

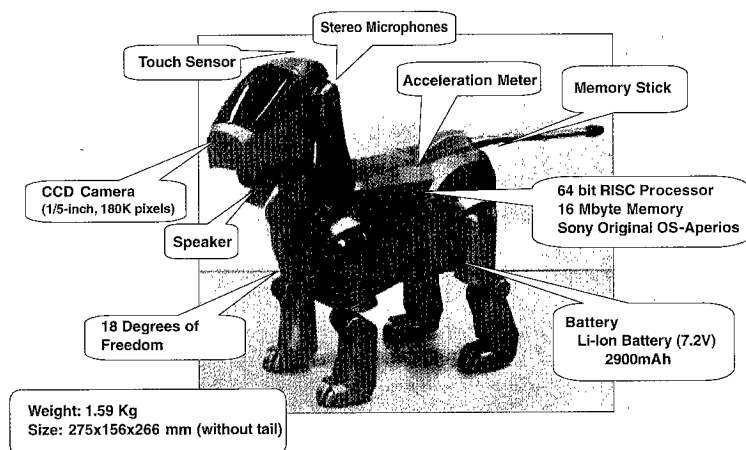
5

世紀は自律型ロボットの時代

Robot Entertainment System AIBOの開発

藤田 雅博

ソニー (株)
デジタルクリーチャーズラボラトリー



©ソニー

20世紀の終盤を迎えて、コンピュータは我々の生活に深く根付いた感がある。たとえば、さまざまな家庭電気製品に組み込まれたコンピュータは、人に代わって細かな制御や調節をしてくれる。また、銀行のトランザクションのような膨大な量のデータを処理したり、巨大なデータベースを管理するコンピュータ、あるいは、個人で文書作成やゲームをするいわゆるパーソナルコンピュータやゲームマシンなど、現在、社会のありとあらゆるところにコンピュータ技術が利用されている。

そのコンピュータ技術の進歩と普及は1990年代になってネットワーク社会という新しい社会構造を創造している。情報の発信や共有の方法、あるいは、物を売る方法やお金の流れまで、ネットワークの時代になって大きく変わろうとしている。ネットワーク上に仮想的な世界を構築し、そこで商店街を開いたり、あるいはネットワーク上に話し相手やゲームの相手を探し、インタラクションを行うような世界がもう始まっている。

我々はコンピュータ技術の進歩が生み出すもう一つの可能性として自ら考えて行動する自律型ロボットがあると考えている。コンピュータの性能が急激に向上し、一部の認識や制御といったことが実時間処理で実現可能となりはじめている。このようなコンピュータ技術の発展を背景に、自律型ロボットに必要な技術が実用化され成果として見えはじめている。たとえば、ここで紹介するAIBOやホンダヒューマノイド¹⁾などがそれにあたる。コンピュータ、センサ、アクチュエータなどハード的な性能の向上と、実世界、実時間の認識、制御あるいは知識処理といっ

たソフト的技術の発展が進み、21世紀は自律型ロボットの時代になるといわれている。

自律型ロボットの従来の応用分野は主として介護ロボット、救助ロボット、危険作業ロボットなど人を助けるということを目的としたものであった。現在これらの実用化の大きな壁となっているものが安全性や信頼性である。すなわち、これらのロボットは与えられたタスクを確実に、かつ安全に達成する必要があり、現状存在する認識、制御技術よりも高度の技術が必要となる。我々の提案する Robot Entertainment という自律型ロボットの新しい応用分野においては、従来の介護、救助といったことを行うロボットほどこれらの安全性と信頼性は求められない。認識や制御技術は現在の技術レベルのものでも実用に耐えるものである。また、現在研究が活発化している学習や進化、あるいは感情を扱うような技術は Robot Entertainment という分野でも非常に重要であり、これらも実用可能なものが存在する。

また研究開発のスタイルも従来の自律型ロボットのそれと比較して特

徴のあるものになる。すなわち、従来は視覚、聴覚といった個々の認識、制御技術を深く追求する、といった研究のスタイルであったが、Robot Entertainment において重要なことはこれらの技術を用いていかに自律型ロボット全体を構成するか、ということである。限られた計算資源を用いて実世界で動作する自律型ロボットをいかに構成するか、という統合の技術が重要なものになる。

