

Title	Global Facial Modelと画像の粗さ測度に基づく表情顔 画像からの表情情報抽出に関する研究
Author(s)	高橋, 公生
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1134
Rights	
Description	Supervisor:小谷 一孔, 情報科学研究科, 修士

修士論文

Global Facial Model と画像の粗さ測度に基づく 表情顔画像からの表情情報抽出に関する研究

指導教官 小谷一孔 助教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

高橋 公生

1998年2月13日

要旨

本研究の目的は、顔画像から表情変化によって発生する物理的な表情特徴を抽出する手法を構成することである。一般に、人は顔を認識する際に目と口に注目するという傾向があることが知られている。その傾向を確認するために、人が表情認識を行う際にある部分に注目するのか否か、また注目する部分があるとしたらどの部分であるか、その傾向を測定する。その結果を用い、顔画像から物理的な表情特徴を抽出する手法を構成する。

抽出手法は、顔画像の2次元平面上の濃淡値を高さとして画像を3次元曲面とみなし、その形状特徴を Hausdroff 次元であらわし、計算された次元値によって画像の領域を分割する。分割の後、領域の位置、大きさと Hausdr 次元値のパターンの比較によって目領域の候補(目領域候補ペア)を抽出し、それぞれの領域ペアとの位置関係から口領域を推定する。

次に、抽出された目、口領域から表情情報を抽出する。表情変化をより詳細に捉えるため、それぞれの顔部品の形状を輪郭線として抽出する手法の構成を目指す。

目次

1	はじめに	1
1.1	研究の背景	1
1.2	研究の目的	2
1.3	本論文の概要と構成	2
2	ヒトの表情認識行動	4
2.1	表情認識行動の特徴	5
2.2	表情認識行動特性測定	5
2.2.1	実験方法及び条件	6
2.2.2	測定結果	14
2.3	まとめ	22
3	顔画像領域分割	23
3.1	領域分割手法の一般	23
3.2	顔画像の特徴抽出, 解析	24
3.2.1	従来研究の問題点	24
3.2.2	フラクタル特徴解析による顔画像の領域分割	25
3.3	Hausdroff (フラクタル) 次元による曲面形状 (パターン) 解析	26
3.3.1	Hausdroff 次元	26
3.3.2	画像における Hausdroff 次元推定	27
3.3.3	画像において Hausdroff 次元が示す特性	29
3.4	Hausdroff 次元による領域分割	38
3.4.1	対象画像範囲	38

3.4.2	領域分割アルゴリズム	38
3.5	領域分割の限界	41
4	Global Facial Modelに基づく顔部品推定	52
4.1	Global Facial Model (GFM)	52
4.1.1	顔部品の位置関係の測定	53
4.1.2	推定領域設定	56
4.1.3	GFM の設計	56
4.2	目領域候補ペア推定	59
4.2.1	目領域推定アルゴリズム	59
4.2.2	目領域推定実験条件	61
4.2.3	実験結果	65
4.2.4	考察	65
4.3	口領域推定処理	69
4.3.1	口領域の推定アルゴリズム	69
4.3.2	口領域推定実験条件	71
4.3.3	実験結果	71
4.3.4	考察	71
4.4	回転・大きさの変化に対するロバスト性の検討	74
4.5	まとめ	77
5	顔部品画像からの物理的的表情情報抽出	78
5.1	従来研究とその問題点	78
5.2	フィルタ処理による物理的的表情情報抽出	78
5.3	実験条件	79
5.3.1	輪郭抽出法に関する検討	79
5.3.2	フィルタの種類に関する検討	81
5.3.3	フィルタのパラメータ決定に関する検討	84
5.4	顔部品の輪郭抽出	88
5.4.1	目の輪郭抽出	88
5.4.2	口の輪郭抽出	91

5.5	輪郭抽出結果	92
5.6	回転・大きさの変化に対するロバスト性の検討	94
6	顔部品輪郭形状情報を用いた表情判別	97
6.1	顔部品輪郭形状情報の記述	97
6.1.1	表情識別実験の方法と条件	100
6.1.2	実験結果	102
6.2	まとめ	103
7	むすび	104
	謝辞	107
	文献	108

第 1 章

はじめに

1.1 研究の背景

人間の記憶・計算能力を補い，増幅するための支援ツールとしてのコンピュータが，コミュニケーションツールの一つとして重要な位置を占めつつある．音声，文字といったコミュニケーションメディアは，既に電話や電子メールという形で実現されており，さらに音声や文字以外に画像を用いたテレコミュニケーションの機会が増えている．これは，コンピュータをはじめとするコミュニケーションツールの普及とそれを取り巻く情報環境の高度化によるものである．

コミュニケーションによる人間への情報入力，五感（視覚，聴覚，嗅覚，触覚，味覚）といわれる各感覚器官を経由する [15]．そして入力される情報は，言語的なものと非言語的なものに大別できる．ノンバーバルコミュニケーションではメッセージのうち，38[%]が音声，55[%]が表情によるものであると心理学者である Mehrabian [1] は実験的に示しており，また表情は感情の伝達のみならず，相手の同定と状態の把握，心理的距離の設定を可能とするなど，コミュニケーションにおける重要な役割を果たしている．また表情は音声と異なり，相手が見える状態であれば中断なく情報を受けことができ，相手からの意識的な情報発信を必要としないという特徴がある．このような表情情報を伝達するだけでなく，コミュニケーションツールが利用できるようになれば，

- 1) コミュニケーション効率の向上，
- 2) 非言語情報が持つ機能の利用，

を高度に実現することができるであろう．そして表情情報を扱うためには，表情情報を抽

出す表情解析技術の開発が必要である。

1.2 研究の目的

本研究の目的は、顔画像から表情変化によって発生する物理的な表情特徴を抽出する手法を構成することである。一般に、人は顔を認識する際に目と口に注目するという傾向があることが知られている。その傾向を確認するために、人が表情認識を行う際にある部分に注目するのか否か、また注目する部分があるとしたらどの部分であるか、その傾向を調査する。その結果を用い、顔画像から物理的な表情特徴を抽出する手法を構成する。

抽出は、顔画像の2次元平面上の濃淡値を高さとして画像を3次元曲面とみなし、その形状特徴を Hausdroff 次元であらわし、計算された次元値によって画像の領域を分割することから始まる。分割の後、領域の位置、大きさと Hausdroff 次元値のパターンの比較によって目領域の候補(目領域候補ペア)を抽出し、それぞれの領域ペアとの位置関係から口領域を推定する。

次に、抽出された目、口領域から表情情報を抽出する。表情変化をより詳細に捉えるため、それぞれの顔部品の形状を輪郭線として抽出する手法の構成を目指す。特に画像の大きさや傾きの正規化を必要としない手法の構成を目指すため、それらを必要としない、画像のフィルタリングによる手法を構成する。

1.3 本論文の概要と構成

本論文は、全6章からなる。以下に各章の内容を記す。

第2章では、表情の物理的特徴抽出手法構成の指針を得るために、ヒトの表情認識における傾向を測定する。測定は顔画像提示中の視点停留点をアイマークレコーダを用いて求めることによって行う。そして停留点の分布を求め、傾向について検討する。

第3章では、第2章の結果から、目、口の顔部品から表情の物理的特徴を抽出する手法について述べる。本研究では、画像の濃淡パターンの特徴を画像の粗さとして扱うことにより領域分割する。この画像の粗さについての説明と特性について述べる。そして画像の粗さ測度を用いた画像の領域分割アルゴリズムについて述べる。また画像の粗さ測度の計算手法の特性と、領域分割手法の特性について検討する。

第4章では、分割された顔部品候補領域から目、口領域を得るために用いる Global Facial

Model の設計を行う．そして GFM を用いた顔部品領域の抽出手法を構成し，表情画像に適用した結果について述べる．

第 5 章では，目，口の部分顔画像から物理的表現情報を抽出する手法の構成と，その適用結果について述べる．

第 6 章では，むすびとして本研究のまとめを行う

第 2 章

ヒトの表情認識行動

顔画像から表情情報を抽出するためには、心理的な表情特徴の解析と物理的な表情特徴の解析という2つのアプローチが必要である(図 2.1)。心理的な表情特徴の解析では、ある表情を表す印象語を選定し表情を分類する。そして同じ心理的特徴を示す表情画像に常に現れる物理的特徴を抽出し、表情空間を構築して心理的な表情特徴を抽出する。また物理的な表情特徴の解析では、表情情報が豊富に現れる部分を選定し、構造を解析することで、表情の種類、強度の情報を抽出する。この心理的、物理的特徴の関係を解析することで表情情報の抽出ができる。

本研究では、物理的な表情特徴を抽出するために必要な画像解析手法の構成にとりくむ。

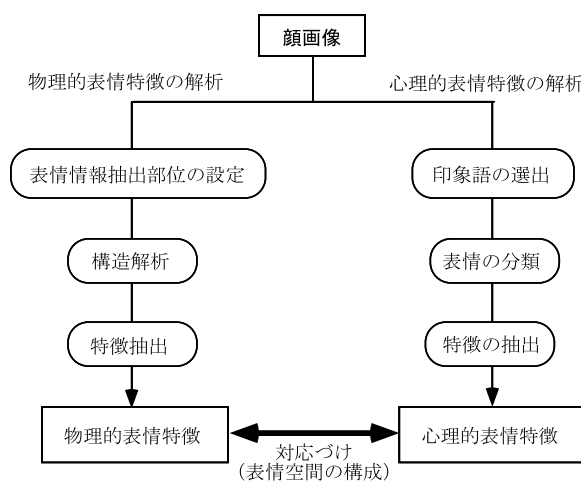


図 2.1: 表情解析のアプローチ

2.1 表情認識行動の特徴

入力画像は表情変化や頭部の移動，背景の変化等によってその濃淡パターンは様々に変化する．その中から表情情報を抽出するためには，顔画像の中で表情情報が現れる部分を求めることが必要である．

従来，視点を統制した状態で表情判断を行わせることによって表情情報が現れる部位を求めるという研究がなされている．その結果，口，目に注視した場合は [80%] ~ [100%] で基準と同じ表情判断がなされたという報告がある [1]．しかし注視点を統制することは自然な状態とはいえず，また基準に対する表情判断は，表情認識とは異なる知覚が行われていることも考えられる．

ところで，ヒトの身体的な構造から，視力は網膜中心部の狭い範囲でのみ大きく，周辺部では急激に低下する．すなわち我々は網膜中心部分でしかものを詳しく見ることができない．このため眼球は絶えず上下左右に運動し，見ようとする外界の物体の像を網膜の中心に持ってくる必要がある [5]．この性質を利用して，小濱ら [16] はカタカナを表示した画像の認識過程における注視点の位置と認識結果との関連を実験的に求めており，その結果，注視点を中心としてその周囲の比較的狭い範囲から正確な情報を獲得することが明らかにされた．

そこで，本研究では表情の認識における視点の位置を求め，注視点を求めることによってより自然な状態で表情情報が現れる部位を求めることができると考えられる．

2.2 表情認識行動特性測定

本節では，画像解析手法を構成するために，まず表情情報が豊富に現れる部分，すなわち表情認識においてヒトが注目する部分があるのか否か，その傾向を測定する．一般に，ヒトが表情を認識する際には，目や口という顔器官（以下顔部品）に注目することが知られている．表情認識におけるヒトの傾向について，表情顔がモニタ上に提示される場合の傾向は未だ示されていないようである．そこで，本研究では，コンピュータのモニタ上に提示した顔画像の表情認識におけるヒトの傾向を測定し，その結果から物理的特徴を抽出する部分を決定する．

表 2.1: 視点追跡実験利用機器

名称	用途	型番	製造者
アイマークレコーダ	視点位置測定	両眼眼球運動測定装置	竹井機器工業
処理用コンピュータ	視点位置計算・記録	PC-9821Ap3	NEC
VHS ビデオデッキ	視点画像記録	SLV-BS511	SONY
画像提示用モニター	表情画像表示	54T	NANA O
画像提示用コンピュータ	表情画像提示	PowerMacintosh 9600 / 200MP	Apple Computer

2.2.1 実験方法及び条件

測定は、モニターに表情画像を提示し、画像提示中の被験者の視点をアイマークレコーダを用いて測定し、注視点の分布を求めることによって行う。また被験者には表情画像の表情を回答させ、注視点の分布との関連を求める。

【A】使用機器・実験システム

実験に用いた機器は表 2.1に示す。実験システムの構成を図 2.2に示す。

被験者の視点は、被験者が装着するゴーグルに取り付けられたセンサーによって測定される。またゴーグルに取り付けられたビデオカメラによる映像を用いて、視点位置のキャリブレーションを行う。視点の測定は、62.5[点/秒]で行われる。また被験者が正面を見たときに、画像の中心を見るように高さを調整した。

【B】被験者

被験者は、20代の大学院生10名(男性9名,女性1名)で、女性を含む9名が眼鏡装着者であった。

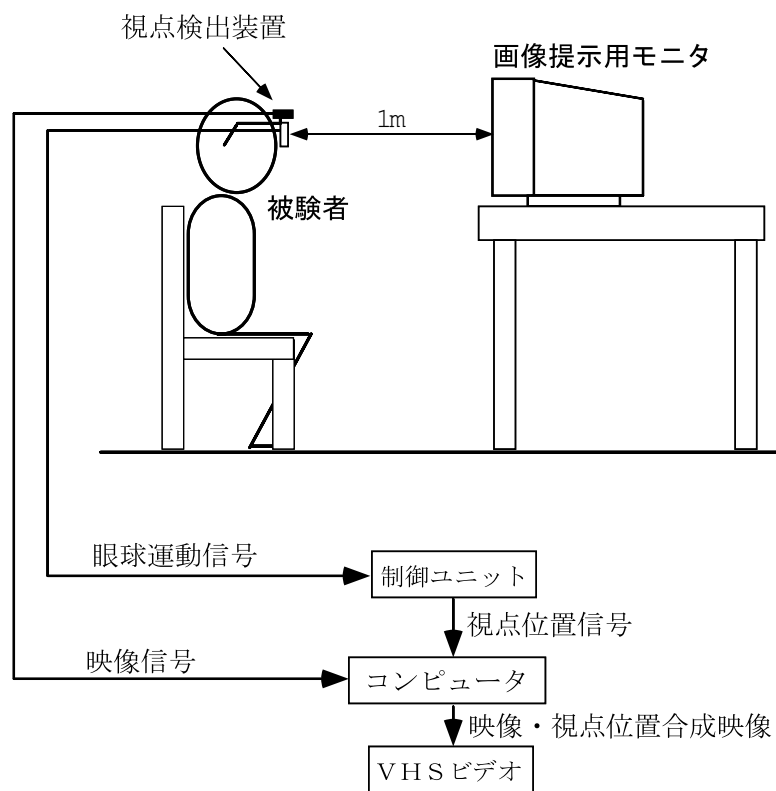


図 2.2: 実験システム構成図

表 2.2: 画像提示順・画像表情

提示番号	1	2	3	4	5	6
画像番号	11	12	13	21	22	23
人物	1	1	1	2	2	2
表情	無表情	笑い	怒り	無表情	笑い	怒り
提示順	7	8	9	10	11	12
画像番号	31	32	33	41	42	43
人物	3	3	3	4	4	4
表情	無表情	笑い	怒り	無表情	怒り	笑い

【C】提示画像

提示した画像は，20代の男性2名女性2名の合計4名，それぞれの人物について無表情，笑い，怒りの3表情（合計12枚）である．提示した画像を図2.3~図2.6に示す．画像の大きさは実際の人物の大きさと同じ程度の大きさに合わせ，552 × 680[pixels]とした．なお，画像の色数はフルカラーとした．

【D】提示画像順

提示画像は，表2.2に示す順に提示した．それぞれ提示時間は2秒間とした．ヒトは画像の認識を行うためには数百ミリ秒程度の提示で充分であり，長時間提示した場合，初期の認識が終了した後は，画像認識とは異なる状態で画像を観察しているという報告がある[16]．今回はヒトの表情認識における傾向を測定することが目的であるため，短時間提示とする．また予備実験では画像提示を1秒間としていたが，事前情報を与えないという目的から，画像提示前の視点位置を指定しておらず，また画像提示のタイミングを一定にしていなかったために画像を見落とすという例が見られたため，2秒間とした．

图 2.3: 提示画像 (人物 1)

图 2.4: 提示画像 (人物 2)

图 2.5: 提示画像 (人物 3)

图 2.6: 提示画像 (人物 4)

【 E 】 実験進行

実験の進行は(図 2.7)に示す。被験者は順に提示される画像を観察し、画像提示後に表情を回答する。表情の回答は複数選択ではなく、印象語を自由に記述させた。そして回答が終了したことを確認してから、次の画像を提示した。

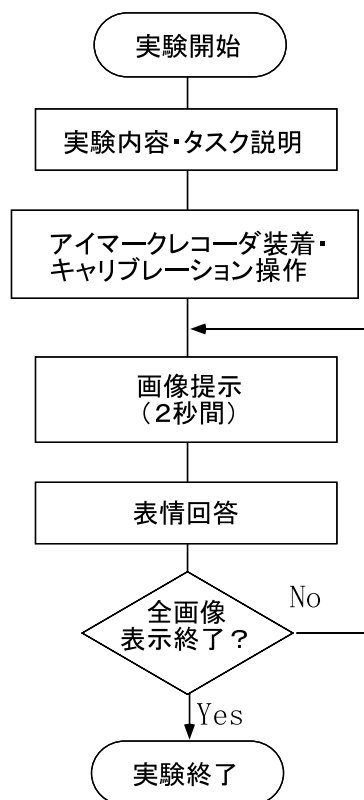


図 2.7: 表情認識傾向測定実験の進行

2.2.2 測定結果

本節では、アイマークレコーダを用いた視点測定によって得られた視点位置から、注視点を計算によって求め、顔を領域に分割し、それぞれの領域への注視点の分布を求める。

【A】注視点の定義

アイマークレコーダによって測定された視点は、それだけでは注視点を示しておらず、視点から注視点を求める必要がある。本研究では、ヒトの視覚特性と類似の研究における定義から [5] [17] [18]，視角 1 [度] の範囲に連続して 200[ミリ秒] 以上の間視点が存在する場合、その中心を注視点とした。

【B】顔部品領域の設定

注視点の分布状態を求めるため、顔画像を顔部品領域とそれ以外の領域に分割した。提示した画像は、大きさや傾きの正規化を行っていないため、領域の大きさや位置を変える必要がある。そのため、画像上の特徴点を基にしたパラメトリックな領域の設定を行う。

設定した領域は、目、鼻、口、頬とした(図 2.8)，それらを決定する座標は式 (2.1)~式 (2.16)によって求められる。基準となる点は、目や口の両端、小鼻の左右など、比較的明瞭な点を用いることとした。また基準点は画像ごとに手作業で求めている。以上の結果決定された領域の例を図 2.9に示す。

