



サイバースペースで感情を認識・生成する  
インタラクティブキャラクターの研究

1999年

---

土佐 尚子

# 目次

あらまし

第1章 序論.....	1
1.1 本研究の背景.....	1
1.1.1 従来のヒューマンインタフェース.....	1
1.1.2 ノンバーバルインタフェース.....	1
1.2 本研究の目的.....	3
1.3 本研究の概要.....	4
第2章 感情インタフェース研究の課題と本研究の取り組み.....	5
2.1 まえがき.....	5
2.2 本研究の位置付け.....	5
2.2.1 感情インタフェースの重要性.....	5
2.2.2 アーティスティックなアプローチの重要性.....	6
2.3 本研究の課題とアプローチ.....	7
2.3.1 感情の認識.....	7
2.3.2 感情の生成.....	7
2.3.3 トータルなシステムの構築.....	8
2.3.4 感情移入の実現.....	9
第3章 感情コミュニケーションを行なうキャラクターの生成.....	11
3.1 まえがき.....	11
3.2 インタラクティブティを持った感情表現の設計.....	11
3.3 感情表現媒体としての「ひとがた」.....	11
3.4 ひとがたキャラクター「ニューロベイビー」.....	13
3.4.1 コンセプト.....	13
3.4.2 システム構成.....	13
3.4.3 ニューロベイビーのデザイン.....	15
3.4.3.1 感情モデルのデザイン.....	15
3.4.3.2 音声と感情の関係の学習.....	16
3.4.3.3 顔のモデリング.....	19
3.4.3.4 ニューロベイビーの音声出力.....	19
3.4.4 ニューロベイビーの実行時の処理.....	19
3.4.4.1 声の感情認識.....	20
3.4.4.2 顔の表情合成.....	21
3.4.4.3 効果音の出力.....	22

3.4.4.4	感情の時間的推移を考慮したイベント生成.....	22
3.4.5	ニューロベイビーの汎用性.....	23
3.4.6	まとめ.....	24
<b>第4章</b>	<b>感情を用いた異文化間のコミュニケーション.....</b>	<b>28</b>
4.1	まえがき.....	28
4.2	異文化間のコミュニケーションにおける感情の役割.....	28
4.3	コミュニケーションの仲介役としての感情キャラクター.....	29
4.4	感情キャラクター「ネットワークニューロベイビー」.....	29
4.4.1	コンセプト.....	29
4.4.2	ネットワーク型ニューロベイビーシステム.....	30
4.4.3	握手デバイスシステム.....	32
4.4.4	アイトラッキングシステム.....	33
4.4.5	ネットワークシステム.....	34
4.5	まとめ.....	35
<b>第5章</b>	<b>マルチモーダルに感情を認識・生成する</b>	
	<b>擬人化キャラクターの生成.....</b>	<b>36</b>
5.1	まえがき.....	36
5.2	マルチモーダルな感情認識と感情生成.....	36
5.3	マルチモーダルな感情モデルのデザイン.....	37
5.4	音声とジェスチャーから感情認識する	
	マルチモーダルキャラクター「ミック」.....	37
5.4.1	コンセプト.....	37
5.4.2	取り扱う感情.....	38
5.4.3	処理プロセス.....	39
5.4.3.1	基本方針.....	39
5.4.3.2	音声特徴抽出.....	43
5.4.3.3	感情認識.....	45
5.4.4	ジェスチャーによるインタラクション.....	50
5.5	まとめ.....	52
<b>第6章</b>	<b>声の感情と言葉の意味に対応する</b>	
	<b>「インタラクティブポエム」.....</b>	<b>53</b>
6.1	まえがき.....	53
6.2	インタラクション.....	53
6.3	ソフトウェア構成.....	55
6.4	処理プロセス.....	59

6.5	ハードウェア構成.....	62
6.6	まとめ.....	62
第7章	感情に反応するインタラクティブシアターの構成1 .....	63
7.1	まえがき.....	63
7.2	インタラクティブシアターの位置付け.....	64
	7.2.1 従来メディアに関する考察.....	64
	7.2.2 新しいメディアとしてのインタラクティブシアター.....	66
7.3	インタラクティブシアターの機能.....	69
	7.3.1 ストーリー生成.....	69
	7.3.2 シーン生成.....	71
	7.3.3 キャラクター生成.....	73
	7.3.4 インタラクション.....	74
7.4	システム構成例.....	75
	7.4.1 特徴.....	75
	7.4.2 ソフトウェア構成.....	75
	7.4.3 ハードウェア構成.....	77
7.5	インタラクティブストーリーの構成例.....	78
	7.5.1 コンセプト.....	78
	7.5.2 具体的なストーリー.....	78
	7.5.3 インタラクション.....	79
7.6	まとめ.....	80
第8章	感情に反応するインタラクティブシアターの構成2 .....	81
8.1	まえがき.....	81
8.2	第1次システムの評価と問題点.....	81
8.3	第2次システムの概要.....	82
	8.3.1 改良点.....	82
	8.3.2 ソフトウェアシステム構成.....	84
	8.3.3 ハードウェアシステム構成.....	88
8.4	インタラクション技術.....	91
	8.4.1 ジェスチャー認識.....	91
	8.4.2 感情認識.....	93
8.5	インタラクティブストーリーの構成例.....	96
	8.5.1 インタラクティブストーリー.....	96
	8.5.2 インタラクション.....	96
8.6	まとめ.....	101

第9章 結論.....	102
9.1 本論文のまとめ.....	102
9.2 本論文で実現できたこと.....	103
9.2.1 感情を介したインタフェースの実現.....	103
9.2.2 感情移入の実現.....	104
9.2.3 能動的なストーリー体験の実現.....	105
9.3 今後の研究課題.....	107
9.3.1 感情インタフェースの課題.....	107
9.3.2 次世代映画メディア実現のための課題.....	107
9.3.3 感動工学の実現に向けて.....	109
謝辞.....	109
参考文献.....	111
本論文に関する研究業績.....	117
その他の研究業績.....	129

## あらまし

本論文は、人間同士のコミュニケーションにおいて極めて重要であるノンバーバルコミュニケーションを人間とコンピュータの間でも実現しようとする手法についての研究をまとめたものである。本研究の特徴は、(1) ノンバーバルな情報の中でも特に感情に注目し、感情をインタフェースとした種々のインタラクションシステムのデザイン設計、システム構築を行い、感情インタフェースの重要性を示したこと、および、(2) デザイン設計、システム構築の過程で随所にアーティスティックなアプローチを適用したことである。

1章では、本研究の背景・目的・概要について述べる。まず従来のヒューマンインタフェースの研究を概観した後、ノンバーバルインタフェースの重要性および本論文の目的・概要を述べる。

2章では、本研究の位置付けと感情インタフェース研究の課題と本研究のアプローチについて述べる。

3章では、人間とコンピュータの感情によるコミュニケーションに注目し、音声処理技術と、ニューラルネットワークを用いて、人間の声に含まれる感情に反応するひとがたキャラクター「ニューロベイビー」について述べる。

4章では、異文化間のコミュニケーションにおける感情の役割と、ノンバーバルコミュニケーションの仲介役としての感情キャラクター「ネットワークニューロベイビー」を実現し、米国と日本を結び感性翻訳実験を試みた研究について述べる。

5章では、感情キャラクターを進化させたより人間的なマルチモーダルキャラクター「ミック」の生成を取り扱う。擬人化された個性の生成を目的とした感情認識のモデルと声とジェスチャーによるマルチモーダルなインタラクションを認識する手法について述べる。

6章では、感情インタフェースとバーバルインタフェースを融合した研究として、言葉の意味と声の感情を読み取り即興詩を作る「インタラクティブポエム」システムについて述べる。

7章では、上記で研究した感情に反応するキャラクターが物語りに登場して、観客自身が主人公を演じることで、没入感と感情移入を感じると共に、ドラマが主体的にインタラクティブに展開するインタラクティブシアターシステムのコンセプトと、第一次システムについて述べる。

8章では、第一次システムを改良し、より自由なインタラクティブ性を実現した第2次インタラクティブシアターシステムシステムの構成と、システムを用いたコンテンツ例として、「ロミオ&ジュリエット黄泉にて」について述べる。

9章では、本研究の総括と今後の展望について述べる。

# 1 序論

## 1.1 本研究の背景

### 1.1.1 従来のヒューマンインタフェース

従来のコンピュータとのインタラクションの特徴は、高速性・的確さ・容易性の3要素であった。しかしながら、コンピュータソフトウェア設計の分野でも、処理中心から情報中心へ、コンピュータから人間中心へ、形式中心から内容中心へとパラダイムシフトが起こりつつある。コンピュータの使い方は、情報の自動化から人間との共同作業に変わってきた。従って、コンピュータとのインタラクションの機会が確実に増えており、人間とコンピュータ間のインタフェースの重要性が増大している。

従来のマウス・キーボード・音声認識・画像認識といった人間とコンピュータ間のインタラクションのための技術は、論理的で人間にとって信頼でき正確ではあるが、冷たく難しく、専門家のみが使う物、愛着・親しみの持てない堅物としてのコンピュータのイメージを人々に植え付けている。抽象的表現で言うと、人間ならコミュニケーションの際、必ずやりとりされる感性が欠けていたのである。近年、認識技術が進み、顔表情などの感性情報を認識する研究が行なわれているが、まだ記号のような感性情報の表面を読み取るものである。そこでは本来、感性がもつ深さと広がりを取り扱う能力に欠け、一般の人々が使えるコミュニケーション技術にはまだ未熟の観がある。コンピュータに感情などの感性を生成させる研究もなされているが、部分的なものが多い。なによりも、コミュニケーションを行うためのトータルなシステム作りが欠けていると考えられる。

### 1.1.2 ノンバーバルインタフェース

#### (1) 人と機械のコミュニケーション

人間のコミュニケーションには、バーバルな要素とノンバーバルな要素があり、それを時と場合によって使い分けたり統合して用いている[1]。しかしながら従来、人間とコンピュータとのインタラクションについては、ほとんどの場合、クリエイティブな作業さえもバーバルなインタラクションに基づいて行なわれていると言える。具体的に従来のインタラクションテクノロジーとして使われてきたのは、ハードウェアとしては、マウス・キーボード・ディスプレイなどであり、ソフトウェアとしては、コマンドやコンピュータグラフィックスを始め、知的なインタフェースとして、文字認識[2]・音声認識[3]・カメラ画像を用いた画像によるパターン認識[4]等があげられる。

## (2) ノンバーバルインタフェース

ノンバーバルインタフェースとは、図形・色・動き・雰囲気・音楽・声質・表情・感性といった言葉以外のメッセージにより、コンピュータとインタラクションを行なうインタフェースである[5]。1960年頃から重要視されるようになったインタフェースであるが、現在研究されている物では図形言語であるアイコンを始めマッキントッシュなどが採用している直感的ユーザインタフェース、画像認識、身ぶり認識、顔の表情認識などがあげられる。ノンバーバルインタフェースは、今後マルチメディア、バーチャルリアリティーなど画像や音声を空間的かつ時間的に扱うメディアで一層必要とされるだろう。

## (3) ノンバーバルインタフェースの重要性

ノンバーバルコミュニケーション研究に携わる研究者によると、コミュニケーションで交されるメッセージ全体のうち65%は、ノンバーバル言語が担っているとしている[6]。またMehrabianの研究によると、人の感情や態度、性格に関する知識のうち、言葉により獲得されるのは7%に過ぎず、55%は身体動作から、また38%は周辺言語から得られるとのデータを示している[7]。

バーバルコミュニケーションは、言語という論理的媒体をコミュニケーション手段として用いているので、言葉による論理的な表現のみで、人間のコミュニケーションが持つ微妙な深い機能を全て網羅することはできない。ノンバーバルモードでは、言葉の論理的内容によって、フィルターをかけられることなく感情成分を引き出すことができるので、状況や他人に関する自分の感情・感性を引き出すことができる。従ってより深い全体的なコミュニケーション、言い換えれば人間的なコミュニケーションが行えるといえよう。

## 1.2 本研究の目的

本研究では、感性の生成と認識を含めたトータルなシステムのデザイン設計とシステムの実装を達成することを目的とする。今まではバーバルなインタフェース[5]に関する研究は数多いが、ノンバーバルな情報を含めたインタフェースに関する研究は少なかった。行なわれている研究も顔表情やジャスチャーなどに関するものが多く、ノンバーバルな情報の中でも非常に重要な働きをする「声」に含まれる「感情」をインタフェースとして取り上げた研究は、ほとんど存在しないといえる。本研究ではノンバーバルな情報の中で、特に感情を生成するインタフェースに注目して研究を行うことにした。感情といえども、一種の人間の思考過程である。思考を工学的デザイン化する方法[8]に基づき研究をすすめた。

ノンバーバルインタフェースを人間の感覚や感性の技術研究というとらえ方をすると、現在ではバーチャルリアリティや、他のコンピュータテクノロジーが発達して、いろいろなハードウェアやソフトウェアの研究が出現してきてはいる。しかしながら、それらは、人間が日常生活で営んでいる五感の機能の再現を狙っているものが多い。残念ながら、我々が日常的に感じている五感の方がリアリティーがあり、テクノロジーが再現した感覚に満足できず、感性をくすぐらないのが現状である。さらに、我々自身が昔と比べてテクノロジーに慣らされているので、感性や知覚の探究が後退している印象がある。ゲームによる対話型エンターテインメントのアプローチもあるが、まだ条件反射的・体感的であり、人間の情感を引き出すといったところまで至っておらず、消費される娯楽となっている。これらを解決し、情感、しいては感動を引き出すためには、アーティスティックなアプローチを取り入れることがよいのではないかと考えられる。アートは、個性的主観が意識的に陶冶・練磨された技術と融合して、精神の充足の高みを指向する独創的な生産活動である。アートのアプローチにより創造されたコンピュータキャラクターが感情を判断するモデルを持ち、能動的に人々とノンバーバルコミュニケーション、バーバルコミュニケーションを対話的に行うシステムの実現が本研究の目指したものである。

### 1.3 本論文の概要

本論文は、感情を中心としたノンバーバルインタフェースのデザイン設計法と具体的なシステム構築例について述べたものである。本論文の特徴は、(1)感情インタフェースの種々のレベルのシステムを構築したことと、(2)デザイン設計には工学的手法と同時にアートの手法を用いている点にある。以下に本論文の構成と、各章の概要を述べる。

第2章では、まず本研究の位置付けを示すため、感情インタフェースの重要性とアーティスティックなアプローチの重要性について述べる。次に、本研究で取り上げる具体的な課題とアプローチについて述べる。

第3章では、インタラクティビティ＝コミュニケーションと考え、インタラクティビティを持った感情表現のモデル化を試みる。又、人間が感情移入をする媒体のモデルとしての「ひとがた」について重点をあてる。自己の分身であり自分に最も近い他者をキーワードに、話しかける人の声から感情抽出をし、インタラクティブに答えるひとがたキャラクター「ニューロベイビー」との新しいコミュニケーションの形態を述べる。

第4章では、感情を用いた異文化間のコミュニケーションについて述べる。まず、異文化間のコミュニケーションにおける感情の役割とは何かを考察する。次に、コミュニケーションの仲介役としての感情キャラクターの位置付けとその実例として、感情キャラクター「ネットワークニューロベイビー」を用いて異文化間感情翻訳実験を行なった内容について述べる。

第5章では、感情に反応する擬人化キャラクターの生成について述べる。感情キャラクターからマルチモーダルな擬人化キャラクターへ進化させ、より高度なコミュニケーションを生成するため、感情モデルのデザインの複雑化を計る。その実例として、音声とジェスチャーからマルチモーダルに感情認識するマルチモーダルキャラクター「ミック」について述べる。

第6章では、バーバルインタフェースとノンバーバルインタフェースの融合を図る。コンピュータと感情のやりとりのみではなく、言葉の意味も同時に理解して、連歌形式で人間と即興詩を作れるコンピュータ詩人「インタラクティブポエム」システムについて述べる。

第7章では、擬人化キャラクターを包む世界に、インタラクティブストーリーを導入し、その結果として感情移入型インタラクションの実現を目指したインタラクティブシアターについて述べる。特にインタラクティブな物語の進行をコントロールするスクリプト設計の基本的概念について述べる。

第8章では、第7章で述べた第1次システムの問題点を改良した第2次システムについて述べる。改良した点は、複数人の参加が可能なシステムの実現と、任意の時点でのインタラクション(anytime interaction)が可能なメカニズムの実装、より複雑なインタラクティブシナリオの実現を可能にしたスクリプト記述である。

第9章では、本研究の総括と今後の展望について述べる。

## 2 感情インタフェース研究の課題と 本研究の取り組み

### 2.1 まえがき

本章では、本研究の位置付けとして、感情インタフェースの重要性とそれを取り扱うことの重要性について述べる。さらに感情インタフェースを実現するための課題と具体的なアプローチについて述べる。

### 2.2 本研究の位置付け

#### 2.2.1 感情インタフェースの重要性

ノンバーバルインタフェースで取り扱われる対象としては、先に述べたように図形・色・動き・雰囲気・音楽・声質・表情・感性などがあるが、これは低次の感覚的なものから、高次の感性的なものまで極めて広いダイナミックレンジを持っている。従来の感性処理の名の基に取り扱われてきたのは、低次の感覚的なものであり[9]、人間とコンピュータ間のノンバーバルインタフェースとしてはより高次のものが望ましい。一方で、感性のような高次の情報は人により解釈が違っていたり、個人性に依存する部分が多いため、現時点では取り扱いが困難である。従ってその中間に位置する感情をインタフェースとすることは現時点では適切と考えられる(図2.1)。しかしながらこの分野の研究は、顔表情を対象とした感情分析[10]、感情認識[11]、感情合成[12]の研究は多いものの、声に含まれる感情を対象とした研究例は少ない。さらに、それらを統括し、かつ、バーバルな機能が一緒になったトータルなシステムはまだ実現されていない。

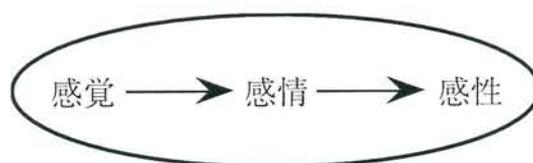


図2.1 感情インタフェースの高度化

## 2.2.2 アーティスティックなアプローチの重要性

従来のヒューマンインタフェースの研究はいずれも工学的アプローチに基づいて行われてきた。それに対し、本研究のアプローチはアート的方法論を取り入れていることに特徴がある。別の言い方をすると、本研究は、芸術の創造性を活用した人間の知覚の探究や感性の拡張の技術研究ということができる。芸術家が、創作を開始したとき、頭の中に明確な最終イメージを持っていることはほとんどない。したがって作品へのプロセスは、ゴールへの到達ではなく、むしろゴールへの探究なのである。探究の方法は、工学における、客観的定量化の方法とは異なり、個人の感じる主観に基づいて判断がなされる。感じ方の切り口の深さ、新しい価値観の定義が評価される。芸術家は、その時代性と人間的な意識を反映する作品を呈示するので、その考え方を取り入れることにより、新たな感性インタフェースを発想することができる。現在はアナログからデジタルへ情報の数値化が進んでおり、感性についてもそれを数値化するシステムを考える時代に突入している。これは感性工学の研究とも接点があり、深い意味で人間の感性をインタフェース化する研究として位置付けられる。このような位置付けで感情の認識と生成を取り扱う。

以上の言い方は抽象的であるが、実は極めて現実的なアプローチと言える。例えば、ブレンダローレルは、「劇場としてのコンピュータ」という本で、現代のコンピュータ技術における情報処理機能とユーザの関係を、演劇を観る観客／ステージ／舞台裏の三つの要素に分け、ステージの上の演技の展開を支える照明や音響などの舞台裏の機能を、いかにして観客の無意識下に置くかがコンピュータインタフェース技術の主要な役割であると主張している[13]。劇場を生成する映像の表現力とコンピュータの情報操作能力を結びつければ、コンピュータが人間の思考やイマジネーションを高めることができるのではないか。これは工学的アプローチとアートのアプローチを統合したアプローチのひとつの例といえよう。感性情報処理の研究に見られるように、現在は、今まで研究されていなかった情緒・感性についての研究が、認識技術・人工知能・ニューロコンピュータ・脳科学などの領域で行われ、人間の精神活動の様々な部分についての分析と研究が行なわれ始めている。これらの研究成果は将来的にはノンバーバルインタフェースに生かしていくことが可能であるが、現時点では感情などを取り扱う技術とアーティスティックアプローチが融合することによって初めて具体的なノンバーバルインタフェースが実現できると考えられる。

## 2.3 本研究の課題とアプローチ

### 2.3.1 感情の認識

感情認識を実現するための技術を考えてみる。現在使用できる認識技術は、ニューラルネットワーク、HMM等があげられる。しかしながら具体的な研究例は極めて少ない。しいて例をあげれば、MITメディアラボのコンピュータによる感情理解の研究[14][15][16][17]、顔の表情から感情を認識・合成する研究としては、東京大学原島研究室で行なわれた顔の表情合成[12]からその「印象」をコンピュータで解析する研究[18]を初め、成蹊大学森島研究室で行なわれている、CGで作られた顔の表情をニューラルネットワークを用いて感情認識を行なうシステム[19]や、東京理科大学原研究室で行なわれている人間の顔表情を持つロボットが表情筋の動きによって感情的表情を生成するシステム[20][21]、オプティカルフローにより表情筋の動きを検出して感情を認識する方法[22]、顔表情から感性情報・性別・年齢を分析する研究[23]をなどがある。

### 2.3.2 感情の生成

感情生成の方法論は、心理学の分類から引用してくるのが、従来のパターンであった。例えばEkmanによると、恐れ、驚き、怒り、嫌悪、悲しみ、喜び、の6つの基本情報がそれぞれ特定の刺激事象で生じると共に、これらの個々の感情は個別の表情と生理的徴候のパターンを備えている[24][25]。感情喚起過程(図2.2)においては、特定の刺激の知覚が特定の顔面・姿勢反応を生気させ、引き続いて特定の自律神経系の活動と感情経験を引き起こす[26]。

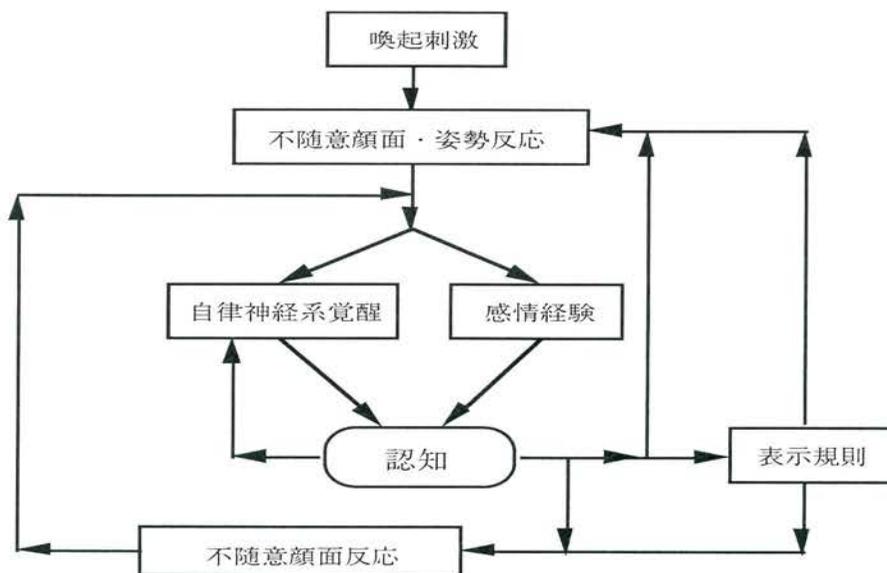


図2.2 Ekmanの感情喚起過程

海外の工学的研究では、MIT人工知能研究所のBrooks が率いるヒューマノイドロボットプロジェクト「Cog」[27][28]は哲学者のダニエルデネット(タフツ大学)と組み感情の生成モデルを考えている。そして、カーネギーメロン大学の Joseph Bates が率いる Oz プロジェクト[29]やニューヨーク大学のメディアラボのPerlin が率いるインタラクティブアクターの研究[30][31]は、ディズニーのアニメーションの動きと人工知能からその感情モデル生成を作り出し、スタンフォード大学の Hayes-Roth が率いる バーチャルシアタープロジェクト[32][33]は、人工知能研究の立場から感情生成モデルを考えている。日本ではコンピュータが人間の顔の特徴をとらえて似顔絵を生成するシステムの研究[34][35]などがある。しかし、それらは、感情パターンと感情パターンの間を埋める中間感情や、時間軸を伴った複合感情を記述するには至っておらず、感情が浅く記号化されて扱われ、駆け引きのような微妙な感情コミュニケーションが抜け落ちている。これに対しここでは、人間を深い洞察で描くことのできるアーティスティックなアプローチを用い、新たな感情の領域の生成と複合感情を伴ったシステム設計を考案する。感情表現の分類と定義は、小説・詩・ドラマ・色彩・形態・音色から分析することができる。時間軸を伴った感情表現の行為・臨在感は映画・演劇から分析することが可能である。フランスの映画評論家アンドレ・バザンは、歴史的な著作「映画とは何か」の中でプレゼンス(臨在感)という用語を用いた[36]。映画を見ている観客は、自分がスクリーンの映像と同じ時空間内に存在しているように感じるものであるが、プレゼンスという言葉はそういった感覚を表している。映画は、幻想をもう一段高め、観客は、まるでそれが目の前で起こったことのように感じる。このリアリティーは、技術の力と人間の創造力によって支えられているのである。映画の方法論を取り入れることは、時間軸を持つアートの方法論を用いることでもあり、これによって人間的な深みを持ったインタフェースが実現されると期待される。

### 2.3.3 トータルなシステム構築

次に重要なのは、感情の認識と生成を組み合わせ、総合システムを構築することである。人間のコミュニケーションは、双方向的なものであることが重要である。そこでここでは、感情を軸として感情の認識と生成に基づいた人間とコンピュータの双方向コミュニケーションシステムを構築することを目指した。上記に述べるように、感情の認識・生成の個別研究はあるが、これらを統合したシステムは見当たらない。人間とコンピュータ間のコミュニケーションの統合システム構築の例としては、擬人化されたキャラクターの表情を使って、ユーザーにシステムの状態を伝える東芝の開発した TOSBUG II[37]や、Sony CSL の Talkman[38]などがあり、ユーザーとの対話を円滑にしている。しかしながらここでは、人間からコンピュータへのインタフェースとしては、感情は使われていない。感情をインタフェースとした双方向コミュニケーションの可能な統合システムを構築して始めて、感情をインタフェースとすることの有効性が実現できるのであり、その点に本研究の重要性がある。

本研究ではまず、インタラクション＝コミュニケーションと考える。感情検出の方法としては、血流変化のような生理データの計測による方法などがある[26]が、インタフェースとして用いることは困難である。そこで、人間にとって自然なコミュニケーションの媒体である音声に含まれる感情面を用いたノンバーバルインタフェースの実現を図る。次にネットワークで結んだ異文化間での感情翻訳実験の試みを行ない、感情を異文化間のインタフェースとすることの有効性の検討を行なう。さらに、ノンバーバルインタフェースとバーバルインタフェースを統合したシステムと、マルチモーダルなインタラクションにより複合感情の実装を実現する。最終的に、それら全てを時間軸と共に包括して記述することにより、感情によって生成されるインタラクティブ映画システムの実現について検討する。

## 2.3.4 感情移入の実現

もう一つ重要なことは、感情移入の導入である。いかにすばらしい仮想世界を作っても、そこに我々が感情移入しなければ、それはうわべだけのものである。人々はそのような仮想世界に長く留まろうとは思わないだろう。感情移入とは、他人と同じように感じることを意味する。いいかえれば、もうひとりの人間の心の状態を疑似体験することである[39]。筆者は感情移入技術を実現するために、まず、インタラクティブコンピュータキャラクターを使って人間の感情移入が実現できないかと考えた。キャラクターの設計思想の基本は自己の分身のように親近感があり、自分の事を良く理解してくれるように見え、かつ客観的な行動を行なう存在である。コンピュータ上のキャラクターとしては、私達にとって身近で一番コミュニケーションをする生物である「人間」を選び、心理状態・感情表現・性格付け・知性・行為さらにはこれらを統合したコミュニケーション行為をテクノロジーを用いてどこまでリアルに個性を持って表現できるのかという興味のもとに研究を進めた。インタラクティブキャラクターやエージェントの研究は、最近多く行われているが[40][41][42][43]、感情移入の実現を目指している研究例はない。

また、本研究では仮想世界を人間が寝ている間にみる夢に例え、夢の世界をリアルに体験できないかということを目指した。我々は夢を見る場合、常にその夢の中のできごとに主体的に関わっている。我々自身が主人公であり、歩いたり、話をしたり、喜んだり、悲しんだりして、実際に汗をかいたりもする。この時、我々は、それを意識しているのだろうか？それとも、深層心理が働いているのだろうか？一方、我々が小説や、映画を見る時、その作品に感動していると、心はその世界に入っている。「2001年宇宙の旅」という超大作SF映画を制作したスタンリー・キューブリック監督も、「人間が映画を見ている時、その体験は何よりも夢に近いものである。現実ではとても無理なことを映像として表現する際、基本となるものは、現実としての次元や空間である。事実から物語りのアイデアが展開する」とインタビューで述べている[44]。しかし、我々はどんなに感情移入していようと、現実と非現実の壁があり、その世界に直接踏み込めない。本研究ではインタラ

クシヨンの導入により、この壁をやぶってみようと考えた。映画の世界に入って、その場の雰囲気  
を堪能したり、キャラクターと話してみたり、話す内容によって、ドラマが変わってしまうようなこ  
とが可能システムを文脈性と創造性を生かした工学知の技術[45]に基づいて研究した。これらの  
意識・無意識に行なわれるコミュニケーションを、映画的演出[46]とインタラクティブストーリー  
を用いて、アートの感性をシステムデザインの中に生かしていくという方法論をとることとする。

### 3 感情コミュニケーションを行なう キャラクターの生成

#### 3.1 まえがき

まず、感情をベースとした人間とコンピュータのインタラクションについて考察してみる。感情を含めたノンバーバルインタラクションを人間とコンピュータの間で実現しようとしている典型的な例はインタラクティブアート[47]である。しかし、従来のインタラクティブアート[48]はパターン化した現代美術風の観念的で冷たい装置のような作風が多かった。それに対し筆者は、インタラクティブティの機能を、生き生きとしたリアルな対応のできる生物の生成と、その生物とのコミュニケーションに置き換えて考えると新しい可能性が生まれてくると考えた。そこで、私達にとって身近で一番コミュニケーションをする生物「人間」を選び、心理状態・感情表現・性格付けや知能・行為とコミュニケーションをテクノロジーを用いてどこまでリアルに個性を持って表現できるのかという興味のもとに研究を始めた。

#### 3.2 インタラクティブティを持った感情表現の設計

インタラクティブティ＝コミュニケーションと仮定した場合、感情移入が必然的に関わってくる。人と人はもちろんのこと、人と物でも同じである。インタラクションは、人間が能動的な気持ちにならないと起こらない。人間が能動的になる時は、好ましい感情の時である。ならば、だれもが、好ましい気分になる状況を作る必要がある。筆者は、その状況を、具体的に大人が赤ん坊に対する関係の中に見い出そうとした。例えば、図3.1のように人間の赤ん坊も生まれたばかりでは、感情表現が興奮と関心の2種類しかないが、成長と共に増えていく[49]。また、生まれた時からだれに教わるわけでもないが、赤ん坊の仕草も感情と密接に関係していることが解る(図3.2)[49]。

#### 3.3 感情表現媒体としての「ひとがた」

人間が、老若男女問わず「ひとがた」の物を好むのはなぜだろうか。昔から、土偶からロボットにいたるまで、人は自分と同じ形を作り、愛着を覚え感情移入をしている。近年の「ひとがた」ロボットには人間に似たものという願望あるいは期待が込められている。とすれば、ロボットこそ人間と機械との「関係」をもっとも集約して考えさせる存在といえる。まさに、ロボットは機械人形である。人間と機械との問題は、人間と人形との「関係」の問題でもある。現代では、人形またはロボットは、物理的制限を超え、コンピュータによって、仮想的なキャラクターとして存在する。

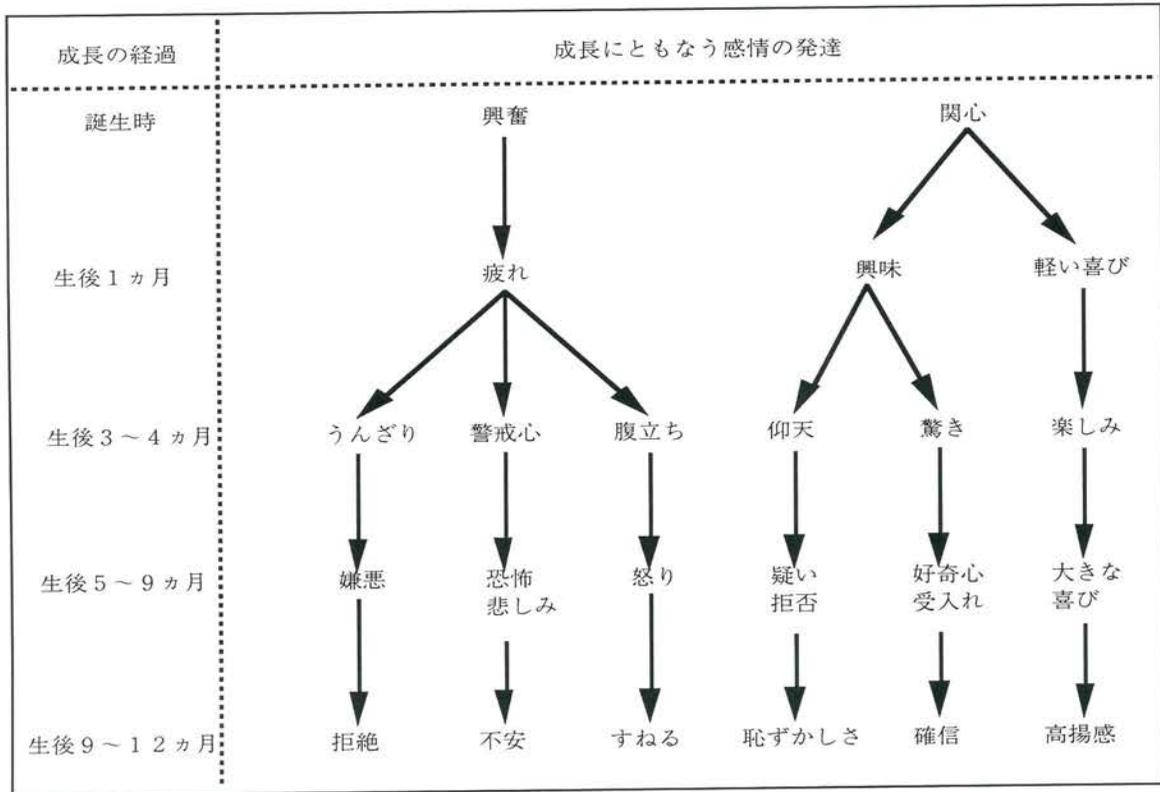


図3.1 0歳児の感情の発達表

注目点	声	顔の表情	手のうごき	体全体の動き
その意味				
そろそろ親と交流したい		目をみひらく 眉毛を吊り上げる 口を開ける 顔を輝かせる	指を開き、軽く曲げる	顔をあげる
いますぐ親と交流したい	クスクス笑い 片言のお喋り クスンクスンと泣く	微笑 親の顔を見つめる 親の目をじっと見る	親へと手を伸ばす	親のほうを向く 手足を滑らかに回すように動かす
そろそろ一人になりたい	くずり泣く しゃっくり 鼻を鳴らす くしゃみ	瞬きが増える 舌を突き出す 口を曲げる しかめたように眉が下がる 口をしっかりと閉じる	手を合わせる 手を握りしめる 手を耳や首や口へ持っていく	脚で蹴る 首を下へ回す 脚を突っ張る 両手を横に突っ張る
もう一人きりにしてほしい	泣く くずる 騒ぐ 唾を出す	泣き顔 しかめっ面	親へと手を突き出す 握りこぶしで叩く	背を丸める 首を振る 頭上で手を叩く

図3.2 赤ちゃんの仕草の読み方

### 3.4 ひとがたキャラクター「ニューロベイビー」

#### 3.4.1 コンセプト

この忙しい情報化社会の中で、人間ほど気をつかわなくて良くて、ペットほど面倒を見なくてよい、ある程度自分自身で判断できる何かを、ふと求めたことはないだろうか。ニューロベイビーは、自己の分身のようであり自分に最も近い他者がコンセプトである。ニューロベイビーは人間が声をかけても、その意味はわからない。しかし声の抑揚から、感情を判断しそれに反応することができる[50]。

#### 3.4.2 システム構成

ニューロベイビー[51][52]のシステム構成を図3.3に示す。システム設計は、富士通研究所との共同研究で行った。ニューロベイビーは、音声解析部、画像生成部、音声出力部の3つのサブシステムからなる。音声解析部から画像生成部及び音声出力部へは、音声から感情を認識した結果である感情変数と呼ばれるデータが送られる。感情変数および感情モデルについては、別途に述べる。

音声解析部ではマイクに向かって声を発した人間の音声波形を分析してニューラルネットワークに与え、その出力として感情変数を得る。ニューラルネットワークには、あらかじめ人間の声の調子と感情変数との関係をいくつか学習させてある。画像生成部では感情変数を入力として、赤ちゃんの表情のCGを出力する。感情変数のいくつかの値に対応する顔は、キーフレームのみ予めデザインしておく。例えば楽しそうな声に対する期待出力として学習させた感情変数の値に対しては笑い顔が対応するようにする。実行時には、入力された感情変数の値に応じて、予めデザインした顔の補間を行い表情を生成する。画像生成部には、IRIS4D、音声出力部では、FM-TOWNSを使い、表情に適應するよう登録された効果音を出力する。

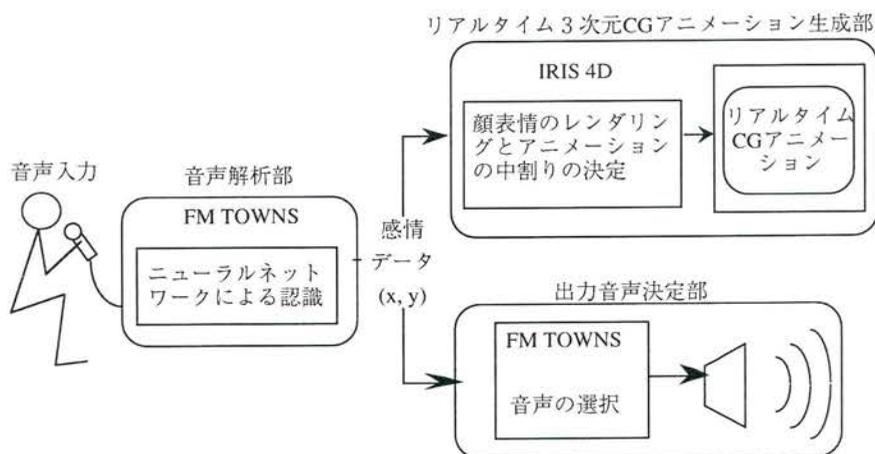


図3.3 ニューロベイビーのシステム構成

図3.4は、オーストリアのリンツで行われたアート&テクノロジーの国際会議「アルスエレクトロニカ」[53]での展示風景である。

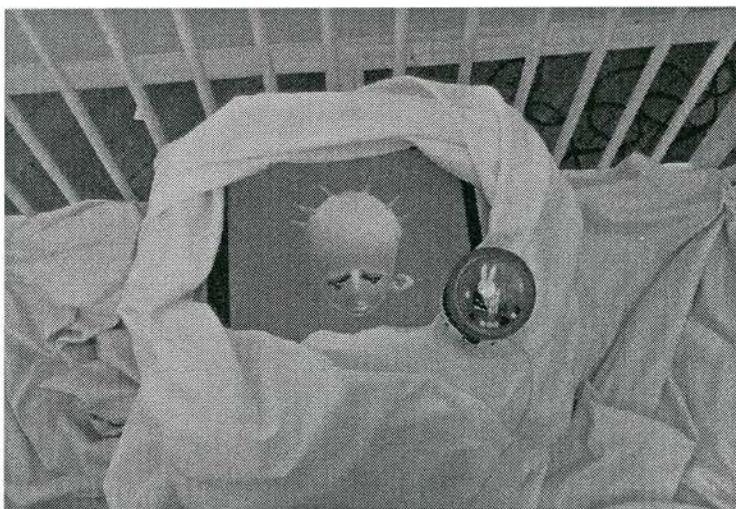


図3.4 ニューロベイビーとの対話

### 3.4.3 ニューロベイビーのデザイン

図3.5にニューロベイビーのデザインの過程を示す。デザインの主な作業は典型的音声の例を与え、感情変数を通じて顔モデルとの対応をつけることである。

#### 3.4.3.1 感情モデルのデザイン

ニューロベイビーでは、人間の感情を数値データとして扱うための感情モデルを導入している。仮定として、人間の感情が多次元の距離空間である「感情座標系」あるいは「感情空間」を形成する。この仮定に従えば、喜怒哀楽などの感情タイプは、それぞれ感情空間の特定の領域に配置できる。感情空間の次元を理想的にはいくつにすべきか現時点で決定できないが、現システムでは、感情空間を2次元平面にうまく投影できたと仮定し、投影された感情平面上で、感情モデルをデザインする。図3.5の(1)で示すように、感情平面の点、例えば $(x_1, y_1)$ などに喜怒哀楽のようないくつかの感情タイプを配置する。