このように、適用できる技術レベルの広さと統合技術に重点を置くことによって、自律型ロボットを家庭内に普及させ人間と自律型ロボットの共存する世界を早期に実現できることが期待される。ここでは、この Robot Entertainment という世界の実現への第1歩として我々が開発してきたペット型自律ロボット、さらに最近、限定台数ではあるが初の家庭用の商品として売り出した AIBO ERS-110 について、そのインタフェースに焦点を当てて述べる。さらに、この自律型ロボットをコンピュータ社会へのインタフェースとして位置付けた場合についても簡単に述べる。

Robot Entertainment: ペット型ロボット

前述のように、Robot Entertainment とは自律型ロボットの応用をエンタテイメントにフォーカスし、個人が自律型ロボットと共存する世界を早期に築こうとするものでもある。我々はこの Robot Entertainment の可能性を検討するためにいくつかの自律型ロボットを試作してきた。その試作を通じて大きな可能性があると思われているものがペット型ロボットと呼ばれる自律型ロボットである。まず、1997年に開発発表した4脚型自律ロボットに関して述べる^{2), 3)}。

このロボットは、図-1に示すように4つの脚部、頭部、尻尾を持ち、各脚部は3自由度、頭部に3自由度、尻

実	世	界	に
近	づ	く	イ
ン	タ	フ	ェ
ー	ス	技	術

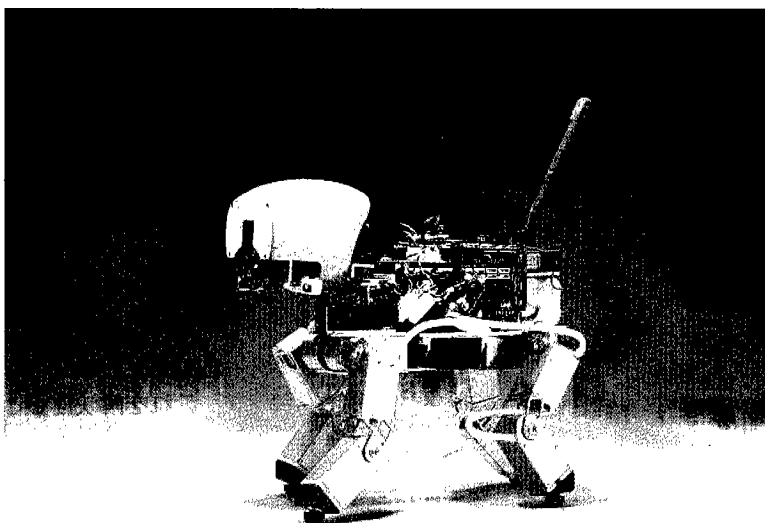


図-1 4脚ペット型ロボット (試作)

©ソニー

尾部に1自由度の合計16自由度から構成される。また、センサとして頭部に超小型カラーCCDカメラ、ステレオマイク、タッチセンサを搭載、胴体部には3軸の加速度センサを搭載している。さらに、頭部にはスピーカを搭載し音を出すことが可能である。CPUにはMIPS R3000系を画像処理用に1つ、またその他の制御、音処理用に1つ搭載している。センサを用いたヒューマンインタフェースとして以下の点に特徴を持たせた。

まず、音処理として、音声認識の導入は望まれるものであるが、あえて音声認識を導入しなかった。理由は以下のとおりである。

1. 実使用環境のもとでは、自分のモータノイズや周囲の人間の声などが妨害音となるため、認識率が問題になる。
2. 自然言語のインタフェースを人間に開放した場合、たとえば、車の運転時のように両手がふさがれ、発話によるインタフェースが唯一の方法という状況の場合は、ユーザは音声認識がしやすいような協力的な発話をする。しかし、そうでない場合、人間は自然な発話ができると思ってしまい、自然な発話に対する認識率が十分でないと、逆に“不自然さ”を

感じてしまう場合がある。現状の音声認識のレベルはまだ自然な発話を自由に許すレベルに達していない、と我々は考えた。

そこで、我々は、口笛のような特殊な信号のみをインタフェースとして用い、音階言語と呼ぶ音階によるインタラクションを試みた。口笛のように定常的な周波数成分を持つ信号は、ホワイトノイズや会話時の周波数が上下する声との分離が容易なため、実環境下での認識率を高く保つことができる。我々は、“口笛などは発話に比べて不自然なインタフェースであるが、人間は不自然であることをはじめから意識しているため認識率が保たればそれを受け入れる傾向がある”，という仮説のもとにこのような方法をあえてとっている。