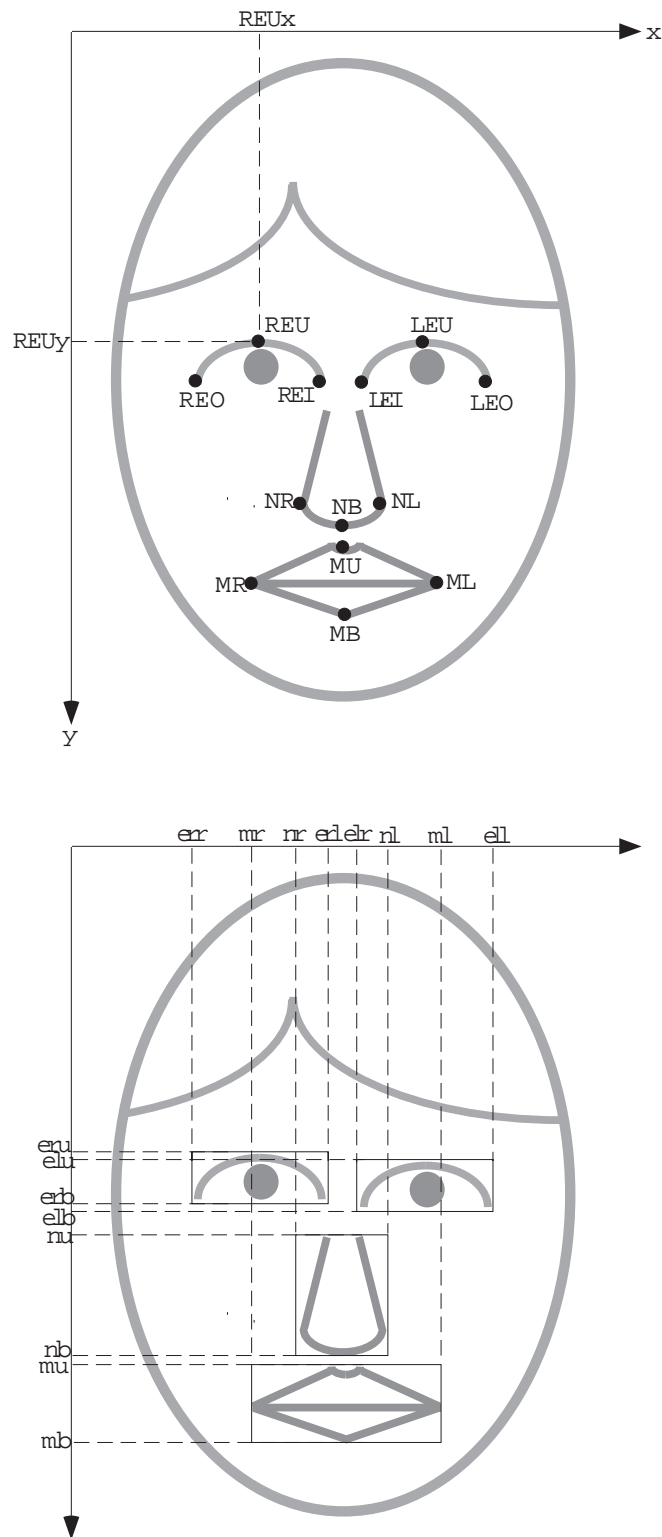


図 2.8: 顔部品領域の基準点と顔部品領域設定

$$el_u = LEU_y - 1/10 \max(LEO_x - LEI_x) \quad (2.1)$$

$$e_l b = e_l u + 2/3 \max(LEO_x - LEI_x, REI_x - REO_x) \quad (2.2)$$

$$e_l l = LEO_x + 1/6(LEO_x - LEI_x) \quad (2.3)$$

$$e_l r = LEI_x - 1/6(LEO_x - LEI_x) \quad (2.4)$$

$$e_r u = REU_y - 1/10 \max(REI_x - REO_x) \quad (2.5)$$

$$e_r b = e_r u + 2/3 \max(LEO_x - LEI_x, REI_x - REO_x) \quad (2.6)$$

$$e_r l = REI_x + 1/6(REI_x - REO_x) \quad (2.7)$$

$$e_r r = REI_x - 1/6(REI_x - REO_x) \quad (2.8)$$

$$n_u = \max(e_l b, e_r b) + 5 \quad (2.9)$$

$$n_b = NB_y + 1/2(MU_y - NB_y) - 5 \quad (2.10)$$

$$n_l = NL_x + 1/6(NL_x - NR_x) \quad (2.11)$$

$$n_r = NR_x - 1/6(NL_x - NR_x) \quad (2.12)$$

$$m_u = n_b + 10 \quad (2.13)$$

$$m_b = MB_y + 1/7(LEI_x - REI_x) \quad (2.14)$$

$$m_l = ML_x + 1/7(LEI_x - REI_x) \quad (2.15)$$

$$m_r = ML_x - 1/7(LEI_x - REI_x) \quad (2.16)$$

図 2.9: 表情認識傾向分析のための顔部品領域 (例)

【B】注視点分布

目，鼻，口，頬を対象とした注目点の分布状態を図 2.10 に示す．また表情の回答と，提示画像の表情（正解）を表 2.3 に示す．表情の正答率が 70 % を越えており，提示された画像の表情をほぼ正しく認識していることがわかる．注目点の分布は，領域への注目点の密度によって示している．これから，目や口に比較して鼻領域に注目していることが示されている．しかし，ビデオテープに記録された視点の移動を観察すると，視点の移動が左右の目の間を移動し，鼻領域を経由して口領域内を移動する運動が見られ，また左右の目領域を探るように往復移動し，その際に鼻領域を通過する運動が見られた（図 2.12）．その結果，計算上注目点が鼻領域に存在することになったと考えられる．現在は視点の移動状態を考慮していないため，鼻を注目した結果現れる注目点と，鼻領域上を移動する際に計算上発生した注目点との分離ができない．また表情変化による鼻画像の変化は少ないことから，鼻領域を除去して分布密度を求めた（図 2.11）．これより，目領域への注目が最も高く，次に口領域に注目する傾向があることがわかる．

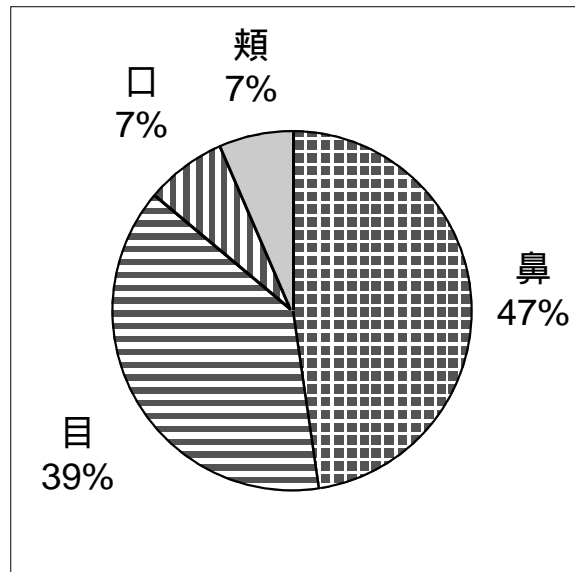


図 2.10: 顔部品領域への注視点分布 (全領域)

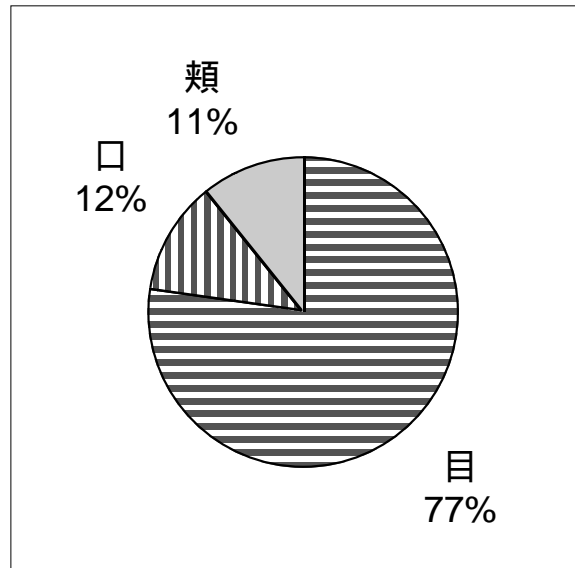


図 2.11: 顔部品領域への注視点分布 (鼻領域を除く)

図 2.12: 表情認識における視点の軌跡

表 2.3: 提示画像の表情と被験者による表情の回答

画像番号	表情 (正解)	被験者 1	被験者 2	被験者 3	被験者 4	被験者 5
11	無表情	悲しみ	無表情	無表情	無表情	無表情
12	笑い	笑い	笑い	笑い	笑い	笑い
13	怒り	怒り	悲しみ	怒り	怒り	無表情
21	無表情	悲しみ	無表情	無表情	怒り	無表情
22	笑い	笑い	笑い	笑い	笑い	笑い
23	怒り	怒り	怒り	怒り	無表情	怒り
31	無表情	悲しみ	悲しみ	無表情	微笑	無表情
32	怒り	笑い	怒り	怒り	怒り	怒り
33	笑い	怒り	笑い	笑い	笑い	笑い
41	無表情	微笑	無表情	無表情	無表情	無表情
42	怒り	悲しみ	嫌悪	悲しみ	悲しみ	怒り
43	笑い	笑い	笑い	笑い	笑い	笑い
画像番号	表情 (正解)	被験者 6	被験者 7	被験者 8	被験者 9	被験者 10
11	無表情	無表情	無表情	無表情	困惑	無表情
12	笑い	笑い	笑い	笑い	笑い	笑い
13	怒り	怒り	悲しみ	怒り	悲しみ	怒り
21	無表情	無表情	無表情	無表情	困惑	無表情
22	笑い	笑い	笑い	笑い	笑い	笑い
23	怒り	怒り	怒り	悲しみ	怒り	怒り
31	無表情	無表情	無表情	無表情	悲しみ	無表情
32	怒り	怒り	笑い	怒り	怒り	怒り
33	笑い	笑い	笑い	笑い	笑い	笑い
41	無表情	無表情	無表情	悲しみ	悲しみ	無表情
42	怒り	悲しみ	悲しみ	怒り	怒り	怒り
43	笑い	笑い	笑い	笑い	笑い	笑い

2.3 まとめ

本章では、物理的特徴抽出手法を構成するために、顔画像中に表情情報が豊富に現れる部分を実験によって求め、その結果について述べた。本研究では、アイマークレコーダを用いた視点追跡によって画像中の注目点を測定し、その分布の傾向を求めた。

4名(男性2名,女性2名)の人物の3表情(無表情,笑い,怒り)の画像を用い、10名(男性9名,女性1名)の被験者に対して表情の回答をさせ、画像提示中の視点の位置を測定し、視覚特性から決定される時間と位置の範囲から注目点を計算によって求めた(図2.10)。その結果、鼻に注目する傾向がみられたが、視点の移動の傾向から、鼻を注目した結果発生する注目点と、目から口に視点が移動する際に、計算上発生した注目点が存在することが考えられた。鼻領域は表情変化の影響を受けにくいいため、鼻を対象外として分布傾向を求めたところ、目、口領域へより注目している傾向が見られた(図2.11)。また表情の正答率を見ると、笑いの表情は全回答中笑い以外の表情と回答したものが1例だけであった。また怒りの表情を怒り以外の表情とした回答は14例、無表情を無表情以外とした回答は12例あった。これらから、笑いはほぼ正しく認識できているが、無表情と怒りは悲しみや困惑として認識するという傾向が見られた。これは、怒りや悲しみは笑いに比較して明確な区別が困難であるということを示している。また、正答率が70%と高いことから、目、口に注目することで正しく表情が認識できていることが示された。

以上より、目、口から物理的特徴を抽出する手法を構成することとする。

抽出する物理的特徴として、心理的な研究から、顔画像における眉・目・口の曲線の湾曲度(開き具合を示す)「湾曲・開示度」と、眉、まなじりの持ち上がりを示す「傾斜」の2つが表情認識において重要な役割を果たしているという報告がされている[4]。これらから、本研究では物理的な表情情報として顔部品の形状情報を用いることとする。

第 3 章

顔画像領域分割

本章では、顔部品抽出処理のための、顔画像の粗さ測度を用いた領域分割処理について述べる。

3.1 領域分割手法の一般

画像から目的とする対象を抽出するためには、画像から目的とする対象を領域として、それ以外の背景と分離する事が必要である。

対象の形状や濃淡パターン等があらかじめわかっている場合、それらをテンプレートとして用意し、入力画像とのマッチングを行うことで対象の抽出が可能である。しかし顔画像の場合、個人間で濃淡パターンは異なり、また表情変化によって多様に変化する。このためこれらに対する正規化が必要であり、顔画像解析では大きな問題として存在する。

これとは別に、領域分割によって対象を抽出することができる。これは、対象と背景が異なる特徴量を利用し、その非均一な部分を領域の境界として求めるものである。

通常、隣接した画素の集合が一つの連結した領域として知覚されるためには、その領域が均質であることが必要である。領域の均質性は、(1) 個々の画素の性質、(2) 局所的な画素集合の性質、などが関係している [6]。そして目的とする対象とそれ以外との区別が可能である特徴量を用いることが領域分割には必要である。

画像から顔部品領域を抽出する手法として、色情報を用いる手法、テンプレートとのマッチングによる手法等が提案されている。次に、それらについて特長と問題点を比較する。

3.2 顔画像の特徴抽出，解析

本節では，従来の顔部品の位置推定手法とその問題点について述べ，次に本研究で用いる，フラクタル特徴解析に基づく顔部品領域の抽出手法について述べる．

3.2.1 従来研究の問題点

【A】色情報を用いる手法

色情報を用いて肌色領域を画像中から抽出し，赤色領域を求めて唇を抽出し，低輝度領域から目，眉領域を抽出する手法が提案されている [8]．色情報を用いることで処理の高速化が図れ，表情や個人性による画像の変化にもロバストである．しかし目領域と眉領域の区別ができないため，頭部は直立しているという制限が必要である．

【B】パターンマッチングによる手法

対象とする画像の濃淡パターンをテンプレートとして用意し，それとのマッチングによって目的とする部分画像を抽出する手法が提案されている [14]．顔部品全体をテンプレートとするのではなく，各部分特徴の相互配置関係をも含めることによって，回転の影響やある程度の大きさの変化，個人性にも対応できるが，目的とする対象以外にもマッチングをすることがあり，マッチングを行うだけではそれらの区別は不可能である．

【C】フラクタル次元を用いる手法

画像曲面のフラクタル性を測度として用いて領域分割し，その中から目や口といった目的とする領域を推定する手法が本田ら [1] [2]．によって提案されている．この手法は，顔画像の2次元平面上の濃淡値を高さとして見なし，これによって構成される3次元曲面を考え，その形状の特徴 (activity) をフラクタル次元値 (Hausdroff 次元値) として扱っている．

フラクタル次元値 (Hausdroff 次元値) は明るさや画像の線形変換 (回転・拡大縮小) の影響を受けないことから，入力画像に対する正規化が不要であるという特長を持っているが，分割領域は顔部品以外の不要領域を含んでおり，それらから目的とする領域を選別する処理が必要である．

3.2.2 フラクタル特徴解析による顔画像の領域分割

本研究では、既に本田らによって提案されている、画像曲面のフラクタル性を測度として用いて顔画像を領域分割し、その中から目や口といった目的とする領域を推定するアプローチを用いる [11] [12]. この手法は、顔画像の2次元平面上の濃淡値を高さとして見なし、これによって構成される3次元曲面を考え、その形状の特徴 (activity) をフラクタル次元値 (Hausdroff 次元値) で表す. そして計算された次元値を測度として画像を分割・統合することによって顔部品や眉、顔の輪郭などの顔部品の候補となる領域を得る. フラクタル次元 (Hausdroff 次元) は、線形変換 (回転, 拡大・縮小) の影響を受けないので、顔画像ごとの回転や大きさへの正規化が不要であり、撮影条件に対し柔軟で精度の高い顔部品の位置推定が行えるとしている (図 3.1).

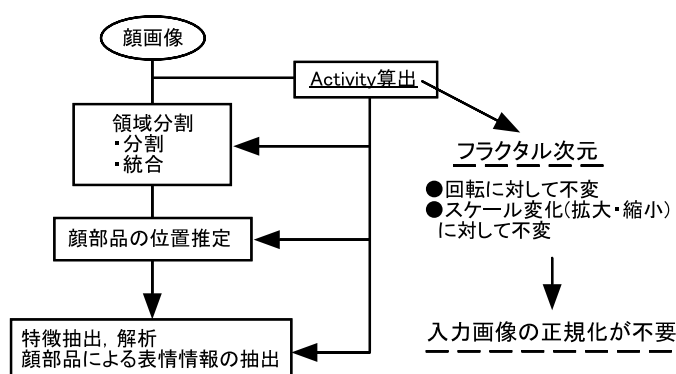


図 3.1: フラクタル特徴解析に基づく顔画像の特徴抽出, 解析のアプローチ

しかしフラクタル次元 (Hausdroff 次元) を用いた領域分割の結果、顔部品領域以外に輪郭や衣類といった不要領域が残っており、領域分割処理の後に、不要領域を除去し、目、口領域を推定する処理が必要である。

提案手法では領域分割によって得られた顔部品候補領域から目的とする目、鼻、口領域を抽出する際に、顔の横幅を推定して不要領域を除去し、目、鼻、口領域間に存在する次元値の大小関係を用いて領域種類の推定を行っている。しかし顔の横幅推定を行うため、頭部が傾いた場合には横幅の推定を誤るという問題があった。また領域の推定のチェックとして、目、鼻、口がT字型の配置をしているという制約を用いているが、誤推定を避けるために頭部はほぼ正立という制限が残っていた。

本研究では、機械によるコミュニケーションの効率化・高度化を大きな目的としている。

想定している状況から，顔画像の撮影において頭部の回転や大きさに対する制限を設けない解析手法が必要である．フラクタル特徴解析による顔画像の領域分割手法は，顔画像への正規化が不要であるという点から，この条件に沿っている．以上より，本研究ではフラクタル特徴解析による顔画像の領域分割手法を用い，その後の顔部品領域推定において，提案手法では制限されている，回転へのロバスト性をさらに高めることを目指す．

3.3 Hausdroff(フラクタル)次元による曲面形状(パターン)解析

一般にフラクタル次元は，図形の自己相似性や複雑さを表す尺度として知られている．本研究では，画像の濃淡値が作る画像曲面の変化の複雑さ(粗さ)を表す尺度として用いる．本節では，まずフラクタル(Hausdroff)次元の基礎的な知識について述べ，次に画像上でのフラクタル(Hausdroff)次元の計算手法について述べる．

3.3.1 Hausdroff 次元

フラクタル [2] は，狭義には自己相似(部分と全体が相似)な図形，講義には Hausdroff 次元が位相次元よりも大きい図形として定義される．フラクタル次元は，フラクタル図形に与えられる図形の複雑さを表す尺度として用いることができ，数学的には Hausdroff 次元のことである．自然に存在する図形や，離散的な画素値の配列による画像などには自己相似性が成り立つものはほとんどなく，このような図形に対しては自己相似次元の拡張として Hausdroff 次元が定義される．

集合 E を，半径 $r(m)$ ($m = 1, 2, \dots$) の球 $V_{r(m)}$ で被覆，すなわち， $\cup V_{r(m)} \supseteq E$ とする．

$$H_D(E) = \lim_{\max\{r(m)\} \rightarrow 0} \inf_{\{V_r\}} \sum r(m)^D \quad (3.1)$$

とおくと， $H_D(E)$ は D の関数として単調減少である．このとき， E の Hausdroff 次元 Hd は $H_D(E)$ がはじめて有限の値を取るときの D の値として定義される．すなわち，

$$H_D(E) = \infty \quad \text{when} \quad 0 \leq D < Hd \quad (3.2)$$

$$H_D(E) = 0 \quad \text{when} \quad Hd < D < Hd \quad (3.3)$$

である．Hausdroff 次元を定義通りに計算するのは困難であるが，通常は

$$\lim_{s \rightarrow 0} R(s) = -\log N(s) / \log s \quad (3.4)$$

として計算する．Hausdroff 次元には自己相似の概念は入っていないため，任意の図形に対して Hausdroff 次元を定義することができる．

3.3.2 画像における Hausdroff 次元推定

前節での Hausdroff 次元の定義に従い，画像の Hausdroff 次元を求める．ここでは，図 3.2 に示すように，画像を 2 次元平面上に展開された等間隔な格子点 (x, y) に画素値を高さ $z = (x, y)$ として配置した時に，それらが構成する曲面を考える．定義から，画像曲面を 1 辺の長さが s の単位面積素（正方形パッチ）で被覆したとき，そのパッチの数を $N(s)$ とし， $(\log(s), \log N(s))$ のプロットの分布から極限値を推定する手法について言及してきた．しかし実際には画像曲面の総面積 $A(s)$ を何らかの方法で求め，次式から Hausdroff 次元を推定する手法が提案されている．

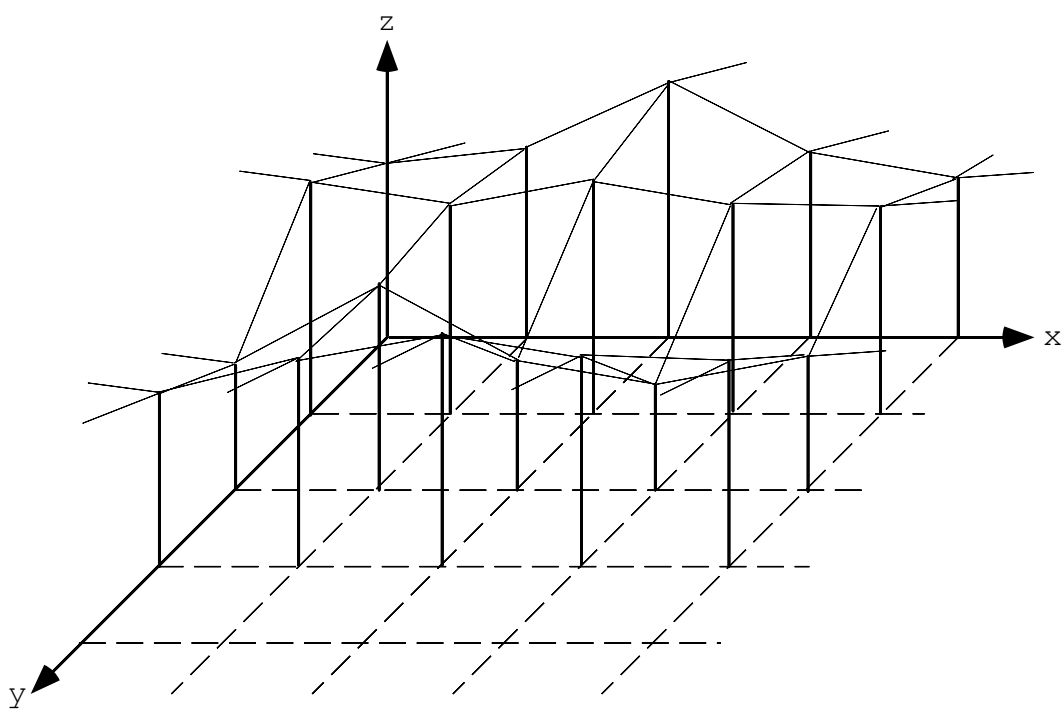


図 3.2: 画素値が構成する画像曲面

本研究では, peleg [3] らによって提案されている Blanket-Covering 法により $A(s)$ を求め Hausdorff 次元値を計算する. Blanket-Covering 法では, 画像サイズを w , スケールを s とするとき, 上面 U_s と下面 L_s によって構成される膜の体積の変化を以下のように帰納的に求め, Hausdorff 次元を推定する.

$$U_0(x, y) = L_0(x, y) = f(x, y) \quad (3.5)$$

$$U_s(x, y) = \max[U_{s-1}(x, y) + 1, \max_{|(l,m)-(x,y)|=1} \{U_{s-1}(l, m)\}] \quad (3.6)$$

$$L_s(x, y) = \min[L_{s-1}(x, y) + 1, \min_{|(l,m)-(x,y)|=1} \{L_{s-1}(l, m)\}] \quad (3.7)$$

$$A(s) = \frac{1}{2s} \sum_{xy \in w} [U_s(x, y) - L_s(x, y)] \quad (3.8)$$

3.3.3 画像において Hausdorff 次元が示す特性

一般に Hausdorff 次元は, 図形の線形変換(回転, 拡大・縮小)の影響を受けないという特徴を持っている. しかし, 本研究では Blanket-Covering 法によって Hausdorff 次元を近似して求めているため, 一般に言われている特徴が保存されているかははっきりしていない. そこで, Blanket-Covering 法により計算される Hausdorff 次元が線形変換に対してどのような特徴を持っているのかを検討する. 本田 [12] は, 8×8 [pixel] の DCT 基底画像に対してこれらの特徴について検討しているが, 自然画像に対しては検討していないため, 本研究では自然画像を対象として検討する [13].