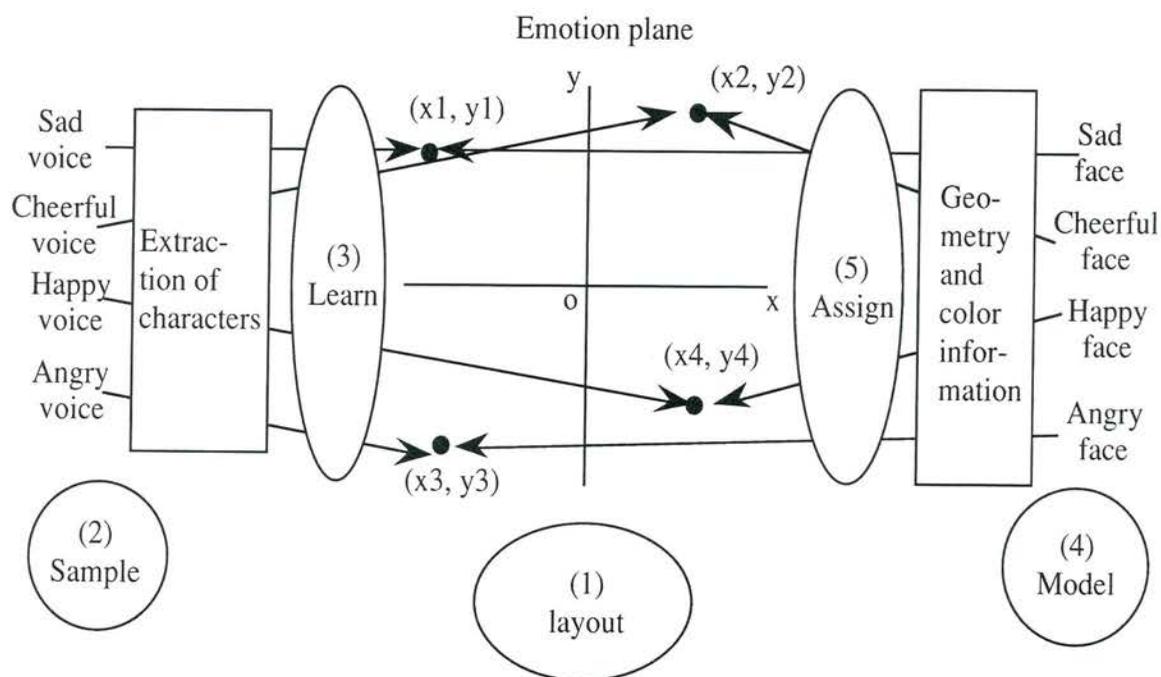


図3.5 ニューロベイビーのデザイン手順

### 3.4.3.2 音声と感情の関係の学習

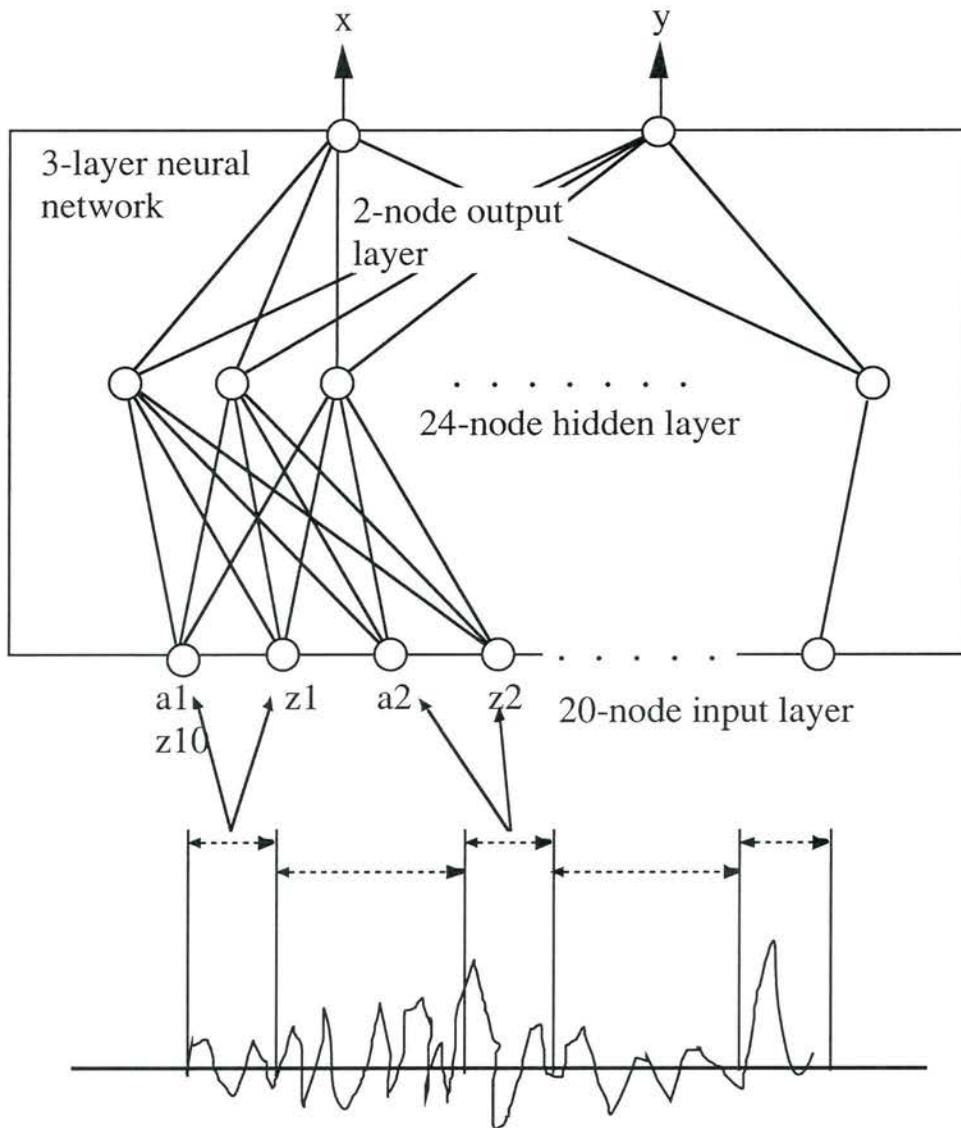
音声解析部では、実行時に人間の音声感情変数  $(x, y)$  に変換することによって感情を認識する。そのための準備として、入力音声と感情変数との関係を作っておく必要がある。音声と感情の関係についての研究は意外に少なく、音声合成における感情表現を扱った文献 [54] がある程度である。

筆者らは新たに、現実の音声波形をサンプルし、ニューラルネットワークに学習されることによって音声から感情を認識させる方式を考案した。この学習作業は、ニューラルネットワークに、この声は怒っている声、この声は楽しい声、この声は哀しい声というように感情を憶えさせることに相当する。これは、図 3.5(2),(3)に示されている部分である。

学習方法の一例を図 3.6 に示す。まず、マイクで数百 msec 分の音声波形を取り込んでデジタル化する。次にその中の 10msec 分の波形について最大の振幅値  $a$  と波形のゼロ交差回数  $z$  を調べる。これを 10 組み分取り出し、計 20 の数値列  $((a_i, z_i), i = 1, 10)$  をニューラルネットワークの入力パターンとする。そしてその音声に対応させたい感情平面上の点座標  $(x, y)$  を期待出力とする。これを各感情タイプについてそれぞれ数回ずつ音声をサンプルしてニューラルネットワークに与える。

学習アルゴリズムは、バックプロパゲーション法 [55] を使用している。学習させる感情変数は、4 つの感情タイプが感情座標平面上に適切に配置されるように人間が決定する。使用したニューラルネットワークの入力層ノード数は 20、中間層ノード数は 24、出力層ノード数は 2 である。このようにしていくつかの組みの音声波形と感情タイプを学習したニューラルネットワークを使うと、任意の音声入力に対して感情変数が得られる。また、ニューラルネットワークの性質により、曖昧な音声でも人間の期待する感情変数が得られる。

Expected output emotion variables for the input voice



A sampled wave pattern of input voice

図 3.6 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

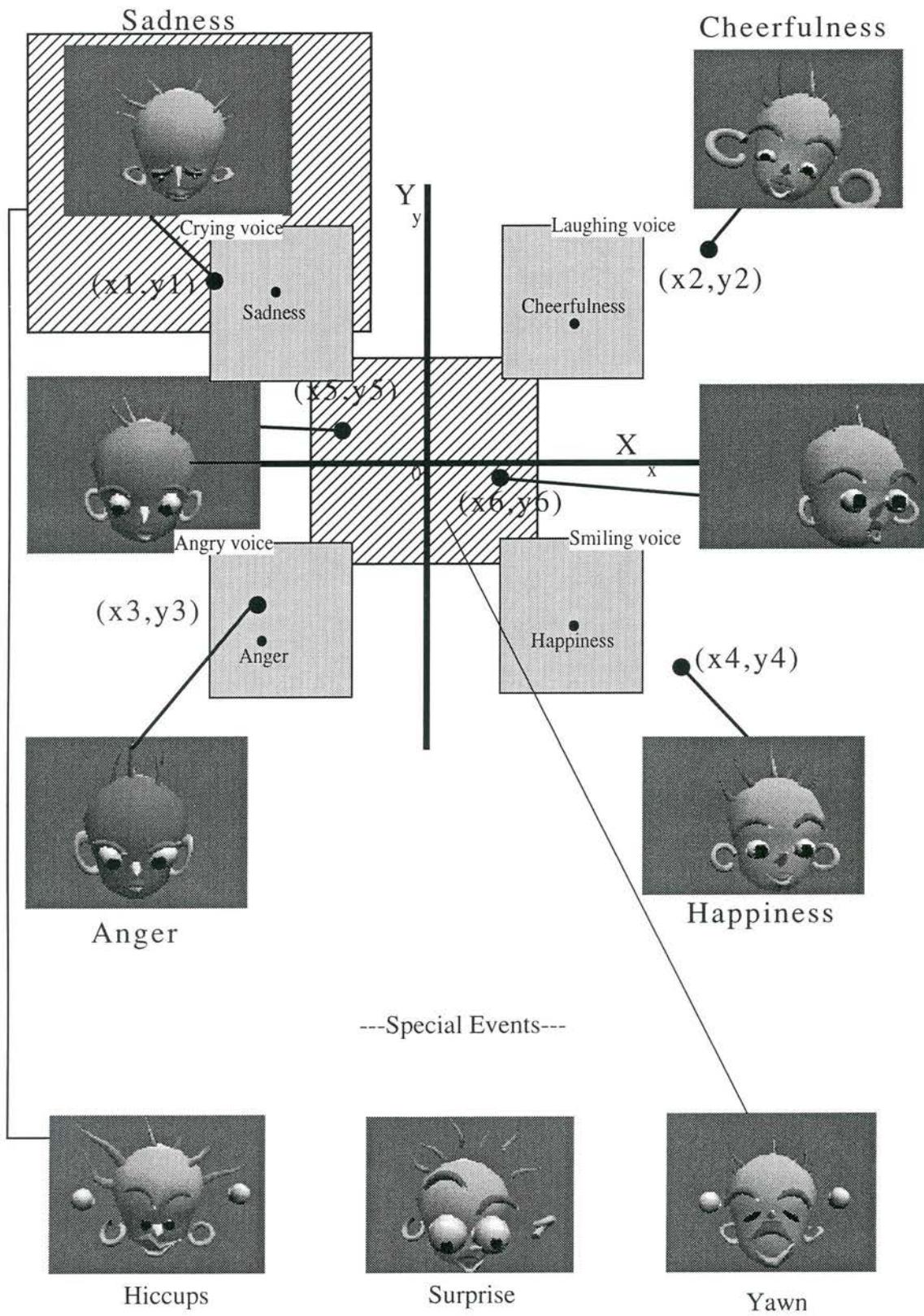


図 3.7 顔の表情と感情タイプの反応

### 3.4.3.3 顔のモデリング

ニューロベイビーの表情生成のモデリングについて述べる。今回作成したCGの赤ちゃんの基本形状と、感情パターンの反応を図3.7に示す。形状は、頭・目・唇などの18のパーツからなり、合計1000ポリゴン以内のモデルである。あらかじめ、それぞれの感情に対応する表情を作っておき、任意の感情座標に配置する。すべてのパーツの頂点座標は、配置した感情座標と対応して記憶される。なお、人間の顔面のいろいろな表出ジェスチャーはチャールズ・アウトベルトのパントマイム芸術に関する著書から参考にした[56]。

### 3.4.3.4 ニューロベイビーの音声出力

音声出力部では、表情と同期して赤ちゃんの声を発するように、音声データとして取り込み、例えば、図3.5に示すように感情平面上の適した領域に張り付ける。実際のマップは複雑で、取り込んだ音声の種類は50種類になる。

### 3.4.4 ニューロベイビーの実行時の処理

ニューロベイビー実行時の処理を図3.8に示す。まず、ユーザがマイクに向かって発する声を、音声解析部がニューラルネットワークを用いて感情変数に変換する。感情変数を受け取った画像生成部は、それに応じて表情を生成し、アニメーションを表示する。感情変数は、音声出力部にも渡され、生成された表情と同期する音声を発生する。

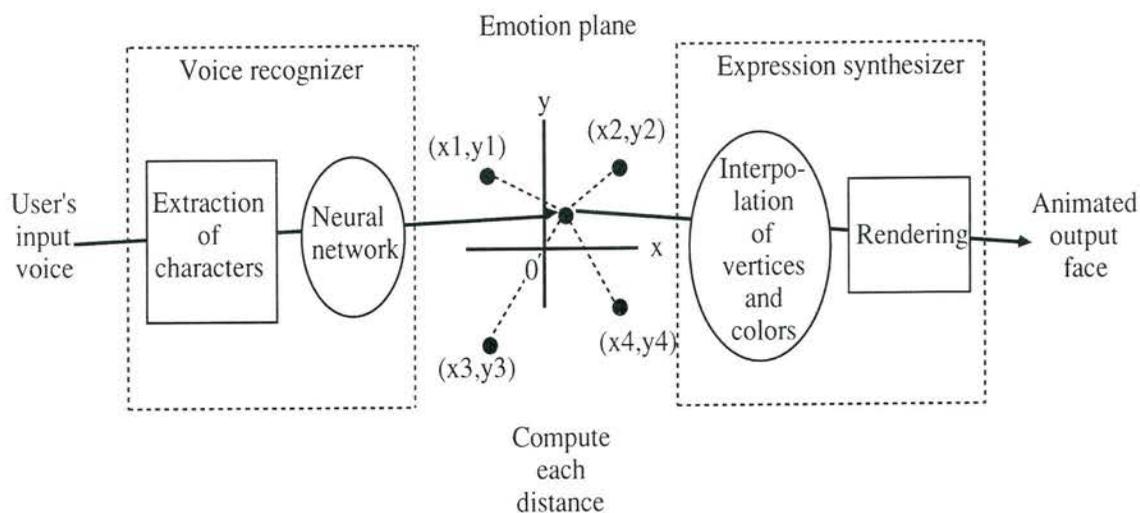


図3.8 処理プロセス

### 3.4.4.1 声の感情認識

音声解析部の処理は、まずマイクから取り込んだ音声波形から特徴量を計算する。これは、図3.4に示した学習段階で行ったやり方と全く同じである。すなわち、10 msec分の波形について特徴量として最大振幅  $a$  とゼロ交差回数  $z$  を求め、これを10回くり返す。

1回の特徴量抽出には約300 msecかかる。特徴量を受け取った学習済みのニューラルネットワークは感情変数  $(x, y)$  を出力する。現状の特徴量計算の方法では、感情変数が急激に変化する場合がある。これをそのまま画像生成部および音声出力部に渡すと、顔の表情や選択される音声も急激に変化して、不安定な結果となる。これを防ぐ対策として一種の緩衝効果を間にいれる。具体的には、図3.9で示すように1回の感情変数の変化量がある一定の距離  $D$  以内に押さえる。直前に出力した感情変数を  $(x_0, y_0)$  とし、次にニューラルネットワークから得られた感情変数を  $(x_{1t}, y_{1t})$  としたとき、距離  $d$  を

$$d = \sqrt{(x_{1t} - x_0)^2 + (y_{1t} - y_0)^2} \quad (3.1)$$

によって求め、 $d$  が予め定めた値  $D$  よりも大きい場合は出力する感情変数  $(x_1, y_1)$  を、

$$x_1 = x_0 + (D/d)(x_{1t} - x_0) \quad (3.2)$$

$$y_1 = y_0 + (D/d)(y_{1t} - y_0) \quad (3.3)$$

によって求める。

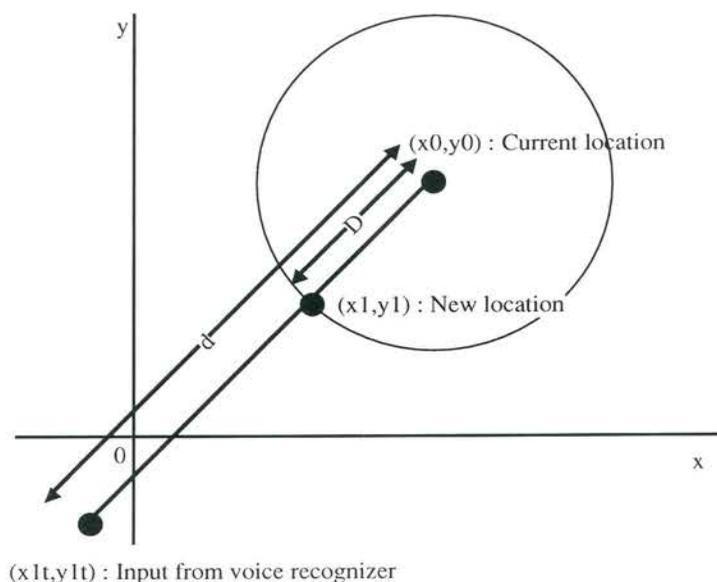


図3.9 感情変数の制限

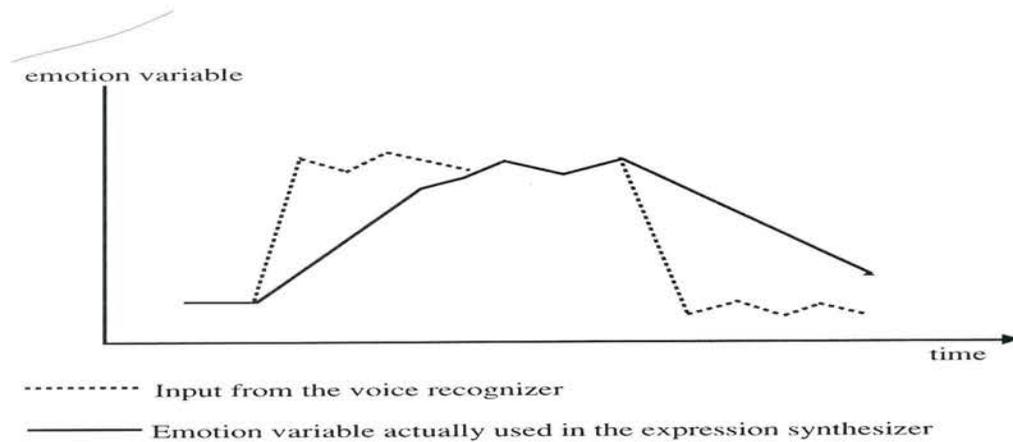


図 3 . 1 0 感情変数の制限の効果

このようにして得られた感情変数は画像生成部と音声出力部に送られる。この緩衝作用の効果を図 3 . 1 0 に示す。横軸は時間方向、縦軸は感情平面内での位置を模式的に 1 次元で示している。現実の世界でも、外部の状況をそのまま感情の変化に対応させる人間はいない。認識された外部の状況とそれによって引き起こされる感情との間には、何らかの緩衝作用があると考えられる。

また、ここでの変化量の制限値  $D$  を変えることによってニューロベイビーの性格を変えることができる。 $D$  を大きくすれば、感情変化の激しい性格になり、小さくすれば感情を表に出さない性格あるいは鈍感な性格になる。

### 3 . 4 . 4 . 2 顔の表情合成

実写の顔表情の合成はいろいろと研究されている[57][58]。筆者はデフォルメされた顔表情の合成を試みた。ニューロベイビーの実行時には、感情変数  $(x,y)$  の値に応じて、実際にデザインした各表情を重み付けして合成したような表情が出力される。 $(x,y)$  から  $(x_i,y_i)(i=1,2,\dots,n)$  への距離をそれぞれ計算し、距離に反比例した重みを付けて各タイプの表情を合成する。実行時における表情の合成は、各感情タイプに対するパーツの頂点座標が、求められた重みに従って合成され、正規化される。このようにして各パーツの形状と位置を求めれば、感情変数の値に応じて補間された表情を合成することができる。この処理を連続して行い、アニメーションとして表示する。この例を図 3 . 7 に示す。しかし、実際には、音声解析部からの感情変数が 1 ~ 2 秒間に 1 回の頻度でしか送られてこない。顔の表情が滑らかに変化するように見せる工夫が必要である。直前に入力された感情変数  $(x_0,y_0)$ 、新たに入力された感情変数  $(x_1,y_1)$  の時、この 2 点を結ぶ線分を  $n$  等分する点を求める。実際の画像生成は、これらの点列に対応して図形配置情報を補間して顔のパーツの形状を求め、点列の順序に沿って連続的に表示すれば、表情は滑らかに変化する。ここで、実際の  $n$  の値としては 5 を用いている。

### 3.4.4.3 効果音の出力

音声出力部では、3.4.3.4で述べたように予めCDからサンプリングした2～3秒間の音声[59]を、感情変数の値に応じて図3.7で示したマップに従って選択して出力する。感情変数が、配置した短形領域に入ったときにその音声を出力するようにしている。

### 3.4.4.4 感情の時間的推移を考慮したイベント生成

以上のような処理を行うことによって、ニューロベイビーは一通りの変化のある表情を生成することができる。しかし、この手法ではある時点での感情だけによって表情が生成され、感情の時間的な推移が考慮されていない。ある時点における実際の人間の感情は、それまでの感情と大いに関連がある。例えば、「飽きる」という感情はしばらくの間変化が起こらなかった場合に現われてくる感情であり、時間軸を考慮しないで扱うことはできない。

そこで、ニューロベイビーでも感情変数の時間的な推移に対していくつかの条件を設け、その条件に当てはまる場合には、予めデザインした特定の表情や音声をだすことができるようにした。このような反応をイベントと呼んでいる。具体的には試作システムでは次のようなイベントを用意し、図3.5のように配置した。

- (1) 「哀しみ」あるいは「怒り」を含む領域に一定時間以上留まった場合 ---> 泣く
- (2) 原点0を含む領域に一定時間以上留まった場合 ---> 退屈であくびをする

また、感情平面の中間付近に留まっていて一定時間以上音声入力がない場合は次のようなイベントを発生する。

- (3) 中間付近の「哀しみ」に近い部分
  - － シャッキリをする
- (4) 中間付近の「怒り」に近い部分
  - － あいさつをする
- (5) 中間付近で「楽しみ」に近い部分
  - － 口笛を吹く
- (6) 中間付近で「喜び」に近い部分
  - － 笑い転げる

このほか、感情変数の値に関係なく発生する次のような特別なイベントがある。

- (7) 急激な立ち上がりの音声波形が入った場合
  - ー びっくりする

### 3.4.5 ニューロベイビーの汎用性

先に示した多くの例はニューロベイビーの一実現法であり、基本システムを変更せずに多くのバリエーションが実現可能である。具体的には、

- (1) 顔の種類、形状、色
- (2) 感情タイプの決定と配置
- (3) 学習させる音声パターン
- (4) 感情の変化に対する鋭敏さ
- (5) イベントの種類及び発生条件
- (6) 使用する効果音

のような要素を自由に設定することができる。特に(1)の形状モデルは、顔である必要すらなく、幾何学形状のような抽象的のもの。あるいは現実世界に存在する物体をデータ化したものなど、任意の形状を用いて構わない。ここに、ニューロベイビー以外のキャラクターの例3点(図3.11、図3.12、図3.13)に示す。最近盛んに研究されているモデルベーストコーディング[12]で行われているように、実際の人間の顔をモデルとして用いるのも興味深いかもしれない。

このようにして、ニューロベイビーでは、自分好みの外観や性格をもった仮想生物を創造することが可能である。もちろん、優れたものを造るにはアートのセンスと粘り強さが必要である。

### 3.4.6 まとめ

人間の感情モデルを導入した「ニューロベイビー」について述べた。今後の課題として

- (1) 感情モデルの改良
- (2) 学習方法の改良
- (3) 感情認識率の向上と評価法の確立
- (4) 感情平面を複数持つことによる奥深い性格の実現
- (5) 使いやすい製作環境の実現
- (6) 子供の情操教育への応用の検討

などが挙げられる。

ニューロベイビーのおもしろさは、人々とのインタラクションにある。世界中の人々が使える共通言語技術が確立されていない今、インタラクションに役立つ可能性のあるものは、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楽典等の感性言語があげられる。このように、人間と人間を結ぶ感性コミュニケーションエージェントの役割は、これからの時代に必要となってくるだろう。

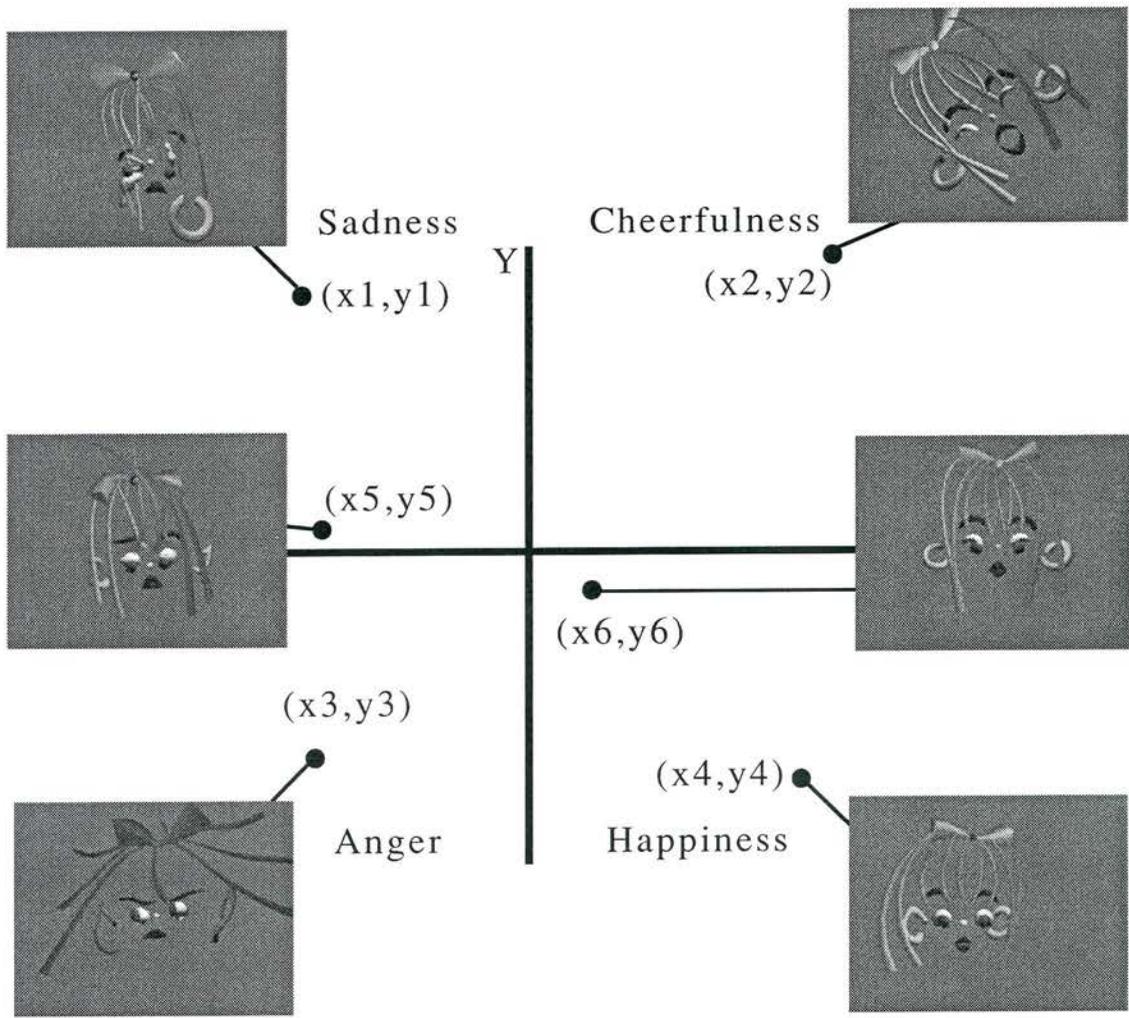


图 3 . 1 1 Neuro-Girl's Emotions

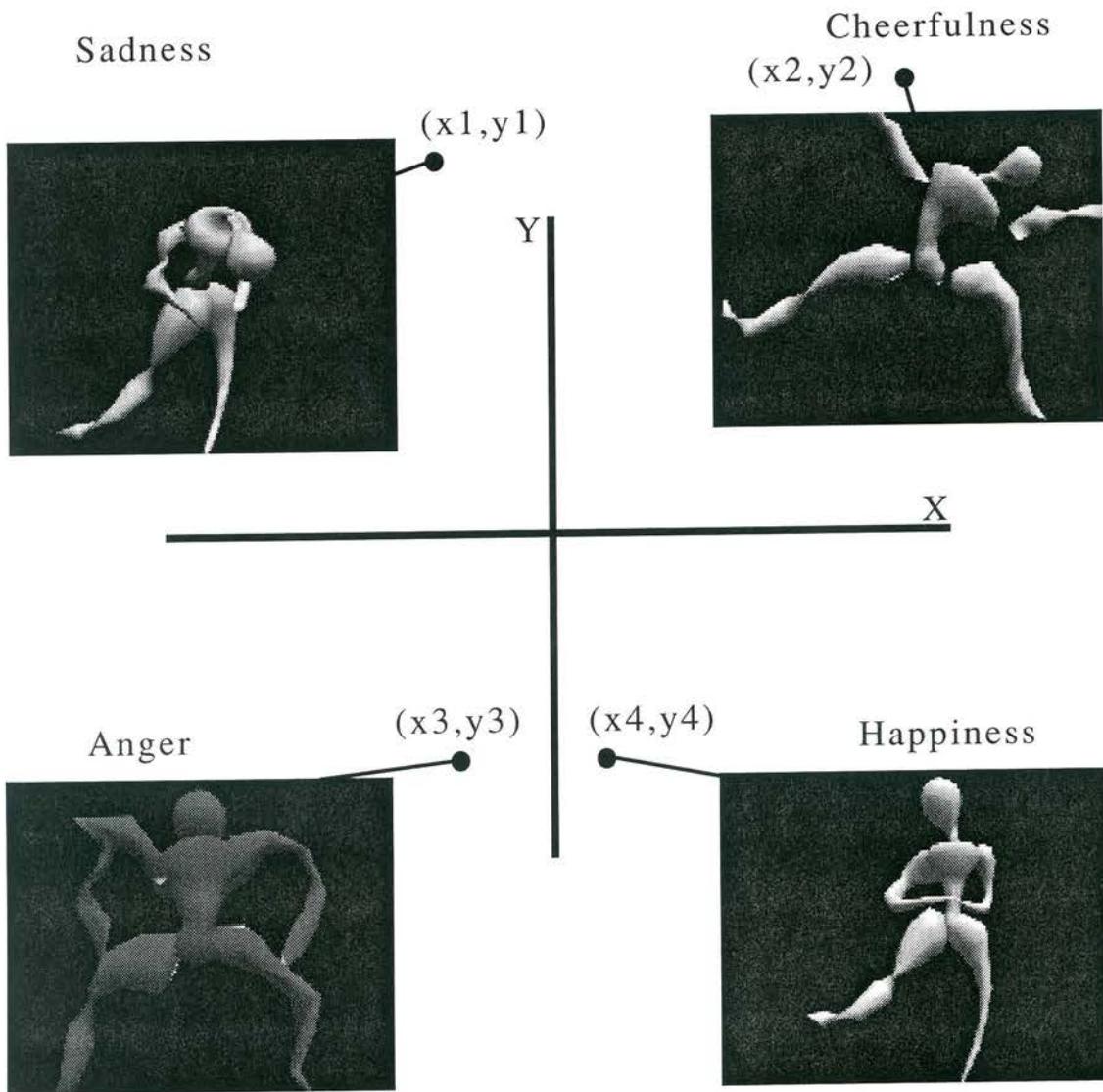


图 3 . 1 2 Emotional Body's Emotions

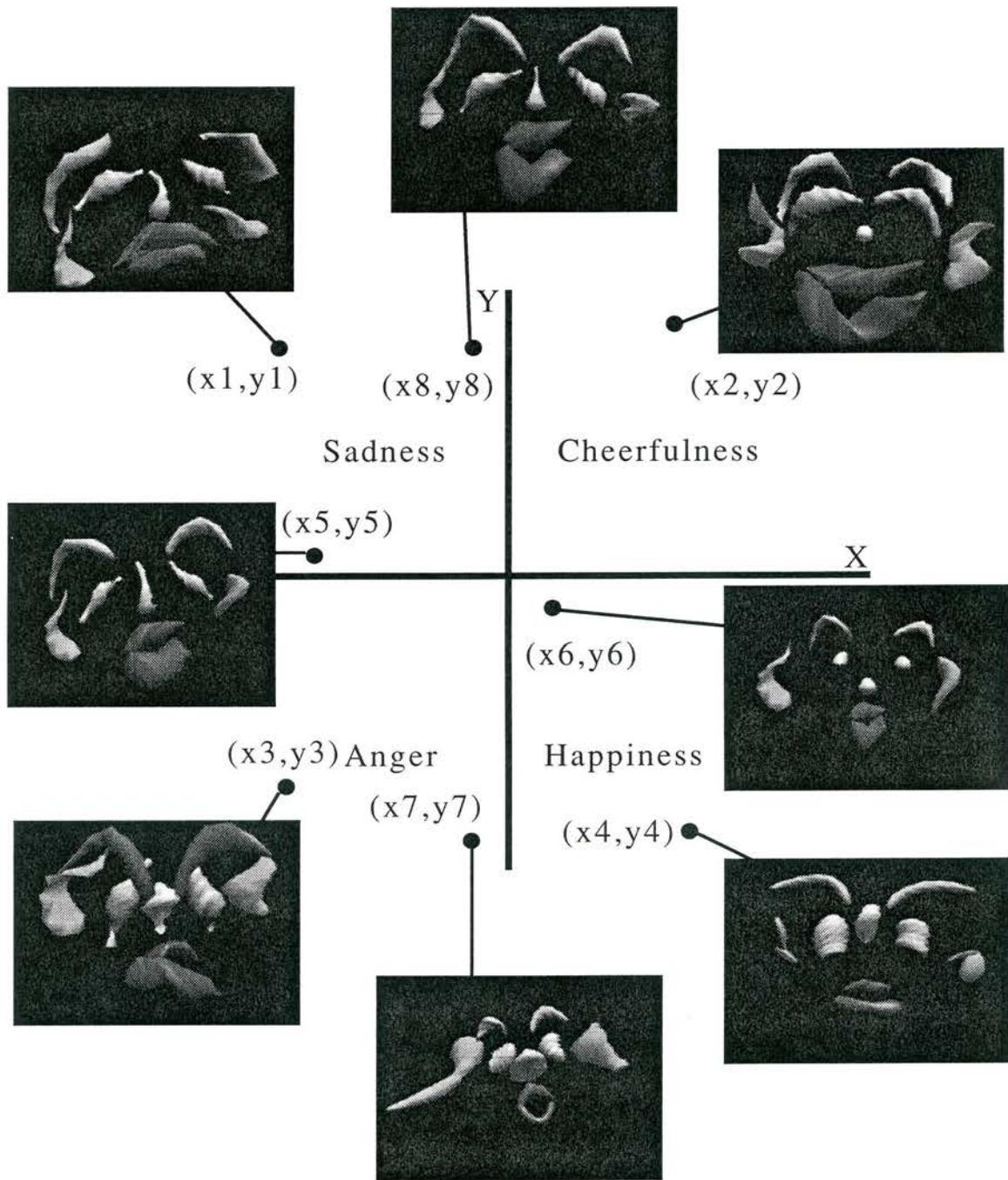


图 3 . 1 3 Neuro-Mask's Emotions

## 4 感情を用いた異文化間のコミュニケーション

### 4.1 まえがき

従来、異文化間のコミュニケーションでは文化の違いから相互理解がままならぬことがしばしばあった。一般的にそれは「外国語の会話能力」であることを連想されるが、実は、言語を超えるさまざまなコミュニケーションの問題を抱えているといえよう。その中でも非言語メッセージは、文化によって、特に会話中に現れる差異や、生活様式における行為に差異がある[1]。筆者らは、これらの点に注目し、感性エージェントに感性翻訳させるという試みを研究した。

### 4.2 異文化間のコミュニケーションにおける感情の役割

文化とは文化人類学者、社会学者、心理学者らによって定義されているが、それらに共通な考え方として、価値観・信念・態度・自己概念が文化の心理的側面を形成することが指摘されている。以上のことから、心理的側面が文化によってかなり違い、重要な価値を占めていることがわかる。一方、人間がコミュニケーションをする時は各人はそれぞれの文化を有しており、両者が同一であれば、最小限の誤解の範囲ですむが、両者がまったく異なった観点からみているとすれば、誤解の可能性が大きくなる。異文化間コミュニケーション過程において各コミュニケーターの内部ノイズの要因がコミュニケーションを妨げているといえる。すなわち、文化はコミュニケーションにおける

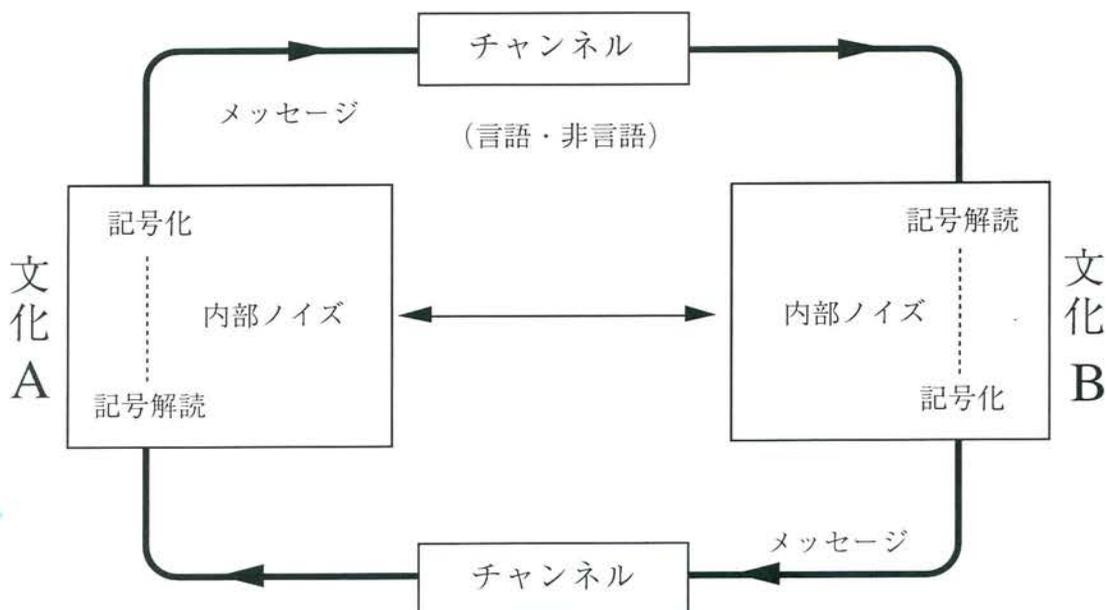


図4.1 異文化間コミュニケーション

「ノイズ」と考えることができると共に、ノイズの中で感情の割合は大きいといえる。

両者がこの自文化中心の姿勢をくずさないかぎり、相互の意志疎通は成立しない。この意味でも、適切な感情表現は、意志の疎通を成立するために大きな役割を担うことがわかる。

### 4.3 コミュニケーションの仲介役としての感情キャラクター

コミュニケーターの間で共通の言語コードが存在しない場合、あるいは通訳者がいない場合、非言語メッセージが唯一のチャンネルとなる。特に、共通の言語があっても、片方または双方のコミュニケーターがその言葉を得意としない場合、ノンバーバルコミュニケーションへのウエイトが増すことになる。以上のことから、ノンバーバルコミュニケーションを理解する感情キャラクターが異文化間のコミュニケーションに役にたつことがわかる。

この場合、感情キャラクターは、非言語対話のインタフェースとして用いられる。具体的な例をあげると、日本人の利用者は初対面の人に対し、顔の表情などが一定で感情をあまり表にださない。しかし、声の調子は好意的である場合、米国の感性キャラクターが好感度を増幅して米国人に伝えるのである。逆に米国人の利用者がそれを見て表した感情データを日本の感情キャラクターに送られる。それぞれの感情キャラクターは利用者の国の国民性に対応した反応を相手国に伝える。それぞれの非言語対話が翻訳されて伝えられ異文化間の対話が可能となる。

### 4.4 感情キャラクター「ネットワークニューロベイビー」[60]

#### 4.4.1 コンセプト

ネットワークを用いて日米間を接続した両端にニューロベイビー[62]を置き、外国人との非言語対話のインタフェース（感情を翻訳するキャラクター）として用いる。国や民族ごとの感情タイプを正しく設定することにより、図4.1に示すように、それぞれの国の人の感情表現が相手の国の感情に翻訳されて伝えられることにより、文化を越えた感情コミュニケーションができる。これは、文化の異なる人々間の相互理解を深める新しいコードになる可能性を秘めている。

