OUKA
Osaka University Knowledge Archive

Title	色彩画像を用いた顔情報処理に関する研究
Author(s)	呉, 海元
Citation	大阪大学, 1997, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3132617
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

色彩画像を用いた顔情報処理に関する研究

呉 海元

大阪大学大学院基礎工学研究科

博士論文

平成９年９月

色彩画像を用いた顔情報処理に関する研究

呉 海元

大阪大学大学院基礎工学研究科

博士論文

平成９年９月

内容梗概

本論文は，色彩画像を用いた顔情報処理の研究について述べるものである．コンピュータをより有用な道具にするために，人間の行動や表情を理解し，それに応じて反応する能力を備えなければならない．その一環として，表現力が豊かで，感情が最も顕著に表れる顔を自動的に捉え，その特徴，動き，表情等を認識することは極めて重要である．

複雑な背景を持つ画像からの顔検出は，ヒューマンインタフェースの分野における重要な研究課題である．従来の研究では画像中の顔の姿勢や大きさなどに関する制約が多く，未知の環境，特に屋外の自然環境において安定に動作する実用的なシステムはまだ見られなかった．又，従来の個人識別や，表情認識等の顔認識に関する研究は大きさと位置が正規化された正面顔画像を対象としていた．従って，それらの手法を実際の画像に適用しようとする時，入力画像の中から顔を見つけ，処理に耐え得る正面顔画像を積極的に獲得する処理が必要である．

本論文は，色情報や顔器官の幾何関係の知識等を階層的かつ全面的に利用して，複雑な背景から，大きさ，位置，姿勢，個数等がともに未知である顔の発見をはじめ，正面顔画像の獲得，顔器官の位置ならびにその輪郭，特徴点の抽出まで，自然な状態でいる人に対しても応用可能な顔認識システムの実現を目指す．

まず，色情報をうまく表現するために，心理学者Fransworthが提案した均等知覚色空間を用いる．多くの顔画像より均等知覚色空間での肌の色の分布，髪の色の分布を調査し，肌の色，髪の色を表現するモデルをそれぞれ構築する．それらを用いて，色彩画像の各点の肌色らしさと髪色らしさを求める．又，ファジイ表現を導入した頭部形状モデルを前もって構築しておく．それを利用して，2項ファジイパターン照合の手法により，肌色/髪色類似度マップを用いて入力画像から，顔が存在する確率が高い部分 (以下，顔候補と呼ぶ) を検出する．

そして，顔候補領域内の肌色領域，髪色領域，頭部領域 (肌色領域，髪色領域を併せた領域) の面積，重心，慣性主軸という一次，二次モーメントを用いて，カメラに対する頭部の三次元姿勢を高速かつ安定に推定する．その結果をもとに能動

的カメラシステムや複数台のカメラを用いた協調システムなどを用いて，正面顔をカメラで追跡する．

最後に，正面顔画像を対象として，色情報(肌と髪の色)および，エッジ情報(強度と方向)を総合的に考慮した加算投影法により，目，眉等の顔器官を抽出する．それらを顔器官に関する幾何関係モデルと比較することにより，顔候補の真偽を検証する．又，得られた各顔器官の大まかな位置を初期位置とした複数本のスネークを配置し，顔器官の輪郭に収束させ，顔器官の特徴点を獲得する．

それぞれの手法に基づいて，実時間に動作する顔検出システム，姿勢推定システム，そして顔部品の輪郭及び特徴点の抽出システムを構築した．実験によって提案した手法の有効性を確認した．