対象画像は, 一般的な画像を対象とし, 図 3.3 に示す画像を用いた. 画像は著作権フリーの素材を集めた画像集から, 回転に対し対称となりにくかつ画像の濃淡パターンが異なるものを選び, カラー画像をグレースケールに変換して用いた. 画像は $(700[\text{pixel}] \times 700[\text{pixel}])$ (濃淡値 1) のテクスチャの中央に, 図 2(a), (b) に示すテクスチャ $(200[\text{pixel}] \times 200[\text{pixel}])$ (256 階調) を張り付けたものを作成した. これは Hausdorff 次元値の計算において, 画像端の処理が計算結果に影響しないようにするためである. 以上のようにして作成したモデル画像のそれぞれに回転, 大きさを変化させた画像を作成し, それぞれの Hausdorff 次元値を計算することによって検討する.

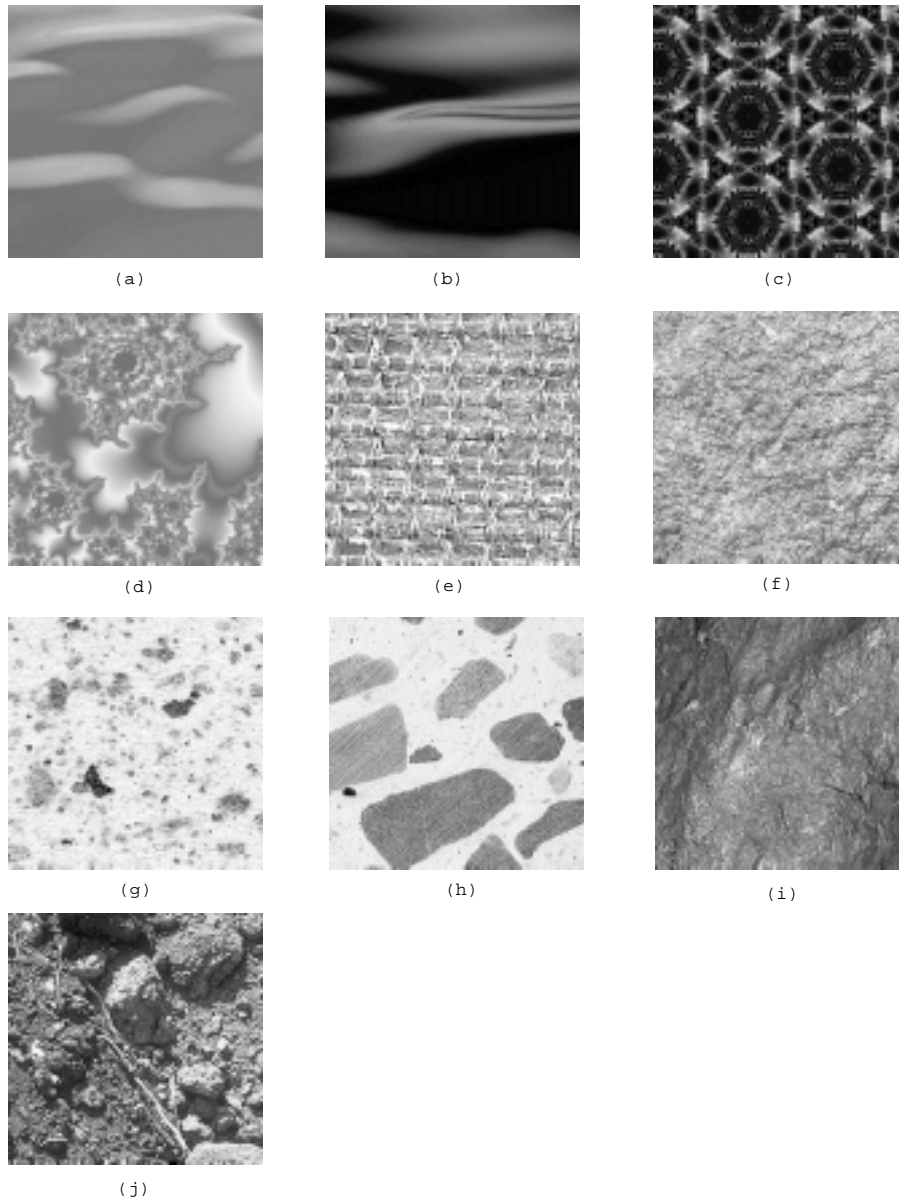


図 3.3: 検討に用いた画像

【A】回転に対する特性

Blanket-Covering 法による Hausdroff 次元の計算の，画像の回転に対する特性を調べるために，図 3.3の中央に張り付けた画像を回転させたものそれぞれについて Hausdroff 次元値を算出した．対象画像は，モデル画像のそれぞれに対して，中央に張り込んだ画像を 0[度]～90[度] まで 10[度] おきに回転させたものを作成し，それぞれに対して Hausdroff 次元値を計算した．なお，次元値を求めたスケールは 50 とした．

次元値をプロットしたものを図 3. 4に示す。

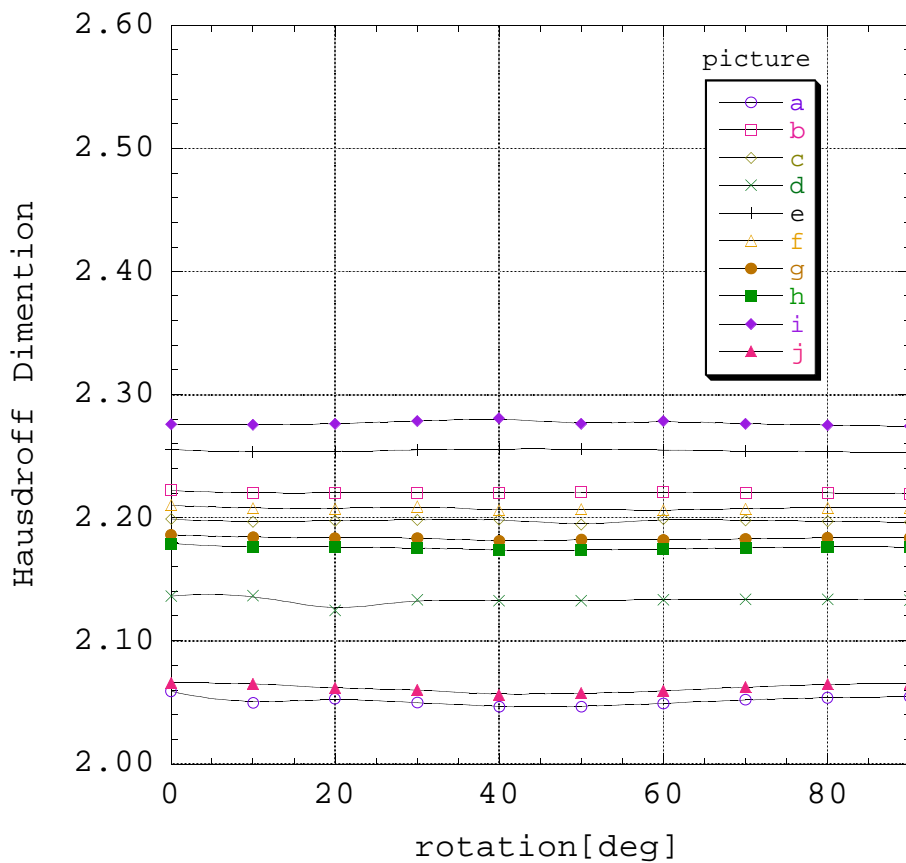


図 3. 4: 回転に対する次元値の違い

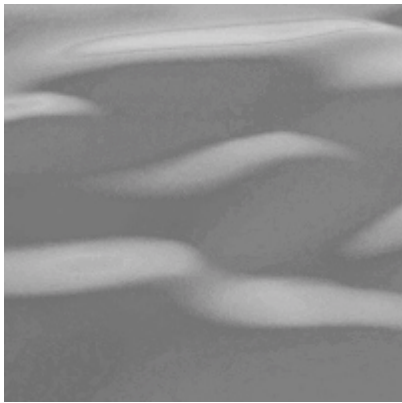
図 3.4 から, Blanket-covering 法を用いて計算したフラクタル次元値は回転に対し, サンプル画像 (a) については最大 5.2[%], サンプル画像 (b) については 0.37[%] の範囲に収まっている. これにより, Blanket-covering 法による Hausdorff 次元値は, 画像の回転に対してはほぼ影響されないという結果を得た.

【B】大きさの変化に対する特性

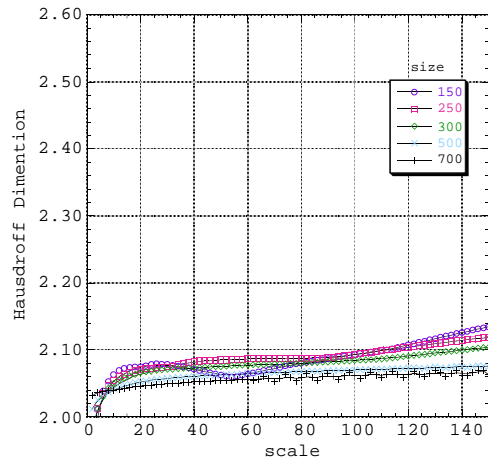
次に, Hausdorff 次元値に現れる画像の特徴が, 画像の大きさの変化に対しどの程度保存されるかを実験により検証する. 検証は, 作成したモデル画像の解像度を变化させた画像に対し次元値を算出し, 比較検討することで行う. Hausdorff 次元値を計算する場合, あるスケール値に対する次元値の形で計算を行うが, 今回はスケール値についての特性も同時に調べるため, ($s[2, 4, \dots, 16]$) の範囲でそれぞれ次元値を計算した.

次元値をプロットしたものを図 3.5~ 図 3.9 に示す.

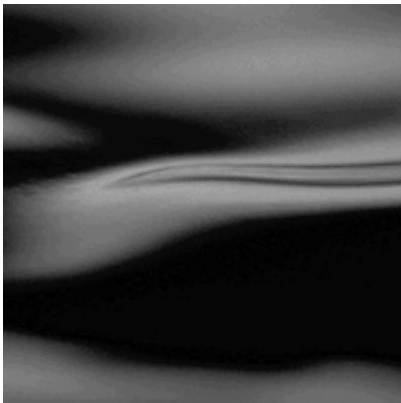
実験の結果, 今回用いたモデル画像の場合, 原画像に対して半分程度の解像度まで低下させても, サンプル画像 (a) で最大 2.6[%], サンプル画像 (b) で最大 1.5[%] の変動範囲に収まるが, それ以下の解像度になると値が大きく変動する傾向がある見られた. これより, ある程度解像度が保証されている場合, 次元値は大きさの変化に対して影響を受けないということがいえる. また, 実際の解析において, 次元値を求めるために必要となる解像度を解析的に決定することができると考えられる.



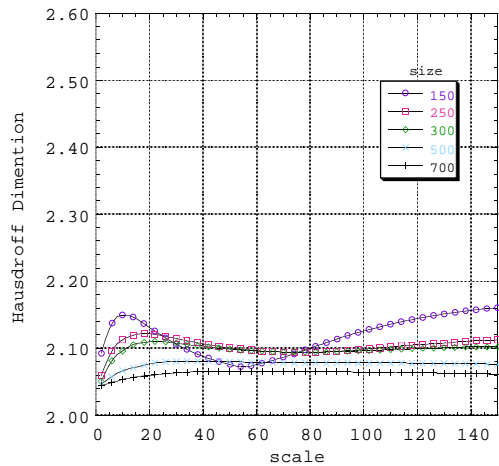
サンプル画像 a



Hausdorff次元値(サンプル画像a)

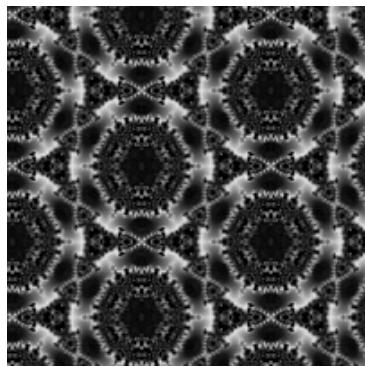


サンプル画像 b

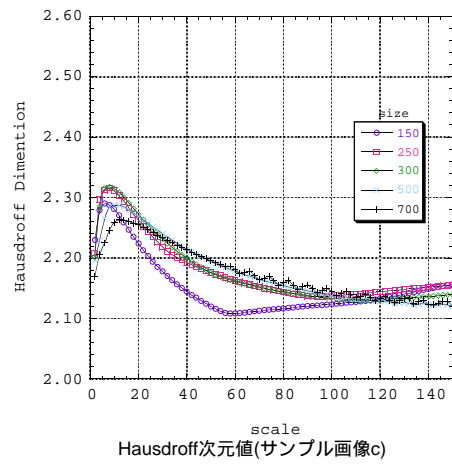


Hausdorff次元値(サンプル画像b)

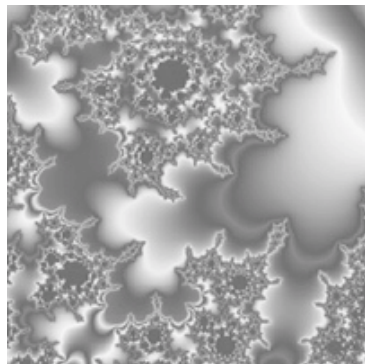
図 3.5: 解像度変化に対する次元値の変化 (1)



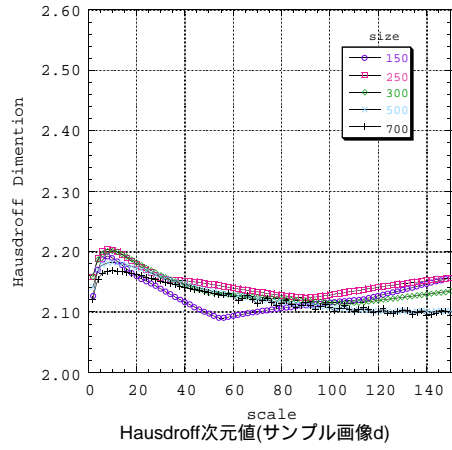
サンプル画像 c



Hausdroff次元値(サンプル画像c)



サンプル画像 d

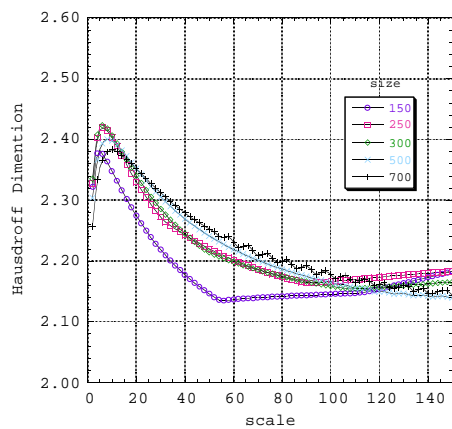


Hausdroff次元値(サンプル画像d)

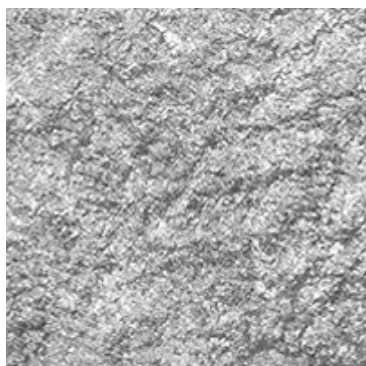
図 3.6: 解像度変化に対する次元値の変化 (2)



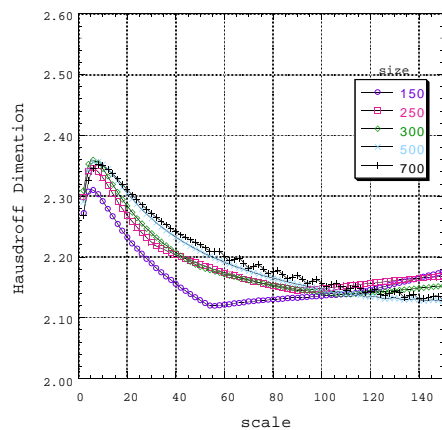
サンプル画像 e



Hausdroff次元値(サンプル画像e)

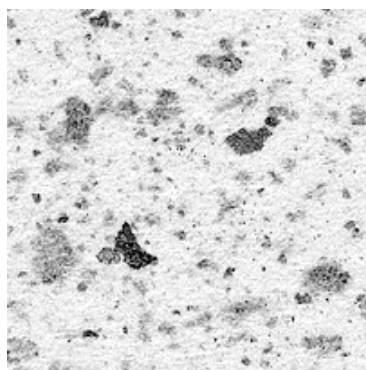


サンプル画像 f

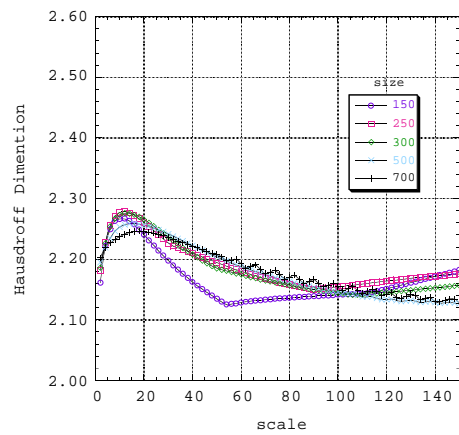


Hausdroff次元値(サンプル画像f)