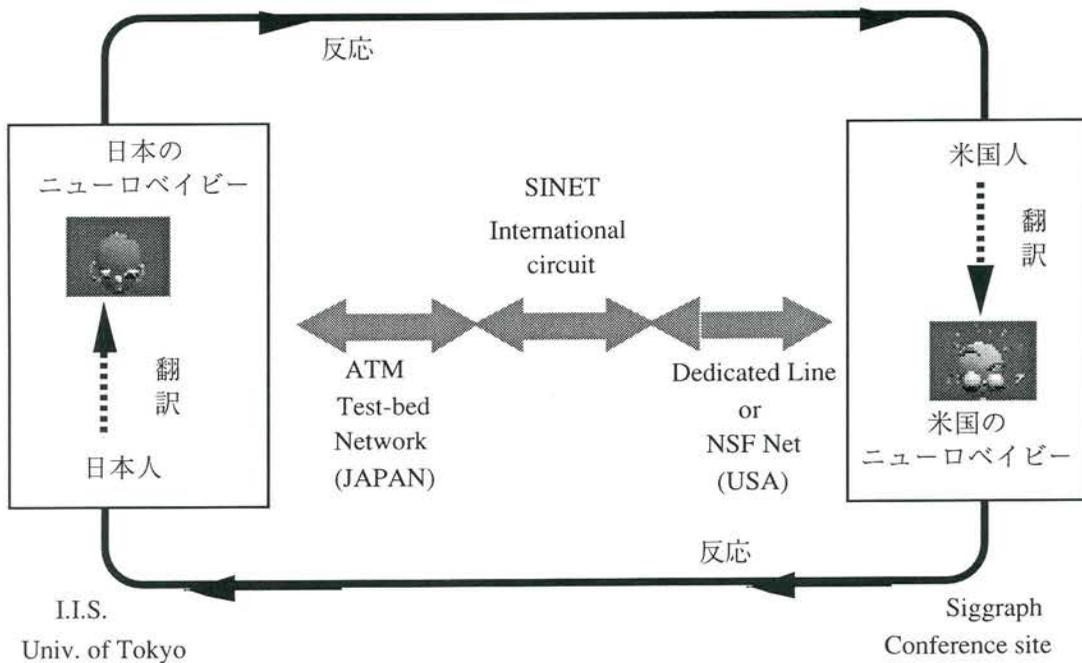


図4.1 ネットワーク型ニューロベイビー

#### 4.4.2 ネットワーク型ニューロベイビーシステム[61]

米国ロスアンゼルスでのSIGGRAPH'95の会場と、共同研究先の東京大学生産技術研究所を結んで公開実験を行った(図4.3, 図4.4)。双方の会場にニューロベイビーのシステムを設置し、通信回線にてリアルタイムに感情のやりとりを行う。図4.2に処理プロセスを示す。本システムでは、新たにハンドシェイクマシン(握手を感知するセンサーデバイス)で、握り方の強さをデジタルデータ化して、感情の分析に用いる(図4.5, 図4.6)。又、アイトラッキングの装置(図4.7)は、ニューロベイビーの視覚となり、相手のいる方向に向かって顔を向けるようになる。図4.8にネットワークを示す。実験では、日本の利用者の感情データが米国のニューロベイビーに送られ、逆に米国の利用者のデータが日本のニューロベイビーに送られた。それぞれのニューロベイビーは、利用者の国民性に対応した反応(表情と握手)を相手に伝えた。

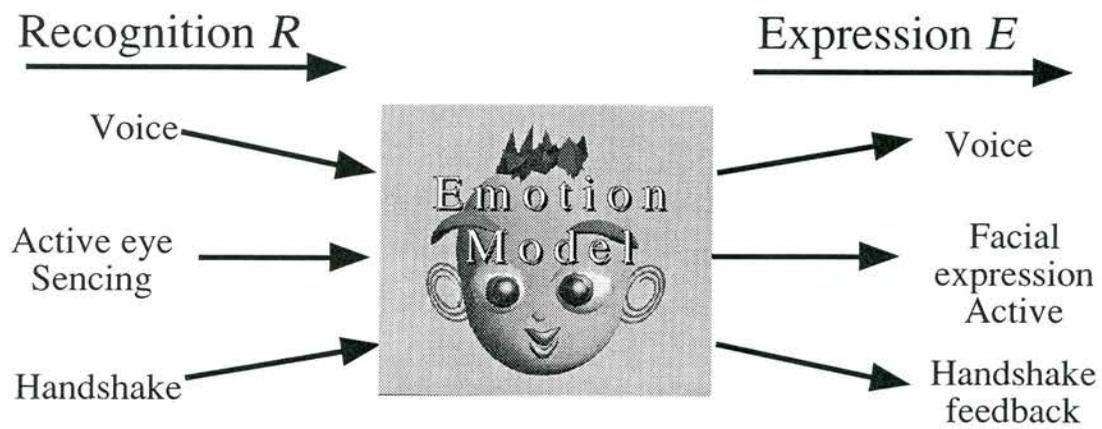


図4.2 ネットワークニューロベビーの処理

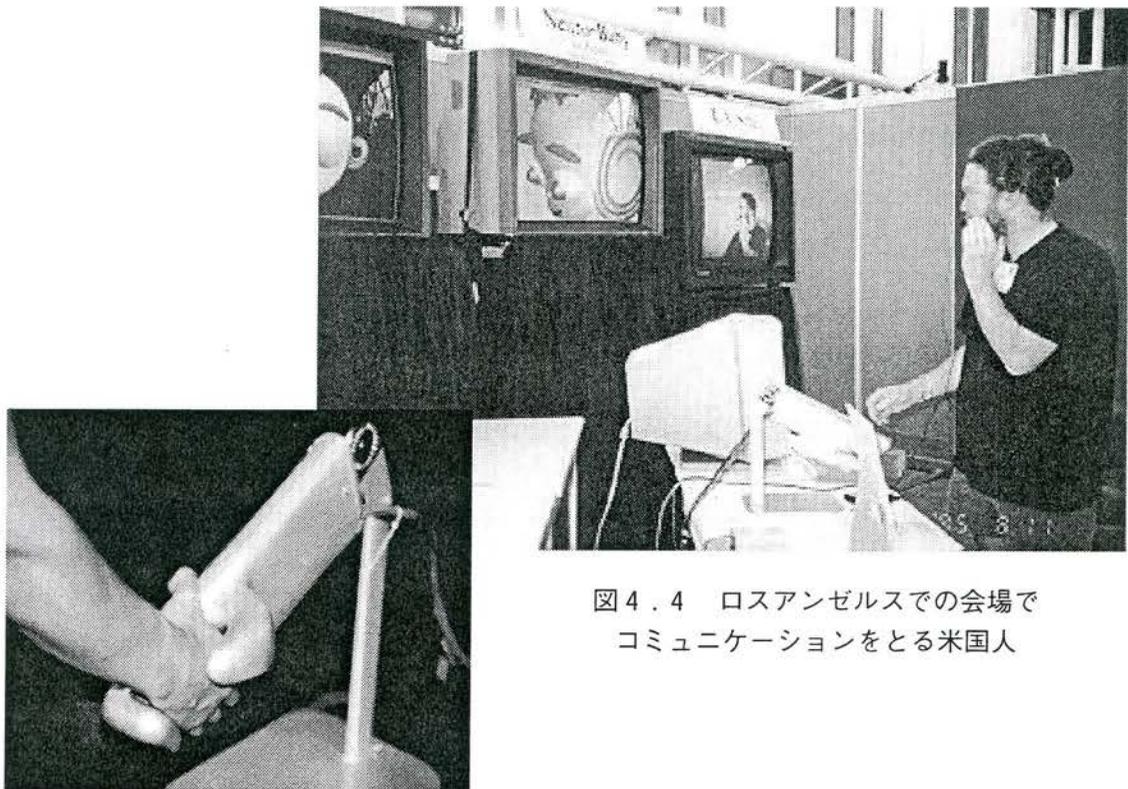


図4.4 ロスアンゼルスでの会場で  
コミュニケーションをとる米国人

図4.3 握手マシンと握手する

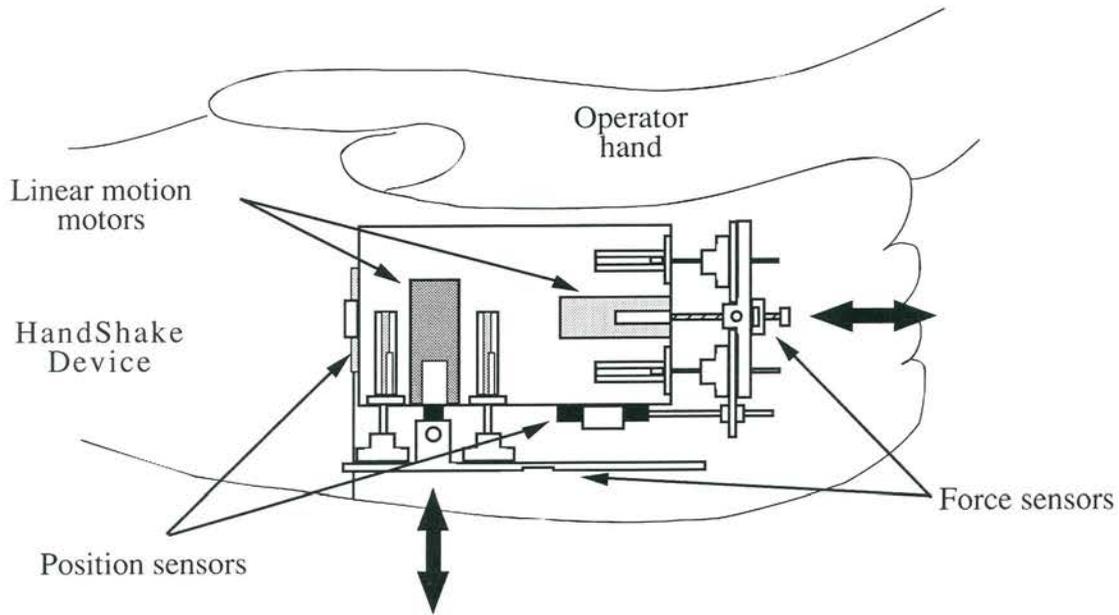
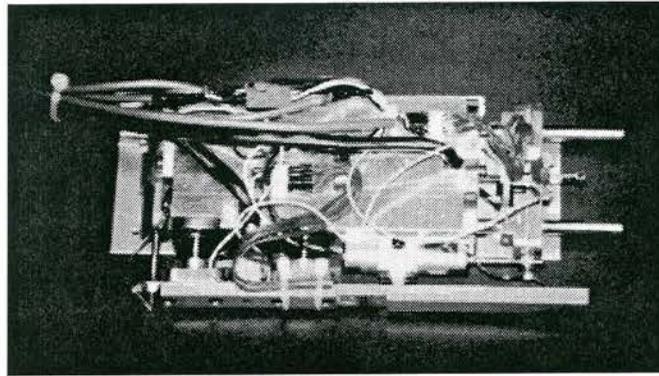


図 4 . 5 握手デバイスの設計

#### 4 . 4 . 3 握手デバイスシステム [63][64]

ニューロベイビーの手になる握手デバイスは、図 4 . 5 に示すように、FRP 製のカバーに包まれて、手の中に配置されている。握られる部分はマスター部、指の部分がスレーブ部である。マスター・スレーブの各システムには力センサーが組み込まれており、話者の握手の強さや、他国から送られてきた握手の強さをフィードバックして握り返すことができる。ここで得られた握手の強さのパラメータは直接デバイスを通して、感情モデルに送られ、CGを通して話者にフィードバックされることになる。

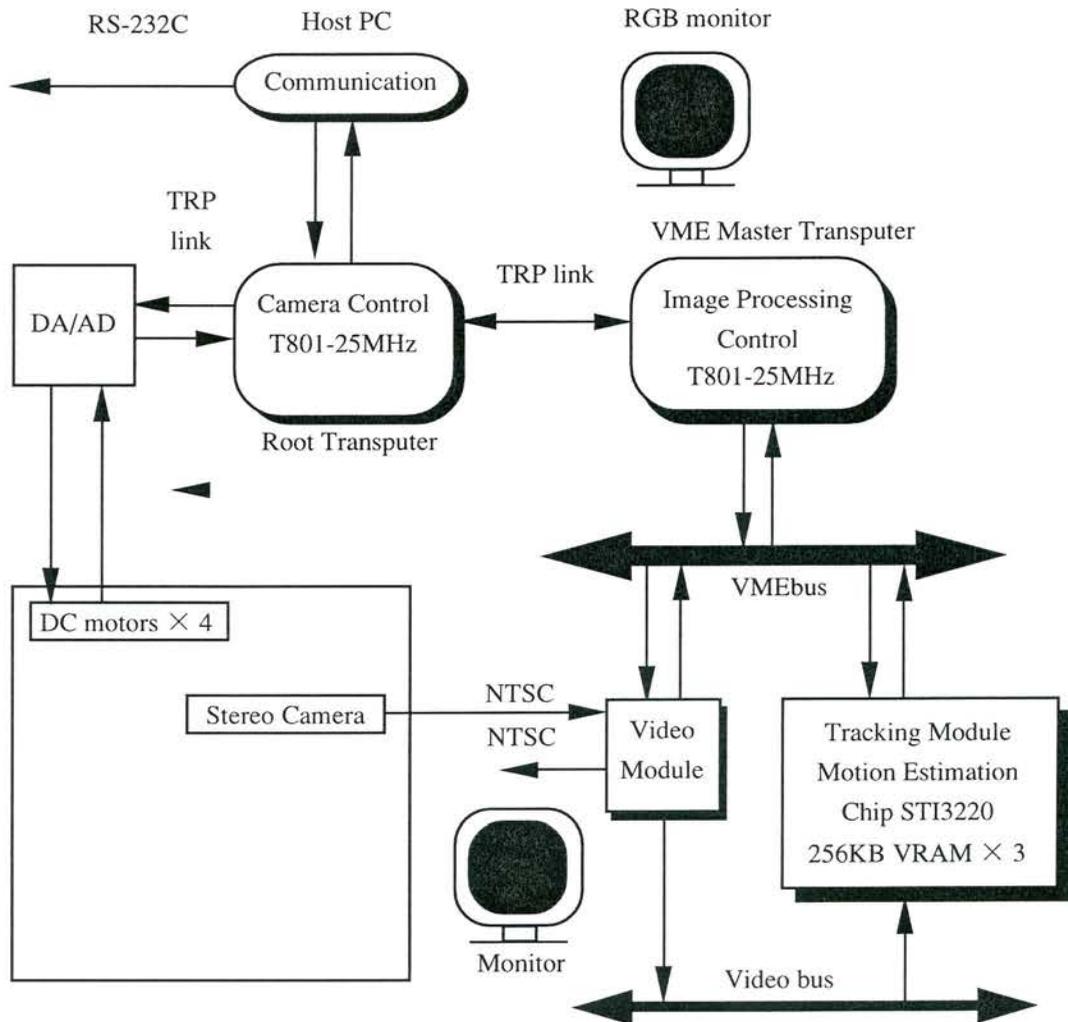


図 4.7 アイトラッキングシステム

#### 4.4.4 アイトラッキングシステム[64]

従来のアイトラッキングシステムは、静止したものや静かな動きを扱っていたが、このシステムでは、実時間で移動物体の運動パラメータを抽出することを目指している。自国のニューロベイビーとインタラクションしている話者を忠実にトラッキングし、計測を続ける視覚システムで、ニューロベイビーの視界となる。抽出された話者の位置は、ニューロベイビーのコントローラーへ送られ、ニューロベイビーは、話者の方を忠実に向くことができる。

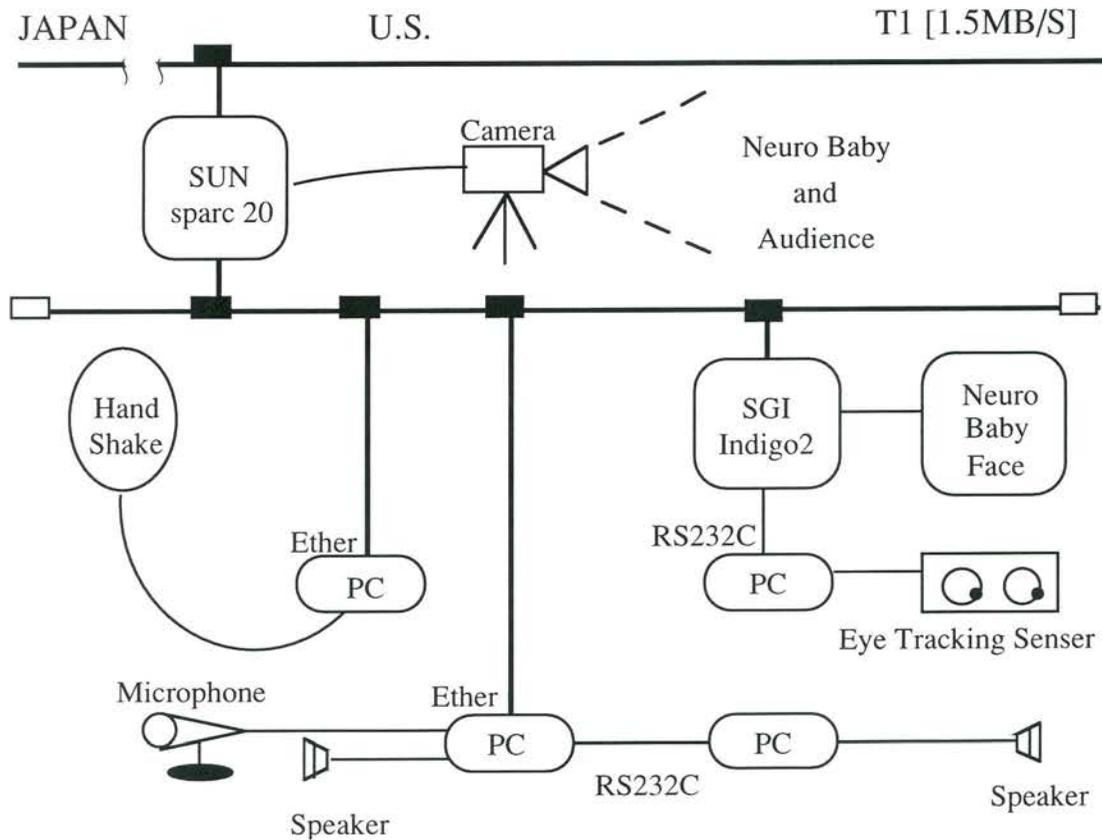


図4.8 ネットワークシステム

#### 4.4.5 ネットワークシステム

各ニューロベイビーのデータと共に、各会場にカメラを設置し、その状況映像をM-JPEGで送り、テレビ電話を同時に行なう。回線は、東京大学生産技術研究所(IIS)から学術情報センター(NACSIS)を通じて、米国ストックトンを経由して会場につなげる。IISとNACSIS間はATMを用いた70Mbpsの専用回線を用い、NACSIS、ストックトン間はインターネットのひとつであるSINET(6Mbps)を使用する。図4.8にネットワークを示す。

## 4.5 まとめ

非言語による異文化対話を実現する感性キャラクターを用いるコミュニケーション実験で、さまざまな文化的相違による入力の問題点が発見できた。例えば、握手マシンへの入力の違いである。日本人にとって握手は手を握ることであるが、米国人の場合、強く握った手を上下に振ることが多い。実験に使った握手マシンは、こうした入力を想定して設計していなかったため、トラブルを生じたり、ニューロベイビーを泣かせてしまったりということがあった。一方、ニューロベイビーの視覚となったアイトラッキングと、声の抑揚に基づいた異文化対話は、ある程度のコミュニケーションが成立していたといえる。触覚フィードバックは遠隔地でも身近に感じられるノンバーバルインタフェースなので、これを視覚情報、音声情報によるコミュニケーションと融合させることは、相手に対する親近感が増し、非常に有効であると考えられるし今後の課題であろう。非言語による異文化対話という研究を実現するためには、人間という複雑系の分析・理解を行ないつつ、文化の差を埋めていかなければならない。それは長い道のりであるが、今後の研究として極めて大きな意義を持つ。

## 5 マルチモーダルに感情を認識・生成する 擬人化キャラクターの生成

### 5.1 まえがき

筆者らが、ニューロベイビーで実現しようとしたものは、音声に含まれる感情によるコミュニケーションである。感情によるコミュニケーションは、非言語コミュニケーション [65] という広い文脈の中でとらえることができる。筆者らは、ニューロベイビーをベースに感情音声とジェスチャーによるマルチモーダルな非言語コミュニケーションによる人格を持つコンピュータキャラクター「ミック」の研究を開始した。

### 5.2 マルチモーダルな感情認識と生成

人間は感情を表現する時、ひとつのサインだけでなく、さまざまなノンバーバルな出力を総合的に行なう。それは時に複雑であり、「顔で笑って心で泣いて」という場合や、好意的な言葉を言っているが、表情は堅いというケースがある。コンピュータによる感性エージェントも、より自然な感情表現を生成するため、そして人間が感情移入しやすいように、複数の感覚の入出力があることが望ましい。システムとしては、聴覚は音声認識、視覚は画像認識・文字認識、触覚はVRによる触覚認識などを用い、これらをバランス良く統合してリアルタイムに感情を生成できるシステムであることが望ましい。マルチモーダルな感情ではないが、マルチモーダル・インタフェースの実現例としては、音声認識、簡単な発話理解、磁気センサーを用いた手による指示動作の認識、大型ビデオディスプレイなどを組み合わせたMITメディアラボのPut-That-There[66]や、キーボードによる自然言語入力、ビデオディスクによる画像情報、タッチパネルによる指示動作を統合したイタリアのIRSTで開発されたフレスコ画の情報検索システムAlFresco[67]がある。さらにNECの島津らは、自然言語入力と指示入力が混在するマルチモーダル入力を解析する汎用の文法記述形式として、マルチモーダル DCG(Define Clause Grammar)を提案している[68]。ほとんどの場合、指示動作はマウスやタッチパネルによって入力されるが、アイトラッキングや顔画像の処理によって、視線を認識することによっても入力される[69]。

### 5.3 マルチモーダルな感情モデルのデザイン

マルチモーダルな感情モデルをデザインするには、言語行動とノンバーバル行動の両者を用いる必要がある。しかし人間は時に、ノンバーバル行動と言語行動にときどき矛盾が見い出される。一貫していれば、その人は誠実で信頼できる人だと推論されるが、一貫していなければ、その人は非誠実で虚偽的だと推論されるのである。心理学者のローテンバークらの研究によると、実際に、真実を話すときより、嘘をつくときの方が言語行動とノンバーバル行動の間により大きな非一貫性が示されることが見い出されている。そして、ノンバーバル行動によって本当の気持ちが「漏れ出す」と考えられている。尚、本当の気持ちが漏れ出しやすいのは視線や表情といったコントロールしやすいノンバーバル行動よりも、むしろ手足の微妙な動きに表れることが見い出されている[26]。具体的な複合感情モデルの設計としては、まず複数の入力を用いて感情表現のずれを表現することが可能だろう。例えば、動作認識によりおとなしく座っているが、顔表情は怒っているなどの複合感情を取り扱うことが可能となる。筆者は、このようなマルチモーダルなインタラクションをする方法でマルチモーダルな感情の認識・生成を設計することにした。

### 5.4 音声とジェスチャーから感情認識する マルチモーダルキャラクター「ミック」

#### 5.4.1 コンセプト

「ミック」[70][71]はニューロベビーの進化形で、以下の点でニューロベビーから進歩している。

##### (1) 非言語コミュニケーションの能力の増大

ミックはニューロベビーと同様に感情に反応するキャラクターであるが、認識出来る感情を4つから8つ（喜び、怒り、驚き、悲しみ、愛想をつかす、からかう、怖れる、普通）に増やした。

##### (2) キャラクターの反応パターンの高度化

全身像のCGを作り、単に顔の表情による反応だけではなく、全身でのアクションにより反応である感情表現をすることにした。

##### (3) 感情認識機能の高度化

8つの感情の各々に対応したニューラルネットを用意し、これらを並列に設置することにより、感

情を認識する方法を採用し、認識性能の向上を図ると共に、音韻バランスのとれた多数単語を多数話者が種々の感情で発声した音声データを訓練データとして用いることにより、不特定話者・内容独立型の感情認識を可能とした。

#### (4) マルチモーダルコミュニケーション

音声によるコミュニケーションの機能に加え、画像認識機能を取り入れることにより、人間の簡単なジェスチャーにミックが反応する機能を追加した。

### 5.4.2 取り扱う感情

ミックは、8つの感情を声の抑揚から認識し、それに反応できる。例えば、だれも話しかけないと居眠りをし、だれかが話しかけると、機嫌が良いときは、「こんにちは」、悪い時は、「バイバイ」と返事をする。低い声でばかりにすると怒り、からかうと逆立ちをする。口笛を吹いてあげると、エキサイトしてジャンプをし、人間が不機嫌な顔をして咳払いをすると、悲しくなって手で顔を覆い後ろをむいてしまう。たまに愛想をつかし人間に愚痴をこぼす。ミックの感情モデルは、図5.1の2次元座標に表わされる。X軸は快・不快、Y軸は感情の強弱である。

具体的な8つの感情および、それに対応した入力音声の特徴（\*で示した）は以下の通りである。

#### (1) Joy (幸福, 満足, 楽しい心地よい, 笑顔)

\* 高く強い声, 口笛など。

#### (2) Anger (立腹憤り, 不機嫌)

\* 低く強い声で叱る。

#### (3) Surprise (驚く, ショック, びっくりする)

\* 突然、大きな声で叫ぶ。

#### (4) Sadness (哀しい, 泣く, 寂しい)

\* 低く弱い声, 咳払いなど。

#### (5) Disgust - - - \* 愛想をつかした疲れた声。

#### (6) Teasing - - - \* 高く軽い声でからかう。

#### (7) Fear - - - - \* 低く低い声で、怖がらせるような声。

#### (8) Neutral - - - \* 普通の声。

### 5.4.3 処理プロセス

本章では、音声に含まれる感情の認識の基本方針と処理の内容について述べる。また、認識結果の感情に対応したミックの反応パターンの生成法についても述べる。

#### 5.4.3.1 基本方針

ミックは、その原形であるニューロベイビーから何段階かの変遷を経て現在の形に成長した[51][60]。現在のミックでは、人間とエージェント間のより洗練されたインタラクションを実現するために必要な技術を開発することを狙っている。このために、以下に述べる項目について検討した。

##### (1) リアルタイム処理

人間とエージェント間のリアルタイムのインタラクションを実現するためには、音声特徴抽出・感情認識・応答生成の各部分をリアルタイムで動作させる必要がある。本システムでは、音声処理およびシステム構成の部分の構成をリアルタイム向けにすることにより、リアルタイム処理を実現している。

##### (2) 種々の感情の取り扱い

何種類のかつどのような感情を認識対象とするかは、重要なかつ困難な問題である。これまでに、感情を扱った種々の研究がなされているが、そこで扱われている感情の例を以下に示す。

- a. 怒り、悲しみ、幸せ、喜び[51]
- b. 普通、喜び、退屈、悲しみ、怒り、恐れ、からかい [72]
- c. 怒り、恐れ、悲しみ、喜び、軽蔑 [73]
- d. 普通、幸せ、悲しみ、怒り、恐れ、退屈、軽蔑 [74]
- e. 恐れ、怒り、悲しみ、幸せ [75]

我々の以前の研究では、aに示してある4つの感情を扱った [51]。最初のバージョンのニューロベイビーを各種の機会に展示した経験、および感情の種類が増えることが人間とエージェントのインタラクションの内容の豊かさにつながるという考察の上に立って、研究では、5.4.2で述べた計8つの感情を扱うこととした。

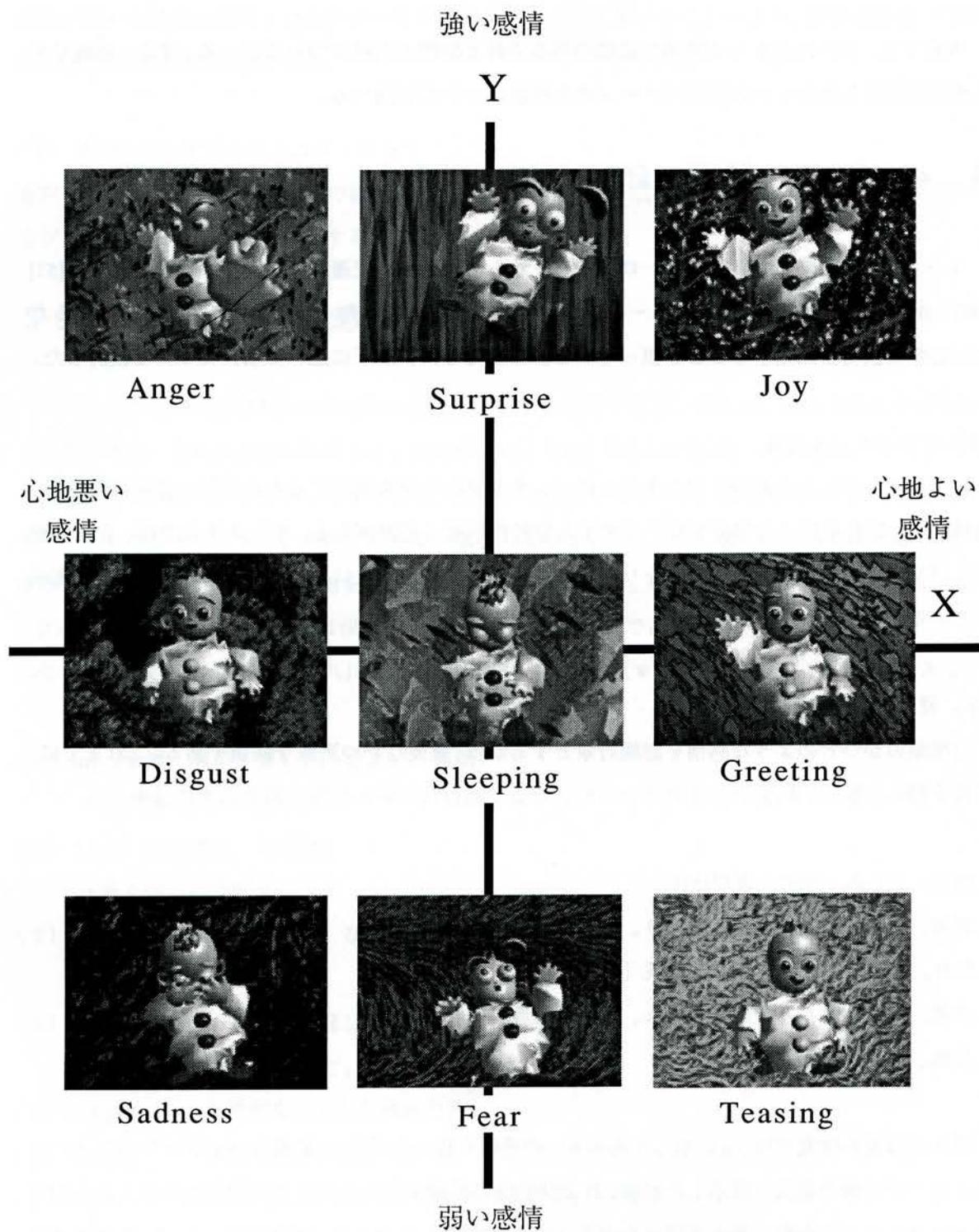


図 5 . 1 ミックの代表的な感情表現

### (3) 精密な音声処理

感情認識のためにどのような音声特徴量を用いるかは重要な問題点である。1つの考え方として、感情認識に用いる音声特徴量は音声認識に用いる音声特徴量とまったく異なるようにすべきだというものがある。これは、音声認識においては、周波数スペクトラムなどの音韻特徴が重要な役割をしているのに対し、感情認識においては、音声の韻律特徴が主要な役割をしているという一般的な考え方に基づいている。しかしながら、また別の見方がある。すなわち、音声を発声する場合は、音韻特徴と韻律特徴は分かち難く結び付いており、韻律特徴の制御のみによって感情を表現することは困難であるため、韻律特徴と同時に音韻特徴も用いるべきであるという考え方である。本研究では後者の立場に立ち、2つの特徴パラメータ、すなわち音韻特徴を表現するパラメータと韻律特徴を表現するパラメータを両方用いることにした。

### (4) 不特定話者コンテキスト独立型感情認識

音声認識や感情認識において、不特定話者を扱えるというのは重要な機能である。実用的な観点からすると、話者が変わる度に大変な学習処理を行う必要があるというのは望ましくない。別の観点からすると、人間は不特定話者の声に含まれる意味内容と感情を同時に認識出来るという事実がある。また、感情認識においてはコンテキスト独立性は重要な機能である。われわれの日常の会話では、同じ単語や文章に異なった感情を乗せることがひんぱんに生じる。本研究では、認識のためのアーキテクチャーとしてニューラルネットを用いると共に、大量の学習サンプルを用いた学習処理を行うことにより、不特定話者に対応できかつコンテキスト独立型の感情認識を実現することをめざした。

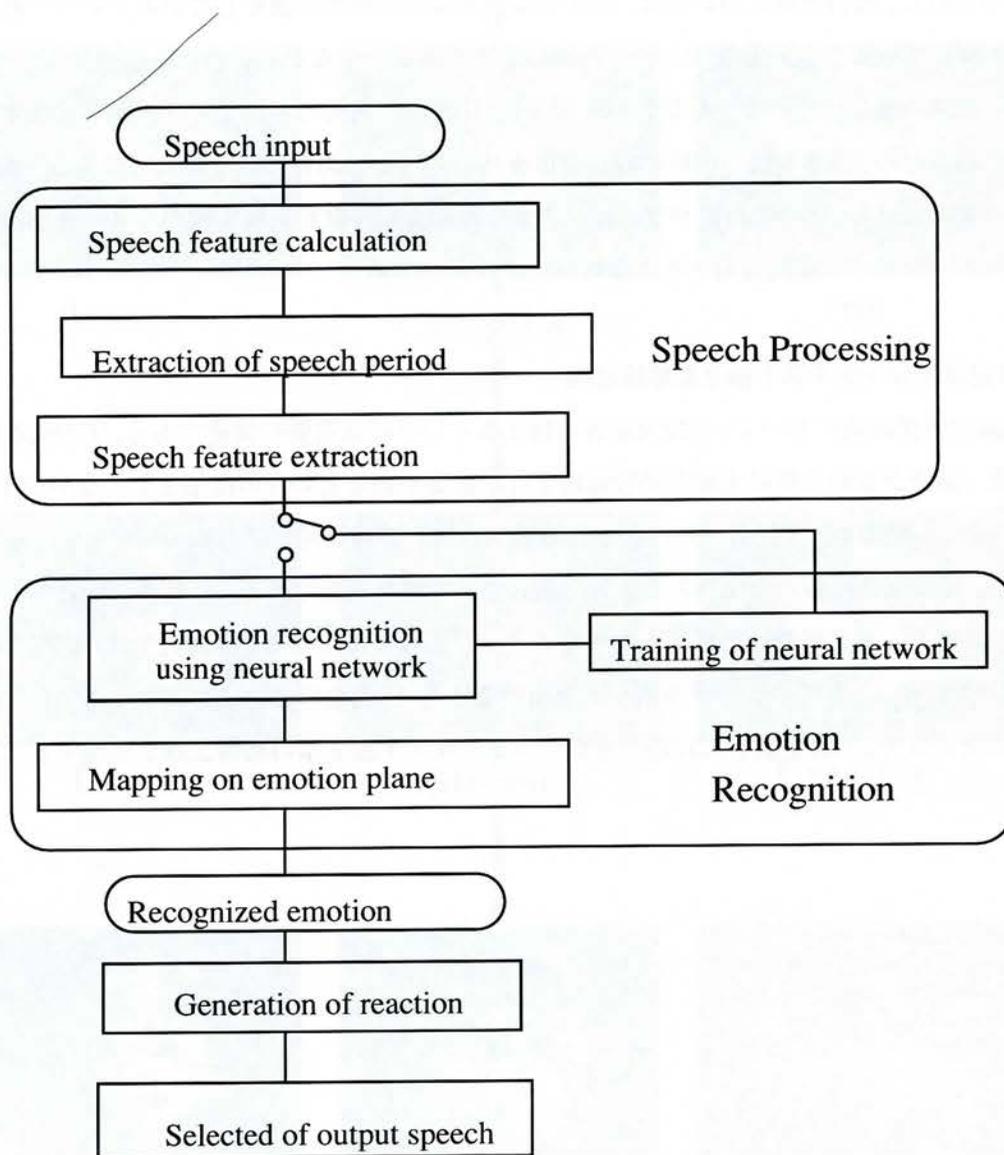


図 5.2 ミックの処理プロセス

図5.2は処理の流れのブロック図である。全体の処理は3つの部分、すなわち、音声特徴抽出部、感情認識部、および反応生成部から構成されている。音声特徴抽出部では、まず入力音声から音声特徴パラメータがリアルタイムで抽出される。次に、音声パワーを用いて音声区間の抽出が行われる。抽出された音声区間を用いて、入力音声のそれぞれに応じた音声特徴量が決定される。この音声特徴量は、感情認識部への入力として用いられる。感情認識部では、2段階の感情認識が行われる。第1段階ではニューラルネットが用いられる。第1段階は、8つの感情のそれぞれを認識出来るよう学習された8つのニューラルネットが並列におかれた構成になっており、音声特徴抽出部の出力が8つのニューラルネットに同時に入力される。次に、第2段階として論理部があり、8つのニューラルネットの出力を論理処理して、2次元の感情平面への写像を行う。2次元の感情平面上には8つの感情をあらかじめ適切に配置しておく。2次元の感情平面上の認識感情の位置およびその動きに応じて、ミツクの反応、すなわち顔の表情と体の動きが生成される。これらの顔の表情および体の動きは、アーティストの直感と感性と経験によって注意深く事前に定めておく。反応パターンはコンピュータグラフィクスで表示されると共に適切な音声もしくは音楽が出力される。

以下に、それぞれの処理の詳細、および感情認識を行うためのシステム構成を述べる。

## 5.4.3.2 音声特徴抽出

### (1) フレーム毎の特徴パラメータ抽出

感情認識のために2種類の特徴パラメータが用いられる。1つは音韻特徴を表わすパラメータであり、もう1つは韻律特徴を表わすパラメータである。音韻特徴パラメータとしてはLPCパラメータが用いられる [76]。LPCパラメータは、音声特徴を表現するための代表的なパラメータであり、音声認識においてよく用いられる。これに対し、韻律特徴としては3種類のパラメータが用いられる。すなわち、エネルギー、音韻特徴の時間変化、およびピッチである。エネルギーおよびピッチに対応するパラメータとしては、LPCパラメータから求められる音声パワーおよびピッチ情報を用いる。また、LPCパラメータの時間変化を表わすパラメータを時間変化パラメータとして用いる。

入力音声の音声特徴パラメータの計算は次の手順で行う。アナログ音声は、まず6kHzの低域通過フィルタを通した後、11 KHz、16bitでサンプリングされ、デジタル音声にy変換される。デジタル音声は、256個のサンプリング点を含むフレームの連続として表現され、各フレーム毎にLPC分析が行われ、以下の特徴パラメータが求められる。

$$\left. \begin{array}{l}
 \text{音声パワー} : P \\
 \text{ピッチ} : p \\
 \text{LPCパラメータ} : c_1, c_2, \dots, c_{l_2} \\
 \text{時間変化パラメータ} : d
 \end{array} \right\} \quad (5.1)$$

tフレームに関する音声特徴パラメータは以下で表現される。

$$F_t = (P_t, p_t, d_t, c1t, c2t, \dots, c12t) \quad (5.2)$$

このパラメータの時系列が音声区間抽出部へ送られる。

## (2) 音声区間抽出

図5.3に音声区間抽出と音声特徴抽出の処理を示してある。音声区間抽出では、音声パワー情報を用いて以下のようにして音声区間を抽出している。音声パワーはあらかじめ定めたいき値  $P_{th}$  と比較される。 $P_{th}$  より音声パワーが大きい区間が連続してある程度続くと音声が存在すると判断される。またその後で音声パワーがいき値より小さい区間がある程度続くと音声を終了したと判断される。背景雑音が音声抽出に大きな影響を与えるため、システムが動作開始した段階で背景雑音レベルを測定し、これに基づいて  $P_{th}$  を定めている。

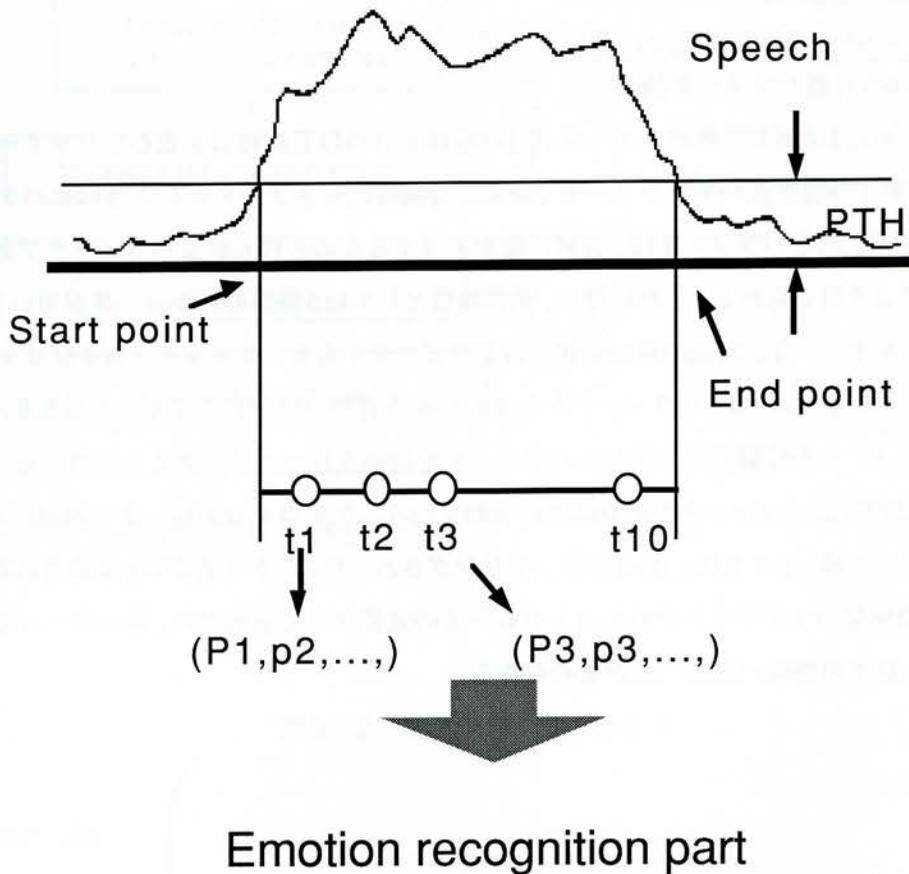


図5.3 音声区間抽出と音声特徴抽出の処理

### (3) 音声特徴抽出

抽出された音声区間の全体から、等間隔になるように配置された10フレームを取り出す。この抽出された10フレームを、 $f_1, f_2, \dots, f_{10}$ とする。これらの10フレームの特徴パラメータをまとめることにより、音声の特徴量は150次元 ( $15 \times 10$ ) の特徴ベクトル、

$$FV = (F_1, F_2, \dots, F_{10}), \quad (5.3)$$

として表現される。ここで、 $F_i$ は、 $i$ フレームの特徴パラメータである。 $FV$ は、感情認識部への入力として用いられる。

## 5.4.3.3 感情認識

感情認識は困難な研究対象である。その主な原因として、人間は赤ん坊の時には感情の認識に基づいて主として行動しているが(赤ん坊は意味内容の理解を始める前に感情の認識ができるといわれている)、大人の場合は主として音声の中に含まれる意味内容の理解に基づいてコミュニケーションを行っているという点にある。このために、音声認識研究においては、長い間音声に含まれる感情の情報を単なる変動もしくは雑音として扱ってきた。さらに問題を複雑にしているのは、人間が発声する音声において意識的な場合も無意識的な場合も含めて、意味内容と感情が分かち難く結び付いている点である。特に、コンテキストは極めて重要な役割をしており、コンテキストが我々が無意識に表出し知覚する感情のレベルを左右していることが多い。このことは言い替えれば、音声に含まれる感情表現の量が状況に極めて依存していることを示している。もちろん、最終的には無意識の感情表現と意味内容表現がミックスされている場合にも感情の認識ができることが望ましいが、当面は無意識的な感情表現を認識することは上記の理由によって困難であると考えられる。したがって、ここでは無意識的な感情表現がされた音声を取り扱うのではなく、意識的に感情表現を伴って発声された音声を取り扱いの対象とする。次に、認識のアルゴリズムであるが、現在音声認識で主として用いられているのはニューラルネットワークとHMM(隠れマルコフモデル)の2種類である。現在の音声認識における主流のアルゴリズムはHMMであるが、ここでは以下の理由によってニューラルネットを用いることとした。

