

エネルギー科学研究科
エネルギー社会・環境科学専攻修士論文
題目： 動的表情認識による
感情推定手法に関する基礎研究

指導教官： 吉川 榮和 教授

氏名： 國 弘 威

提出年月日： 平成 11 年 2 月 8 日 (月)

論文要旨

題目：動的表情認識による感情推定手法に関する基礎研究

吉川 榮和 研究室

國弘 威

要旨：

コンピュータに代表される機械が社会に浸透するにつれて、人間と機械とのインタフェース、すなわちヒューマンマシンインタフェースが重要視されるようになってきた。本研究では、ヒューマンマシンインタフェースを人間にとってより自然なものとするために、人間の表情を認識し、感情を推定することができるシステムを試作することを目的として研究を進めた。

本論文では、はじめに、人間同士のコミュニケーションを参考に、表情をヒューマンマシンインタフェースに利用することの有用性及び重要性について述べたあと、本研究に関連する従来研究について考察している。そして、これらを背景として、表情の認識に動的な顔画像を用いること、及び感情の推定のためにリアルタイムで処理を行うことの必要性を述べている。

次いで、本研究で利用している基本 6 表情や表情記述法 FACS などの、表情及び感情に関連する心理学的知見についてまとめたあと、本研究で用いる感情推定手法について、全体の処理の流れを述べ、本手法を構成する、特徴量抽出部、表情認識部、視線方向検出部、感情推定部の 4 つについてそれぞれ説明を加えた。

さらに、上記の感情推定手法に基づいて試作したリアルタイム感情推定システムについて、まずシステム構築上の着眼点を挙げ、それを考慮したシステムの構成を示している。試作システムは感情推定手法の各部に対応する 4 つのサブシステムから構成され、各サブシステムにおいてどのような処理を行っているかを具体的に述べている。特徴量抽出サブシステムでは顔の変形を表す特徴量を、視線方向検出サブシステムでは視線方向をそれぞれ画像処理により抽出しており、表情認識サブシステムではファジィ推論により表情の認識を行う。

そして、実際に試作したシステムを用いて表情認識及び、視線方向検出の評価実験を行って、その結果を考察することにより、システムがどの程度の性能を備えているかを評価している。また、処理速度の計測を行い、試作したシステムがリアルタイムで動作することを確認している。

最後に、結論として本研究を総括し今後の課題を挙げている。

目次

第 1 章 序論	1
第 2 章 研究の背景と目的	3
2.1 研究の背景	3
2.1.1 相互適応型インタフェースの必要性	3
2.1.2 人間同士のコミュニケーション	4
2.1.3 ヒューマンマシンインタフェースへの表情の利用	6
2.2 従来の研究	6
2.3 研究の目的	7
第 3 章 感情推定手法	9
3.1 心理学における表情研究	9
3.1.1 感情の分類	9
3.1.2 表情筋と顔の 3 領域	12
3.1.3 基本 6 表情	13
3.1.4 表情記述法 FACS	15
3.2 感情推定手法	20
3.2.1 感情推定手法の概略	20
3.2.2 特徴量抽出部	20
3.2.3 表情認識部	22
3.2.4 視線方向検出部	22
3.2.5 感情推定部	22
第 4 章 リアルタイム感情推定システムの試作	27
4.1 システム構成	27
4.1.1 システム構築上の着眼点	27
4.1.2 ハードウェア構成	28
4.1.3 ソフトウェア構成	29

4.2	特徴量抽出サブシステム	31
4.2.1	顔領域の抽出	31
4.2.2	顔器官領域の抽出	38
4.2.3	特徴点の抽出	43
4.2.4	特徴量の算出	47
4.3	表情認識サブシステム	49
4.4	視線方向検出サブシステム	50
4.5	感情推定サブシステム	55
第 5 章	試作システム評価実験	57
5.1	表情認識実験	57
5.1.1	実験の目的	57
5.1.2	実験方法	57
5.1.3	実験結果及び考察	60
5.2	視線方向検出実験	68
5.2.1	実験の目的	68
5.2.2	実験方法	68
5.2.3	実験結果及び考察	70
5.3	感情推定に関する考察	78
第 6 章	結論	79
	謝辞	81
	参考文献	82

目次

2.1	相互適応型インタフェース	3
2.2	ノンバーバル情報の占める割合	5
3.1	表情の類似性評定による感情の位置づけ	11
3.2	次元評定法による感情の位置づけ	11
3.3	混合感情	12
3.4	主な表情筋	13
3.5	顔の3領域	14
3.6	「怒り」と「幸福」の表情の組み合わせ	14
3.7	基本6表情	15
3.8	感情推定手法の概略	21
4.1	ハードウェア構成	28
4.2	ソフトウェア構成	30
4.3	感情推定処理の概略	30
4.4	特徴量抽出サブシステムの概略	32
4.5	抽出する特徴点と特徴量	33
4.6	入力画像の座標系	34
4.7	顔領域の抽出方法	35
4.8	顔領域の探索領域	36
4.9	暗色部分の割合を求める領域	38
4.10	口領域の抽出方法	39
4.11	口領域の探索範囲	40
4.12	眉・目領域の抽出方法	42
4.13	眉・目の端点	44
4.14	眉・目の端点の選択条件(1)	44
4.15	眉・目の端点の選択条件(2)	44
4.16	眉・目の端点の候補点の分類方法	45

4.17 特徴点の抽出方法	46
4.18 眉の特徴点の抽出方法	48
4.19 目の特徴点の抽出方法	48
4.20 口の特徴点の抽出方法	48
4.21 表情認識に用いたメンバシップ関数	49
4.22 視線方向の検出方法	53
4.23 瞳の中心点の抽出方法	54
4.24 眼球の回転角と瞳の中心点の関係	55
5.1 表情表出動画像の作成	58
5.2 表情認識実験の実験風景	59
5.3 実験時の照明環境	59
5.4 入力画像例 (1)	63
5.5 入力画像例 (2)	64
5.6 特徴量抽出結果 (入力画像 1)	65
5.7 表情認識結果 (入力画像 1)	65
5.8 特徴量抽出結果 (入力画像 2)	66
5.9 表情認識結果 (入力画像 2)	66
5.10 被験者 N の「驚き」の表情	67
5.11 特徴点の抽出失敗例	67
5.12 視線方向検出実験用動画像の作成	69
5.13 視線方向検出実験における注視点	69
5.14 視線方向検出実験の実験風景	70
5.15 視線検出実験結果 (被験者 A、右目、同経度点を連結)	72
5.16 視線検出実験結果 (被験者 A、右目、同緯度点を連結)	72
5.17 視線検出実験結果 (被験者 A、左目、同経度点を連結)	73
5.18 視線検出実験結果 (被験者 A、左目、同緯度点を連結)	73
5.19 視線検出実験結果 (被験者 I、右目、同経度点を連結)	74
5.20 視線検出実験結果 (被験者 I、右目、同緯度点を連結)	74
5.21 視線検出実験結果 (被験者 I、左目、同経度点を連結)	75
5.22 視線検出実験結果 (被験者 I、左目、同緯度点を連結)	75
5.23 視線検出実験結果 (被験者 N、右目、同経度点を連結)	76
5.24 視線検出実験結果 (被験者 N、右目、同緯度点を連結)	76

5.25 視線検出実験結果 (被験者 N、左目、同経度点を連結)	77
5.26 視線検出実験結果 (被験者 N、左目、同緯度点を連結)	77

表 目 次

2.1	ノンバーバル情報	4
3.1	顔の上部において特徴的な表情動作	16
3.2	顔の中央部において特徴的な表情動作	16
3.3	顔の下部において特徴的な表情動作	17
3.4	Action Unit (1)	18
3.5	Action Unit (2)	19
3.6	基本 6 表情を表す AU の組み合わせ	19
3.7	顔の 3 領域の組み合わせによって表される感情	24
4.1	顔の上部のファジィルール	51
4.2	顔の中央部のファジィルール	51
4.3	顔の下部のファジィルール	51
4.4	試作システムで顔の 3 領域の組み合わせから推定できる感情	56
5.1	表情認識実験結果	62
5.2	1 フレームの認識に要する時間 (msec)	67

第 1 章 序論

現在、我々の周囲は様々な機械で満ち溢れている。家庭においても、職場や学校においても、パソコン、ビデオ、ファックス、電子レンジ…、数え上げればきりが無い。大量消費社会の中で、機械は絶え間なく進歩し多機能・高機能化する一方で、その恩恵を受けることが徐々に困難になっていくという矛盾が生じている。大抵の場合、基本的な機能だけで満足してしまい、便利な機能にはその存在すら気付かないことも多い。これまで、人間と機械とがインタラクションを行う場合、歩み寄るのは人間の側であり、人間は機械を理解し、機械に歩調を合わせようと努めてきた。しかし、そこに限界が見えはじめてきた現在、歩み寄るのは機械の方である。機械が人間を理解することにより、機械と人間との協調を試みる必要がある。より人に優しい機械の出現が望まれているのである。

それでは、どうすれば機械は人間を理解することができるのだろうか。そのヒントは人間同士のコミュニケーションの中に見い出すことができる。人間は相手を理解することにより円滑なコミュニケーションを行っている。その際、お互いの間で意図的、非意図的にかかわらず、様々な情報を発信し、そして受信している。そこでは、ことばもちろん重要であるが、ノンバーバルメッセージと呼ばれる、ことばによらないメッセージ、すなわち、表情、身振り、視線、声の抑揚から、服装や化粧に至るまでの多種多様なメッセージが、コミュニケーションにとって重要であることが知られている。このノンバーバルメッセージを機械と人間との間でもやり取りすることができれば、機械の親和性は飛躍的に向上すると考えられる。機械が人間に直感的に理解しやすい形のメッセージを発信し、また、人間が発したメッセージを受け取り、理解するのである。『2001年宇宙の旅』の HAL-9000、『スターウォーズ』の C3-PO と R2-D2、鉄腕アトム、ドラえもん etc、考えてみると、人間が思い描いてきた未来の機械像は、みなこのノンバーバル・メッセージの送受信機能を備えている。

本研究では、ノンバーバルメッセージの中でも最も内面状態を表すと言われる、顔表情の認識を研究対象とする。すなわち、機械を人間にとってより自然で負担のかからないものにするを旨として、ヒューマンマシンインタフェースに表情を利用する。また、人間は相手の表情変化を動的に読み取ることによって多くの情報を得ていることを考え、動的な表情を扱うことにする。そこで本研究では、まず心理学分野で

得られている知見を利用して、ビデオカメラなどによって撮影された動的な顔画像から表情を認識し、感情を推定する手法を提案する。そして、それに基づいてリアルタイムで表情を認識し、感情を推定するシステムの試作を行うことを目的としている。

以下に、本論文の構成について述べる。まず、第2章では、研究の背景について述べたあと、本研究に関連する従来研究についてまとめ、それらを背景とした本研究の目的を示す。第3章では、心理学分野における表情及び感情に関する知見についてまとめたあと、本研究で用いる感情推定手法について述べる。次いで、第4章では、第3章で述べた感情推定手法に基づいて試作したリアルタイム感情推定システムの詳細を述べる。第5章では、試作したシステムの評価実験について述べる。最後に、第6章で本論文のまとめと今後の課題を挙げる。

第 2 章 研究の背景と目的

本章では、人間同士のコミュニケーションにおける顔表情の役割とその重要性をもとに、ヒューマンインタフェースへの表情利用の有効性について述べたあと、従来の研究について触れ、それらを背景とした本研究の目的を述べる。

2.1 研究の背景

2.1.1 相互適応型インタフェースの必要性

コンピュータに代表される機械が人間の生活に浸透し、人間と機械とが接する場面が増加するにつれて、人間と機械とのインタフェース、すなわちヒューマンマシンインタフェースが重要視されるようになってきた。現在のところ、人間と機械とがコミュニケーションを行う場合には、歩み寄るのはほとんどが人間の側であり、人間にとって快適であるとはとても言えないのが実状である。ヒューマンマシンインタフェースを人間にとってより自然で負担の軽いものにするためには、これまでのように人間の側から機械に働きかけるだけでなく、機械の方でも人間の状態を察知し、それに即した対応ができるような、相互適応型のインタフェースを構築する必要がある (図 2.1)。そのためには人間同士のコミュニケーションが良い手本となる。

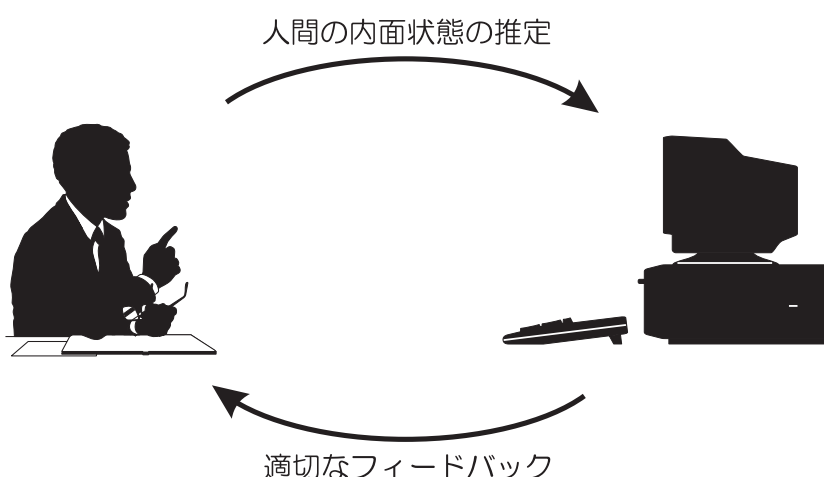


図 2.1: 相互適応型インタフェース

2.1.2 人間同士のコミュニケーション

人間と人間とが向かい合ってコミュニケーションを行う場合、「ことば」以外にも様々な情報が複数の媒体を通じて伝達される。それらはノンバーバル・メッセージと呼ばれ、表 2.1 のようなものがある^[1]。コミュニケーションにおいて伝達される全メッセージのうち、このノンバーバル・メッセージがどの程度の割合を占めているかを示したのが図 2.2 である。たとえば Birdwhistell は、人間同士の間で交わされるメッセージの 65% はノンバーバル・メッセージが担っているとしており、Mehrabian に至っては、メッセージの 93% がノンバーバルなものであるとしている。もちろん、文数や語数といった測度が見えるバーバル・メッセージとは違い、ノンバーバル・メッセージではその解釈が受け手に任せられているため、これらの数字は主観的なものでしかない。しかし、ノンバーバル・メッセージがコミュニケーションにおいて重要であることは、経験的にも明らかであると言えるだろう。

このように、人間はノンバーバル・メッセージをお互いに送受信することにより、はじめて円滑なコミュニケーションを行うことができる。そのため、人間と機械との間でもノンバーバル・メッセージを伝達することができれば、より自然なヒューマンマシンインタフェースが実現すると考えられる。

ノンバーバル・メッセージの機能としては、

表 2.1: ノンバーバル情報

ノンバーバル情報		属性など
身体動作	表情	
	視線	視線交差、凝視、無視
	瞳孔	散大、縮小
	身振り	
	口唇の動き	口話を含む
	姿勢(構え)	
周辺言語	声質	声域、アクセント、発音、ピッチ
	発声法	
	特徴性	笑い、泣き、うめき、叫び、つぶやき
	限定性	強勢、大小、長短
	遊離性	つなぎ、間、沈黙
対人距離		個人空間、距離、位置

1. 明示的情報の伝達

バーバル・メッセージが利用できない状況である場合や、バーバル・メッセージでは表現不可能か、十分正確に表現できない明示的情報を伝達する。

2. 発信者の内面状態の伝達

感情や、バーバル・メッセージを生成する心的過程で生じるイメージを伝達する。

3. 発信者のバックグラウンド情報の伝達

発信者の民族、文化的背景、出身地、家柄、所属社会・所属集団、社会的地位、学歴、職業などを伝達する。

4. 発信者の性格に関する情報の伝達

発信者の気質や性格を伝達する。

5. メタコミュニケーション機能

コミュニケーションの内容を規定したり、コミュニケーションの進行を調節したりする。

などがあるが^[1]、ヒューマンマシンインタフェースにおいては2.の発信者の内面状態の伝達が、最も有用な機能ではないだろうか。すなわち、人間から伝達されるノンバーバル・メッセージから、機械が人間の内面状態を読み取ってそれに則した処理を行い、また、機械の内部状態をノンバーバル・メッセージとして人間が直観的に理解しやすいかたちで提示することにより、人間の負担を大幅に軽減できると考えられる。

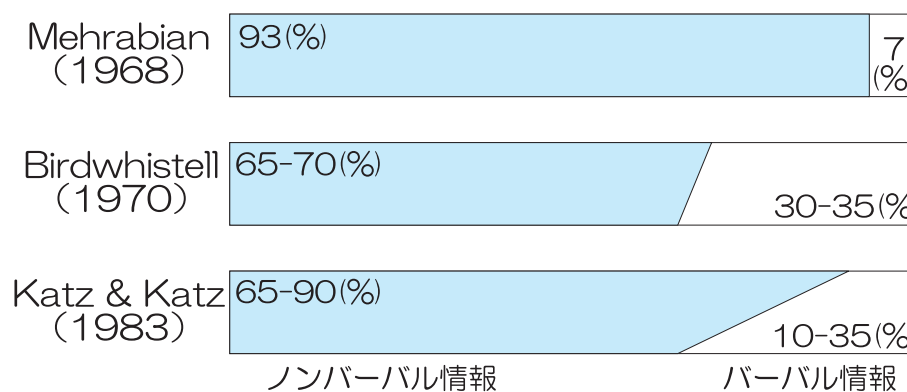


図 2.2: ノンバーバル情報の占める割合

2.1.3 ヒューマンマシンインタフェースへの表情の利用

ノンバーバル・メッセージのなかでも、表情は相手の内面状態を知る上で特に重要である。人間は相手の表情を読み取り、そこから相手の内面状態を推定することにより、円滑なコミュニケーションを行っている。人間同士のコミュニケーションにおいて、表情は必要不可欠なものである。そこで、表情をヒューマンマシンインタフェースに利用することにより、機械が人間にとってより身近で親しみやすい存在になると考えられる。

ヒューマンマシンインタフェースに表情を利用する場合のアプローチ方法としては、表情メッセージの伝達される方向から、

- 機械から人間へ表情により情報を提示
- 機械が人間の表情を認識

という2通りが考えられるが、本研究では後者の方法をとる。次節では、この分野で行われてきた研究についてまとめる。

2.2 従来の研究

表情をヒューマンマシンインタフェースに利用するためには、機械が自動的に人間の表情を認識しそれを解釈する必要がある。表情の認識に関しては、心理学及び工学の両分野で盛んに研究が行われてきているが、心理学分野の研究については次章の最初にまとめることとし、ここでは工学分野でこれまでに行われてきた研究について述べる。

まず、表情の認識及びその解釈過程は次の3段階に分けることができる。

第1段階 顔画像から顔の変形を検出する。

第2段階 顔の変形と表情(基本表情)を対応づける。

第3段階 表情から感情を推定する。

第1段階では顔画像から顔がどのように変形しているかを検出し、第2段階ではその顔の変形がどのような表情に対応しているかを調べ、第3段階でその表情がどのような感情を表しているかを推定する。

第1段階の顔の変形を検出する手法については、様々な試みが行われているが、その手法は次の3つに大別できる。

1. 顔画像から目や口などの器官 (顔器官) を抽出し、その各器官の変化をとらえる方法
2. 顔画像のオプティカルフロー*を用いる方法
3. 顔全体を1つのパターンとして取り扱う方法

1.の方法として、顔画像からラプラシアンフィルタにより抽出したエッジを利用して顔器官の抽出を行うもの^[2]や、動的輪郭モデル (SNAKES) を用いるもの^[3]、まばたきを手掛かりとするもの^[4]、テンプレートマッチングを行うもの^[5]などがある。また、カラー情報を用いた研究も行われており、HSV 表色系や YIQ 表色系が利用されている^[6, 7]。2.の方法としては、オプティカルフローから表情筋の動きを推定する方法^[8]や、オプティカルフローと顔の筋肉モデルを利用して、表情変化に伴う顔の各部位のエネルギー分布のパターンを求める方法^[9]などがある。3.の方法としては、表情の変形に伴う空間周波数成分の変化を利用するもの^[10]や、ウェーブレット変換を利用したもの^[11]などが提案されている。

第2段階の顔の変形と表情とを対応づける手法としては、ニューラルネットワークを用いるもの^[12]や、また、時系列情報を考慮したものとして、隠れマルコフモデルを用いて基本6表情を認識するもの^[13]や、リカレントニューラルネットワークを用いる方法^[14]などが提案されている。基本表情としては、3.1.3で述べる基本6表情、もしくはその一部が用いられる場合がほとんどである。

第3段階の表情からの感情の推定については、現在のところほとんど報告が行われていない。表情までを対象としていて感情については考慮しないか、第2段階で得られた基本表情がそのまま感情に対応しているとするかのどちらかである。

2.3 研究の目的

前節で述べたように、表情の認識に関しては多くの研究が行われており、様々な手法が提案されている。しかし、次に挙げる3点を考慮しているものは少ない。

* 輝度パターンの見かけ上の動き

- 動的な表情変化の認識

人間同士のコミュニケーションにおいて、人間は相手の表情を動的にとらえており、時間とともに微妙に変化していく表情の流れから相手の内面状態を読み取っている。そのため、ヒューマンマシンインタフェースに利用する際にも、動的表情認識、すなわち、動的な顔画像から表情の変化やその速さを検出することが必要であると考えられる。そのためには、次で述べるリアルタイム性が必要となる。

- リアルタイム性

表情をヒューマンマシンインタフェースに利用することを考えた場合、リアルタイム性が非常に重要である。表情の認識処理をリアルタイムで行うことができないければ、人間の内面状態の変化に即座に対応することができず、適切なタイミングでフィードバックを行うことができない。従来の認識手法のなかには、認識処理に1フレーム当たり数十秒から数分を要するものもあり、ヒューマンマシンインタフェースへの利用には適さない。また、高速なものでは毎秒数フレーム程度の認識を行えるものもあるが、それでも表情の動的変化を捉えるのに十分であるとは言えない。

- 複雑な感情の推定

前節でも述べた通り、従来の研究は基本表情の認識を目的とするものがほとんどであり、そこから複雑な感情を推定することについては考慮していない。

本研究では、これらの問題点を克服し、人間と機械との間により良いコミュニケーション方法を実現することを目指しており、その基礎的研究として、リアルタイムで顔の動画画像から表情を認識し、感情を推定できるシステムを試作することを目的とする。なお、本研究では視線方向も表情の一部としてとらえ、利用することにする。

第 3 章 感情推定手法

ここでは、まず心理学分野における表情及び感情に関する研究について述べたあと、その知見に基づく感情推定手法について述べる。なお、「表情」及び「感情」という言葉は日常でも広い意味・用法によって使用されるため混同されやすい。本研究では、「感情」が心的過程であるのに対し、「表情」は単に外見的なものであるとする。すなわち、人間の内面で生起している「感情」によって、顔面の形状変化である「表情」が引き起こされる。

3.1 心理学における表情研究

心理学の分野では、古くから表情及び感情に関する様々な研究が行われており、本研究で用いる感情推定手法ではその成果を積極的に活用している。本章では、心理学分野における表情及び感情に関する研究についてまとめる。

3.1.1 感情の分類

心理学において、感情は他の心的過程と比較すると十分な研究成果をあげているとは言えないようである。それは、感情研究には次のような感情特有の研究の難しさがあることに由来すると考えられる^[15]。

- 人の繊細で微妙な感情を客観的に捉えることは困難である。特定の感情状態にある人が客観的に内観もしくは自己観察することは難しく、その行為自体によって感情状態が変化してしまう。
- 知覚、精神発達、記憶などの心的過程は実験的研究が可能であるのに対し、感情の場合は倫理的問題が生じるため、真の恐怖、悲しみ、あるいは幸福といった感情を、実験的に誘起することが許されない。

感情の定義についても研究によって若干の差異がみられるようであるが、一般的には、感情は「人があらゆる事物、事象に触れたときに生じる情感の総称」であり、次のようなものをすべて含む^[15]。

- 好み：快・不快を表す比較的穏やかな主観的反応。
- 気分：現行の心的過程の基底を形作る長期にわたる感情。
- 情動：複雑な感情の組み合わせで、生理的喚起(覚醒)を伴った表出反応を含む。自律神経活動や行動から客観的に捉えられる。しばしば感情と同義に用いられる。

一方、表情の分析においては、通常の場合、感情をいくつかのカテゴリーに分類し、それぞれに表情やその要素を関連させていくという手法がとられる。以下で、表情分析の研究において感情を分類する手法について述べ、人間の感情がどのように分類されるかについてまとめる。

まず、感情を分類する手法の一つに、ある表情とそれから想起されるある感情語を関連させて分類していくものがある。この感情語によるカテゴリー判断によって、数多くの分類法が提案されたが、Ekmanらはそれらを比較検討してまとめ、感情カテゴリーとして「嬉しさ、驚き、恐れ、怒り、悲しみ、嫌悪、興味」の7つを抽出した。この7つの分類であれば比較的混同が少ないとされている^[16]。

次に、表情の形態的な特性を元に分類していく手法がある。例えば、RussellとBullockが多次元尺度構成法により表情の形態的類似性をプロットした結果が図3.1である^[17]。この方法では、感情カテゴリーは「退屈、眠り、満足、嬉しさ、興奮、驚き、恐れ、怒り、嫌悪、悲しみ」の順に円環上に配置される。「ニュートラル」の表情が図の中心にはなっておらず、感情としては「ニュートラル」である表情が必ずしも形態的な平均になっていないことが注目される。

また、感情を定量的に評価するために、快-不快などいくつかの表情認知次元を仮定して、感情カテゴリーをその相空間上に位置づけようとする、次元評定法と呼ばれる手法がある。認知次元の探求については数多くの研究が存在するが、主要な認知次元としては、第1次元には快-不快、第2次元には注意・覚醒度をとる場合が多い。感情カテゴリーが、快-不快、覚醒度・注意活動の2つの認知次元に対してどのように位置づけられるかについては、RussellとBullockが、図3.2のようにまとめている^[17]。

以上のように、感情の分類については様々な提案がなされているが、感情の持つ主観性・曖昧性により、詳細な分類法については未だ確立されていない。しかし、基本的な分類法としては、「怒り(anger)、嫌悪(disgust)、恐怖(fear)、幸福(happiness)、悲しみ(sadness)、驚き(surprise)」の6カテゴリーに分類することが広く定着しており、基本6感情と呼ばれている。そして、これらのそれぞれの感情のときに表出される表情が基本6表情であり、3.1.3で詳しく述べる。実際に感情が引き起こされる場合には、