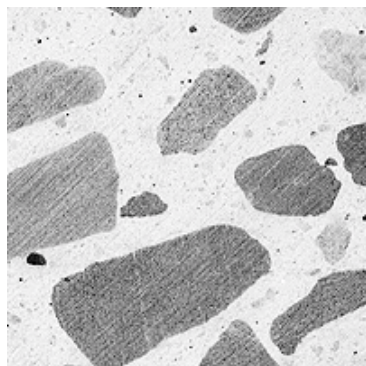
図 3.7: 解像度変化に対する次元値の変化 (3)



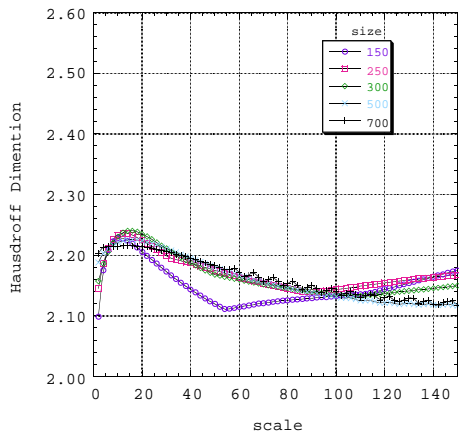
サンプル画像 g



Hausdroff次元値(サンプル画像g)

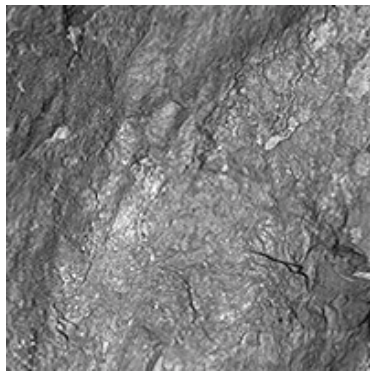


サンプル画像 h

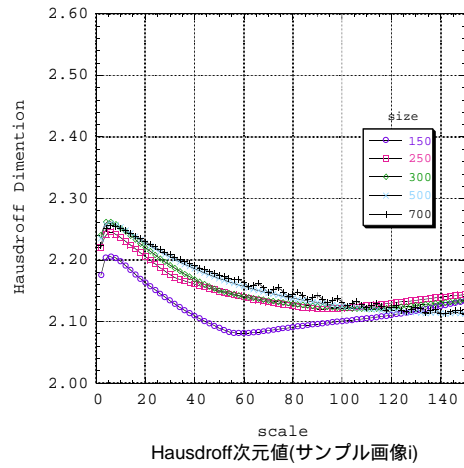


Hausdroff次元値(サンプル画像h)

図 3.8: 解像度変化に対する次元値の変化 (4)



サンプル画像 i



サンプル画像 j

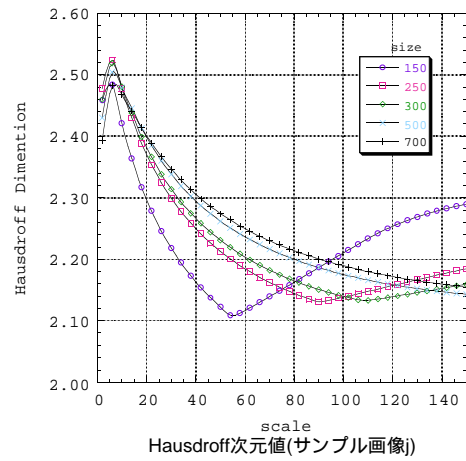


図 3.9: 解像度変化に対する次元値の変化 (5)

3.4 Hausdroff 次元による領域分割

本研究では表情の物理的特徴を抽出するため，領域分割によって得られた顔部品候補領域から顔部品領域を抽出する処理を行う．本研究では，領域分割には本田ら [11] [12]が提案しているフラクタル解析による手法を用いている．本節では，提案手法にて設定している対象画像の範囲，領域分割アルゴリズム，その特性について述べる．

3.4.1 対象画像範囲

本研究では，機械を介したコミュニケーションの効率化・高度化を実現することを大きな目的としており，そこで考えられる状況としては，機械の前に人が座っており，利用者の顔画像を機械が捉え，その画像から表情情報を抽出するというものである．そこで，本研究では正面顔画像を対象とする．画像の種類は，頭部～首を含む画像とし，首を傾げる動作，頭部の回転は許容する．ただし，髭のある人物画像やメガネを装着した人物画像，髪の毛が顔を部分的に覆っている人物の画像は領域分割の結果，複数の顔部品が一つの領域として分割されてしまうため，各顔部品の抽出が困難である [12]．そこで，これらの画像は髭やメガネなどの別の特徴が付加されたものと考え，本研究では対象外とする．

3.4.2 領域分割アルゴリズム

提案手法による領域分割結果の一例を図 3.10，図 3.11に示す．領域の分割には，測度としてスケール $s = 50$ における Hausdroff 次元値を用い，分割アルゴリズムには quad-tree アルゴリズムが用いられている．図から，目，鼻，口といった顔部品領域が，顔輪郭や頭髮といった領域とは独立して良好に分割されていることがわかる．本研究では，この領域分割の結果を用い，以後の顔部品推定処理を行う．

なお，Blanket-Covering法は画素値の上限（最大値），下限（最小値）の変化（置き換え操作）によって次元値を計算している．そのため，目や口などの画素数が充分でないと，画像曲面の特徴が十分に求められない．そこで本研究では，背景を無地にし，顔部分を $35[\text{mm}]$ のカラーリバーサルフィルムを用いて撮影し，フィルムスキャナによってデジタルサイズ（ $2700[\text{dpi}]$ ）する．デジタルサイズされた画像の顔部分を $1300 \times 1600[\text{pixel}]$ のサイズで切り出し，輝度成分のみを用いて $8[\text{bit}]$ 階調で表現することで，各顔部品の画素数を確保する．

図 3.10: 画像粗さ測度を用いた領域分割 (上: 原画像, 下: 領域分割結果)

図 3.11: 画像粗さ測度を用いた領域分割 (上: 原画像, 下: 領域分割結果)

3.5 領域分割の限界

提案手法では，頭部の傾きの正規化を行っていないため若干の頭部の傾きが発生しているが，良好に領域分割が行えている．しかしそれ以上の回転が起きた場合においても領域分割が可能か否かについては示されていない．また，大きさが変化した場合について，ある程度小さな画像を用いた場合でも良好に領域分割ができることが示されているが，その限界は示されていない．そこで，顔画像の大きさを変化させ，また回転させ，そのような場合でも領域分割が可能か否かを確認する．

検討は，正立状態で良好に領域分割が行えている画像（図 3.12）に対し，回転，大きさの変化を施した画像を作成し，それらに対して領域分割処理を行い，結果を比較することで行う．

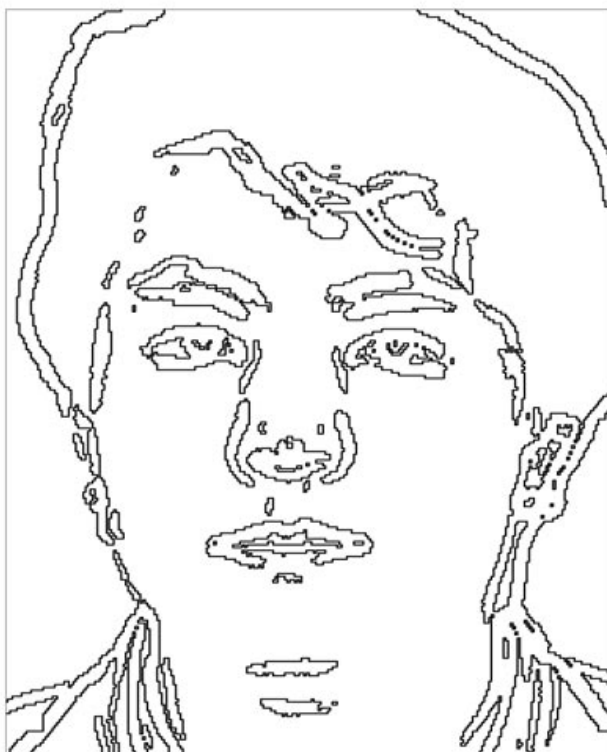


図 3.12: 画像粗さ測度を用いた領域分割 (上: 原画像, 下: 領域分割結果)

【A】回転

回転について、図 3.12 の画像を 45, 90, 135, 180[度] 回転した画像を作成し、それぞれについて領域分割処理を行った。原画像と領域分割結果をそれぞれ図 3.13~ 図 3.16 に示す。回転させた画像は、均一な輝度値を持つ画像の中央部に、原画像を回転処理したものを張り込むことで作成した。90 [度] 回転画像は、画像の縦横サイズを交換している。

領域分割結果から、顔画像が回転した場合、領域分割された領域形状に若干の違いがあるものの、それぞれの顔部品領域は独立して良好に分割されていることがわかる。

これから、提案手法による画像の領域分割は、頭部の回転に対して影響を受けないことが確認された。これから、首を傾げた場合についても、領域分割が良好に行えることが示された。

【B】大きさ

顔画像の大きさが変化した場合の領域分割が可能な範囲の検討について、図 3.12 の画像を 80, 70, 60, 50[%] に縮小した画像を作成し、それぞれに対して領域分割処理を行った。なお、縮小した画像は、均一な輝度値を持つ画像の中央に張り込むことで作成している。

原画像と領域分割結果をそれぞれ図 3.17~ 図 3.20 に示す。領域分割は、顔画像が縮小した場合でも、それぞれの顔部品領域が良好に分割されていることがわかる。そして画像の縮小率が 50[%]、顔の横幅の画素数が約 120[*pixels*] 程度になった画像ではそれぞれの顔部品領域が分離せず、融合している。

これから、顔画像の大きさが変化した場合でも、両目の内側で約 120[*pixels*] 程度の解像度があれば顔部品領域の分離が可能であることが確認された。



図 3.13: 45[度] 回転画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)

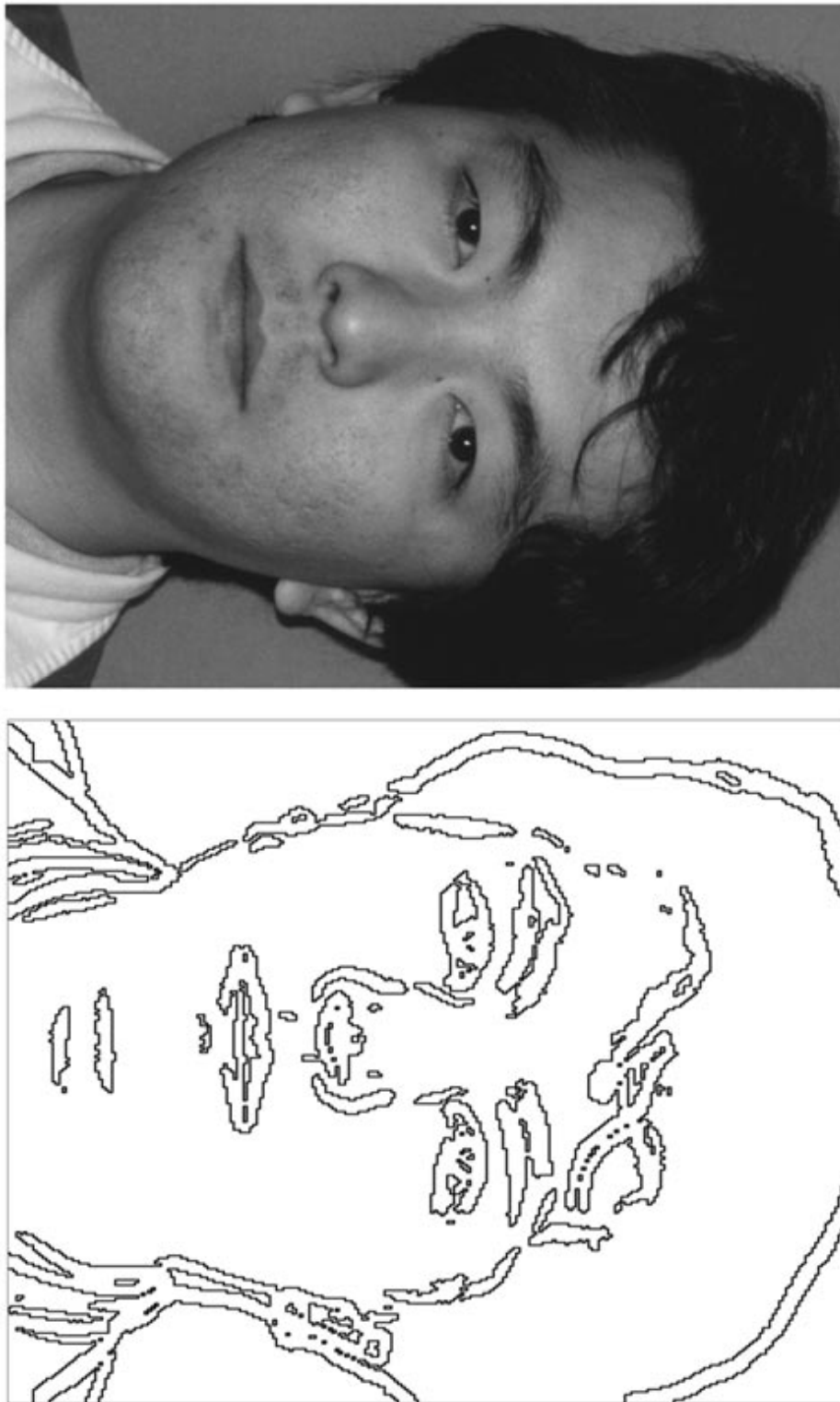


図 3.14: 90[度] 回転画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)



図 3.15: 135[度] 回転画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)

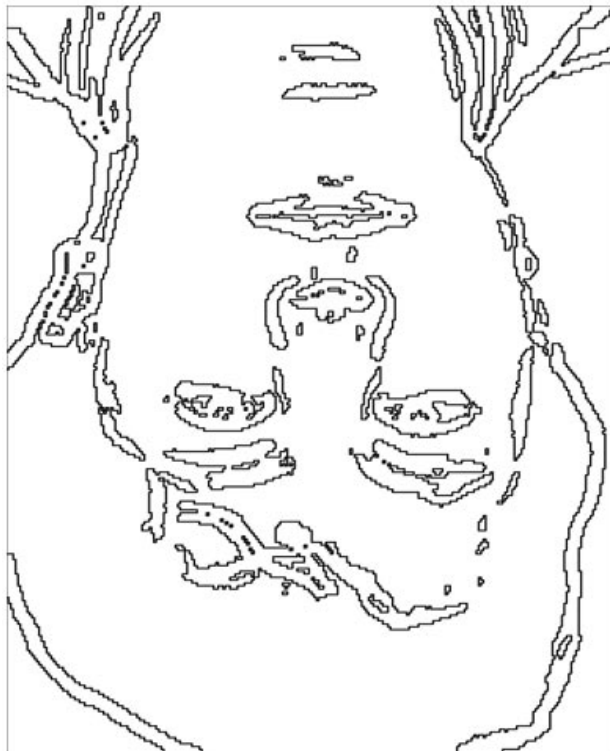
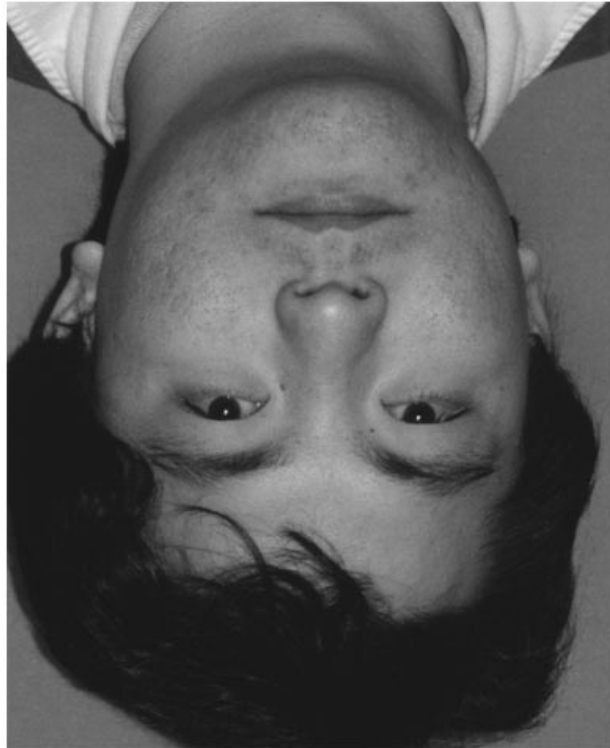


図 3.16: 180[度] 回転画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)

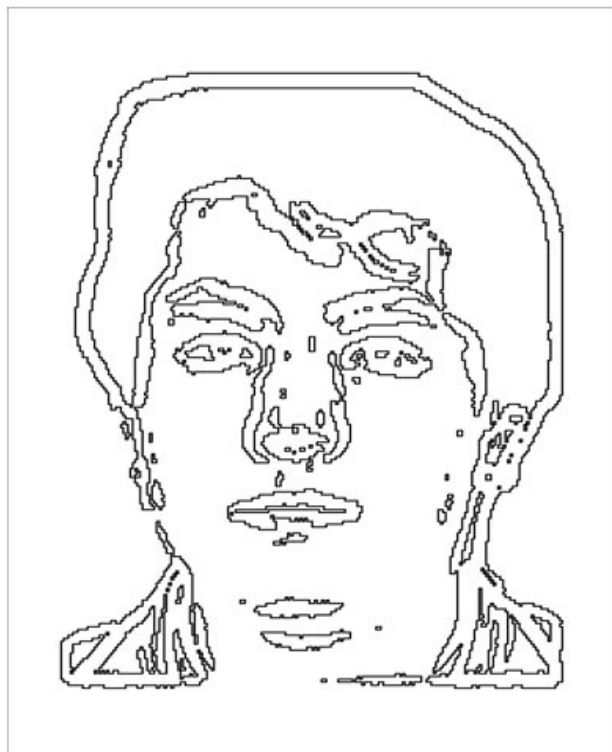


図 3.17: 20[%] 縮小画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)

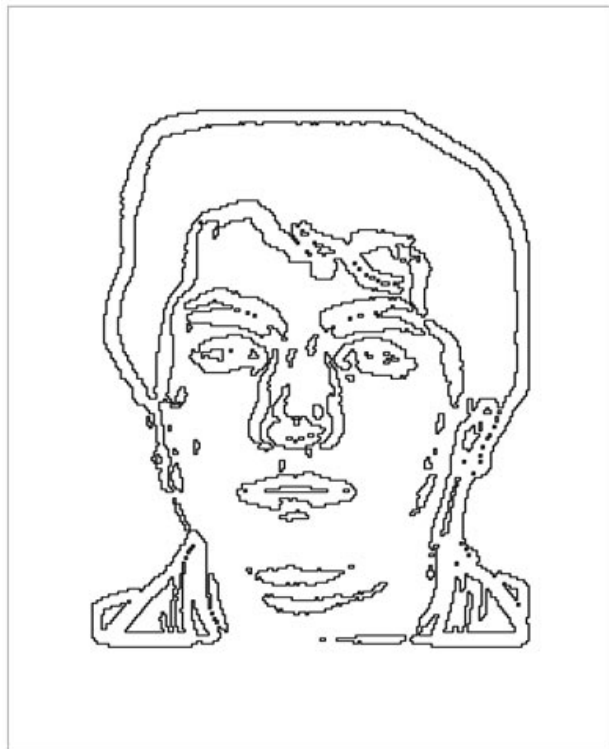
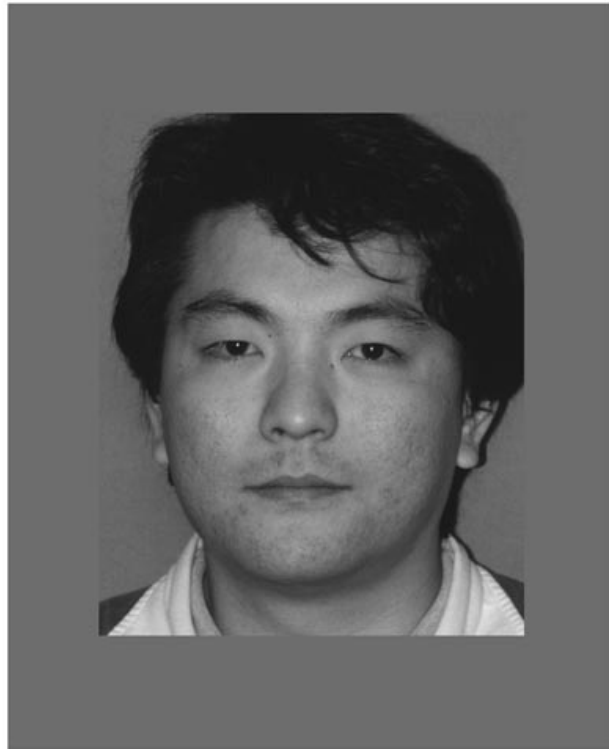