- a) 我々の目的はコンテキスト独立型感情認識である。HMMは、意味内容の認識には適しているが、コンテキスト独立型の認識にはニューラルネットが適していると考えられる。
- b) HMMは、認識対象の構造がある程度明確になっている場合に適している。意味内容の認識の場合は、音韻連鎖という音声構造が事前にわかっているためにHMMが適当である。しかしながら、感情認識の場合はまだ感情の構造が明らかでないため、ニューラルネットを用いる方が適当と考えられる。

(1) ニューラルネットの構造

感情認識のためのニューラルネットの構造を図5.4に示す。このネットワークは8つのサブネットワークの集合とそれらのサブネットワークの出力を統合する論理部から構成されている。8つの各々のサブネットワークは8つの感情（怒り、悲しみ、喜び、恐れ、驚き、愛想をつかす、からかい、および普通）のそれぞれにあらかじめチューンしてある。

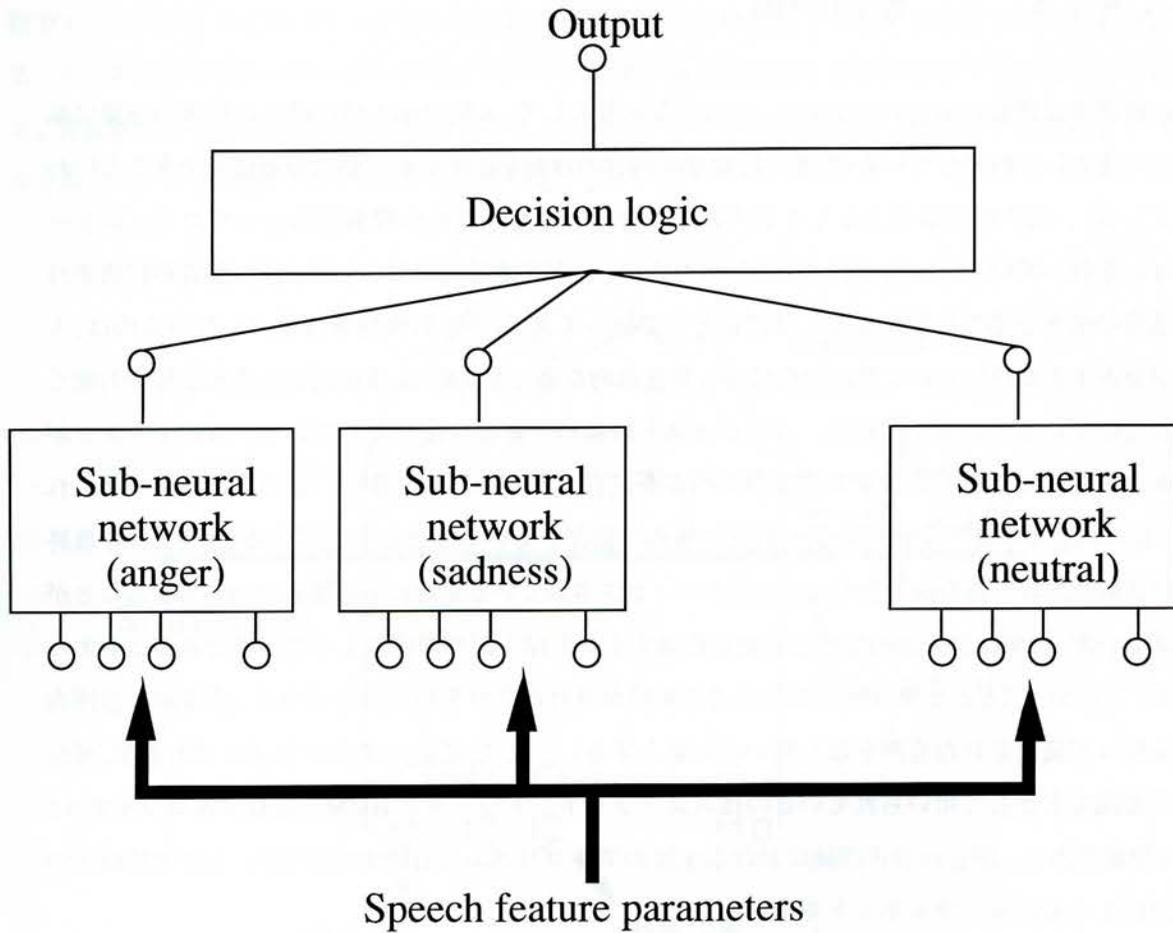


図5.4 感情認識部の構造

各々のサブネットワークの構成を図5.5に示す。基本的には、これらのサブネットワークは同じ構造をしている。すなわち、3層構造であり、入力層は音声特徴量の次元に対応した150個のノードよりなり、中間層は20ないし30のノードからなり、そして出力層は1個のノードから構成されている。感情認識の困難さは、認識すべき感情によって大きく異なっているため、1個のニューラルネットを用意するより、各々の感情に対応したニューラルネットを用意しておいて、これらをそれぞれの感情にチューンした方がいいと考えられるため、このような構造を採用した。このことは予備実験によって確かめられた。すなわち、怒り、悲しみのような負の感情が比較的認識しやすいのに対し、喜びのような正の感情の認識は困難であって、1つのニューラルネットを用いた場合は学習が収束しなかった。また、中間層のノードの数のような個々のサブネットワークの細部構造は、学習が収束しやすいように個々の感情毎に変えてある。8つのサブネットワークの出力を統合して最終的な感情認識結果を得るために、サブネットワークの後段に論理部をおいている。その内容については後述する。

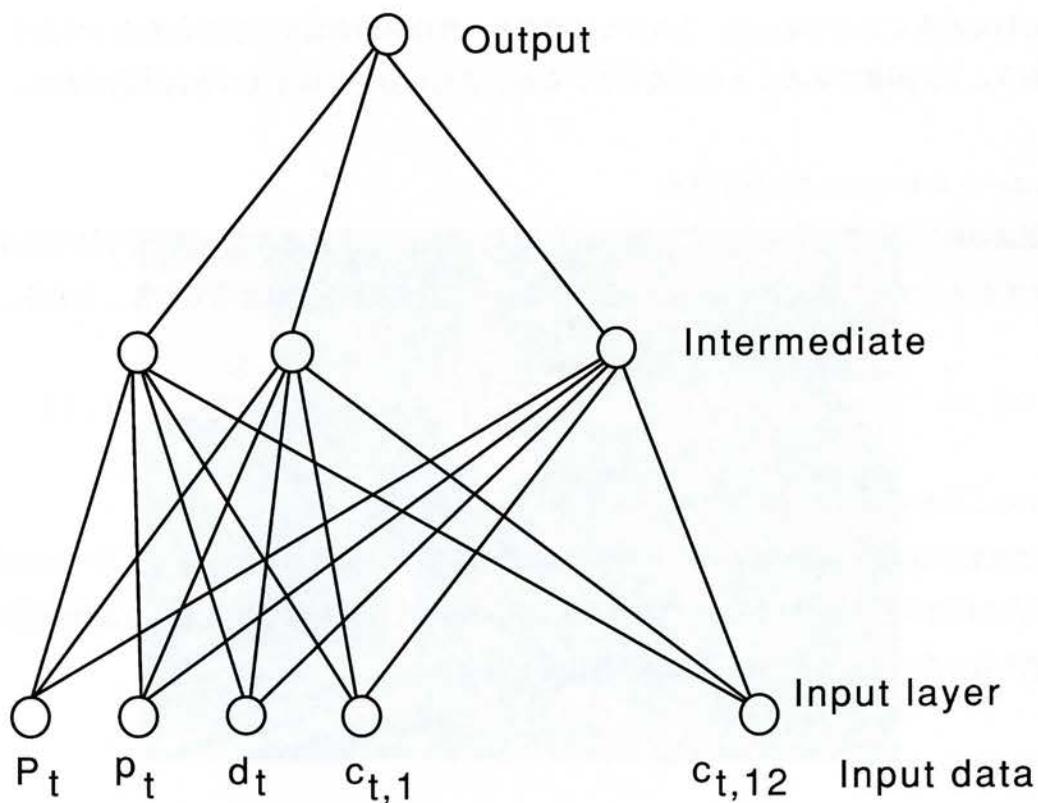


図5.5 各々のサブネットワークの構成

## (2) ニューラルネットの学習

感情認識を行うためには上に述べたニューラルネットをあらかじめ学習させておく必要がある。我々の目標は不特定話者、コンテキスト独立型の感情認識であるため、以下のような音声サンプルを学習データとして用意した。

単語：100個の音韻バランスがとれた単語

話者：10名（5名の男性と5名の女性）

感情：普通、怒り、悲しみ、喜び、恐れ、驚き、愛想をつかさ、からかい

音声サンプル：各々の話者が8つの感情で100個の単語を発声する。したがって、全体で各話者毎に800個の音声サンプルが得られる。

この学習データを用いて種々の予備実験を行った。その結果、男性、女性をまとめたニューラルネットを用意するより、男性、女性それぞれにチューンしたネットワークを用意する方が学習、認識共に有利であることがわかった。このことは男性と女性の感情表現が本質的に異なっており、まとめて扱うことが困難であることを示唆しているが、詳しい分析等は今後の検討課題である。

## (3) ニューラルネットによる感情認識

感情認識の際には、音声特徴抽出部で得られた音声特徴量が、上に述べた方法で学習が行われた8つのサブネットワークに入力される。その結果として、8つの実数値が得られる。これを、

$$V = (v_1, v_2, \dots, v_8) \quad (5.4)$$

と表現することとする。

感情認識部の性能を評価するための小規模の認識実験を行った。8つのサブネットワークの出力の内、最大の出力を与えるサブネットワークに対応した感情を認識結果とする簡単な論理を用いて認識実験を行ったところ、約60%の感情認識結果が得られた。

## (4) 感情平面への写像

上に述べたように、感情認識部の出力はベクトル、 $V = (v_1, v_2, \dots, v_8)$  で表現される。次に  $V$  から最終的な感情認識結果を得ることが必要である。我々のこれまでの研究では、感情認識結果は2次元の平面上の点として表現してきた[2]。これまでの研究における結果・経験に基づき、本研究では感情平面  $E$  の上に8つの感情を図5.1に示すように再配置することとした。

したがって、VからEへの写像を行うことが必要になる。この写像を行う際にもニューラルネットを用いるということが考えられるが、今回は時間の関係もあって、以下のような簡単な論理を用いて写像を行うこととした。

m1およびm2を8つの実数値、 $v_1, v_2, \dots, v_8$ の最大値および第2位の値とする。また、 $(x_{m1}, y_{m1}), (x_{m2}, y_{m2})$ をm1、m2に対応する感情位置であるとする。このとき、最終的な感情認識結果、 $(x, y)$ を次式で求めることとする。

$$x = c * x_{m1} + (1 - c) * x_{m2}, y = c * y_{m1} + (1 - c) * y_{m2} \quad (c: \text{定数}) \quad (5.5)$$

5.4.3.2から5.4.3.3までの処理によって「ミック」の感情認識が行われる。図5.6にユーザーが感情音声によりミックとコミュニケーションしている様子を示す。

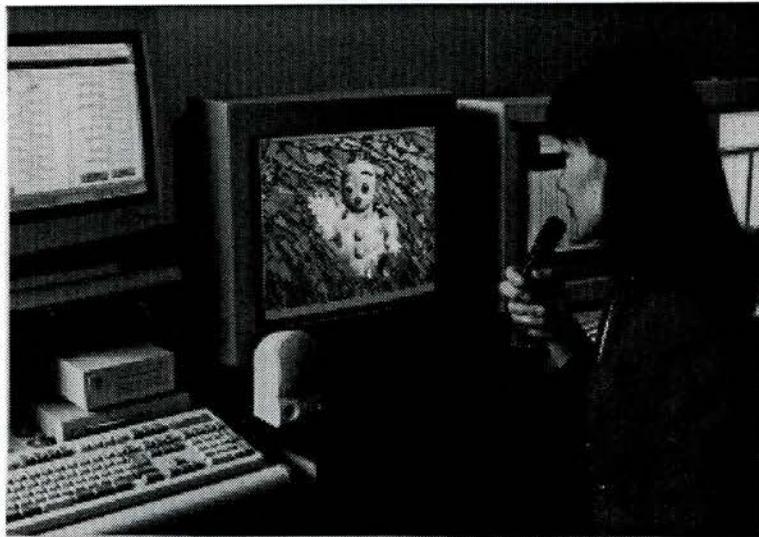


図5.6 ミックとコミュニケーションの例

#### 5.4.4 ジェスチャーによるインタラクション

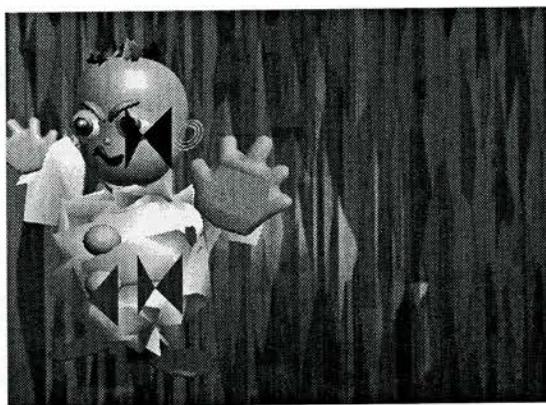
非言語コミュニケーションの機能の一つとして、ジェスチャーによるインタラクションの機能をミックに追加した。カメラからの2次元空間の画像認識では、MITメディアラボで行なわれている Pfunder Sytem[77]がある。筆者は、それを応用してシステムに組み込んでいる。人が手を伸ばしたりするボディアクションで感情を表現すると、ミックは人と同じ感情を共有しようとして同じように手を伸ばしたアクションをする。人がミックに近づくと、時にはさわられるのを嫌うかのように逃げ出したり、また時には人とのふれあいを求めてすりよってくる。また、人はスクリーンの上の蝶や蜂に変身してミックとのインタラクションを楽しむことも出来る。例えば観客が蜂になってミックを刺そうとしたとしよう。図5.7の様に、彼は、蜂から逃れようとし、刺されるといかにも痛そうな表情をする。最後に図5.8にシステム構成図を示す。



観客が蜂になってミックを刺そうとする



蜂から逃げようとするミック



蜂から刺されて痛そうな顔をするミック



図5.7 ジェスチャーによるインタラクション例

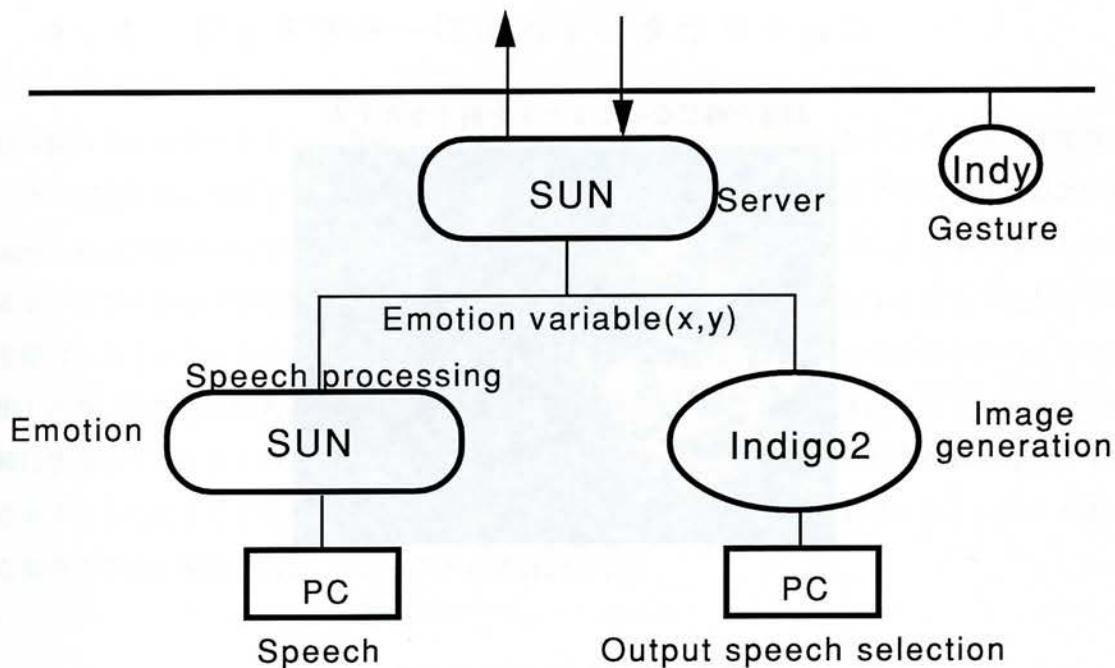


図5.8 ジャスチャー認識機能を持つ「ミック」のシステム構成

## 5.5 まとめ

本論では、新しいタイプのキャラクター「ミック」を紹介した。基本的コンセプトを述べると共に、感情音声とジェスチャー認識によるマルチモーダルな感情の取り扱いの実装を試み、その具体的内容について述べた。マルチモーダル・インタフェースにとって、言葉によるモダリティと同じくらい、人間の表情や動作などの行為によるモダリティが重要であることは言うまでもない[5]。コンピュータとのインタラクションを効率的で快適なものにし、人間の能力とコンピュータの性能を十分に発揮できるようにするためには、ユーザーに認知的負担を与えないようなインタラクションが必要である。そのためには

- (1) 身体的表現・操作による意識・無意識的表現を直接伝達できるモダリティがサポートされていること。
- (2) 人の直感や感性に直接訴えかけることができるモダリティを持つこと。
- (3) それらのモダリティが、広帯域で十分な情報を即時に伝達できること。
- (4) 身体的束縛が十分に少ないこと。

などの条件が満たされている必要があり、今後の研究課題である。

## 6 . 声の感情と言葉の意味に対応する 「インタラクティブポエム」

### 6 . 1 まえがき

マルチモーダルキャラクター「ミック」では、複合感情は生成できたが、意味は認識しなかった。ところが人間は会話に意味を求めるので、言葉がわからない「ミック」に必ず話しかけていた。たとえ相手が人形でも、物でも、話しかけて、自分で意味を作り上げてしまう能力は子供も持っている。しかし、ノンバーバルコミュニケーションだけでは、コミュニケーションの正確さ、しいては奥の深さが欠けてしまう。そこで、言葉の意味を踏まえたコンピュータキャラクターとのコミュニケーションの設計を考えてみた。日本の情緒ある「あうん」の会話をコンピュータとリアルタイムに行うのは極めて困難である。土台となるような状況の枠を設計することを考えていたとき、我々の古来の文化である連歌を思いついた。そして連歌形式で、コンピュータ詩人と人間が即興詩を作る「インタラクティブポエム」という考えにたどり着いた。本来「詩」は、詩人が表現したい「メッセージ」を、「言葉の力」を用いて個人の世界を作りあげたものであり、それを読んだり聞いたりすることにより、その世界を我々は享受している。「インタラクティブポエム」[78][79]は、享受するだけではなく、能動的に人間がコンピュータと創作することにより、思いもよらぬハプニング(偶然性)、誰がコンピュータと創作するかによる個性の違いが出てくると期待される。これらはインタラクシオン作用により開放的になる。つまり、「対話」の機能を取り入れることで感性によるコミュニケーションの媒体ができあがる。

### 6 . 2 インタラクシオン

大きなスクリーンにギリシャ神話の音楽の女神「ミューズ」の顔が浮かび上がる。ミューズは、まるで一緒に歌うように、人間と対話しながら詩を紡ぎ出す。ミューズはユーザーに対して、短い詩的な言葉を感情を込めて語りかける。それを聞いてユーザーはその詩の世界に入り、同時に、ユーザー自身も詩的な言葉を、ミューズに語り返す。この詩的な言葉をやりとりする対話プロセスを通じて、「インタラクティブポエム」は、インスピレーション、フィーリング、そして、感性に満ちた即興詩の世界を、ユーザーとコンピュータが一緒に創りあげることができる。(図6.1)

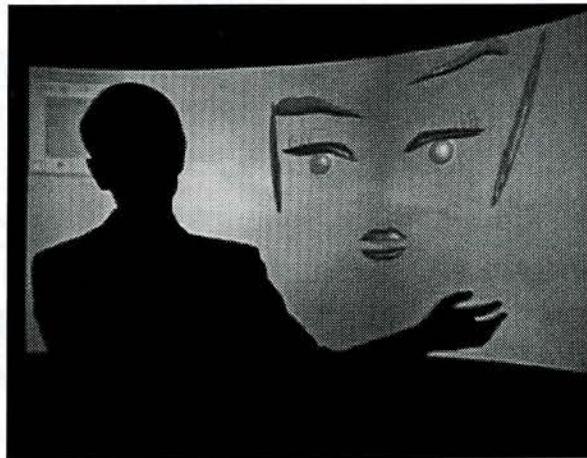


図6.1 インタラクティブポエムを行う観客

### 6.3 ソフトウェア構成

インタラクティブポエムのシステムは、システム制御、音声認識、イメージ生成、および、音声出力の4種類のユニットから構成される。(図6.2)

システム制御ユニットは、データベースを使用して、システム全体の動作をコントロールする。このシステムにおいて、もっとも重要な課題は、インタラクティブな詩を作成することである。最初にインタラクティブポエムを記憶しているデータベースの内容を説明する。従来の詩は、詩的なフレーズが連続したものと考えられる。つまり、従来の詩の基本的構成は、単純な状態遷移ネットワークによって表現が可能である。このネットワークでは、各フレーズは状態に対応し、各々の状態には1つの状態が接続する。(図6.3)

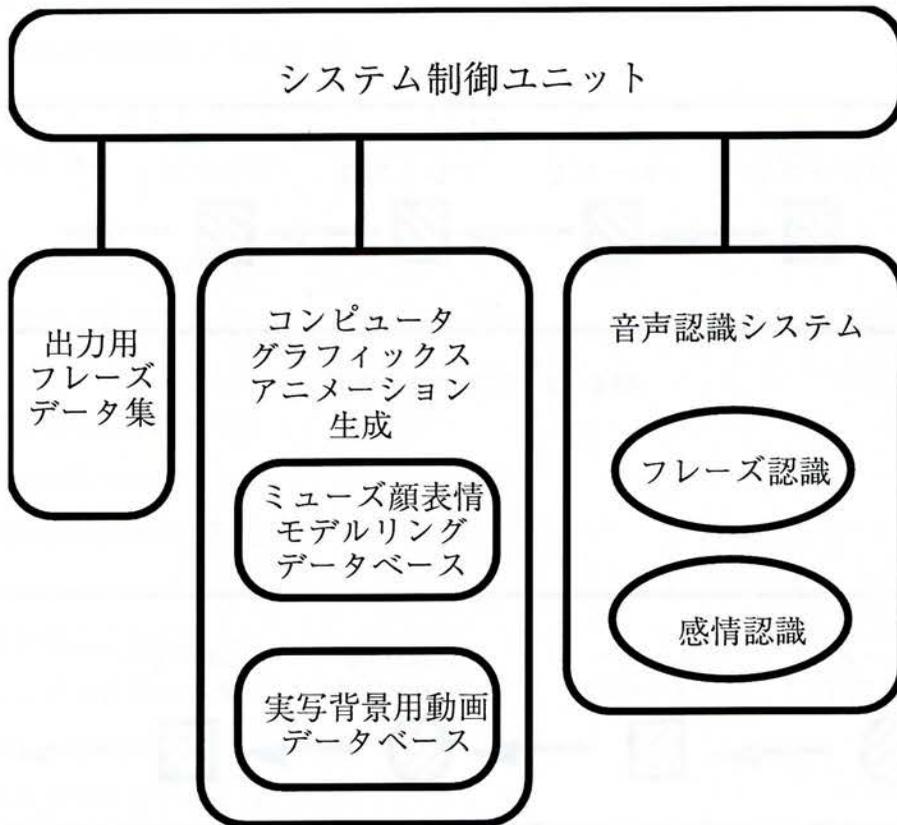


図6.2 インタラクティブポエムのソフトウェア構成システム

インタラクティブポエムの基本形式は、この遷移ネットワークによって表現される。この違いは、コンピュータの発声するフレーズとユーザーの発声するフレーズが交互にあらわれることである。これは、コンピュータとユーザーが、あらかじめ決められた詩のフレーズを交互に読むという単純な対話に対応する。(図6.4) 本インタラクティブポエムシステムでは、より高度な対話機能を実現するために、コンピュータの発声するフレーズに接続可能な複数のフレーズを用意してある。これらのフレーズは、コンピュータのフレーズとの音韻的なまた意味的なつながりを考慮して注意深く作成してある。この遷移ネットワークはデータベースに記憶され、全体のプロセスを制御するために使用される。この仕組みにより、ユーザーは自分の感情・感性に合ったフレーズを選択しコンピュータに語りかけることにより、オリジナルな詩の世界に自分自身の感情・感性を加えることが出来る。(図6.5)

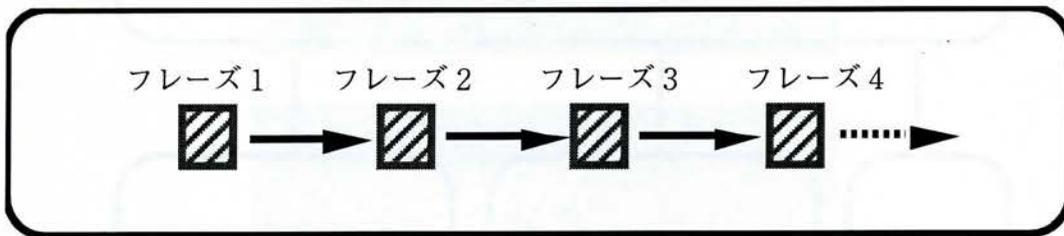


図6.3 従来の詩の構造

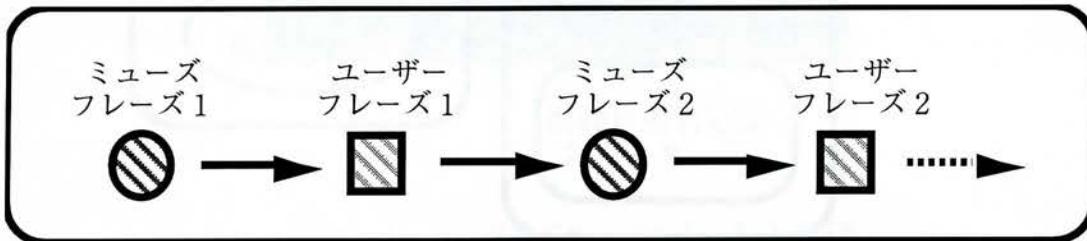


図6.4 インタラクティブポエムの詩の構造 (A)

詩のフレーズ (a) -----ユーザーのフレーズ (b) ----- ミューズのフレーズ

谷川俊太郎「あなた」より[80]

- 0 あなたはだれ (b)
- 1 わたしではないあなた (b)
- 2 あのひとでもないあなた (b)
- 3 もうひとりのひと (b)
  
- 4 わたしとおなじようなみみもち (a)
- 5 わたしとはちがうおとをきくひと (b)
- 6 わたしとそっくりの (a)
- 7 じゅっぽんのゆびもち (b)
- 8 わたしにはつかめないものを (a)
- 9 つかもうとするひと (b)
- 10 あなた (b)
  
- 11 あなたはたっている (a)
- 12 まなつのひをあびて (b)
- 13 うみにむかって (a)
- 14 わたしにせをむけて (b)
- 15 あなたはみつめる (a)
- 16 とおいすいへいせんを (b)
  
- 17 あなたのころには (a)
- 18 わたしのみたことのないまちの (b)
- 19 わたしのあるいたことのない (a)
- 20 こみちがつうじている (b)
- 21 そのこみちに (a)
- 22 しずかにゆきがふりつもり (b)
- 23 わたしのあったことのないひとが (a)
- 24 こっちへはしってくる (b)

図6.5 インタラクティブな詩の例

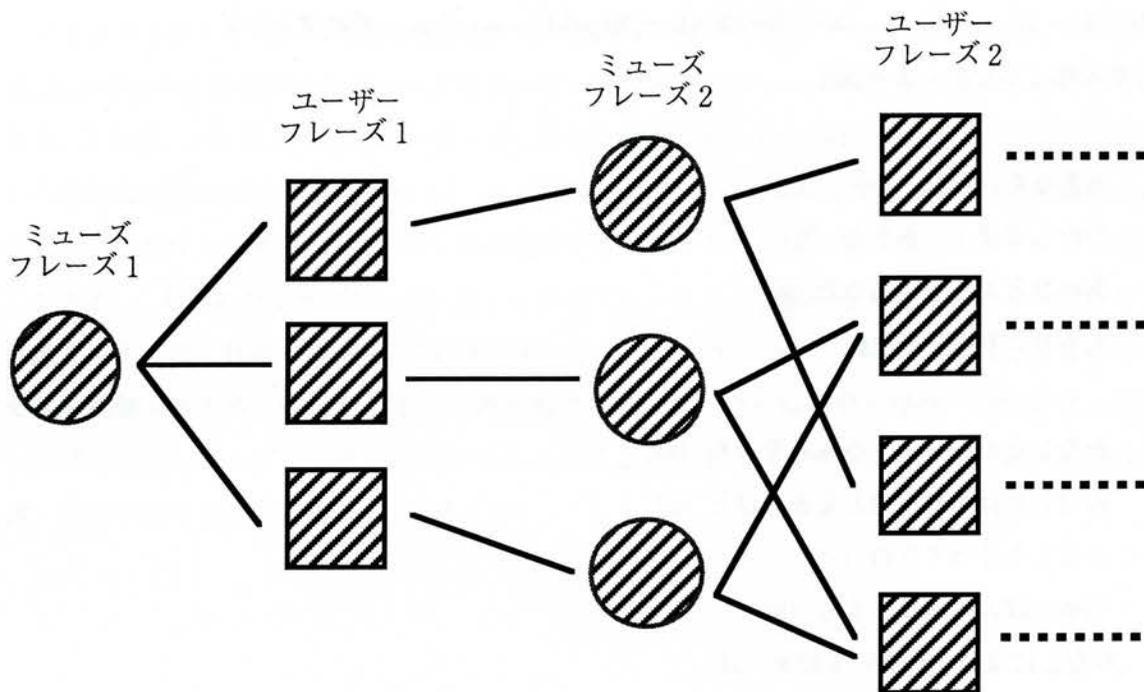


図 6.5 インタラクティブポエムの構造 (B)

音声認識ユニットは、意味認識と感情認識の2種類の認識機能を持つ。ユーザーが発声するフレーズの意味内容を認識するため、HMM (hidden Markov model) に基づいた不特定話者音声認識を用いる。発声される各フレーズは、音韻系列の形で表現され、認識用の辞書に記憶される。(図 6.6) また、同時にユーザーの感情状態を認識するために、感情認識を行なう。感情認識のための基本アーキテクチャーとしては、ニューラルネットワークを用いる。このニューラルネットワークは、多数の話者が8種類の感情で多数の単語を発声した音声データを用いて学習している。この学習により、不特定話者、内容独立型の感情認識が実現される。(図 6.7)

ユーザーが発声するフレーズに対するコンピュータキャラクターミュージズの反応は、音声と映像によって表現される。音声出力ユニットでは、ミュージズの発声するフレーズに対する音声データが記憶されており、必要に応じて再生される。コンピュータグラフィックス生成ユニットは、映像表現をコントロールする。映像は、MUSEの表情と背景シーンの2種類から構成される。ミュージズの表情は、ユーザーの感情に対するミュージズの反応を表わす。これらの表情は、8種類の感情に対応した3次元モーフィングアニメーションによって表される。(図 6.8) また、詩の世界の雰囲気表現するために、何種類かの背景シーンを記憶しておき、遷移ネットワークの状態に応じて適切な背景映像を表示する。

## 6.4 処理プロセス

(1) MUSEがフレーズを発声すると同時に、認識プロセスが起動される。ユーザーが発声すると、その言葉は、コンピュータの意味認識機能によって認識される。この時、遷移ネットワークの次の状態に対応した辞書サブセットを用いて認識を行なう。同時に、ユーザーの言葉に含まれている感情が、コンピュータの感情認識機能によって認識される。

(2) 認識結果と遷移ネットワークに基づいて、システムの反応が決定される。MUSEの表情は感情認識の結果に応じて変化し、MUSEが発声するフレーズは意味認識の結果と遷移ネットワークに基づいて決定される。背景シーンは、遷移ネットワークの状態に応じて変化する。

(3) このようにして、MUSEとユーザの間の詩的なインタラクションが行なわれる。

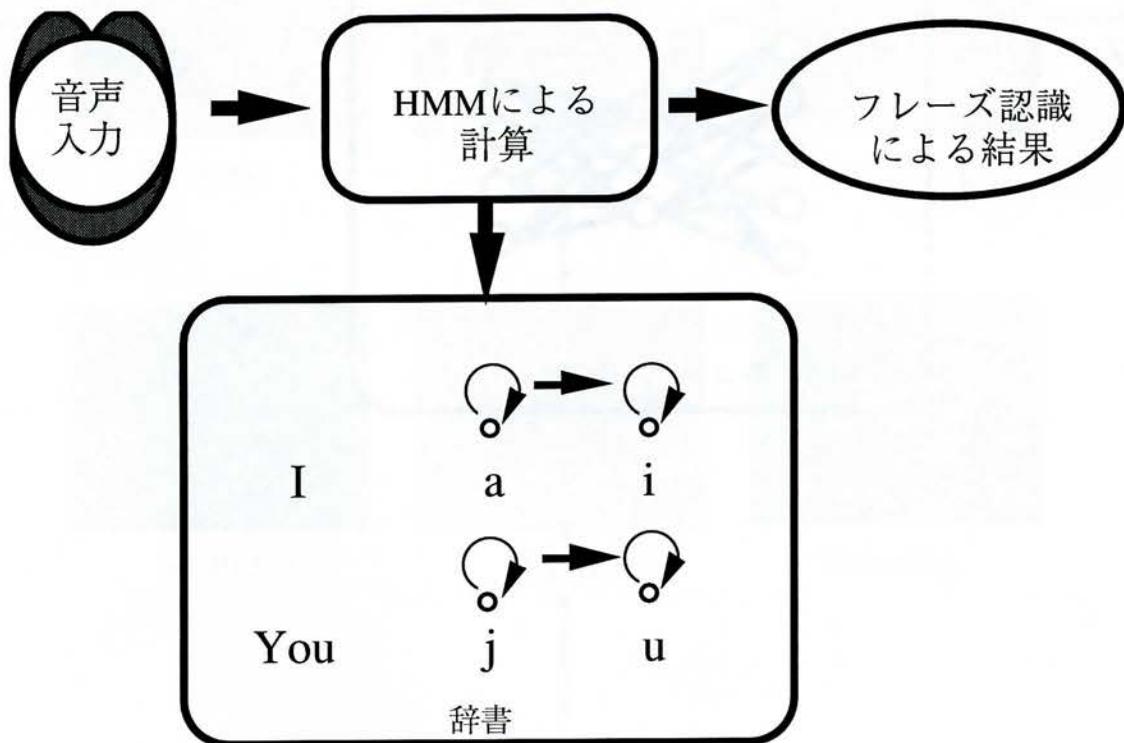


図6.6 詩のフレーズ認識

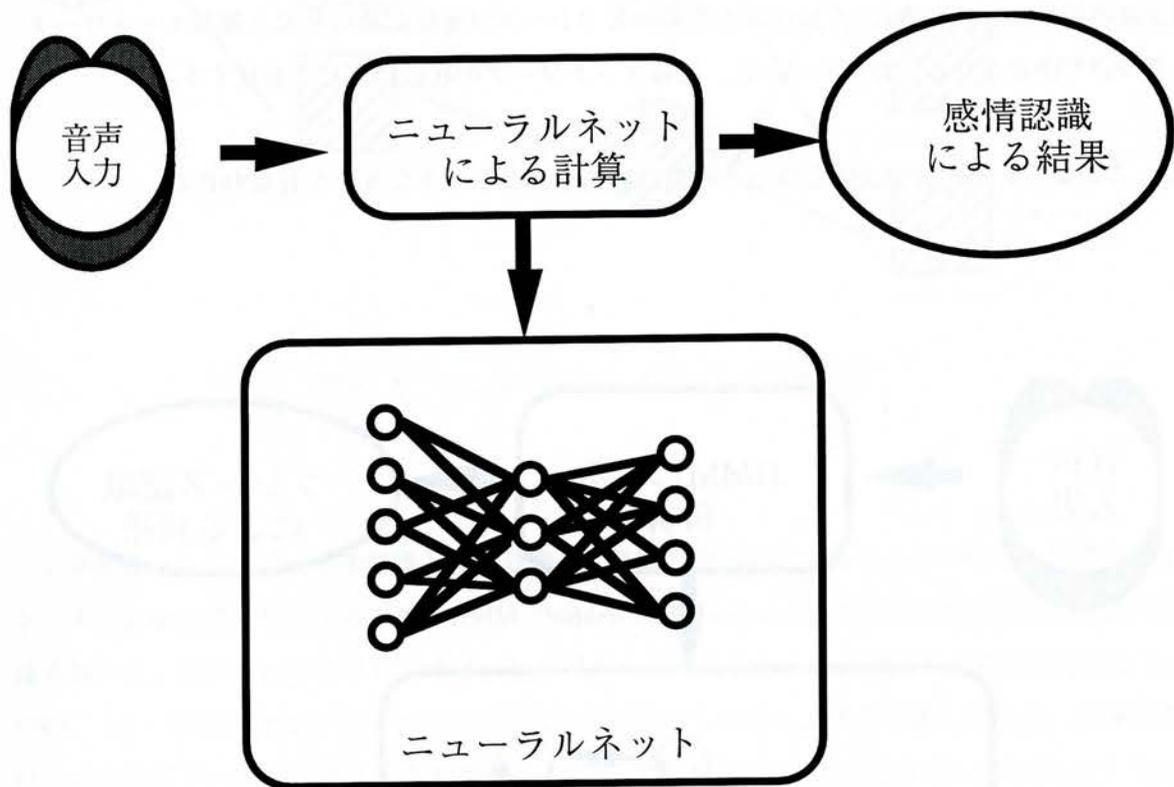


図6.7 感情認識システム

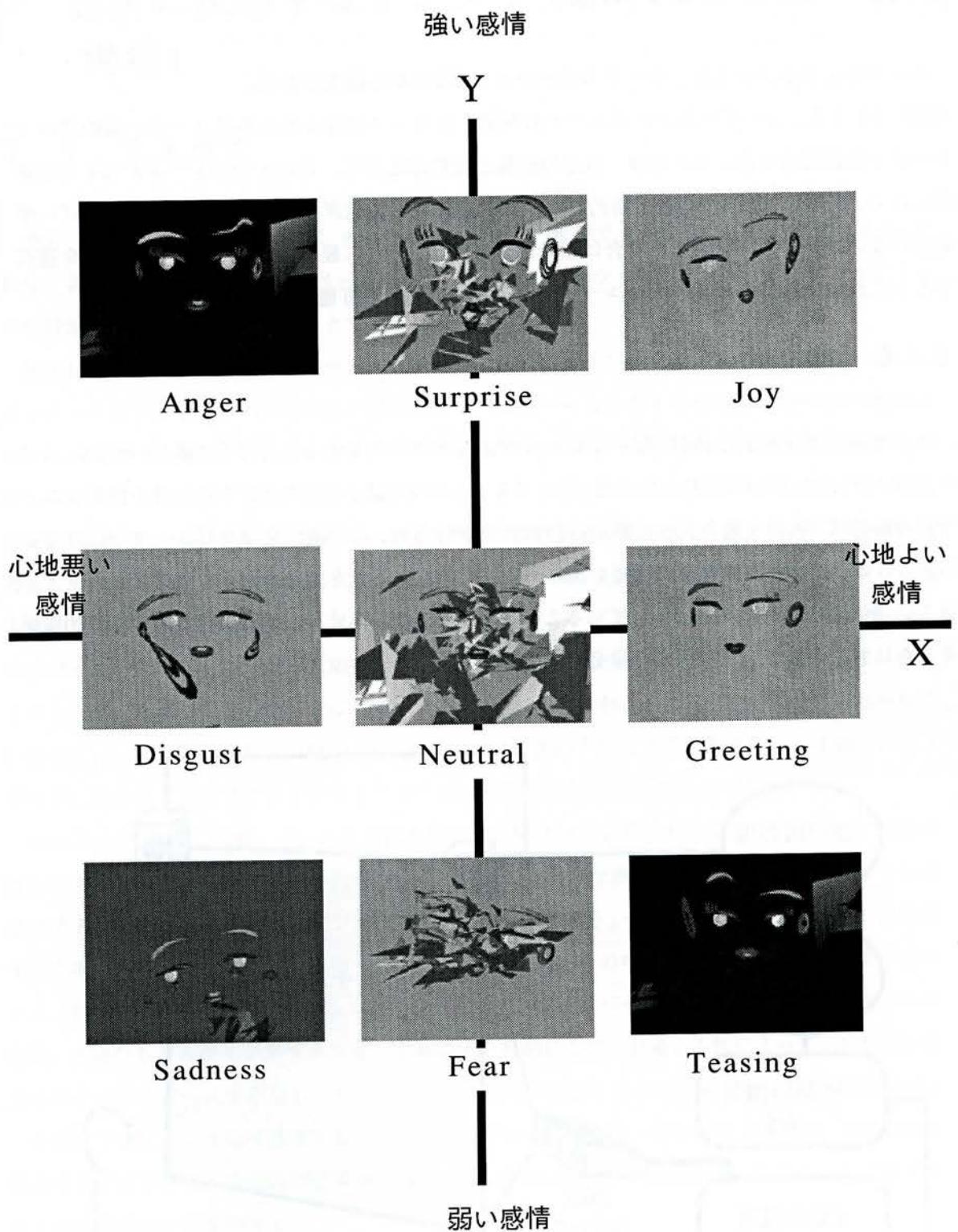


図6.8. ミューズのフィーリング表現例

## 6.5 ハードウェア構成

ハードウェアシステムは、ワークステーションとPCから構成される。  
(図6.9) また、ユーザーとコンピュータの対話をスムーズに行なうためにユーザー用の詩のフレーズを画面表示する。ユーザは、自分の感情と感性に応じて、これらのフレーズから1つを選択したり、あるいは、自分自身で詩的なフレーズを作成して発声することができる。ユーザーが辞書にないフレーズを発声した場合は意味認識機能は、それに最も近い、既存のフレーズを選択する。これによりコンピュータとユーザーの即興的な対話も可能になる。