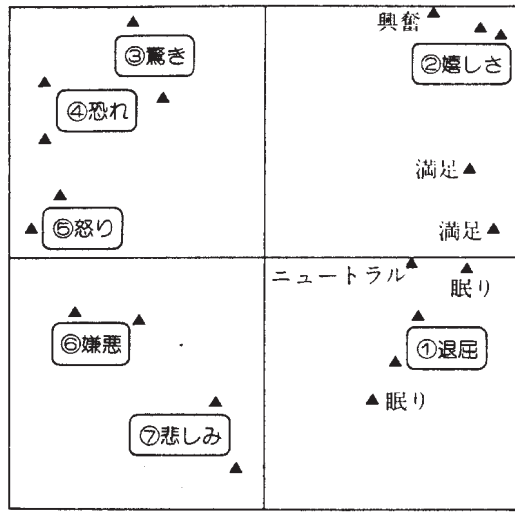


図 3.1: 表情の類似性評定による感情の位置づけ

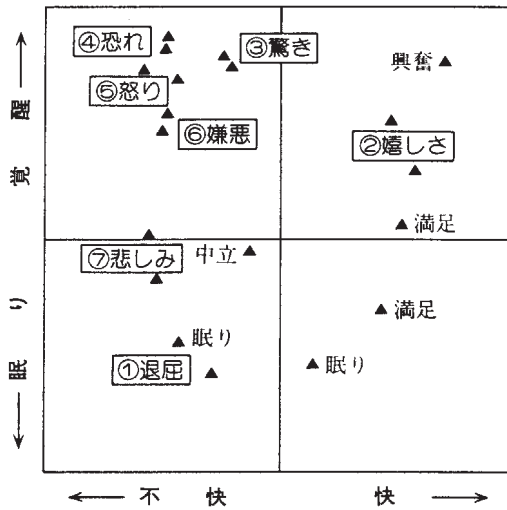


図 3.2: 次元評定法による感情の位置づけ

基本感情が単一に現れることは少なく、むしろいくつかの基本感情が混ざり合った混合感情として現れる。混合感情についても様々な分類法があり、例えば図 3.3 のような 33 種への分類が Plutchik により提案されている^[18]。この分類法では、円の中心部に位置する 8 つの基本感情の組み合わせが混合感情となる。まず、基本感情の一つ外側の円環上にあるのが、隣接する基本感情同士を混ぜ合わせた感情であり、第一次混合感情という。例えば、「恐れ」と「驚き」を混合すると「畏怖」となる。次に二番目の円環上にあるのが、一つ離れた基本感情同士の混合感情、第二次混合感情である。例えば、「恐れ」と「悲しみ」の組み合わせは「絶望」である。同様に、第三次混合感情、すなわち二つ離れた基本感情同士の混合感情が、一番外側の円環上に示されている。

3.1.2 表情筋と顔の 3 領域

表情変化は、表情筋と咀嚼筋の収縮に基づくものである。咀嚼筋は主に口の開閉を行う。表情筋は主なものだけでも 20 種類以上あり、それらの収縮・弛緩が複雑に組み合わせられて表情を形成している。なかでも重要なものは、前頭筋、皺眉筋、眼輪筋、口輪筋、頬骨筋、口角筋などである^[8, 16]。図 3.4 に各筋肉の位置関係を示す。

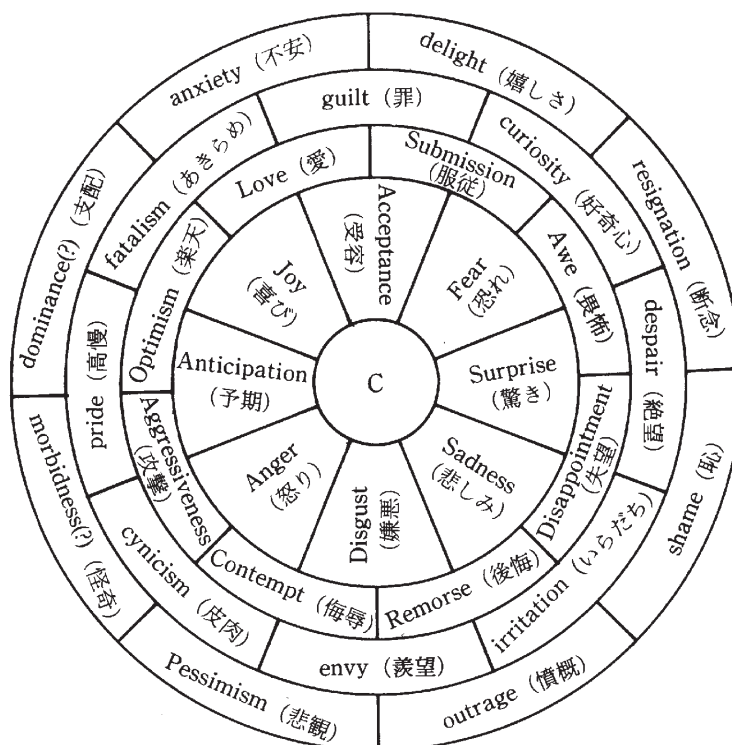


図 3.3: 混合感情

また、顔は表情筋の影響範囲・移動部分によって互いに独立した動作が可能な次の3領域に分割できる^[19]。

- 上部：眉と額の部分
- 中央部：目と鼻の付け根の部分
- 下部：口と頬の部分

これらの3領域の表情の組み合わせによって、混合感情を表す表情や感情の隠蔽を示す表情となる。たとえば、上部と中央部が「驚き」の表情を、下部が「幸福」の表情を示しているとき、それは単なる「驚き」の感情ではなく、うれしさを伴った驚きの感情を表している(図 3.5)。また、上部と中央部が「怒り」の表情を、下部が「幸福」の表情である場合の表情が図 3.6である。これは苦笑いの表情であり、怒りを隠そうとしている表情でもある。また、怒りとともに勝利感を感じていることも考えられる^[19]。

3.1.3 基本 6 表情

前述した通り、基本 6 表情は基本 6 感情、すなわち「怒り (anger)、嫌悪 (disgust)、恐怖 (fear)、幸福 (happiness)、悲しみ (sadness)、驚き (surprise)」の 6 つを表す表情であり、表情に関連する研究において最も広く用いられている表情の分類法である。基本 6 表情は人種や文化、性別などに関係なく、ほぼ万国共通であることが実験的に確

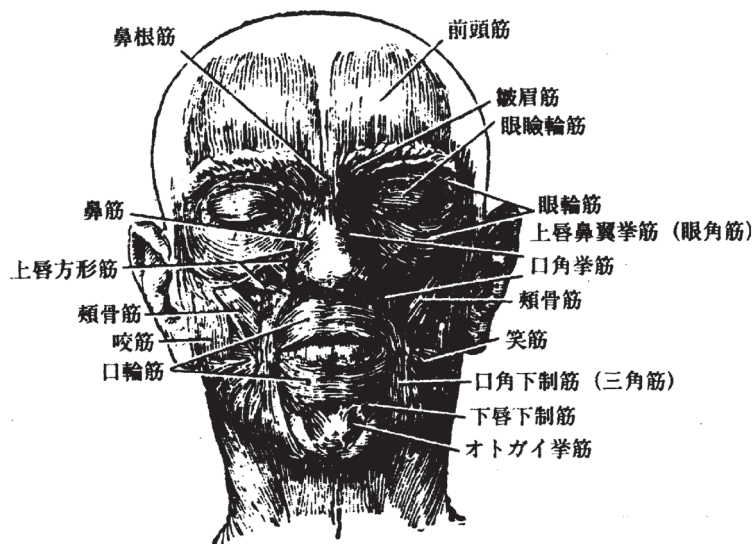


図 3.4: 主な表情筋



図 3.5: 顔の 3 領域



図 3.6: 「怒り」と「幸福」の表情の組み合わせ

かめられている。各表情にはそれぞれに特徴的な表情動作が存在し、それによって識別することができる。図 3.7に、基本 6 表情の典型的な表情形態を示す^[19]。また、表 3.1～表 3.3に各表情の形態的特徴をまとめる。

3.1.4 表情記述法 FACS

Facial Action Coding System(FACS)^[20]は Ekman と Friesen によって提案された組織的な表情記述法であり、表情に関連する研究において最も広く利用されている。FACS では、表情を Action Unit(AU) と呼ばれる解剖学的に独立し、かつ視覚的に識別可能な表情動作の最小単位によって記述する。AU は表 3.4及び表 3.5のように合計 58 個ある。なお、表中に AU の番号として No.66 までがあるが、途中で欠番が存在するため、AU の個数としては 58 個である。この AU の組み合わせにより人間のあらゆる表情が記述可能であると言われている。たとえば、頬を持ち上げる動作を表す AU6 と唇の端

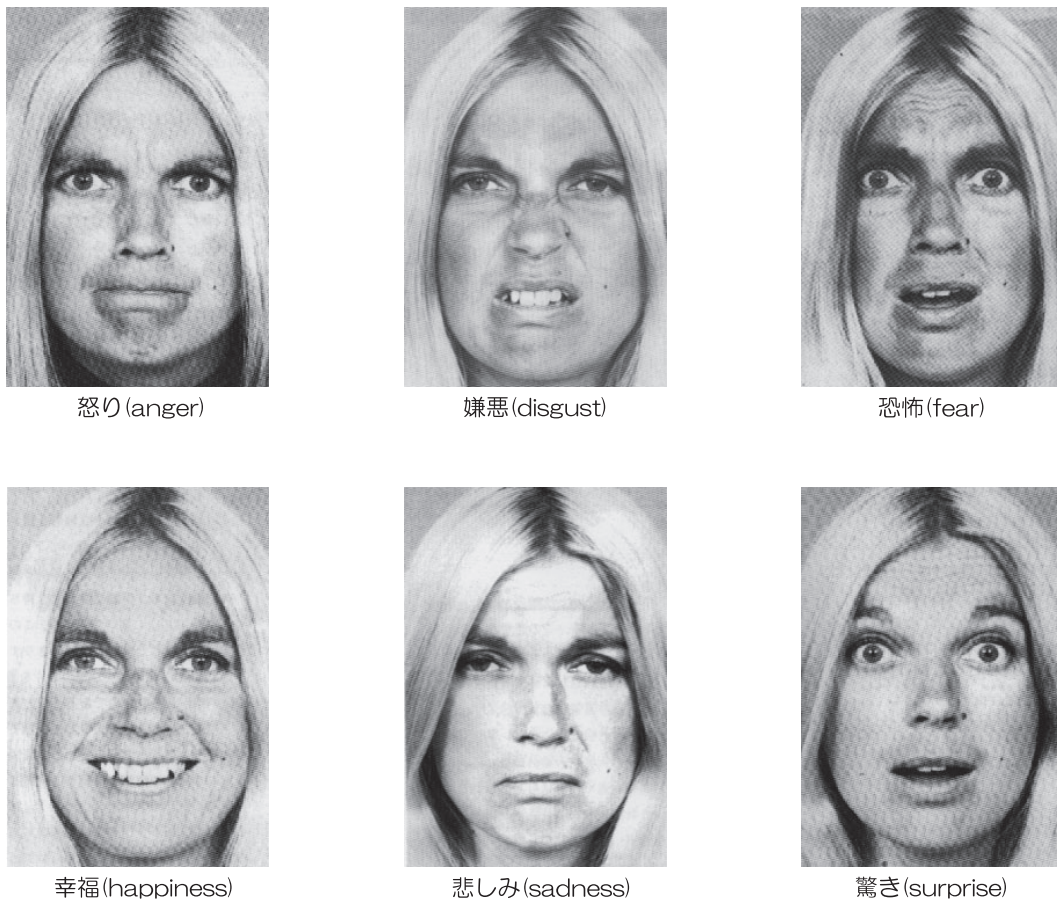


図 3.7: 基本 6 表情

表 3.1: 顔の上部において特徴的な表情動作

表情	特徴的な動作
驚き	眉が湾曲し全体が引き上げられる。そのため、眉の下の皮膚が引っ張られて目立つようになる。額に水平の皺が現れる場合もある。
恐怖	眉はやはり全体が引き上げられ額に水平の皺ができる。また、眉の内側が引き寄せられる。
嫌悪	眉が下がる。
怒り	眉が下がり、内側に引き寄せられ眉間に皺ができる。
幸福	特徴的な眉はない。
悲しみ	眉の内側の両端が上がり引き寄せられる。

表 3.2: 顔の中央部において特徴的な表情動作

表情	特徴的な動作
驚き	目は大きく見開かれる。上瞼は引き上げられ、下瞼は弛緩する。白目が虹彩の上にも現れるのが特徴的である。
恐怖	目は大きく見開かれて緊張する。上瞼は引き上げられるが、下瞼が緊張するのが特徴的である。
嫌悪	目の開き具合が狭くなり、目の下に窪みや皺が現れる。
怒り	上瞼が下がり、下瞼が緊張する (1)。それに加えて、下瞼がやや持ち上がる場合もある (2)。
幸福	目の下や目尻に皺が現れる。
悲しみ	上瞼の両外側に三角形のような領域が現れ、下瞼はやや持ち上がる。また、視線が下を向くことが多い。

表 3.3: 顔の下部において特徴的な表情動作

表情	特徴的な動作
驚き	顎は下方に下がり口が開く。唇と歯は引き離され、開いた口には緊張がないのが特徴である。
恐怖	口が開く場合 (1) は、唇が緊張している。唇が押し広げられ、口が閉じたようになる場合 (2) は、唇は角が後方に引かれて緊張する。この唇の広がりが大きいほど、恐怖も強いものとなる。
嫌悪	上唇を持ち上がり、鼻の付け根に皺ができる。下唇を持ち上げられ前に出される場合 (1)、下唇が下げられて口が開き、突き出される場合 (2)、両唇が押しつけられ、唇の一方の端だけが隆起する場合 (3) の3パターンある。(3) は軽蔑の表情である。
怒り	口が閉じる場合 (1) と開く場合 (2) がある。口が閉じる場合には、上下の唇が強く押しつけられる。口が開く場合は、なにか叫んだりしているときのものである。
幸福	強さに応じて3つのパターンがある。いずれも唇の両端が後方に引かれ持ち上がり、鼻から口の両端へ特徴的な皺ができる。唇の形が幸福の強さを示し、(1) 唇がくっついている場合、(2) 唇は離れるが歯は離れない場合、(3) 口が開き歯が離れる場合の順に幸福が強くなる。
悲しみ	唇の両端が下がる。また、唇が弛緩するのは泣き出す前に特徴的な表情である。

を引っ張り上げる動作を表す AU12 を組み合わせると、「幸福」の表情となる。前項で述べた基本 6 表情を表すには、表 3.4 及び表 3.5 に示した AU が全て必要となるわけではなく、15 個程度の AU で十分であると言われている。基本 6 表情を AU の組み合わせで表したものを表 3.6 にまとめる。

ただし、FACS は本来、人間が表情を記述し分析することを目的として提案されたものであり、AU には定性的な表現が用いられている。また、AU には AU1 「眉の内側を上げる」のようにある動作を行っている状態を表す静的なもの、AU46 の「まばたく」のように動きを表すものが混在しており、計算機上で処理を行う場合にはこれらを同様に扱うことはできない。

表 3.4: Action Unit (1)

顔の上部の AU		顔の下部の AU	
No.	動作	No.	動作
1	眉の内側を上げる	9	鼻に皺を寄せる
2	眉の外側を上げる	10	上唇を上げる
4	眉を下げる	11	鼻唇溝を深める
5	上瞼を上げる	12	唇端を引っ張り上げる
6	頬を持ち上げる	13	唇端を鋭く上げて頬を膨らます
7	瞼を緊張させる	14	えくぼを作る
41	瞼を力なく下げる	15	唇端を下げる
42	薄目	16	下唇を下げる
43	瞼を閉じる	17	下顎を上げる
44	細目	18	唇をすぼめる
45	まばたく	20	唇を横に引っ張る
46	ウインクする	22	唇を突き出す
		23	唇を固く閉じる
		24	唇を押さえつける
		25	顎を下げずに唇を開く
		26	顎を下げて唇を開く
		27	口を大きく開く
		28	唇をかむ (吸い込む)

表 3.5: Action Unit (2)

その他の AU		頭と目の位置の AU	
No.	動作	No.	動作
8	唇どうしを接近させる	51	左を向く
19	舌を見せる	52	右を向く
21	首を緊張させる	53	頭を上げる
29	下顎を突き出す	54	頭を下げる
30	下顎を横へずらす	55	左へ傾ける
31	歯を食いしばる	56	右へ傾ける
32	唇を噛む	57	前へ出す
33	息を吹きかける	58	後へ引く
34	頬を息で膨らます	61	左を見る
35	頬を吸い込む	62	右を見る
36	舌で頬や唇を膨らます	63	上を見る
37	舌で唇をなめる	64	下を見る
38	鼻孔を開く	65	斜視
39	鼻孔を狭める	66	内斜視

表 3.6: 基本 6 表情を表す AU の組み合わせ

表情	AU の組み合わせ
驚き	1, 2, 5, 26
恐怖	1, 2, 4, 5, 20, 26
嫌悪	4, 9, 10
怒り	4, 5, 7, 10, 23
幸福	6, 12
悲しみ	1, 4, 15

3.2 感情推定手法

3.2.1 感情推定手法の概略

2.2で述べたように、表情を認識し、感情を推定する過程は次の3段階に分けることができる。

1. 顔画像から顔の変形を検出する。
2. 顔の変形と表情(基本表情)を対応づける。
3. 表情から感情を推定する。

本研究で用いる感情推定手法は、このそれぞれの段階に対応する3つの部分に、視線方向を検出する部分を加えて、以下の4つの部分から成る。

1. 特徴量抽出部
2. 表情認識部
3. 視線方向検出部
4. 感情推定部

感情推定手法の処理の流れを図3.8に示す。まず、特徴量抽出部で顔画像から顔の変形を表す特徴量を抽出する。そして、表情認識部で、その特徴量がどんな表情を表しているかを推定する。一方、視線方向検出部では、顔画像から視線方向を検出し、最後に表情と視線方向の時間的推移から感情を推定する。以下に、各部分について詳しく述べる。

3.2.2 特徴量抽出部

特徴量抽出部では、ビデオカメラやVTRなどによる動的な顔画像から顔の変形を表す特徴量を抽出する。抽出する特徴量としては、次の2点を満足している必要がある。

1. FACSのAUに対応すること
2. 抽出が容易であること

次の段階の表情認識部において、特徴量は基本6表情を認識するために用いられる。そのため、基本6表情を識別できるだけの精度で顔の変形を捉えることができなければならない。そこで、特徴量は3.1.4で述べた表情記述法 FACS の AU に対応させることにする。AU の組み合わせにより様々な表情を記述することができるので、特徴量を AU に対応させることにより、基本6表情の識別ができることはもちろん、既知である表情と AU の組み合わせとを対応づけるルールを表情認識部で用いることができるという利点がある。

また、2. に関して、本研究では表情のヒューマンマシンインタフェースへの利用を目的としているため、リアルタイムでの処理が前提となっている。特徴量の抽出は主に画像処理によって行うことになるが、その際にも時間のかかる複雑な処理ではなく、より高速な手法を選択する必要がある。

試作システムでは、まず眉、目及び口上に設定した18個の特徴点を抽出し、その各特徴点間の距離などから特徴量を算出する方法をとっている。詳細は4.2で述べる。

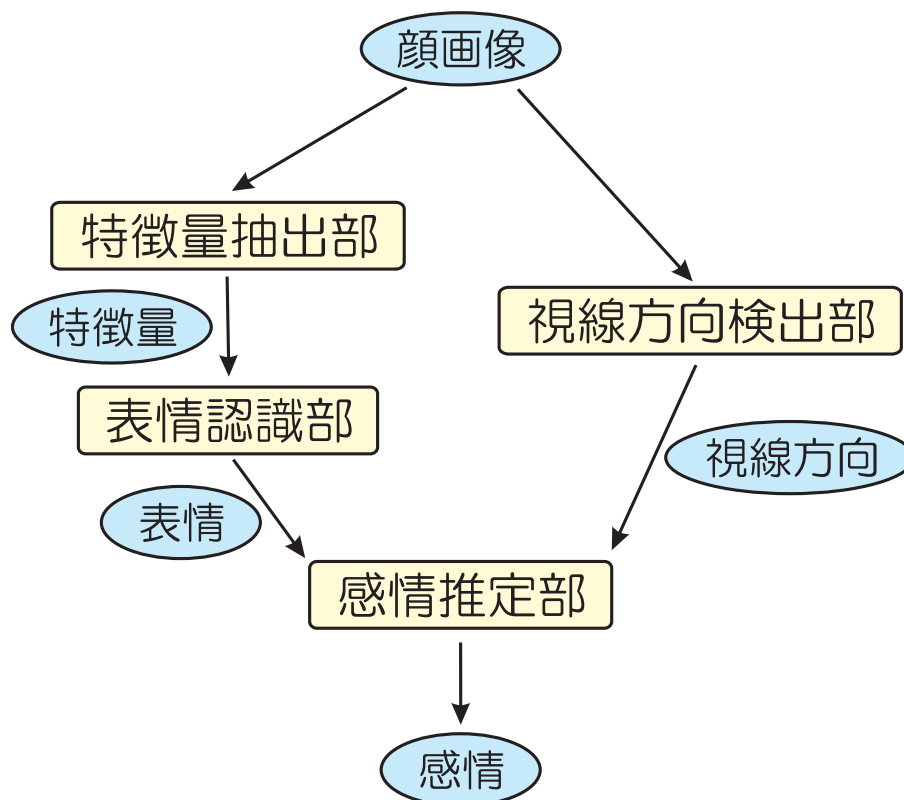


図 3.8: 感情推定手法の概略

3.2.3 表情認識部

表情認識部では、特徴量抽出部で得られた特徴量から、どんな表情を表出しているかを認識する。3.1.2で述べた顔の3領域、すなわち顔の上部、中央部、及び下部はそれぞれ独立した動作が可能であるので、その3領域のそれぞれについて表情の認識を行う。表情の分類としては基本6表情に中立の表情を加えた計7種類を用いる。

具体的な方法としては、ニューラル・ネットワークやパターン・マッチングを利用する方法なども考えられるが、特徴量はAUに対応しているので、3.1.4で例を挙げたようなAUの組み合わせと表情との対応関係から、ルールベースで識別する方法が適していると考えられる。試作システムでは、ルールに基づくファジー推論により認識を行っている。詳細については4.3で述べる。

3.2.4 視線方向検出部

視線方向検出部では、ビデオカメラやVTRなどによる顔画像から視線方向を検出する。マウスに代わるようなポインティング・デバイスとして視線方向を用いる場合には、非常に高い検出精度が求められるが、感情の推定に用いるのであればそれほどの精度は必要ない。AUでも、上下左右方向の区別しかしておらず、上下方向、左右方向とも3段階程度(上中下、左中右)の検出ができれば、十分感情の推定に利用できると考えられる。

また、視線方向を検出するには特徴量抽出部と同様に画像処理を行う必要があるが、リアルタイム処理を行うために時間のかかる処理は避けなければならない。試作システムでは、瞳が黒くて丸いことを利用して視線方向の検出を行っており、4.4で詳細を述べる。

3.2.5 感情推定部

表情認識部では顔の各領域の表情が、また、視線方向検出部では視線方向が、時間とともに変化していく様子が得られる。したがって、感情推定部では次のような情報を利用することができる。

- 表情の種類
- 表情の表出順序・タイミング
- 表情の表出時間

- 表情の表出頻度
- 視線方向の変化

以下では、上記の各情報からどのような感情が推定できるかを述べる。

表情の種類

表情認識部では、顔の3領域のそれぞれについて表情の認識を行う。同時刻におけるその3領域の表情の組み合わせは、複雑な混合感情を表すことができると考えられる。

3つの領域すべてが同じ表情を表しているのであれば、組み合わせた表情はもちろんその表情に対応する感情を表す。異なる表情同士で組み合わせると、3.1.2の例のように混合感情を表す。同様な例を表3.7にまとめる^[19]。なお、表中の下部の表情の部分に示されている数字は、表3.3の数字と対応している。

表情の順序・タイミング

短時間に複数の表情が表出されたときの感情は、それぞれの表情に対応する別々の感情というよりも、それらが混合した感情であると考えられる。例えば、驚きの表情のすぐあとに幸福の表情が表出された場合、それは幸福を伴った驚きの感情である。また、以前に強い感情が生起していたならば、それ以後の感情にもその強い感情が少なからず影響する場合がある。

表情の表出されるタイミングも、感情を推定するための手掛かりとなり得る。本来、最も重要なのは発話や身体動作と表情とのタイミングであるが、本研究ではそれらは対象としていないため、顔の3領域の表情同士のタイミング、及び表情と視線方向とのタイミングのみを扱うことになる。部分的な表情変化のタイミングが違えば、異なった感情を表す表情となる例として、「幸福」の表情について次のような実験結果が報告されている^[21]。

- 口が目よりも先に動く ⇒ 「快の笑い」と認識される
- 目が口よりも先に動く ⇒ 「不快の笑い」と認識される

表情の表出時間

表情の表出時間は、次の3つの時間によって規定される^[19]。

- 開始時間：表情が顔に現われるのにどれくらいの時間がかかるのか

表 3.7: 顔の 3 領域の組み合わせによって表される感情

上部	中央部	下部	意味
驚き	中立	中立	不信・疑念、感嘆(頭の動きあり)、句読法(短時間)
驚き	中立	嫌悪	疑念に基づく不信感、疑いを内包した感嘆(頭の動きあり)
中立	驚き	中立	関心、句読法(瞬時)
中立	中立	驚き	呆気、哑然
驚き	驚き	中立	疑問を示す驚き
中立	驚き	驚き	仰天
驚き	中立	驚き	茫然とした驚き、感興の沸かない驚き
恐怖	中立	中立	心配、懸念、抑制された恐怖
中立	恐怖	中立	わずかな恐怖、抑制された恐怖
中立	中立	恐怖(1)	心配、懸念
中立	中立	恐怖(2)	抑制された恐怖、恐怖に言及(短時間)
恐怖	恐怖	中立	強い懸念
中立	恐怖	恐怖(1)	戦慄、ショック
中立	恐怖	恐怖(2)	身の毛もよだつ恐怖感、冷淡な感じの恐怖
驚き	恐怖	恐怖(1)	恐怖を伴う驚き
驚き	恐怖	恐怖(2)	恐怖を伴う驚き
驚き	驚き	恐怖(1)	恐怖を伴う驚き
驚き	驚き	恐怖(2)	恐怖を伴う驚き
恐怖	恐怖	驚き	怯え、恐怖を伴う仰天、懐疑的恐怖
中立	中立	嫌悪(1)(2)	見せかけの嫌悪、句読法(短時間)
中立	中立	嫌悪(3)	軽蔑
驚き	驚き	嫌悪(1)(2)	嫌悪を伴う驚き、嫌悪の強調
驚き	嫌悪	嫌悪(1)(2)(3)	不信感
恐怖	恐怖	嫌悪(1)(2)	対象に畏怖を感じる恐怖
怒り	中立	中立	抑えた怒り、苛立ち、真剣、集中、句読法(短時間)
怒り	怒り	中立	抑えた怒り、苛立ち、真剣、集中、句読法(短時間)
中立	中立	怒り(1)	かすかな怒り、抑えた怒り、肉体的努力、集中
怒り	怒り	嫌悪(1)(2)(3)	嫌悪を伴う怒り
怒り	怒り	驚き	当惑した驚き、ショック
怒り	怒り	恐怖	途方に暮れた恐れ、集中している時の恐れ
驚き	驚き	幸福	嬉しさを伴う驚き、思いがけない喜び、驚きから幸福への移行中(短時間)
中立	幸福	嫌悪(3)	冷笑、傲慢
怒り	怒り	幸福	嬉しさを伴う怒り
恐怖	恐怖	幸福	嬉しさを伴う恐怖
悲しみ	中立	中立	弱い悲しみ、抑えた悲しみ
悲しみ	悲しみ	中立	中程度の悲しみ
中立	中立	悲しみ	軽い苦痛、反抗
悲しみ	悲しみ	恐怖(1)(2)	恐怖、不安を伴う悲しみ
怒り	怒り	悲しみ	怒りを伴う悲しみ
悲しみ	悲しみ	怒り	怒りで悲しみを隠す
悲しみ	悲しみ	嫌悪	嫌悪を伴う悲しみ
悲しみ	嫌悪	嫌悪	嫌悪を伴う悲しみ
悲しみ	悲しみ	幸福	幸福で悲しみを隠す

- 持続時間：表情が現われてから消滅し始める、あるいは別の表情に変化し始めるまでに、その表情はどのくらい長く顔に現われているか
- 消滅時間：表情が消えるのにどのくらいの時間がかかるのか

表情の開始時間、持続時間、消滅時間が異なれば、その感情の強度なども変化する。表情の持続時間に関する一般的な規則はないが、ごくありふれた表情では数秒ほどであることがほとんどである。5秒や10秒も続く場合はその感情はかなり激しいものである。なかには、微表情のように1秒にも満たないものもある。微表情はある感情が表出されたが、それを隠そうと抑制したために、ごく短時間表出されただけですぐに消滅するか、別の表情に変化した場合の表情である。微表情の持続時間はせいぜい0.2秒程度であり、人間同士のコミュニケーションでも見落とされる場合が多いが、偽りの表情を見破るよい手掛かりとなる。

