音量 |
| 外界の情報 | ユーザの位置 (方向, 距離), 周囲の騒音レベル パーソナルロボットの移動 (ユーザへの接近・離脱, 追尾), 音声合成装置による発話 (音量を変化) |

ユーザは、パーソナルロボットと音声による簡単な対話ができるという説明から、明示的入力である音声入力とパーソナルロボットからの音声出力 (発話) にだけ注意を傾けている。一方、*Chaser* は、アクティブインタフェースの第一の能動性 (2.3節参照) に従い、無意識の入力であるユーザの発話音量、および外界の情報であるユーザの位置 (方向, 距離), 周囲の騒音レベルを各種センサ (3.5節参照) によって積極的に獲得する。そして、アクティブインタフェースの第二の能動性に従い、パーソナルロボットがより有利な状況でユーザに接することを試みる。この場合では、第一の能動性によって獲得された情報を基に、インタラクションに有利になるようにアクション (パーソナルロボットの移動, 発話音量の変化) を起こす。これらの能動性によって、ユーザからの明示的入力である音声入力をより正確に解釈しようとし、かつ、出力である発話をより正確にユーザに伝達しようとする。具体的には、以下ようになる。

● 騒音レベルの変化

騒音レベルが上昇すると、ユーザの位置 (方向, 距離) を基にしてパーソナルロボットがユーザに接近し、音声認識率の低下を防ぐ。同時に、音声合成による発話音量を大きくして、ユーザに聞き取りやすくする。

騒音レベルが下降すると、ユーザの位置 (方向, 距離) を基にして、音声認識率が下がらない程度の距離までパーソナルロボットがユーザから遠ざかり、ユーザに接近しすぎることによる不快感を与えないようにする。同時に、音声合成による発話音量を小さくして、ユーザに不快感を与えないようにする。

● 発話音量の変化

騒音レベルの変化の場合と同様に、ユーザの発話音量が小さくなるとユーザに接近し、ユーザの発話音量が大きくなるとユーザから遠ざかる。

● ユーザの位置変化 (方向, 距離)

音声対話中にユーザが移動すると、ユーザの位置 (方向, 距離) を基にしてパーソナルロボットがそのユーザを追尾し、ユーザと一定の距離 (ユーザが移動する直前まで保っていた距離) を維持する。

以上のようにして、アクティブインタフェースを用いることにより、ユーザに負担や不快感を与えることなく音声認識に有利な状況を作りだし、同時にユーザに適切な音量で音声情報

を伝達することのできる音声対話システムを構築する。このように、環境変化などにユーザが対応するのではなく、インタフェースが対応することで、ユーザの負担を減少させる。

3.4 パーソナルロボットアーキテクチャ: *ASPIRE*

Chaser を実装するプラットフォームとなるパーソナルロボットは、アクティブインタフェースを実装するのに十分な性能を持っていなければならないと考えられる。我々は非人工的かつ未知な環境 (雑踏など) でも反応性良く動作することのできるパーソナルロボットのアーキテクチャとして *ASPIRE* (*ASynchronous, Parallel, Interrupt-based and REsponsive architecture*) を提案している [10]。 *ASPIRE* は、パーソナルロボットにリアクティブ性 (外部からの刺激に対して反射的に動作を行う性質 [11]) とリアルタイム性を統合した性質であるレスポンス性を持たせることに重点を置いて設計されている。そしてそのプロトタイプとしてパーソナルロボット *ASPIRE-I* (Fig. 3 参照) を実装しており、音声対話システム *Chaser* は *ASPIRE-I* 上に設計・実装を行う。

ASPIRE-I は、VMEバスで結合した分散共有メモリを持つ機能別並列計算機として実装されている。頭 (メインモジュール)、足 (モータモジュール)、目 (画像処理モジュール) などに機能分散された各モジュールはそれぞれプロセッサを持っており、各モジュール内で独自の判断・処理が可能である。

ASPIRE の最大の特徴はすべての I/O およびモジュール間に徹底的に割り込み信号線を張りめぐらすことにある (Fig. 4 参照)。この割り込み信号線を用いて、必要な時にのみ即座にイベントを伝達することができ (モジュール間割り込み伝播時間 2.5 [μsec])、即時性と並列性を最大限に活用することができる。また、割り込みとイベントは一対一に対応し、通常処理と例外処理を統一して扱うことができる。そして、通常処理や例外処理の区別なく、単なるイベントとして統一的かつ反応性良く処理を行うことが可能となる。これによって *Chaser* で用いるすべてのセンサの処理をレスポンス性を行うことを可能にし、ユーザにストレスをできるだけ与えないようなインタフェース