目 次

第1章 はじめに **1**

1.1 ヒューマンインタフェースにおける顔の役割 1

1.2 顔合成・認識の技術動向 2

1.3 本論文の構成 5

第2章 顔と顔器官の検出の従来技術と提案手法 **7**

2.1 顔検出の現状と問題点 7

2.2 顔器官の同定の現状と問題点 9

2.3 提案手法の着目点と概略 11

第3章 ファジイパターン照合を用いた色彩画像からの顔検出 **13**

3.1 色情報による肌色領域と髪色領域の抽出 15

3.1.1 F氏均等知覚色空間 15

3.1.2 肌の色分布モデル 17

3.1.3 髪の色分布モデル 20

3.1.4 肌色領域と髪色領域の抽出 21

3.1.5 肌色類似度マップと髪色類似度マップ 22

3.2 頭部形状モデル 25

3.3 ファジイパターン照合による顔候補の抽出 26

3.3.1 顔候補の抽出 26

3.3.2 肌色/髪色面積の割合のファジイ集合 27

3.3.3 2項ファジイ関係に基づくファジイパターン照合法 28

3.4 顔候補の初期検証 30

3.5 顔の検出実験 32

3.5.1 顔候補の検出実験 32

3.5.2 顔候補の検証実験 35

3.6 まとめ 36

第4章　色彩画像からの人物の頭部の三次元姿勢推定　37
4.1　単軸回転における姿勢推定 38
4.1.1　Z軸回りの回転角度の推定 39
4.1.2　Y軸回りの回転角度の推定 40
4.1.3　X軸回りの回転角度の推定 43
4.2　3軸複合回転への姿勢推定の拡張 44
4.2.1　Z軸回転の推定に対する3軸複合の影響 44
4.2.2　Y軸回転の推定に対する3軸複合の影響 45
4.2.3　X軸回転の推定に対する3軸複合の影響 47
4.3　実験及び考察 48
4.3.1　頭部の三次元姿勢推定実験 48
4.3.2　能動的カメラシステムによる正面顔の獲得シミュレーション . 52
4.4　まとめ 55

第5章　顔器官による顔候補の検証　57
5.1　加算投影法の改良 57
5.1.1　肌色類似度マップと髪色類似度マップ 58
5.1.2　エッジ強度マップとエッジ符合マップ 58
5.1.3　動的閾値 59
5.2　色情報を用いた顔位置と大きさの評価 60
5.3　色とエッジ情報を用いた顔器官位置の決定 61
5.3.1　顔器官の存在しうる領域 62
5.3.2　眉と目 63
5.3.3　鼻と口 63
5.4　顔器官幾何関係モデルによる顔候補の検証 65
5.5　顔器官の決定と顔候補の検証の実験結果 67
5.5.1　顔位置の決定 67
5.5.2　顔器官の位置決定 67
5.5.3　顔候補の検証 68
5.6　まとめ 69

第6章　動的輪廓モデルを用いた顔器官の輪廓と特徴点の抽出　71
6.1　顔器官輪郭の抽出 71
6.1.1　動的輪郭モデルのエネルギ関数 71

6.1.2 顔輪郭について 72
6.1.3 眉について 73
6.1.4 目について 74
6.1.5 鼻について 74
6.1.6 口について 75
6.2 特徴点の決定 75
6.3 顔器官の輪廓と特徴点の抽出の実験結果 76
6.4 まとめ 77

第7章 おわりに **79**

謝辞 **81**

参考文献 **83**

研究業績 **91**

図目次

3.1 CIE色度図上に表示するMacAdamの偏差楕円. 16
3.2 F氏均等知覚色空間の色度図. 17
3.3 肌の色分布モデルおよび髪の色分布モデルを構築するためのサンプル顔画像. 19
3.4 肌の色分布モデル. 19
3.5 髪の色の明るさのヒストグラムと髪の色と無彩色の距離のヒストグラム. 21
3.6 入力画像から推定された肌色類似度マップと髪色類似度マップ. 22
3.7 異なる肌の色分布モデルを用いて推定された肌色類似度マップの比較結果. 24
3.8 基本頭部形状モデル. 26
3.9 肌色類似度マップと髪色類似度マップより評価された照合度マップ. . 30
3.10 顔候補の検証. 31
3.11 実験結果のまとめ(静止画像). 33
3.12 顔候補の抽出の実験結果. 34
3.13 顔候補の検証結果. 35

4.1 顔の姿勢の回転要素 39
4.2 推定された慣性主軸の例 40
4.3 各重心間の水平方向距離 41
4.4 Y軸回転におけるθ_yの推定式 42
4.5 重心間の垂直方向距離 43
4.6 X軸回転におけるS_yの変化 44
4.7 Y軸回転による観測される慣性主軸の変化 45
4.8 慣性主軸と重心の間角ϕ 46
4.9 各θ_yのX軸回転によるS_xの変化 47

4.10 各θ_xのY軸回転によるS_xの変化 48
4.11 姿勢推定の結果 50
4.12 入力画像系列と合成された顔モデル結果 51
4.13 ガイドされたカメラからの画像系列 53
4.14 カメラの移動 53