具体例としては、蒲地らはモーフィングによって作成した変化速度の異なる動画刺激を用いて表情認知実験を行い、表情認知は変化速度に影響を受け、しかもその影響の受け方は表情により異なるという結果を得た^[22]。特に、「驚き」の表情は速い速度で変化するほど強度が強いと判断され、逆に「悲しみ」の表情は速度の変化には影響を受けなかった。

表情の表出頻度

一定時間内にひとつの表情が繰り返し現われた場合、その表情に対応する感情が生じやすい状態であると言える。すなわち、その人はある「気分」の状態にある。例えば、頻繁に微笑みを浮かべていたとすると、それは幸福な気分であると考えられる。

視線方向の変化

表情が感情の質と強さの両方を表すのに対し、視線方向は主に感情の強さを表す。それは、視線方向の対象への固定、すなわち注視に最もよくあらわれ、注視時間が長いほど感情が強い。また、悲しみの表情において、視線を下に向け伏し目がちにすると、悲しみの表情が一層強まる。

まとめ

以上のように、表情及び視線方向の時間的変化を利用することにより、基本6感情だけでなく、より複雑で微妙な感情を推定できると考えられる。表情がどのように変化

すればどのような感情であるのか、といった対応関係については、現在のところ、断片的には知られているものの、体系的な研究はなされていないようである。したがって、その対応関係を調査する必要があると考えられるが、本研究ではまず、動的な表情を感情推定に利用することが可能なシステムを構築することに焦点をおいており、表情と感情との関連についてまとめることは今後の課題とする。

第 4 章 リアルタイム感情推定システムの試作

本章では、第 3 章で述べた感情推定手法に基づいて試作した、リアルタイム感情推定システムについて述べる。

4.1 システム構成

4.1.1 システム構築上の着眼点

本研究で試作したシステムは動的な表情の認識を行うため、ハードウェア及びソフトウェアともに、静的な表情認識システムにはない制約が生じる。

まず、入力画像として動的な顔画像が必要となる。そこで、ビデオカメラで撮影した顔画像、もしくはビデオテープに録画された顔画像を入力画像として用いることにした。また、入力画像としてはカラー画像を用いる。従来は、システムや認識アルゴリズムが複雑になるといった理由から、表情認識には濃淡画像が用いられることが多く、カラー画像はあまり使用されていなかった。しかし、現在ではカラー画像を扱えるシステムが汎用化しており、色情報を用いると認識の際の特徴抽出が容易になると考えられるので、本研究では入力画像としてカラー画像を用いて、積極的に色情報を利用することにした。

また、ヒューマンマシンインタフェースへの適用を前提としているため、認識に要する時間をできるだけ短縮し、リアルタイムで動作することが必要である。そのため、高速に画像を取り込み加工することができるハードウェアとソフトウェアが必要となる。認識手法に関しても、複雑で時間のかかる処理は避け、なるべく高速な手法を採用しなければならない。画像処理においてもそれほど高い精度を必要としない場合には縮小画像を用いるなど、なるべく小さな画像を扱うことにより処理の高速化をはかる必要がある。しかし一方で、十分な認識精度も確保しなければならない。

以上のことをふまえて試作したシステムのハードウェア構成、及びソフトウェア構成を述べる。

4.1.2 ハードウェア構成

試作システムは以下のハードウェアにより構成されている。

- グラフィック・ワークステーション O2 R10000/195MHz
(シリコングラフィクス製、以下、O2)
- グラフィック・ワークステーション OCTANE R10000/250MHz
(シリコングラフィクス製、以下、OCTANE)
- デジタルビデオカメラ DCR-VX1000 (ソニー製、以下、VC)
- デジタルビデオカセットレコーダー WV-D9000(ソニー製、以下、VTR)
- ディスプレイ(シリコングラフィクス製)

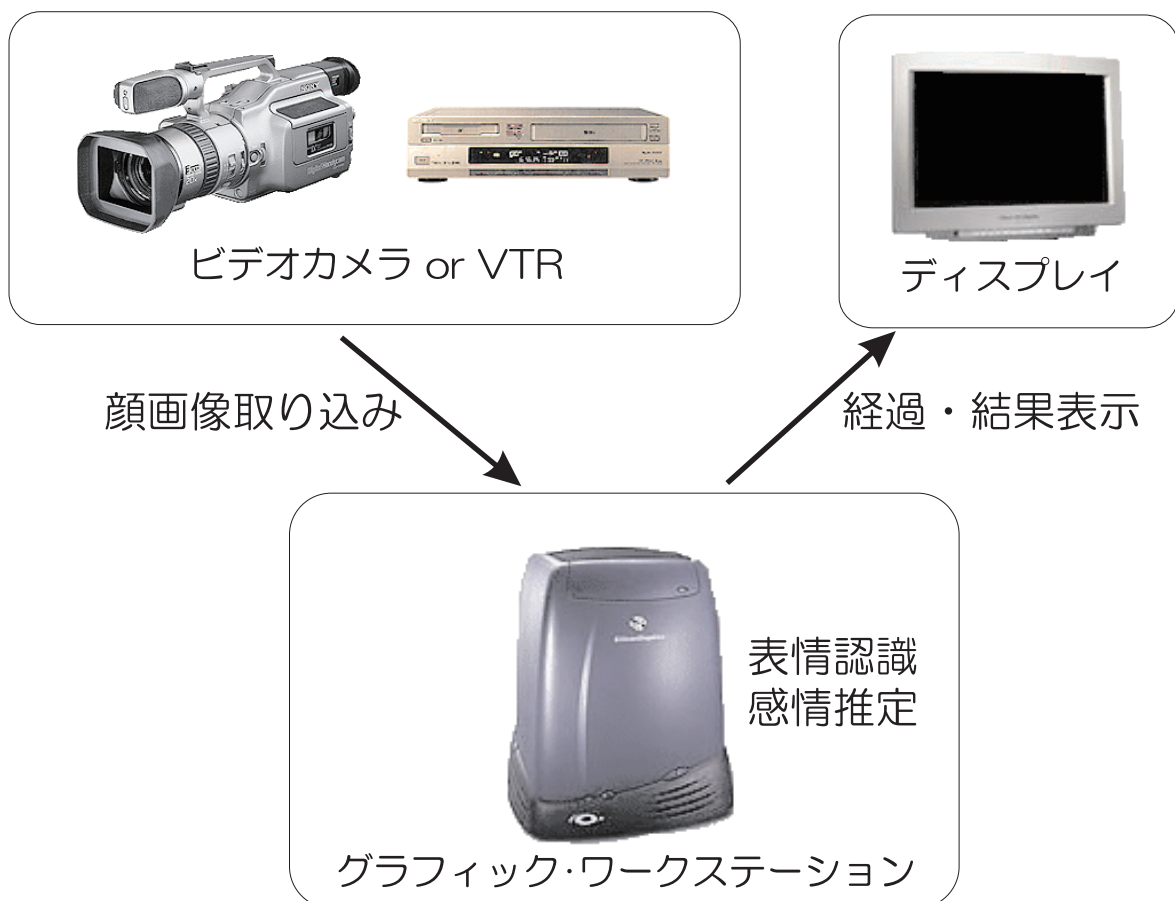


図 4.1: ハードウェア構成

ハードウェア構成を図4.1に示す。VCまたはVTRから出力される顔画像はNTSC方式のアナログ・ビデオ信号であり、O2に内蔵されているビデオシステムMVP(Multiport Video Processor for the O2 system)に入力される。MVPではその顔画像をデジタル信号に変換しビデオメモリに蓄えることによりO2で処理できるようにする。そして、O2で表情認識及び感情推定の処理を行い、その処理過程と結果をディスプレイに出力する。OCTANEはビデオ入力のためのハードウェアを備えていないが、O2に比べ処理速度が速いため、後述する評価実験での処理時間の計測に用いる。

4.1.3 ソフトウェア構成

試作システムは以下のソフトウェアにより構成されている。

- SGI ビデオライブラリ (以下、VL)：ビデオデバイス用 C 言語ライブラリ
- Open GL：画像表示用 C 言語ライブラリ
- 本研究で作成したリアルタイム感情推定プログラム

ソフトウェア構成を図4.2に示す。VCまたはVTRからMVPに入力された画像を、VLを用いてR、G、B各8ビットのRGBデータに変換し、メモリに格納する。そのRGBデータに対して本研究で作成した感情推定プログラムで表情認識や感情推定の処理を行い、Open GLを用いてその過程と結果をディスプレイに表示する。

感情推定プログラムは、3.2で述べた感情推定手法の各部に対応した、次の4つのサブシステムからなる。

1. 特徴量抽出サブシステム
2. 表情認識サブシステム
3. 視線方向検出サブシステム
4. 感情推定サブシステム

感情推定処理の大きな流れを図4.3に示す。特徴量抽出サブシステムでは顔画像のRGBデータから特徴量を抽出し、表情認識サブシステムでその特徴量から表情の認識を行う。一方、視線方向検出サブシステムでは、特徴量抽出サブシステムの処理途中で得られる目領域の画像から視線方向を検出し、最後に、感情推定サブシステムで、表情と視線方向の時間的推移から感情を推定する。以下の各節では、これら4つのサブシステムについて、それぞれ詳しく述べる。

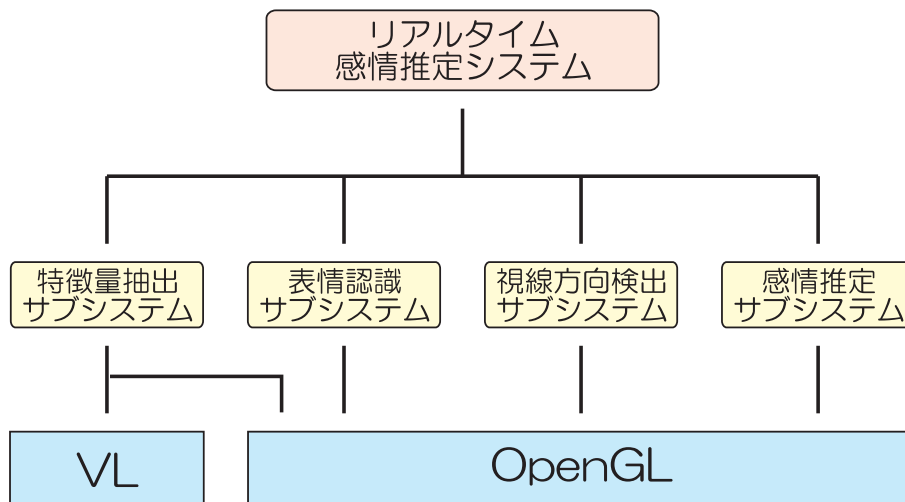


図 4.2: ソフトウェア構成

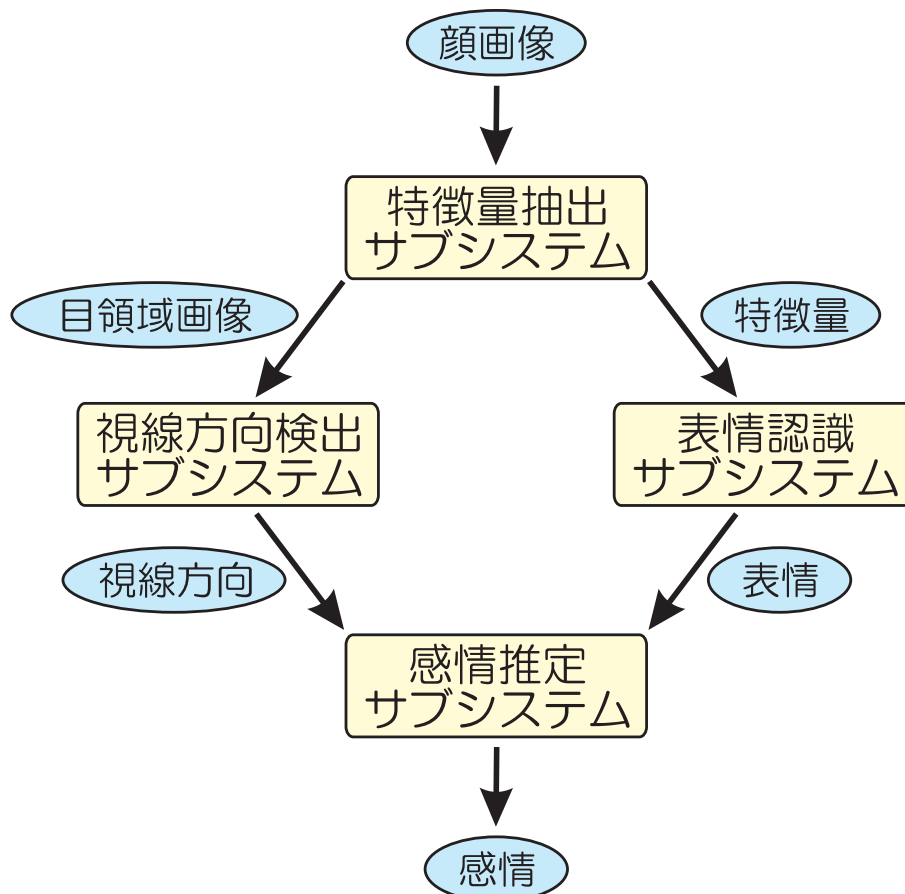


図 4.3: 感情推定処理の概略

4.2 特徴量抽出サブシステム

特徴量抽出サブシステムでは、顔画像から顔の変形を表す特徴量を抽出する。その際、まず顔器官(眉、目、口)上に設定した特徴点を抽出し、特徴点間の距離などから特徴量を算出する方法をとっている。特徴量抽出サブシステムで行う処理の流れを図4.4に示す。図のように、特徴量抽出サブシステムで行う処理は次の4つに分けることができる。

1. 顔領域の抽出
2. 顔器官領域の抽出
3. 特徴点の抽出
4. 特徴量の算出

まず、顔画像から顔を含む矩形領域(顔領域)を抽出し、次に顔領域から眉、目及び口を含む矩形領域(顔器官領域)をそれぞれ抽出する。つづいて各顔器官領域から特徴点を抽出し、最後に特徴点の座標から特徴量を算出する。

また、抽出する特徴量に関しては、3.2.2で述べた条件を考慮して、図4.5に示すような16個の特徴量を抽出することにした。特徴量を算出するための特徴点としては右眉・左眉について各3点、右目・左目・口について各4点の計18点を選んだ。

以下に、上記の各処理について詳しく述べる。なお、画像中の点の座標を用いることがあるが、その際には図4.6に示すような、入力画像の左上隅を原点とするxy座標系を用いる。

4.2.1 顔領域の抽出

ここでは、入力画像から顔の位置とその大きさを検出し、入力画像から顔を含む矩形領域を切り出す。大まかな流れとしては、まず、前のフレームの顔領域の抽出結果から顔領域を探索する領域を決定し、次にその探索領域から肌色領域を抽出する。そして、肌色領域のうち最も顔らしいものを顔領域として抽出する。以下で、抽出方法について具体的に述べる。また、図4.7に抽出方法を示す。

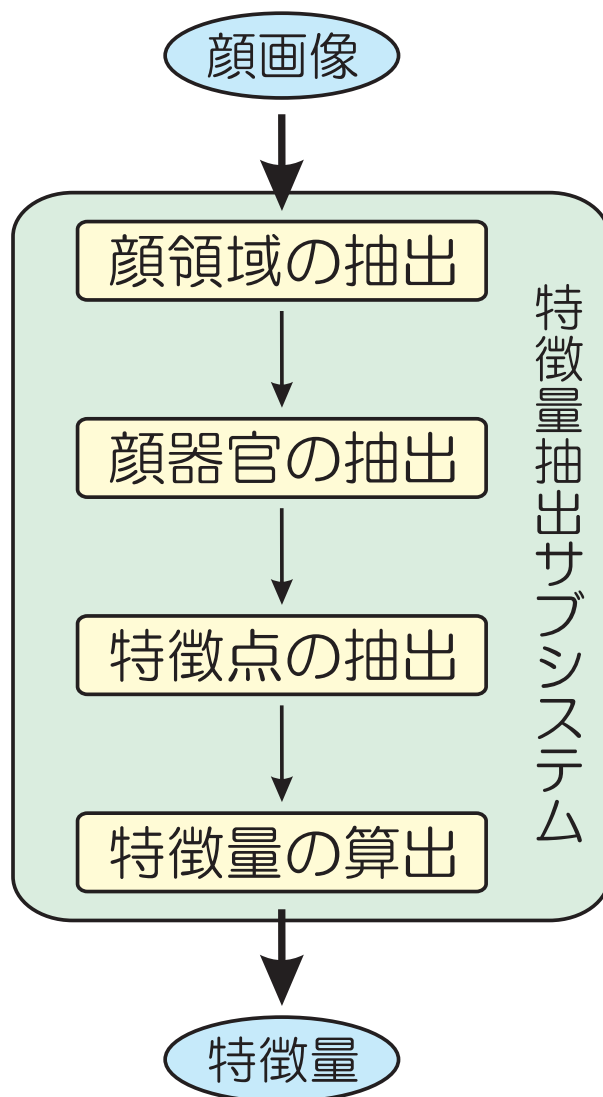


図 4.4: 特徴量抽出サブシステムの概略

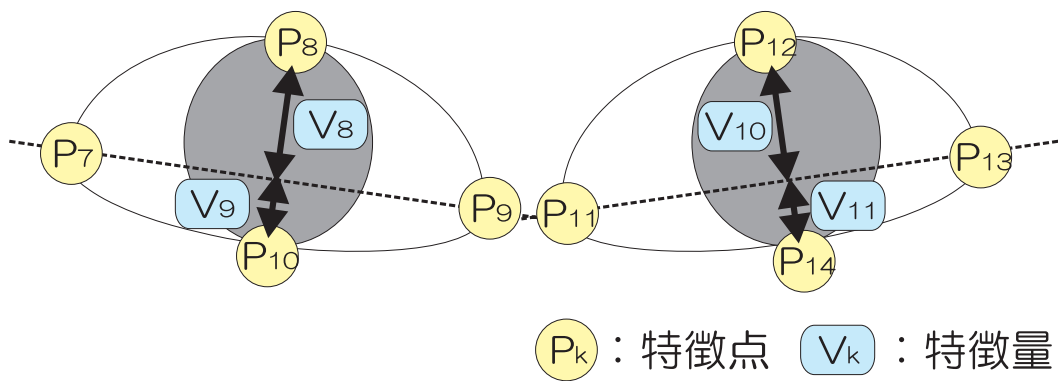
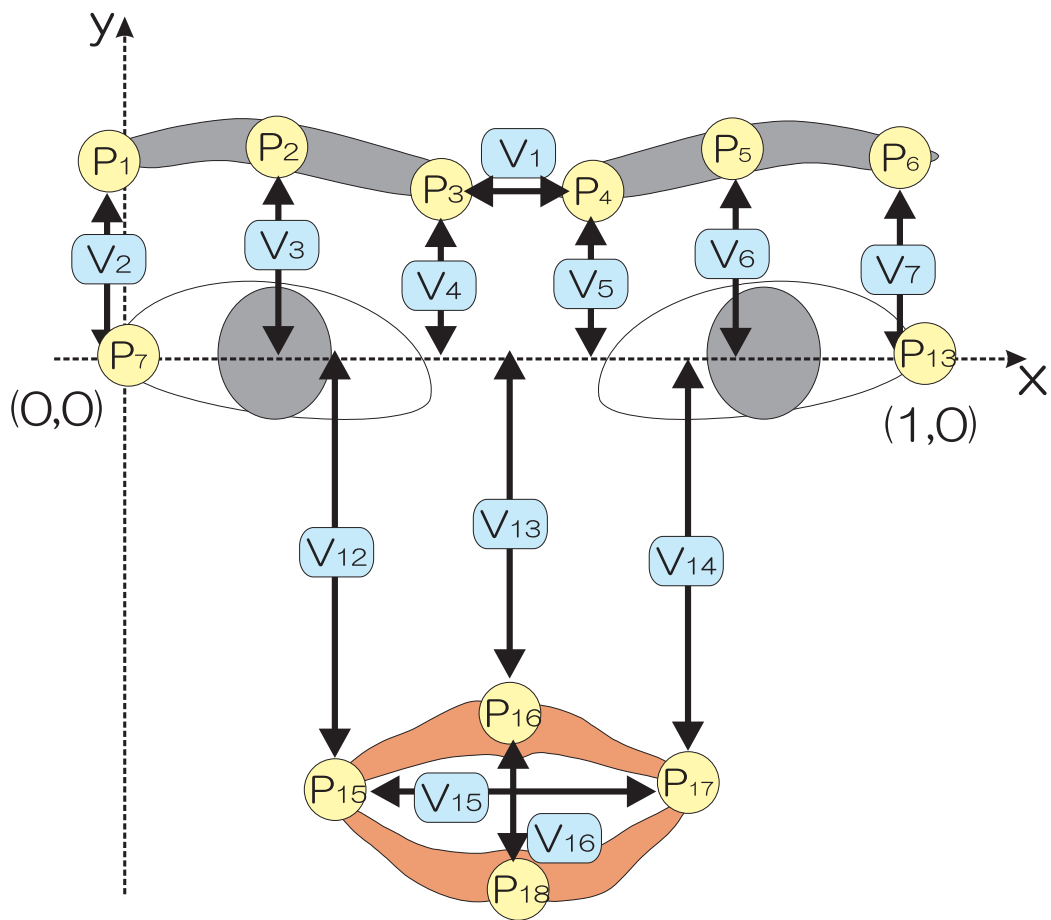


図 4.5: 抽出する特徴点と特徴量

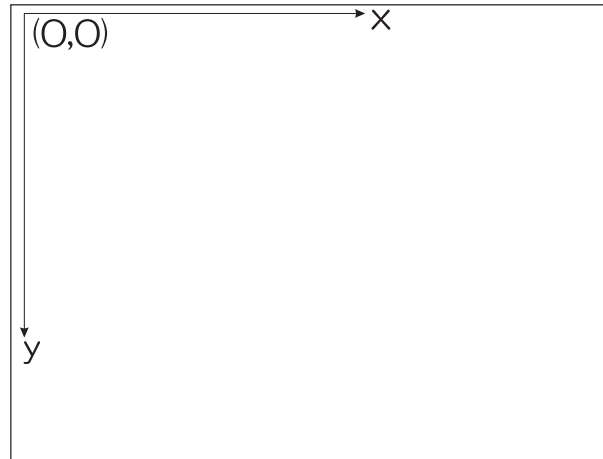


図 4.6: 入力画像の座標系

顔探索領域の決定

顔の探索領域の決定には、1つ前のフレームの顔領域の抽出結果、すなわち顔領域の位置と大きさを用いる。実際の顔の幅を w_r [cm]、最大移動速度を v_r [cm/sec] とすると、画像中での顔の最大移動量 d [pixel] は、前フレームの顔の幅 w_p [pixel]、及び前フレームからの経過時間 t [sec] により、次式のように求められる。

$$d = \frac{w_p}{w_r} v_r t \quad [\text{pixel}] \quad (4.1)$$

したがって、前フレームの顔領域の左上隅の座標が (x_p, y_p) 、大きさが幅 w_p 、高さ h_p であった場合、探索領域は左上隅の座標 $(x_p - d, y_p - d)$ 、幅 $w_p + 2d$ 、高さ $h_p + 2d$ の矩形領域となる (図 4.8)。試作システムでは顔の幅を 15cm、最大移動速度を 50cm/sec と仮定して計算を行っている。また、 t については各フレームが取り込まれる際の時刻を記録しておき、前フレームと現在のフレームの時刻の差をとっている。なお、前フレームで顔領域の抽出に失敗している場合には入力画像全体を探索領域とする。

このように探索領域を限定することにより、一度顔領域の抽出に成功すれば、背景に別の肌色領域が現われた場合にも、顔領域を追跡しつづけることができる。また、顔領域の抽出処理に用いる画像の画素数を削減できるので、高速化にもつながる。

肌色領域の抽出

次に、4.2.1の方法で決定した探索領域の画像から、次のように肌色領域を抽出する。

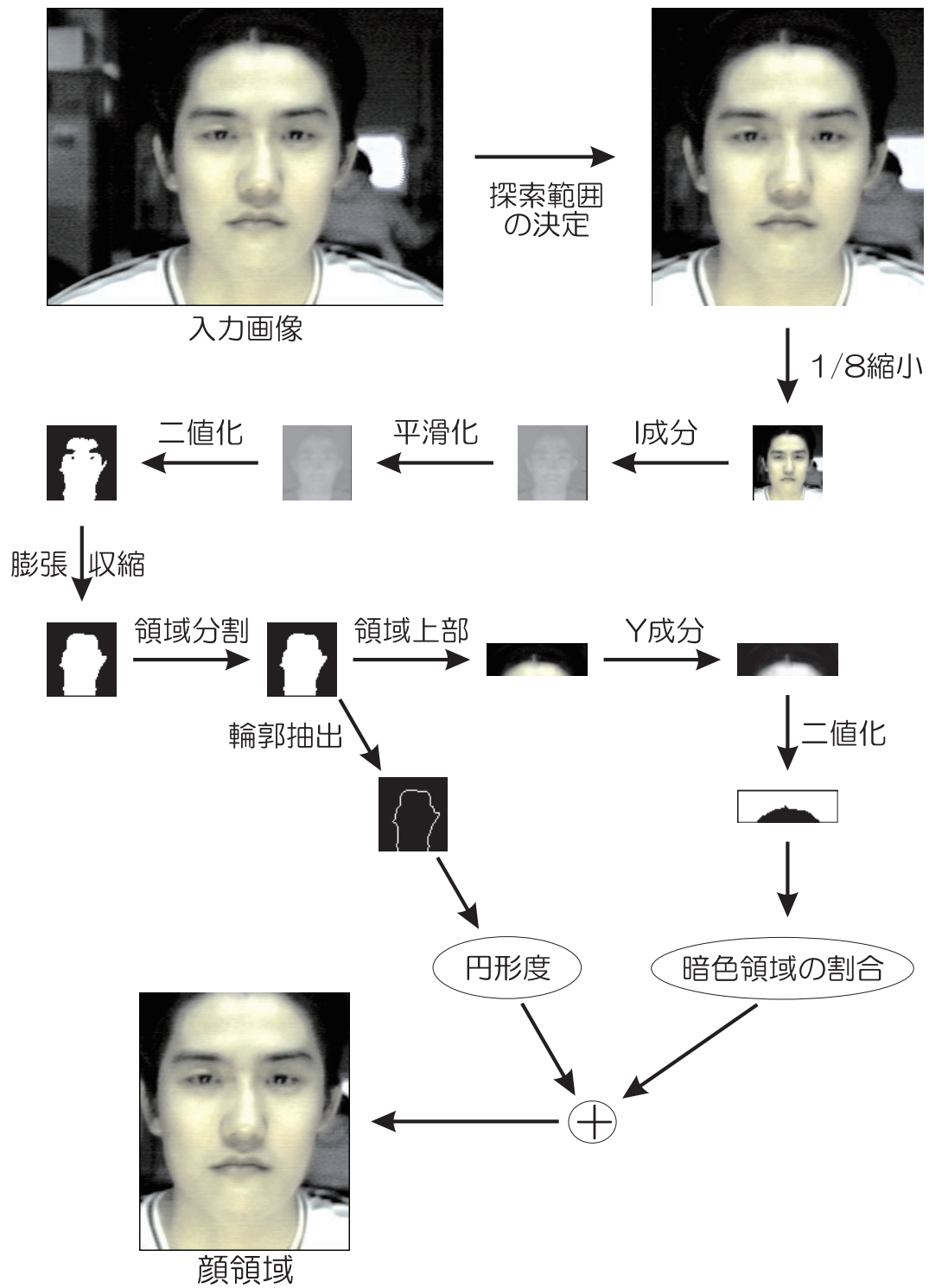


図 4.7: 顔領域の抽出方法

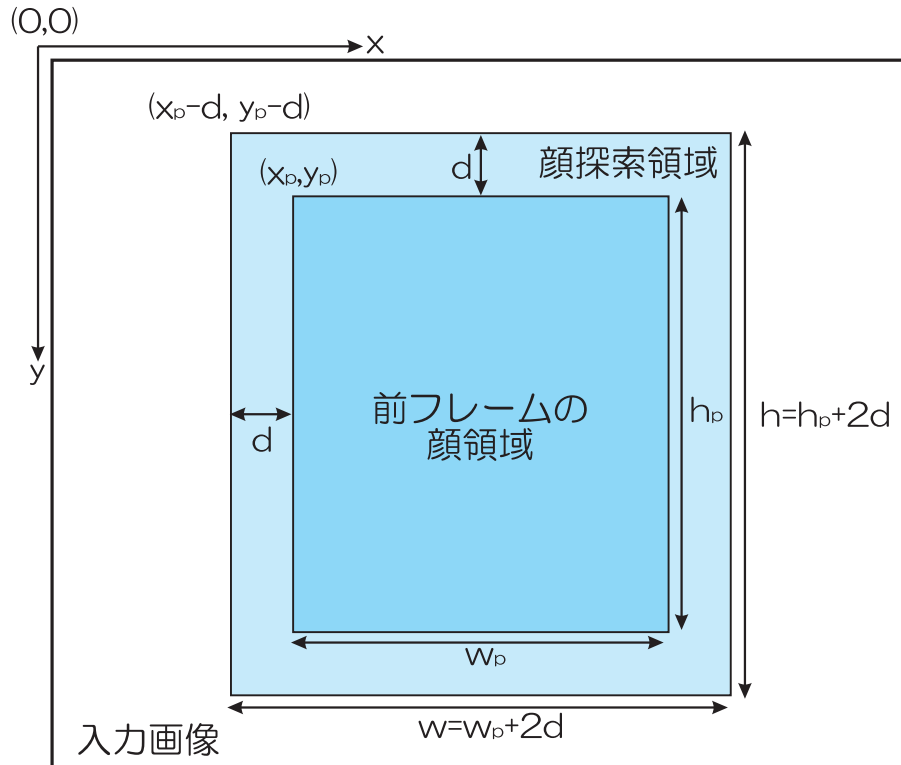


図 4.8: 顔領域の探索領域

1. 探索領域の画像を 1/8 に縮小する。これにより、処理を大幅に高速化できるとともに、雑音成分を除去することができる。
2. 1/8 に縮小した画像から、YIQ 表色系の I 成分を算出する。I 成分は肌色領域で大きい値をとることが知られており^[6]、RGB 値から次式により算出できる。

$$I = 0.596R - 0.275G - 0.321B \quad (4.2)$$

3. I 成分画像に平滑化を行い、雑音成分をさらに削減する。平滑化には 3×3 メディアンフィルタを用いる。 3×3 メディアンフィルタは各画素を中心とする 3×3 画素の領域の中央値をその画素値とするものである。
4. 3. の画像を二値化し、I 成分の大きい部分を抽出する。その際の閾値は現在のところ、顔画像の撮影環境に合うように手動で行っている。
5. 二値化した画像に膨張・収縮処理を施すことにより、微小な領域や細かな凹凸を除去する。収縮→膨張→膨張→収縮の順に処理を行う。