図 3.18: 30[%] 縮小画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)

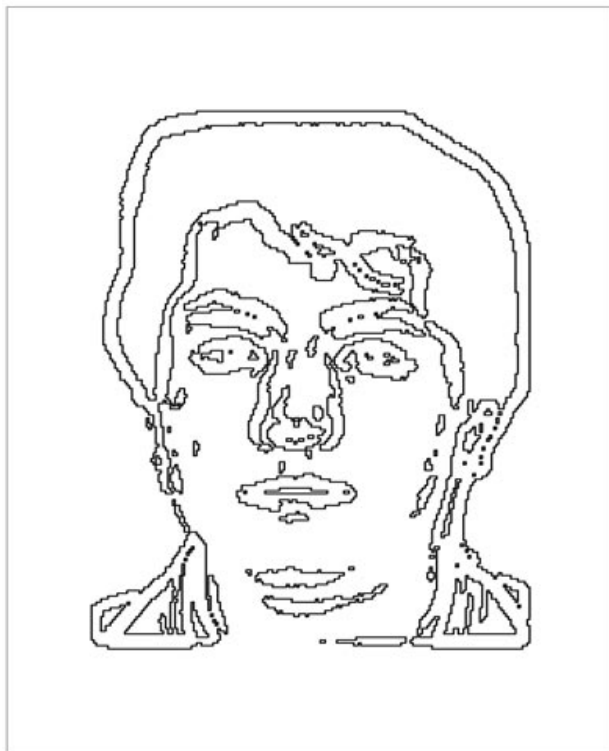


図 3.19: 40[%] 縮小画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)

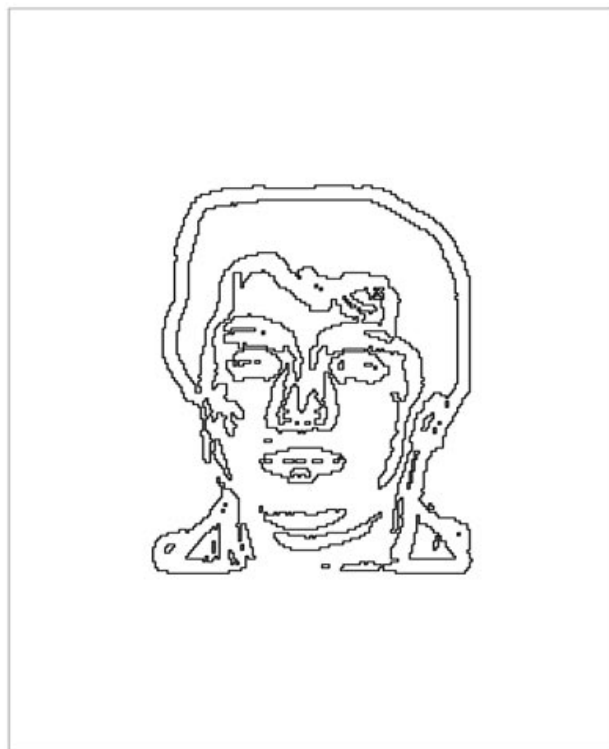


図 3.20: 50[%] 縮小画像に対する領域分割結果 (上: 原画像, 下: 領域分割結果)

第 4 章

Global Facial Model に基づく顔部品 推定

領域分割によって得られた顔部品領域を含む複数の領域(顔部品候補領域)から目的とする顔部品領域を得るために,本研究では Global Facial Model(GFM [14])と, Hausdorff 次元値パターンを用いる.本章では,顔部品の配置に関するモデルである GFM の構成と,それに必要な顔部品の位置関係の調査について述べる.

4.1 Global Facial Model(GFM)

Global Facial Model (GFM)は,顔部品の配置に関するトポロジカルなモデルのことである.一般には目,鼻,口以外に,顔の輪郭なども含んだモデルを示す.前章までに示した顔画像の領域分割の結果,目,口領域が他の領域と独立して分離しており,またその形状が顔部品のおおよその大きさと位置を情報として持っているため,GFMを用いることによって,それぞれの領域の持つ大きさや位置の情報を有効に利用した顔部品領域推定が可能であると考えられる.モデルの設計において,厳密なモデルを設計することによって高精度な抽出が期待できるが,そのためにはモデルに与える情報を高精度かつ安定に抽出することが必要となる.本研究では目と口領域の抽出を目的としていること,画像の領域分割の結果,目と口領域の位置と大きさがおおむね安定して得られていることから,目と口のみをモデルの構成要素とし,領域分割の結果得られる領域の位置と大きさの情報をを用いるモデルを構成することとする.

4.1.1 顔部品の位置関係の測定

まず，GFM を構成するために必要な顔部品の位置・大きさの関係を求める．顔画像は，顔部品抽出に用いている画像と同じ条件で撮影したものを扱い，図 4.1 に示す基準点を手作業で求めた．基準点は，目や口の上下左右端点といった，画像上で特定の容易な点を選んだ．対象とした画像は 20 代の男性 19 枚，女性 6 枚で，表情は無表情のみを対象とした．

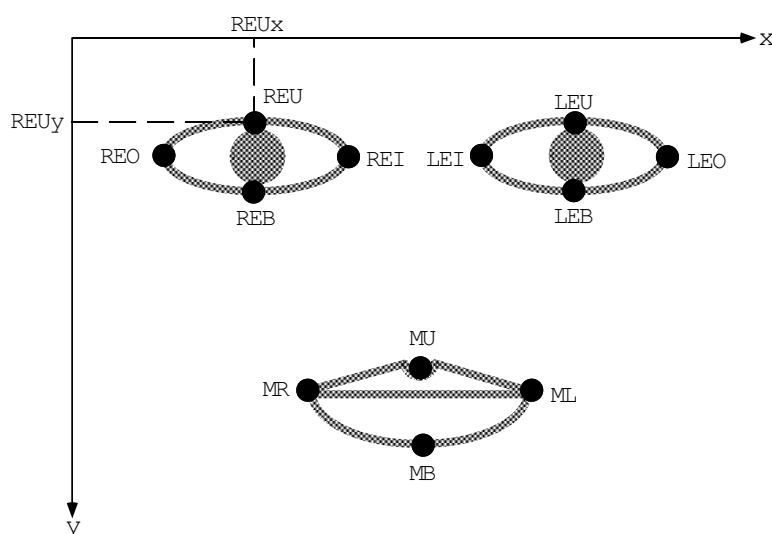


図 4. 1: 顔部品大きさ・位置関係測定のための基準点

次に，図 4. 1 に示す，顔部品の大きさと顔部品の基準点を式 4. 1~ 式 4. 1d によって求めた．

$$REC_x = 1/2(REI_x - REO_x) \quad (4. 1)$$

$$REC_y = 1/2(REU_y - REB_y) \quad (4. 2)$$

$$LEC_x = 1/2(LEO_x - LEI_x) \quad (4. 3)$$

$$LEC_y = 1/2(LEU_y - LEB_y) \quad (4. 4)$$

$$MC_x = 1/2(ML_x - MR_x) \quad (4. 5)$$

$$MC_y = 1/2(MU_y - MB_y) \quad (4. 6)$$

$$d(p1, p2) = \sqrt{(p1_x - p2_x)^2 + (p1_y - p2_y)^2} \quad (4. 7)$$

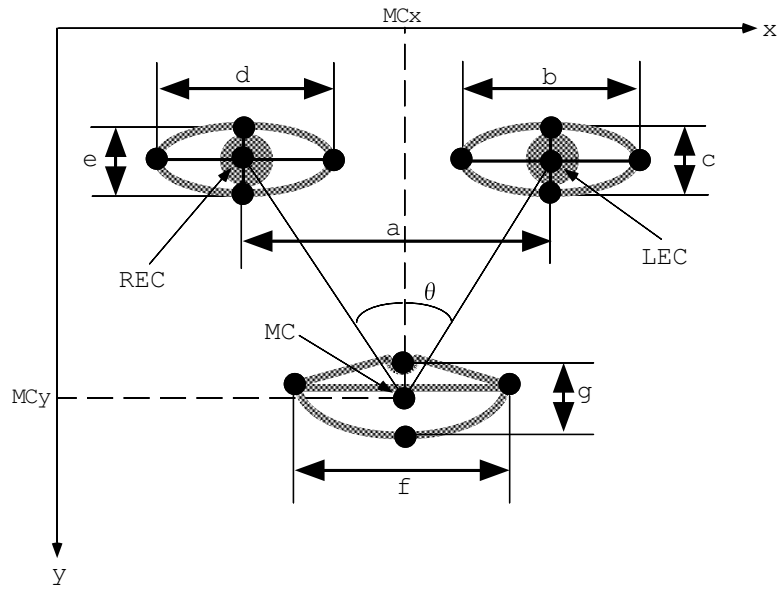


図 4.2: 顔部品大きさ・位置関係測定のための基準点

$$a = d(REC, LEC) \quad (4.8)$$

$$b = d(LEI, LEO) \quad (4.9)$$

$$c = d(LEU, LEB) \quad (4.10)$$

$$d = d(REU, REB) \quad (4.11)$$

$$e = d(REU, REB) \quad (4.12)$$

$$f = d(ML, MR) \quad (4.13)$$

$$g = d(MU, MB) \quad (4.14)$$

$$\theta = 2\pi - \arctan \frac{(MC_y - REC_y)}{(MC_x - REC_x)}$$

$$\arctan \frac{(MC_y - LEC_y)}{(LEC_x - MC_x)} \quad (4.15)$$

$$(4.16)$$

図 4.4 に、測定によって求めた顔部品間の位置・大きさの関係を求めた部位を示し、表 4.1、表 4.2 にそれぞれの値を示す。表では、左右それぞれの目の左右端点を結ぶ直線の中点間の距離（図中、 a ）を基準とし、それに対する比率を示している。

表 4.1: 顔部品の大きさの関係

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>
平均	1	0.450	0.150	0.446	0.146	0.750	0.279
分散 ($\times 10^{-2}$)	0	0.046	0.088	0.063	0.092	0.271	0.137
最小	1	0.409	0.094	0.397	0.082	0.629	0.198
最大	1	0.506	0.206	0.508	0.190	0.855	0.343

表 4.2: 両目, 口の位置関係

	θ
平均	42.59
分散	4.33
最小	41.53
最大	43.64

4.1.2 推定領域設定

領域分割の結果得られた領域はそれぞれ複雑な形状をしており，また頭部の回転によって水平，垂直方向の大きさが変化する．GFM を顔部品推定処理に適用するため，領域の大きさと位置を決定する必要がある．本研究では分割された領域を内包する矩形領域を考え（図 4.3(a)），中心から矩形の頂点までの長さを領域の大きさとして考える．また Hausdroff 次元値を求める際に，回転による領域面積の変化を抑えるために，領域の大きさを半径とする円形領域を用いる（図 4.3(b)）．実際には処理の簡単のために，円形領域を内包する正方形領域を用いて処理を行っている．なお領域の位置は，分割領域を内包する矩形領域の中心座標を用いて表す．

4.1.3 GFM の設計

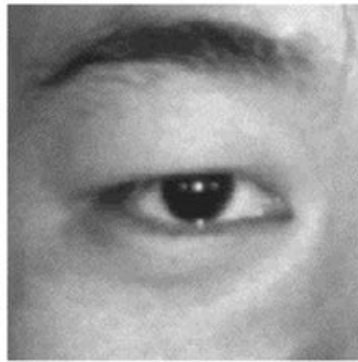
これまでに，顔部品の大きさ，位置関係の測定，対象領域の大きさ，領域形状設定を行った．本節では，実際に顔部品領域推定処理を行う際に用いる GFM を構成する．本研究では，GFM を構成する要素として左右の目，口を用いており，基準に対する領域サイズの許容範囲と，位置の許容範囲をそれぞれ決定することでモデルを設計してゆく．

本研究では，顔部品間のそれぞれの大きさ，位置関係を図 4.4（左示すように表す．

次に，顔部品領域のそれぞれの大きさ，位置の許容範囲を決定するため，目領域の大きさと中心距離の関係を求める．次に，口領域と左右目領域とがなす角度を設定する．以上より GFM を構成してゆく．

なお，先に設定した領域の大きさは，たとえば目領域では領域の左右端点の距離ではなく，分割領域を内包する矩形領域の対角線長の $1/2$ である．表 4. から，目の横幅を w ，縦幅を h とすると，ほぼ $w = 3h$ という関係がある．そこでこの場合に領域の大きさ r を求めると， $r = \sqrt{(h^2 + (3h)^2)}$ から $r = 1.16w$ となり，実際に分割された目領域のの横幅 w に比較して領域の大きさは約 16 [%] 大きくなる（図 4.3 (b).) これを考慮し，表 4. から，GFM のパラメータを式 4.17~式 4.19 のように設定する．それぞれの値の範囲は，測定の結果得られた最小値と最大値を含む範囲とした．パラメータは図 4. の e_r, e_l, e_b, θ を用い， e_r を基準として示している．

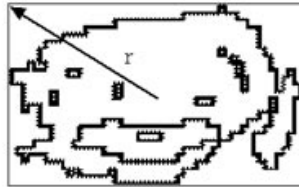
$$0.9e_r \leq e_l \leq 1.1e_r \quad (4.17)$$



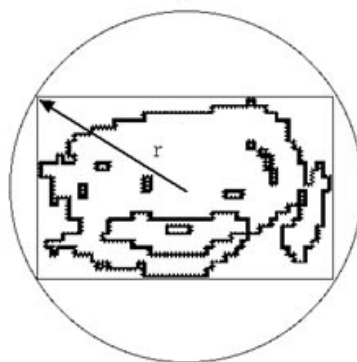
現画像 (右目部分)



分割領域



(a) 矩形領域と領域サイズ r



(b) 円形領域

図 4.3: 分割領域からの顔部品推定処理領域の決定

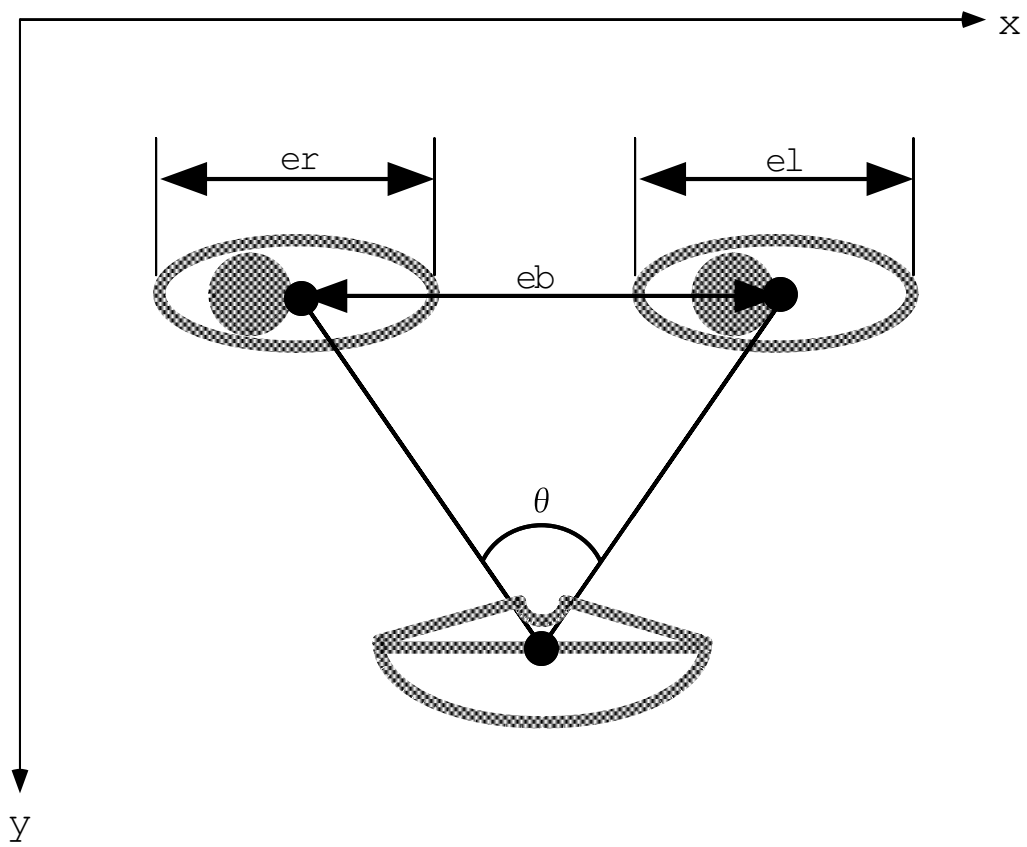


図 4.4: 顔部品大きさ・位置関係

$$2.0er \leq eb \leq 3.0er \quad (4.18)$$

$$40 \leq \theta \leq 45 \quad (4.19)$$

以上から求められた GFM を用い、次節から顔部品領域種類の推定を行う。

4.2 目領域候補ペア推定

本節では、顔部品候補領域から 2 つの目領域を抽出する処理について述べる。本研究では、GFM を用いて分割領域を比較して目領域を抽出するため、GFM に合致する目領域候補は 2 つの領域がペアとなる。目領域候補ペアは、分割された顔部品候補領域群に対し、GFM に合致するものを抽出するため、複数存在することが考えられる。そのため本研究では、目領域の候補となる一对の領域を「目領域候補ペア」と呼ぶこととする。

目領域候補ペアは、GFM に合致する領域の組み合わせを求め、さらに、目領域候補ペアのそれぞれに対し、ペアを構成する 2 つの領域の Hausdroff 次元値の変化パターンの差を求めて、その差が小さい順に順位づける。このようにすることで、それぞれの目領域候補ペアの位置、大きさに「目領域らしさ」の情報を加えて出力する。

4.2.1 目領域推定アルゴリズム

まず、顔部品候補領域から目領域の候補となる領域ペアを求める処理について述べる。

目領域の候補領域ペアを抽出する処理の流れを図 4.5 に示す。本研究では、左右の目の大きさと形状はそれぞれ似ていると仮定する。

【1】GFM による目領域候補ペア抽出

領域分割の結果得られた顔部品候補領域から領域を一つ選び、それを目領域候補 1 とする。次に、目領域候補 1 以外の領域を一つずつ選び、大きさと位置関係が GFM に合致するか否かチェックする。合致しない場合はその領域を候補外とし、次の領域を選んで再び大きさと位置関係をチェックする。このようにして目候補領域 1 に対する目領域候補 2 を選定し、目候補領域ペアとする。以上の処理を、すべての顔部品候補領域が目候補領域 1 として処理されるまで繰り返す。この処理によって得られる目領域候補は複数あることが考えられるため、次に 2 つの領域の類似性を求めて次の口領域の推定を行う。

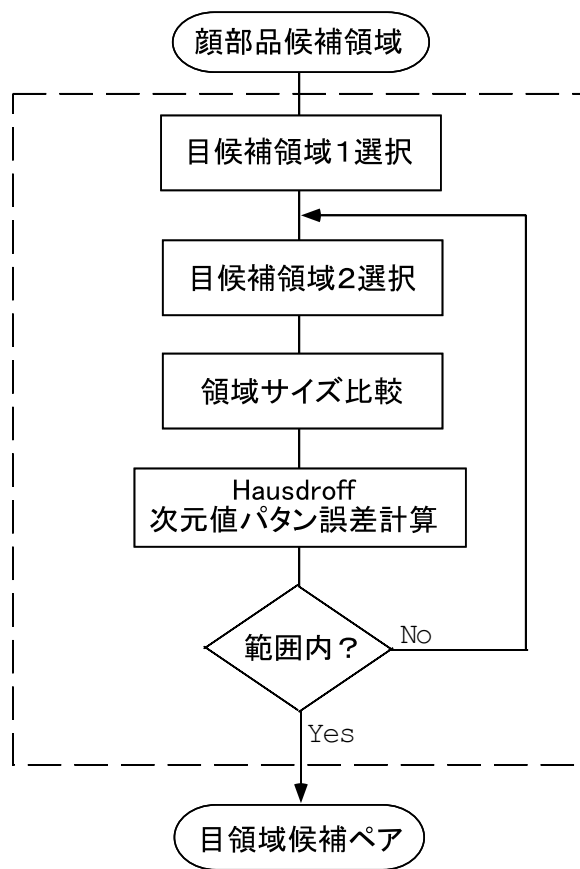


図 4.5: 目領域候補ペア抽出処理の流れ

【 2 】 領域ペアの類似性比較

次に，目領域候補ペアが複数存在する場合，それらの領域の類似性を求め，類似性の高いものをより「目領域らしい」として，顔部品領域推定を行う．

2つの領域の類似性を求める手法として，相互相関を求める手法が一般的である．しかし本研究でその類似性を求める目領域候補ペアは，領域の大きさが一定ではなく互いに異なることがあるため，相互相関を求めることができない．そこで本研究では，部分画像の Hausdroff 次元値のパターンを部分画像の特徴量として用い，領域の類似性を比較する．