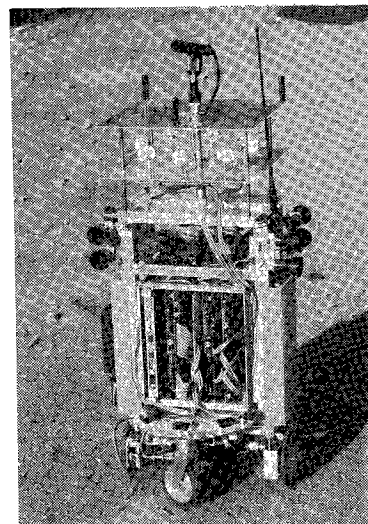


Fig. 3 Personal robot *ASPIRE-I*

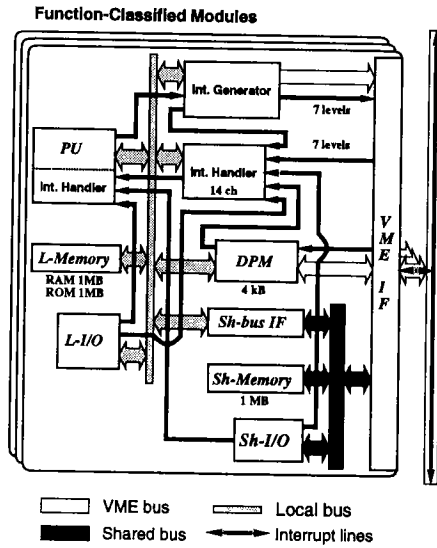


Fig. 4 Block diagram of ASPIRE-I

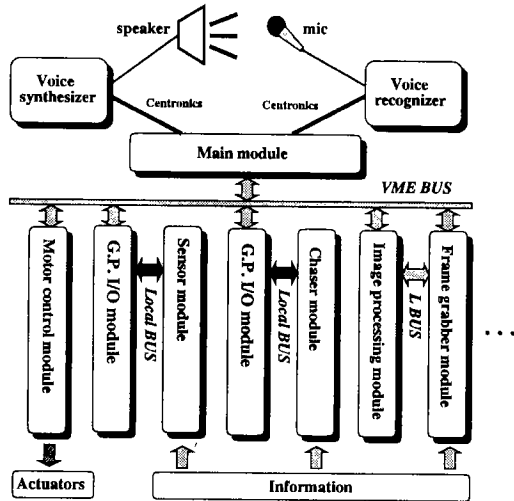


Fig. 5 Function-classified modules

の構築を可能にする。

現在、ASPIRE-Iには以下の機能別モジュールが実装されている (Fig. 5 参照)。ただし、今回実装した Chaser では画像処理モジュールは使用していない。

- メインモジュール：MC68030, MC68882, イーサネット
- モータモジュール：TMP68301, DC モータフィードバック制御 (PWM 制御)
- センサモジュール：TMP68301, 超音波距離センサ, 赤外線距離センサ, タッチセンサ, 光ファイバジャイロ
- Chaser モジュール：TMP68301, MC68882, 音源方向定位センサ, 熱源方向定位センサ (3.5.3項参照)
- 画像取込モジュール：640 × 480, 24bit color 2 planes
- 画像処理モジュール：T805 × 5
- 通信モジュール (メインモジュールに含まれる)
- 入出力モジュール：TMP68301, Local bus
- 音声認識モジュール：MN1901611BAT1

● 音声合成モジュール：VCS11A-CIF

ASPIRE-Iのこれらのモジュールには、マイクロカーネル技術を応用したマルチスレッドベースのマルチタスク・リアクティブ・リアルタイム OS μ -PULSER が載っている [12]。 μ -PULSER では、ハードウェア割り込みとソフトウェア割り込みを Direct Interrupt (DI) という機構で統合している。ASPIRE と μ -PULSER によって ASPIRE-I のレスポンスな動作が可能となり、ユーザに反応性の悪さに起因するストレスを与えることのないインタフェースを構築することが可能となる。