6. 5. の画像に対して領域分割を行う。ここで、領域分割とは互いに連結している画素集団ごとに分割することである。

以上の処理により、顔領域の候補となる肌色領域が抽出できる。

顔領域の選択

最後に、顔領域の候補の中から最も顔らしいと思われるものを顔領域として選択する。選択の基準としては次の2つを用いる。

- 円形度 (c)

円形度は領域の周囲長 l と面積 S から、

$$c = \frac{S}{l^2} \quad (4.3)$$

により求められ、円で最大値 $1/4\pi$ をとる。顔領域は楕円形に近いので、円形度は比較的大きい値をとる。

- 領域の上部で暗色領域の占める割合 (d)

顔の上部には頭髪がある場合が多いので、顔領域の上部には暗色領域が存在すると考えられる。そこで、顔領域の上部に図 4.9 のような矩形領域を設定し、その領域の全画素数 N と暗色画素数 n から、

$$d = \frac{n}{N} \quad (4.4)$$

により、暗色領域の占める割合を求める。

これらの値を各候補領域について求め、その和、

$$s = \alpha c + \beta d \quad (\alpha, \beta \text{ は定数}) \quad (4.5)$$

をその領域の顔らしさの度合とする。なお、試作システムでは $\alpha = 4\pi$ 、及び $\beta = 0.5$ としている。そして、顔らしさが最大となる領域を顔領域とみなし、入力画像からその領域に対応する部分を含む矩形領域を切り出す。ただし、顔らしさの最大値が一定値に満たない場合は、抽出失敗とみなす。

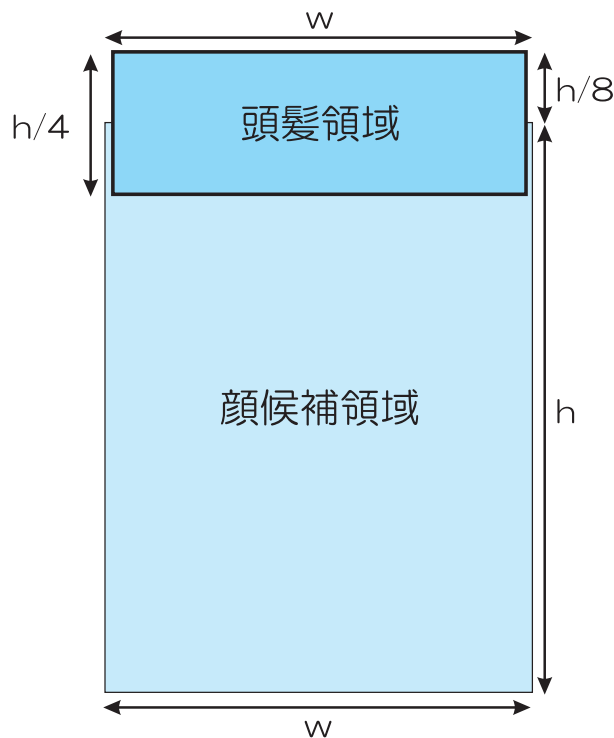


図 4.9: 暗色部分の割合を求める領域

4.2.2 顔器官領域の抽出

ここでは、顔領域の画像から右眉、左眉、右目、左目及び口の5つの顔器官を検出し、各顔器官を含む矩形領域を入力画像から切り出す。まず、抽出が最も容易な口領域の抽出を行い、口領域の位置を参考にして眉及び目領域の抽出を行う。顔領域の抽出と同様に、ここでも高速化のためにまず顔領域画像を1/2に縮小し、その縮小画像に対して処理を行う。

口領域の抽出

口領域の抽出は肌と唇の色の違いを利用して行う。具体的な方法を以下に示す。また、図 4.10に抽出方法を図示する。

1. 口は顔領域の下方の中央部にあるので、顔領域画像の1/2縮小画像から図 4.11の部分を用いて口の探索領域として切り出す。
2. 背景部分を除去するためのマスク画像を次の手順で作成する。
 - (a) 1. で切り出した画像からI成分を算出する。

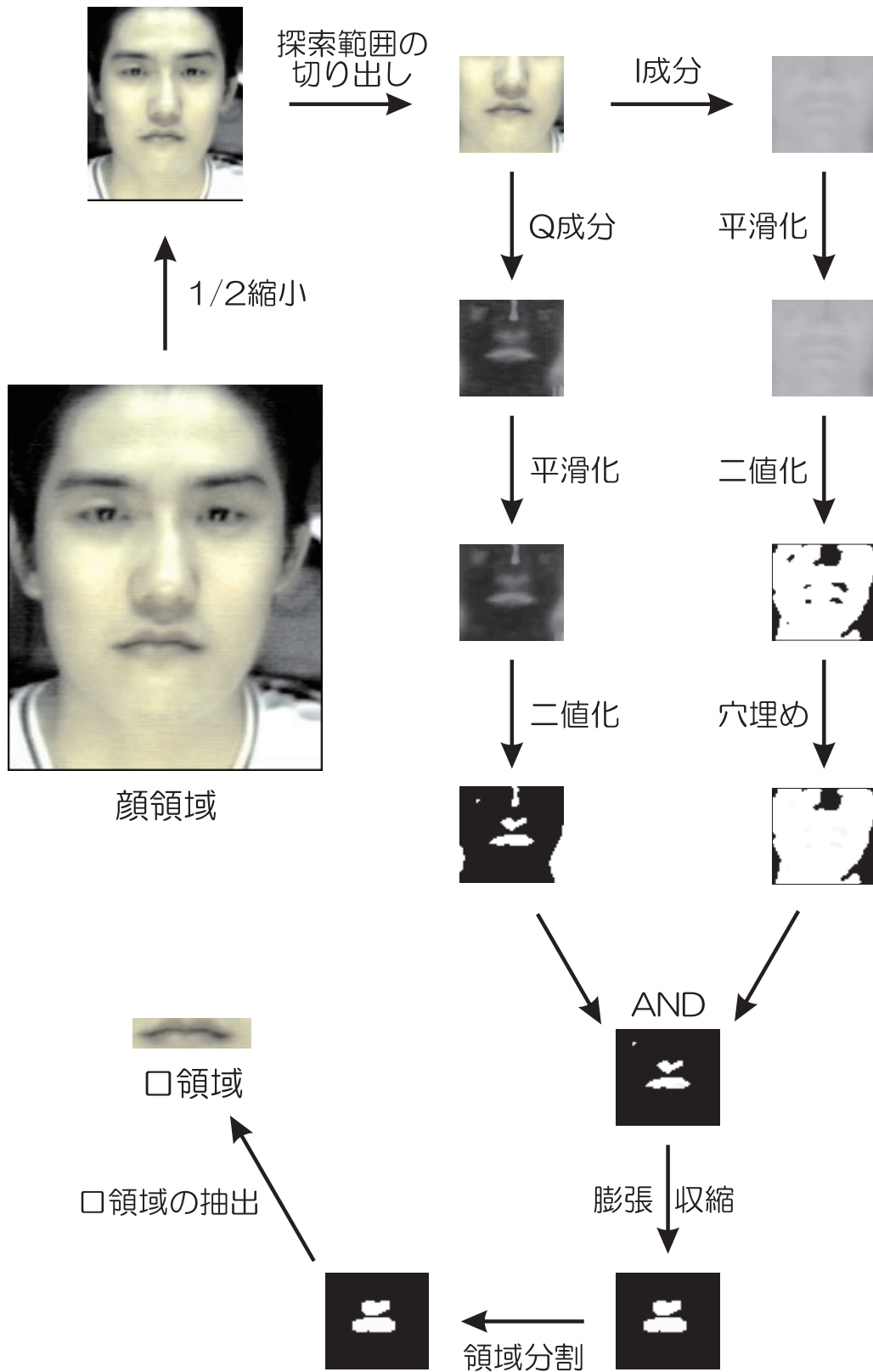


図 4.10: 口領域の抽出方法

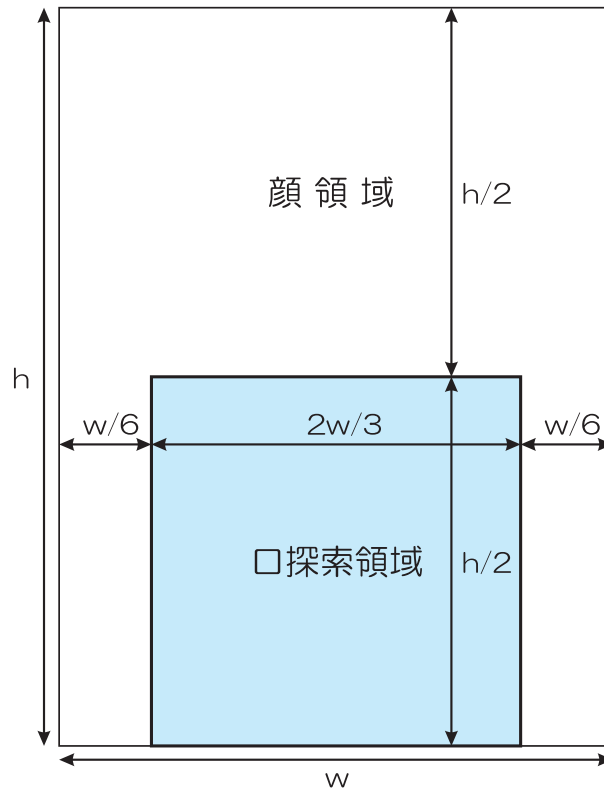


図 4.11: 口領域の探索範囲

- (b) I成分の画像に 3×3 メディアンフィルタによる平滑化を行い、雑音成分を除去する。
 - (c) 平滑化した画像を二値化することにより、I成分の大きい部分 (肌色部分) だけを取り出す。
 - (d) 領域分割を行う。
 - (e) 最大の面積を持つ領域だけを残し、他の領域を除去する。
 - (f) 残った領域の内部に穴が開いている場合があるので、その穴を埋める。
3. 1. の画像から YIQ 表色系の Q 成分を算出する。Q 成分は唇領域で大きい値をとり、RGB の各値から次の式により求めることができる。

$$Q = 0.212R - 0.523G + 0.311B \quad (4.6)$$

4. Q 成分画像を 3×3 メディアンフィルタにより平滑化し、雑音成分を除去する。

5. 4. の画像を二値化し Q 成分の値が大きい部分を取り出す。これにより、唇の部分とそれ以外の唇と色が近い部分が抽出される。
6. 膨張・収縮処理を行い、雑音成分を除去する。
7. 6. の画像と 2. で作ったマスク画像との論理積 (AND) をとり、背景の影響を取り除く。
8. 領域分割を行う。
9. 各領域のうち、面積、高さ、幅が口として不適当なものを除き、残った領域のうちで最大の面積を持つものを選ぶ。
10. 9. で選んだ領域に対応する部分を含む矩形領域を、入力画像から口領域として切り出す。

眉・目領域の抽出

眉及び目領域の抽出方法を図 4.12 に示す。眉及び目領域の抽出は、図 4.13 のような眉・目の端点を見つけることにより行う。具体的な方法は以下の通りである。

1. 眉・目は顔領域の上部にあるので、顔領域画像の 1/2 縮小画像の上半分を切り出す。
2. 1. の画像から YIQ 表色系の Y 成分を算出する。Y 成分は輝度を表し、RGB 成分から次式により求めることができる。

$$Y = 0.299R + 0.587G + 0.114B \quad (4.7)$$

3. Y 成分画像に 3×3 メディアンフィルタをかけて平滑化し、雑音成分を除去する。
4. 3. の画像を二値化することにより、Y 成分の値が小さい部分、すなわち暗色部分を取り出す。
5. 二値化した画像から輪郭線を抽出する。
6. 輪郭線上の点の曲率を計算し、次の条件に当てはまる点を眉または目の端点の候補とする。

- 曲率が極大となっていること

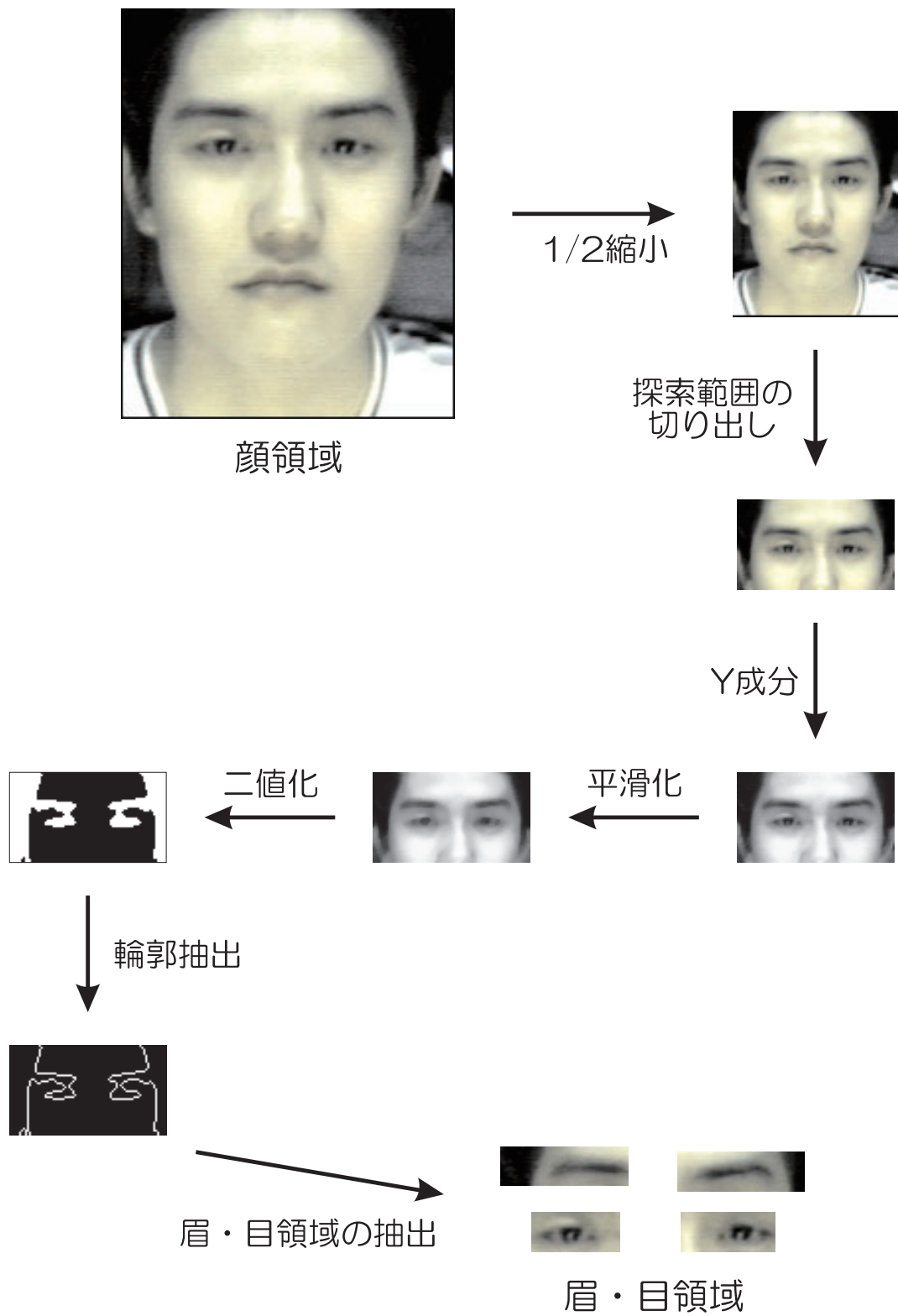


図 4.12: 眉・目領域の抽出方法

- 曲率が一定値以上の大きさであること
- 領域の凹んだ部分ではなく、凸の部分であること (図 4.14)
- 縦方向ではなく横方向に凸であること (図 4.15)

7. 口領域の位置に基づいて、候補点を次のように分類する (図 4.16)。ただし、 x_c を候補点の x 座標、 x_m を口領域の重心の x 座標、 w を顔領域の幅とする。

- $x_c - x_m < -\frac{w}{2} \Rightarrow$ 候補から除外
- $-\frac{w}{2} \leq x_c - x_m < -\frac{w}{4} \Rightarrow$ 右眉の右端 or 右目の右端の候補
- $-\frac{w}{4} \leq x_c - x_m < 0 \Rightarrow$ 右眉の左端 or 右目の左端の候補
- $0 \leq x_c - x_m < \frac{w}{4} \Rightarrow$ 左眉の右端 or 左目の右端の候補
- $\frac{w}{4} \leq x_c - x_m < \frac{w}{2} \Rightarrow$ 左眉の左端 or 左目の左端の候補
- $\frac{w}{2} \leq x_c - x_m \Rightarrow$ 候補から除外

また、右端点の範囲にある左に凸な領域の候補点、及び、左端点の範囲にある右に凸な領域の候補点も候補から外す。

8. 候補点の位置関係から、各候補点がどの端点に対応するかを調べる。このとき、

- 眉の端点同士、及び目の端点同士は y 座標の値が近い
- 目の端点より眉の端点の方が上方にある
- 端点の位置はほぼ左右対称である

などの顔の構造が有する特徴を利用する。

9. すべての端点が揃わなかった場合は、見つかった端点の位置から 8. のルールを使用して求める。例えば、右眉の左端が見つからない場合、左眉の右端が見つかっていれば、それと左右対称な点を右眉の左端とする。

10. 各器官の両端点を含む領域に対応する部分を入力画像から切り出す。

4.2.3 特徴点の抽出

ここでは、各顔器官の画像から図 4.5に示した 18 個の特徴点を抽出する。なお、特徴点は入力画像の左上隅を原点とする座標で表される。ここでは高い精度が要求され

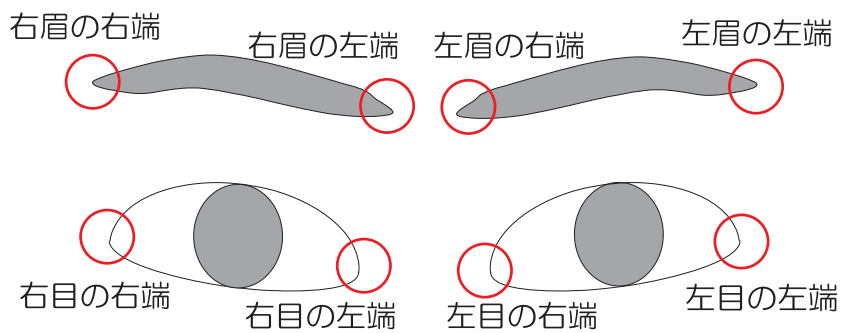


図 4.13: 眉・目の端点

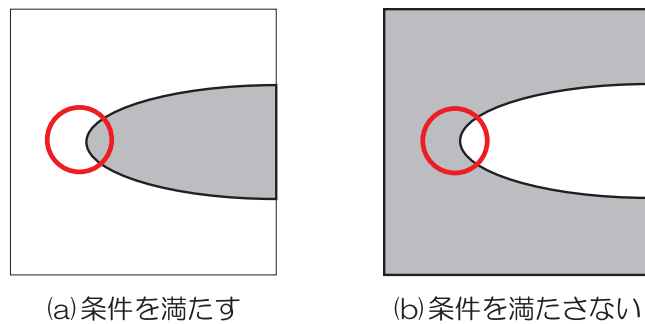


図 4.14: 眉・目の端点の選択条件 (1)

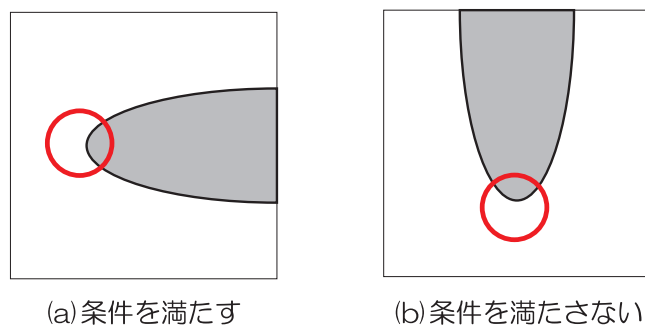


図 4.15: 眉・目の端点の選択条件 (2)

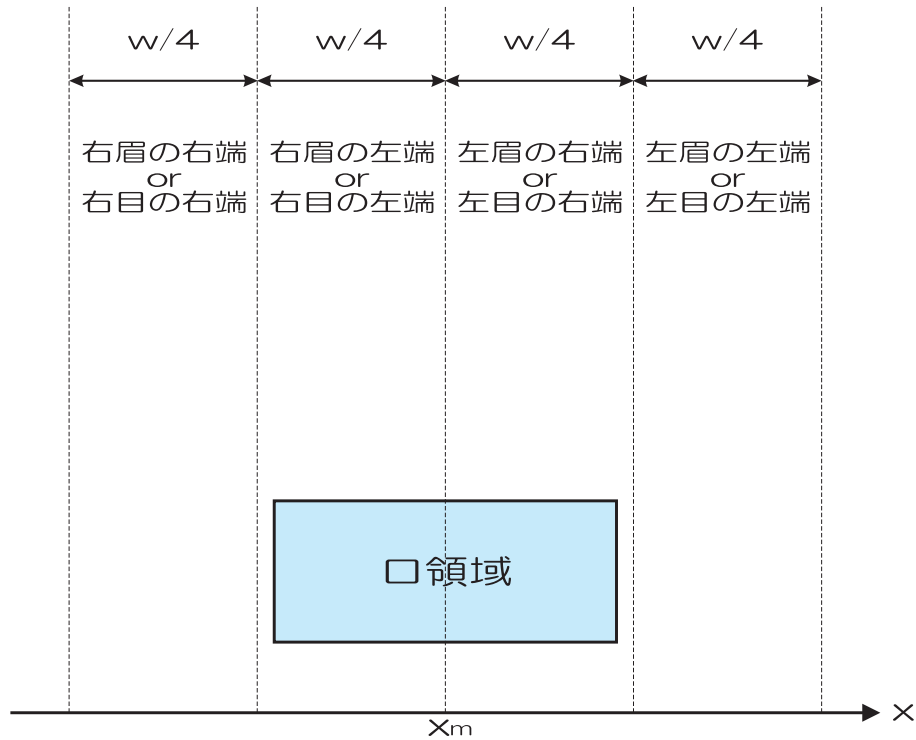


図 4.16: 眉・目の端点の候補点の分類方法

るため、顔領域や顔器官の抽出の際に行ったような画像の縮小は行わず、入力画像から切り出した各顔器官の画像をそのまま用いて処理を行う。具体的な方法を以下に述べる。また、図 4.17に抽出方法を示す。

1. 眉・目領域画像から Y 成分を、口領域画像から Q 成分を算出する。
2. 1. で得られた画像を 3×3 メディアンフィルタにより平滑化し、雑音成分を取り除く。
3. 平滑化した画像を二値化する。眉・目領域は Y 成分が小さい部分を、口領域は Q 成分が大きい部分を取り出す。
4. 領域分割を行い、最大の面積を持つ領域だけを残し、他の領域は除去する。
5. 残った領域の左右の端点を特徴点として抽出する。
6. 両端点以外の特徴点はそれぞれ次のようにして抽出する。
 - 眉の特徴点 (図 4.18)
両端点の垂直二等分線上にある、領域中央の点を特徴点とする。