Blanket-Covering法による Hausdroff 次元は，一般画像においてはスケール s に関する関数として考えることができる．第3章で示したように，Hausdroff 次元は画像の濃淡パターンの粗さを示しており，スケールに対する Hausdroff 次元値の変化パターンは，濃淡画像の一つの特徴として考えることができる．顔部品画像において，左目画像と右目画像では Hausdroff 次元値の変化パターンは非常に似ていることが示されている [9]．そこで Hausdroff 次元値のパターンを比較することで目領域候補ペアの類似性を求めることができる．今回，2つの領域の Hausdroff 次元値の変化パターンの2乗誤差を計算し，差の小さい領域ペアから順に「目領域らしい」ペアとして出力する．比較する際のスケールは， $(s = 10, 12, \dots, 50)$ の範囲とした．これは，Blanket-Covering法があるスケールにおける画表面積の対数比をプロットして求めるため，点数が少ないと計算誤差が大きくなるためである．

4.2.2 目領域推定実験条件

前節までに示した領域推定処理を用い，複数人の表情画像に対して目領域候補ペアの抽出実験を行う．実験に用いる画像は，領域分割において顔部品領域が独立して分割された画像を用いる．画像の撮影条件，画像の作成条件は領域分割におけるものと同じである．

対象とした画像は，20代の男性3名，女性2名の，無表情2種類，笑い，怒り，悲しみの5表情，合計25枚である．対象とした画像を図4.2.2～図4.2.6に示す．



图4.5: 顔部品抽出対象画像(人物1, 人物2)



図4.5: 顔部品抽出対象画像 (人物1, 人物2)



図4.7: 顔部品抽出対象画像 (人物5)

表 4.3: 目領域ペア候補抽出結果

人物	顔部品候補 領域数 (平均個数)	目領域ペア 候補数 (平均個数)	目領域推定 成功率 [%]
人物 1	46.2	7.6	80
人物 2	55.0	8.4	60
人物 3	43.0	6.6	60
人物 4	67.6	10.4	80
人物 5	55.4	5.8	60
平均	53.4	7.76	68

表情	顔部品候補 領域数 (平均個数)	目領域ペア 候補数 (平均個数)	目領域推定 成功率 [%]
無表情 1	54.8	6.0	80
笑い	61.8	7.6	10
怒り	52.8	10.4	80
悲しみ	46.6	5.4	60
無表情 2	51.2	9.4	100
平均	53.4	7.76	66

4.2.3 実験結果

目領域候補ペア領域の抽出結果を表 4.3 に示す。領域分割の結果得られた顔部品候補領域の数，GFM 適用の結果得られた目領域候補ペアの中に，抽出されるべき目領域が含まれているか否かを示す。また目領域候補ペアの抽出例を図 4.6 に示す。

4.2.4 考察

実験の結果，対象人物が異なる場合でも，平均して 68[%] の画像において，目領域候補ペアとして目領域を抽出することができた。表情別の結果から，笑いの表情では他に比較して 10[%] の抽出率と，低い値となった。それ以外の表情では 60[%] ~ 100[%] と，表情の

図 4.6: 目領域ペア候補抽出結果

図 4.7: 目領域ペア候補失敗例 1

図 4.8: 目領域ペア候補失敗例 2

違いによらず目領域の抽出が行えている。

抽出に失敗する原因として、顔部品領域の分割の結果、目の端が一部分欠けた状態となり、目領域の大きさが小さくなったことが挙げられる(図 4.7, 図 4.9)。また設定された、眉を中心とする正方形領域の中に目が入っており、次元値パターン比較の結果眉を含む方がより「目らしい」と判定された例もあった(図 4.8)。これらは、領域サイズの大きさを小さく設定すること、領域分割をより正確に行うことで解決できると考えられる。

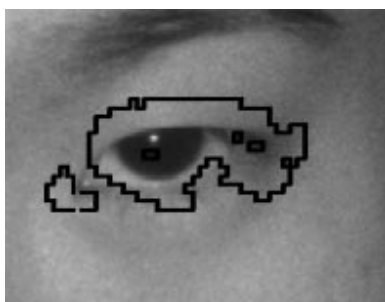


図 4.9: 領域分割失敗例

4.3 口領域推定処理

本節では、顔部品候補領域から口領域候補を抽出する処理について述べる。これまでに、GFM を顔部品候補領域群に対して適用し、これに合致する領域のペアを目領域候補ペアとして求めた。このペアについて、その大きさと位置のパラメータを GFM に与え、その結果得られる範囲内に口領域が存在するか否かをチェックする。これによって、目領域候補ペアに対して口領域が存在した場合、目領域・口領域として出力する。そして最も「目領域らしさ」の高いものを最終的な顔部品として出力する。以下に、口領域推定処理について述べる。

4.3.1 口領域の推定アルゴリズム

口領域の推定処理の流れを図 4.10に示す。

前節にて得られた目領域候補ペア領域に対応する口領域を推定する。口領域の推定では、目領域候補ペア領域の位置と大きさをパラメータとし、GFM を用いて口領域の存在

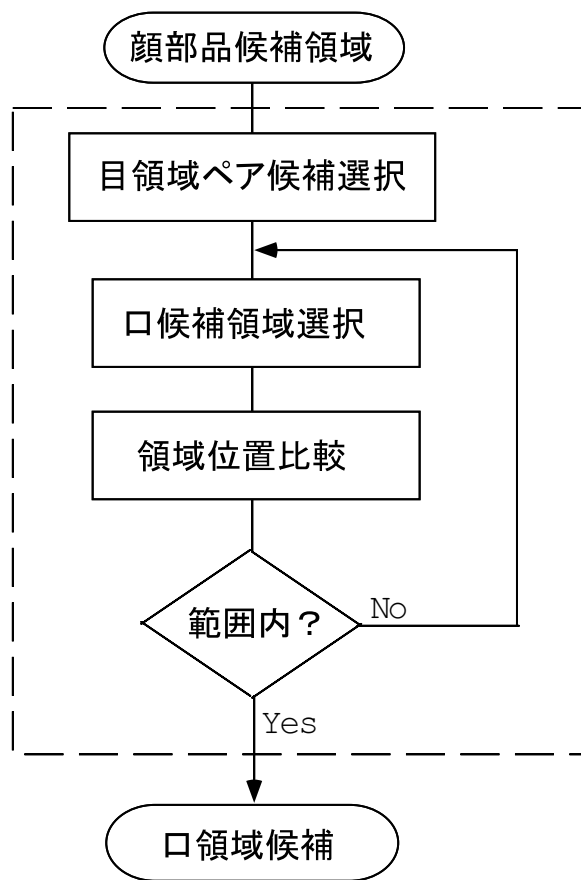


図 4.10: 口領域候補推定処理の流れ

範囲に領域が存在するか否かチェックする。領域が存在しない場合，その目領域候補ペアは目領域ではないと判定し，他に目領域候補ペアが存在する場合はそれらの領域ペアに対して口領域推定処理を行う。また口候補領域が存在する場合，それに対応する目領域候補ペア領域を目領域とし，口候補領域を口領域として出力する。以上の処理を，目領域候補ペアが複数ある場合はすべての候補ペアに対して処理する。すべての目領域候補ペアに対する口領域の推定処理が終了した後，目，口領域が揃って抽出された候補の中で，最も目領域ペアの「目領域らしさ」が高いものを最終的な目，口領域として出力する。

以下に，口領域推定処理の流れについて述べる。

4.3.2 口領域推定実験条件

前節までに示した領域推定処理を用い，複数人の表情画像に対して口領域の抽出処理を行う。処理に用いる画像は，領域分割において顔部品領域が独立して分割された画像を用いる。画像の条件は領域分割におけるものと同じである。

対象とした画像は，目領域候補ペアの抽出で用いたものと同じ，20代の男性3名，女性2名の，無表情2種類，笑い，怒り，悲しみの5表情，合計25枚である。

4.3.3 実験結果

口領域候補の抽出結果を表4.4に示す。GFM適用の結果得られた目領域候補ペアに対し口領域が存在するか否かチェックする。そして口領域が存在するものの中で目領域候補ペアの類似性が最も高い組が正しく両目，口領域であったものを正解とする。抽出例を図4.11に示す。

4.3.4 考察

実験の結果，正しく目領域・口領域が抽出できたものは約50[%]と低い抽出率となった。原因として，領域分割の結果，口領域が複数に分割(4.12)されたため，GFMの範囲外として除去されてしまったことが挙げられる。また口が大きく開いている場合(図4.7)，GFMの範囲外になって抽出に失敗する結果となった。

これらから，顔部品候補領域の分割が正確に行われる必要があること，また無表情顔のみを用いて設計したGFMに対し，目の開閉や口の開閉等の変化量を考慮したGFMへの

表 4.4: 口領域候補推定結果

人物	目領域ペア 候補数 (平均個数)	口領域推定 成功率 [%]
人物 1	7.6	60
人物 2	8.4	40
人物 3	6.6	60
人物 4	10.4	60
人物 5	5.8	40
平均	7.76	52

表情	目領域ペア 候補数 (平均個数)	口領域推定 成功率 [%]
無表情 1	7.6	60
笑い	8.4	10
怒り	6.6	60
悲しみ	10.4	60
無表情 2	5.8	80
平均	7.76	54

図 4.11: 口領域推定結果例

改良が必要と考えられる。

口領域が複数領域として分割された場合，あらかじめ複数領域を統合し，単一領域としてから領域推定処理を行う，または GFM に適合する口候補となる複数領域を統合して口領域とするなどによって解決が可能であると考えられる。



図 4.12: 複数領域に分割された顔部品（口）

4.4 回転・大きさの変化に対するロバスト性の検討

本研究にて想定してる状況では頭部の移動・回転が発生し得るため，これらに対する輪郭抽出のロバスト性が必要である．これについて，対象画像に対して回転・大きさの変化を加えた画像を作成し，輪郭抽出処理を行うことで，提案手法の回転・大きさの変化に対するロバスト性を検証する．

対象とした画像は，顔部品抽出の回転・大きさの変化に対するロバスト性を検証した際に用いた画像である．この画像に対して，10[度]の回転，20[%]の縮小を行った画像を用い，輪郭抽出処理を施した．原画像と輪郭抽出処理結果を図 4.13，図 4.14に示す．

図 4.13: 回転画像に対する顔部品抽出結果 (10[deg] 回転)

図 4.14: 縮小画像に対する顔部品抽出結果 (20[%] 縮小)

4.5 まとめ

本章では、分割された顔部品候補領域群から、目領域・口領域を抽出する手法について述べた。本研究では顔部品領域の推定手法として、目・口の位置と大きさの関係を求め、顔部品のトポロジカルなモデルである Grobak Facial Model (GFM) を構成し、それを用いて顔部品領域の推定を行った。

まず、GFM を構成するために、25 枚の無表情顔画像に対し、目、口に設定した基準点を手作業で求め、それらから計算される大きさ、位置関係を求めた。そして本研究で用いている領域の設定手法に合わせ、GFM を設計した。

次に、5名の無表情2種類、笑い、怒り、悲しみの表情、合計25枚の表情画像を Hausdorff 次元値を測度とした顔画像の領域分割手法によって領域分割を行い、その結果得られた顔部品領域群に対して GFM を用いて顔部品推定処理を行った。その結果、全画像において目領域候補ペアが抽出でき、その中で抽出されるべき目領域のペアが含まれていたものは 68[%] と、無表情以外の表情画像においても目領域ペアの抽出が可能であることが示された。

目領域候補ペアの位置・大きさパラメータと GFM を用いた口領域の推定では、口領域が正しく抽出できたものは 50[%] 程度と低い値となった。これは、口領域は目領域とは異なり、上唇と下唇などの複数の領域に分割される場合が多く、また大きく口が開いている場合、GFM に合致しないために除去されてしまったことなどが挙げられる。これに対して、分割された領域情報をそのまま用いるのではなく、ある範囲内にある複数の小領域は統合してから領域抽出処理を行う、または複数の領域が口領域として抽出された場合、それらの位置と大きさのチェックを行い、条件を満たす場合は統合するなどの手法を用いることによって、より正確な領域の抽出が可能であると考えられる。

第 5 章

顔部品画像からの物理的表情情報抽出

本章では，顔部品画像から表情の物理的特徴を抽出する処理について述べる．

5.1 従来研究とその問題点

これまでに，2次曲線によって顔部品形状をモデル化し，エッジ情報を用いて形状情報を抽出する手法 [21] や，Deformable Templates [20] を用いた形状情報の抽出が試みられている．しかし，表情変化による顔部品の濃淡パターンの変化は複雑であり，2次曲線を用いたモデルでは対応しきれない．またエネルギー最小化原理を用いた顔部品の輪郭追跡手法 [19] が提案されているが，解析の初段階ですべての特徴点を手作業で与えなければならず，処理の自動化にとって大きな問題が残っている．

5.2 フィルタ処理による物理的表情情報抽出

本研究では，形状や濃淡パターンのモデルを用いない，画像のフィルタ処理による輪郭形状抽出手法の構成を試みる．顔部品画像に対し，縦，横方向のグラディエントフィルタを適用し，閾値処理によって，輪郭に対応する部分の縦，横成分の2値画像を得る．次に縦，横成分を合成する．合成画像には輪郭要素以外の不要な要素が残っているので，合成画像に対して収縮処理を行い，目画像では領域形状情報を用いた処理，口画像では最大の面積を持つ要素のみを残すという処理によって不要要素を除去し，輪郭要素を得る．輪郭要素に対して膨張処理を行い，目は細線化，口はラプラシアンフィルタを適用して輪郭形

状を得る．これによって，表情変化に対して柔軟な物理的特徴の抽出が期待できる．

5.3 実験条件

本節では，顔部品輪郭抽出のためのフィルタ処理，2値化，2値化画像の論理計算における条件，パラメータの設定について検討する．

顔部品形状抽出処理に用いる顔部品画像は，領域分割，顔部品推定を行った結果得られた顔部品領域画像とする．顔部品画像は，顔部品分割，顔部品推定の際に用いた正方形であり，その大きさは領域分割の結果によるため，一定ではない．

5.3.1 輪郭抽出法に関する検討

本節では，輪郭抽出を行うために用いるフィルタの種類について検討する．輪郭を抽出するためには，目的とする対象領域（目，口）が他の肌領域と分離して出力される必要がある．顔部品画像の濃淡値は，図 5.1 に示すように，肌部分で細かく変化しており，顔部品の領域は肌部分と比較して低い濃淡値を持っている．これより，肌部分と顔部品領域との境界を得ることで顔部品領域を得ることができる．この境界を求める手法として，本研究ではエッジ検出手法による手法を試みる．

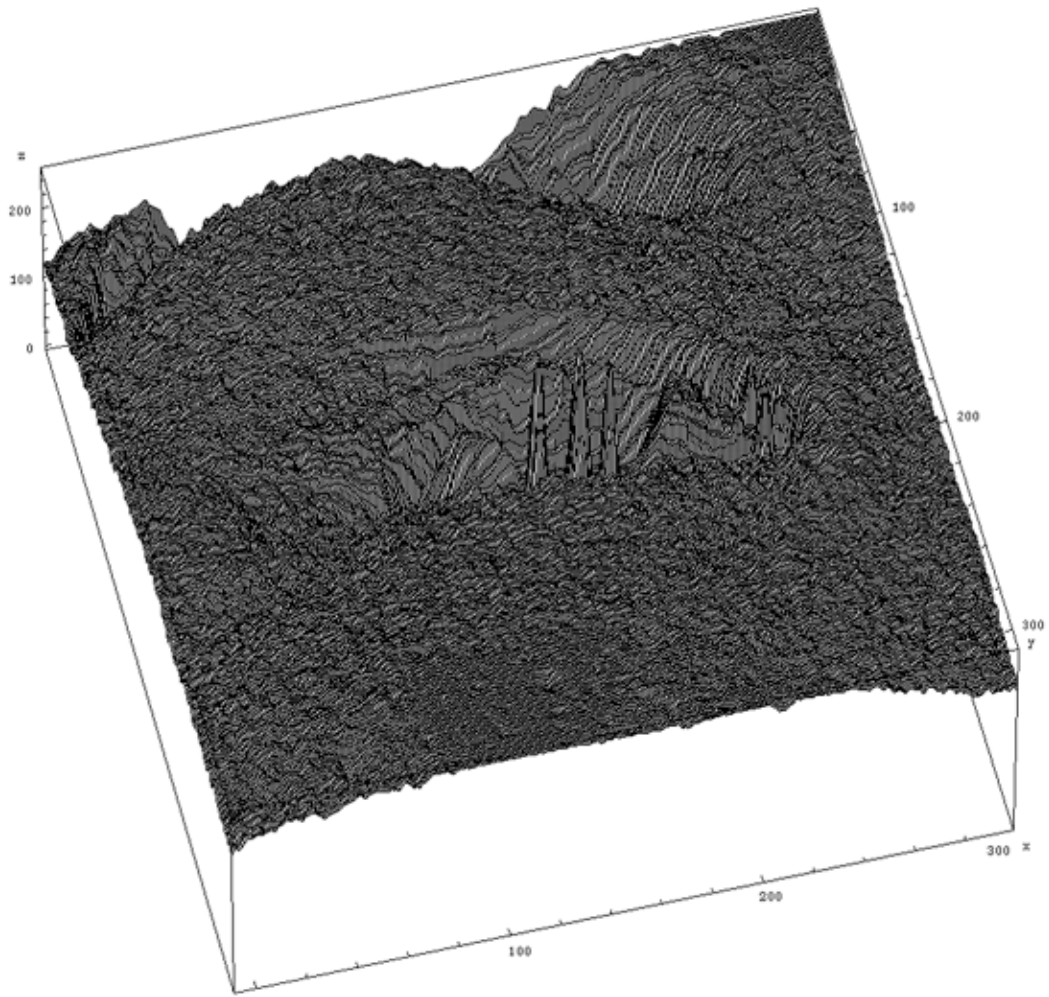


図 5.1: 顔部品画像の濃淡値変化 (画像サイズ: 309 × 309[pixels])

5.3.2 フィルタの種類に関する検討

エッジ検出フィルタとして、ソーベル、ラプラシアン、フーリエ変換による方法等が挙げられる。本節では1次微分に相当するグラディエントフィルタ、2次微分に相当するラプラシアンについて実験し、本研究に適した境界抽出手法を求める。

【1】グラディエントによる境界抽出

グラディエントは1次微分に相当し、 x 方向微分 δx と、 y 方向の微分 δy の2つのオペレータから構成される。3×3の場合の x 方向1次微分にあたるオペレータ δx と、 y 方向の1次微分にあたるオペレータ δy を表5.1に示す。

【2】ラプラシアンフィルタ

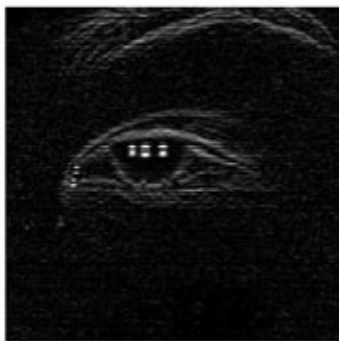
ラプラシアンは2次微分に相当し、そのオペレータは表5.2に示す。

以上のフィルタを用いて領域の抽出を行う。図5.2に目の部分画像を対象にした結果を、図5.3.に口の部分画像を対象にした結果を示す。

以上の結果から、グラディエントフィルタの適用結果の方に、顔部品と肌の境界に出力が現れる傾向が見られる。これより、顔部品の輪郭抽出にはグラディエントフィルタを用いることとする。



原画像



3×3グラディエントフィルタ
(水平方向)適用結果



3×3グラディエントフィルタ
(垂直方向)適用結果

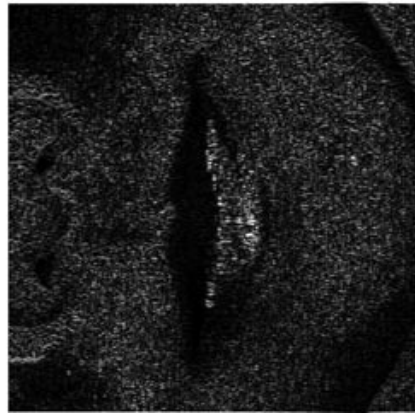


3×3ラプラシアンフィルタ
適用結果

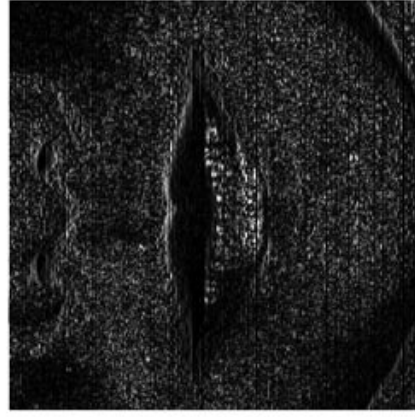
図 5.2: エッジ検出フィルタ適用結果 (対象: 目部分画像)



原画像(423×423[pixels])



3×3グラディエントフィルタ
(垂直方向)適用結果



3×3グラディエントフィルタ
(水平方向)適用結果



3×3ラプラシアンフィルタ
適用結果

図 5.3: エッジ検出フィルタ適用結果 (対象：口部分画像)

-1	0	1	-1	1	-1
-1	0	1	0	0	0
-1	0	1	1	1	1

表 5.1: 3 × 3 のグラディエントフィルタ (左 : x 方向 , 右 : y 方向)

0	1	0
1	-4	1
0	1	0

表 5.2: 3 × 3 のラプラシアンフィルタ

5.3.3 フィルタのパラメータ決定に関する検討

次に , フィルタのパラメータについて検討する .