## 6.6 まとめ

ユーザーとコンピュータが、短いフレーズや文章をやりとりすることで「対話」を行ない、「インタラクティブシステム」について述べた。そして、この対話から、オリジナルな詩の世界にユーザー自身の感情・感性を取り入れた新しい詩の世界が作られ、その創作をコンピュータが助けている。これは、つまり人間自身の情操を高め、広げていく楽しさを持ちながら、コンピュータと心地よい関係の状態が作り上げられていることになる。このコンピュータと人間の間の心理的関係をさらに育て深めていく関係が今後必要である。

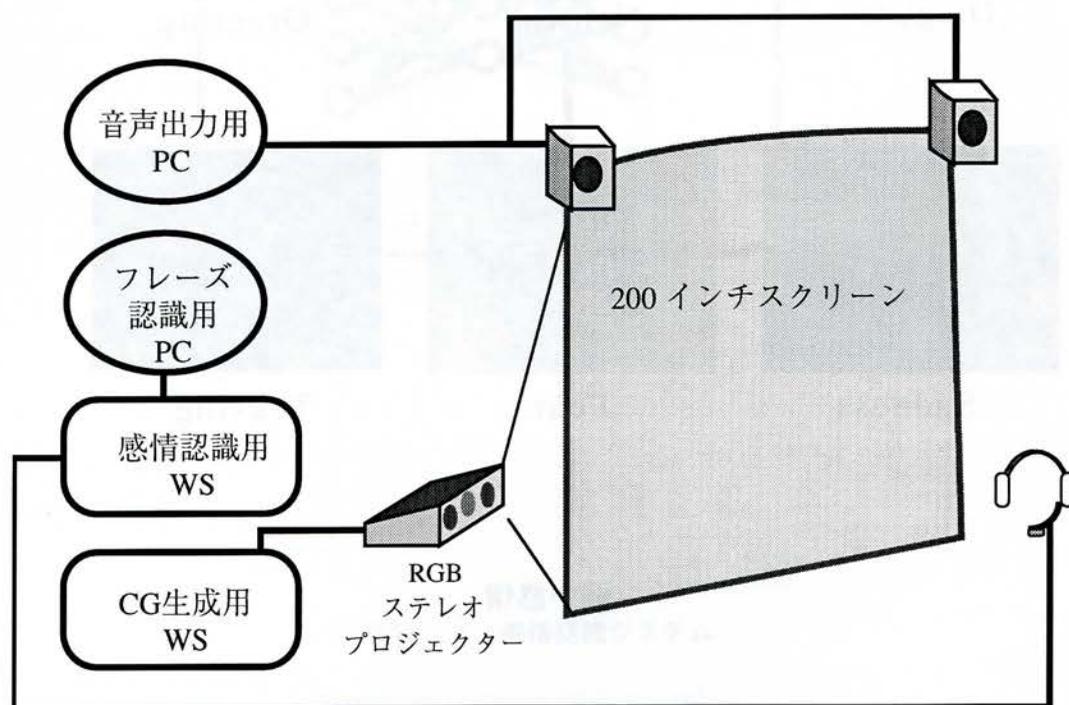


図6.9 インタラクティブポエムのハードウェア構成

## 7 感情に反応するインタラクティブシアターの構成1

### 7.1 まえがき

最近、コミュニケーション、エンターテイメントなどの世界で大きな変化が起こりつつある。これは一言でいえば、「サイバースペースの中でのコミュニケーション、エンターテイメント」という分野が出現しつつあるということである。

例えば、通信の世界ではインターネットという新しいコミュニケーションの場が生じつつある。インターネットは全世界の人をつなぐ巨大なサイバースペースと考えられる。人はその中で他の人とコミュニケーションをしたり、ショッピングをしたり、また新しい情報を求めて情報の海の上でネットサーフィンを行う。次に巨大なエンターテイメント市場である映画業界を見てみよう。最近の映画は、デジタル技術やコンピュータグラフィックス技術を取り入れ、新しい世代の映画へと移行しつつある。デジタル技術やコンピュータグラフィックス技術は、従来の映画にはなかった超現実的な世界、すなわちサイバースペースを作り出す能力を与えてくれる。一方、テレビゲーム、特にロール・プレイング・ゲーム (RPG) は人が仮想世界、サイバースペースにおける主人公になってストーリーを楽しむことを可能にしてくれる。RPGが子供を中心として熱狂的に受け入れられているのは、ストーリー展開により我々をサイバースペースの中に引き込む力を持った小説というメディアに新たにインタラクティブストーリーという能力が付加されたためと考えられる。

このような、通信、映画、ゲームなどにおける新しい流れを組み合わせることによって、将来の新しいメディアを作り出すことが出来ないかというのは大きな夢である。ここでは、このような将来のメディアをインタラクティブシアターと呼ぶことにしよう。インタラクティブシアターは超現実的な映像やシーンからなるサイバースペースを作りだし、その中に人が入っていくことを可能にする。またその中で他の人とコミュニケーションしたり、サイバースペースの中のキャラクターや動物・植物たちと人がインタラクションすることを可能にしてくれる。これによって、人が自分自身のサイバースペースを創造し、さらにはその世界の主人公となって活躍することが可能になる。

本論文ではこのような可能性をもったインタラクティブシアターがいかなるものか、またそれを構築するにはどのような条件が必要かを考察する。また、現在我々が具体的に進めているプロジェクトの内容について説明する。

## 7.2 インタラクティブシアターの位置付け

まず、本章では、インタラクティブシアターのコンセプトを明らかにする。インタラクティブシアターは、従来の映画、小説、テレビゲーム、さらには通信などを統合した新しい形のメディアである。これは以下で説明するように、現在のメディアとの関連からすると「観客参加・体験型の映画もしくは劇場」というのが最も近い説明になる。インタラクティブシアターという呼び名はこの基本的なコンセプトに基づいている。

インタラクティブシアターは以下の要素で構成されている。

- (1) インタラクティブな話の展開を可能とするインタラクティブストーリー
- (2) 主人公となってインタラクティブストーリーの世界を体験する参加者
- (3) 主人公とインタラクションを行ないつつストーリー展開に参加するキャラクター

### 7.2.1 従来メディアに関する考察

メディアが人々に受け入れられるには、人間の関心をとらえ人間を引き込む力を持つことが必要である。従来の各種のメディアでは、以下のような手法により、読者・観客をメディアの作り出す世界に引き込む、すなわち感情移入を行わせようとしてきた。代表的なものを以下にあげる。

#### (1) 小説・詩

小説や詩は、言語の力のみによって読者を作者の描き出す世界の中に引き込み、それを体験させようとしてきた。この鍵は人間の想像力にある。人間は言語のみから頭の中で架空のイメージを構築する力、すなわち想像力を持っている。小説や詩は出来るだけ余計なものを排して、言葉の力のみによって人間の持つこの想像力に訴え、人に感動を与える役割をしてきた。人間の持つこの力は偉大であり、インタラクティブストーリーの展開においてもこれを利用することが大切である。

#### (2) 映画

映画は、大画面に迫力ある映像を映し出し、かつ同時に大音量のサウンドを流し、人間の視覚・聴覚に訴えることにより、人間を架空の世界に引き込もうとしてきた。特に最近はCG技術の急速な発展により、架空の世界、実際にはありそうもない出来事をリアルな映像として作成することが可能になってきた。これは人間の想像力に訴えるというよりは人間の視覚・聴覚に圧倒的な情報を

送り込むことによって人間を架空の世界に引き込もうとしていることになる。時にこれは成功し、また失敗する。例えば小説の映画化の場合によくあることであるが、小説に比較して映画がまったくつまらないという場合が多々ある。これは、小説の読者が頭の中で作り上げた架空世界に対し、映画の中の世界がそれに及ばない場合に人が感じる感覚である。反対に人の想像力を超えるような映像を作り上げることのできる監督は、才能があると認められることになる。

### (3) テレビゲーム

テレビゲームには大きく分けて2種類の製品がある。アクションゲーム (ACG) とロールプレイングゲーム (RPG) である。これらのゲームについてインタラクションの内容、レベルの面から分析する。

a) ACG: アクションゲームは、ボタン入力によりゲームの主人公、乗り物などの動きをコントロールし敵を倒すというのが基本である。したがってボタン入力の正確さと素早さが鍵となる。すなわち運動神経の素早さが勝負をわけることになる。インタラクションは極めて単純であり、人間の身体面に快感・快楽を与えることを目的としている。ただし、このような単純なインタラクションが人間に快感を与えることが出来るというのは重要な事実である。

b) RPG: ロールプレイングゲームは、小説の世界をゲームに仕立てたものということができる。基本的なストーリーが設定してあり (世界を滅ぼそうとする悪に対して主人公が仲間と協力して立ち向かうという筋をファンタジー仕立てにしてある場合が多い)、人はゲームの主人公を操ってストーリー展開をコントロールすることが出来る。主人公が敵と対戦しつつ成長するという、小説によくあるタイプの物語性と、敵との対戦時のアクションゲームとしての要素、さらには謎をときながらストーリーを展開していくという謎解きなどの種々の要素が相互に補いあって新しいゲームのジャンルを作り上げ、子供を中心として熱狂的な多くのファンを生み出している。このように、種々の要素の複合体であるという性格は持つが、基本的には小説にインタラクションを持ち込んだ新しいメディアであるといっている。したがって、人間の身体面に快感を与えるだけでなく、人間の精神面に訴える力を持っている。またインタラクションは、素早さを必要とするものではなく、ストーリー展開の中で判断を行う必要がある場合にこれを用いるという形式になっている。

## 7.2.2 新しいメディアとしてのインタラクティブシアター

映画の持つ迫力ある映像・音響情報をベースとしたストーリー展開の能力、文字（言語）という単純なメディアを使いながら人間の想像力に訴える小説の持つ力、そしてテレビゲームのもつインタラクシヨンの能力、これらを組み合わせることにより従来なかった新しいメディアを作り上げることが可能ではないだろうか。仮想世界を構築し、その世界と観客とのインタラクシヨンを行なわせようとする試みは種々行なわれている。人間との行動レベル、感情レベルでのインタラクシヨン可能なコンピュータキャラクターの生成[81][82]や、インタラクティブアート[83]などがそれにあたる。しかしながら、これらのインタラクシヨンは短時間的なものであり、ストーリーの導入はなされていない。また映画の領域では観客参加型の映画の概念は従来からあるが、実現例はプリミティブなレベルにとどまっている。RPGは、インタラクティブシアターの概念に最も近いが、インタラクシヨンがボタン入力により行なわれ、人間同士のインタラクシヨンと異なるという本質的な相違がある。これに対しインタラクティブシアターは、以下のような特徴を持つ。

### (1) 映像・音によるサイバースペースの構築

CGの活用、CGと実写映像の統合、さらには立体映像・立体音場の創出によりこれまでになかった仮想世界、人間の想像力を刺激する世界（サイバースペース）を創出することが可能となる。

### (2) サイバースペースへの没入

しかもその仮想世界が自分とは別にあるのではなく、自分がその仮想世界の中へ入っていくことを可能にしてくれる。

### (3) サイバースペースの中のストーリーの体験

仮想世界の中に入っても単なる傍観者ではおもしろくない。仮想世界の主人公となって仮想世界のストーリーを体験することにより、従来なかった新しい世界が広がると考えられる。

### (4) サイバースペースとのインタラクシヨン

仮想世界の中の住人、映画の世界であれば主人公を取り巻く種々のキャラクタと声や身振り・手ぶりでインタラクシヨン（コミュニケーション）を行うことによりストーリー展開を楽しむことが出来る。

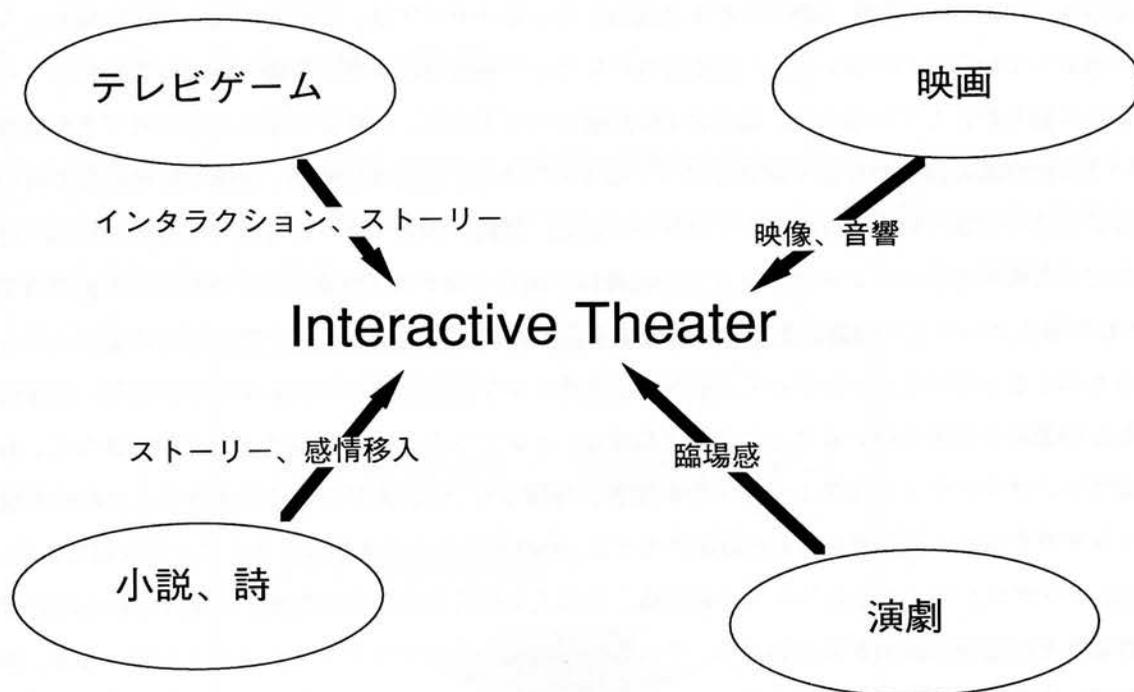


図7.1 従来のメディアとの比較

1. サイバースペース生成技術
  - ストーリーの生成
  - シーンの生成
  - キャラクターの生成
2. インタラクション技術

図7.2 インタラクティブシアターの実現に必要な技術

また、このような機能を実現するには、工学的側面とアートの側面の融合が重要であると考えられる[84]。すなわち、映画、RPGのような成功しているメディアは、芸術的側面と工学的側面をうまく融合しているものが多い。例えば最近のハリウッド映画がかったの斜陽と呼ばれた時代がうそのように盛りかえしているのは、最新のCG技術という工学と、これまで積み上げられてきた映画という芸術作成の技術がうまく統合されているからである。最近RPGも、映画と同様総合芸術的な様子を見せてきている。丹念なストーリーの設定・展開、ストーリーに合致した音楽の作成・付与などの芸術的側面とインタラクションの高速性や面白さなどの工学的側面、さらには最近3DのCGの導入という工学的側面がうまく融合することにより大人から子供まで熱狂的に受け入れられるものとなっている。したがって、インタラクティブシアターにおいても単に音声処理・画像処理などの最新の技術を用いることによる「高度な」インタラクションに重点をおくのではなく、あくまでインタラクティブなストーリーを参加者に体験させ、仮想世界の中に没入させるための手段がインタラクション技術であるとの認識のもとに、他の要素とのバランスをとることが大切である。一方、インタラクティブなストーリーの作成、見る人を引き込む映像の作成、映像と合致した音響の作成などは芸術的側面が大きいので、それらの技術にたけたアーティスト（ストーリー作家、映像作家、作曲家など）の参加が必要である。

以上述べた従来のメディアとインタラクティブシアターとの関係を図7.1に示す。上で述べたインタラクティブシアター構築のための基本要素の内、最も大切なのはサイバースペースの構築機能とインタラクション機能である。また、サイバースペースの構築機能は、ストーリーの生成機能、シーンの生成機能、サイバースペース中のキャラクターの生成機能より構成される(図7.2)。したがって、以下ではこれらの基本機能の内容について考察する。

## 7.3 インタラクティブシアターの機能

### 7.3.1 ストーリー生成

ここでは、インタラクティブなストーリー展開を特徴とするシステムにおける具体的なストーリーの設計・制作法について考察する。インタラクティブなストーリーの展開については以下のような方式が考えられる（図7.3）。

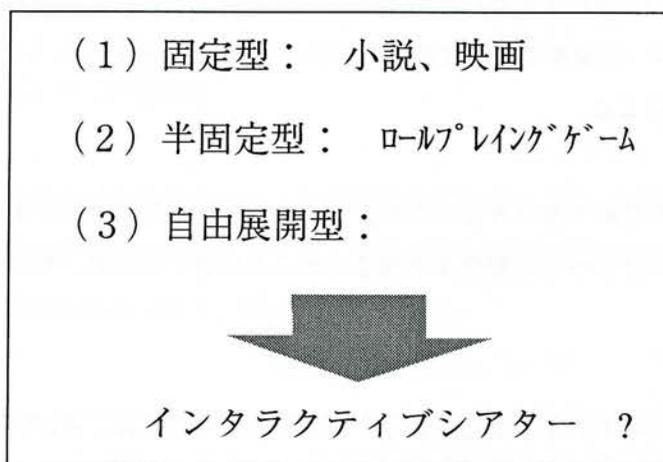


図7.3 ストーリーの生成

#### (1) ストーリー固定型

固定型は、従来の小説・映画などのような、作家から読み手・鑑賞者への一方向的な情報の伝達メディアにおいて用いられてきた手法である。固定型は、インタラクションという機能を持たないメディアにおいては必然である。その反面、いかに人の気持ちを引き込み感情移入をさせ、人を最後まで引っばっていくかは、作家・監督が心を砕いてきた部分であり、文学・映画といった領域が確立している。

#### (2) ストーリー半固定型

半固定型とは、ストーリーの大筋があらかじめ決定されているものをさす。その意味では小説・映画などの固定型と同様である。ただし、途中のストーリー展開にある程度の自由度（もしくは複数のストーリー展開）が設定されており、参加者はその範囲内でインタラクションによりストーリー展開の自由を楽しむことが出来る。現在のRPGはほぼこの形式であると考えられる。この方式の利点は、従来の小説・映画の制作において積み重ねられてきたノウハウなどがほぼそのまま使え

る点にある。したがって、参加者を引き込めるようなストーリー展開をあらかじめ作り上げておき、なおかつ全体的なストーリー展開に影響を及ぼさない範囲で、ストーリーに幅を持たせるという作り方をすることが出来る。

### (3) ストーリー自由展開型

自由展開型は、ストーリーがあらかじめ決まっておらず、参加者のインタラクションに依存していかようにも展開していくものをいう。このような場合、制作側は、ストーリーの展開される世界を構築しておき、後は参加者の自由意思にまかせるという方策をとることになる。一見、この自由展開型が最も自由度が高く、参加者に満足度を与えられそうであるが、必ずしもそうとはいえない。これは以下のような理由による。

a) 参加者の行動があらかじめ予想できないため、すべての可能性を考慮した作りをしておく必要がある。これはある意味では1つの世界全体を構築することに等しいため、不可能とはいえないまでも極めて困難である。

b) プレイする側からいっても、完全な自由が与えられるということは、逆にいえばどのようにふるまってもいいかわからないということになりかねない。これは、右も左もわからない外国に初めていった時に我々がどうしていいかとまどうのと同様の状況である。

したがって、完全な自由展開型は、将来的にはありえても当面はありえないと考えられる。上記の考察の結果、ストーリー半固定型が現実解であると考えられる。半固定型を採用した場合でも以下のような項目が重要である。

#### (1) 脇役の自然なガイド

現在のRPGでも、マニアは別として多くのプレイヤーは、しばしば道標を失ってどうすればいいかわからなくなることがある。このことは、ほとんどすべてのRPGに対応してガイドブックが出版され、ガイドブック自身がベストセラーとなったりする現実をみると納得できるであろう。したがって、インタラクティブシアターにおいてもガイドは必要である。ガイドブックをつくるわけにもいかないから、付き添いとなるべき脇役のキャラクターを用意して、キャラクターの自然なガイドによりストーリーの展開を助ける方策が必要である。

#### (2) 要所では自分で判断

しかしながら、何から何までガイドが案内してくれては、インタラクティブストーリーの面白さは半減する。ストーリー展開が変わるような要所では、参加者の判断によってストーリー展開が選

扱われるような方式になっている必要がある。したがってインタラクションには以下2つの種類があると考えられる。

(1) ストーリー展開の大きな岐路となる時点でのインタラクション。このインタラクションの結果によってそれ以降のストーリー展開が変わる。

(2) ストーリー展開には関係しないインタラクション。他の登場人物との会話などを楽しんだり、いわばインタラクションそのものを楽しむ部分。

### 7.3.2 シーン生成

サイバースペースにおける各種のシーンを生成するにあたっては、これまでの各種のメディアが行ってきたシーン生成の方式を振り返って見る必要がある。従来のシーン生成法は大きくわけて以下の2つの方式に分類される(図7.4)。

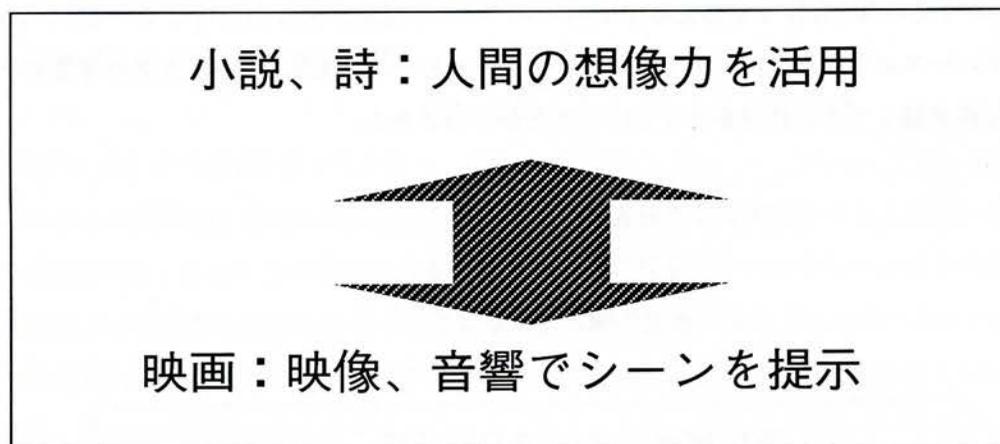


図7.4 シーンの生成

#### (1) 小説・詩における方式

小説や詩は読者に文字のみによって訴える必要がある。もちろん挿絵などが用いられる場合もあるが、これはあくまで補助的な役割と考えられる。文字という情報量の少ないメディアを用いるがゆえに、人間の想像力に訴えることが出来る。すなわち、人間の頭の中にサイバースペースを言葉の力によって作り上げようとするのが小説や詩におけるシーン生成の手法である。

## (2) 映画における方式

映画においては、圧倒的な映像・音響によってシーンを生成し観客に与える。すなわち、観客の視覚・聴覚を映画が与える映像・音響によっていわば占拠して、観客に仮想世界を体験させようとするのが映画の方式である。これはある意味では、観客の想像力のはばたく範囲を限定してしまうことにもなりかねない。

最近のシーン生成の方式は、映画の方式をベースとしてそれを延長する方向にいとっていると考えられる(図7.5)。すなわち、テレビゲームは従来の素朴な2DのCGから、2Dでも高精彩のグラフィックスを取り入れたり、実写を取り入れたり、さらには3DのCGを用いたりという方向に進んでいる[85]。一方、映画の世界では、CGを活用したり、デジタル技術を取り入れたりするいわゆるバーチャルリアリティ技術を用いることにより、従来は不可能とされたリアリスティックな映像をCGで作成することが出来るようになった。これらは全体として、リアルなさらには超リアルな映像・音響シーンを作成する方向に向かっているということが出来る。しかしながら、いかに新奇な映像であっても人間はすぐにそれに慣れてしまう。モーフィング技術が一時映画やCMなどでもはやされたが、現在では見る側がそれに慣れてしまい、新鮮さを与えなくなったため、最近あまり用いられなくなったのがよい例である。これに対して、人間の想像力はそれをうまく刺激してやれば、いかなる映像・音響をも超えるサイバースペースを構築出来る力を持っている。すなわち、サイバースペースの生成にはバーチャルリアリティ技術のみでは限界があり、人間の想像力をうまく活用して感情移入できる状況を作り上げることが大切である。

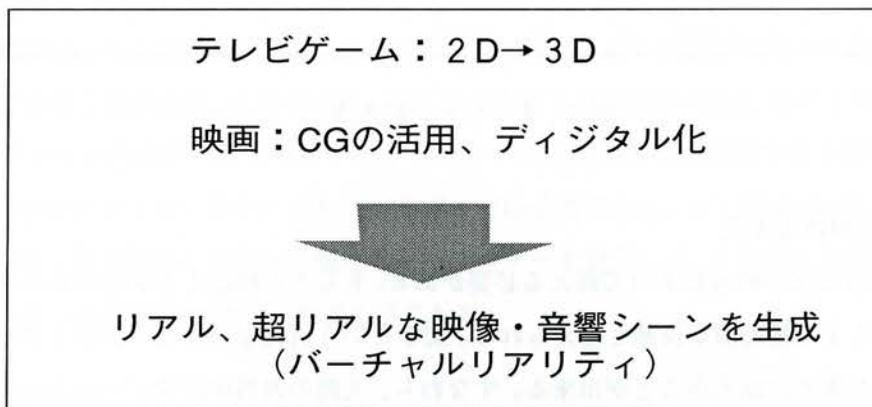


図7.5 シーン生成の動向

### 7.3.3 キャラクター生成

インタラクティブシアターを作成するに当たって最も重要かつ困難な仕事は、人とインタラクションができ、人をサイバースペースの世界に導き入れてくれる存在をどのようにして作り出すかという点にある。人間は年齢・性別にかかわらず、人間に似せて作られたものに強く魅せられてきた。古代の土偶や、江戸時代のからくり人形、さらには、現代の着せ替え人形やロボットに至るまで、人はそれらに感情移入をしてきた。このことから明らかなように、インタラクティブシアターの実現のためには、人間に似せたキャラクターを生成し、インタラクティブシアターの中で重要な働きをさせることが必要である。キャラクター生成に関しては以下のような点が重要である。

#### (1) 人型キャラクター

いわゆる人型キャラクター (human-like character) は必ずしも姿形が人間にそっくりであることを意味しているのではない。アニメの登場人物は必ずしも実際の人間とそっくりの姿形をしているのではない。かなり大胆なデフォルメを施してある。さらには、我々はディズニーのキャラクターのような動物のキャラクターにも感情移入してしまう。これは、外見ではなくその動作・仕草に生き生きとした (life-likeness) や実在感 (believability) を感じるからである[86][87][88]。このような意味での人型キャラクターの生成が必要である。

#### (2) 実写キャラクター対CGキャラクター

キャラクターは通常CGで生成される。これと対照的な方法として実写フィルムを用いる方法がある。この方法では、あらかじめ考えられるキャラクターの反応パターンやシーンのすべてを人間の俳優に演技させて撮影し用意しておくことになる。人間の俳優が演技している訳であるから、生き生き度や実在感は極めて優れている。しかしながら、これはインタラクションがうまくいっている場合である。我々は人間同士のインタラクション、コミュニケーションに慣れているため、どこかにずれがあるとたちまち不自然さを感じてしまう。この不自然さをなくすためには、きめ細かく反応パターンやシーンを用意すると共に、自然な誤りのないインタラクションを実現する必要がある。

#### (3) 自律キャラクター対アニメーション

CGキャラクターの反応パターン、動作シーケンスを用意する際、すべてあらかじめアニメーションとして用意しておく方法と、キャラクターに自律性を持たせることによって、ある程度の動作・反応は自動的にこなすようにする方法がある。自律キャラクターに関する研究はかなり行なわれているが[86][89][90][91]、まだ十分な自律性を持つには至っていないのが実情である。したがって、アニメータによるアニメーション作成を基本としつつ、それに徐々に自律性を付け加えて行くというアプローチが望ましい。

### 7.3.4 インタラクション

キャラクターと人間とのインタラクションでは以下のような点が重要である。

#### (1) マルチモーダルインタラクション

人間同士のコミュニケーションは、マルチモーダルインタラクションによって行われている。具体的には、対面型のコミュニケーションにおいては、音声の他に顔の表情、身振り、手ぶりなどによりコミュニケーションが行われている。したがって、動作もしくは音声といった単一のモダリティのみを使うのではなく音声と動作といった複数のモダリティを使うことが望ましい。このことがより自然なインタラクションの実現、ひいてはより深いレベルの没入感、感情移入につながると考えられる。

また、このような人間同士のインタラクション、コミュニケーションにおいて通常用いられるモダリティの他にも、インタラクティブシアターを新しいメディアと位置付けるなら、別のモダリティを利用することも可能である。特に音声認識・画像認識がいまだ十分成熟した技術ではなく、人間の能力に比較するとコンピュータで処理した場合は格段に認識誤りが生じることを考慮すると、より確実な手段を用いることも必要である。このような手段としてボタン入力・フットセンサー・モーションキャプチャーなどを使うことが考えられる。

#### (2) ノンバーバルインタラクション

人間のコミュニケーションにおいては言語的な（バーバル）情報と非言語的な（ノンバーバル）情報が送受されている。非言語的な情報としては個人性に関する情報や感情・感性的な情報がある[5]。残念ながら、従来のコミュニケーション技術の研究においては、言語情報の送受をサポートする技術の研究のみに重点が置かれてきた。しかしながら、感情・感性的な情報の送受は我々のコミュニケーションにおいて重要な役割を果たしている。仮想世界のキャラクターと人間の非言語コミュニケーションを実現するためには、キャラクターが感情の認識・表現を行える機能をもつ必要がある。

筆者らは、人間の音声に含まれる感情を認識してそれに反応する機能をもったキャラクターを作り出す研究を行ってきた[70]。展示等を通してわかったことは、人が性別や年齢、さらには言語の壁を越えて、このキャラクターとのインタラクションに没入するということである。この経験からすると、感情によるコミュニケーションがインタラクティブシアターの成功の鍵であると考えられる。

## 7.4 システム構成例

以上の考察に基づいて、現在、インタラクティブシアターのシステム開発を進めた。以下に具体的内容を述べる。

### 7.4.1 特徴

#### (1) 臨場感のある映像表現

アーチ型のスクリーンに立体映像を投影することにより臨場感に富んだ環境設定を行ない観客をインタラクティブシアターの世界に引き込むことを狙った。

#### (2) 自然なインタラクション

音声認識、ジェスチャー認識機能を用いることにより、人間にとって自然な音声、身振りによるインタラクション実現を図った。

#### (3) マルチストーリー用CGアニメーション

インタラクティブストーリーはストーリー展開が複雑になるため、それに応じた膨大なアニメーションを用意する必要がある。CGキャラクターに自律性を持たせればこの問題を回避出来るがアニメーションの自然性が損なわれる。ここではあくまでアニメーションの完成度を優先し、すべての必要なアニメーションを事前に用意する方式をとった。

### 7.4.2 ソフトウェア構成

ソフトウェア構成を図7.8に示す。

#### (1) スクリプトマネージャー

ストーリー作家の作る台本、シナリオから具体的なインタラクティブなシナリオを生成し、インタラクティブストーリーの展開全体をコントロールする部分である。具体的には、各シーンの構成要素の定義、各シーン間の遷移の制御などを行なう。

(2) シーンマネージャー

各シーンを記述・コントロールするマネージャーである。各シーンは背景と登場人物から構成される。背景としては、CGで構成された背景、実写をベースとした背景を混在させて用いている。また、登場人物については、CGで各キャラクターを生成した。

(3) インタラクションマネージャー

スクリプトマネージャー、およびシーンマネージャーの下にあって、各シーンにおけるインタラクションを制御する。音声・動作によるインタラクションを用いることとし、音声認識・画像認識機能を採用すると共に、それらを組み合わせてマルチモーダルインタラクションの機能を実現した。

(4) 各種のハンドラー

シーンマネージャー、もしくはインタラクションマネージャーの下にあって、実際に各種の入力装置や出力装置を制御する機能を持つ。具体的には以下のようなハンドラーを用意した。

- a) 音声認識ハンドラー
- b) 動作認識ハンドラー
- c) キャラクタハンドラー
- d) サウンドハンドラー
- e) イメージハンドラー

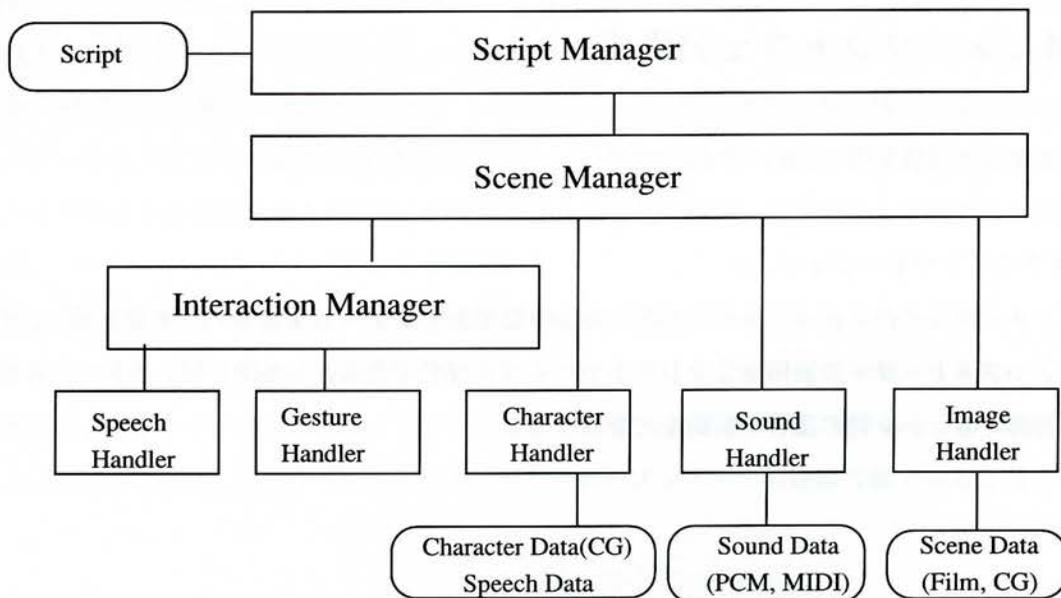


図7.8 インタラクティブシアターのソフトウェア構成

### 7.4.3 ハードウェア構成

図7.9にハードウェア構成を示す。CG生成用高速WS（Onyx Infinite Reality）を中心に、音声認識処理用WS、動作認識処理用WS、音声・サウンド出力用WSによりシステムを構築した。また、映像出力は2台のプロジェクターによりアーチ型のスクリーンに投影される。

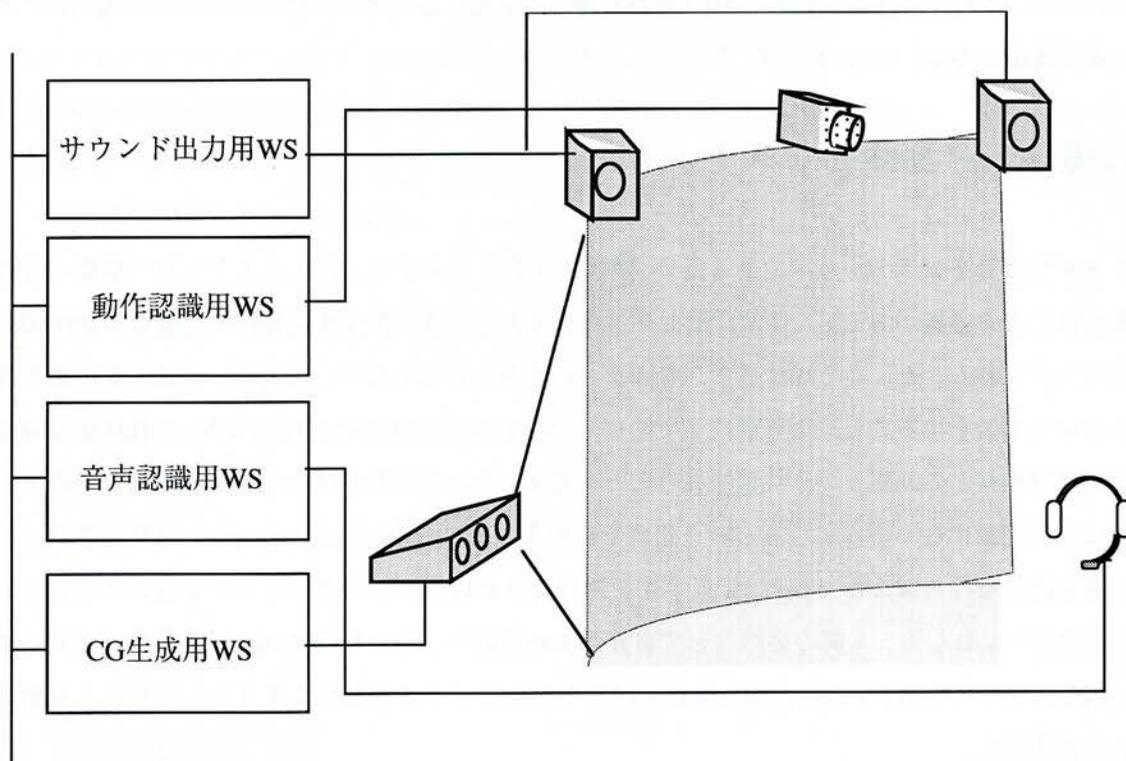


図7.9 インタラクティブシアターのハードウェア構成

## 7.5 インタラクティブストーリーの構成例

### 7.5.1 コンセプト

我々日本人は、子供の頃、昔話を聞かされて育つ。語り継がれた昔話にはのほのほとした印象があるのは、大人の子供に対する愛があるからであろう。原書では、必ずしも暖かい話ではなく、話も単純ではない。隠された意味も様々な解釈があり奥深い。だから時代を超えて語り継がれ、いぶし銀のように魅惑的である。インタラクティブシアター用脚本は、日本古来の「浦島太郎」と、西洋の「不思議の国のアリス」から現代風に脚色した。採用した根拠は、「浦島太郎」は、現実の世界から竜宮城という仮想世界へ旅だって、再び戻ってくるという物語の構造を持ち、極めてVRを用いた物語の表現に適している。「不思議の国のアリス」も、また同じ様な構造を持ち、物語のビジュアル的表現が効果的である。このような理由から、この2つの物語を脚色してストーリーを作成することにした。おとひめ様を「ミューズ」(美の女神)に変更するなどの現代風アレンジを施し、タイトルを「Wonderland」とした。

### 7.5.2 具体的なストーリー

インタラクティブストーリーであるため、随所にインタラクションによるストーリー展開の分岐点を設けておく必要がある。人は常に選択を迫られており、選択を行うことにより新しい状況の中へ投げ入れられる。そしてそれは必ずしも好ましい状態とは限らない。例えば、浦島太郎は亀を助けるという選択をすることにより龍宮城での楽しい一時を過ごすことが出来るが、それは結局現実世界の時間推移との乖離という代償を払うことになる。また玉手箱を開けるという行為を選択することにより一瞬にして老人になるという悲劇を味わうことになる。しかしながら、この昔話は、決して亀を助けなければよかったとか、玉手箱を開けなければよかったということを言おうとしているのではない。むしろ、人間の必然として自分が決断を行い、そしてその結果必ずしも好ましい結果に結び付かなくても、それに直面しなければならないということを伝えようとしていると解釈することが出来る。

### 7.5.3 インタラクション

観客は、立体視用液晶シャッターめがねをかけ、マイクを持ちスクリーンの前に立つ。映像と音で提示されるストーリーを見ながら、映画の中のキャラクターなどとのインタラクションを行ないつつ、ストーリー展開に参加しそれを体験する。上記のようなストーリー展開の分岐点では音声や動作によりインタラクションを行なう。もちろん、龍宮城にとどまる、玉手箱をあけないなどの選択をすることが可能であり、それによってそれ以降のストーリーの内容は原作とは異なったものとなる。図7.10、図7.11にインタラクションの様子を示す。

城の中では、インタラクティブポエム、インタラクティブミュージックなどのインタラクティブアートを導入している。インタラクティブポエムは、音声認識により観客とコンピューターキャラクターが、あらかじめフレーズごとに切ってある詩のフレーズを選んで、連歌のように掛け合いで即興詩を詠んでいくものである。インタラクティブミュージックは、観客がポップスの曲に合わせて、キーボードを即興演奏することで、そのひき方（高低、強弱、音数）によって、インタラクティブにダンシングシンガーのダンスが変わるものである。例えば、高音域で強くひくと、力強く踊り、単音でゆっくりした退屈な演奏をすると、座りこんだり、眠りだしたりする。

また、インタラクションに基づいたストーリーの展開を実現するため、随所に分岐点が設けてあり、例えば下記のようなものがある。

- (1) いじめられている兎を助けるか、否か
- (2) 兎を助けなかった場合、あやまるか、否か
- (3) 城で、どのアトラクションを選ぶか
- (4) 城にとどまるか帰るか
- (5) 玉手箱を開くか、否か

これらは、誘惑である。その誘惑にのった結果、期待通りの結果になるとは限らないが、誘惑とは魅力的な物である。そこに観客の心理が働き、そこで起こる物事を主観的に受け止め、物語を体験する。

