

別刷

計測と制御

Vol.34 No.4 1995

音声で人間とコミュニケーションできる
ロボティクスアート：ニューロベイビー

と さ なお こ
土 佐 尚 子



社団 法人 計測自動制御学会

特集 人間と共に存するロボットの新技術

人にサービスする新機械—マシナサビエンス—をめざして

事例紹介

音声で人間とコミュニケーションできる ロボティクスアート：ニューロベイビー

と さ な お こ *
土 佐 尚 子 *

1. ニューロベイビー

1.1 アーティスティックコンセプト

私は、多様化する現代社会の中で、情報の波にもまれ消耗し、時間に追われるよう働き、束の間に幸せに漂っている現代人（とくに都会人）と共に存し、潤いのあるコミュニケーションができる賢い電子ペットを創りたいと考えていた。試行錯誤の結果、生まれたものがニューロベイビーである。名前の由来は、人間の赤ちゃんの顔をもつ電子生物の誕生と、その頭脳であるニューロコンピュータ自体も生まれたばかりの技術（当時）ということをかけている。

彼らは、コンピュータの中で生息し、音声で人間とコミュニケーションをとる。クリーンでかわいらしい。かまうのが面倒になら電源を消すといなくなったり、かまいたくなったら電源をONにすればよい。いつでも死に、いつでもよみがえるデジタルな生命である。しかも人間をモデルにしているので、好みに応じて性格を決めることができ、知識を学習できる機能をもつ。人間臭く天の邪鬼な性格に育てることもできる。近い将来、誰もが平等にニューロベイビーのような物をもち、育てる（カスタマイズ）ことができる。他人のニューロベイビーを傷つけたり、殺害したら、犯罪になるのだろうか？ ニューロベイビーは、人々と共に存し親しみのもてるペットであり、成長する知、情、意をもつ電子生物である。人類の近未来における、生命に対するアンチテーゼとして、起爆剤として現代社会に生きている¹⁾。

1.2 インタラクティブ＝コミュニケーション

インタラクティブという機能をリアルな対応のできるムービーの中の生物とのコミュニケーションにおきかえて考えた。私達にとって最も身近で、同格なコミュ

ニケーションをとる生物「人間」を選び、富士通研究所の技術者と共同研究を行った。図1は、ヨーロッパで行われた芸術とテクノロジーの国際会議であるアルス・エレクトロニカ、米国のCG国際会議 ACM・SIGGRAPH ('93年)、大阪のサム・ミュージアムで展示した方法である。

1.3 性格のデザイン

デジタルな人間の赤ん坊の顔をもち、人々の声の抑揚を耳から聞き取り、喜怒哀楽などの感情を人工知能（学習されたニューラルネットワーク）から判断する。たとえば悲しくなって泣く、樂しくなって笑う、機嫌が悪くなって怒る、大きな声に驚く、しゃっくりをする、退屈であくびをする、口笛を吹き、人々にあいさつなどを話しかけたりすることができる。

3次元リアルタイムコンピュータグラフィックスの、自由変形(free form deformation)ソフトウェアにより任意の顔の表情を増やしたり、ニューラルネットワークの学習データを変えることによりニューロベイビーの性格を変えることができる。

ハードウェアは、音声入力と出力に富士通製パーソナルコンピュータ FM-TOWNS を1台ずつ使用。画像生成用にグラフィックワークステーション SGI の Indigo 2 を使用している。

ニューロベイビーのデザインは、つきの5つのステップからなる。



図1 Neuro-Baby

*武蔵野美術大学造形学部映像学科 小平市小川町1-736

ATR 映像情報研究部 京都府相楽郡精華町光台2-2

キーワード：ニューロベイビー (neuro-baby), 人工生命 (artificial life), インタラクティブアート (AI based interactive art), 感情モデル (emotional model), 人間のような行為 (human like behavior).