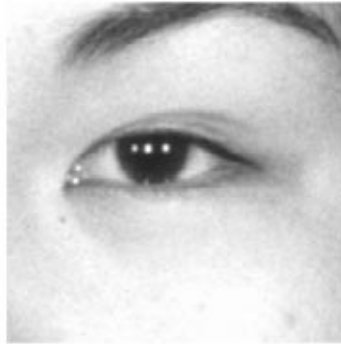
【1】マスクサイズ

ここでは , グラディエントフィルタのマスクサイズについて検討する . フィルタの種類
の検討にて用いた画像を対象に , 図 5. 1~ 図 5. 4に示す 3 × 3 , 5 × 5 , 7 × 7 のグラディ
エントフィルタをそれぞれ適用し , その出力結果から , フィルタのについて検討する . 目
の部分画像を対象にした結果を図 5. 3に , 口の部分画像を対象にした結果を図 5. 4に示す .

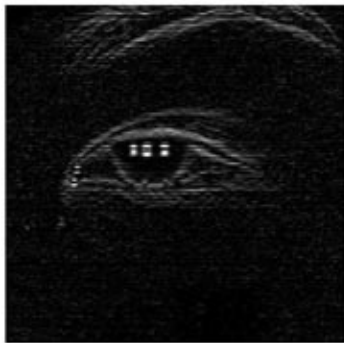
目を対象とした結果から , 5 × 5 のグラディエントフィルタの適用結果が最も目領域の
輪郭に沿って高い出力が出ており , 3 × 3 , 7 × 7 のサイズはそれぞれ輪郭近辺の出力が
小さいか大きすぎることから , 目の部分画像の処理に用いるフィルタのサイズは 5 × 5 と
する .

-2	1	0	1	2	-2	-2	-2	-2	-2
-2	1	0	1	2	-1	-1	-1	-1	-1
-2	1	0	1	2	0	0	0	0	0
-2	1	0	1	2	1	1	1	1	1
-2	1	0	1	2	2	2	2	2	2

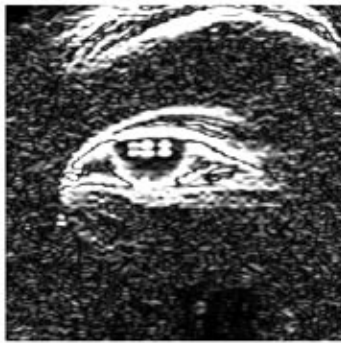
表 5.3: 5 × 5 のグラディエントフィルタ (左 : x 方向 , 右 : y 方向)



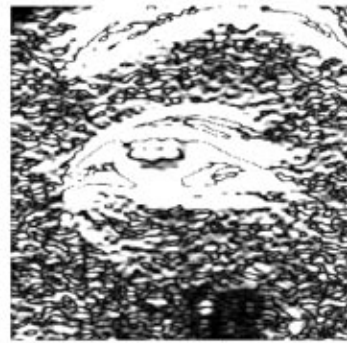
原画像



3×3グラディエントフィルタ
(水平方向)適用結果



5×5グラディエントフィルタ
(水平方向)適用結果



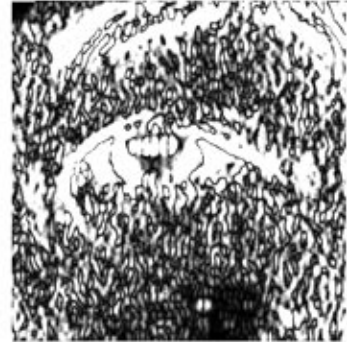
7×7グラディエントフィルタ
(水平方向)適用結果



3×3グラディエントフィルタ
(水直方向)適用結果



5×5グラディエントフィルタ
(水直方向)適用結果



7×7グラディエントフィルタ
(水直方向)適用結果

図 5.3: グラディエントフィルタ適用結果 (対象: 目部分画像)



原画像



3×3グラディエントフィルタ
(垂直方向)適用結果



5×5グラディエントフィルタ
(垂直方向)適用結果



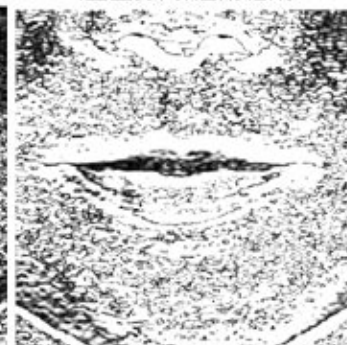
7×7グラディエントフィルタ
(垂直方向)適用結果



3×3グラディエントフィルタ
(水平方向)適用結果



5×5グラディエントフィルタ
(水平方向)適用結果



7×7グラディエントフィルタ
(水平方向)適用結果

図 5.4: グラディエントフィルタ適用結果 (対象: 口部分画像)

-3	-2	1	0	1	2	3	-3	-3	-3	-3	-3	-3	-3
-3	-2	1	0	1	2	3	-2	-2	-2	-2	-2	-2	-2
-3	-2	1	0	1	2	3	-1	-1	-1	-1	-1	-1	-1
-3	-2	1	0	1	2	3	0	0	0	0	0	0	0
-3	-2	1	0	1	2	3	1	1	1	1	1	1	1
-3	-2	1	0	1	2	3	2	2	2	2	2	2	2
-3	-2	1	0	1	2	3	3	3	3	3	3	3	3

表 5.4: 7 × 7 のグラディエントフィルタ (左 : x 方向 , 右 : y 方向)

口を対象とした結果から , 5 × 5 のグラディエントフィルタの適用結果が最も口領域の輪郭に沿って高い出力が出ており , 3 × 3 , 7 × 7 のサイズはそれぞれ輪郭近辺の出力が小さいか口の輪郭を越えて出力が大きくなっていくことから , 口の部分画像の処理に用いるフィルタのサイズは目の場合と同じく 5 × 5 とする .

【 2 】 2 値化閾値

本研究では , グラディエントフィルタによる処理結果から顔部品の抽出を行うために , 画像の 2 値化を行う . 本研究では 2 値画像の値として , 0 と 255 を用いる . 画像の 2 値化は次に述べる収縮処理によって , 顔部品とそれ以外の肌の部分に現れる不要要素が除去できる値を選ぶ必要がある . 本研究では , 入力画像中の画素 (i, j) での濃淡値を $f_{eye}(i, j)$ (目部品画像) , $f_{mouth}(i, j)$ (目部品画像) とするとき , 2 値化による出力 $g_{eye}(i, j)$, $g_{mouth}(i, j)$ は式 5. 1 ~ 式 5. 2 に従って求め , 2 値化を行った .

$$g_{eye}(i, j) = \begin{cases} 255 & \text{for } f(i, j) \geq 230 \\ 0 & \text{other wise} \end{cases} \quad (5. 1)$$

$$g_{mouth}(i, j) = \begin{cases} 255 & \text{for } f(i, j) \leq 230 \\ 0 & \text{ot her wi se} \end{cases} \quad (5. 2)$$

【 3 】 膨張 , 収縮処理

2 値化の結果から , 不要部分の除去を行うために収縮処理を行う . 入力画像中の画素 (i, j) での濃淡値を $f(i, j)$ とすると , 収縮処理による出力 $g_x(i, j)$ は , 式 5. 3 ~ 式 5. 4 から計算される . また膨張処理による出力 $h_x(i, j)$ は , 式 5. 5 ~ 式 5. 6 から計算される .

$$m(i, j) = \sum_{k=-1}^1 f(k, l-1) + f(k-1, l) + f(k+1, l) + \sum_{k=-1}^1 f(k, l+1) \quad (5.3)$$

$$g(i, j) = \begin{cases} 255 & \text{for } m(i, j) = 2040 \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

$$m(i, j) = \sum_{k=-1}^1 f(k, l-1) + f(k-1, l) + f(k+1, l) + \sum_{k=-1}^1 f(k, l+1) \quad (5.5)$$

$$h(i, j) = \begin{cases} 0 & \text{for } m(i, j) = 0 \\ 255 & \text{otherwise} \end{cases} \quad (5.6)$$

【 4 】 2 値画像の合成

次に， x 方向， y 方向のフィルタリング，収縮処理を行った画像を合成する．入力画像中の画素 (i, j) での濃淡値を $f(i, j)$ ， $g(i, j)$ とするとき，合成による出力 $h(i, j)$ は式 5.7 に従って求め，画像を合成する．

$$h(i, j) = \begin{cases} 0 & \text{for } f(i, j) = g(i, j) = 0 \\ 255 & \text{otherwise} \end{cases} \quad (5.7)$$

5.4 顔部品の輪郭抽出

本節では，検討したフィルタとパラメータを用いた，表情顔部品画像の輪郭抽出処理について述べる．

5.4.1 目の輪郭抽出

検討の結果決定したフィルタ，パラメータを用いて，目の輪郭抽出を行う．処理の流れは図 5.5 に示す．ここでは画像に対して水平，垂直方向のグラディエントフィルタを適用した画像を閾値処理によって 2 値化し，縦，横の 2 値化画像を論理演算によって合成する．その結果から不要部分を除去するために収縮処理を行う．目の部分の 2 値画像が複数に分割する場合があるため，不要要素除去では顔画像を領域分割した際の領域形状情報を

用い、領域外要素を除去することとした。不要要素除去結果を膨張処理し、細線化処理によって輪郭線を抽出する。

抽出対象の画像と、抽出された輪郭の例を図 5.6 に示す。

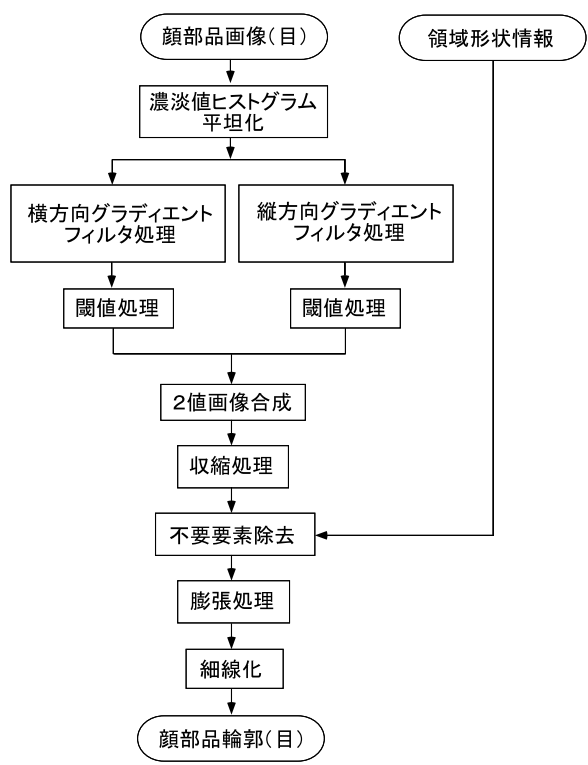


図 5. 5: 顔部品輪郭画像抽出処理 (目)

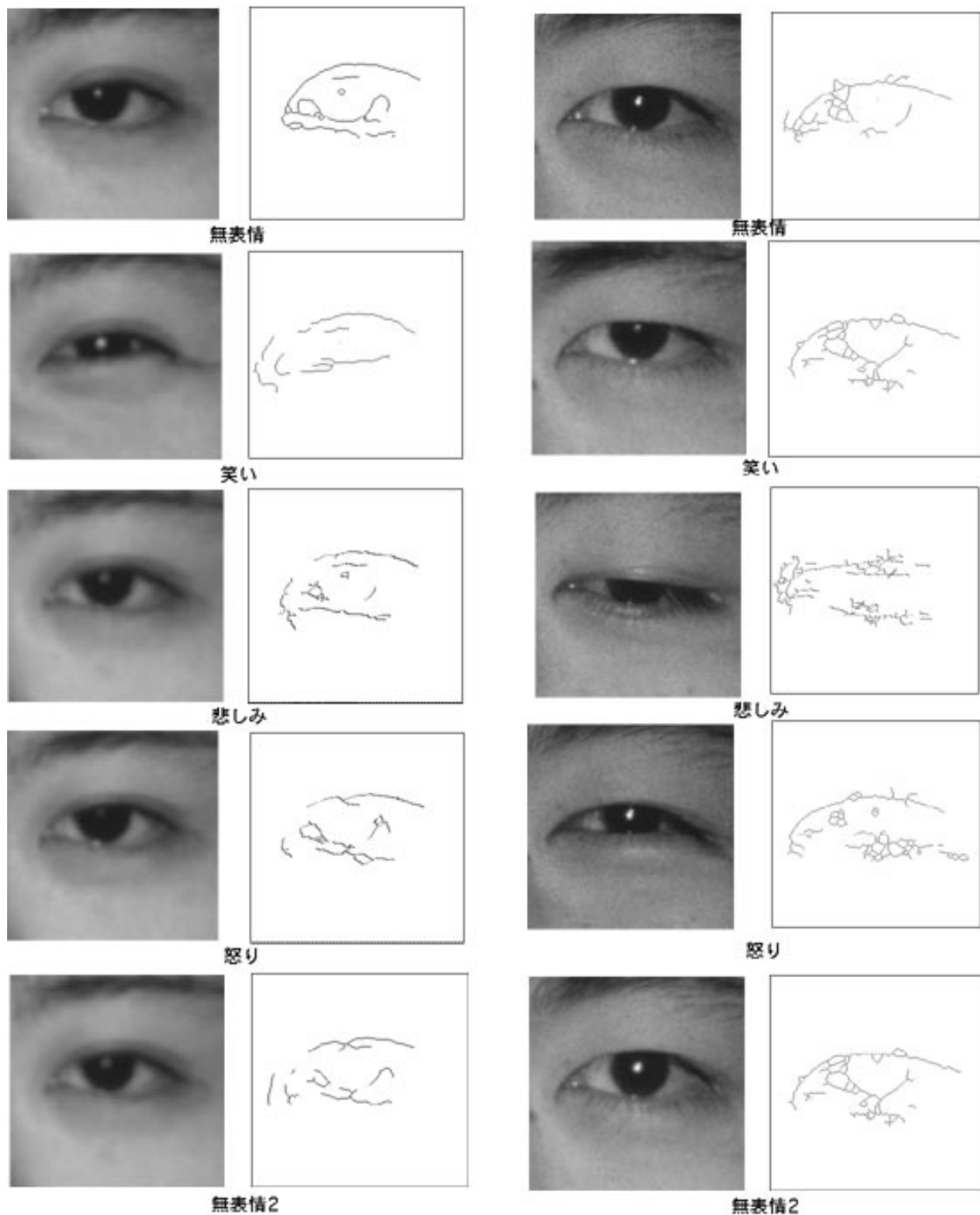


図 5.6: 輪郭抽出処理結果(目)

5.4.2 口の輪郭抽出

検討の結果決定したフィルタ、パラメータを用いて、口の輪郭抽出を行う。口の輪郭抽出の流れを図 5.7 に示す。ここでは、目の輪郭抽出と同様に水平、垂直方向のグラディエントフィルタを適用した画像を閾値処理によって 2 値化し、縦、横の 2 値化画像を論理演算することによって合成する。その結果から不要部分を除去するために収縮処理を行い、不要領域除去を行う。不要領域の除去は、それぞれの領域の面積を計算し、最大のもののみを残し、それ以外を除去するようにした。その結果を膨張処理し、ラプラシアンフィルタを適用して輪郭線を抽出する。

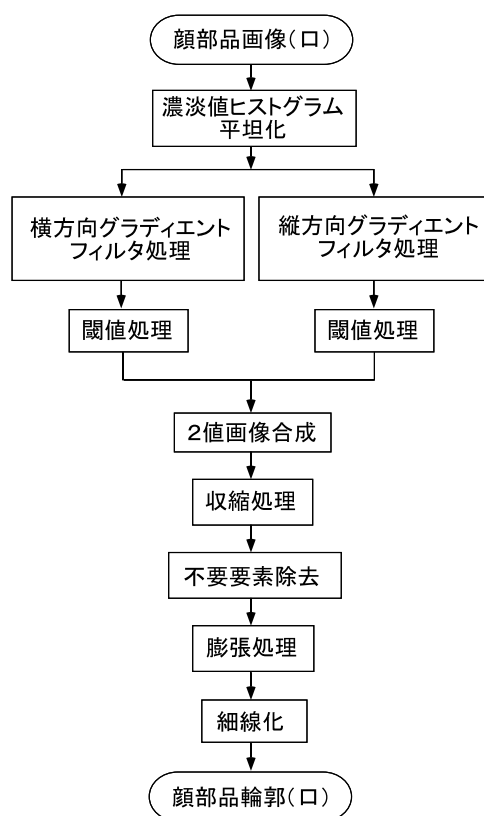


図 5.7: 顔部品輪郭画像抽出処理 (口)

5.5 輪郭抽出結果

抽出対象の画像と，抽出された輪郭の例を図 5.8に示す．

图 5.8: 輪郭抽出処理結果例 (口)

輪郭抽出処理の結果，一名の顔画像を除いた20枚の画像で，輪郭線が抽出できた．抽出できなかった原因として，原画像がぼやけており，フィルタ処理の後の収縮処理でも顔部品領域が肌領域と融合したまま分離できなかったことが挙げられる．

口の輪郭抽出では，下唇部分が途切れたことによって形状が不正確になったもの，上唇と下唇が分離したものは，不要領域除去によって唇が一部欠損し，不完全な領域抽出となったものなどがあつた．しかし画像がぼやけていたもの以外は輪郭線の抽出ができ，表情変化や個人性によらない特徴の抽出が可能であることが示された．

目の輪郭抽出では，目を伏せる，目を閉じる等によって睫毛が目の輪郭にかかる場合，睫毛の影響によって輪郭が複数現れる結果となつた．しかし目が開いている場合，個人性によらず比較的良好な輪郭抽出が行えた．しかし目の外側の輪郭が欠損し，部分的な抽出に留まっている．これはグラディエントフィルタの適用による顔部品と肌の境界が，目の外側部分の出力が低く，閾値処理の閾値以下になつたことが原因であると考えられる．

5.6 回転・大きさの変化に対するロバスト性の検討

本研究にて想定している状況では頭部の移動・回転が発生し得るため，これらに対する輪郭抽出のロバスト性が必要である．これについて，対象画像に対して回転・大きさの変化を加えた画像を作成し，輪郭抽出処理を行うことで，提案手法の回転・大きさの変化に対するロバスト性を検証する．

対象とした画像は，顔部品抽出の回転・大きさの変化に対するロバスト性を検証した際に用いた画像である．この画像に対して，10[度]の回転，20[%]の縮小を行った画像を用い，輪郭抽出処理を施した．原画像と輪郭抽出処理結果を図5.9，図5.10に示す．結果から，画像の回転，大きさの変化に影響されずに輪郭抽出ができていることがわかる．