さらに、ステレオマイクによる左右の位相比較により音源の方向推定をし、注目する音（定常的な音）のみの位相を比較するため、ホワイトノイズなどの影響を受けにくい音源推定が可能である。このような方法で、口笛を吹いて、ロボットの注意を引いたり、座れ、といった命令を出すことができるようにデザインされている。

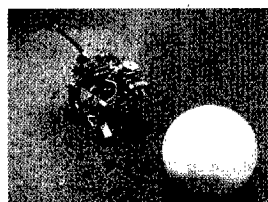
もう一つの重要なセンサ処理が画像処理である。まず、カラーカメラ自身が大きさと価格の問題を持っていた。我々は、マルチチップモジュールの技術を用いてCCD、カメラ制御、タイミングジェネレータなどをベアチップで1つのモジュールとし、さらにCCDチップが小型化されてきたことを利用して、非球面の1枚玉レンズをCCDの上に配置しモジュール化することにより超小型のカラーCCDカメラモジュール（Micro Unit CCD (MUC)）を開発した。

この試作機ではこのMUCを用いて色の検出を行い、ロボットと外界のインタラクションを行う方針をとった。色の検出は撮影した物体の各ピクセルのデータが色空間のある領域にあるかどうかを判別する方法をとっている。これにより、色のついたボールを追いかけたり、肌色の手を認識したりすることができる。

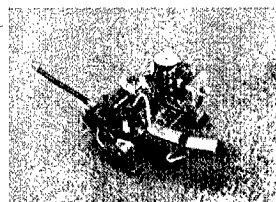
もう一つ、実世界で行動するロボットとのインタラクションで重要なものに“触る”ということがあげられる。これは、頭や足先にタッチセンサを配置し、頭を叩かれたことや、撫でられたことを認識できるようにすることで、後述の感情モデルを動かすことによって、怒ったり、喜んだりすることができる。

以上は、インタラクションの1つ1つの方法に関して述べてきたが、自律型ロボット、特にペット型ロボットにおける最大の問題は、個々のインタラクションとともに全体として、“いかに生きているかのように見せるか”，という点である。我々は、“いかに生きているかのように見せるか”，という問題を、“いかに複雑な動き、行動を見せるか”，という問題に置き換えて、これを解決するという方法をとった。それらは以下のようにまとめられる。

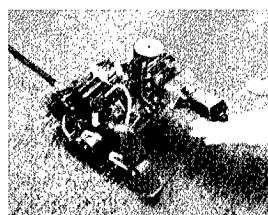
- (1) 4つの脚部と全部で16自由度という機構系による複雑度：
通常の知能ロボット研究がタイヤ



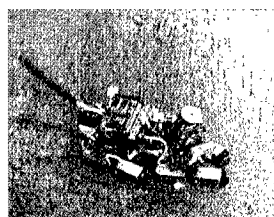
(1)



(2)



(3)



(4)

©ソニー

図-2 さまざまな行動をするペット型ロボット

型ロボットを使っているのに対して、4脚型という機構系にすることで動きに複雑さを増すことができる。4脚歩行は、3脚立脚による静的安定性を持つ歩行形態があるため2脚などに比べた場合、実現が簡単であると同時に、すわった姿勢で両前腕が地面に拘束されずに自由に動かすことができるため、動きでさまざまな表現ができるという特徴がある。

(2) 外見：

本物の犬や猫のような外見が当然選択肢に上ったが、本物と比較されることにより動きのぎこちなさが強調される、という点を考慮し、あえて縫いぐるみのようなものをかぶせずに、メカニカルなデザインを強調した。この考えは、前述の音声認識を採用しなかった理由と同じコンセプトと考えることができる。本物の犬や猫のような毛を持った外見では、もっと自然な動きや反応が要求される。しかし、現状の実現レベルがそこまで達していない、という理由である。

(3) 反射と熟考行動：

行動に複雑度を持つ方法として行動を起こすためのMotivationをいくつも持ち、それらを競合、協調させるという方法をとった。その1つの考え方に反射的行動と熟考的行動の競合がある。これは、外界からの刺激に対して時間軸上で速い行動と遅い行動と考えることができる。

(4) 本能、情動と理性：

もう1つの考えとして本能的な行動、情動的な行動、そして理性的な行動の競合という考え方ができる。そのために本能、感情モデルを導入し、前述した音の命令による行動と競合させている。