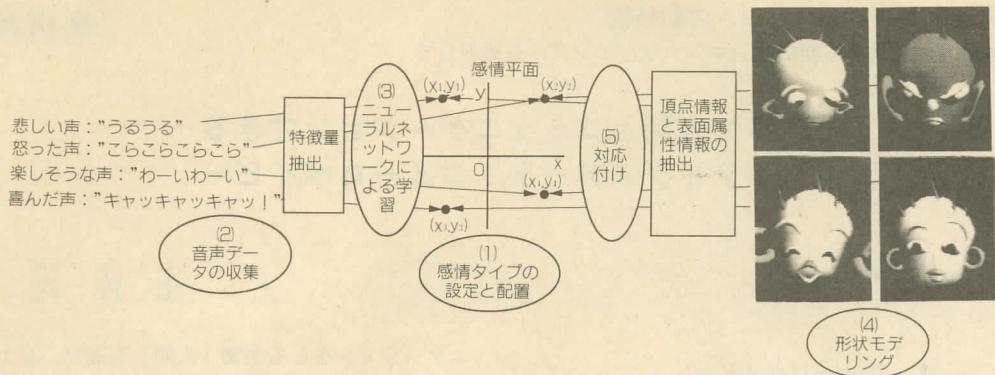


図2 ニューラルペイビーのデザイン手順

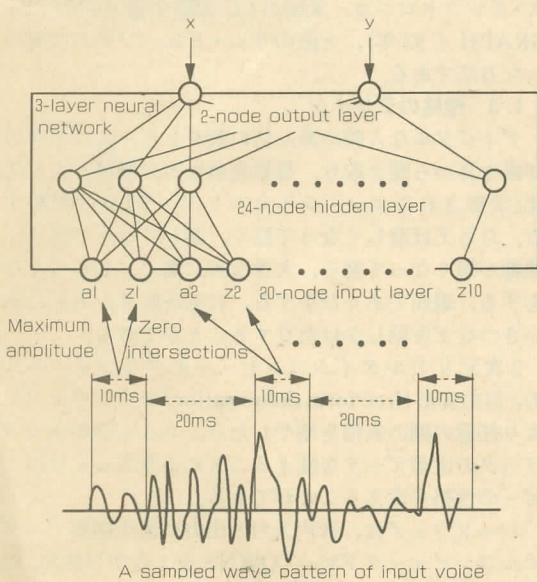


図3 ニューラルネットワークによる音声と感情との関係の学習法の一例

1. 感情タイプの設定と配置
2. 学習用音声データの収集
3. 音声と感情タイプの関係の学習
4. 表情の形状モデリング
5. 表情と感情タイプの対応付け

1.4 音声と感情の関係の学習

音声と感情との関係の学習で、私達はニューラルネットワークの学習機能を利用して音声から感情を抽出する手法を開発した。まず、感情タイプをいくつか設定し、感情座標に配置する。当時の感情タイプは、われわれ日本人特有の定版の感情表現である「喜怒哀楽」の4種類を用いている。つぎに図2の左端で示したように、それらの感情タイプを表現していると思

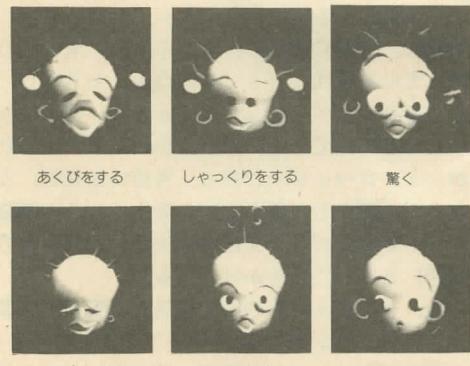


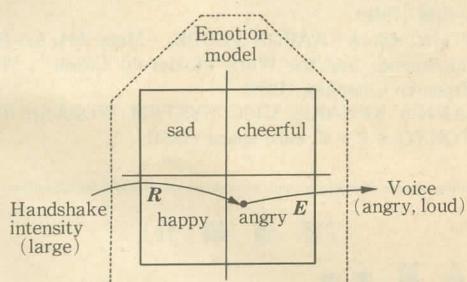
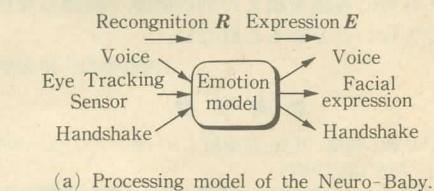
図4 イベントの種類と対応する表情変化

われる典型的な声を決め、それぞれについて複数の人間の声をサンプルする。さらに、サンプルした波形から特徴量を抽出する。特徴量としては、ある時間区間における波形の最大振幅 a と、ゼロ交差回数 z を用いる。ニューラルネットワークには、続けて何回か抽出した特徴量 (a, z) と期待する感情タイプに感情変数 (x, y) との関係を学習させる。学習アルゴリズムは、バックプロパゲーションを使用する。学習方法を図3に示す。また、人間の感情は、時間軸を伴って感情が変わる。この感情変化をイベント(図4)と称して、ニューラルネットワークだけでは散漫になる表情変化に、めりはりをもたせている²⁾。