3.5 Chaser のセンサ

Chaser では、ASPIRE-I がユーザに正確に接近したり遠ざかったりする必要がある。これを実現するためには、あらかじめインタラクションの対象となるユーザの位置 (方向, 距離) を定位する必要がある。しかしながら、ASPIRE-I では距離を測定するためのセンサは標準装備しているが、ユーザの方向を測定するための特殊なセンサは装備していない。そこでユーザの方向を定位するためのセンサを ASPIRE-I 上に新たに実装する。

3.5.1 音源方向定位センサ

ユーザが発する音を用いてユーザのいる方向を定位するためのセンサとして音源方向定位センサを実装する。音源方向定位センサは、一辺が 30 [cm] の正三角形の頂点に配置された三つのマイクロホンに到達する音の時間差から三角測量を行い音源の方向を算出する。

この手法は音源定位の方向定位として知られている手法であり、より高度な定位を実現したシステムも報告されている [13]。実際のサンプリングは 1° 単位で Fig. 6 のようなデータが得られる。

3.5.2 熱源方向定位センサ

ユーザが発する熱を用いてユーザのいる方向を定位するためのセンサとして熱源方向定位センサを実装する。人体が発する熱をセンシングするセンサとして、焦電型の赤外線熱源センサを 8 個 (8 方向) 用いている。この赤外線センサは微分型のセンサであり、移動する熱源に対しては速やかに反応し、静止した熱源には反応しにくくなっている。この熱源方向定位センサからは 45° 単位で Fig. 7 のようなデータが得られる。

3.5.3 Chaser モジュール

上記のセンサを制御するモジュールとして、ASPIRE-I 上に新たに Chaser モジュールを設計・実装する。Chaser モジュールには、前述した音源方向定位センサと熱源方向定位センサ (8 方向) のインタフェースおよび浮動小数点コプロセッサ (MC68882) を実装する。各センサに ASPIRE に従って張られている割り込み信号線とコプロセッサによって、レスポンスにユーザの方向定位をすることを可能にする。

3.5.4 Chaser で用いるセンサ系

すでに ASPIRE-I に実装されていたものも含めて、音声対話システム Chaser で使用するセンサ系は以下ようになる。

- 音声認識装置 (MN1901611BAT1)
- 音声合成装置 (VCS11A-CIF)
- 騒音レベル監視モニタ [14]
- 音源方向定位センサ

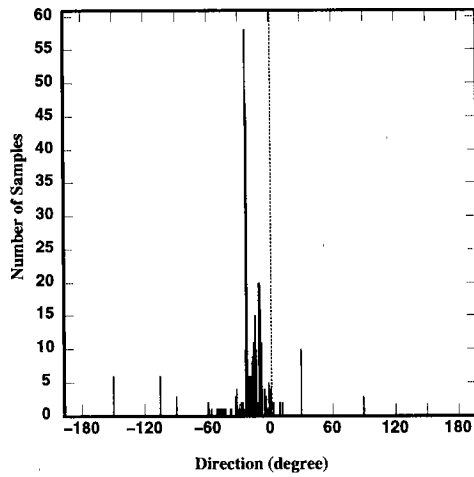


Fig. 6 Information of sound source sensor

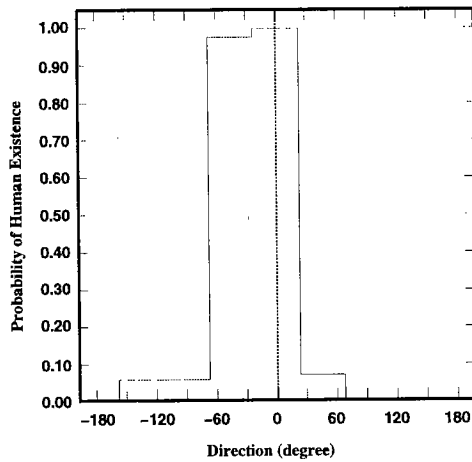


Fig. 7 Information of heat source sensor

- 8 方向赤外線熱源方向定位センサ
- 4 方向超音波距離センサ (ワイドレンジ, 中～長距離)
- 8 方向赤外線距離センサ (ナローレンジ, 短～中距離)
- 8 方向タッチセンサ
- 光ファイバジャイロスコープ (HOFG-3)
- 左右輪二つの距離計