(5) 各部位の独立性：

最後に頭部、脚部、尻尾部に独立の動きのMotivationを持たせるという方法をとった。これは、たとえば頭部はボールを追いかけて、脚部はそれ

とは別の方向に歩行のための動きをし、尻尾部はそれとは独立に感情にあわせて上下動をする、といったことである。これにより、ロボットの動きは組合せ的な数で増やすことができる。

以上のようなセンサ処理や複雑な動き、行動を作り出す方法により、お手をしたり、伏せをしたり、あるいはボールを好きなものとして追いかけたり、頭を叩くと怒ったり、また撫でれば喜んだり、といったインタラクションを持つペット型ロボットを実現している。図-2は、上記のような方法によって、さまざまな反応や行動をするペット型ロボットの動き、行動例を示している。

のペット型ロボットと同じである。すなわち、人間は、音、画像、触る、といった方法でロボットとインタラクションを行う。ERS-110で実装された認識、制御技術も前述のものと同様であり、“いかに複雑な動き、行動を見せるか”という問題への置き換えも同じ戦略を用いて解決している。ただし、自由度に関しては口と尻尾にもう1自由度を増やし、合計18自由度でさらに表現力を増やすようにした。

さらに、AIBO ERS-110は学習や成長といった機能が追加されている。これは、長時間の観察において複雑な動き、行動をいかに増やすか、という問題に対する解である。

A IBOのインタフェース

我々は、1999年5月にEntertainment Robot: AIBO ERS-110の商品化を発表した。日米で5000体の限定数の販売であったが、日本は20分、アメリカは4日で限定数の申し込みをすべて完了した。

AIBO ERS-110の基本的なデザインコンセプトとインタフェースは前述

実	世	界	に
近	づ	く	イ
ン	タ	フ	エ
ー	ス	技	術

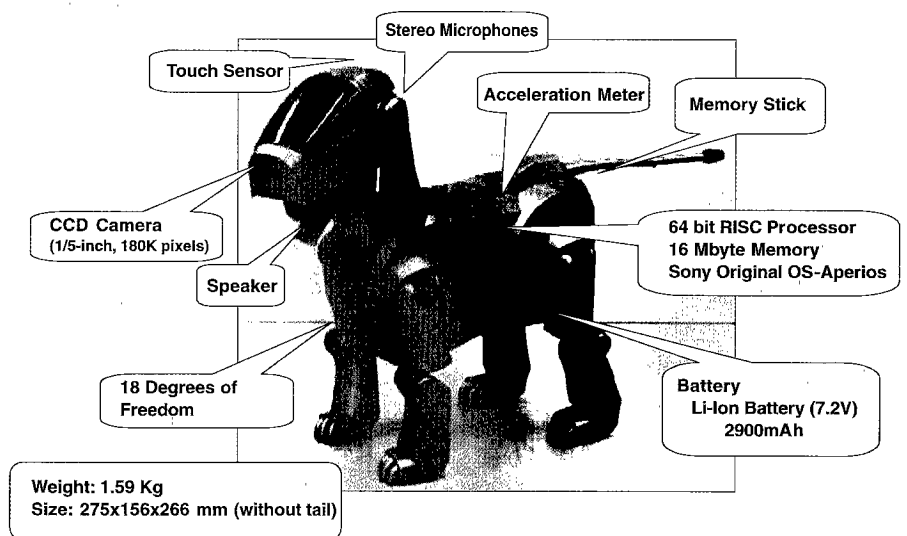


図-3 AIBO ERS-110

©ソニー

インタフェースとしての自律型ロボット

以上は自律型ロボット、特にペット型ロボットに対するインタフェースという切り口で説明をしてきた。この章では、AIBOのような自律型ロボットをコンピュータシステムのインタフェースとして考えるとどのような展開があるか、ということに関して述べる。なお、ここで述べることはRobot Entertainmentという応用分野に限ったものではない。しかし、ここでも安全性や信頼性などの問題からやはりRobot Entertainmentという分野が有望であることは同じである。

自律型ロボットはそのものがコンピュータシステムである。すなわち、キーボードやマウスの代わりに音や画像入力を用いてインタフェースされるコンピュータと考えることができる。さらに、自律型ロボットがネットワーク技術により通常のコンピュータと結合されると、自律型ロボットがそのコンピュータへのインタフェースの役割を果たす、と考えることができる。