2. 「ネットワーク・ニューラルペイビー」 — ACM SIGGRAPH'95 Invitation Event —

富士通研究所との初期バージョンを経て、現在さらに改良し、ネットワークバージョン“ネットワーク・ニューラルペイビー”を東京大学生産技術研究所の原島文雄教授、橋本秀紀助教授、瀬崎 薫助教授、東京大

学工学部の原島 博教授の各研究室と共同研究をしている。言葉によるコミュニケーションは、伝達する内容は明確になるが、現時点では、言葉の壁と文化の壁を作ってしまうことになる。しかし、キャラクターの方から人間に語りかけてきたら、人間は自然に言葉でコミュニケーションしたがる習性をもつようである。現に、私は話しができないニューロペイビーに日本語や英語で語りかけている老若男女をたくさん見た。そこで“ネットワーク・ニューロペイビー”では、各人に對してニューロペイビーの反応のカスタマイズ化を行う。今まで万人向けだったものが、“あなたに”反応するニューロペイビーになる。今までの機能は、耳と口のみであったが、目と手をもつことになる(図5)。これを、世界最大のコンピュータグラフィックスの国際会議である ACM-SIGGRAPH '95 で発表する。会場のロスアンゼルスと東京を学術情報センターのネットワークである SINET (6 MBbps になる予定) で結び、ネットワークイベントを行う。米国人用のニューロペイビーが人々と触覚バーチャルリアリティを用いて握手をし、その強さから、対応方法を判断し、相手



(b) Emotional States of the Neuro-Baby.

図5 Concept of the Neuro-Baby

の方を向き、声から感情を読み取り、ネットワークを通じて日本用のニューロペイビーに感性度を伝える(図6)。つまり、感性における日米の Culutreal・Exchange ができる。たとえば、日米でさまざまなゲームや視覚言語などを用いた、共通なコードによるコミュニケーションを設定し、ニューロペイビーがインターフェースとなり、ホストを務めることができる。相手とダイレクトに接しない感性エージェントを用いたコミュニケーションは、スーパー情報ハイウェイにおける新しいトピックスである。その近未来的ビジョンを実験的にデモンストレーションする。

3. アーティストとテクノロジーを結ぶメディアアート

私のバックグラウンドは、映画、ビデオ、CG といったメディア・アートである。ブレンダ・ローレル³⁾が、アリストテレスの演劇論である詩学から導いた劇的表現をコンピュータを舞台とし、新しいインターフェースの考え方として「劇場としてのコンピュータ」で論じた。彼女の本を読んだとき、私はメディア・アートからそれを導き出すことができ、見方を変えれば、すでに提示していることに気がついた。この場合、アーティストの考えをコンピュータに入力し、構造化した創造支援インターフェースとなる。実は、この方法は AI 的手法に似ている。たとえば、ハロルド・コーエンという1928年英国生まれの画家がいる。彼は、米国で抽象絵画の大家であり、彼自身の描き方の形式と方法をコンピュータに教え、コンピュータが彼の分身として自動的に描き、彼はその制御と仕上げを行う。このアーティスト支援コンピュータの名前を「アローン」という。アローンは、コーエンとスタンフォード大学のコンピュータ・サイエンス研究所との共同開発によるもので、アカデミックな米国 AI 研究者にも高く評価されている。それは、ボストンのコンピュータミュージアムに展示されている⁴⁾。また、20世紀の現代音楽を代表するひとりであり、ルーマニア生まれ、フランスに亡命・帰化した建築家で、作曲家のヤニス・クセナキス(1922年生まれ)は、「ユーピック」という人間が2次元ディジタイザーに描いたドローイングを楽譜としてコンピュータが読み取って作曲するシステムを作り上げた。この絵画的楽譜の書き方は、まず好きな波形