3.6 Chaserにおける出力

音源方向定位センサから音源の方向を求めることができるが、壁からの反射音などを拾ってしまう可能性がある。また、熱源センサにより人体の温度を関知してその方向が分かるが、人間以外の熱源 (コンピュータなど) に反応する可能性がある。ユーザのいる方向は音源方向と熱源方向に関する二種類の情報を融合することにより、より高い精度で決定できると考える。Fig. 8 は、前述の音源センサの情報 (Fig. 6) と熱源センサの情報 (Fig. 7) を掛け合わせたデータである。Chaser モジュールは、この掛け合わせたデータからユーザのいる方向を求める。

周囲の騒音レベルが上昇すると、騒音レベル監視モータはモータモジュールに対してユーザに接近するようにイベントを起こす。(同様に、ユーザの発話音量が下降した場合は音声

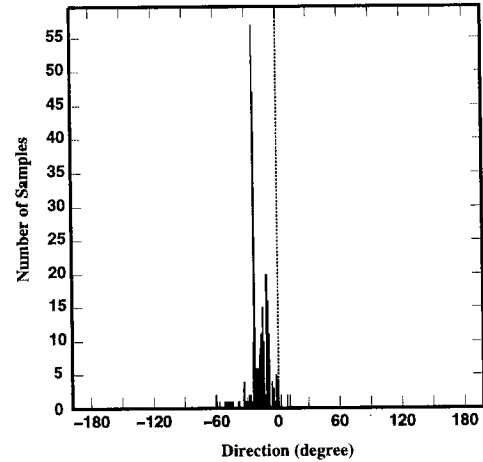


Fig. 8 Combined information

認識モジュールが、ユーザの位置が遠ざかる方向に変化した場合はセンサモジュールがユーザに接近するようにイベントを起こす。) モータモジュールは、Chaserモジュール、センサモジュールと協調しながらユーザに接近を行う。Chaserモジュールはユーザのいる方向を定位し、センサモジュールは超音波距離センサと赤外線距離センサを用いてユーザとの距離を測定する。センサモジュールは、ワイドレンジ (約 30°) で中～長距離 (約 50 [cm]～5 [m]) を測定する超音波距離センサとナローレンジ (約 5°) で短～中距離 (約 5 [cm]～1 [m]) を測定する赤外線距離センサという 2 系統の距離センサを制御している。超音波距離センサは大略的に障害物 (対象物) を発見するのに使用し、赤外線距離センサは発見した障害物 (対象物) との距離を正確に測定するのに使用する。モータモジュールは、Chaserモジュールで求めた方向にセンサモジュールで測定した距離を基にしてパーソナルロボットを前進させる。モータモジュールはモータに直結した距離計を基に、センサモジュールと協調しながらユーザに接近する。ここで、この距離計の値だけを元に移動すると時間と共にタイヤのスリップなどによる誤差が積算するので、センサモジュールで制御している光ファイバジャイロスコープを用いて、一定期間 (100 [msec]) ごとにグローバルな座標値 (マップ) の補正を行う。

また、周囲の騒音レベルが下降すると、ASPIRE-I は同様にユーザから遠ざかる。

4. Chaserの評価

アクティブインタフェースを応用することによってユーザにより有利な状況を生み出すことが可能かどうかの検証をChaserを用いた実験により行う。

ここでは、以下の二つの項目について実験・評価を行う。一つはインタラクションの対象となるユーザの方向を定位しそのユーザに接近する実験、もう一つは騒音レベルに応じてASPIRE-Iがユーザに接近・離脱した場合の音声認識率に関する実験である。

4.1 ユーザの方向定位に関する実験

ユーザの方向定位に関する実験では、*Chaser*に搭載された音源方向定位センサおよび熱源方向定位センサを、それぞれ単独で利用した場合と両方のセンサを掛け合わせて利用した場合について、正しく方向定位ができユーザに接近を行うことができるかどうかの検証の実験を行う。

実験方法

音源および熱源を四つの Case に分けてパーソナルロボットの周囲に配置し、パーソナルロボットが正しくユーザに接近できるかどうかの評価を行う。

Case 1: 単一音源, 単一熱源

ロボットの周囲に人間が1人だけ存在し、その人が音を発生させる (音源1, 熱源1)。

Case 2: 単一音源, 複数熱源

ロボットの周囲に2人の人間が存在し、どちらか一方の人だけが音を発生させる (音源1, 熱源2)。

Case 3: 複数音源, 単一熱源

ロボットの周囲に人間1人とテープレコーダを1台配置し、音は人間とテープレコーダの両方から発生させる (音源2, 熱源1)。

Case 4: 複数音源, 複数熱源

ロボットの周囲に2人の人間が存在し、両方の人が同時に音を発生させる (音源2, 熱源2)。

この実験では、パーソナルロボットが移動した方向に対して、ずれ角 $\pm 15^\circ$ の範囲内にユーザがいる場合に正しくユーザに接近を行ったとする (サンプリング数: 各30回)。