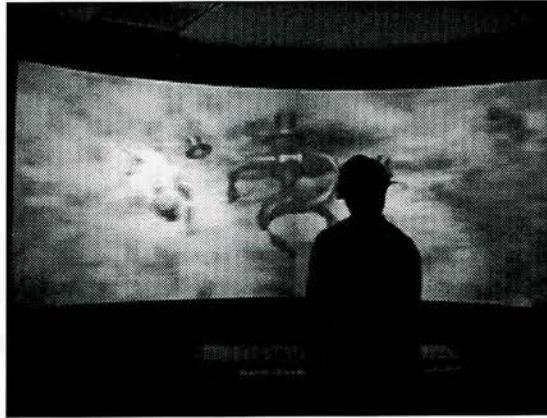


図7.10 インタラクティブシアター「ワンダーランド」より  
「インタラクティブミュージック」のシーン

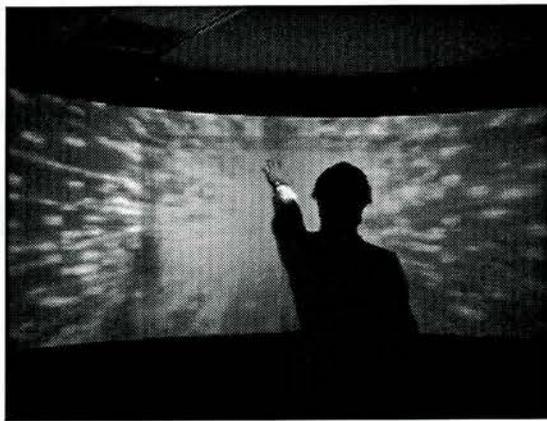


図7.11 インタラクティブシアター「ワンダーランド」より  
たまたま箱に触れるシーン

## 7.6 まとめ

通信・放送・映画・演劇・テレビゲーム等の種々のメディアを統合した新しいメディアの1つと考えられる、インタラクティブシアターについて述べた。インタラクティブシアターはサイバースペースの中に人が入り、サイバースペースの他のキャラクターとインタラクションをしながらストーリー展開を楽しむという新しいメディアである。このような新しいメディアが従来のメディアとどのような関係にあるのか、またそのようなメディアを実現するために必要な技術はなにか、具体的なシステム構成はどのようなものか等について述べた。また、具体的な実現例として、筆者らが作成した第1次システムの構成、および日本のおとぎ話に題材をとったインタラクティブストーリーの構成・内容について述べた。

## 8 感情に反応するインタラクティブシアターの構成2

### 8.1 まえがき

19世紀末にルミエールらによって発明されて以来[92]、映画は技術およびコンテンツ制作の面で進化を繰り返し、現在では、アートからエンターテインメントまでの広い領域を含む複合芸術としての地位を確立している。また最近の映画はデジタル技術やコンピュータグラフィクス技術を取り入れ、新しい世代の映画へと移行しつつある。デジタル技術やコンピュータグラフィクス技術は、従来の映画にはなかった超現実的な世界、すなわちサイバースペースを作り出す能力を与えてくれる[93][94][95]。しかしながら、従来の映画は、仮想の世界（サイバースペース）とそこにおけるストーリーを観客に一方的に与えるという形態をとっていた。これに対し、インタラクション技術を取り入れると、観客自身が主人公となって、サイバースペースに入り、主体的にストーリーを体験することが可能になる。このような観点から、筆者らは従来の映画にインタラクション技術を導入したインタラクティブ映画システムの検討を行なっている。すでに、その第1次システム[96][97]を試作した。第1次システムをベースとして、改良版である第2次システムを構築した。ここでは、第1次システムの評価および問題点と、それを改良した第2次システムの構成を述べる。

### 8.2 第1次システムの評価と問題点

本システムは、開発以来約半年にわたって所内の研究者や当所への見学者など約50名に体験してもらった。これらの体験者の感想などに基づき第1次システムの問題点をまとめると以下のようになる。

#### (1) サイバースペースへの参加者数

第1次システムでは、1人の参加者が主人公となってストーリーを体験するという形態をとっていた。しかしながら、サイバースペースはネットワーク上に構築されるわけであるから、サイバースペースには1人ではなく複数人が同時に参加してストーリー展開を体験できることが望ましい。

#### (2) インタラクションの頻度

インタラクションが原則的にはストーリーの変化時点のみに限られていた。そのため、それ以外の部分ではストーリーは映画と同様あらかじめ決められたリニアな時間軸に沿って進んでいた。こ

のため、参加者が「観客」になってしまい、インタラクションの必要な場面で参加者としてインタラクションに加わるのを忘れてしまうという欠点を持っていた。

## 8.3 第2次システムの概要

### 8.3.1 改良点

上記の点を考慮して第2次システムでは、以下の点を改良することとした。

#### (1) 複数人参加型システム

システムの将来の狙いは、ネットワークを介した複数人参加システムであるが、今回はその手始めとして、LANで接続された2システム間での複数人参加型システムを試作した。

#### (2) 任意の時点でのインタラクション (anytime interaction) の実現

参加者とシステムのインタラクションの頻度を上げるため、任意の時点で本人とサイバースペースのキャラクタとのインタラクションが可能な仕組みを取り入れることとした。原則として、このインタラクションは即興的なインタラクションであって、ストーリー展開に影響を及ぼさない。このようなインタラクションを story unconscious interaction (SUI) と呼ぶことにする。これに対し、ストーリーの分岐点におけるインタラクションであって、ここでの結果がそれ以降のストーリー展開を決定するものを story conscious interaction (SCI) と呼ぶことにする。

#### (3) その他

参加者のアバターの表示、照明が暗い条件下でのジェスチャー認識等を可能にするため、磁気センサを用いて参加者の動きを取り込む方式を採用した。

これらの機能を取り入れた第2次システムの概観を図8.1に示す。

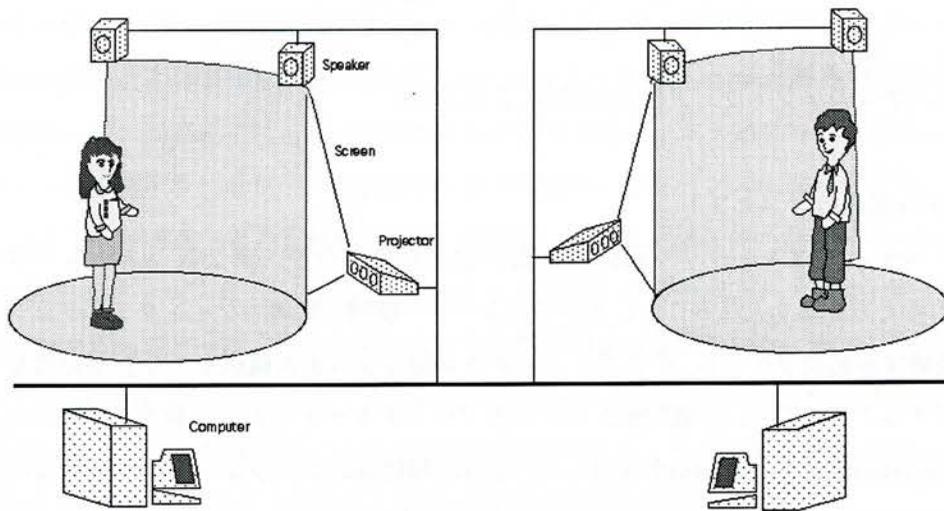


図 8.1 第 2 システム構成図

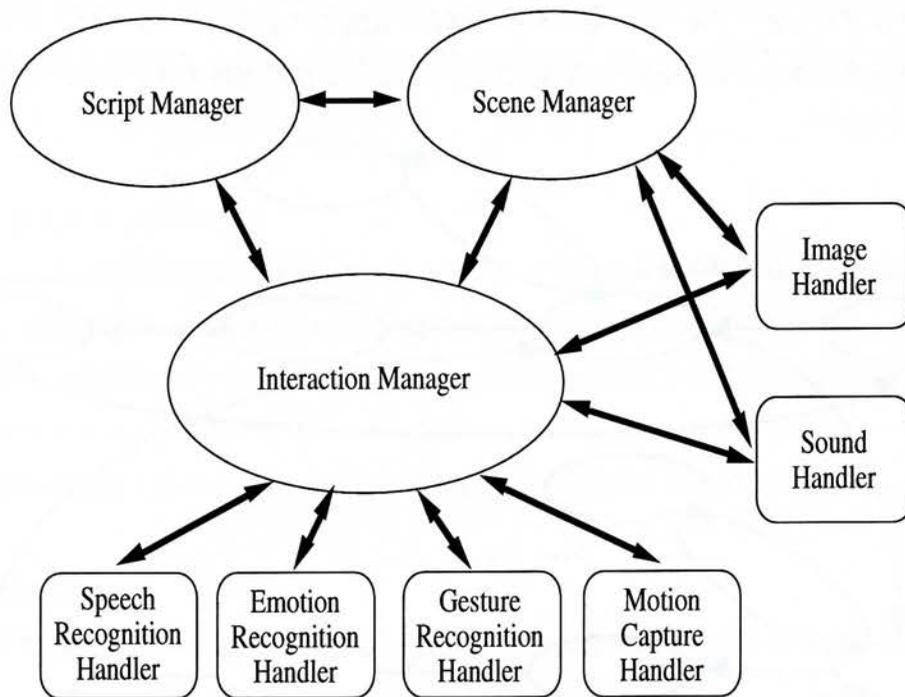


図 8.2 第 2 次システムのソフトウェア構成

## 8.3.2 ソフトウェアシステム構成

第2次システムのソフトウェア構成を図8.2に示す。

### (1) システム構成のコンセプト

第1次システムではストーリー展開に重点がおかれていたが、第2次システムでは anytime interaction の考え方を取り入れることによってストーリー展開と即興的なインタラクションとをバランス良く制御する必要が生じる。そこでトップダウン的なシステム構成から分散制御的なシステム構成へ移行することとした。分散制御システムのアーキテクチャーとして種々のものがあるが、ここでは action selection network [98] を採用した。これは複数個のエージェント間で活性値の送受が行なわれ、活性値がいき値を越えたエージェントが活性化し、そのエージェントに付随したプロセスが動作するというものである。

### (2) スクリプトマネージャー

スクリプトマネージャーの役割は第1次システムと同様であり、シーン間の遷移を制御する。インタラクティブストーリーは、種々のシーンの集合およびシーン間の状態遷移図で表現される(図8.3)。スクリプトマネージャーは、この状態遷移図を記憶しておき、シーンマネージャーから送られてくるインタラクション (SCI) の結果に応じて、シーン間の遷移をコントロールする。

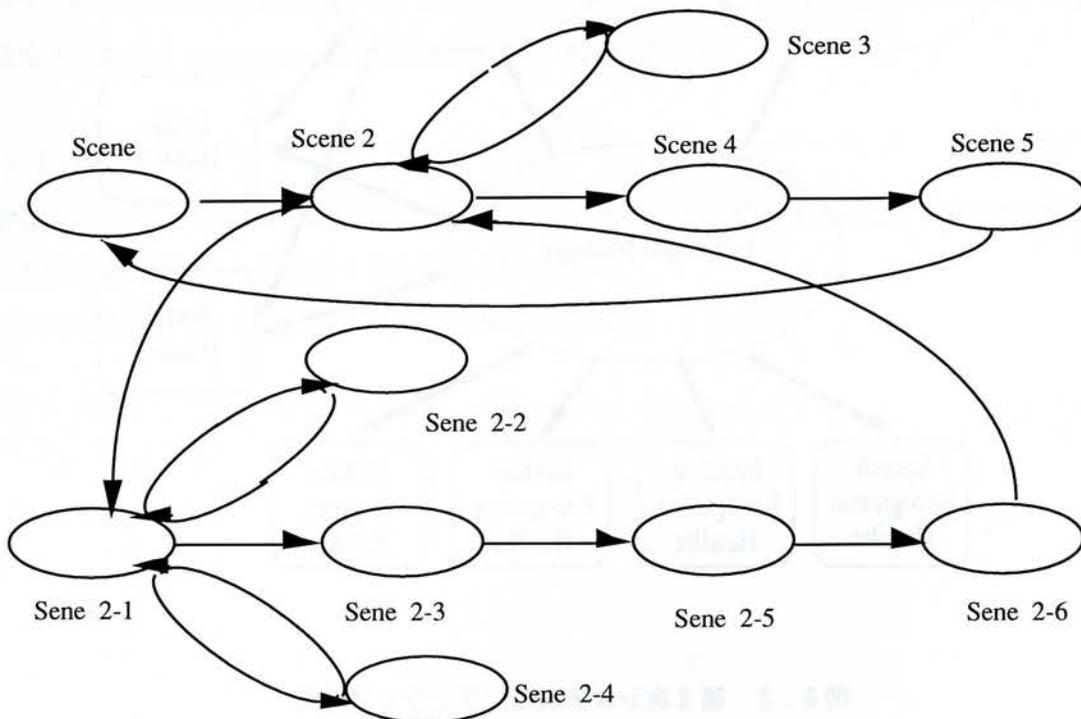


図8.3 シーン間遷移図

### (3) シーンマネージャー

シーンマネージャーはシーンの記述およびシーン内のストーリー進行を管理する。シーン内のストーリー進行に関係ある出来事を event と呼び、シーンマネージャーは event 遷移の制御を行なう。あらかじめ各シーンの内容記述は event network として蓄えてある。シーンマネージャーは、スクリプトマネージャーから指示されたシーンの記述データを参照し、各シーンの生成を行なう。シーン毎の event は、a) シーン映像、b) 背景音楽、b) 効果音、c) キャラクターのアニメーションおよび台詞、d) 参加者とキャラクターのインタラクション (SUI) から構成されている。第1次システムでは、シーンマネージャーは、これらを出力する時間管理を行っていた。しかしながら、第2次システムでは、anytime interaction の考え方を取り入れているため絶対的な時間の管理は出来ない。可能なのは相対的な時間の管理もしくは時間順序の管理である。そこでここでは action selection network の考え方を取り入れることとした。動作の概要は以下の通りである。

- 1) event は他の event および外部と活性値の送受を行なう。
- 2) 活性値の累積値がいき値を越えると event が活性化する。
- 3) event の活性化に伴い event に伴う action が発現する。また、他の event に活性値が送られると共に、その event の活性値はリセットされる。

活性値の送られる方向、強さなどを定めておくことにより、event の発生順序をあらかじめ定めておくことが出来ると同時に、event の発生順序にゆらぎやあいまいさを導入することも可能である (図8.4)。

### (4) インタラクションマネージャー

anytime interaction の実現には、インタラクションマネージャーが最も重要な働きをする。anytime interaction は、感情認識と action selection network を組み合わせることにより実現する。具体的には、各キャラクター (参加者のアバターもキャラクターの1つと考える) に感情状態を割り当て、参加者とのインタラクション、およびキャラクター相互のインタラクションがキャラクターの感情状態を決定し、それに応じて各キャラクターの反応が決まるという構造を考える (図8.5)。

#### a) 感情状態の定義

参加者( $i=1,2,\dots$ )の時刻  $T$  における感情状態および、その強度を以下のように定義する。

$$E_p(i,T), sp(i,T) \text{ where } sp(i,T)=0 \text{ or } 1 \quad (8.1)$$

(0 は入力がない場合、1 はある場合をさす。)

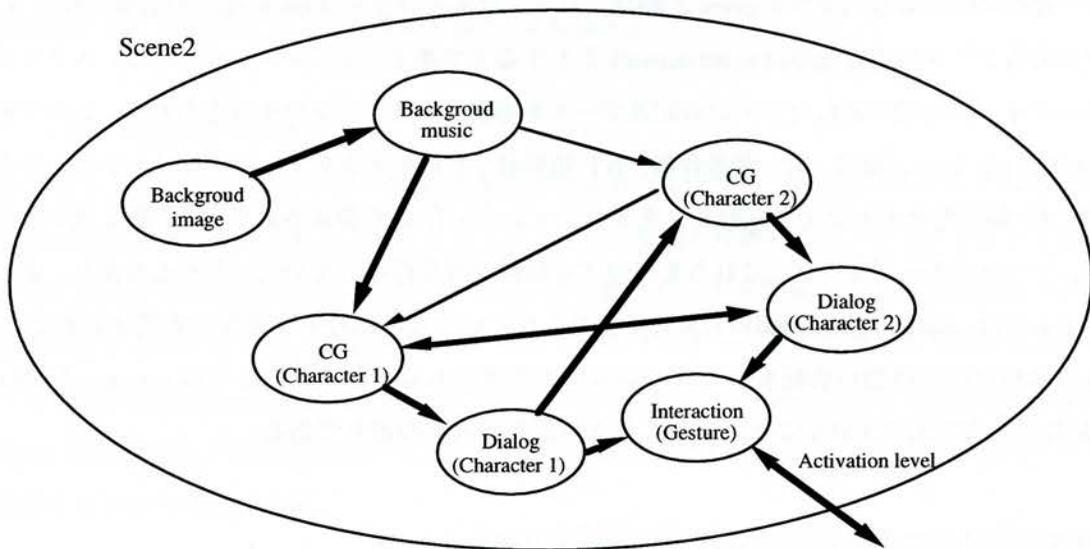


図8.4 シーンマネージャーの構成

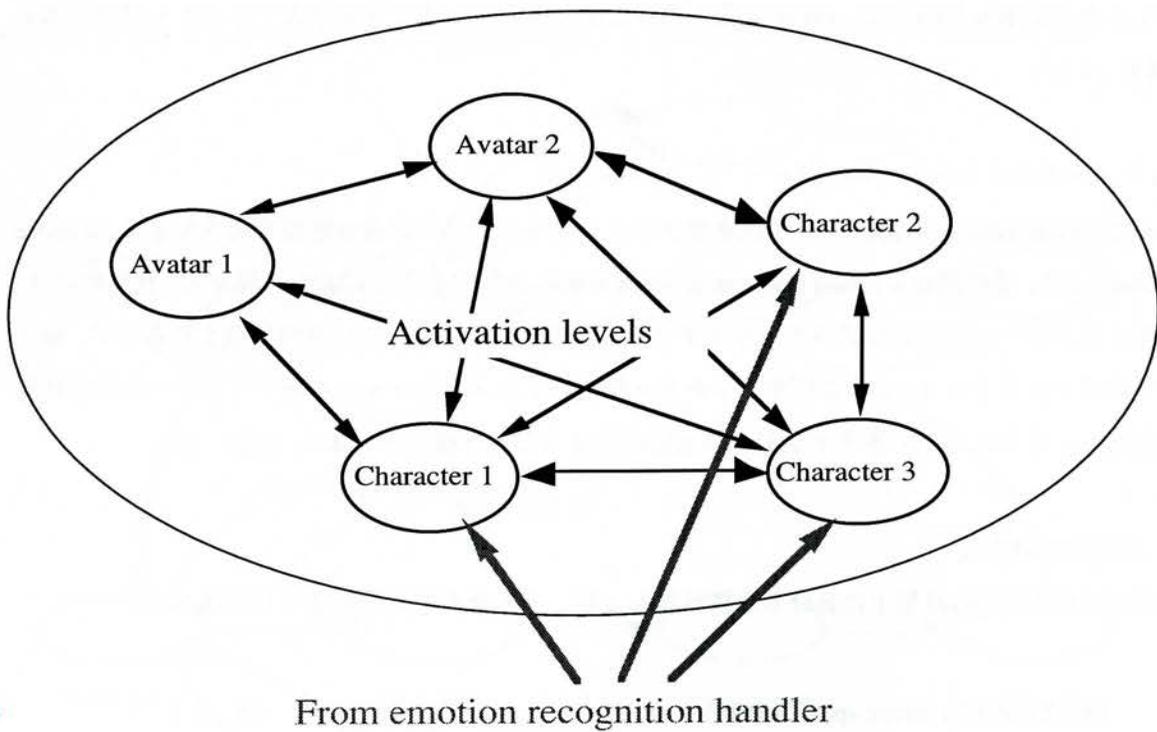


図8.5 インタラクションマネージャーの構成

同様に、キャラクター( $i=1,2,\dots$ )の時刻  $T$  における感情状態、および、その強度を以下のように定義する。

$$Eo(i,T), so(i,T) \quad (8.2)$$

b) キャラクターの感情状態の決定

簡単のため、キャラクターの感情状態は参加者からの感情認識結果が得られた場合、その感情状態によって決定することとする。

$$\{Ep(i,T)\} \rightarrow \{Eo(j,T+1)\} \quad (8.3)$$

感情認識結果が入力されると、各キャラクターに活性値が送られる。

$$sp(i, T) \rightarrow sp(i, j, T) \quad (8.4)$$

$sp(i, j, T)$  は、参加者  $i$  の感情認識結果に基づいてキャラクター  $j$  に送られる活性値である。キャラクター  $j$  の活性値は送られてくる活性値の総和となる。

$$so(j, T+1) = \sum sp(i, j, T) \quad (8.5)$$

c) アクションの発現

活性値がいき値を越えたキャラクターは、アクション  $Ao(i, T)$  をおこす。アクションは感情状態によって定まる。具体的にアクションとは、参加者の感情に応じた、キャラクターの動作、台詞によるリアクションをさす。同時に、他のキャラクターに活性値  $so(i, j, T)$  が送られる。

$$\begin{aligned} & \text{if } so(i, T) > TH_i \\ & \text{then } Eo(i,T) \rightarrow Ao(i, T), Eo(i, T) \rightarrow so(i, j, T) \\ & so(j, T+1) = \sum so(i, j, T) \end{aligned} \quad (8.6)$$

この仕組みにより、キャラクター同士の相互作用が発生し、感情認識結果とキャラクターのリアクションが1対1に対応する単純なインタラクションに比較して多様なインタラクションが可能となる。action selection network で表現されたインタラクションマネージャーの構成を図8.5に示す。

### 8.3.3 ハードウェアシステム構成

図8.6に第2次システムのハードウェア構成を示す。映像出力サブシステム、音声・感情認識サブシステム、動作認識サブシステム、サウンド出力サブシステムより構成される。

#### (1) 映像出力サブシステム

CG生成用高速WS 2台(Onyx Infinite Reality および Indigo2 IMPACT)を映像出力のためのWSとして用いている。Onyx上には、スクリプトマネージャ、シーンマネージャ、インタラクティブマネージャ、映像出力ハンドラーの各ソフトウェアがインプリメントされている。キャラクターの映像はCGアニメーションデータとしてあらかじめ蓄えておき、リアルタイムでCGが生成される。背景のCG映像もデジタルデータとして蓄えておき、リアルタイムで背景映像が生成される。背景の一部は実写映像を用いており、これは外付けのLDに蓄えておく。これら複数キャラクターのCG、背景のCG、背景の実写映像はOnyx および Indigo2 のビデオボードでオーバーラップ処理が行なわれる。

臨場感に富んだ映像生成のために、CG映像は立体表示させる。また、参加者を映像で取り囲みインタラクティブ映画の世界に没入させるため、アーチスクリーンを採用した。あらかじめ、左眼用、右眼用の2種の映像データをWSで作成しておき、立体視コントローラを介してこれらを混合すると共に、2台のプロジェクターによりアーチ型のスクリーンに投影する(図8.7)。ただし、Indigo2側は処理速度等の問題のため立体視は採用しておらず、また映像出力も通常的大型ディスプレイで行なう。

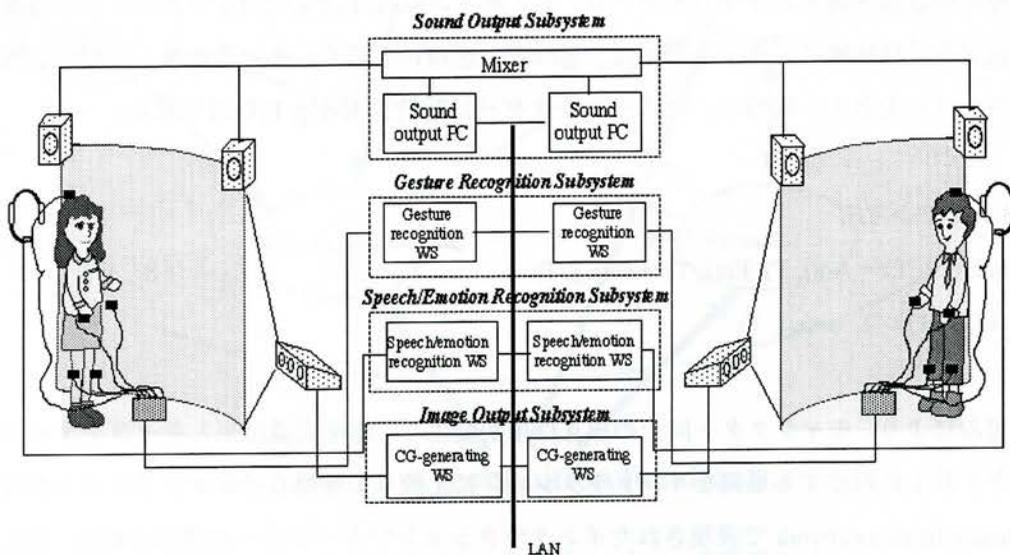


図8.6 第2次システムのハードウェア構成

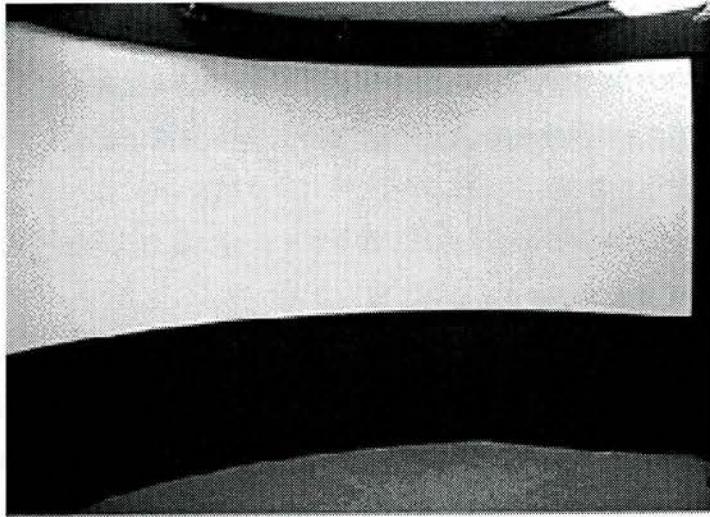


図8.7 アーチスクリーンと立体視用プロジェクター

## (2) 音声・感情認識サブシステム

音声および感情認識は2台のWS (SUN SS20×2) で実行される。これらのWS上には音声認識ハンドラー、感情認識ハンドラーもインプリメントされている。マイクから入力された音声はSUN内蔵の音声ボードによりAD変換され、音声認識ソフト、感情認識ソフトにより音声認識、感情認識が実行される。2台のWSが2人の参加者それぞれの音声入力を処理する。音声認識はHMMに基づく不特定話者、連続音声認識機能を持つ[99]。また、感情認識はニューラルネットを用いて行なわれる[54]。

## (3) 動作認識サブシステム

動作認識は2台のSGI Indyで実行される。Indy上には動作認識ハンドラーもインプリメントされている。それぞれのWSは、2人の参加者の体に装着された磁気センサーからの出力を取り込み、アバター制御用のデータとして用いると共に、ジェスチャーの認識も行なう。ジェスチャー認識は磁気センサーの出力である位置データの時間変化を特徴量とし、HMMを用いて行なっている。

## (4) サウンド出力サブシステム

サウンド出力サブシステムは複数台のPCより構成されている。同時に出力する必要があるサウンドは、背景音楽、効果音、キャラクター毎の台詞である。効果音、キャラクターの台詞はデジタルデータとして蓄えておき、必要に応じてDAを行なう。これらのサウンドの同時出力をサポートするため、複数チャンネルの同時DAが可能ないように複数台のPCを用意してある。また、背景音楽はあらかじめ外付けのCDに記録しておき、その出力制御をPCから行なう。これら複数のチャンネルより出力されたサウンドはコンピュータ制御可能なミキサー(ヤマハO2R)でミキシングされ出力される。

## 8.4 インタラクション技術

### 8.4.1 ジェスチャー認識

ジェスチャー認識については種々の研究が行なわれている。それらは体全体の大きな動きの認識[9]と手先の細かな動きの認識[11]の2種類に大別される。しかしながら、日常のコミュニケーションで生じる身振り・手振りはこの中間的なものが多い。また、このような動作により非言語的情報が表出・認識される。映画の世界に没入するためには非言語コミュニケーションが重要であり、そのためにもジェスチャーの認識は必要である。以上のような考察の基にHMMを用いたジェスチャー認識を取り入れることとした。

#### (1) ジェスチャーのデータ取得

オクルージョンの問題や、ジェスチャーの持つ3次元空間的特徴からすると、カメラ入力を用いることはデータ取得段階でかなりの困難さを持つ。したがってここでは、磁気センサーを体の適切な位置に装着し、複数の磁気センサーからの入力情報をジェスチャーに関するデータとして用いた。本センサーでは、3次元の位置座標およびセンサーの回転角度の情報が得られる。これらのうち、今回は3次元位置座標を用いることとした。

磁気センサー*i*の時間*t*における3次元位置座標を( $x_i(t)$ ,  $y_i(t)$ ,  $z_i(t)$ )とする。N個の磁気センサー(1,2,...,N)を用いることとした場合、以下のデータを時間*t*における特徴パラメータとして用いた。

$$\begin{aligned} v(t) = & (v_{x1}(t), v_{y1}(t), v_{z1}(t), v_{x2}(t), v_{y2}(t), \\ & v_{z2}(t), \dots, v_{xN}(t), v_{yN}(t), v_{zN}(t)) \end{aligned} \quad (8.7)$$

(ただし、 $v_{xi}(t) = x_i(t) - x_i(t-1)$   $v_y, v_z$  も同様)

入力の特徴量は  $v(t)$  の時系列として与えられる。

#### (2) HMMによる認識と学習

認識処理ではHMMを用いて特徴ベクトルの生成確率を算出する。HMMの構成は図8.8に示したようにleft-to-right型とする。

HMMは以下で定義できる。

- a) 状態:  $s = \{s_1, s_2, \dots, s_N\}$
- b) 出力シンボル:  $O_1, O_2, \dots, O_M$
- c) 状態遷移確率:  $a_{ij}$  (状態  $s_i$  から状態  $s_j$  に遷移する確率)

d) 出力確率:  $b_j(O)$  状態  $s_j$  においてベクトル  $O$  を出力する確率

このとき、あらかじめ  $O$  の学習データが与えられたとき、最適な状態遷移確率、および出力確率を求めておく、認識時には、出力されたシンボルの系列が与えられたとき、これを出力する最適の状態遷移およびその場合の出力確率が得られる。したがって、カテゴリー毎に学習した HMM を用意しておく、観測されたシンボル確率が与えられたとき、それを出力する最適の HMM が決定される。

### (3) 学習データ

あらかじめ、6種の感情（喜び、怒り、驚き、悲しみ、軽蔑、恐れ）に対応したジェスチャーを定めておき、これを学習データとして用いる。学習データは5人の被験者から一人、各ジェスチャーあたり10サンプルずつ取得し、上記の学習アルゴリズムを用いて学習を行なった。

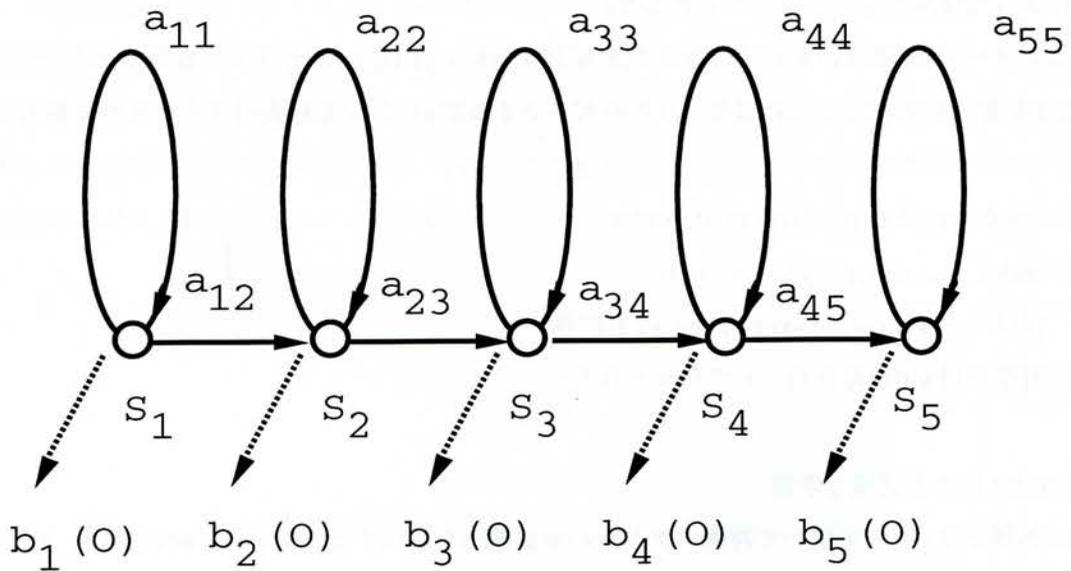


図 8.8 left-to-right HMM の構成

## 8.4.2 感情認識

anytime interaction を実現するための鍵として感情認識を採用した。通常の声認識では、入力できるのは基本的には認識対象の語彙に限定されるため、任意の時点で任意の音声入力を行なわせようとする anytime interaction の考え方には適合しない。そこで、不特定話者、内容不依存タイプの感情認識を用いることによりこの問題の解決を図った。認識アルゴリズムとしてはニューラルネットを採用することとし、大量の学習データを用いることにより、安定した性能を得ることを狙った。認識対象の感情は、「喜び、怒り、驚き、悲しみ、軽蔑、からかい、恐れ、普通」の8種類である。

図8.9は処理の流れのブロック図である。音声特徴抽出部、感情認識部から構成されている。

### (1) 音声特徴抽出

感情認識のために、音韻特徴を表わすパラメータと韻律特徴を表わすパラメータを用いる。音韻特徴パラメータとしてはLPCパラメータを用いる。韻律特徴としては、エネルギー、音韻特徴の時間変化、およびピッチを用いる。入力音声デジタル化された後、適当な長さのフレーム毎にLPC分析が行われ、以下の特徴パラメータが求められる。

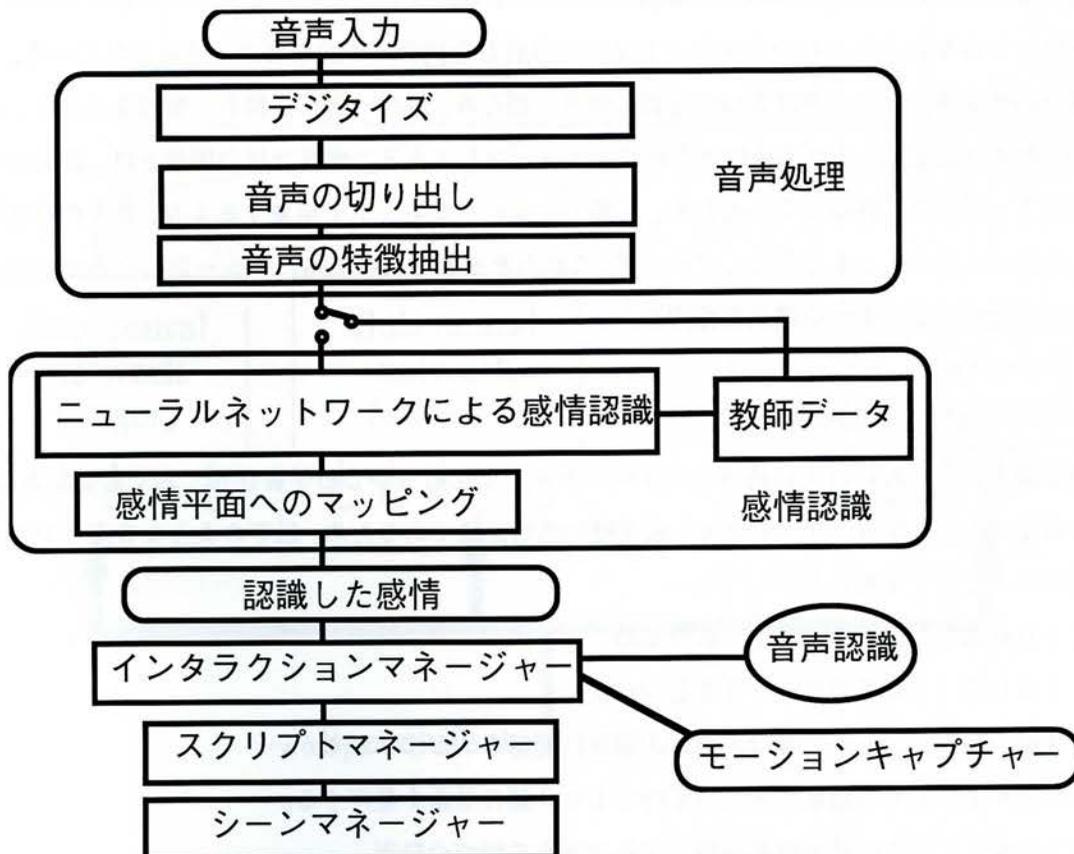


図8.9 感情認識のフローチャート

$$F_t = (P_t, p_t, dt, c1t, c2t, \dots, c12t) \quad (8.8)$$

ただし、 $P_t, p_t, dt, (c1t, c2t, \dots, c12t)$ は、 $t$ フレームに関する音声パワー、ピッチ、時間変化パラメータ、LPCパラメータである。次に、このパラメータの時系列から音声パワーを用いて音声区間が抽出される。抽出された音声区間の全体から、等間隔になるように配置された10フレームを取り出す。これらの10フレームの特徴パラメータをまとめることにより、音声の特徴量は150次元(15×10)の特徴ベクトル、

$$FV = (F_1, F_2, \dots, F_{10}), \quad (8.9)$$

として表現される。ここで、 $F_i$ は、 $i$ フレームの特徴パラメータである。FVは、感情認識部への入力として用いられる。

## (2) 感情認識

感情認識のためのニューラルネットの構造を図8.10に示す。このネットワークは8つのサブネットワークの集合とそれらのサブネットワークの出力を統合する論理部から構成されている。8つの各々のサブネットワークは8つの感情(怒り、悲しみ、喜び、恐れ、驚き、愛想をつかす、からかい、および普通)のそれぞれにあらかじめチューンしてある。感情認識の困難さは、認識すべき感情によって大きく異なっているため、1個のニューラルネットを用意するより、各々の感情に対応したニューラルネットを用意しておいて、これらをそれぞれの感情にチューンした方がいいと考えられるため、このような構造を採用した。

## (3) ニューラルネットの学習

感情認識を行うためには上に述べたニューラルネットをあらかじめ学習させておく必要がある。我々の目標は不特定話者、コンテキスト独立型の感情認識であるため、以下のような音声サンプルを学習データとして用意した。

単語：100個の音韻バランスがとれた単語

話者：100名(50名の男性と50名の女性)

感情：普通、怒り、悲しみ、喜び、恐れ、驚き、愛想をつかす、からかい

音声サンプル1：各々の話者が8つの感情で100個の単語を発声する。

音声サンプル2：各々の話者が各感情毎に各母音を5回づつ発声

この学習データを用いて予備実験を行った。その結果、男性、女性をまとめたニューラルネット

を用意するより、男性、女性それぞれにチューンしたネットワークを用意する方が学習、認識共に有利であることがわかった。

#### (4) ニューラルネットによる感情認識

感情認識の際には、音声特徴抽出部で得られた音声特徴量が、上に述べた方法で学習が行われた8つのサブネットワークに入力される。その結果として、8つの出力が得られる。最大の出力を与える感情が認識結果である。

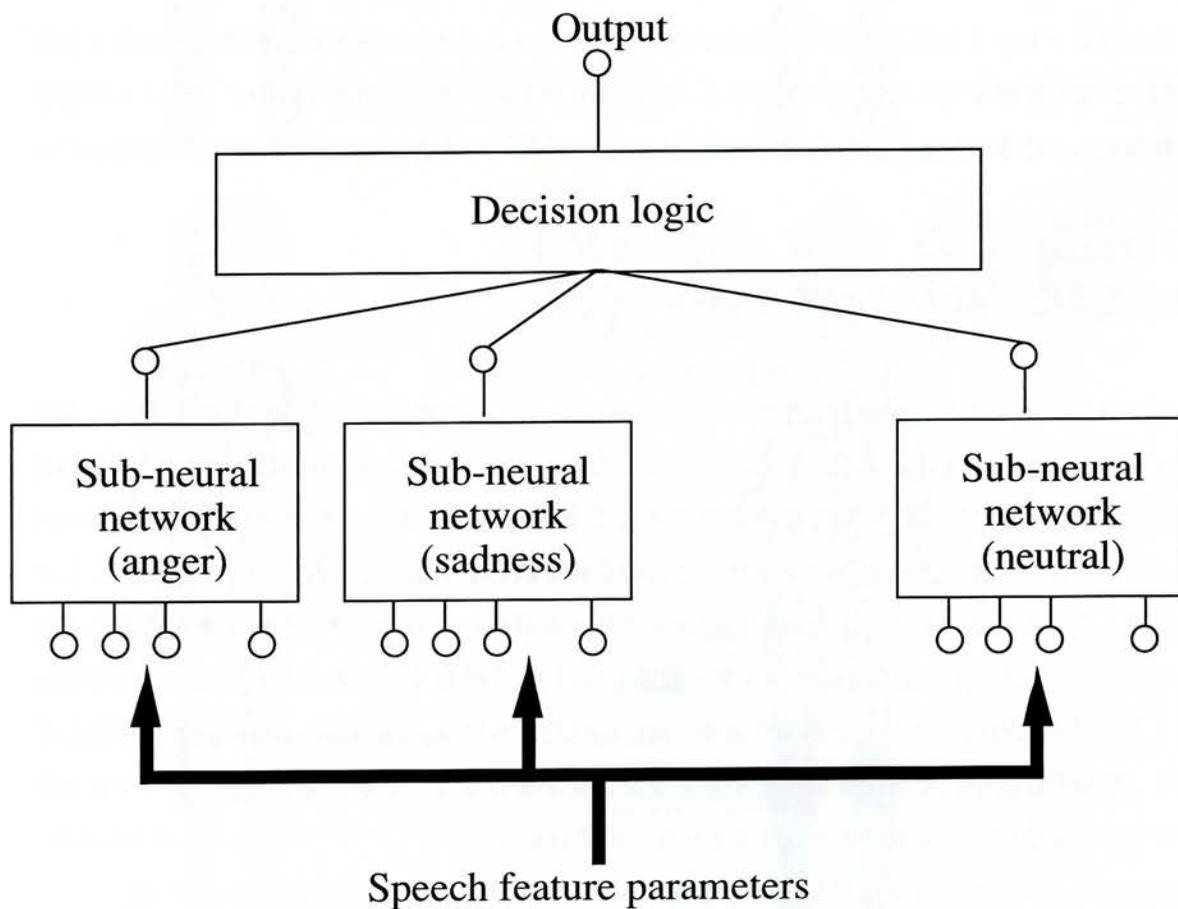


図8.10 ニューラルネットによる感情認識

## 8.5 インタラクティブストーリーの構成例

### 8.5.1 インタラクティブストーリー

以上述べたシステムの上に具体的なインタラクティブストーリーを構築した。ストーリーのベースとしてシェークスピアの「ロミオとジュリエット」[100]を採用した。これは以下の理由による。