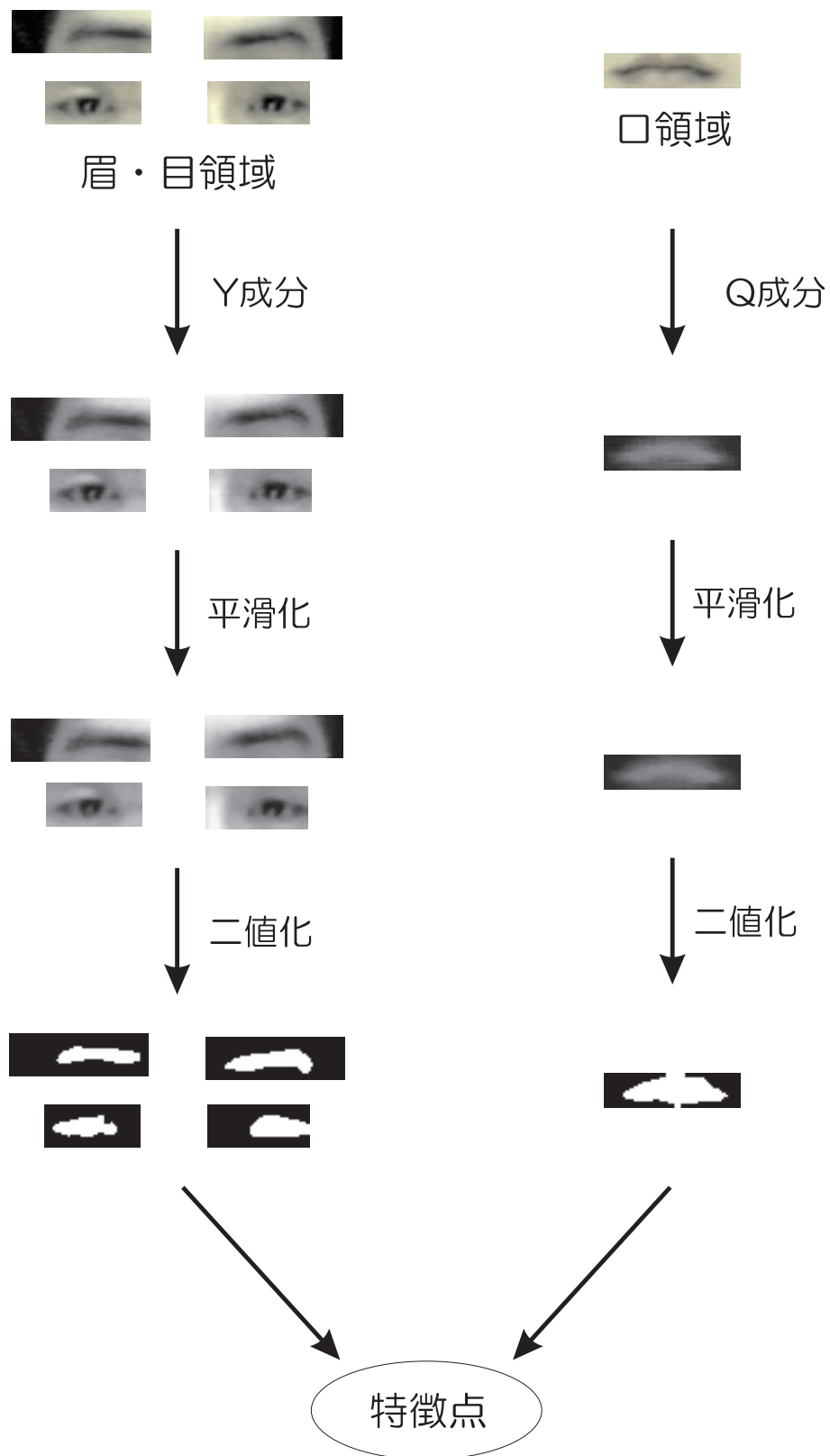


図 4.17: 特徴点の抽出方法

- 目・口の特徴点 (図 4.19、図 4.20)

領域をすべて含む矩形領域を設定し、その中心を通り y 軸と平行な直線と矩形領域の上下の辺が交わる点を特徴点とする。

4.2.4 特徴量の算出

ここでは、特徴点の座標から特徴量を算出する。まず、表情による位置の変化がほとんどない特徴点 P_7 及び P_{13} を基準として、各特徴点の座標を図 4.5 のような座標系に変換する。顔は常に正面を向いていることを仮定しており、これにより顔の大きさと傾きを補正できる。

具体的には、特徴点 P_7 及び P_{13} の座標をそれぞれ (x_7, y_7) 、 (x_{13}, y_{13}) 、補正する特徴点 P_k の座標を (x_k, y_k) とすると、補正後の P_k の座標 (x, y) は次式により表すことができる。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{l} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_k - x_7 \\ y_k - y_7 \end{pmatrix} \quad (4.8)$$

θ : $\overrightarrow{P_7P_{13}}$ が x 軸となす角

$$l = \sqrt{(x_{13} - x_7)^2 + (y_{13} - y_7)^2}$$

ここで、

$$\cos \theta = \frac{x_{13} - x_7}{l}, \quad \sin \theta = \frac{y_{13} - y_7}{l} \quad (4.9)$$

であるから、結局、式 4.8 は次のようになる。

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{l^2} \begin{pmatrix} x_{13} - x_7 & y_{13} - y_7 \\ -y_{13} + y_7 & x_{13} - x_7 \end{pmatrix} \begin{pmatrix} x_k - x_7 \\ y_k - y_7 \end{pmatrix} \quad (4.10)$$

次に、変換した座標系上での各特徴点の座標から次のようにして図 4.5 の 16 個の特徴量を算出する。

- V_1, V_{15} : 2 つの特徴点の x 座標の差
- $V_2 \sim V_7, V_{12} \sim V_{14}$: 特徴点の y 座標の絶対値
- V_{16} : 2 つの特徴点の y 座標の差
- V_8, V_9 : 直線 P_7P_9 と特徴点との距離
- V_{10}, V_{11} : 直線 $P_{11}P_{13}$ と特徴点との距離

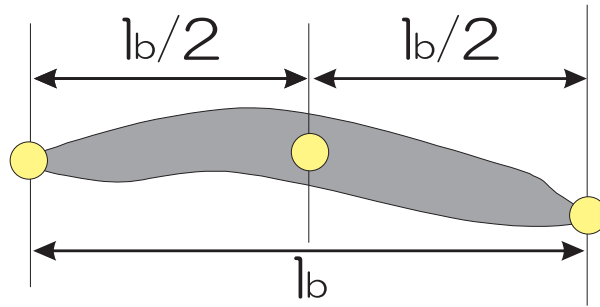


図 4.18: 眉の特徴点の抽出方法

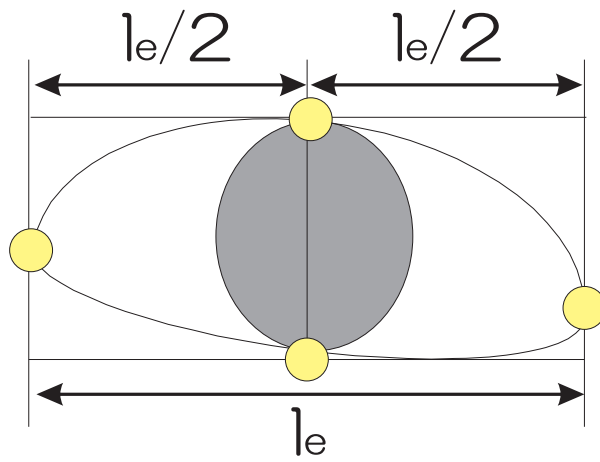


図 4.19: 目の特徴点の抽出方法

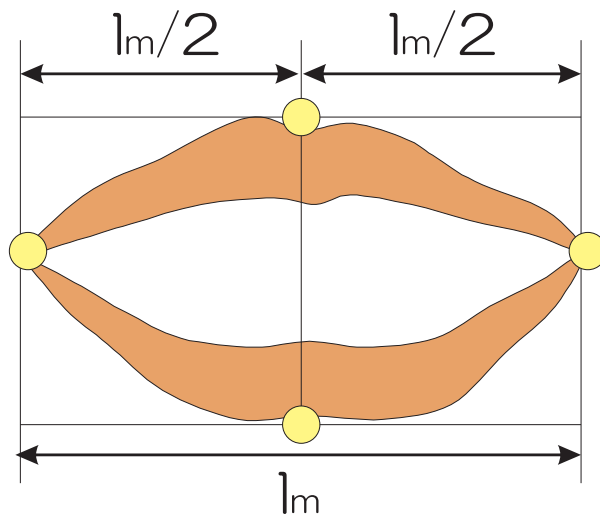


図 4.20: 口の特徴点の抽出方法

4.3 表情認識サブシステム

表情認識サブシステムでは、特徴量抽出サブシステムで得られた特徴量から顔の3領域のそれぞれについて基本6表情の認識を行う。具体的には、ファジィ推論を用いて特徴量からの表情認識を行っている。すなわち、各表情に特徴的な表情動作を特徴量によって表現し、それをもとにファジィルールを作成してファジィ推論器を構成し、ファジィ推論を行う。

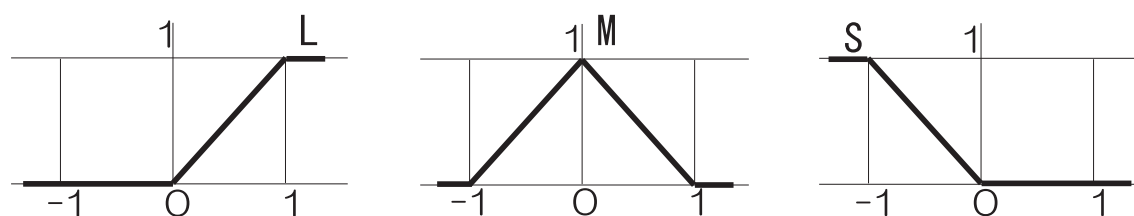
各ファジィルールには、ある特徴量が中立の表情と比較して大きい (**L**) か、小さい (**S**) か、あるいは同じくらい (**M**) であるかという条件のときに、各表情が強い (**ST**) か、あるいは弱い (**WK**) かが記述されている。例えば、表情 exp_k では、特徴量 V_i が大きく、 V_j が小さくなるという場合には、それに対応するファジィルールは、

$$\begin{aligned} &IF V_i \text{ is } \mathbf{L} \text{ and } V_j \text{ is } \mathbf{S} \\ &THEN exp_k \text{ is } \mathbf{ST} \text{ and } exp_{others} \text{ are } \mathbf{WK} \end{aligned} \quad (4.11)$$

とした。なお、各ファジィ集合のメンバシップ関数は図 4.21 に示すものを用いた。

ファジィ推論器への入力には特徴量を正規化したものを用いる。具体的には、中立の表情での特徴量を V_{k-n} とすると、ファジィ推論器への入力 x_k を次式により求める。

$$x_k = \begin{cases} \alpha_{k-p}(V_k - V_{k-n}) & (V_k \geq V_{k-n} \text{ のとき}) \\ \alpha_{k-m}(V_k - V_{k-n}) & (V_k < V_{k-n} \text{ のとき}) \end{cases} \quad (4.12)$$



(a) 前件部



(b) 後件部

図 4.21: 表情認識に用いたメンバシップ関数

ただし、 α_{k-p} 及び α_{k-m} は特徴量によって異なる定数である。中立の表情に比べて大きい場合と小さい場合では、可能な変化量が異なるので、それぞれの場合に異なる定数を用いる。また、このときの、 V_{k-n} の値は、中立の表情の画像からシステムにより実際に求めた特徴量の値を用いる。

この入力値から出力値を算出する方法は次の通りである。まず、この入力値から、各ルールの適合度 w_i を次式により求める。

$$w_i = \prod_{k=1}^{n_{in}} \mathbf{A}_k(x_k) \quad (4.13)$$

ただし、 n_{in} は入力の特徴量の個数、 \mathbf{A} はルールの前件部に記述された各特徴量に対応するファジィ集合 (\mathbf{L} 、 \mathbf{M} 、 \mathbf{S} のいずれか) のメンバシップ関数である。次に、ルールごとの表情 j についての推論結果 $\mathbf{B}_{i,j}$ を次式により求める。

$$\mathbf{B}_{i,j} = w_i \mathbf{C}_{i,j} \quad (4.14)$$

ただし、 \mathbf{C} はルールの後件部に記述された各表情に対応するファジィ集合 (\mathbf{ST} 、 \mathbf{WK} のいずれか) である。次に、ルール全体の表情 j についての推論結果 \mathbf{B}_j を、

$$\mathbf{B}_j = \bigcup_i \mathbf{B}_{i,j} \quad (4.15)$$

として求め、最後にこのファジィ集合のメンバシップ関数の重心をとることにより非ファジィ化し、表情 j についての出力値 y_j を求める。すなわち、

$$y_j = \frac{\int s \mathbf{B}_t(s) ds}{\int \mathbf{B}_t(s) ds} \quad (4.16)$$

がファジィ推論器の出力値となる。この値が大きいほど強い表情である。

このようなファジィ推論器を、顔の3領域のそれぞれについて作成した。各領域において推論に用いたファジィルールを表4.1、表4.2、及び表4.3にまとめる。表には各表情が \mathbf{ST} である条件が示されている。なお、顔の上部には「幸福」の表情に特徴的な表情動作がないので、「幸福」については推論を行わない。

4.4 視線方向検出サブシステム

視線方向検出サブシステムでは、特徴量抽出サブシステムにおいて入力画像から特徴量を抽出する過程で切り出される目領域の画像から視線方向を検出する。まず、目

表 4.1: 顔の上部のファジィルール

表情	特徴量						
	V ₁	V ₂	V ₃	V ₄	V ₅	V ₆	V ₇
怒り	S	S	S	S	S	S	S
嫌悪	M	S	S	S	S	S	S
恐怖	S	L	L	L	L	L	L
悲しみ	S	M	M	L	L	M	M
驚き	M	L	L	L	L	L	L
中立	M	M	M	M	M	M	M

表 4.2: 顔の中央部のファジィルール

表情	特徴量							
	V ₄	V ₅	V ₈	V ₉	V ₁₀	V ₁₁	V ₁₂	V ₁₄
怒り	S	S	S	S	S	S	M	M
嫌悪	S	S	S	S	S	S	S	S
恐怖	L	L	L	S	L	S	M	M
幸福	M	M	M	S	M	S	S	S
悲しみ	M	M	S	S	S	S	L	L
驚き	L	L	L	M	L	M	M	M
中立	M	M	M	M	M	M	M	M

表 4.3: 顔の下部のファジィルール

表情	特徴量				
	V ₁₂	V ₁₃	V ₁₄	V ₁₅	V ₁₆
怒り	M	M	M	M	S
嫌悪	S	S	S	M	M
恐怖	M	M	M	L	L
幸福	S	S	S	L	M
悲しみ	L	M	L	L	M
驚き	M	M	M	M	L
中立	M	M	M	M	M

領域の画像から目の輪郭線を抽出し、次にその輪郭線から瞳の中心点を検出する。そして最後に、瞳の中心点と目の両端点との位置関係から視線方向を推定する。図 4.22 にこの視線方向検出方法を示し、以下で詳しく述べる。

目の輪郭線の抽出

まず、目領域の画像から次の手順により目の輪郭線を抽出する。

1. 目領域の画像から YIQ 表色系の Y 成分を算出する。
2. Y 成分画像を 3×3 メディアンフィルタにより平滑化し、雑音成分を除去する。
3. 平滑化した画像を二値化し Y 成分の値が小さい部分を取り出す。これにより、瞳の部分や目の上端が作る影などが抽出される。
4. 膨張・収縮処理により、細かな凹凸を除去する。
5. 残された領域の輪郭線を抽出する。

瞳の中心点の抽出

次に、目の輪郭線から、瞳が円形であることを利用して瞳の中心点を抽出する。

まず、図 4.23 のように輪郭線上のある点 (x_k, y_k) とその点から両側に一定距離だけ離れた 2 点 (x_p, y_p) 及び (x_m, y_m) の計 3 点を通る円を求める。その円の方程式は次式のようなになる。

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \quad (4.17)$$

ただし、

$$x_c = x_k + \frac{(y_m - y_k)\{(x_p - x_k)^2 + (y_p - y_k)^2\} - (y_p - y_k)\{(x_m - x_k)^2 + (y_m - y_k)^2\}}{2\{(x_p - x_k)(y_m - y_k) - (x_m - x_k)(y_p - y_k)\}}$$

$$y_c = y_k + \frac{(x_p - x_k)\{(x_m - x_k)^2 + (y_m - y_k)^2\} - (x_m - x_k)\{(x_p - x_k)^2 + (y_p - y_k)^2\}}{2\{(x_p - x_k)(y_m - y_k) - (x_m - x_k)(y_p - y_k)\}}$$

$$r = \sqrt{(x_c - x_k)^2 + (y_c - y_k)^2}$$

である。ここで、特徴点 P_7P_{13} 間の距離と比較して、この円の半径 r が、次の条件を満たすとき、円の中心点 (x_c, y_c) を瞳の中心の候補点とする。

$$0.03 \leq \frac{r}{P_7P_{13}} \leq 0.12 \quad (4.18)$$

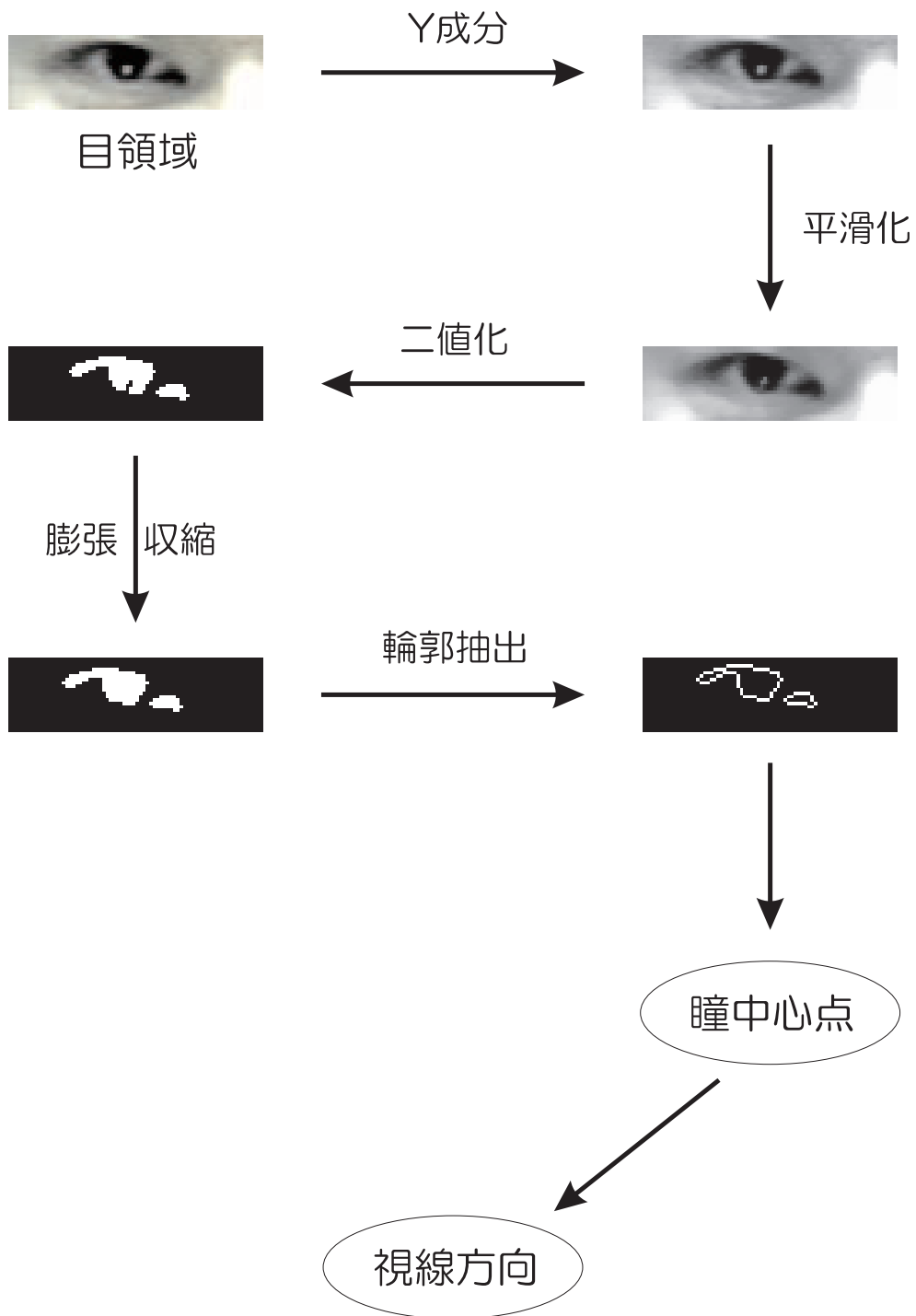


図 4.22: 視線方向の検出方法

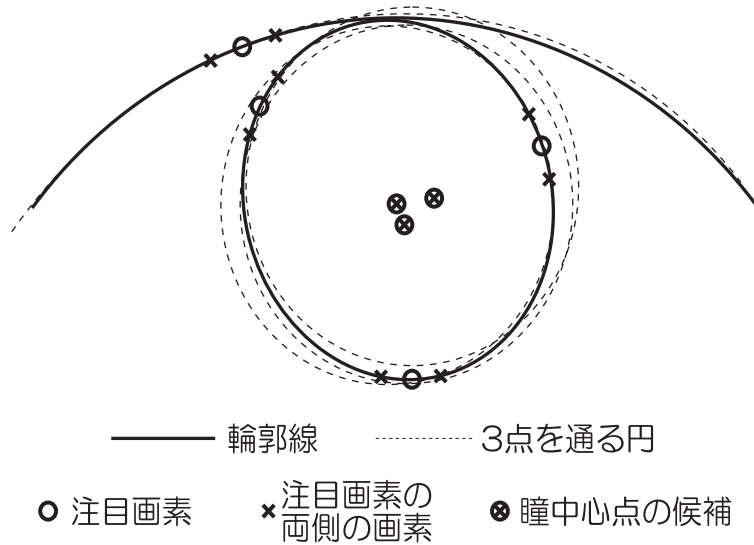


図 4.23: 瞳の中心点の抽出方法

この処理を輪郭線上のすべての点について行い、得られた候補点の座標の平均を算出し、瞳の中心点の座標とする。

視線方向の推定

まず、得られた瞳の中心点の座標を式 (4.10) を用いて図 4.5に示した xy 座標系に変換する。基準点 P_7 及び P_{13} と眼球との位置関係は変化しないので、この座標系では瞳の中心点は眼球の回転によってのみ移動する。このとき、図 4.24のように真正面を向いている場合からの眼球の回転角を θ [rad]、図 4.5の座標系での眼球の半径を r_p 、瞳の中心点の移動量を d とすると、次式が成り立つ。ただし、眼球は完全な球体であると仮定する。

$$\theta = \sin^{-1} \frac{d}{r_p} \quad (4.19)$$

また、実際の P_7P_{13} 間の距離を w_r 、眼球半径を r_r とすると、

$$r_p = \frac{r_r}{w_r} \quad (4.20)$$

であるから、結局、

$$\theta = \sin^{-1} \frac{dw_r}{r_r} \quad (4.21)$$

が成立する。したがって、図 4.5の座標系における、真正面を向いているときの瞳の中心点、及び推定された瞳の中心点をそれぞれ (x_c, y_c) 、 (x_p, y_p) とすると、視線方向、す

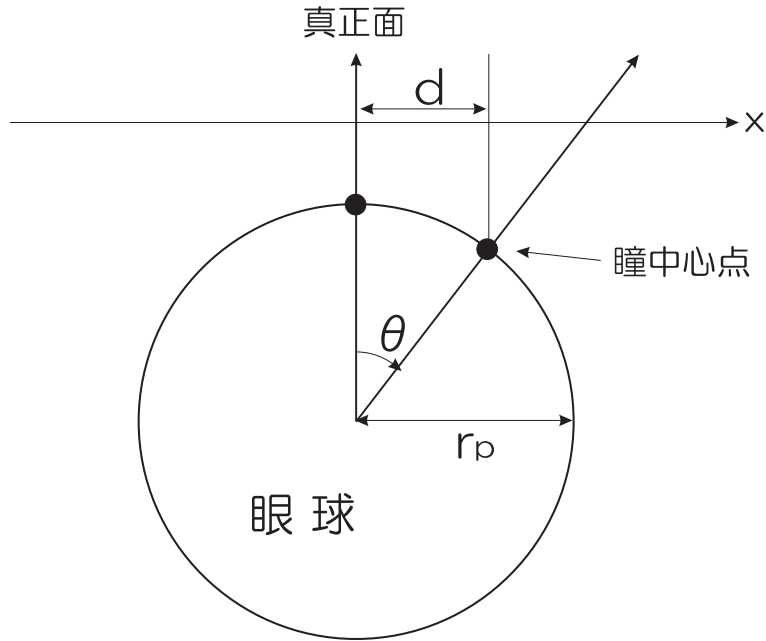


図 4.24: 眼球の回転角と瞳の中心点の関係

なわち眼球の x 軸及び y 軸方向の回転角 θ_x 及び θ_y は、

$$(\theta_x, \theta_y) = \left(\sin^{-1} \frac{w_r(x_p - x_c)}{r_r}, \sin^{-1} \frac{w_r(y_p - y_c)}{r_r} \right) \quad (4.22)$$

となる。

4.5 感情推定サブシステム

感情推定サブシステムでは、表情認識サブシステムで得られる表情と、視線方向検出サブシステムで得られる視線方向の時間的推移から、基本 6 感情だけでなくより複雑な混合感情を推定する。3.2.5 に推定し得る感情について述べたが、現在のところ試作システムにはそのごく一部しか実装されていない。試作システムで推定できる感情は次の通りである。

- 顔の 3 領域の組み合わせから、感情を推定する。推定できる感情は表 4.4 の通りである。これらは、表 3.7 の中から、比較的分かりやすいものを選んだ。
- 異なる 2 種類の表情が短時間のうちに続けて現れた場合、感情はその 2 種類を混合したものであるとする。試作システムでは、「驚き」のすぐあとに「幸福」が現われた場合、「うれしさを伴った驚き」の感情とする。

表 4.4: 試作システムで顔の 3 領域の組み合わせから推定できる感情

上部	中央部	下部	感情
怒り	怒り	怒り	怒り
怒り	中立	中立	抑えた怒り
嫌悪	嫌悪	嫌悪	嫌悪
恐怖	恐怖	恐怖	恐怖
中立	幸福	幸福	幸福
中立	中立	幸福	幸福
悲しみ	悲しみ	悲しみ	悲しみ
悲しみ	中立	中立	弱い悲しみ
驚き	驚き	驚き	驚き
驚き	驚き	中立	疑い

第 5 章 試作システム評価実験

本章では第 4 章で述べた試作システムに対して行った評価実験について述べ、最後に感情推定に関する考察を行う。

5.1 表情認識実験

5.1.1 実験の目的

表情を表出している動画像から、試作システムを用いて表情の認識を行い、意図した表情と認識結果を比較することにより、どの程度正しく表情が認識できるかを調べる。また、認識の途中段階である特徴量の抽出についても正しく抽出が行われているかを確認する。また、処理に要する時間を計測し、リアルタイムで動作することを確認する。

5.1.2 実験方法

表情表出動画像の作成

まず、被験者に表情を表出してもらい、その様子をビデオカメラで撮影する。撮影に際しては、カメラの設置位置や照明などを次の条件を満たすように整える。

1. 顔が真正面を向いていること
2. 顔に均一に照明が当たっていること
3. 顔ができるだけ大きく映っていること

1. については、試作システムは顔が正面を向いていることを前提としていることによる。2. については、照明が均一ではない場合は顔面上に影ができてしまい、認識が困難になるからである。3. については、顔を構成する画素数が少ないほど細部の抽出が困難になり、認識率が低下すると考えられるからである。

これらの条件を考慮して、実験は図 5.1 のように被験者の顔の真正面に、約 80cm 離れた位置にビデオカメラを設置して行った。実際の実験風景を図 5.2 に示す。また、照

明環境は図 5.3のように、上方には 3 本の蛍光灯を左右対称になるように配置し、下前方にも蛍光灯を設置した。なお、顔面上での明るさは 170[lux] であり、そのうち上方からのものが 110[lux]、下方からのものが 60[lux] であった。

被験者に表出してもらった表情は中立の表情と、基本 6 表情のうち驚きと幸福の、計 3 種類とした。驚きと幸福以外の表情は、意図的に表出するのが困難であり、俳優としての訓練を受けた者などでない限り、自然な表情とはならない場合がほとんどである。そこで、今回は驚きと幸福のみに限定した。また、中立の表情は、表情認識を行う際の基準となる特徴量を求めるために用いる。

被験者には事前に図 3.7の典型的な表情の写真、及び表 3.1～表 3.3の各表情に特徴的な動作のリストを簡略化したものを見せ、簡単なレクチャを行った。同時に、これらの典型的表情にはあまりこだわらず、自分がその感情を感じている場面を想像して表情をつくるように求めた。表情を表出してもらった順番は、

中立 → 幸福 → 中立 → 驚き → 中立 → 幸福 → 中立 → 驚き → 中立

の順に中立を挟んで各表情を二度ずつとした。

被験者は、正常な表情表出を行うことができると思われる、本研究室の男子学生 3 人である。眼鏡をかけている場合は外してもらい、また、髪の毛が顔にかからないようにした状態で撮影を行った。