4.2 認識率に関する実験

*Chaser*では、音源方向定位センサと熱源方向定位センサの情報からユーザの方向を測定し、*ASPIRE-I*が測定した方向に向かってユーザに接近することで認識率の低下を防ぐ。この機能が有効に働くかどうかを検証するために以下の実験を行う。

実験方法

実験で使用する騒音のサンプルは、FM ラジオの局間ホワイトノイズを録音し再生を行う。騒音レベルは収録音を再生する際の音量によって変化させ、JIS C-1502 の A 特性に準拠した騒音計を用いて測定する。約 -50 [dB]~ 0 [dB] までの範囲を用いる。

認識に用いた単語候補数は11個で、音声入力には辞書中から10個の単語を順番に選択・発声し、これを三人の声で録音したものを使用する (合計30回)。

実験1

比較的静かな環境 (-70 [dB]) でパーソナルロボットを音源に向かって移動させ、マイク・音源間距離と認識率との関係について実験を行う。

実験2

騒音レベルとマイク・音源間距離の両方を変化させ、認識率をある一定値にするように騒音レベルとマイク・音源間距離の相関を調べる実験を行う。

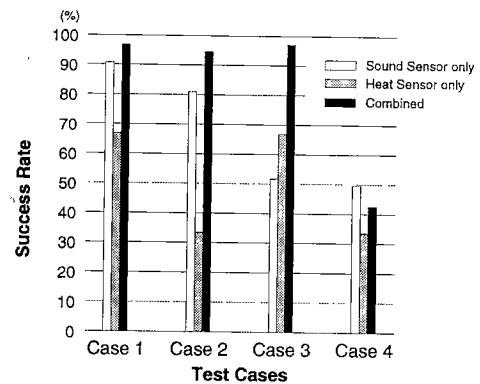


Fig.9 Success rate of approaching the user

4.3 結果および考察

ユーザの方向定位に関する実験

ユーザの方向定位に関する実験の結果は、Fig.9 のようになった。

音源センサのみによる方向定位は、Case 1やCase 2のように単一音源の場合に有効であり、平均で85%程度の成功率でユーザに接近することができた。しかし、Case 3やCase 4のように音源の数が複数になった場合、平均で50%まで成功率が落ちてしまった。

熱源センサのみによる方向定位は、Case 1およびCase 3のように単一熱源の場合に有効であるが、一つのセンサからは 45° の範囲までしか絞れないため、平均で67%程度の成功率にとどまっている。これがCase 2やCase 4のように複数の熱源になると、33%まで落ちてしまう。このことから、熱源センサの情報だけを頼りにユーザの方向を定位することは、あまり有効でないことが分かる。

音源センサの情報と熱源センサの情報の両者を掛け合わせた場合は、Case 4を除くすべての場合に有効であった。Case 1~3までは平均で95%の成功率で定位および接近を実現することができた。Fig.9を見ると、音源センサ、熱源センサをそれぞれ単独で用いるよりも、両者を掛け合わせて利用する方が、常に高い性能を示している。特に複数音源、単一熱源の場合 (Case 3) では、音源センサのみによる定位が52%、熱源センサのみによる定位が67%と低い値しか得られていないのに対して、両者を掛け合わせた方法では96%と高い値を得ている。このように、熱源センサで大略的な範囲を限定し、その範囲内で音源センサによって角度を絞る方法は有効であることが分かる。

しかしながら、ロボットの周囲に複数のユーザが存在するCase 4の場合では、どの方法を用いても正確な定位は実現できなかった。現在のセンサだけでは、このような状況から対象となるユーザの方向を定位するには限界があることが分かった。

認識率に関する実験1

認識率に関する実験1の結果は、Fig.10 のようになった。この結果をみると、マイクと音源との距離が20~40 [cm] で認識率に急激な変化があることが分かる。40 [cm] より近くなるにつれて急に認識率が上昇している。これは、使用した音声認識装置の雑音と有意な音を区別するためのフィルタの音圧レベル