(1) 主人公が二人（ロミオ、ジュリエット）いるため、複数人参加型のシステムに適したストーリーである。

(2) 誰でも良く知っているストーリーであり、参加者が容易にインタラクティブストーリーの中に没入し、相手や、ストーリー中のキャラクターとのインタラクションを行なってくれる。

核となるプロットは以下の通りである：ロミオとジュリエットは、彼等の悲劇的な死の後「黄泉の国」に魂と記憶を失った状態で送られる。黄泉の国のキャラクターのガイドやアドバイスに助けられながら、彼等は徐々に自分自身を再び見出し、最後に魂を再び手に入れて現世へと帰っていく。

### 8.5.2 インタラクション

本システムは二人の人が同時に参加できる。上のストーリーの例では一人がロミオをもう一人がジュリエットの役割を演じる。図8.6に示した2台のシステムは別の場所に設置されており、LANを介して接続されている。(同じ場所に設置することも可能である。この場合には、参加者＝パフォーマーによって演じられるストーリー進行の全体を観客が楽しむというパフォーマンスとして上演することも可能である。)各々の参加者はスクリーンの前に磁気センサーとマイクを装着した状態で立つ。ロミオ役は3Dの液晶シャッター眼鏡をつけ3Dの映像を楽しむことが出来る。基本的にはストーリー進行はシステムが制御するが、先に述べたようにanytime interactionの機能を採用しており、参加者は自由にキャラクターとインタラクションを行なうことが出来る。参加者のインタラクションの頻度によって本システムはストーリーの進行を楽しむストーリーベースのシステムにも、即興的なインタラクションを楽しむインタラクションシステムにもなる特徴を持っている。

#### (1) アバター

映画というフィクションの中で、自分があたかも、その中で存在しているようにイメージするためには、自然に自分が投影でき、コントロールできる分身が必要になる。我々のアバターのデザ

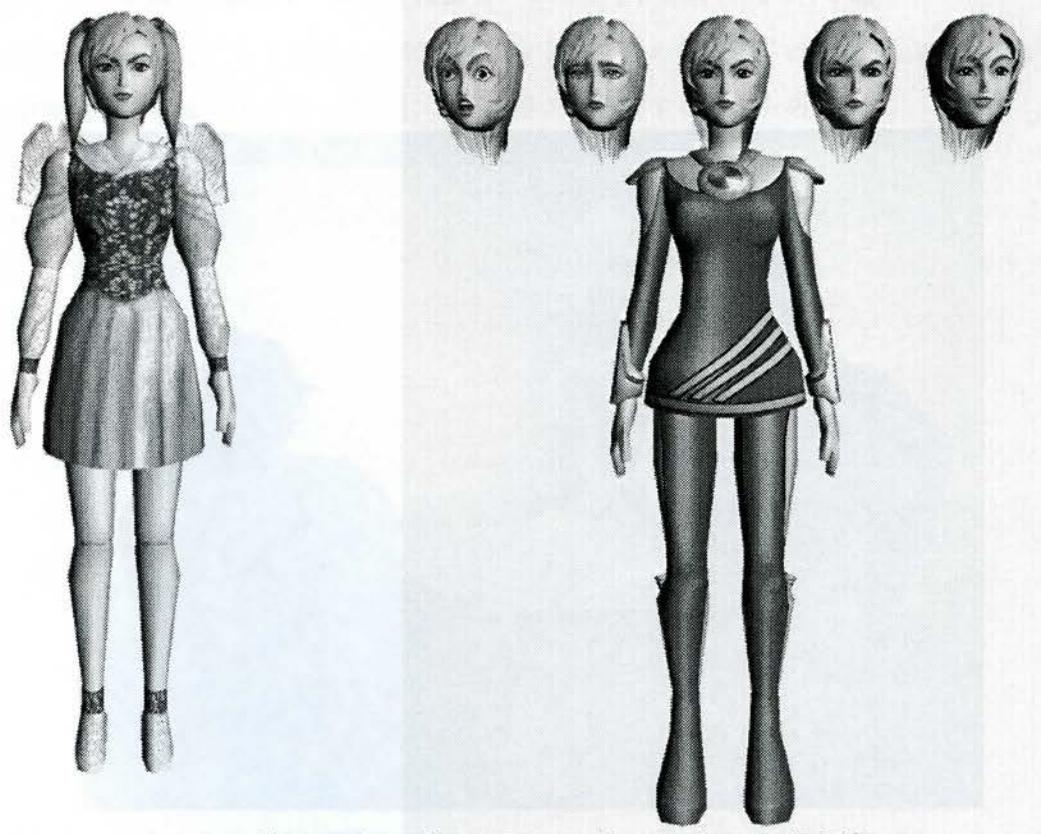
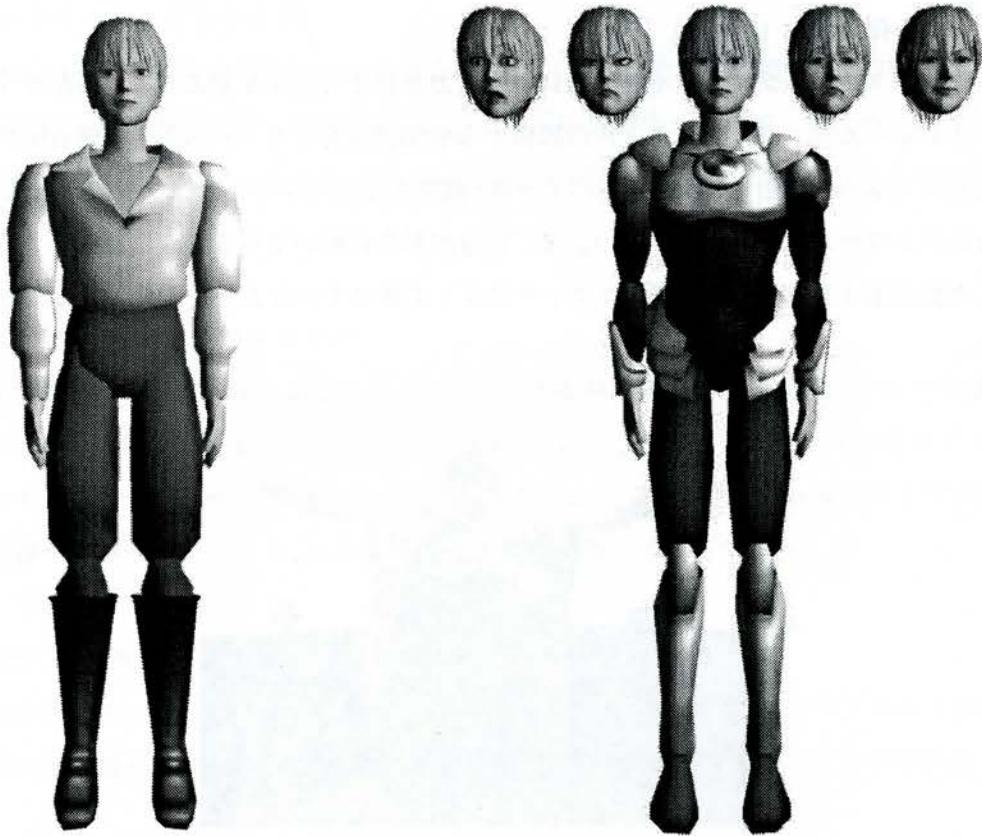


図8.11 「ロミオ」と「ジュリエット」の-avatarと顔表情  
 通常は、内側の衣装を着ているが、異次元の世界に行った場合は、SF系の衣装になる。

インは、人間を採用した（図8.11）。

サイバースペースでは、ジェスチャー認識により重力を超えて空中を飛んだり、体がガリバーのように大小になったり、自分の発する声の抑揚から感情認識をして、アバターの顔表情が感情に応じて変化するように設計している。アバターは、観客が磁気センサーウェアを装着して頭・体・手・足をコントロールする（図8.12）。又、自分が本当に動いているように感じるように、自分が動くとき動作音が仕組みられており、ロボットになって動いているような気分になるように設計している。



図8.12 「ロミオ」アバターを磁気センサーで制御する観客



図8.13 「ロミオ」と「ジュリエット」になった観客が感情コミュニケーションに没入している様子

## (2) インタラクティブシナリオ

インタラクティブポエムのように一つ会話ごとにインタラクティブに枝分かれする構造を持つ。主役級のアバターが気分にあった台詞を演じ、その反応にてキャラクター同士や、アバターを含めた感情によるコミュニケーションにより、物語りが演劇的に生成されるスタイルを採用した(図8.13)。

## (3) ジェスチャーインタラクション

ロミオ&ジュリエットアバターは、サイバースペースの中で任意の感情キャラクターや魂といったオブジェクトに触ることができ(図8.14)(図8.16)、状況によってさまざまなリアクションが返ってくる。また、サイバースペースの特徴を生かしてロミオ&ジュリエットアバターは、仮想の武器サイバーガン(図8.17)を持ち、相手と戦うこともできる。

## (4) 追従インタラクション

ロミオ&ジュリエットの感情を認識して、感情キャラクターがロミオ&ジュリエットに近づき、なぐさめたり、怒ったり、ほおをピシヤリと叩いたり、抱き締めたりすることができる。

## (5) 接触インタラクション

活動的なシーンでは、図8.18に見られるように騎士パリスの剣に刺された場合、ロミオ&ジュリエットが装着しているバイブレーターにより、痛さを感知することもある。

図8.14~18にシステムと参加者がいくつかのシーンでインタラクションを行なっている様子を示す。

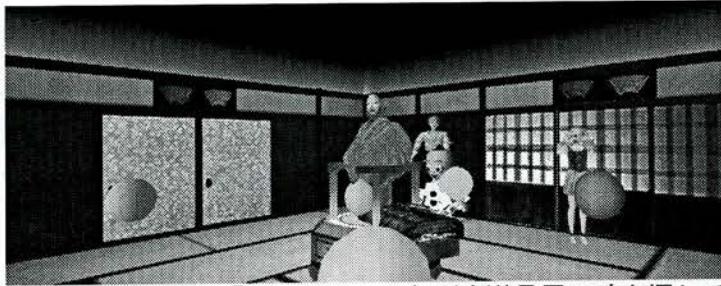


図 8.1 4 「ロミオ」と「ジュリエット」が古道具屋で魂を探している様子

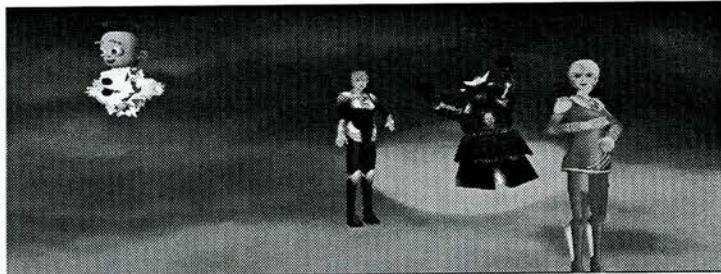


図 8.1 5 「ジュリエット」が現世の父と再会して話している様子



図 8.1 6 「ロミオ」と恋敵「パリス」との決闘の場面をジュリエットが止める様子



図 8.1 7 「パリス」の仇を打つためジェスチャー認識のサイバーガンで戦う「ロミオ」

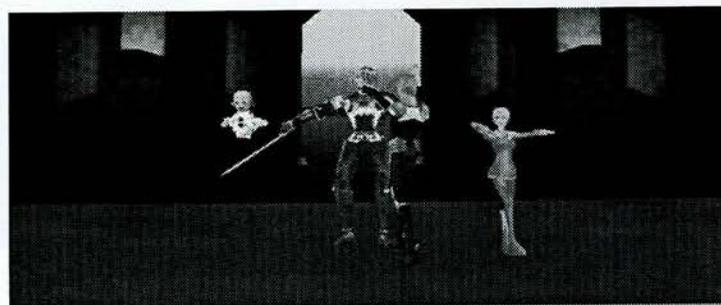


図 8.1 8 「パリス」の剣の近さを、接触インタラクションで感知する「ロミオ」

## 8.6 まとめ

映画とインタラクション機能を統合することによって実現されるインタラクティブ映画は新しいメディアになる可能性を持っている。本論文では、インタラクティブ映画のコンセプト、それに基づいて我々が開発した第1次システムの問題点について簡単に述べると共に、第1次システムの問題点を解決した第2次システムについて述べた。本システムでは二人の人がネットワークを介して、インタラクティブストーリーの進行に参加できる。また、いつでもストーリー中のキャラクターとインタラクション出来る機能を持たせている。このインタラクションの設計が要である。参加者は、このインタラクションを行なうことにより、物語りの世界に没入していくことができる。登場人物の感情キャラクターに全て感情モデルを持っている。そして、全てのインタラクションはロミオ&ジュリエットの感情認識を核にして、反応するようにできていることで観客の気持ちを引き付けることが成功できたといえる。ストーリーの設計の別の方法としては、人工生命の技法のひとつで、ハードウェア行動の進化モデルでオートマトンを自動生成し、進化させる仕組みに基づき、アニメ・キャラクターの行動の生成やストーリーの自動生成を行なう研究[101]もある。この方法では違う問題が発生してくる。自動生成された物語りはすべておもしろくなるのか？という問題である。いづれにしろ、進化のためはかなり洗練された論理的条件を人間があらかじめ決めておくことが必要になる。現時点では、いかなるダイナミクスを設計しようとも、人間が決めなければいけないことが多いのが現状である。しかし、そこに予測のつかない異分子である人間が複数入り仮想世界でコミュニケーションをする形は、まさに現実と非現実を行ったり来たりしているような実感がある。

これからの課題としては、人間が無意識で行なう行動の認識をどのように取り入れるかということである。声や、動作認識は、意識的に行なうことが多い。無意識情報の中により感情成分が表れていることは、心理学の研究からも明らかである[102][103]。又、インタラクティブシネマの将来の形としては、VRML等のサイバースペースの中だけで、米国人と、フランス人同士がインタラクティブシネマを興じるということも可能になるであろう。

## 9 結論

### 9.1 本論文のまとめ

人間同士のコミュニケーションにおいてはバーバルな情報と同様にノンバーバルな情報が極めて重要な役割を果たしている。人間とコンピュータのコミュニケーションにおいてもノンバーバルな情報を取り扱えるようになることは、人間とコンピュータの間の健全な関係のために重要なことである。ノンバーバルな情報の取り扱い、マウス・アイコンなどの低次の人間の感覚を利用したデバイスの実現から、さらには情緒・情感などの高次の感性的なものをいかに扱うかまで、多岐にわたっている。最近「感性情報処理」の名のもとに、これらを取り扱おうとする研究が盛んになってきたが[105]、それらは感覚を取り扱い始めたばかりで、高次の感性情報処理の扱いはまだまだである。本研究では感性情報の取り扱いに向けた第一歩として感情を取り上げ、感情を介した人間・コンピュータ間のインタフェースを作り上げることを目指した。その際、

- (1) 感情を介した双方向コミュニケーションの可能な総合システムを構築すること
- (2) システム構築のデザイン設計においてアートのアプローチを導入することを目指して研究を行った。

まず、第1章では、序章として研究の背景、特に従来のヒューマンインタフェース、さらにはノンバーバルインタフェースについて概要を述べた。

第2章では、本研究で取り上げる具体的な課題とアプローチについて述べた。特に感情インタフェースの重要性とアーティスティックなアプローチの重要性について述べた。

第3章では、インタラクティブティ=コミュニケーションと考え、インタラクティブティを持った感情表現のモデル化を試みた。又、人間が感情移入をする媒体としての「ひとがた」について重点をあてた。自己の分身であり自分に最も近い他者をキーワードに、話しかける人の声から感情抽出をし、インタラクティブに答えるひとがたキャラクター「ニューロベイビー」との新しいコミュニケーションの形態を述べた。

第4章では、感情を用いた異文化間のコミュニケーションについて述べた。まず、異文化間のコミュニケーションにおける感情の役割とは何かを考察した。次に、コミュニケーションの仲介役としての感情キャラクターの位置付けと、その実例として感情キャラクター「ネットワークニューロベイビー」を用いて異文化間感情翻訳実験を行なった内容について述べた。

第5章では、感情に反応する擬人化キャラクターの生成について述べた。感情キャラクターから

マルチモーダルな擬人化キャラクターへ進化させ、より高度なコミュニケーションを生成するため、感情モデルのデザインの複雑化を図った。その実例として、音声とジェスチャーから複合的に感情認識するマルチモーダルキャラクター「ミック」について述べた。

第6章では、バーバルインタフェースとノンバーバルインタフェースの融合を図った研究について述べた。コンピュータと感情のやりとりのみではなく、言葉の意味も同時に理解して、連歌形式で人間と即興詩を作るコンピュータ詩人「インタラクティブポエム」システムについて述べた。

第7章では、擬人化キャラクターを包む世界に、インタラクティブストーリーを導入し、その結果として感情移入型インタラクションの実現を目指したインタラクティブシアターについて述べた。特にインタラクティブな物語の進行をコントロールするスクリプト設計の基本的概念について述べた。

第8章では、第7章で述べた第1次システムの問題点を改良した第2次システムについて述べた。改良した点である複数人の参加が可能なシステムと、任意の時点でのインタラクション(anytime interaction)が可能なメカニズムの実装、より複雑なインタラクティブシナリオの実現を可能にしたスクリプト記述について重点的に述べた。

## 9.2 本論文で実現できたこと

### 9.2.1 感情を介したインタフェースの実現

本研究の成果のひとつは感情を介した人間・コンピュータ間のインタフェースを実現したことである。まず、第3章に述べたように人間の感情を認識し、それに反応するコンピュータキャラクター「ニューロベイビー」を実現した。また第5章で述べたように「ニューロベイビー」における感情インタフェースを高度化し、感情の数の増加、反応パターンの高度化・マルチモーダル化を図ったマルチモーダルキャラクター「ミック」を構築した。これらは、いずれも感情インタフェースの実現例としては最初のものである。また、「ニューロベイビー」・「ミック」は国内外の数々の技術展示会・アートの展覧会に展示・招待展示され、実際に体験した見学者から高い評価を得た。これらの結果から感情インタフェースが文化を問わず極めて有効なものであるという評価をすることができるであろう。

異文化間感情翻訳という大きな問題に挑んだ「ネットワーク・ニューロベイビー」は、これら二つに比較すると、感情認識技術が未熟であるなどのため、感情翻訳をリアルに実現するまでには至らなかったが、そのコンセプトは、多くの人々から評価された[61]。

## 9.2.2 感情移入の実現

次に本研究で実現できたことは、感情移入の実現である。従来の人間とコンピュータのインタフェースでは「効率」が評価基準であった。しかしながら人間同士のコミュニケーションでも明らかなように、「感情インタフェース」では効率は評価基準として適切ではない。むしろ、感情インタフェースにより、人間がインタラクションにいかにか夢中になり没入できるか、さらにはコンピュータキャラクターにいかにか感情移入できるか、が適切な評価基準といえるだろう。

筆者はまず感情をインタフェースにすることそのものが感情移入生成の鍵であると考え「ニューロベイベー」を構築した。予想通り、いやはるかに予想を超えて、人々はニューロベイベーとの対話に夢中になり感情移入をした。これは最初の狙いが正しかったことを示している。さらにその試みをベースに、ニューロベイベーを高度化してマルチモーダルキャラクター「ミック」を構築した。ミックでは以下のような方法を取り入れることにより、更に高度の感情移入実現を図った。

- (1) 取り扱える感情の数を4つから8つに増加した。
- (2) ニューロベイベーが顔だけのキャラクターだったのに対し、全身のモデルを作成し顔と動作のアニメーションによって感情に対する全身的な反応を表現できるようにした。
- (3) 入力を声だけでなく、人間の手のジェスチャーからも行なえるようにし、マルチモーダリティを持たせ複合感情を実現し、疑似的に触るなどのノンバーバルインタフェースを実現した。

展示会などでの体験者・見学者からは、「ニューロベイベーよりソフィスケートされたインタラクションを楽しめた」、「ニューロベイベーより生命感があり、人間的な親しみを覚えた」などの感想が得られ、当初の目標を達成することができたと考えられる。

次の段階としてバーバルな要素の付加を試みた。すなわち、ニューロベイベー・ミックが言葉の意味内容を無視したインタラクションであったのに対し、言葉の意味をも取り込んだインタラクションを実現しようとした。しかしながら、あくまで狙うのは感情移入であり、ホテルの予約や情報案内などは適切ではない。そこで言葉の中でも特に感情・情緒・感性と深く関わっている詩を対象として取り上げ、人間と詩を連歌的に創作しうるシステム「インタラクティブポエム」を構築した。単に詩を朗読しあうことで、どの程度の感情移入が実現できるかという危惧もあったが、結果ははるかに予想を超えて成功をおさめた。数多くのアート展示・技術展示に出展・招待出展し、多くの人々に体験してもらったが、いづれも「コンピュータ詩人ミューズとの詩の掛け合いに没入した」「黙読するよりもはるかに詩の世界に感情移入できた」との感想が圧倒的であった。特に、イン

タラクティブポエムのオリジナルの詩は日本人の詩人のものであるが、それを英訳した英語版インタラクティブポエムを米国・ヨーロッパで展示した結果、現地の人々にも深い感銘を与えた。これらの事実から、ノンバーバルインタフェース、バーバルインタフェースの融合が文化を越えて深い感情移入を実現できることを実証した。

「インタラクティブシネマ」では、更に深い感情移入生成を実現するためドラマの導入を試みた。人が小説を読み、映画を観る時、それらの対象に深く感情移入することは周知の事実である。しかし、我々はどんなに感情移入していようと、現実と非現実の壁がありその世界に直接踏み込めない。本研究ではインタラクションの導入により、この壁をやぶりその世界に入る没入感を実現した。映画の世界に入って、その場の雰囲気や堪能したり、感情キャラクターと話してみたり、話す内容によって、ドラマが変わってしまうようなことが可能なシステムを文脈性と創造性を生かした工学的技術[45]の考え方をベースとして実現した。映像の世界観に没入して主体的に体験するというのが、インタラクティブシネマ研究の動機である。インタラクティブシネマを数々の場に出展し、人々の評価を仰ぐことは今後の課題であるが、ATRへの見学者に体験してもらった感想によれば「自分の話した台詞に対し、感情キャラクターが打てば響くような反応をした時に感動する」「物語りのサイバースペースが実現できている」「どきどき、わくわくする」「自分の台詞を喋るのが待ち遠しい」など、いままでになかったサイバースペースでの深い感情移入が実現されたと確信している。

### 9.2.3 能動的なストーリー体験の実現

サイバースペースの中でのストーリー体験という考え方は、テレビゲームやテーマパーク等におけるバーチャルシアターなどを含めて考えると、これまでも存在した考え方であるが、その在り方について体系的な考察が加えられた例は少なく、従来の小説などの物語展開 (storytelling) を越える新しい物語が生まれる可能性があるという抽象的なレベルのものにとどまっておき、具体的なシステムの提示には至っていない。

我々は、誕生以来100年を経て種々の物語展開技法が生み出され、メディアとしての位置を確立している映画[92]をその出発点として、サイバースペースの中でのストーリー体験が可能なシステムを具体的に構築することをめざした。また、この新しいメディアであるインタラクティブ映画は、その運用まで含めて考えると、インタラクティブ映画システムとそれを用いて主人公となってストーリーを体験する参加者から構成される。インタラクティブ映画システムは、下記の要素から構成されるとして研究を行った。

(1) 参加者とのインタラクションに基づいて途中経過・結末などが変化するインタラクティブなストーリーを生成する機能。具体的には、ストーリーはシーンの連続として構成される。

(2) スクリーンまたはディスプレイに投影・表示される映像と、映像に関連する音声・サウンドで構成されるシーンを生成する機能。

(3) ストーリー中の主人公以外の役割を演じる単数もしくは複数のキャラクターおよびその動作を生成・制御する機能。

(4) 参加者とキャラクターもしくはシーン中のオブジェクトとのインタラクションを可能にするインタラクション機能。

文献[104]では、サイバースペースの中でのストーリー体験システムの特徴を *immersion* (没入感) と *interactivity* であると述べているが、インタラクティブ映画ではより具体的に以下のような点を実現することを目指した。

#### (1) サイバースペースの構築と参加者のサイバースペースへの没入

CGの活用、CGと実写映像の統合、さらには立体映像の創出により、臨場感に富んだサイバースペースを構築し、参加者にあたかもその中にいるかのような感覚を与える。

#### (2) インタラクションを介したサイバースペースの中でのストーリー体験

サイバースペースの住人、すなわち参加者を取り巻く種々のキャラクターと音声や、身振り・手振りでインタラクションを行なうことによりストーリー展開を体験することができる。

#### (3) 感情キャラクターの個性

ストーリーがおもしろくて、それを演じるアクターの表情・行動・言動が実現されていなければならない。インタラクションに使用する技術がうまく作動してないと、それは、途中でさえぎられ、好ましくない感情が生成される。インタラクションテクノロジーと、そのパターンの組み合わせがドラマの善し悪しを大きく変える。この点に最大の注意を払ってシステムを構築した。

## 9.3 今後の研究課題

### 9.3.1 感情インタフェースの課題

従来の情報科学が扱ってきた情報は、論理的な情報であり、論理的対象であるので、「客観性、一意性、再現性」が保証されていた。しかし、感情情報は、「主観性、多義性、あいまい性、状況依存性」といった属性を持つ情報である[105]。主観的であるということは、個に対する依存性が高いということであり、好き嫌いの問題がでてくる。また、ネットワークニューロベビーで試みた異文化間のコミュニケーション実験でわかったように、感情情報は受け手によって異なった意味（多義性）を持つ。さらに、環境などに大きく左右される状況依存性を持つ。感情は無意識の情報も含まれているので、なぜその感情が生成されるのかという感情発生の要因を研究するべきである。以上のような理由から感情情報は記号ではなく、個性的、異文化間の幅を解釈する尺度を選べるようにするべきである。具体的には、感情生成の要因に対応した自然な入力、感情情報のデータベース化、状況による任意の感情の生成などを研究する必要がある。

### 9.3.2 次世代映画メディア実現のための課題

現在は、ディズニー・ユニバーサルスタジオといった映画と結びついたエンタテインメント産業やナムコ・セガ・任天堂といった大手ゲーム産業は、すでにインタラクティブ性をキーワードに新たな感性を刺激するテクノロジー産業を展開しようとしている。ロールプレイングゲームや、CD-ROMベースのインタラクティブシネマに見られるように物語り中心のインタラクティブ記述のメディアが広く受け入れられつつある。しかし、現在産業化されているゲームテクノロジーは、感性を本当にコンピュータが理解するテクノロジーには至っていないのが実情である。人間が考えたシナリオに沿ってコンピュータがリアクションすることにより、実時間でインタラクティブ性が実現されている。これからは、コンピュータとのインタラクション部分に感性や自律性を導入することが必要である。それを実現するテクノロジーは、人工知能・各種認識技術・感性工学といった情緒をダイナミクス化したコンピュータテクノロジーであろう。

これからの課題は、各々シーンの臨場感を高めるように、インタラクティブな状況記述をしていくことである。これは、決して現実と非現実の境がわからなくなるような、リアルな現実ではない。筆者が描きたいのは、人間と人間の間に横たわっている未知の心地よいイマジネーションの生成である。その体験は、決してハッピーエンドとは限らない。自律キャラクターが人間を裏切ったり、人間がおもわず自分の悩みをつぶやいたり、恐ろしかったり、哀しかったりするかもしれない。しかしその体験をした後は、なにか充実感が感じられるデジタルカタルシスである。

それを高めるために、先端技術を用いてシーンの状況表現を行なっていくことが必要である。それは最終的に、人間の未知の部分の研究につながっていく。機械は、生命体にいちばん似たときに、いちばん美しくなるという。CGキャラクターも、人間くささのような物が感じられた時感動する。根本的に表現と、実現する技術をテクノロジーというわざで実現できた時、本当に感動できるであろう。直観するとき起こる主要な感情の一つが、美的満足感である[106]。美こそ真理であり、真理こそ美である。このように美と調和の感覚が、芸術・科学に活気をもたらしてくれる。科学も感受性が必要なのである。

### 9.3.3 感動工学の実現に向けて

今後、感性・情緒・深層心理といったカオス的な分野をコンピュータが理解する研究は進み、ますます増大していくであろう。しかし、人間の心身が、画一形式化不可能の側面を持つことも否定できないのも事実である。それには、今までのように人間の知覚の探究や感性の拡張の研究を技術的アプローチだけで進めていくのは充分とはいえない。芸術は、人間の知覚の探究や感性の拡張を、精神的・感性的に表現してきた歴史がある。また、すぐれた芸術は、その時代の最先端の道具を使って創作されている。コンピュータ全盛時代の未来は、感性を論理の世界に変換した後、あやふやな感性を統一する精神的なもの、哲学をコンピュータにどう構築させるかが問題になってくる。具体的には工学知の技術[45]をはじめ、人工知能技術、創発性を持つ人工生命技術[87]、オーサリングシステム技術を、どのようにクロスオーバーさせてシステム設計するかといったことが未来の研究課題になってくる。21世紀を担う若い世代は、特にコンピュータネットワークを介したグローバルな感動を欲している。だから21世紀には、感動生成のテクノロジーを我々は手に入れることが期待される。我々が、すばらしい小説や映画を見る時、その作品に感動していると、心は感情移入してその世界に入っている。しかし、まだ工学的に感情移入を作るための技術は完成されていない。それは、創作者の個人的な私見や精神的・哲学的な問題を含むからだ。今後は、これらを工学的に設計する研究をすべきである。感性をコンピュータで分析する研究以上にこれらの感性をいかにまとめあげるかということが、個性の創出につながる。現段階では、これは人間にしかできない聖域である。芸術とテクノロジーはマルチメディアの時代に来て、芸術の方法論をコンピュータを通して工学化することが可能になった。21世紀には、我々の脳内イメージをコンピュータで映像化することも実現できるであろう[107]。そうすると、今まで、表現方法を知らなかった人々もコンピュータを用いてイメージを表現できるようになる。この技術は、人にイメージを伝達するための豊かなメディアとして人類に貢献することができる。それは新たな映画の未来[108]、であり社会システムの構造を変えるほどのインパクトを持ち、現代社会における秩序や情報の連鎖、関係性を変える物になるだろう。

## 謝辞

本論文をまとめるにあたり、懇切丁寧な御指導、御鞭撻を賜りました  
東京大学大学院工学系研究科電子情報工学専攻 原島 博 教授 に心より感謝致します。先生には、研究の概念の重要性、研究に対する着目点、そして研究の進め方など、数多くのことを教えて頂きました。

本論文の査読をこころよくお引き受けくださいました、  
東京大学生産技術研究所 坂内 正夫 教授  
東京大学人工物工学研究センター 河口 洋一郎 教授  
東京大学生産技術研究所 橋本 秀紀 助教授  
東京大学工学部附属総合試験所 廣瀬 通孝 助教授  
東京大学大学院工学系研究科電子情報工学専攻 相澤 清晴 助教授  
に、深く感謝いたします。

本論文は、筆者が武蔵野美術大学で教鞭をとっていた時期に行なった富士通研究所ヒューマンインターフェース研究部との共同研究、そして東京大学生産技術研究所との共同研究、その後、筆者が株式会社 国際電気通信基礎技術研究所 (ATR) 知能映像通信研究所に移り、そこでの研究をまとめたものであります。

本研究の機会を賜るとともに、本論文の執筆の機会を与えて頂くとともに常日頃らご指導、ご助言を頂いている国際電気通信基礎技術研究所 葉原 耕平前副会長 (現、ATR 顧問)、ATR 知能映像通信研究所 酒井 保良会長、中津 良平社長、に心より感謝致します。

また、ATR において研究する機会を与えて頂いたATR 人間情報通信研究所 東倉 洋一前社長 (現、NTTコミュニケーション科学研究所所長)、ATR 人間情報通信研究所第6研究室 下原 勝憲前室長 (現、NTTコミュニケーション科学研究所グループリーダー)、ATR 人間情報通信研究所元客員研究員 トーマス レイ博士 (現、オクラハマ大学教授) に厚くお礼申し上げます。

さらに、本研究に関して共同研究をする機会を与えて頂いた富士通研究所ヒューマンインターフェース研究部元部長 森田 修三博士 (現、富士通株式会社 パーソナルシステム研究所所長代理)、富士通研究所ヒューマンインターフェース研究部元主任研究員 村上 公一氏 (現、富士通株式会社 電脳メディア開発部長)、東京大学生産技術研究所 橋本 秀紀 助教授、瀬崎 薫 助教授 に厚くお礼申し上げます。

メディアアート研究に関して指導、ご助言を頂きました武蔵野美術大学造形学部 映像学科 吉田 直哉 元主任教授、齊籐 嘉博 主任教授、山本 圭吾 教授、そして、コンピュータと芸術教育に関して深い示唆をいただいた武蔵野美術大学造形学部 共通デザイン研究室 小井戸 満 教授に厚くお礼申し上げます。

本研究において開発した各種システムの実装にあたって多くのご尽力をいただいた、富士通研究所ヒューマンインターフェース研究部元研究員 柿本 正憲氏 (現、シリコングラフィックス株式会社)、東京大学大学院工学系研究科電子情報工学専攻博士課程学生 国井 康晴博士 (現、中央大学理工学部講師)、プログラムの作成にあたって多大なご協力を頂いた株式会社ビクタウンズ 佐籐 薫 社長、株式会社CSK 越知 武氏、NTTソフトウェア株式会社 鈴木 秀昭氏、ATR 知能映像通信研究所客員研究技術員 松瀬 尚氏に、ここに謹んで感謝の意を表します。

最後に、本研究の遂行、および論文の執筆にご協力頂いたATR 知能映像通信研究所の皆様、また実験にご協力頂いた皆様に感謝致します。

## 参考文献

- [1]原岡一馬編、「人間とコミュニケーション」ナカニシヤ出版 pp.65-67 (1990).
- [2]白井良明編、「パターン理解」、オーム社 (1987).
- [3] K.Nagao and A.Takeuchi, "Speech dialogue with facial displays: Multimodal human-computer conversation," Proc.32nd Annual Meeting of Assoc. for Comput.Linguistics, pp.102-109(1994).
- [4] 末永、間瀬、福本、渡部、「Human Reader:人物像と音声による知的インターフェース」、信学論, Vol.J75-D-II, No.2, pp.190-202(1992).
- [5]黒川隆夫著、「ノンバーバル・インタフェース」オーム社 (1994).
- [6] Birdwhistell,R.L., "Kinesics and Context," Univ. of Pennsylvania Press (1970).
- [7] Mehrabian,A., "Communication without words," Psychol. Today, Vol.2, No.4, 52-55 (1968).
- [8] 吉川弘之著、「概念の設計から社会システムへ」三田出版会 (1990).
- [9] 井上、ほか、「視覚刺激による人間感覚の測定と応用」、自動車技術、Vol.42, No.11, pp.1409 (1988)
- [10] 山口、加藤、赤松、「顔と感性情報と物理的特徴との関連について-年齢/性の情報を中心に-」、信学技報、PRU90-7, pp.17-23 (1990).
- [11] 小林、丹下、原、「人の6基本表情の実時間認識」、日本機械学会ロボティクス・メカトロニクス講演予稿集、pp.873-876 (1995).
- [12] 崔、原島、式部、「顔の3次元モデルに基づく表情の記述と合成」、信学会論文誌、Vol.J73-A, No.7, pp.1270-1280 (1990).
- [13]B. Laurel (遠山峻征訳)「劇場としてのコンピュータ」、凸版出版(1992).
- [14]R. W. Picard, "Affective Computing," The MIT Press, Cambridge, Massachusetts, (1997).
- [15]R. W. Picard and J. Healey, "Affective wearables," In Proceeding of the First International Symposium on Wearable Computers, Cambridge, MA, (1997).
- [16]R. W. Picard; chapter in Hal's Legacy, Ed. by D.Stork, "Does Hal Cry Digital Tears? Computers and Emotions," MIT Press, Cambridge (1996).
- [17]Raul Fernandez, "Stochastic Modeling of Physiological Signals with Hidden Markov Models: A Step Toward Frustration Detection in Human-Computer Interfaces," MIT Electrical Engineering and Computer Science Masters Thesis (1997).
- [18] 原島、永田、金子、「顔の印象学に向けて-コンピュータによる解析-」、文部省重点領域「感性情報の情報学・心理学的研究」第2回全体会議予稿集、pp.96-99 (1995).
- [19]S.Morishima and H.Harashima, "Emotion Space for Analysis of Facial Expression," IEEE International Workshop on Robot and Human Communication, pp.188-193(1993).
- [20]Hiroshi Kobayashi and Fumio Hara, "Analysis of the Neural Network Recognition Characteristics of 6

- Basic Facial Expression," 3rd IEEE International Workshop on Robot and Human Communication, pp.222-227 (1994).
- [21] 小林、丹下、原、「人の6基本表情の実時間認識」、日本機械学会ロボティクス・メカトロニクス講演予稿集、pp.873-876 (1995).
- [22] 間瀬、前田、末永、「表情動画像からの感情の認識の1手法」、信学技報、Vol.PRU89-128, pp.17-24 (1990).
- [23] 山口、加藤、赤松、「顔と感性情報と物理的特徴との関連について-年齢/性の情報を中心に-」、信学技報、PRU90-7, pp.17-23 (1990).
- [24] P.Ekman, "Facial Expression and Emotion," *American Psychologist*, Vol 48, No. 4, pp. 384-392 (1993)
- [25] P.Ekman and W.Friesen, "Facial Action Coding System," Consulting Psychologists Press (1977).
- [26] 吉川佐紀子、益谷 真、中村 真編、「顔と心-顔の心理学入門-」サイエンス社 (1993).
- [27] R. Brooks, C.Breazeal, M. Marjanovic, B. Scassellati, M. Williamson, "The Cog Project: Building a Humanoid Robot," To appear in a Springer-Verlag Lecture Notes in Computer Science Volume.
- [28] R.Brooks, "Prospects for Human Level Intelligence for Humanoid Robots," Proceedings of the First International Symposium on Humanoid Robots (HURO-96), Tokyo, Japan (1996).
- [29] J.Bates, A.Loyall, W.Reilly, "Integrating Reactivity, Goals and Emotions in a Broad Agent, Proceedings of the 14th Annual Conference of the Cognitive Science Society, Indiana (1992).
- [30] K. Perlin, A.Goldberg, "Improv: A System for Scripting Interactive Actors in Virtual Worlds," *Computer Graphics*; Vol. 29, No. 3. (1996).
- [31] K. Perlin, "Real Time Responsive Animation with Personality," *IEEE Transactions on Visualization and Computer Graphics*; Vol 1, No. 1.(1995).
- [32] Hayes-Roth, B. and van Gent. R., "Story-making with improvisational puppets and actors," Stanford Knowledge Systems Laboratory Report KSL-96-05, (1996).
- [33] Hayes-Roth, B., Ball, G., Lisetti, C. Picard, R. W., Stern, A., "Panel on Affect and Emotion in the User Interface," In Proceedings of the Conference on Intelligent User Interfaces (1998).
- [34] 村上他、「似顔絵師システムPICASOにおける感情表現と空間次元の扱い」、信学研報、PRU90-7, pp.17-23 (1990).
- [35] 輿水、村上他、「似顔絵生成システム開発による認知的ビジョンの研究」、重点領域研究平成5年度成果報告書「感性情報処理の情報学・心理学的研究」、pp.103-106 (1994).
- [36] Andre Bazin, "What is Cinema?" Vol.2 Selected and translated by Hugh Gray, Berkeley: University of California Press (1967, 1971).

- [37]竹林洋一、「音声自由対話システムTOSBURG I -ユーザ中心のマルチモーダルインターフェースの実現に向けて -」、信学論、Vol.J77-D-II, No.8, pp.1417-1428 (1994).
- [38]K.Nagao and A.Takeuchi, "Speech dialogue with facial displays: Multimodal human-computer conversation," Proc.32nd Annual Meeting of Assoc. for Comput.Linguistics, pp.102-109 (1994).
- [39]ニコラス・ハンフリー著、垂水雄二訳、「内なる目」-意識の進化論- 紀ノ国屋書店 (1993).
- [40] W. Scott Neal Reilly, Ph.D. Thesis. Technical Report CMU-CS-96-138, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA. (1996).
- [41] P Maes, T. Darrell, B. Blumberg, and A. Pendant, "The A the Computer Animation'95 Conference (1995)
- [42] K. Perlin, "Real time responsive animation with personality," IEEE Transactions on Visualization and Computer Graphics, Vol. 1 ,No. 1 pp.5-15 (1995).
- [43] J. Bates, B. Loyal, S. Reilly, "An architecture for action, emotion, and social behavior," Proceedings of the Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World.
- [44]「ザ・スタンリーキューブリック」キネマ旬報社編 (1981).
- [45] 吉川弘之監修、「技術知の位相」東京大学出版会 (1997).
- [46] スティーブン・D・キャッツ著、津谷佑司訳「SHOT BY SHOT: 映画監督術」フィルムアート社 (1996).
- [47] Hannes Leopoldseder, Christine Schopf, "Cyber arts," Prix Ars Electronica Edition 97, Springer (1997).
- [48] 季刊インターコミュニケーションNo.7 冬号、「特集: インタラクティブアート」NTT出版(1993)
- [49] S. サンガー著 (竹内 均訳)「乳児はなんでも知っている」祥伝社 (1987).
- [50]土佐尚子、「トーキングニューロベビー」、季刊「仏教」特集: 生命、法蔵館 pp.125-133 (1993)
- [51] N. Tosa, et al., "Negro-Character" AAAI '94 Workshop, AI and A-Life and Entertainment (1994).
- [52] 柿本正憲、土佐尚子、森淳一、「ニューロベビー: 仮想世界に棲む感情体」画像ラボ, vol.3, No.10 日本工業出版 (1992).
- [53] N. Tosa, "Talking to NEURO BABY" ARS Electoronica'93 "Genetic Art-Artificial Life" proceeding pp.353-356 (1993).
- [54] 北原、東倉、「音声の韻律情報と感情表現」電子情報通信学会音声研究会研究報告、SP88 - 158 (1989).
- [55] Rumelhart, Hinton, Williams, "Learning Interanal Representations by Error Propagation," Parallel Distributed Processing, vol.I, pp.318-364, MIT Press (1986).
- [56] デスモンド・モリス著 (藤田 統訳)「マンウォッチング」小学館 pp.26 (1980).
- [57] 森島、原島、「テキストおよび音声に基づく顔表情自動合成のためのリアルタイムアニメーションシステムの構成」、第6回ニコグラフ論文コンテスト論文集, pp.144-152 (1990).
- [58] 大場、「多重内挿法MIMe(Multiple Inbetweening Method)によるアニメーション表現」、第6回ニコグラフ論文コンテスト論文集, pp.153-160 (1990).