試作システムによる表情の認識

以上のようにして撮影したビデオから、表情を表出している部分を WS 上に取り込み、横 320 × 縦 240 ピクセルの大きさの画像を、毎秒 10 フレーム程度でハードディスクに保存した。このとき、表情変化を見るために「幸福」及び「驚き」の表情の前後

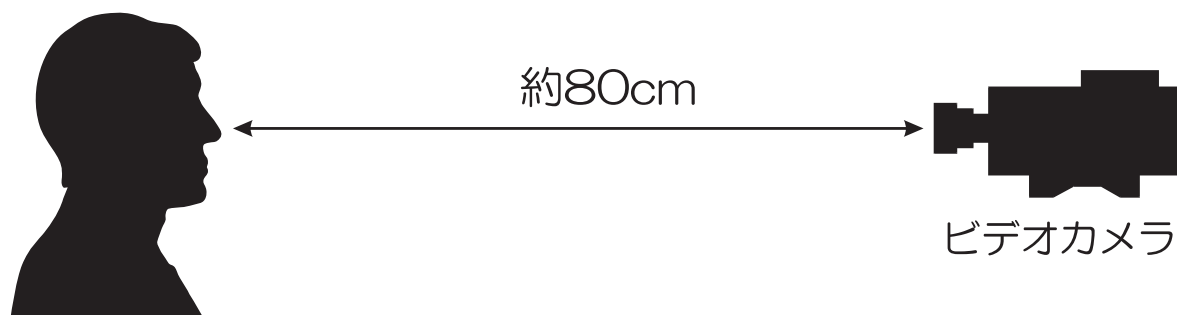


図 5.1: 表情表出動画像の作成

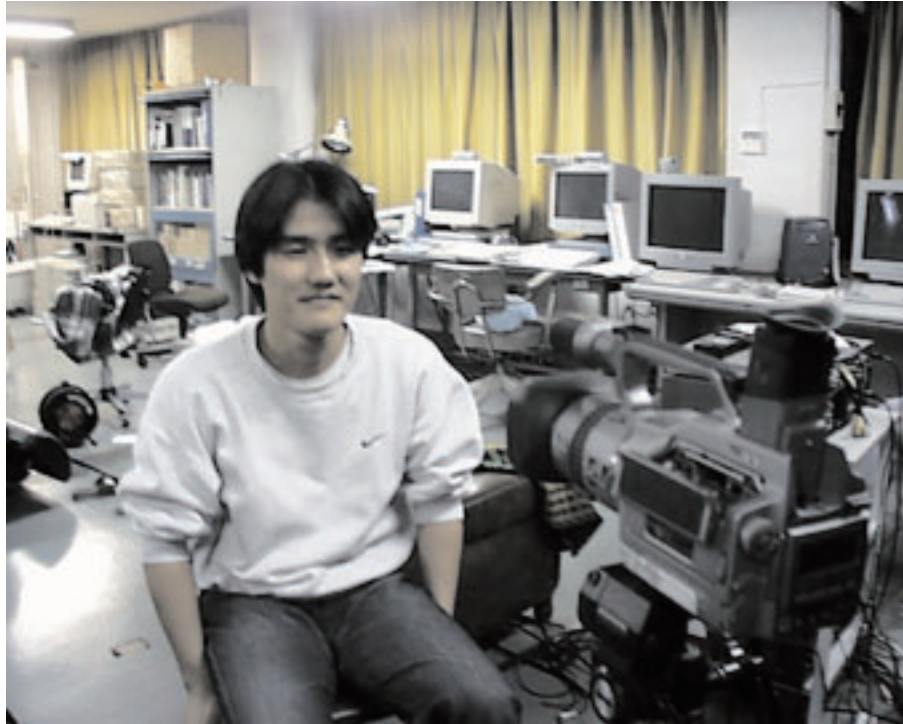


図 5.2: 表情認識実験の実験風景

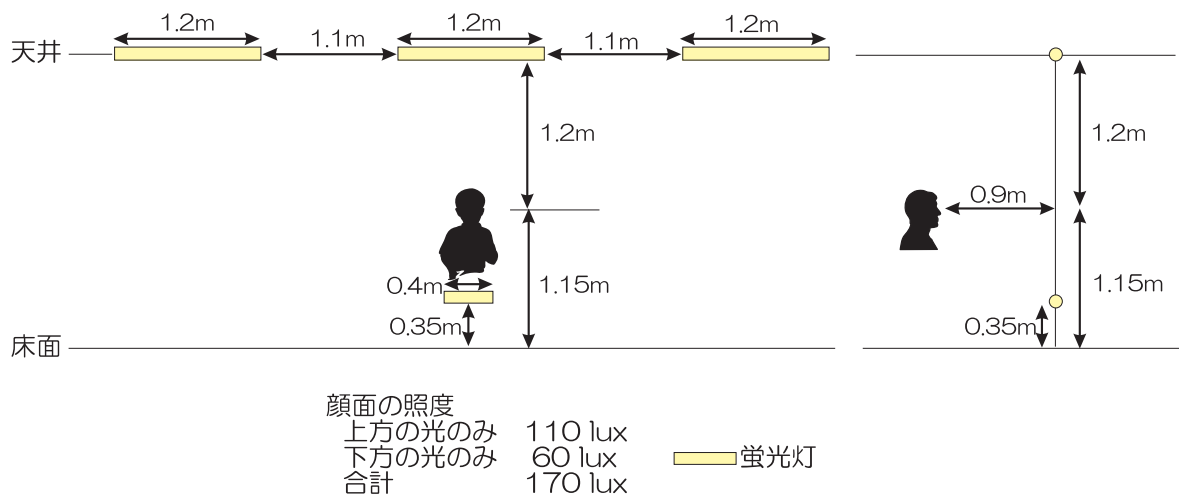


図 5.3: 実験時の照明環境

には必ず中立の表情が含まれるようにした。試作システムによる表情の認識にはこのハードディスク内の動画像を用いる。

つづいて、試作システムにより表情の認識を行った。まず、表情認識サブシステムで必要となる中立の表情の場合の特徴量を、特徴量抽出サブシステムを用いて計測する。これには、撮影した中立の表情のうち最初のものを用いた。この中立の場合の特徴量をシステムに反映させた後、ハードディスクに記録した各表情を表出している動画像を入力画像として、試作システムにより表情の認識を行い、表情認識サブシステムからの出力値、すなわち顔の3領域のそれぞれについての各表情の強度を求めた。

5.1.3 実験結果及び考察

まず、入力画像として用いた画像の例を図 5.4及び図 5.5に示す。以下では、図 5.4の画像を入力画像 1、図 5.5の画像を入力画像 2 と呼ぶ。

入力画像 1 では被験者は「幸福」の表情を表出している。「幸福」の表情で最も特徴的な動作は、顔の下部における「唇の両端を後方に引き持ち上げる」動作である。この画像を入力としたときの、顔の下部における特徴量の抽出結果及び、表情認識結果をそれぞれ図 5.6及び図 5.7に示す。なお、図 5.7には比較的変動の大きかった、「中立」、「幸福」及び「嫌悪」の結果のみを示す。それ以外はみな低い値からほとんど変化しなかった。入力画像 1 からは、5 フレーム目あたりから「幸福」の表情の表出が始まり、43 フレーム目あたりまで持続していることが読み取れるが、特徴量にもその変化が捉えられている。すなわち、唇の両端が上方に持ち上げられることにより、特徴量 12、13 及び 14 が減少し、後方に引かれることにより、特徴量 15 が増加していることが分かる。また、表情の認識結果も実際の表情の変化とともに、「中立」から「幸福」へ変化し、再び「中立」へと戻っている。「嫌悪」の結果が比較的大きくなったのは、「嫌悪」の表情も口全体を上方へ持ち上げる動作を含んでいるためであると考えられる。

入力画像 2 では、「驚き」の表情が表出されている。「驚き」の表情で最も特徴的な動作は「眉全体を上方に引き上げる」動作である。この画像を入力としたときの、顔の上部における特徴量の抽出結果、及び表情認識結果をそれぞれ図 5.8及び図 5.9に示す。入力画像 1 の場合と同様に、図 5.9には比較的変動の大きかった、「中立」、「驚き」及び「恐怖」の結果のみを示す。入力画像 2 では、3 フレーム目付近から驚き始めており、特徴量がそれに伴い増加し、表情認識結果も「中立」から「驚き」へと変化していくことが分かる。また、「恐怖」の表情でも眉を上げる動作が伴うので、表情の強さの出力値が若干増加している。

同様に解析した表情認識結果を表 5.1 にまとめる。特徴量抽出に関しては、処理を行った全フレーム数 (全数) のうち何フレームで特徴量の抽出に成功したか (成功数) を示している。表情認識に関しては、「驚き」もしくは「幸福」の表情が表出されており、かつ特徴量の抽出に成功したフレーム数 (有効数) と、そのうち中立を除いた基本 6 表情のなかでその表情の認識結果が最も大きかったフレーム数 (成功数) を、顔の上部、中央部、下部のそれぞれについて示した。なお、被験者 N において顔の上部の有効フレーム数が 0 となっているが、これは被験者 N の「驚き」の表情に特徴があることによる。被験者 N は「驚き」の表情を表出するときにまったく眉を動かしていなかったため、常に「中立」の値が大きかった。しかし、目と口によって驚きを表現しており、全体として「驚き」の表情に見える (図 5.10)。このように、一見動いているように見える部分でも、実際には全く動いていない場合がある。

表 5.1 から分かる通り、特徴量の抽出で 94%、表情認識で 90%弱と、高い成功率であった。特徴量抽出の過程では、顔領域の抽出を失敗した例は一つもなく、すべて顔器官の抽出段階以降であり、特に眉・目領域の抽出における失敗が目立った。これは、眉と目がもっとも細かい造作であり、頭髮など、誤抽出しやすい対象が存在するため、当然であるとも言える。また、顔器官は抽出できているものの、特徴点の抽出を失敗した例もあった。被験者 I では、「驚き」の表情で目を大きく見開いたときに、瞳だけを暗色領域として抽出してしまい、瞳の端点を特徴点とした場合があった (図 5.11)。現在のところ、試作システムでは特徴点の抽出手法として非常にシンプルなものを用いているため、このような場合にはうまく対処できない。今後、改善していく必要がある。

表情の認識については、表情をはっきりと表出している場合には正しく認識が行えるが、「中立」から変化していく途中などでは、誤った認識を行う場合があった。特に、形態的な特徴が類似している「驚き」と「恐怖」を混同することがあった。これは、認識に用いている現在の特徴量とファジィルールが、必要最小限のものに限られているためと考えられる。特徴量を増やしルールを再考することにより、改善できると思われる。また、表中では失敗例に含めていないが、「中立」の表情の強さの出力値が最も大きい値を示し、「驚き」または「幸福」が二番目であることも多かった。これは、特徴量をファジィ推論器に入力する際に掛け合わせる係数が適当でないためと考えられる。筋肉の動作範囲には個人差があるため、各個人ごとに異なる係数を用いることが必要であると思われる。

また、処理に要した時間を表 5.2 に示す。表の値は横 320 × 縦 240pixel の画像で、顔が画像の縦幅一杯に映っている場合のものである。なお、特徴量の算出に要した時間

は、非常に短時間のため測定できなかった。表から分かるように、1フレームの処理に要した時間はO2で35msec、OCTANEで21msec程度であった。すなわち、それぞれ毎秒28フレーム、及び47フレームの認識を行うことができる。このことから、試作したシステムは十分にリアルタイムで動作することが確認できた。

表 5.1: 表情認識実験結果

被験者	表情	特徴量抽出 成功数 / 全数	表情認識		
			上部	中央部	下部
			成功数 / 有効数	成功数 / 有効数	成功数 / 有効数
A	幸福 1	31 / 34	-	20 / 24	20 / 24
	幸福 2	39 / 39	-	22 / 24	9 / 13
	驚き 1	14 / 14	12 / 12	11 / 12	11 / 11
	驚き 2	20 / 25	10 / 11	8 / 10	12 / 13
被験者 A 計		104 / 112	22 / 23	61 / 70	52 / 61
I	幸福 1	46 / 47	-	38 / 39	36 / 39
	幸福 2	61 / 61	-	52 / 53	52 / 53
	驚き 1	27 / 35	5 / 6	5 / 6	7 / 7
	驚き 2	26 / 33	7 / 9	7 / 9	8 / 11
被験者 I 計		160 / 177	12 / 15	102 / 107	103 / 110
N	幸福 1	58 / 58	-	28 / 38	37 / 40
	幸福 2	45 / 49	-	25 / 32	35 / 38
	驚き 1	51 / 51	0 / 0	23 / 25	21 / 26
	驚き 2	45 / 46	0 / 0	25 / 26	24 / 28
被験者 N 計		199 / 204	0 / 0	101 / 121	117 / 132
幸福 計		280 / 289	-	185 / 210	189 / 207
驚き 計		183 / 204	34 / 38	79 / 88	83 / 96
全計		463 / 493	34 / 38	264 / 298	271 / 303

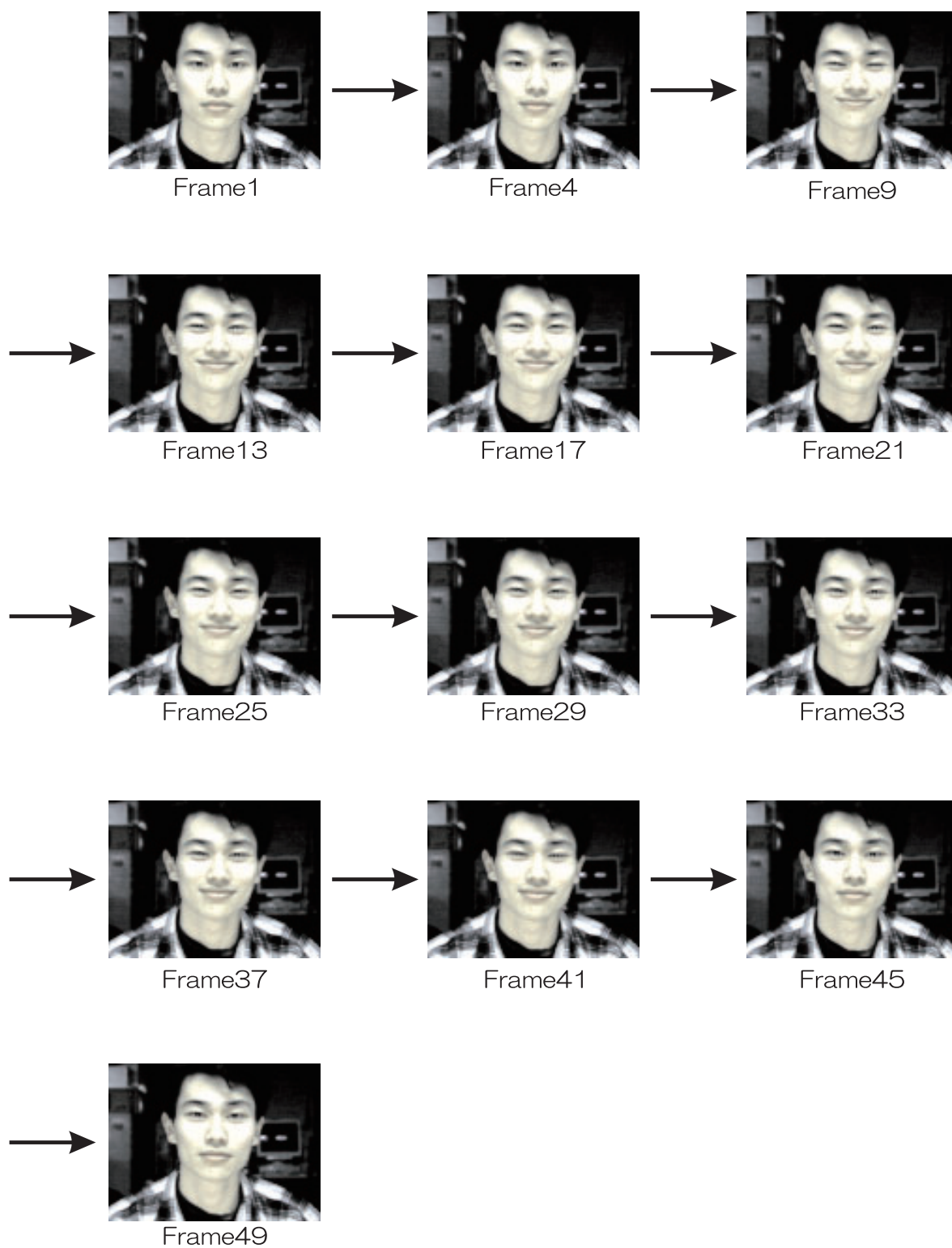


图 5.4: 入力画像例 (1)

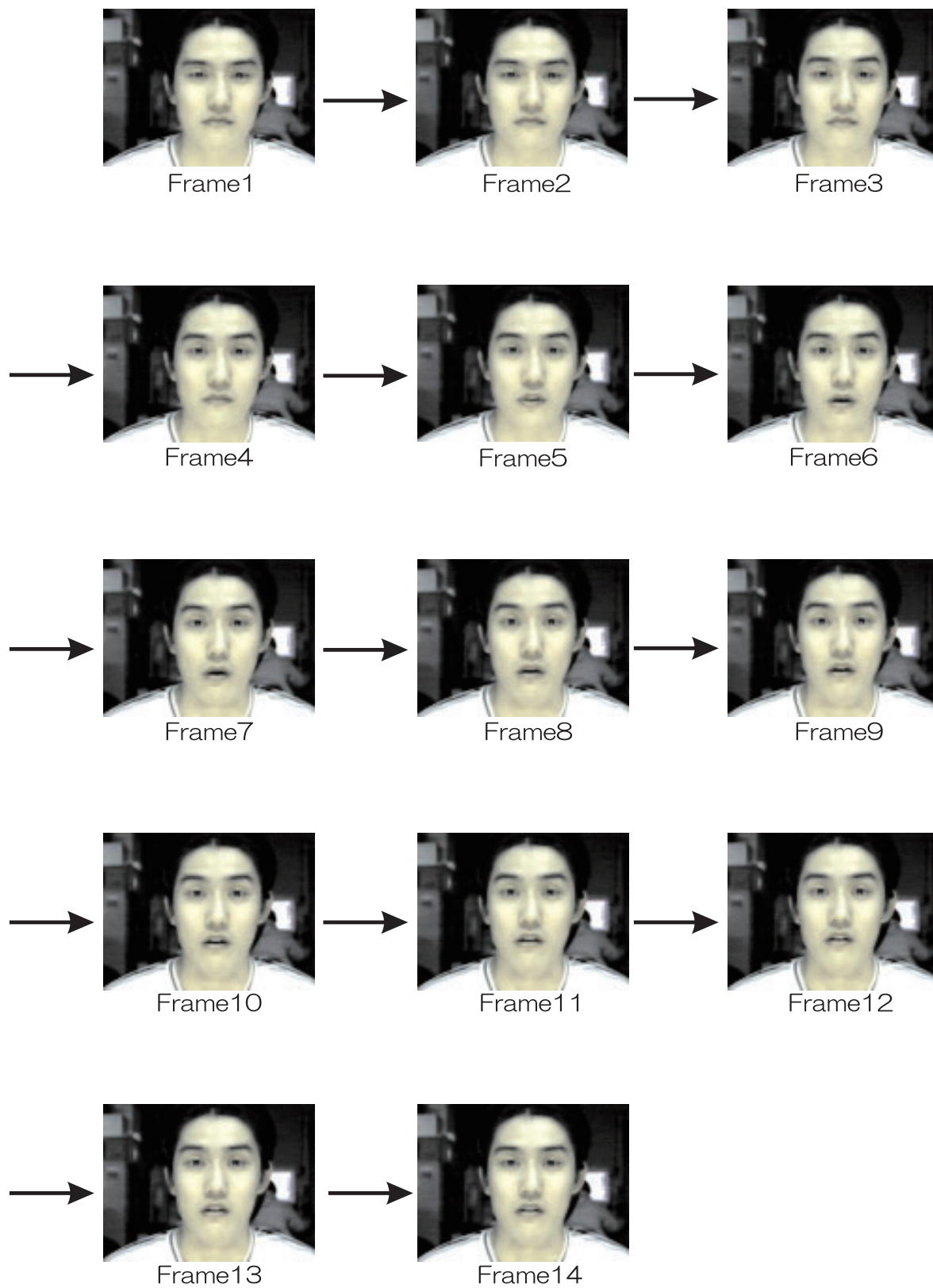


图 5.5: 入力画像例 (2)

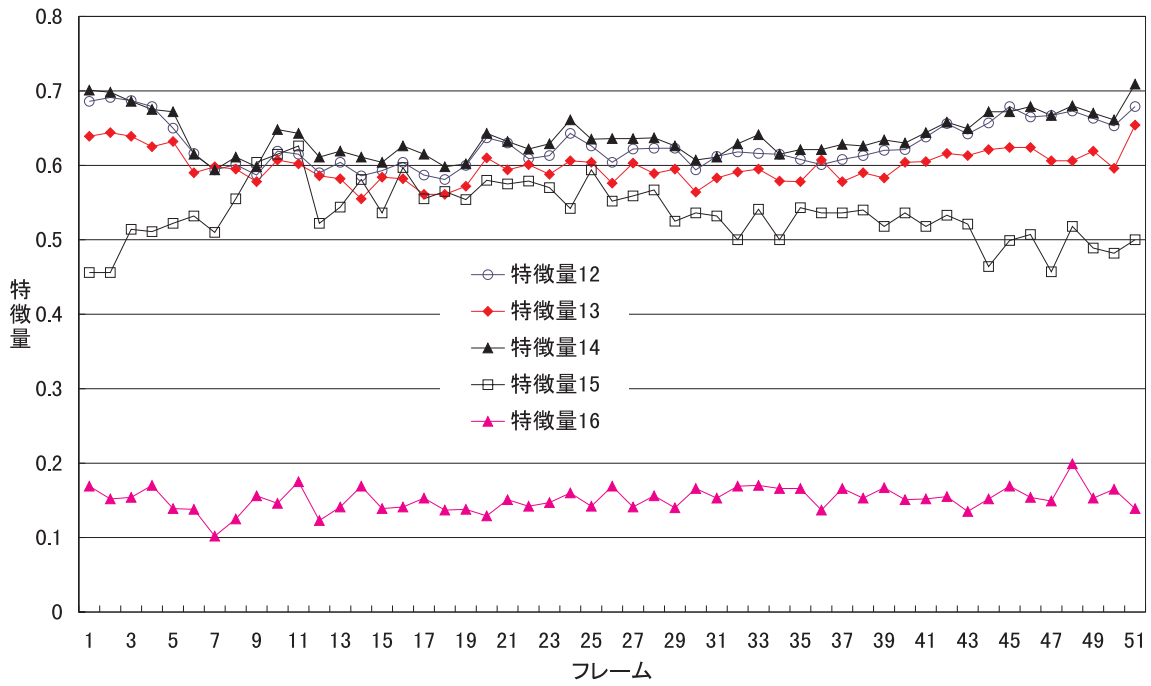


図 5.6: 特徴量抽出結果 (入力画像 1)

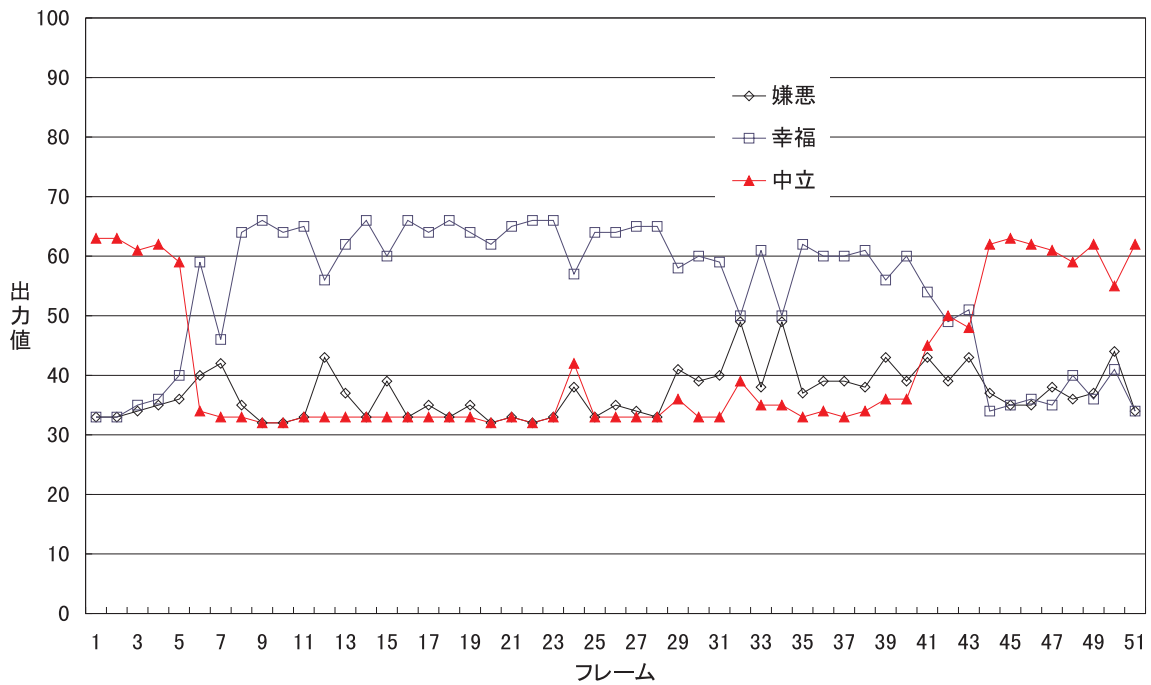


図 5.7: 表情認識結果 (入力画像 1)

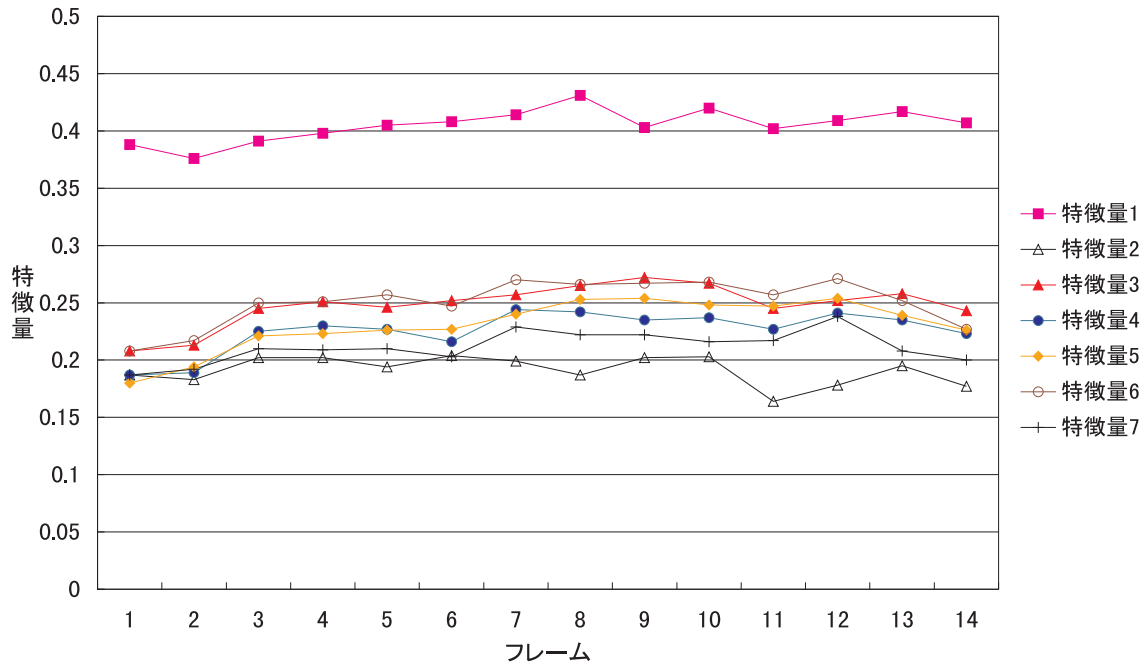


図 5.8: 特徴量抽出結果 (入力画像 2)