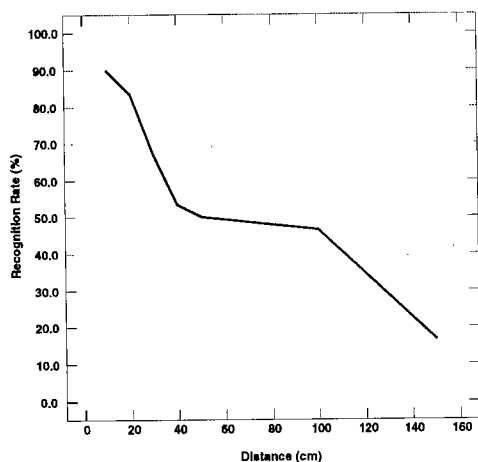


Fig. 10 Recognition rate (Constant noise)

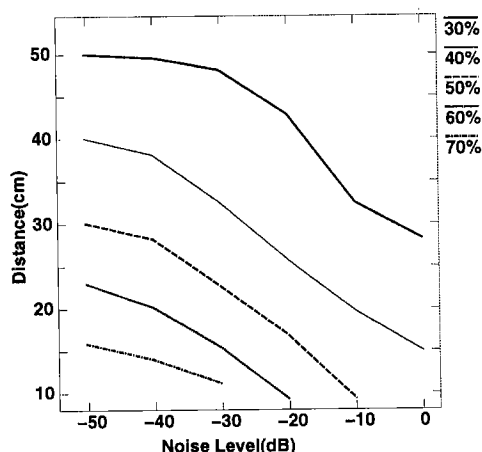


Fig. 11 Recognition rate

の閾値が、マイクまでの距離が40[cm]であるときの音圧レベルの近傍に相当していたためであると考えられる。この *Chaser* による実験結果は、アクティブインタフェースを用いてユーザに近づくことが認識率低下抑制に有効であることを示している。

認識率に関する実験 2

認識率に関する実験 2 の結果は、Fig. 11 のようになった。このグラフより、例えば騒音レベル -40 [dB] で認識率 50% (太破線) を維持するには、*ASPIRE-I* はユーザに約 28 [cm] の距離まで近付ければよいことが分かる。この状態から、もし騒音レベルが -20 [dB] に上昇した場合、*ASPIRE-I* はユーザに約 17 [cm] まで近付くことにより、認識率を維持する。反対に、その状態から騒音レベルが -50 [dB] まで下降した場合、*ASPIRE-I* はユーザから約 31 [cm] の距離まで離れることによって、接近しすぎることによってユーザに与える可能性のある不快感を減少させる。このように、アクティブインタフェースを用いることによって、周囲の騒音レベルが変化した場合でも *Chaser* では自発的にユーザに接近・離脱して、音声認識率を一定に維持することができる。

以上の実験結果は、*Chaser* で提供したユーザの方向定位機能およびユーザへの接近・離脱機能が音声対話において有効であ

ることを示唆するものであると考えられる。

5. おわりに

5.1 結論

本論文では、ユーザがパーソナルロボットと直接向かい合っ
てインタラクションを行う状況において、ユーザにとって負担
が少なく快適なインタフェースについて検討を行い、アクティ
ブインタフェースという新しいインタフェースの概念を提案
した。

そして、アクティブインタフェースの有効性を検討するた
めに、その一実装例として、パーソナルロボットのプロトタイプ
である *ASPIRE-I* にアクティブインタフェースを応用した音声
対話システム *Chaser* を設計・実装した。*Chaser* ではユーザの
音声入力 (明示的入力) だけでなく、音源方向や熱源方向など
のユーザの意識していない情報にも着目した。また、各種セン
サ系を有効に用いて *ASPIRE-I* が自発的にユーザに接近・離脱
することにより音声認識率を一定値に維持することができた。
以上の実験結果より本論文で提案したアクティブインタフェ
ースという概念の有効性の第一段階の検証ができたと考えられ
る。また、本研究では、アクティブインタフェースという新し
い概念の一実装例として音声対話システムを取り上げたが、そ
れ以外のインタフェースにおいても、従来の受動的なインタ
フェースと比較して、より確実に快適なインタフェースを実現
できると考えられる。