- [59] Philips, Sound Ideas, "Sound Effects Library LHH7910Series," CD1001-CD1028.
- [60] N. Tosa, et al., "Network Negro-Baby with robotics hand," *Symbiosis of Human and Artifact*, Elsevier Science B.V. (1995).
- [61] 白石明彦著、「人工生命とは何か」丸善 pp.151-157 (1995).
- [62] 柿本、土佐、森、真田、「音声に反応する表情合成システム：ニューロベビー」情報処理学会第44回全国大会, 4N-9 (1992).
- [63] 橋本、国井、山田、土佐、「仮想環境下における仮想生物に関する研究」文部省重点領域「人工現実感に関する基礎的研究」平成7年度研究成果報告書 pp.43-44(1996).
- [64] 橋本、土佐、国井、「握手マシンの開発および力学情報を用いたコミュニケーション」文部省重点領域「人工現実感に関する基礎的研究」平成8年度研究成果報告書 pp.47-48(1997).
- [65] Shirley Fritz, "Nonverbal Communication," New York:Oxford Univ. Press(1974).
- [66] R.A.Bolt, "Put-That-There: voice and gesture at the graphics interface," *Computer Graphics*, Vol. 14, No.3, pp.262-270 (1980).
- [67] O.Stock, "Natural Language and exploration of an information apace: The AlFresco interactive system," *Proc.12th Int. Joint Conf. on Artificial Intelligence*, pp.972-978 (1991).
- [68] 島津、高島、「マルチモーダル Definite Clause Grammar (MM-DCG)」, *信学論*, Vol.J77-D-II, No.8, pp.1438-1446 (1994).
- [69] M.T.Vo and A. Waibe, "Multimodal human-computer interaction," *Proc. Int. Symp. on Spoken Dialogue*, pp.95-101 (1993).
- [70] Naoko Tosa and Ryohei Nakatsu, "Life-like Communication Agent - Emotion Sensing Character 'MIC' and Feeling Session Character 'MUSE' -," *Proceedings of the International Conference on Multimedia Computing and Systems*, pp.12-19 (1996).
- [71] 土佐尚子、「芸術とテクノロジー」- ファジイ技術を用いたインタラクティブアート - 日本ファジイ学会誌 Vol.9, No.5, pp.648- 665 (1997).
- [72] S. Mozziconacci, "Pitch variations and emotions in speech," *ICPhS 95* Vol. 1, p. 178 (1995).
- [73] K. R. Schemer, "How emotion is expressed in speech and singing," *ICPhS 95*, Vol. 3, p. 90 (1995).
- [74] G. Klasmeyer and W E Sendlmeier, "Objective voice parameters to characterize the emotional content in speech," *ICPhS 95*, Vol. I , p. 182 (1995) .
- [75] S. McGilloway, R. Bowie, and E. D. Bowie, "Prosodic signs of emotion in speech: preliminary results from a new technique for automatic statistical analysis," *ICPhS*, Vol. 1 , p. 250 (1995).
- [76] J. D. Marvel and A. H. Gray, "Linear prediction of speech," Springer-Verbal (1976).
- [77] Christopher Wren, Ali Azarbayejani, Trevor Darrell, and Alex Pentland, "Pfinder: Real-time tracking of the human body," *Integration Issues in Large Commercial Media Delivery Systems (Conference 2615)*. SPIE, (1995).

- [78]土佐尚子、中津良平、「インタクティブポエム」映像情報メディア学会誌, Vol.52 No. 1(1998)
- [79] N.Tosa, R.Nakatsu, "Interactive Poem System, " The 6th ACM International Multimedia Conference Proc. pp.115-118 (1998).
- [80]谷川俊太郎著、「みみをすます」 福音館書店 (1982).
- [81]Pattie Maes et al., "The ALIVE system: Full-body Interaction with Autonomous Agents," Proceedings of the Computer Animation '95 Conference (1995).
- [82] N.Tosa, "Creating a Movie with Autonomous Actor for That Respond to Emotions" SIGGRAPH-97Course Note title is "Interactive Movie, -Techniques, Technology, and Contents" (1997).
- [83]Christa Sommerer and Laurent Mignonneau, "Intro Act & MIC Exploration Space," Visual Proceedings of SIGGRAPH96, pp.17 (1996).
- [84]中津良平、「アーティスティックな要素を取り入れたバーチャルリアリティ技術—アートと工学の融合をめざして—」、日本バーチャルリアリティ学会論文集、Vol.1, No.1, pp.1-9 (1996).
- [85]沖啓介、「知覚と映像のエージェント」、映像学、No.57, pp.19-33 (1996).
- [86]Joseph Bates et al., "An Architecture for Action, Emotion, and Social Behavior," Proceedings of the Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World (1992).
- [87]Ken Perlin, "Real Time Responsive Animation with Personality," IEEE Trans. on Visualization and Computer Graphics, Vol.1, No.1, pp.5-15 (1995).
- [88] 中津良平、土佐尚子、「インタラクティブ映画構築に向けて」電子情報通信学会、画像工学研究会(1997).
- [89]P. Doyle, B. Hayes-Roth, "Computer-aided exploration of virtual environments,"Proceedings of AAAI Workshop on AI and Entertainment, AAAI Press, Menlo Park, CA, (1996).
- [90]F. Sparacino, et al., "HyperPlex:a World of ID Interactive Digital Movies," Proceedings of AAAI Workshop on AI and Entertainment, pp.77-81 (1995).
- [91] Tom Calvert, Armin Bruderlin, John Dill, Thecla Schiphorst, Chris Welman, "Desktop Animation of Multiple Human Figures," IEEE Computer Graphics & Applications, Vol. 13, No. 3, pp 18-26 (1993).
- [92] C.W. ツェーラム著、(月尾嘉男訳)、「映画の考古学」、フィルムアート社 (1977).
- [93] Oren Jacob, "Computer Graphics Story - A Personal Overview of Computer Animation in the Movies," ACM Computer Graphics, Vol. 31, No. 1, pp. 26-28 (1997).
- [94] Graham Walters et al., "The Making of Toy Story," Course Notes of SIGGRAPH 96 (1996).
- [95] Janet H. Murray, "Hamlet on the Holodeck: The Future of Narrative in Cyberspace," Simon & Schuster, New York, (1997) .
- [96] 中津良平、土佐尚子、「Inter Communication Theater —インタラクティブ映画の構築の試み—」、インタラクシオン97論文集、pp.67-68 (1997).

- [97] 中津良平、土佐尚子、「インタラクティブ映画構築に向けて－ Inter Communication Theater の基本概念と構成例－」、信学技報 IE96-113 (1997).
- [98] Pattie Maes, "How to do the right thing," Connection Science, Vol.1, No.3, pp.291-323 (1989).
- [99] Tetsu Shimizu, et al., "Spontaneous Dialogue Speech Recognition Using Cross-Word Context Constrained Word Graph," Proceedings of ICASSP'96, Vol. 1, pp. 145-148 (1996).
- [100] シェイクスピア著、中野好夫訳「ロミオ&ジュリエット」新潮社(1951).
- [101] 下原勝憲著、「人工生命と進化するコンピュータ」工業調査会 pp.118-120 (1998).
- [102] 下條信輔著、「サブリミナル・マインド」-潜在的人間観のゆくえ - pp.37-63 (1996).
- [103] F・D・ピート著、菅啓次郎訳、「シンクロニティ」、朝日出版社(1989).
- [104] Janet H. Murray, "Hamlet on the Holodeck: The Future of Narrative in Cyberspace," Simon & Schuster, New York, (1997).
- [105] 井口征士他著、「感性情報処理」オーム社(1994).
- [106] フィリップ・ゴールドバーグ品川嘉也監修、神保圭志訳、「直観術」、工作舎(1990).
- [107] 吉田直哉著、「脳内イメージと映像」、文芸春秋(1998).
- [108] Morton Leonard Heilig, " The Cinema of the Future," PRESENCE Vol.1 , Number 3, pp279-294 (1992)

## 本論文に関する研究業績

### 著書

- ・ Ryohei Nakatsu, Naoko Tosa, "Interactive Movies," Handbook of MULTIMEDIA COMPUTING, pp.701-712, CRC Press (1999).

### 学術論文（査読有）

- ・ 土佐尚子、「Artificial Life Character の時代」特集：現実感の変容、映像学会誌 pp77-86, No.52 (1994)
- ・ 土佐尚子、中津良平、「感情に反応する自立型バーチャルアクターと仮想世界の生成」日本 VR 学会論文集 Vol. 2, No.1, pp.11- 18 (1997).
- ・ 中津良平、土佐尚子、越知武、鈴木秀昭、「インタラクティブ映画システムのコンセプトと構成例」電子情報通信学会論文誌 D-II Vol.J81-D-II, No.5, pp.944-953 (1998)  
(第13回電気通信普及財団賞（テレコムシステム技術賞）受賞)

### 国際会議（査読有）

- ・ N. Tosa, et al., "Neuro-Character" AAAI '94 Workshop, AI/A-Life and Entertainment (1994).
- ・ N. Tosa. et al., "Network Neuro-Baby with robotics hand," originally appeared in Proceedings of Human-Computer Interaction'95 Symbiosis of Human Artifact(Vol.1pp.77-82) Published by Elsevier Science B.V. (1995).
- ・ N. Tosa, et al., "Network Neuro-Baby with robotics hand", IJCAI '95 Workshop, Entertainment and AI/Alife (1995).
- ・ N.Tosa, R. Nakatsu, "Life-Like Communication Agent- Emotion Sensing Character 'MIC' & Feeling Session Character, 'MUSE' ," Proceedings of The Third IEEE International Conference on Multimedia Computing and Systems pp.12-19 (1996).  
ベストペーパーアワード受賞

- N.Tosa, "The Esthetics of Recreating Ourselves" SIGGRAPH - 96 Course Note No.25, "Life-like Belivable Communication Agents"(1996).
- N.Tosa, R. Nakatsu, "The Esthetics of Artificial Life :Human-like Communication Character, 'MIC'&Feeling Improvisation Character, 'MUSE'," Artificial Life V Proceedings, pp143 - 151(1997).
- N.Tosa, "Creating a Movie with Autonomous Actor for That Respond to Emotions" SIGGRAPH-97Course Note "Interactive Movie, -Techniques, Technology, and Contents" (1997).
- N. Tosa, R. Nakatsu, "Interactive Poem" Proceeding of the AIMI International Workshop on Kansei-The Technology of Emotion pp.54-59 (1997).
- R. Nakatsu and N. Tosa, "Toward the Realization of Interactive Movies - Inter Communication Theater: Concept and System -," IEEE International Conference on Multimedia Computing and Systems Proceedings pp.71- 77 (1997).
- R. Nakatsu and N. Tosa, "Toward the Realization of Interactive Movies- Inter Communication Theater: Concept and System -," IEEE Signal Processing Society Workshop on Multimedia Signal Proceedings, pp.519-524 (1997).
- N. Tosa, "Artistic Communication for A-life and Robotics," International Symposium on Artificial Life and Robotics Proceedings, Vol:2/2 pp.605-609 (1998).
- N.Tosa, R.Nakatsu, "Interactive Poem System, " The 6th ACM International Multimedia Conference Proceedings, pp.115-118(1998).

## アート&テクノロジー国際会議（査読有）

- ・ N. Tosa, "Talking to NEURO BABY" Proceedings of ARS Electoronica'93 "Genetic Art-Artificial Life," pp.353-356 (1993).
- ・ N. Tosa, "Neuro Baby" SIGGRAPH'93 "Machin Culture" Visual Proceedings, p.167 (1993).
- ・ N. Tosa, "Network Neuro-Baby with Handshake" SIGGRAPH'95 Visual Proceedings, p.127(1995).
- ・ N. Tosa, R. Nakatsu, "For Interactive Future Movie :Body Communication Actor, 'MIC' & Feeling Improvement Actor, 'MUSE'," Proceedings of The 7th International Symposium on Electronic Art'96, pp.126-129 (1996).
- ・ N. Tosa, "Interactive Poem" SIGGRAPH'98 Visual Proceedings, p.300 (1998).
- ・ N. Tosa, R.Nakatsu, "Recreating Ourselves" Proceedings of the 2nd International CAiiA Research Conference "Consciousness reframed II," p.51 (1998).
- ・ N. Tosa, R.Nakatsu, "Interactive Poem," Proceedings of The 6th ACM International Multimedia Conference Art Demo, p.38 (1998).

## 学術解説

- ・ 土佐尚子、「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」計測自動制御学会誌 Vol.34, No.4, pp311-314 (1995).
- ・ 土佐尚子、「芸術とテクノロジー」-ファジイ技術を用いたインタラクティブアート-日本ファジイ学会誌 Vol.9, No.5, pp.648- 665 (1997).
- ・ 土佐尚子、中津良平、「感情に反応するインタラクティブアクターと物語りの生成」情報学シンポジウム、日本学術振興会(1997).
- ・ 土佐尚子、中津良平、「インタラクティブポエム」-芸術と科学の融合地点-映像情報メディア学会誌-文化としての映像情報メディア-小特集1998年1月号 Vol:52, No.1 613, pp.60 - 63.

## 研究会等

- ・ 柿本正憲、土佐尚子、森淳一、真田麻子、「音声に反応する表情合成システム”ニューロベビー”」情報処理学会第44回（平成4年前期）全国大会（1992）.
- ・ 柿本正憲、土佐尚子、森淳一、真田麻子、「ニューロベビー：音声から感情を認識して反応する表情合成システム」テレビジョン学会技術報告 Vol.16, No.33, pp7-12（1992）.
- ・ 土佐尚子、「芸術とテクノロジー」情報処理学会グラフィクスとCAD部会研究会 pp.73-80（1992）.
- ・ 土佐尚子、中津良平、「フューチャームービーのためのインタラクティブアクターの生成」日本映像学会全国大会（1996）.
- ・ 中津良平、土佐尚子、「インタラクティブ映画構築に向けて—Inter Communication Theaterの基本概念と構成例—」電子情報通信学会画像工学研先会 IE96-113, pp.31-38（1997）.
- ・ 江谷典子、土佐尚子、中津良平、「インタラクティブなストーリー生成のためのスクリプト制御」1997年電子情報通信学会総合大会 A16-10, pp.382（1997）.
- ・ 中津良平、土佐尚子、越知武、「将来メディアとしての観客参加型映画システム」1997年電子情報通信学会総合大会 SB-7-1, pp.778-779（1997）.
- ・ 土佐尚子、中津良平、「インタラクティブコミュニケーションシアター：Wonderland」-VR.人工知能・感情認識機能を用いた次世代映画研究- 日本映像学会第23回大会（1997）.
- ・ 中津良平、土佐尚子、「Inter Communication Theater -インタラクティブ映画構築の試み-」インタラクティブ'97論文集 pp.67-68（1997）.
- ・ 中津良平、土佐尚子、「Future Movie Systemの基本コンセプトとシステム構成」情報処理学会第54回全国大会講演論文集, Vol. 4, No. 5P-2, pp.4-21 -4 - 22（1997）.
- ・ 中津良平、土佐尚子、越知武、「複数参加型インタラクティブ映画システムの構成」第3回知能情報メディアシンポジウム論文集, pp.203-210（1997）.

- ・ 中津良平、土佐尚子、越知武、「インタラクティブ映画第2次システムの構成」インタラクシオン'98 論文集, pp.93-100 (1998).
- ・ 中津良平、土佐尚子、越知武、「複数参加型インタラクティブ映画システムの構成」1998年電子情報通信学会総合大会 講演論文集, Vol.A-16-28, pp.379 (1998).

## 受賞

- ・ ベストペーパーアワード

N.Tosa, R. Nakatsu, "Life-Like Communication Agent- Emotion Sensing Character 'MIC'&Feeling Session Character, 'MUSE'," The Third IEEE International Conference on Multimedia Computing and Systems (ICMCS)pp.12-19 (1996).

- ・ 1997年ロレアル賞受賞

土佐 尚子、中津良平、「インタラクティブポエム」  
(芸術と科学に関する研究に与えられる表彰)

- ・ 第13回電気通信普及財団賞（テレコムシステム技術賞）受賞

中津良平、土佐尚子、越知武、鈴木秀昭、「インタラクティブ映画システムのコンセプトと構成例」  
電子情報通信学会論文誌 D-II Vol.J81-D-II, No.5, pp.944-953 (1998)

## 展示

### < 1993 年度 >

- 1) システム・作品名：Talking to Neuro-Baby
- 2) 展示会名（主催者）：ARS Electronica'93
- 3) 展示場所、展示期間：Linz, オーストリア
- 4) 招待／自主出品：招待

- 1) システム・作品名：Neuro-Baby
- 2) 展示会名（主催者）：ACM SIGGRAPH93
- 3) 展示場所、展示期間：Los Angels, U.S.A.
- 4) 招待／自主出品：自主

- 1) システム・作品名：Neuro-Baby 「未分化な感情体」
- 2) 展示会名（主催者）：人工生命の美学
- 3) 展示場所、展示期間：東京国際美術館
- 4) 招待／自主出品：招待

### < 1995 年度 >

- 1) システム・作品名：Networked Neuro-Baby with hand
- 2) 展示会名（主催者）：東京大学生産技術研究所研究公開
- 3) 展示場所、展示期間：六本木（東京）6月
- 4) 招待／自主出品：共同研究成果展示

- 1) システム・作品名：Networked Neuro-Baby with hand
- 2) 展示会名（主催者）：ACM SIGGRAPH95
- 3) 展示場所、展示期間：Los Angels, U.S.A.
- 4) 招待／自主出品：招待

- 1) システム・作品名：Neuro-kid "MIC" / Feeling Session "MUSE"
- 2) 展示会名（主催者）：「マルチメディア・ガーデン」（アジア太平洋トレードセンター株式会社）
- 3) 展示場所、展示期間：ATC デザインギャラリー（大阪）11月23日～12月3日
- 4) 招待／自主出品：招待

< 1996 年度 >

1) システム・作品名：Neuro-kid "MIC"

2) 展示会名（主催者）：「ひとがた・カラクリ・ロボット展」

財団法人品川文化振興事業団 O（オー）美術館

3) 展示場所、展示期間：O（オー）美術館 大崎（東京） 2月2日～3月3日

4) 招待／自主出品：招待

1) システム・作品名：Neuro-kid "MIC"

2) 展示会名（主催者）：A-Life V

3) 展示場所、展示期間：奈良 5月16-18日

4) 招待／自主出品：招待

1) システム・作品名：Neuro-kid "MIC"

2) 展示会名（主催者）：IEEE Multimedia Conference Exhibition

3) 展示場所、展示期間：広島、5月

4) 招待／自主出品：自主

1) システム・作品名：Neuro-kid "MIC"

2) 展示会名（主催者）：マルチメディア大航海（ソフトピアジャパン）

3) 展示場所、展示期間：大垣（岐阜）6月

4) 招待／自主出品：招待

1) システム・作品名：Neuro-kid "MIC"

2) 展示会名（主催者）：「NTT プラネット」（NTT）

3) 展示場所、展示期間：名古屋、常設

4) 招待／自主出品：招待

1) システム・作品名：Neuro-kid "MIC"

2) 展示会名（主催者）：「美術家の冒険」-多面化する表現と技法- 国立国際美術館

3) 展示場所、展示期間：国立国際美術館（大阪）、9月5日～11月24日

4) 招待／自主出品：招待

- 1)システム・作品名：Neuro-kid "MIC"
- 2)展示会名（主催者）：「ハイビジョンウイーク」（NHK）
- 3)展示場所、展示期間：NHK スタジオパーク（東京） 11月
- 4)招待／自主出品：招待

< 1997年度 >

- 1)システム・作品名：Neuro-kid "MIC"
- 2)展示会名（主催者）：「Cyber Forum」（PORTUGAL TELECOM）
- 3)展示場所、展示期間：LISBON（PORTUGAL） 4月11日～5月8日
- 4)招待／自主出品：招待

- 1)システム・作品名：Interactive Poem
- 2)展示会名（主催者）：ISEA97
- 3)展示場所、展示期間：Chicago, 9月
- 4)招待／自主出品：自主

- 1)システム・作品名：インタラクティブポエム
- 2)展示会名（主催者）：「デジタルメディア・ワールド」  
(日本コンピュータグラフィックス協会／日経)
- 3)展示場所、展示期間：幕張メッセ、12月
- 4)招待／自主出品：招待

< 1998年度 >

- 1)システム・作品名：Interactive Poem
- 2)展示会名（主催者）：Virtual Environments'98 (Conference and 4th Eurographics Workshop)
- 3)展示場所、展示期間：Stuttgart, Germany 16-18 June 1998
- 4)招待／自主出品：招待

- 1)システム・作品名：Interactive Poem
- 2)展示会名（主催者）：Artificial Life IV
- 3)展示場所、展示期間：L.A. U.S.A. 26-29 June 1998
- 4)招待／自主出品：招待

1)システム・作品名：Interactive Poem

2)展示会名（主催者）：The 6th ACM International Multimedia Conference Art Demo

3)展示場所、展示期間：Bristol, U.K. 26-29 Sept. 1998

4)招待／自主出品：自主

<1999年度>

1)システム・作品名：Interactive Cinema "Romeo & Juliet in Hades"

2)展示会名（主催者）：Trancemedia99（ベルリン映画祭ニューメディア部門）

3)展示場所、展示期間：Berlin, Germany 12-22 Feb. 1999

4)招待／自主出品：招待

## 招待講演

「芸術とテクノロジーLife-Like Autonomous Character "MIC" & Feeling Session Character "MUSE"」フエジイ学会講習会：感性情報を見る、活かす、遊ぶ。(1996).

「サイバームービーにおけるインタラクティブアクターの生成」人工知能学会主催：人工知能セミナー(1996).

「アート&テクノロジーから生まれるマルチメディアの新コンテンツ」通産省九州マルチメディア懇親会の総会における基調講演(1997).

「コンピュータ生命」感性、情緒型ロボット研究会第7回研究会(1998).

Keynote talk: "Creation of Virtual Theater -Interactive Poem and Interactive Cinema," Virtual Environments'98 IEEE YUFORIC Germany'98 (1998).

「くらしとバーチャルリアリティー」VRシンポジウムぎふ'98, 第6回学生対抗手作りバーチャルリアリティーコンテスト:IVRC'98 (1998).

「人間と感性でコミュニケーションできる自律アクターとその劇場空間の研究」電子情報通信学会・画像工学研究専門委員会主催：画像符号化シンポジウム(1998).

## 委員

1996-1998 CG 検定／マルチメディア検定／画像処理検定試験問題作成委員

(財) 画像情報教育振興協会

1996-1998 人工生命を応用した仮想生態系展示システム研究委員

(財) 日本科学技術振興財団

1998 MMCA 市場環境整備事業 (先導的コンテンツ制作及び技術開発支援) 採択審査委員

(財) マルチメディアコンテンツ振興協会

1998 Program Committee of Virtual Environments'98

(4th Eurographics Conference and IEEE YUFORIC Germany)

## その他の研究業績

口頭発表 <全国大会、研究会、講演等>

<研究会>

・土佐尚子、「芸術とテクノロジー」情報処理学会人文科学とコンピュータ部会研究報告(91-CH-11)

<主な講演>

・サッポロテクノパークフェスタ'94セミナー

「メディア学概論 "情報伝達の媒体の原点をさぐる" ('94/9/15)

・光工学会：マルチメディアセミナー 「メディア・アート」 Artificial Life Character ('94/8/20)

・Visual Computing'94 パネルディスカッション「ヒューマン・インターフェース」('94/6/18)

### 作品収蔵美術館

国立国際美術館	(大阪)
O(オー)美術館	(東京)
高松私立美術館	(香川)
イメージフォーラム	(東京)
日本映像カルチャーセンター	(東京)
富山県立近代美術館	(富山)
国際交流基金	(東京)
名古屋県立美術館	(名古屋)
アメリカンフィルム・アソシエーション	(N.Y. U.S.A.)

### 賞歴

1985 SCAN'85 Autumn (ビデオギャラリー SCAN 東京)

2nd インターナショナルビエンナーレビデオ CD'85 (Ljubljana ユーゴスラビア)

ニコグラフ'85 佳作賞 (池袋サンシャインシティ文化会館 東京)

1986 NEW VIDEO JAPAN 招待出品 (ニューヨーク近代美術館)

NCGA'86 (National Computer Graphics Association)

インディペンデントアーティスト部門3位受賞 (アナハイムカリフォルニア U.S.A.)

SIGGRAPH'86 Animation Screening 部門入賞 (ダラステキサス U.S.A.)

国際ハイ・テクノロジーアート展 銅賞受賞 (池袋サンシャインシティ文化会館東京)

日本映像フェスティバル特別賞受賞 (東京)

- 1987 フジTV主催 国際映像&音楽大賞最優秀演出賞受賞 (国立代々木競技場第一体育館東京)  
 第14回 OMNI アートコンテスト優秀賞受賞 (旺文社日本版 OMNI 主催)  
 アートドキュメント'87 (栃木県立美術館)  
 アメリカンフィルム&ビデオフェスティバル ビデオアート部門2位受賞 (N.Y. U.S.A.)  
 The 21st Annual New York Film, Video Exposition 入賞 (メトロポリタン美術館 N.Y. U.S.A.)  
 イメージフォーラム主催 フィルム&ビデオフェスティバル ビデオ部門入賞  
 (渋谷西武 SEED ホール東京)  
 サンフランシスコ インターナショナル フィルムフェスティバル "Golden Gate Awards" 入賞  
 SIGGRAPH-87 ArtShow 入賞 (アナハイムカリフォルニア U.S.A.)  
 ARS Electronica'87 入賞 (LINZ オーストリア)  
 The 2nd Emerging Expression Biennial 入賞 (Bronx Museum N.Y. U.S.A.)  
 ビデオ カルチャー カナダ'87入賞 (トロント カナダ)
- 1988 BACA'S 21st Annual Film-Video Festival 入賞  
 (ジェファーソン マーケット ライブラリー N.Y. U.S.A.)  
 ARS Electronica'88 入賞 (LINZ オーストリア)  
 NEC 主催 パソコンアート大賞大賞授賞
- 1989 ARTEC'89 第一回名古屋国際ビエンナーレ 国内公募部門入賞 (名古屋市科学館展示ホール)
- 1990 SCAN'90 ビデオアート新作公募展入賞 (ハイネッケンビレッジギャラリー 東京)
- 1993 ARS Electronica'93-Artificial Life- 入賞 (LINZ オーストリア)  
 SIGGRAPH'93 Mashine Culture 部門入賞 (アナハイムカリフォルニア U.S.A.)

#### 個展

- 1985 個展 "Video Art Works" (ギャラリー NORTH FORT 大阪)  
 1985 個展 "Film & Video" (QUEST 広島)  
 1987 個展 "VIDEO ART EXHIBITION" (Annex プラザ 名古屋)  
 1990 個展 "Interactive Art & Video" (ギャラリー光彩 東京)  
 1992 個展 "NEURO - BABY'S Birthday !" (ガーディアン・ガーデン 東京)  
 1993 個展 "未分化な感情体" 「Neuro - Baby」 (イムズビル三菱ソシオテックプラザ 福岡)

#### 演劇パフォーマンス

- 1981 脚本 別役 実 劇団葡萄一族公演 "屍" にて役者として参加 (秀巧社ホール 福岡)  
 1984 - 如月小春主宰 劇団ノイズとコラボレーションを始める - 脚本・演出 如月小春 "MORAL"  
 (スタジオ R 東京)

- 1985 脚本・演出 如月小春 "ISLAND" (松竹ベニサンピット 東京)
- 1986 脚本・演出 如月小春 "ISLAND:YOKOHAMA" (神奈川県立青少年センターホール)
- Dancing Voice in SEED "都会の生活" (渋谷西武SEEDホール 東京)
- 脚本・演出如月小春"SAMSA" (FM東京ホール東京)
- 大阪国際演劇祭 招待出品 脚本・演出 如月小春 "MORAL3rd" (近鉄小劇場 大阪)

主な展覧会 CG・ビデオ・フィルム・インタラクティブアート

- 1981 女流画家展 (福岡市美術館)
- 1982 ビデオ・アンデパンダン展 (大阪現代美術センター)
- 芸術機能展 (福岡市美術館)
- 1983 新世代のビデオアート展 (駒井画廊 東京)
- 多種多様態展 (ギャラリーおいし 大阪)
- 多種多様態展 (島田美術館 熊本)
- パーソナル・フォーカス '83 (福岡県立美術館/イメージフォーラム 東京)
- OPEN FILM BOX'83 (秀巧社ホール 福岡)
- 1984 メディア・コレクション (ギャラリーK 東京)
- ビデオ・カクテル (駒井画廊/スタジオ200 東京)
- パーソナル・フォーカス '84 (福岡市民センター/イメージフォーラム 東京)
- 1985 パーソナル・フォーカス '85 (福岡市民センター/イメージフォーラム 東京)
- ビデオ・カクテル2 (ギャラリーNEWS 東京)
- ビデオカルチャー・カナダ '85 招待出品 (トロント カナダ)
- イベント メディア・スクラッチ (ラフォーレミュージアム赤坂 東京)
- 1986 モンベリアール・インターナショナルビデオフェスティバル招待出品
- (モンベリアール フランス)
- オーストラリアン インターナショナルブロードキャスティングプログラム
- インターナショナルビデオアート部門招待出品
- 日本のビデオアート展 "Scanners" 招待出品 (Air Gallery U.K.)
- ビデオ・カクテル3 (原美術館 東京)
- OPEN FILM BOX Vol.9 (福岡県立美術館)
- カメリーノ・インターナショナルビデオフェスティバル招待出品 (イタリア)
- 1987 WAVEFROMS:VIDEO FROM JAPAN (サンフランシスコ U.S.A.)
- Festival International de Films et Videos de Femmes de Montreal 招待出品 (ケベック カナダ)
- 日本のCGアート展 招待出品 (O美術館 東京)
- FILM-TANK 招待出品 (福岡県立美術館)

- 国際ハイ・テクノロジーアート展 招待出品 (NSビル1階 東京)  
 ACM/SIGGRAPH TRAVELING ART SHOW "25years of Computers in the Arts"  
 (カリフォルニア州立大)
- 1990 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招待出品  
 (モンペリアール フランス)  
 ロカルノ・インターナショナルビデオフェスティバル招待作品 (ロカルノ スイス)
- 1991 DIGITAL IMAGE展 (銀座ワシントンギャラリー 東京)  
 移動する遊体展 (小原流会館オハラエスパス 東京)
- 1992 DIGITALIMAGE展 (銀座ワシントンギャラリー 東京)
- 1993 第18回日本映像学会全国大会研究発表展 (武蔵野美術大学 東京)  
 NICOGRAPH'92 特別展示「バーチャルリアリティとインタラクティブCGへの招待」(東京)  
 Discussion of TECHNOLOGY ART Creativity Vol.1 (川崎市民ミュージアム)  
 A-Life WORLD展 (東京国際美術館)
- 1994 NHK 放送文化 WEEK 展 (NHK 放送センター正面玄関ロビー 東京)  
 双方向美術展 (天保山ミュージアム 大阪)
- 1995 The Role of Computer Graphics Technology of the Human Perception of Reality  
 (オハイオ州立大学 U.S.A.)  
 SIGGRAPH'95 Network Event "Network Neuro-Baby with hand" (ロス・アンゼルス U.S.A.)
- 1996 ひとがたカラクリロボット展 (O美術館 東京)  
 Artificial Life V (Nara-Ken Public Hall 奈良)  
 美術家の冒険展 (国立国際美術館 大阪)  
 NHK ハイビジョン WEEK 展 (NHK 放送スタジオパーク 東京)
- 1997 Cyber Forum (Culture Center of Belem Portugal)

作品リスト

< CG & VIDEOGRAPHY >

Misic Image	1982	10min
Beens Dance	1983	6min
ARCHITECTURE	1984	2min
An Expression	1985	9min
TRIP	1985	8min
ESCTACY	1986	7min
ISLAND-image-	1986	7min
ENERGY	1987	3min
Pleasure	1988	3min
Its Grotto be Somewhere	1988	10min
TRANCE	1989	6min
GUSH !	1989	7min

< インタラクティブアート >

RS- 705	1988
Neuro-Face	1990
Neuro-Baby	1992
Neuro-Baby2	1993
Neuro-Girl	1994
Network Neuro-Baby with hand	1995
Neuro-Kids "MIC"	1995-1996
Interactive Poem	1997
Interactive Cinema "Romeo & Juliet in Hades"	1998
Unconscious Flow	1999

< FILMOGRAPHY >

MEDITI	1982	3min
MOVEMENT	1983	5min
意識下の鼓動	1983	15min
AMALGAM	1984	10min
MMMM.....	1985	6min
気	1985	3min
EAT	1986	3min

< CG 静止画 >

プロバガンダ	1986
BIGEYE	1987
EVER	1990
EVE	1991
FLOWERS	1993

作品と作家についての主な文献掲載リスト

・書籍

- びあギャラリーコレクション 「O美術館コレクション」 pp.104 (びあ株式会社)
- 吉田敬一著「教養・コンピュータ」静岡大学教科書 pp.155-1 57 (共立出版株式会社)
- 現代デザイン事典1994年版「インタラクティブアート」 pp.2-3 (平凡社)
- 「思想のソフトウェア」高橋英之著 (株)法蔵館
- 「仮想楽園へ行こう！」 (株)アスキー
- 「人工生命とは何か」白石明彦著 pp.151-156(1995) (株)丸善
- 「仏教」No.24 特集：生命 pp.125-133 (1993) (株)法蔵館
- 図解「人工生命を見る」佐倉統監修、高間康史著 pp.171-175 (1998) (株)同文書院
- "Postmodern Currents" -Art and Artists in the Age of Electronic Media-  
Written by Mergot Lovejoy (1997) Published by Prentice- Hall

・作品掲載画集

- World Graphics Design Now 第6刊 "Computer Graphics" (講談社)
- '89日本アド・プロダクション年鑑 特集コンピュータによる視覚世界 (六耀社)
- CGアートワークス「デジタルイメージ編」 (グラフィック社)
- '93・3月号 日経CG表紙 (日経BP社)
- 飯村隆彦著「'80年代芸術・フィールドノート」表紙CG (朝日出版社)
- 山本純介(現代音楽作曲家)「レクイエム」CDカバー

・展覧会カタログ (国内)

- '89第2回ふくい国際ビデオビエンナーレ pp.135 (福井県立美術館)
- '89第4回現代芸術祭 -映像の今日- (富山県立近代美術館)
- 「ART ON COMPUTER」 (O美術館)
- '92「アニメ進化論」 (O美術館)
- '93 A-Life World展 (東京国際美術館)
- '96「ひとがた・カラクリ・ロボット」展 pp.58-59 (O美術館)
- '96「美術家の冒険」(多面化する表現と手法) pp.49-64 (国立国際美術館)

・展覧会カタログ (海外)

- 国際交流基金海外巡回企画 "Private Vision-Japanese Video Art in the 1980s"
- ARS Electronica '93 -Artificial Life-

SIGGRAPH '93 Visual Proceedings	"Machine Culture"	pp.167
Cyber Forum (Lisbon, Portugal) 1997		pp.20
SIGGRAPH '98 Visual Proceedings	"Sketches" [Art, Design, Multimedia]	pp.300
Transmediale'99 (International Berlin Film Festival New media dept.)		pp.88

・テレビ放映

1993/6/12	ORF	(オーストリア国営放送)「Ars Electronica」紹介
1993/5/16	NHK	サンデー経済スコープ「生物に近づく人工生命」
1993/5/21	NHK	衛星第2ハイビジョン放送「Hi からトーク」-電子雲母-
1992/11/24	NTV	ニュース番組「プラス1」
1993/3/2	TBS	平成ふしぎ探検隊「夢のロボット最前線」
1993/11/10	TBS	ニュース番組「ニュース6:30」
1993/11/25	テレビ朝日	ニュース・ステーション「人工生命特集」
1996/8/6	テレビ朝日	たけしの万物創世紀「ロボット・人工生命特集」
1996/11/6	NHK	ためしてガッテン
1997/3/2	NTV	東芝スペシャル「生命35億年地球の未来に会う旅」
1997/10/18	NHK 教育	未来潮流「機械に心は宿るのか」
1997/12/5	NHK 横浜	「電脳新時代」
1998/1/16	NTV	「ART TIME」

・ビデオソフト

NHK ビジネス情報(1993)

・新聞 (国内)

朝日新聞	文化欄 「人工生命の美学」	1993/7/6 (タ刊)
日経産業新聞	「先端技術-感性工学-」	1993/3/2
日本経済新聞	「電子ペット」	1993/7/7 (タ刊)
毎日新聞	先端技術欄	1992/11/19
読売新聞	1面	1993/5/10 (タ刊)
朝日新聞文化欄	「アートにも人工生命の発想」	1994/11/12 (タ刊)
読売新聞御正月特集号	「メディア・アート」	1995/1/1
読売新聞	私の名前は「ミューズ」	1996/5/15
日経産業新聞	「先端技術-人工生命-」	1996/5/22
日刊工業新聞	「次世代映画は」	1995/5/28

産経新聞	「未来史閲覧」	(1996/7/9)
産経新聞	「次代」探訪 - "感情" 持つ日も近い -	(1996/7/22)
朝日新聞	「美術家の冒険」展 - 枠組みを破る新メディア -	(1996/10/11)
中日新聞	感情劇場' 97 「電腦感情」に挑む	(1997/1/15)
朝日新聞	学芸欄 学研都市探検隊	(1997/3/9)
読売新聞	電子ベビー 21世紀の光	(1997/7/2)
産経新聞	情報通信の最先端 - 関西文化学術研究都市 -	(1998/1/1)
日刊工業新聞	ロレアル大賞に2作品	(1997/6/25)
読売新聞	ロレアル大賞	(1997/6/26)
日本経済新聞	てくのろじー考 - 芸術と科学の間 -	(1997/7/7)
東京新聞	科学 - 科学と芸術の出会い 第1一回ロレアル賞 -	(1997/7/8)
読売新聞	顔欄	(1997/7/29)
The Mainichi Daily News	"People & Events"	(1997/8/7)
朝日新聞	Person- 対話できる「電腦」を目指す -	(1997/9/12)

・ 評論

Design News No.192	1987 "イラスト・リレーション4"
Design News No.193	1987 "イラスト・リレーション5"
Design News No.195	1987 "イラスト・リレーション6"
(Design News - 財団法人日本産業デザイン振興会の出版物 (通産省))	

・ 対談

[特集]メディアとアートの21世紀的關係 - 新しい情報の技術で芸術の概念を変える -  
「自動車とその世界」トヨタ自動車株式会社広報部 pp.24-35 (1996)

・ 雑誌 (国内)

ビデオ・アビック	"Creator's File"	('85/1)
ビデオ・コム	"VIDEOCRITICISM"	('86/1)
月刊ギャラリー	"東京/N.Y. ポップアートの熱い波"	('86/8)
美術手帳	"ハイティック・アートの現在形"	('86/10)
OMNI (日本版)	"OMNIARTCONTEST"	('87/1)
日経CG	"アート&プレゼンテーション"	('87/1)
日経CG	"CGパーソン"	('87/1)
月刊美術	"若手作家特集"	('89/5)

日経 CG	"CGGALLERY"		('89/1)
PIXEL	"ピクセル・ブレイク"		('89/12)
月刊ギャラリー	"メディア・アート"		('90/2)
21世紀版画	"CG批評：肉体とCG - 土佐尚子"	pp.178	('91/3)
PIXEL	"日本のCGを支えるプロフェッショナル"	pp.35	('93/5)
Inter Communication	「Interactive Art」	pp.91	NTT出版 ('93/No.7)
AERA	「人工生命の世界」		朝日新聞社 ('93/7/27)
ファミコン通信	「読み切りショートショート 1990年のゲームキッズ# 13」		(株)アスキー ('93/7/27)
ファミコン通信	「TOPICS」		(株)アスキー ('93/3/26)
月刊アスキー	「ニコグラフ特集」		(株)アスキー ('93/1/1)
週刊宝石	「ハイテク名場面 No.168」		光文社 ('93/4)
THE 21	「今月のKEYWORD」		PHP研究所 ('93/3)
日経 CG	「ニュース・ディスプレイ」		日経BP社 ('92/3)
季刊 かたち	「移動する遊体展」		かたち社 (1992/No.20 春号)
月刊誌	「学校コンピュータ」		(株)少年写真新聞社 (Vol. 22)
PIXEL	「ニコグラフ92」		図形処理情報センタ (1993/1)
ニューズウィーク (日本版)	「人工生命研究特集」	p.70	(1995/12/27)
SCIas サイエンス朝日	「CGキャラクターが笑ってすねる」	pp.33-34	(1996/7/5)
Bart	「喜怒哀楽、快、不快、感情を表す....」	pp.89-90	(1997/2/24)
AXIS SCIENCE&IMAGING	特集：人工生命ネットワークニューロベイビー		(1997/5/1)
BRUTUS	「ノウ・脳・Darling!」	pp.17-18	(1997/7/1)
小学3年生	「感情を持つコンピュータ」		(1997/12/1)
marie claire	「芸術と科学が出会うとき」		(1997/10/1)
美術手帳	「エレクトロニック・アートの現在を知る」	p.214	(1998/3/1)
・新聞 (外国)			
De Telegraaf	"Computers met gevoel zijn in aantocht"		(1996/4/30)
・雑誌 (外国)			
HIGH PERFORMANCE	"View Point"		(1987/11)
NEW MEDIA			(1993/1)
Computer Graphics World			(1993/8)
VIRTUAL	"EMOZIONI NELLE MACCHINE/Calcolare l'ine"		(1997/11/1)