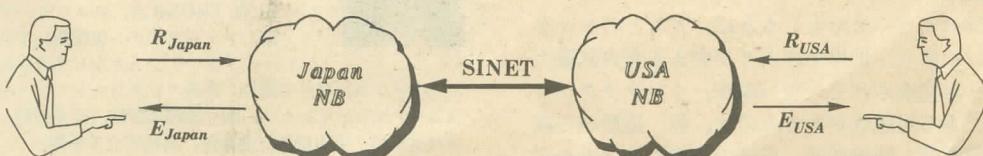


図6 Network Based Neuro-Baby

を描き、エンベロープを描く。そして、音の強弱を指定する。ディジタイザの縦軸は音の高さ、横軸は、音の持続時間である⁵⁾。

このように、インターフェースであり、システム自体が作品であるメディア・アートは、従来の芸術領域を拡張させる方向に、新たな表現を創造しつつある。造形的なイメージの定着から、時間軸と空間的な制約を拡大させながら、版画、写真の発明とその複製技術の発達があり、音の記録生成、映画の誕生による動く映像がはじまる。さらに電波、ラジオ、コンピュータなどの電子映像メディアが、人間の感覚器官へ直接接触していく方向で、表現領域を意識レベルまで到達させようとしている。ゲームによる対話型のエンターテイメントは、消費される娯楽であるが、アートはこれまでの系譜上に新たな視座を築きながら、人間の感覚的伝達と意識の拡張のための姿を垣間見せてくれる。新らしい時代が、社会を変えていく必然的因素であるならば、新たなアートとは、社会システムの構造を変えるほどのインパクトをもち、現代社会における秩序や情報の連鎖、関係性を変えるものである。

4. コミュニケーションの設計

ロボットが自然に能動的に動いてくれることはないと、人間の方から自然に感情移入するように、制作する人間が劇的なシナリオを設定しなければならない。感情移入度の度合いは、人間関係の場合、お互いの好感度によって左右される。それは、心理学でいう同調効果という相手と同じ行為やしぐさをすることによって生まれるものから、もう少し複雑なトリックもある。人は自分に対してどういう態度をとる人に、好意をもつかという実験によると、もっとも好意を感じたのは、付合いの初期に非好意的のうちに好意的になった人で、もっとも好意を感じなかつたのが、初期に好意的のうちに非好意的な人であった。これは、はじめは少しつれないような態度をとつておいた方が、好意を獲得したという満足感が得られるため、より相手に対して好意の度合いが強くなるものと考えられる。このような方法を電子ロボットと人々とのコミュニケーションに応用しシナリオ化することで、人々により感情移入させることが可能である。

4.1 コミュニケーションの基盤、共通言語とは何か

ニューヨペイビーのおもしろさは、人々とのインターラクションにある。世界中の人々が使える音声認識テクノロジーが完成されていない現在、インターラクションに使用できる可能性があるものは、色、造形から成り立つ視覚言語、映像言語、音楽の旋律やリズムから

引き出される感性言語などがあげられる。人それぞれのイメージがあるが、同じ概念を表現したものは、構造的に類似している。この構造的類似性には、多くの場合、表現しようとした概念の視覚的理理解に役立つ共通の直観と思われるものが十分にあることがわかる。つまり、世界中の人々がわかるシンボル、視覚言語を含めた感性言語を、色や芸術、デザイン、映像や音から引き出し、それを感性データベース化できれば、顔の表情以上のボキャブラリーが増えることになり、感性豊かな表現でコミュニケーションすることができる。

5. 最後に

テクノロジーとアートが互いに共通する問題を発見することは、技術的な実用性と人間的な意識が混合しながらシナジー効果を生み出す。便利さと娯楽的要素にあふれる先端技術は、ややもすれば、人間の批判的探求性を奪い、生きる目的を物質的な豊かさにすり変えてしまう。しかし、アートは深遠で普遍的なものをめざす方向に、破壊と創造を繰り返す。アートとテクノロジー、そして世界を見つめる畏怖の念、万物を包み込む宇宙といったものとのバランスを考える必要がある。それは、言語を超えた別の感覚で認識しなければ実感できないことかもしれない。

(1995年2月10日受付)

参考文献

- 1) 月刊「仏教」特集：生命、法藏館「トーキング・ニューヨーベー」、125/133 (1993)
- 2) 画像ラボ「ニューヨーベー（感情モデルを持つ表情合成システム）」、53/57、日本工業出版 (1993)
- 3) B. Laurel (遠山 訳)：劇場としてのコンピュータ、トッパン出版 (1992)
- 4) P. McCorduck : AARON'S CODE, —Meta-Art, Artificial Intelligence and the Work of Harold Cohen—, W. H. Freeman Company (1989)
- 5) IANNIS XENAKIS UPIC SYSTEM WORKSHOP in TOKYO カタログ、euro space (1984)

[著者紹介]

土佐尚子君



1961年10月22日生。アーティスト。武蔵野美術大学造形学部映像学科講師。映画、ビデオ、CGを用いるメディアアートを研究、制作。SIGGRAPH, ARS ELECTRONICA, ニューヨーク近代美術館、メトロポリタン美術館、モンペリアールインターナショナルビデオフェスティバル（フランス）などで発表。作品コレクション：アメリカンフィルム・アソシエイション、国立国際美術館、O美術館、富山県立近代美術館、名古屋県立美術館、高松市立美術館。