5.2 今後の課題

本論文で実装した音声対話システム *Chaser* は、実用性を考
慮した場合その完成度が十分であるとは言い難い。特にユー
ザの方向定位機能に関しては、複数のユーザがロボットの周
囲にいる場合 (複数音源、複数熱源の場合) に十分な性能を
得ることができない。これは音源方向定位センサと熱源方向
定位センサの情報だけでユーザの方向定位を行うことの限界
を示している。今後はこれらの問題に対処するため、画像処
理などを取り入れたシステムへの改良を検討する必要がある
と考えられる。

また、アクティブインタフェースの応用例として、音声対
話システム以外の様々なシステムの実装を行いたいと考えて
いる。

さらに、アクティブインタフェースがもたらすユーザへの
心理的および人間工学的実験と考察を行いたいと考えてい
る。

参考文献

- [1] 石井裕: “コンピュータを用いたグループワーク支援の研究動向”, コンピュータソフトウェア, 日本ソフトウェア科学会, vol.8, no.2, pp.14-26, 1991.
- [2] 阪田史郎: “CSCW におけるマルチメディア技術”, 計測自動制御学会誌, vol.30, no.6, pp.497-504, 1991.
- [3] 山本吉伸, 佐藤充, 開一夫, 山崎信行, 安西祐一郎: “人間とロボットの協調作業は可能か? —オフィスロボットシステム HuRIS—”, 情報 HI 研究会, pp.117-124, vol.92, no.15, HI-41-16, March, 1992.
- [4] 小林保道, 藪内秀隆, 近藤信二, 江口修: “掃除ロボットのインテリジェント化技術”, システム制御情報学会誌, vol.35, no.8, pp.473-480, 1991.
- [5] Yuichiro Anzai: “Towards a New Paradigm of Human-Robot-Computer Interaction,” Proceedings of IEEE International Workshop on Robot and Human Communication, pp.11-17, 1992.
- [6] H.Kobayashi and F.Hara: “Recognition of Six Basic Facial Expressions and Their Strength by Neural Network,” Proceedings

of IEEE International Workshop on Robot and Human Communication, pp.381-386, 1992.

- [7] R. Bajcsy: "Active Perception", Proceedings of IEEE, vol.76, no.8, pp.996-1005, 1988.
- [8] 平岡省二, 二矢田勝行, 木村達也: "ワードスポッティング手法を用いた不特定話者・小教語向け音声認識装置", 信学技法, SP88-18, 1988.
- [9] 竹林洋一, 金澤博史: "ワードスポッティングによる音声認識における雑音免疫学習", 電気通信学会論文集, J74-D-II, no.2, pp.121-129, 1991.
- [10] 山崎信行, 安西祐一郎: "パーソナルロボット用機能別並列計算機アーキテクチャ: *ASPIRE*", 情報処理学会論文誌, vol.37, no.1, pp.81-91, 1996.
- [11] A. Benveniste, P.L. Guernic, Y. Sorel and M. Sorine: "A Denotational Theory of Synchronous Reactive Systems", Information and Computation, vol.99, pp.192-230, 1992.
- [12] 矢向高弘, 菅原智義, 安西祐一郎: " μ -PULSER: パーソナルロボットを構築するためのオペレーティングシステム", 電子情報通信学会論文誌, J77-D-I, pp.207-214, 1994.
- [13] J. Huang, 大西昇, 杉江昇: "生体に示唆を得た音源定位システム - 反響のある環境での単一音源定位 -", 電気情報通信学会論文集 (A), J71-A, no.10, pp.1780-1789, 1988.
- [14] 日本工業規格: "普通騒音計", 日本規格協会, JIS C-1502, 1992.



山崎信行 (Nobuyuki Yamasaki)

1966年5月1日生。1991年慶應義塾大学理工学部物理学卒業, 1993年同大学大学院理工学研究科計算機科学専攻修士課程修了。同年同専攻博士課程入学。現在, 同専攻博士課程在学中。自律移動ロボット, 並列計算機, OS, コンピュータアーキテクチャなどに興味を持つ。(日本ロボット学会学生会員)



安西祐一郎 (Yuichiro Anzai)

1946年8月29日生。1974年慶應義塾大学大学院博士課程修了。1988年より慶應義塾大学理工学部教授。1989年より同大学大学院計算機科学専攻教授兼任。この間, 1981~82年カーネギーメロン大学客員助教授。計算機科学, 認知の情報処理過程の研究に従事。工学博士。(日本ロボット学会正会員)