たとえば、自律型ロボットに人間と同じような会話能力やジェスチャーを使う能力、さらに人間の感情を読んだり、自律型ロボットが感情を表現したり、という能力が備わることにより、ユーザは人間と接するような感覚でコンピュータとインタラクションすることになる。これは、ユーザインタフェースにおける1つの究極の形であると考えられる。前述のペット型ロボットでとっている戦略は実はこの方法に対する現状の技術の未成熟さを考慮してのアンチテーゼ的方法であった。しかし、Robot Entertainmentというビジネス領域が確立されれば、研究開発はこのような自然なインタフェースを目指して加速されるであろう。

ネットワーク化された自律型ロボットはさらに新しい付加価値を生むと思われる。まず単純に計算資源の問題が軽減される。反射行動のような時間的制約のある行動はロボット本体が持つ計算資源で処理されるであろうが、反応時間が遅くてよいものはネットワーク上の計算資源を用いて処理が可能である⁴⁾。さらに、前章までの説明は自律型ロボットに直面してのインタフェースであったが、ネットワーク化により、自律型ロボットの内部に対するインタフェースも可能になる。アールキューブ⁵⁾は、遠隔操作により遠く離れたロボットを動かし、人間の判断でロボットに仕事をさせる方法を提案しているが、これは自律型ロボットの内面に対するインタフェースの1つと考えることができる。なお、ここで人間が遠隔操作するものも自律型ロボットと呼んでいる。人間が遠隔操作するロボットだとしても、自由度の多さ、反射行動などに必要とされる速いフィードバックなどを考えると、なんらかの自律性が操作されるロボットに備わっていなければならないと考えているからである。

自律型ロボットがネットワーク化された場合の最大の特徴は、それが実世界で物理的な仕事を行えるコンピュータである、ということであろう。すなわち、現在の仮想世界でのインタラクションから実世界をも含めた物理的なインタラクションがネットワーク化された自律型ロボットにより実現される。たとえば、現在仮想世界でアバタと呼ばれるエージェントが実世界に出てきていることを想像すればよい。これは、人間が住む現実世界に物理的にインタラクションできる物理的実体を持つエージェントである。これは、現在のネットワークが社会の質的転換を起こしているのと同様に、さらなる社会の質的転換を引き起こすものと思われる。

まとめ

以上、Robot Entertainmentという自律型ロボットの新しい応用分野とその具体的な自律型ロボットとして開発されたAIBOの主な特徴とその可能性に関して述べた。特に、人間がどのようにAIBOとインタラクションするか、そのインタフェースは何か、を簡単に説明した。さらに、それを自律型ロボットとしてではなく、コンピュータシステムのインタフェースとしてみたとき、あるいはネットワークシステムのインタフェースとしてみたときの著者の雑感を述べた。

AIBO ERS-110はRobot Entertainmentというビジネス分野あるいは研究開発分野を切り開く第1歩である。今後、どのようにこの分野を確立していくかが重要である。ここで述べたように、エンタテインメント応用は現状の技術レベルで実現可能なものがある反面、人を楽しませる、という、今までの自律型ロボットの研究では問題にならなかった点も重要になる。また、個人の好みが多様であるように自律型ロボットの形状や応用ソフトも多様化しなければならないであろう。

これらの技術は自律型ロボットに限らず、ネットワーク社会におけるインタフェース技術やエージェント技術と相互作用しあって発展するものと予想する。AIBOがこのような世界の実現を加速できることを期待する。

参考文献

- 1) 広瀬, 竹中, 五味, 小澤: 人間型ロボット, 日本ロボット学会誌, Vol.15, No.7, pp.983-985 (1997).
- 2) Fujita, M. and Kageyama, K.: An Open Architecture for Robot Entertainment, International Conference on Autonomous Agents, pp.234-242 (1997).
- 3) Fujita, M. and Kitano, H.: Development of an Autonomous Quadruped Robot for Robot Entertainment, Autonomous Robots, 5, pp.7-18 (1998).
- 4) Inaba, M.: Remote-Brained Robotics: Interfacing AI with Real World Behaviors, The 6th International Symposium on Robotics Research (ISRR6), pp.335-344 (1993).
- 5) 通産省アールキューブ研究会編, アールキューブ, 日刊工業新聞社 (1996).

(平成11年10月27日受付)