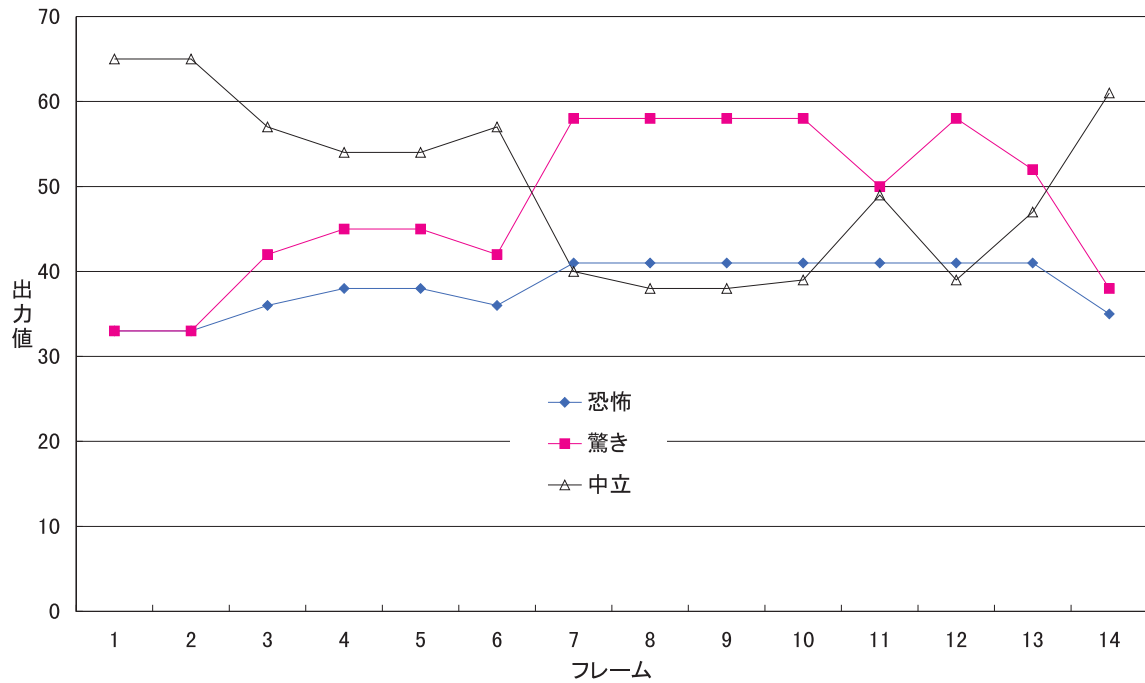


図 5.9: 表情認識結果 (入力画像 2)

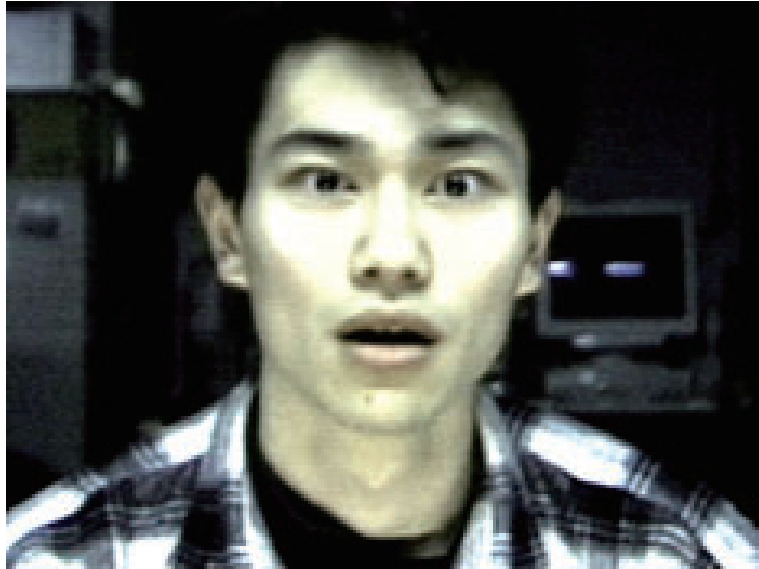


図 5.10: 被験者 N の「驚き」の表情

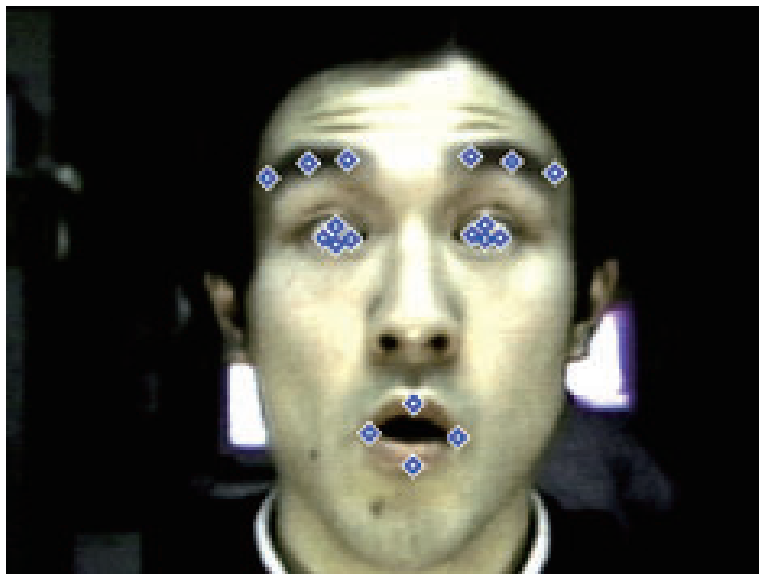


図 5.11: 特徴点の抽出失敗例

表 5.2: 1 フレームの認識に要する時間 (msec)

	顔領域 抽出	顔器官 抽出	特徴点 抽出	特徴量 算出	特徴量 抽出計	表情 認識	視線方向 検出	合計
O2	13.3	7.4	6.1	-	26.8	5.0	3.0	34.8
OCTANE	6.3	5.2	4.6	-	16.1	3.0	2.3	21.4

5.2 視線方向検出実験

5.2.1 実験の目的

ある方向を注視している動画像から、試作システムにより視線方向を検出する。実際の方向と検出結果を比較することにより、視線方向の検出精度について考察する。

5.2.2 実験方法

動画像の作成

まず、被験者がある方向を注視している場面をビデオカメラで撮影する。撮影環境は表情認識実験と同じであるが、視線方向を変える際に顔の向きもそれに追従して変わってしまうことを避けるために、前述した3つの条件以外に、

4. 頭部を固定すること

という条件を加えた。

被験者に注視してもらう点は、図 5.12 のように被験者の前方 1m の距離に設置したボード上に設定した計 35 個の点である。注視点の詳細な位置を図 5.13 に示す。1 番の点が被験者の真正面にくるようにし、上下方向には眼球の回転角がそれぞれ 15° 及び 30° となる点を、また左右方向にはそれぞれ 15°、30° 及び 45° となる点を選んだ。眼球は上下方向にはそれぞれ約 20°、左右方向にはそれぞれ約 30° ずつ回転できることを考慮し、これらの回転角よりも若干大きい範囲を選んだ。また、注視する順番は、

$$1 \rightarrow 2 \rightarrow 1 \rightarrow 3 \rightarrow 1 \rightarrow \dots \rightarrow 1 \rightarrow 34 \rightarrow 1 \rightarrow 35 \rightarrow 1$$

とした。

被験者は、本研究室の男子学生 3 人であり、全員正常に視線移動を行えるが、コンタクトレンズを着用している者も含まれる。

以上のように、被験者が各方向を注視している場面をビデオカメラで撮影した。実際の実験風景を図 5.14 に示す。

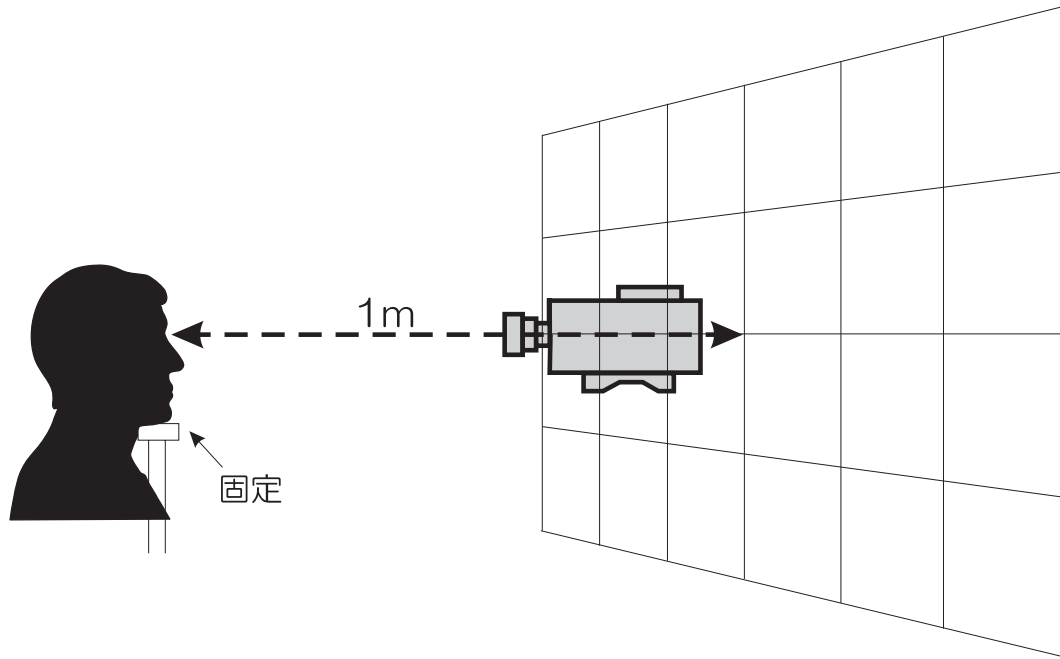


図 5.12: 視線方向検出実験用動画像の作成

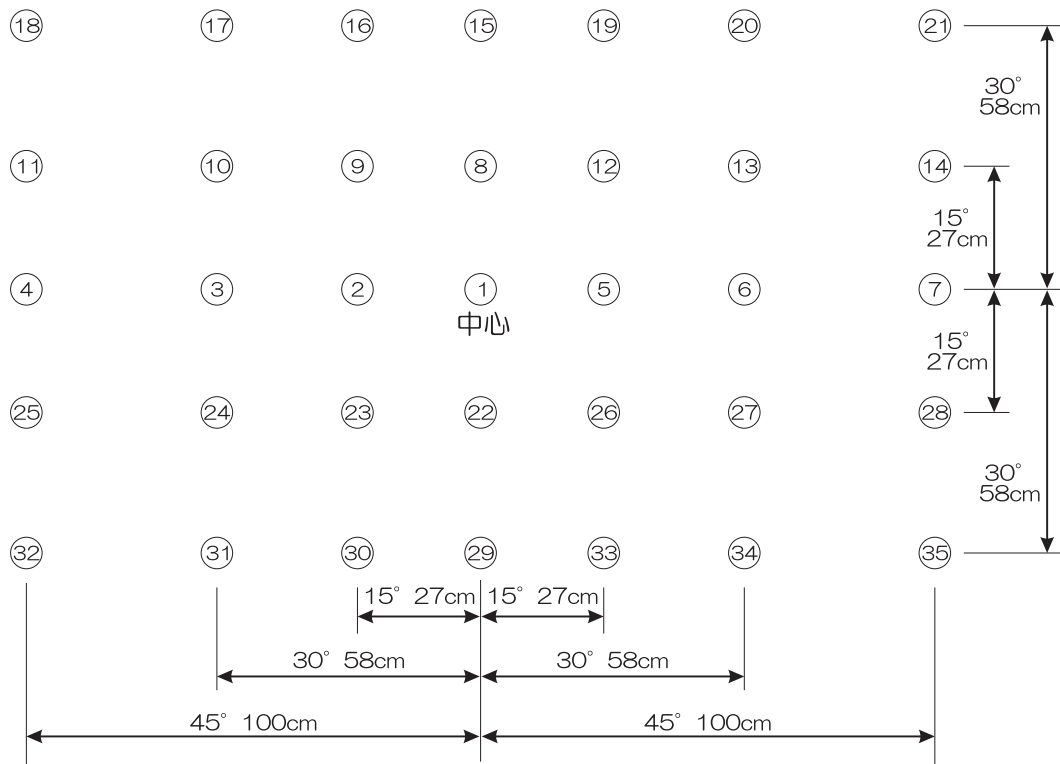


図 5.13: 視線方向検出実験における注視点

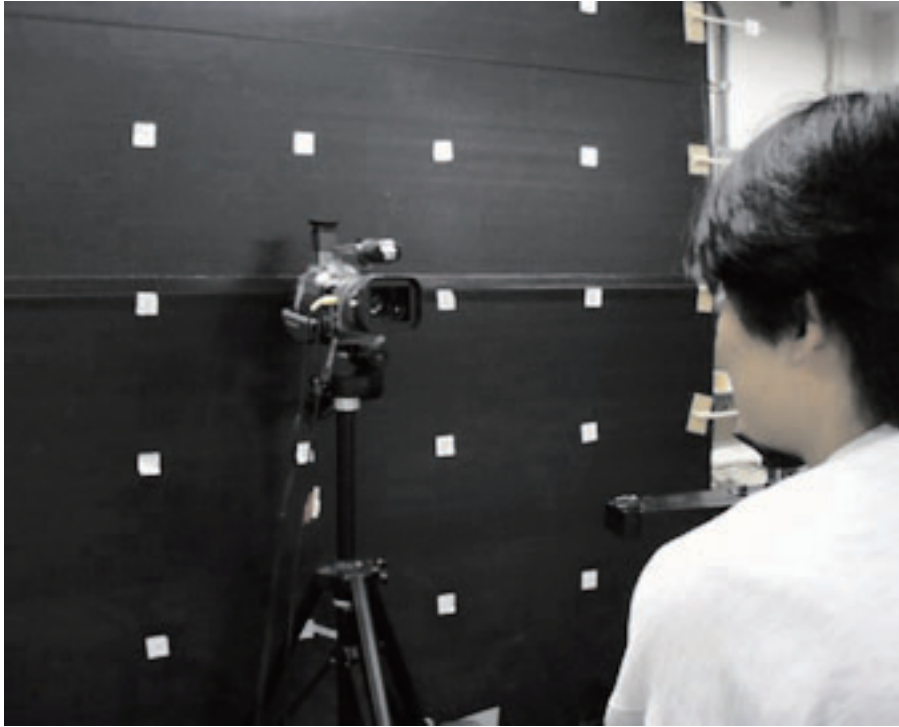


図 5.14: 視線方向検出実験の実験風景

試作システムによる視線方向の検出

まず、表情認識実験と同様に、撮影したビデオから各方向を注視している部分を WS 上に取り込み、ハードディスクに保存した。

視線方向検出サブシステムで用いている手法では、真正面を注視しているときの瞳の中心点の位置、図 4.5 に示す特徴点 P_7P_{13} 間の実際の距離、及び実際の眼球半径があらかじめ必要となる。そこで、瞳の中心点の位置は試作システムにより求め、 P_7P_{13} 間の距離は 10cm とし、また、眼球半径は 12mm という一般的な数値を用いる。

以上を試作システムに反映し、被験者が各方向を注視している動画像から、視線方向の検出を行った。

5.2.3 実験結果及び考察

視線方向検出の実験結果を図 5.15～図 5.26 に示す。ただし、各点の対応関係を理解しやすくするため、左右方向の角度が同じ点を注視した場合の結果を示す点 (同経度点) 同士、及び上下方向の角度が同じ点を注視した場合の結果を示す点 (同緯度点) 同士を線で結んである。図から次のことを読み取ることができる。

まず、中心から近い点では比較的正確に視線方向の検出が行える。真正面を見ているか、左右もしくは上下方向にそれぞれ 15° 離れた方向を見ているかは、ほぼ特定できる。逆に、真正面から上下、左右方向とも 30° 以上離れると、格段に検出精度が落ちる。これは、次の3つの理由が考えられる。

- 正面から大きく離れた方向を見る場合、瞳の大部分が瞼の裏などに隠れてしまい、瞳の中心点の検出精度が低下するため、視線方向の検出精度も落ちる。
- 正面方向から離れるに従い、画面上の瞳中心点の検出誤差に起因する視線方向の検出誤差が増大する。
- 眼球の回転角度の限界は上下方向にそれぞれ約 20° 、左右方向にそれぞれ約 30° ずつであると言われており、それを越える方向では、眼球の回転が追従できない。

また、左側を見るときは左目が、右側を見るときは右目が中心となって対象を見る。この傾向は正面から離れるほど強くなり、ついには片方の目が視界の限界を越え、もう一方の目だけで見るようになる。このことは、実験結果にも表れており、左目では左側に比べて右側の、右目では右側に比べて左側の検出精度が悪くなっている。

以上のように、本研究で試作したシステムでは、正面から大きく離れた方向では正確な視線方向を知ることはできないが、正面方向付近ではある程度の精度で視線方向を検出することができる。したがって、上下左右もしくは正面のどの方向を向いているかという程度なら十分特定することができるので、十分感情推定に利用できると考えられる。

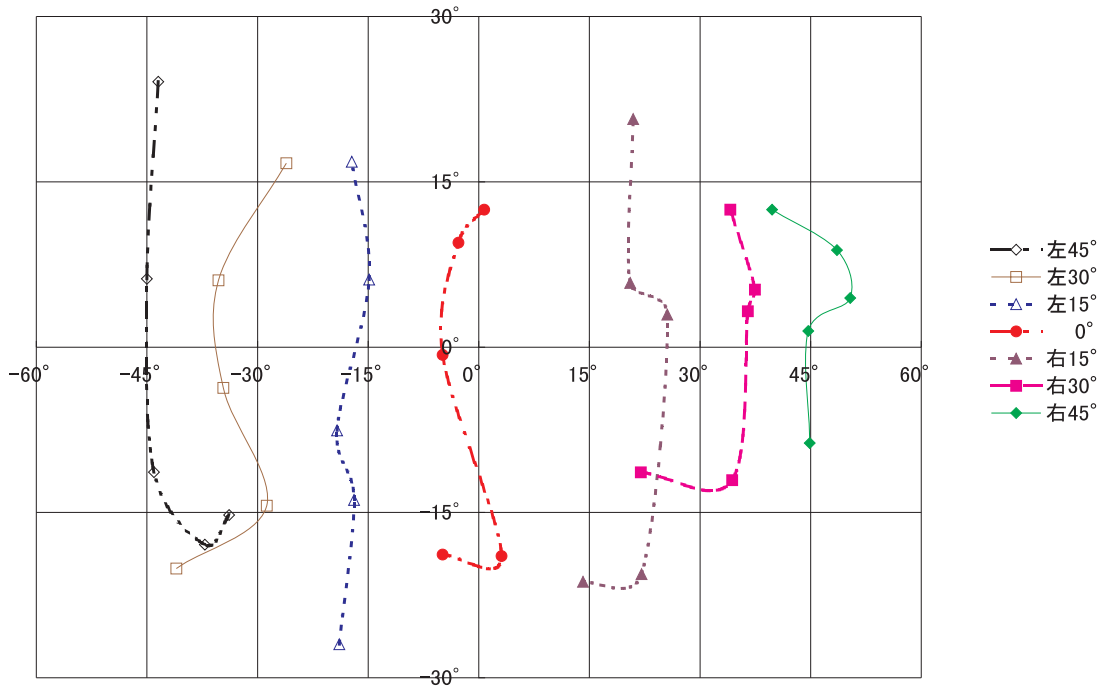


図 5.15: 視線検出実験結果 (被験者 A、右目、同経度点を連結)

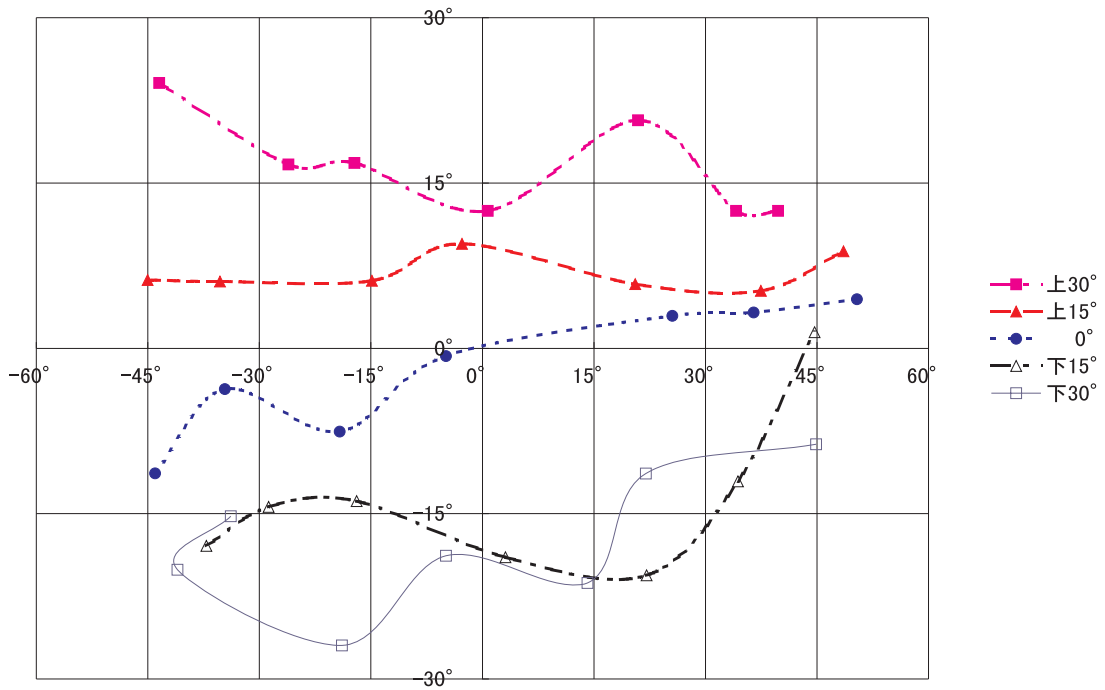


図 5.16: 視線検出実験結果 (被験者 A、右目、同緯度点を連結)

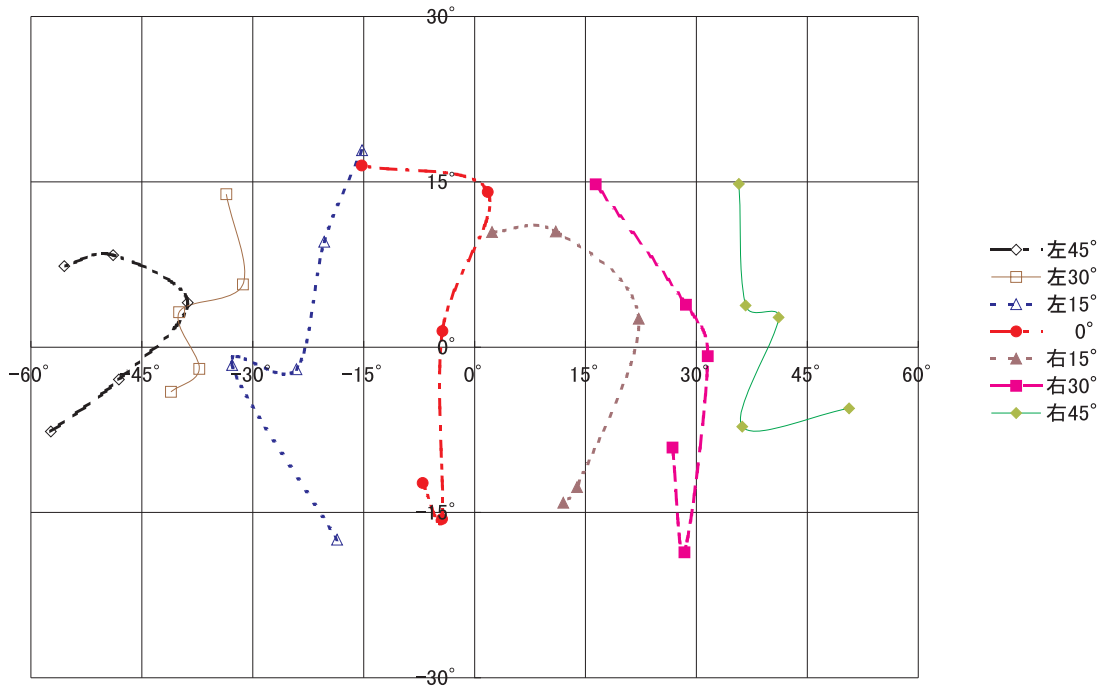


図 5.17: 視線検出実験結果 (被験者 A、左目、同経度点を連結)

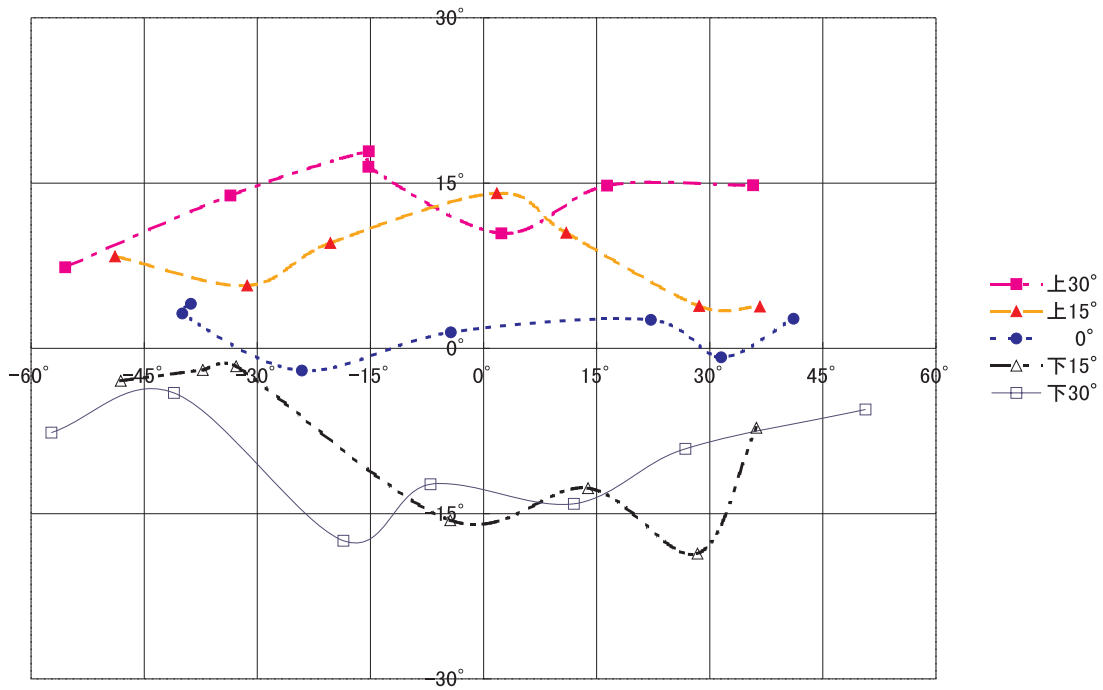


図 5.18: 視線検出実験結果 (被験者 A、左目、同緯度点を連結)