図 5.9: 回転画像に対する顔部品抽出結果 (10[deg] 回転)

図 5.10: 縮小画像に対する顔部品抽出結果 (20[%] 縮小)

第 6 章

顔部品輪郭形状情報を用いた表情判別

本章では，前章までの処理によって得られた表情の物理的特徴（輪郭形状情報）を用いた表情判別を試みる．

6.1 顔部品輪郭形状情報の記述

これまで行ってきた処理によって，顔部品の輪郭形状として表情特徴が得られた．この情報を用いて表情を扱うためには，輪郭形状情報を構造化して扱う必要がある．曲線の特徴を記述する手法として，チェーン符号によるもの，閉曲線の場合は内接円を用いて記述する手法等が挙げられる．本研究では，処理によって得られた輪郭線が途切れる場合があること，また輪郭線が複雑に入り組んでいることから，特徴記述手法の一例として，部品基準点と輪郭上の点の相対位置を用いて輪郭情報を記述する．基準点と輪郭上の点との相対位置を用いることにより，輪郭線の途切れや局所的な形状に影響されにくい特徴の記述が期待できる．

特徴量抽出処理は，(1) 基準点の探索，(2) 輪郭点の探索，(3) 基準点と輪郭点の相対位置の計算，によって行われる．次にそれぞれの処理について述べる

なお，輪郭線を構成する点を以後「輪郭点」と呼ぶ（図 6.1）．

【1】基準点の探索

本研究において，顔部品の基準点は，顔部品の端点に相当する点とする．輪郭線抽出の結果，目では内側の端部，口の場合は左右端部が比較的良好に得られたので，本研究では目については内側端点，口は処理の都合から左側端点とした（図 6.2）．

図 6.1: 特徴記述に用いる点

基準点の探索は，輪郭画像を y 方向の正の向きに探索する処理を，画像の左上の始点から開始することによって行う．そして探索の結果輪郭点が初めて見つかった点を基準点とする．1 列の探索において輪郭点が発見できなかった場合は，次の列の探索を行う．

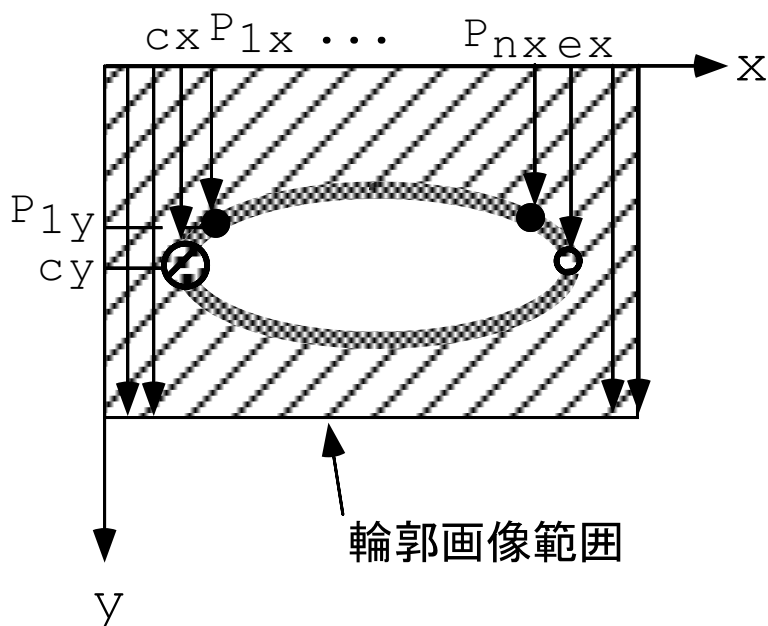


図 6.2: 特徴記述に用いる点

【 2 】輪郭線上の特徴点の探索

輪郭点が検出された後，輪郭形状を表す各輪郭点を探索する処理を行う．各輪郭点の探索は，基準点の探索と同様に 1 列ずつの画素値のチェックによって行われる．

探索が顔部品画像のすべての列に対して行われた場合，それぞれの基準点・輪郭点位置を出力し，それらから輪郭形状の記述を行う．また輪郭線の端となる輪郭点を「終了点」と呼ぶこととする．

【 3 】輪郭点を用いた表情特徴の記述

次に，探索処理によって得られた基準点と輪郭点を用いて表情特徴を記述する．記述は，基準点と各輪郭点との位置関係を用いて記述する．位置関係は，基準点と各列の輪郭点の y 軸方向の差を式 6.4 によって求める．

$$e_i = c_y - P_{iy} \quad (i = 1 \dots n) \quad (6.1)$$

6.1.1 表情識別実験の方法と条件

以上の処理によって記述した表情情報を用いて、表情識別実験を行った。

【A】対象画像

対象とした画像は、これまで輪郭形状抽出処理に用いた顔画像(図4.2.~24.2)のうち、それぞれの画像に対して輪郭線抽出処理を行い、輪郭線が抽出できた4名(男性3名、女性1名)を対象とする。表情は無表情1、笑い、悲しみ、怒り、無表情2とする。無表情1と2は、同じ表情カテゴリに属するが、異なったシーンにて撮影されたものである。

【B】輪郭画像の正規化

今回、輪郭形状情報の記述と、それによる表情識別の可能性を検討するため、原画像の両目内側に設定した基準点(1の基準点探索に用いるのと同じ点)を手作業で抽出し、無表情1を基準にして輪郭画像に対し回転と大きさの正規化を行った。

【C】特徴抽出と記述

特徴の抽出は、先に述べた基準点と各輪郭点の探索によって行う。なお、目については、下瞼側の輪郭線が安定して抽出できなかったことから、上側輪郭線のみを対象にして探索を行った。口については、目とは異なって上下輪郭線が比較的安定して抽出できていることから、上唇側の輪郭線を対象とした探索と、下唇側の輪郭線を対象とした探索を行った。下唇側の探索は、輪郭画像の左下を開始点とし、 y 軸の負の向きに探索を行う。

【D】表情テンプレートの作成

表情判別を行うために、基準となる表情テンプレートを作成する。対象画像のうち、無表情1、笑い、怒り、悲しみの4表情を用いて、4種類のテンプレートを作成した。

なお今回はそれぞれの人物に対して4表情のテンプレートを作成し、それに対して無表情2の表情情報を入力し、判別を行う。

【 E 】 表情判別アルゴリズム

表情判別は，次のアルゴリズムによって行われる．

図 6.3 に示すテンプレートと入力があった場合，基準点 (原点) から x 軸の各点における輪郭点に対し，テンプレートの点と入力点との位置の差 d_{iy} を求める．そして式 6. 4 に従ってテンプレートとの誤差 err を計算し，マッチング度として出力する．なお，輪郭線抽出の結果によっては，輪郭線の存在範囲はテンプレートと入力とが同じであるとは限らないため，マッチング度を求める範囲は，すべてのテンプレートの輪郭線存在範囲と，入力の輪郭線存在範囲の中で最小のものとする．

図 6. 3: 特徴記述に用いる点

$$d_i = T_{iy} - P_{iy} \quad (6.2)$$

$$err = \sum_{i=1}^n d_i^2 \quad (6.3)$$

$$(6.4)$$

以上のようにしてすべてのテンプレートとのマッチング度を求め，マッチング度が最高 (err が最小) となった表情を出力する．

表 6.1: 表情判別実験結果 (左目)

人物	表情			
	無表情	笑い	怒り	悲しみ
1				
2				
3				
4				

: 正しく判定 重複あり : 誤判定

表 6.2: 表情判別実験結果 (口)

人物	表情			
	無表情	笑い	怒り	悲しみ
1				
2				
3				
4				

: 正しく判定 : 誤判定

テンプレートとのマッチングは，目は上瞼側の輪郭線のみを行い，口は上唇側輪郭線と下唇側の輪郭線の2つに対してマッチングを行い，結果を総合して出力する．

6.1.2 実験結果

表 6.1, 表 6.2に，それぞれの人物に対して表情判別実験を行った結果を示す．

表から，目では4例中2例が正しく無表情と判別された．また人物4の結果は，マッチング度が同じ値となったものである．また，口の結果では，正しく無表情と判別されたのは1例のみであった．しかしそれ以外では怒りもしくは悲しみのみであり，変化の大きな笑い表情と判別された例はなかった．

これらから，輪郭形状を用いることで表情判別が可能であることがわかった．

6.2 まとめ

本章では、前章までに述べた方法によって抽出された輪郭形状を用い、表情の判別へと適用した。具体的には、画像処理によって求められた輪郭画像(2値画像)から、顔部品端点を基準点とし、輪郭線上の各点の相対位置を用いて輪郭形状を記述する。この処理を、テンプレートとなるそれぞれの表情カテゴリに属する表情画像に対して行いテンプレートを作成し、入力画像とのマッチング度を求めることで表情の判別を行う。なお、本研究では Ekman らによって提唱されている基本6表情を基にして、(a) 無表情、(b) 笑い、(c) 悲しみ、(d) 怒り、の4表情カテゴリを用いた。

実験の結果、目のみを対象とした判別では、無表情画像を入力した場合、4例中2例が正しく無表情と判別し、無表情と悲しみと同等のマッチング度を示した例が1例、悲しみと判別した例が1例であった。これらから、形状変化の比較的大きな笑いとの区別は可能であることが示され、表情の判別が可能であることが示された。

また口のみを対象とした場合、正しく無表情と判別した例は4例中1例のみであった。しかし他の3例はすべて怒りもしくは悲しみと判別されており、笑いと誤判定した例がなかったことから、笑いとの区別は可能であると言える。

今回は数例のみの画像と表情カテゴリを対象としたが、怒りや悲しみ、無表情といった表情に対する分解能について検討することが今後の課題である。

第 7 章

むすび

本研究では、画像を用いたコミュニケーションツールによるコミュニケーションの効率化・高度化のために必要な顔画像解析アルゴリズムの開発に取り組んでおり、本論文では、顔画像から物理的な表情情報を抽出する手法構成への指針となる、ヒトの表情認識における傾向の調査を行い、その結果を用い構成した、顔画像からの顔部品抽出手法と顔部品の輪郭抽出手法を提案した。

表情認識における傾向の調査は、ディスプレイに提示した表情画像を観察させ、表情を回答させるというタスクを行い、表情観察時における被験者の注視点を求めることを行う。注視点は、アイマークレコーダによって視点を求め、収集された視点情報から視点の停留点を求めることによって得る。そして提示した顔画像に対し、手作業にて求めた基準点から目、口、鼻、頬の部分に領域を設定し、それぞれの領域への注視点の分布傾向を求める。このようにして得られた分布傾向から、ヒトは表情を認識する際に、目、口に注目する傾向があることが確認された。

以上の結果から、目、口から表情の物理的特徴を抽出する手法を提案する。心理的な研究から、目や口の曲線の湾曲度（開き具合）を示す「湾曲・開示度」と、まなじりの持ち上がりを示す「傾斜度」が表情との関連が深いことが示されている。これより、本研究では目、口の部分画像からそれらの輪郭線を求め、物理的特徴とする。

輪郭線を抽出するために、まず顔画像から顔部品の部分画像を求める。顔部品の部分画像を求める手法として、本研究では本田らによって提案されている [11] [12]、画像の粗さ測度を用いた領域分割手法を用い、顔画像を顔部品候補領域として分割する。その後、顔部品間のトポロジカルなモデルである Global Facial Model (GFM) を用い、目、口の部分

画像を抽出する。

画像の粗さ測度として、本田らはフラクタル次元値を用いている。フラクタル次元値とは図形形状の複雑さを表す尺度であり、一般的な画像に対しては Hausdroff 次元と呼ぶ方がふさわしいため、本研究では Haus dr 次元値と呼ぶことにする。

Haus dr 次元は、図形の回転や大きさの変化に対して影響されないという特徴があると言われている。しかし実際の画像解析においてこのような特徴が保存されているかは明らかにされていない部分が少ないため、本研究ではサンプル画像と、それらに対して回転・大きさの変化を行った画像を作成し、実際に次元値を求め、比較検討した。その結果、

- (1) 回転に対してはほぼ影響されず、
- (2) またある程度解像度が高い場合、大きさが変化しても次元値の変化は小さく保たれる。しかし下がりすぎた場合には次元値の値が大きく変動する傾向があることを示した。

また、画像の粗さを用いた顔画像の領域分割において、回転と大きさの変化における限界が示されていないため、これについても画像を作成して領域分割を行うことで検討した。その結果、回転に対しては影響されず、また顔部品が独立して分割されるには、顔の横幅で $12 [pixel]$ を越える解像度が必要であることを示した。

次に、以上の特性を持つ領域分割処理によって得られた顔部品候補領域群から、目的とする目、口領域を得るための手法を提案した。その方法は、

- (1) 顔部品間のトポロジカルなモデルである Global Facial Model を用い、領域の大きさと位置関係をチェックすることによって目領域ペア候補を求める、
- (2) ペアが複数存在した場合、領域の類似度を Haus dr 次元値の変化パターンを用いて比較し、類似度の高いペアをより「目領域らしい」として順位づける、
- (3) それぞれの目領域ペア候補の領域の大きさと位置を GFM に与え、口領域の存在範囲に領域が存在するか否かチェックする。口領域に相当する領域が存在した場合、それらの領域を顔部品領域候補として出力する。
- (4) 顔部品候補が複数組存在する場合、目領域の類似性が最も高いものを顔部品領域として出力する、というものである。GFM は目、口の位置関係のモデルであり、相互関係によってのみ構成しているため、回転や大きさの変化に影響されないという特徴を持つ。本研究では、GFM を構成するために無表情顔画像を対象にし、目、口の位置関係を手作業によって求めた。その結果を用いて GFM を設計し、領域抽出処理を行った。その結果、

個人性，表情によらず顔部品の抽出が可能であることを示した．また，画像の回転や大きさ変化に対しても領域の抽出が可能であることを示した．

顔部品抽出によって得られた顔部品部分画像から輪郭線を求める手法の構成を試みた．部分画像の濃淡値のパターンから，顔部品と周囲の肌とを，縦・横方向のグラディエントフィルタを用いることによって分離し，閾値処理によって画像を2値化する．次にフィルタ処理画像を合成し，収縮処理によって顔部品要素とその他の要素を分離する．次に，目では領域形状情報を用い，口では面積最大の要素以外を除去することによって不要要素を除去する．その結果に対して細線化，ラプラシアンフィルタを適用して輪郭線を得る．

実験の結果，目はすべての画像において，口では原画像がぼやけていたもの以外にて輪郭線が抽出できた．しかし，目が開いている場合には目の内側の輪郭線は抽出できたものの，外側は欠損する結果となり，また睫毛が輪郭にかかる場合，輪郭が乱れ，安定した輪郭抽出は困難であることがわかった．

次に，以上の処理によって得られた物理的特徴を用い，表情の判別に適用した．輪郭画像中に，顔部品の端点を基準点として設定し，輪郭線上の点（輪郭点）との相互位置関係を用いて輪郭形状特徴を記述した．表情判別は，あらかじめ作成したテンプレートとのマッチングを行い，最もマッチング度が高い表情を出力とする．

輪郭の抽出が成功した4名分の画像に対し，無表情，笑い，悲しみ，怒りの4表情カテゴリをテンプレートとして作成し，それに対して別シーンの無表情画像を入力として表情の判別を行ったところ，目のみの特徴を用いた場合，2名において正しく無表情と判別された．また口のみの特徴を用いた場合，無表情と判別されたのは1名のみであった．しかし顔部品形状変化が比較的大きな笑い表情と誤判別された例はなく，表情判別への適用可能性があることが示された．

今後の課題として，

【A】顔部品の位置推定

- 高精度な領域分割手法の構成
- 表情変化を考慮した GFM の設計

【B】表情情報の抽出

- 提案手法以外の，輪郭形状モデルを用いた輪郭線の抽出手法の検討

- 曲線当てはめなどによる輪郭線補完

などが検討項目として挙げられる。

謝辞

本研究を進めるにあたり日頃から熱心に御助言して頂きました本学 小谷一孔助教授に深く感謝致します。本稿は小谷一孔助教授によるご指導なしでは完成することはなく、博士前期課程におけるあらゆる面で多大なるご指導とご支援を賜りました。終始貴重な御意見、御鞭撻を頂きました本学 宮原誠教授、本学 阿部亨助教授に深く感謝致します。

日頃より熱心な議論とご指導を賜り、研究のための計算機環境を整えてくださいました宮原誠・小谷一孔研究室の学生諸氏に深く感謝致します。また表情認識行動特性測定での実験において、快く被験者となってくださいました学生諸氏に感謝致します。

最後に、研究を活動を行うために多大なる支援をくださいました家族、友人に心より感謝致します。

参考文献

- [1] A. Mehrabian 著, 西田 司, 津田 幸男, 岡本 輝人, 山口 輝人 訳: "非言語コミュニケーション", 整文社 (1986).
- [2] B. B. Mandelbrot: "The Fractal Geometry of Nature" Freeman San Francisco (1982)
- [3] S. Peleg, J. Naor, R. Hartl, and D. Avnir: "Multiple Resolution Texture Analysis and Classification" IEEE Trans. Pattern Anal. Mach. Intell., vol. PAM-6, no. 4, pp. 518-523 (1984)
- [4] 吉川 佐紀子, 益谷 真, 中村 真: "顔と心-顔の心理学入門-"サイエンス社 (1993)
- [5] 桶渡 絹二, 大串 健吾, 安田 稔, 斎藤 秀昭: "視聴覚情報概論" 昭晃堂 (1987)
- [6] 谷口 慶治 編: "画像処理工学" 共立出版 (1996)
- [7] 松野 勝弘, 李 七雨, 辻 三郎: "ポテンシャルネットとKL展開を用いた顔表情の認識" 信学会論文誌 D 2 Vol. J77-D 2 No. 8, pp. 1591-1600 (1994)
- [8] 岩澤 昭一郎, 森島 繁生: "モデルフィッティングのための正面顔画像からの特徴点自動抽出" テレビジョン学会技術報告, Vol. 20, No. 41, pp. 43-48 (1996)
- [9] 高橋 公生, 小谷 一孔: "顔画像の粗さ測度による顔部品領域の抽出" 日本顔学会学術大会フォーラム 97, p. 25 (1997)
- [10] 菅生 康子, 松田 圭司, 上野 照剛, 山根 茂: "顔の表情判断に対する注視部位の影響" 日本顔学会学術大会フォーラム 96, p. 20 (1996)
- [11] 本田 明子, 小谷 一孔: "フラクタル分析による顔画像の特徴抽出" 信学会技術報告 CS96-135, IE96-104 (1996)

- [12] 本田 明子： "顔画像のフラクタル解析に基づく領域分割と顔部品位置推定に関する研究 " 北陸先端科学技術大学院大学修士論文 (1997)
- [13] 高橋 公生, 小谷 一孔： "濃淡画像の粗さ測度による画像特徴抽出法の基本特性について " 信学会第 55 会全国大会, 2AB-3, pp.260- 261 (1997)
- [14] Xinguang SONG, Chi-l-Wo LEE, Gang XU, 辻 三郎： "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の抽出 " 信学会論文誌 D- 2, Vol. J77- D 2, No. 8, pp. 1601- 1609(1994)
- [15] 井上 智雄, 岡田 謙一, 松下 温： "ネットワークにおける身振り動作を用いた気持ちの伝達 " 情報処理学会論文誌, Vol. 36, No. 8 pp. 2037- 2044(1995)
- [16] 小濱, 波多野, 斉藤, 吉田： "認知過程における眼球運動の解析 " 信学技報, IE94-18, pp. 9- 16(1994)
- [17] 亀井, 長崎, 関： "手話対話における視線情報解析に関する基礎的検討 " 計測自動制御学会 HI 部会研究会 News&Report, Vol. 13, No. 1, pp. 51- 54(1997)
- [18] 清水嘉重郎： "生物の目とセンサ " 情報調査会 (1985)
- [19] 佐治 斉, 木村 篤史, 太田 寛志, 中谷 広正： "解剖学的知識を用いた顔構成要素の動きの追跡 " 信学会論文誌 D- 2, Vol. J80- D 2, No. 8, pp. 2119- 2128(1997)
- [20] A Yul li eD Cohen and P. Hal li nah "Feat ureExt ract i onfr om Faces Usi ng Def or nabl e Templ at es " Proceedi ngsof CVPR89, pp. 104- 109(1989)
- [21] 中川 督之： "位置と形状の 2 段階探索による顔特徴の自動抽出 " 東京大学工学部修士論文 (1996)