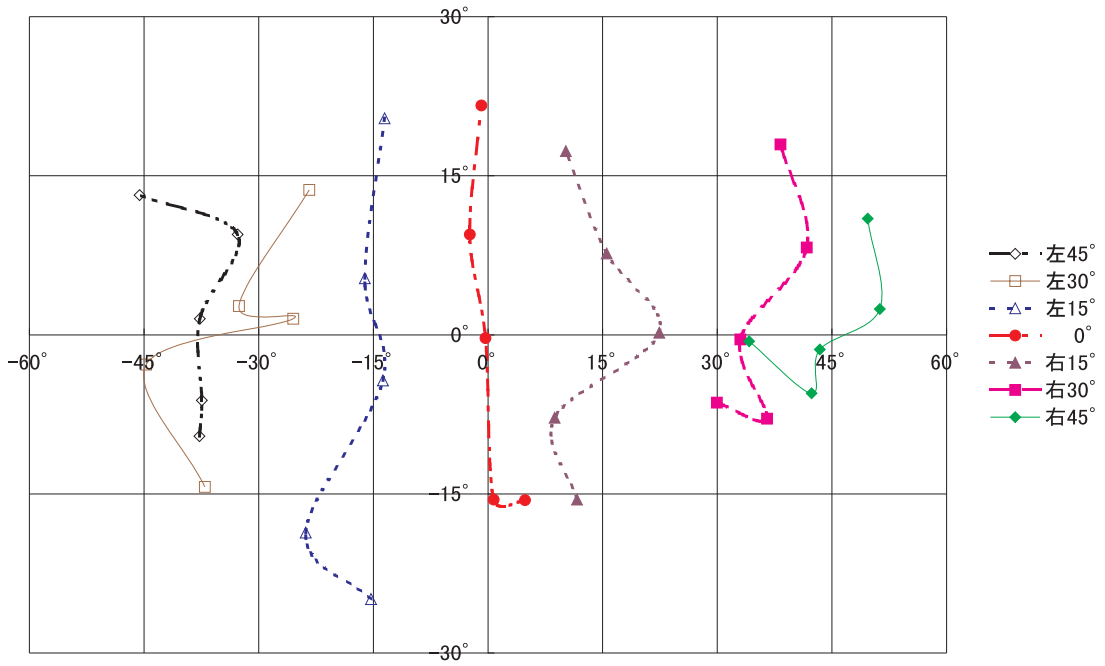


図 5.19: 視線検出実験結果 (被験者 I、右目、同経度点を連結)

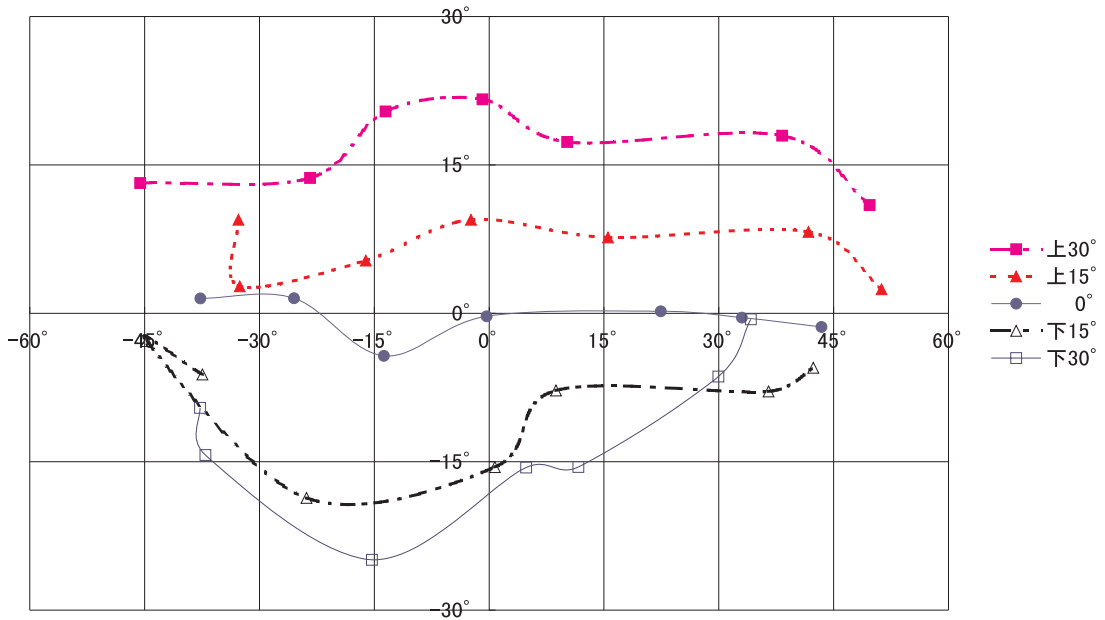


図 5.20: 視線検出実験結果 (被験者 I、右目、同緯度点を連結)

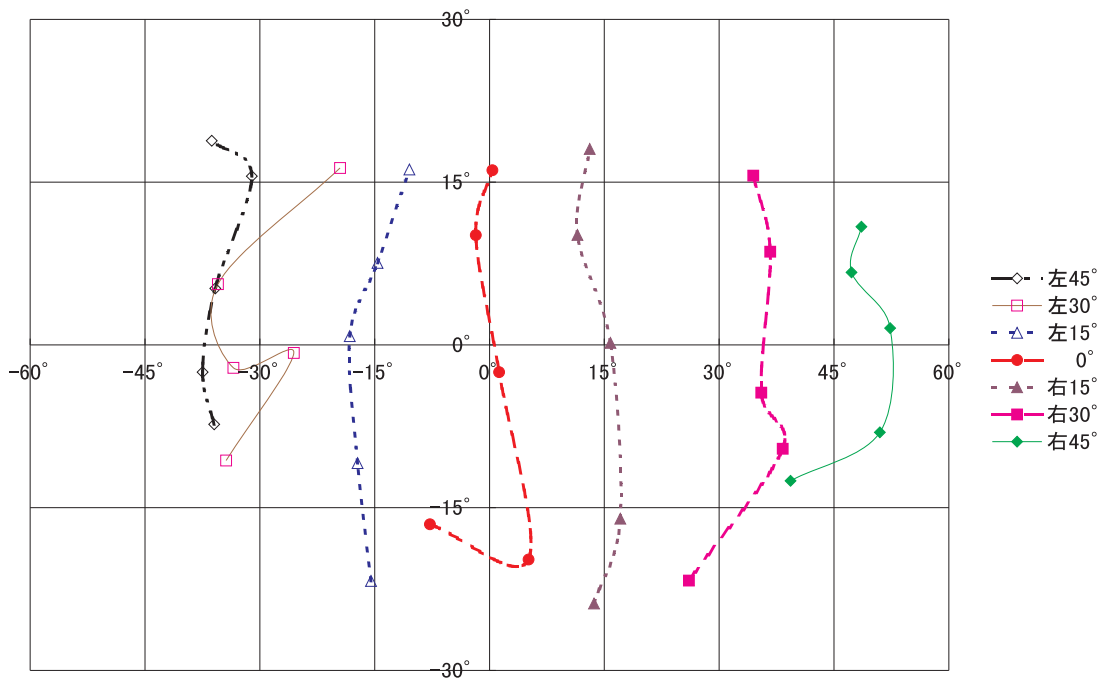


図 5.21: 視線検出実験結果 (被験者 I、左目、同経度点を連結)

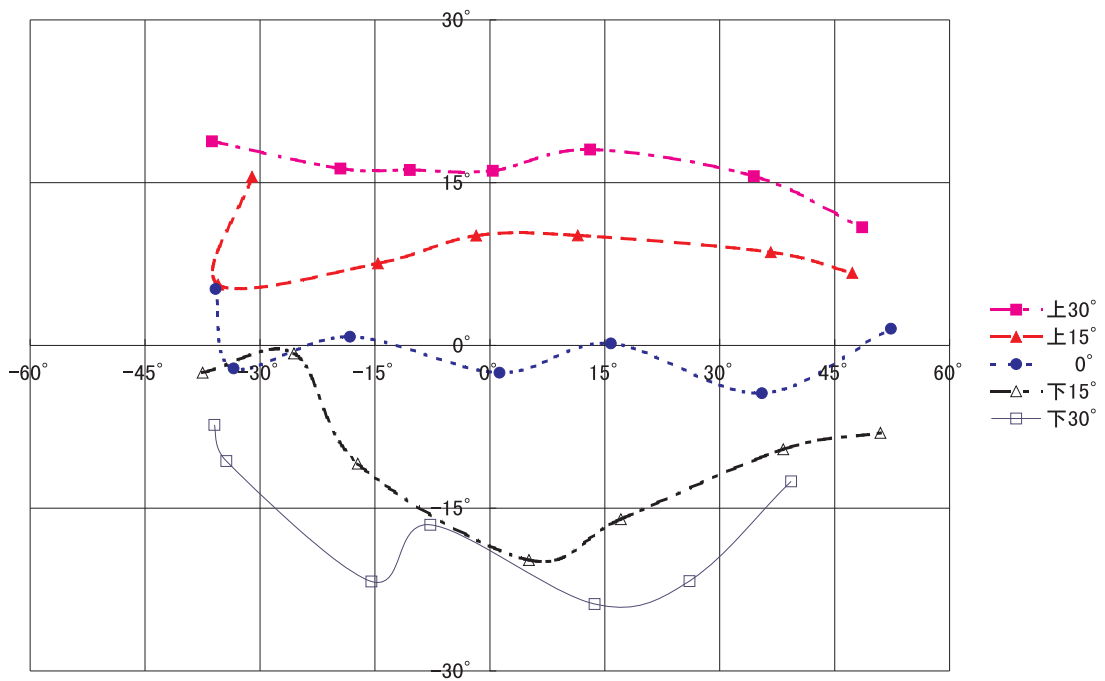


図 5.22: 視線検出実験結果 (被験者 I、左目、同緯度点を連結)

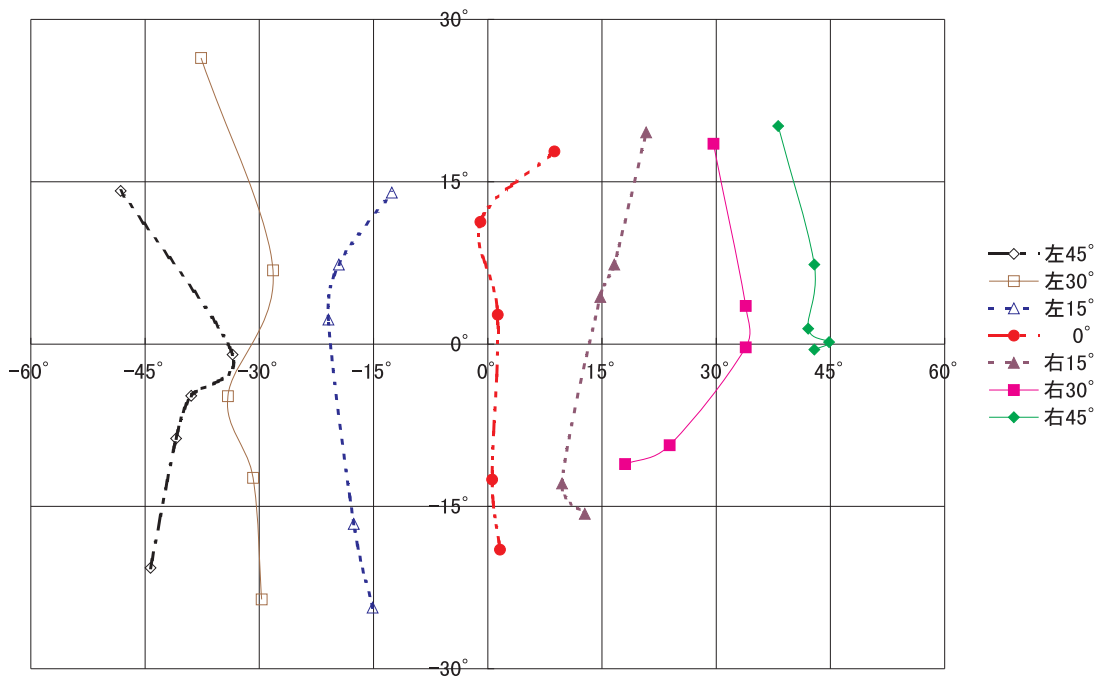


図 5.23: 視線検出実験結果 (被験者 N、右目、同経度点を連結)

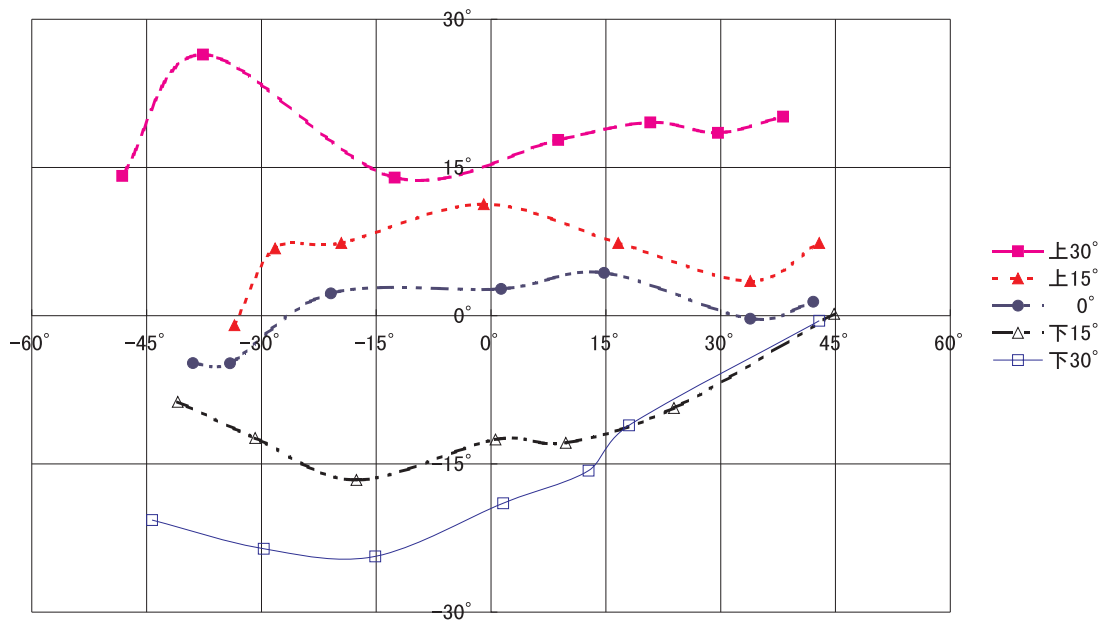


図 5.24: 視線検出実験結果 (被験者 N、右目、同緯度点を連結)

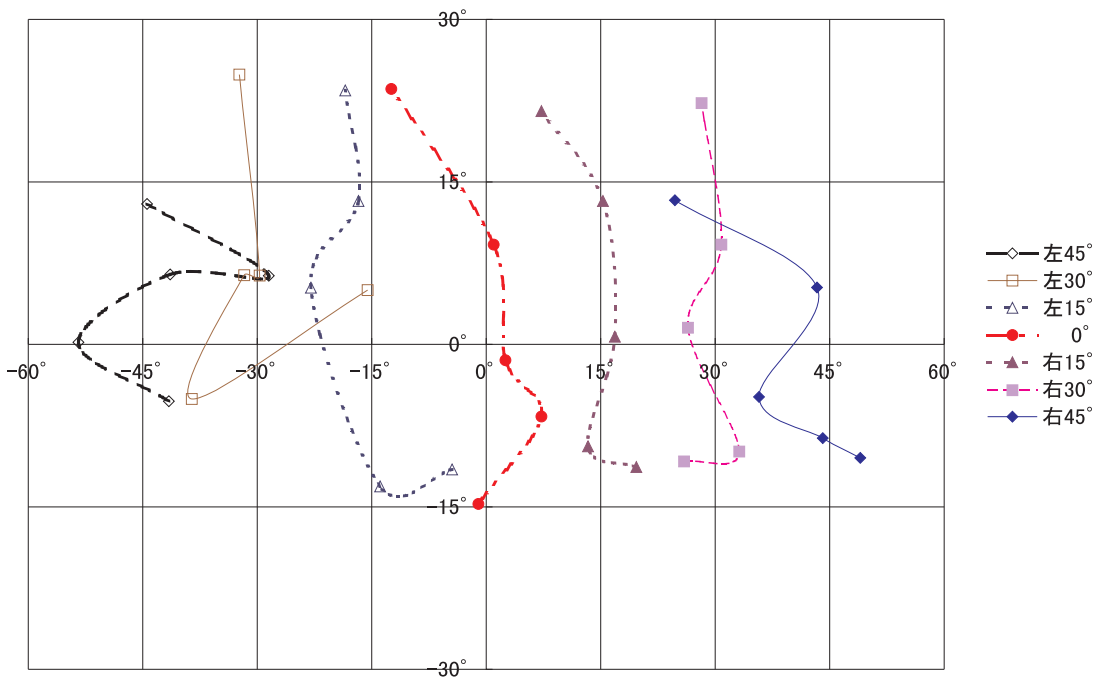


図 5.25: 視線検出実験結果 (被験者 N、左目、同経度点を連結)

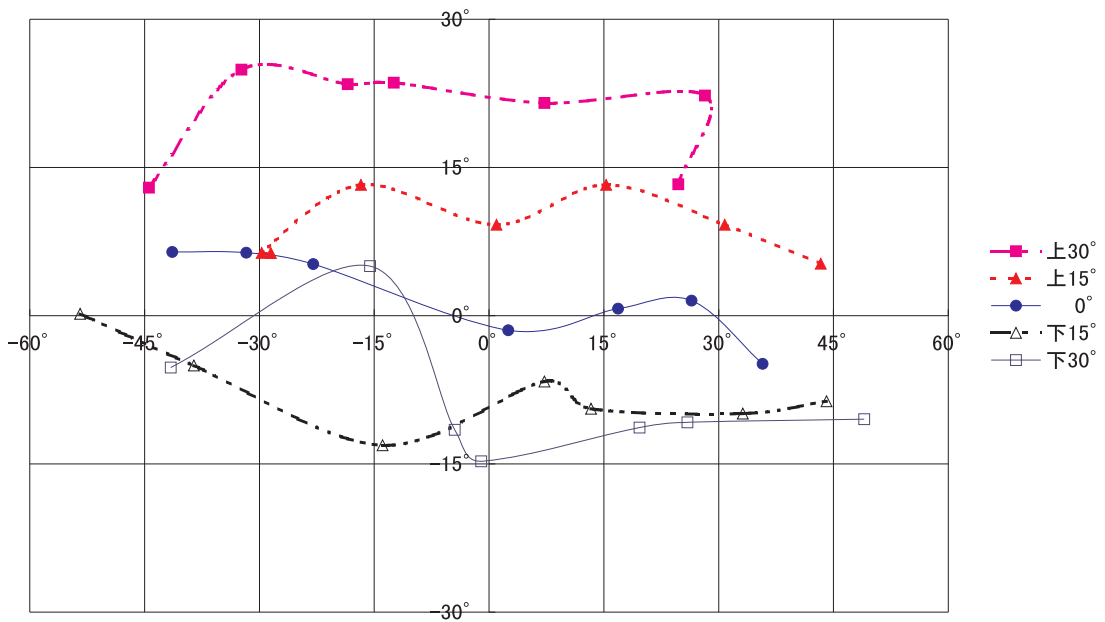


図 5.26: 視線検出実験結果 (被験者 N、左目、同緯度点を連結)

5.3 感情推定に関する考察

試作システムに対して行った、表情認識実験、及び視線方向検出実験の結果、次のことが分かった。

1. 撮影環境を整えた場合、特徴量を 94%の割合で抽出することができる。
2. 特徴量が抽出できた場合、「幸福」及び「驚き」の表情を 90%程度の割合で認識することができる。
3. 正面と上下左右の区別ができる程度の、視線方向の検出ができる。
4. リアルタイムで処理することができる。

以上のように、試作システムでは表情及び視線方向の動的な変化を得ることができるとは。したがって、これらを利用して、3.2.5で述べたような基本 6 感情以外の複雑な感情を推定できると考えられる。ただし、そのためには次に挙げることを今後行う必要がある。

1. 「幸福」、「驚き」以外の表情認識の評価

現在のところ、「幸福」と「驚き」についての評価しか行っていないため、それ以外の「怒り」、「嫌悪」、「恐怖」、及び「悲しみ」についても評価を行う必要がある。これらの表情は意図的な表出が困難であり、「幸福」や「驚き」の場合と同様の方法では評価できないため、評価方法について検討する必要がある。

2. 表情と感情との関連に関するより詳細な調査

現在、表情と感情との対応関係については断片的な情報しか得られていない。感情の推定をシステムに組み込むには、さらに詳細な調査を行い、どのような感情が推定可能か、どのような手法を用いるのが最適か、といったことを吟味しなければならない。

3. 感情推定システムの評価方法の検討

感情の推定には時間的な要素が加わるため、表情認識実験よりもさらに、実験的な表情の表出が困難である。そのため、感情により表出される自然な表情をどのようにして取得するかを検討しなければならない。また、感情は主観的なものであるため、感情の推定結果と被験者の実際の感情をどのようにして比較するかも課題である。

第 6 章 結論

本研究では、ヒューマンマシンインタフェースに表情を利用することにより、人間と機械との間により良いコミュニケーション方法を確立することを目指し、その基礎的研究として、顔画像から表情を認識し感情を推定する手法を提案し、その手法に基づくシステムを試作した。

第 2 章では、お世辞にも使いやすいとは言い難い、コンピュータをはじめとする現在の機械を、人間同士のコミュニケーションを参考にすることで、より身近なものにできることを述べた。人間同士のコミュニケーションでは、言葉以外にもノンバーバルメッセージと呼ばれる、言葉によらない情報伝達が行われている。なかでも、表情は非常に重要な役割を果たしており、人間は相手の表情を動的にとらえることにより、円滑なコミュニケーションを行っている。そのため、ヒューマンマシンインタフェースに表情を利用する際にも、動的な表情認識が必要となる。そこで、本研究は機械が動的な表情認識を行い、人間の感情を推定することを目指していることを述べた。また、ヒューマンマシンインタフェースへの適用を目的としていることから、リアルタイム処理が必要であることについても触れた。

第 3 章では、本研究で用いる感情推定手法についてまとめた。この手法では、心理学分野における研究成果を積極的に利用しており、前半では本研究で利用している表情及び感情に関連する心理学的知見について述べた。具体的には、

- 基本 6 表情
- 表情記述法 FACS
- 顔の 3 領域への分割

などがある。つづいて後半では、本研究で用いる感情推定手法が、

1. 特徴量抽出部
2. 表情認識部
3. 視線方向検出部

4. 感情推定部

の4つの部分に大別できることを述べたあと、各部についてそれぞれ説明を行った。

第4章では、第3章で述べた手法に基づいて試作したリアルタイム感情推定システムについて述べた。試作システムは、感情推定手法の各部に対応した4つのサブシステムによって構成されており、ビデオカメラなどからの動画像を入力画像として、そこからリアルタイムで表情を認識し、視線方向を検出することができる。また、現時点ではごく一部ではあるが感情の推定を行うことができる。

第5章では、試作したシステムを用いた評価実験について述べた。まず、表情認識実験では、試作したシステムによって、比較的表出が容易な「幸福」と「驚き」の表情の認識を行い、この2つの表情が高い確率で認識できることを示した。また、処理に要する時間の測定を行い、システムが十分にリアルタイムで動作可能であることを確認した。次に、視線方向検出実験では、各方向を注視している被験者の顔画像から試作システムにより視線方向の検出を行った。その結果、正面方向と、上下及び左右の区別ができることが分かり、感情推定に利用可能であることを示した。

以上のように、本研究では感情推定を行うための基礎となる、動的な表情を認識できるシステムを試作した。しかし、現在のシステムには次のような問題点がある。まず、顔が正面を向いていなければならず、照明環境も整える必要があるなど、入力画像に多くの制約が生じる。このことを解決するには、新しい画像処理アルゴリズムを考案する必要があるが、リアルタイム処理との兼ね合いが難しい。また、現在のシステムでは、複雑な感情の推定を行うことができない。これにはまず、表情がどのように変化した場合に、どのような感情であるのかを調査しまとめなければならない。いずれにしても、現在のシステムを基礎としてさらに研究を進める必要がある。

新しい世紀に向けて、あたかも人間と対話しているかのように自然に操作できる機械の出現が望まれている。

謝 辞

本研究を進めるにあたり、研究の全般にわたって暖かいご指導を頂きました吉川榮和教授に、深く感謝致します。

本研究を進めるにあたり、直接ご指導頂き、幾度となく適切なお助言を頂きました下田宏助手に、心より感謝致します。

本研究を進めるにあたり、折りにふれ貴重なお助言を頂きました手塚哲央助教授に、心より感謝致します。

本研究を進めるにあたり、数々の助力を頂きました、修士1回生の楊大昭君に深く感謝致します。

また、貴重な時間を割いて実験に参加し、表情筋を酷使して頂いた、被験者の皆様にも心より感謝致します。

最後に、研究を進めるにあたり、何かとお世話頂いた谷友美秘書、ならびに、長い夜を共に過ごした吉川研究室の皆様に、深く感謝致します。

参考文献

- [1] 黒川：ノンバーバルインタフェース，オーム社 (1994)
- [2] 坂井, 長尾, 金出：計算機による顔写真の解析, 信学論, **56-D**, pp.226-233 (1973)
- [3] 大久保, 渡辺, 山田：色情報を用いたオプティカル・スネークによる口唇運動抽出, 第14回ヒューマン・インタフェース・シンポジウム論文集, pp.175-179 (1998)
- [4] 小林, 橋本：まばたきを手掛かりとした顔の特徴点自動抽出とモデリング, 信学技法, **HC94-53**, pp.9-16 (1994-10)
- [5] 宋, 李, 徐, 辻：部分特徴テンプレートとグローバル制約による顔器官特徴抽出, 信学論 (D-II), **Vol.J77-D-II, No.8**, pp.1501-1609 (1992)
- [6] 佐々木, 赤松, 末永：顔画像認識のための色情報を用いた顔の位置合わせ法, 信学技法, **IE91-2**, pp.9-15 (1991)
- [7] 小林, 中村, 森島, 原島：顔画像からの唇の特徴点抽出法, 信学全代春, **D-329** (1990-3)
- [8] 間瀬：オプティカルフロー抽出による表情筋の動作検出, 信学技法, **PRU89-128**, pp.17-24 (1990)
- [9] Essa, Pentland: Facial Expression Recognition using a Dynamic Model and Motion Energy, M.I.T Media Laboratory Perceptual Computing Section Technical Report, **No.307** (1995)
- [10] 坂口, 森島：空間周波数情報に基づく基本表情の実時間認識, 第2回知能情報メディアシンポジウム論文集, pp.75-82 (1996)
- [11] 坂口, 大谷, 岸野：Wavelet 変換による顔動画画像の特徴抽出と表情認識への応用, 信学技法, **IE94-147, PRU94-147** (1995)
- [12] 小林, 原：ニューラルネットによる人の基本表情認識, 計測自動制御学会論文集, **29**, pp.112-118 (1993)

- [13] 坂口, 大谷, 岸野: 隠れマルコフモデルによる顔動画像からの表情認識, テレビ誌, **Vol.49, No.8**, pp.1060-1067 (1995)
- [14] 小林, 原, 池田, 山田: リカレントニューラルネットワークによる動的な基本表情認識, 信学技法, **HC-92-59**, pp.11-16 (1993-01)
- [15] 伊藤ら: 岩波講座 認知科学 6 情動, 岩波書店 (1994)
- [16] 吉川, 益谷, 中村編: 顔と心・顔の心理学入門, サイエンス社 (1993)
- [17] Russell and Bullock: Multidimensional scaling of emotinal facial expressions: Similarity from preschoolers to adults, *Journal of Personality and Social Psychology*, **48**, pp.1290-1298 (1985)
- [18] 福井: 感情の心理学, 川島書店 (1990)
- [19] Ekman and Friesen 著, 工藤訳編: 表情分析入門, 誠信書房 (1987)
- [20] Ekman and Friesen : The Facial Action Coding System, Consulting Psychologists Press (1978)
- [21] 西尾, 小山, 中村: 笑いの表情の分類における注視点の影響, 信学技法, **PRMU97-137, HIP97-18**, pp.53-60 (1997-11)
- [22] 蒲地, 吉川, 赤松: 変化速度は表情認知に影響するか? -動画刺激を用いた顔表情認知の時間特性の解明-, 信学技法, **HCS98-34** , pp.17-24 (1998-10)
- [23] 戸田: 認知科学選書 24 感情, 東京大学出版会 (1992)