5.1 定義された7×3のマスク. 58
5.2 定義された頭部の位置と大きさ. 60
5.3 肌色類似度マップと髪色類似度マップより決定された頭部と顔の位置と大きさ. 61
5.4 エッジ強度マップとエッジ符号マップより決定された顔器官の存在しうる領域. 62
5.5 加算投影による顔器官位置の同定. 64
5.6 顔器官幾何関係モデル. 65
5.7 抽出された顔器官と顔器官幾何関係モデルとの照合. 66
5.8 顔器官の抽出結果. 68

6.1 スネークの初期位置（顔）. 73
6.2 スネークの初期位置（眉）. 73
6.3 スネークの初期位置（目）. 74
6.4 スネークの初期位置（鼻）. 74
6.5 スネークの初期位置（口）. 75
6.6 (a)顔器官を囲む矩形領域，(b)スネークの全ての器官の初期位置. . . 76
6.7 (a)スネークの収束過程，(b)求められた輪郭線と特徴点. 76
6.8 輪郭線と特徴点の実験結果. 77

第1章　はじめに

人間とコンピュータとの双方向情報交換によるインタフェースは，将来の情報システムにおいて核となる技術であり，世界的にその重要性が改めて認識され，活発な研究が行われている．その中ではユーザに関する顔情報を得て計算機への補助入力として利用しようとするものや，セキュリティー分野での監視システムや，TV会議システムなどに必要な顔画像の転送などの研究も多数行われている．

これらの技術を実用化するためには，まず処理対象である画像中の顔を，その個人差やサイズ，向き，個数，背景，照明状況などの環境条件に影響されずに抽出することが必要不可欠である．更に，表情認識や個人識別などの処理に耐え得る大きさ，向きの顔画像を積極的に獲得し，顔と顔器官の輪廓や特徴点を求めることも望まれている．これらの問題は顔画像処理における初期の技術課題である．この課題に対し，これまで様々な研究が行われてきたが，表情認識，個人識別における研究と比べて，まだ方法論は少なく，さらに努力する必要があると指摘されている [R. Chellappa 95].

1.1　ヒューマンインタフェースにおける顔の役割

現代の情報社会においては，コンピュータ，通信技術が急速に発展してきた。これらの利用がビジネスのみならず，日常生活にまで深く浸透してきている．また，メディアとしての多種多様な利用も可能となってきている．

一方，コンピュータと人間との間のインタフェースに関しては，GUI等の工夫がなされてきたが，基本的にはキーボード，マウス，ディスプレイの使用が中心であり，人間がコンピュータ側の都合に合わせるという状況には大きな変化はなかった．

我々人間は，文字だけでなく，表情，音声，身振り，手振りといった様々な手段を通じて，コミュニケーションを行っている．相手がコンピュータである場合も，人間同士のと同じようにコミュニケーションを行うことが実現できてはじめて、真のヒューマンインタフェスを構築したと言えよう。

顔は，我々人間にとって非常に身近な存在であると同時に，表情，仕草，発話を

はじめとする様々な情報の発信を担っており，その取り扱いが極めて重要である．人間と情報システムとの円滑なコミニケーションメデイアを実現するために，コンピュータによる顔情報処理の自動化は欠かせないものである．

1.2 顔合成・認識の技術動向

顔情報の処理に関しては，知的画像符合化の研究を一つの契機として，工学分野だけでなく，文学，心理学，人類学，生理学，医学，芸術分野等との連携も含め，世界的に研究が活発化し，成果の蓄積が進んできている．

• 顔合成システム

コンピュータの性能の著しい向上に伴い，コンピュータグラフィック(CG)による顔表情および動きの合成に関する様々な研究開発が行なわれている．

三次元構造モデルを利用した顔画像合成の技術としては，大きく分けて二つの流れが存在する．一つは顔の皮膚下にある筋肉(表情筋と呼ばれる)およびそれに連動して動く皮膚表面を，物理的なモデルとして取り扱い，筋肉・骨の動きをもとに計算によって皮膚表面にあたるワイヤフレームモデルを動かす方法である．もう一つはワイヤフレームモデルに直接規則的な幾何学的変形を施して表情を作成する方法である．前者としては，トロント大学のTerzopoulosらのグループによる筋肉モデル[D. Terzopoulos 93]や，早稲田大学の橋本らによるスプリングフレームモデル[S.Hashimoto 87]が研究されている．後者としては，東大の原島らのグループによる研究[崔 91]，ＫＤＤの金子らによる研究[金子89]，松下の角らの研究[角 91]等が挙げられる．両者とも一長一短がある．物理モデルでは，各種パラメータの調整が微妙かつ困難であり，最終的な出力である顔画像を自然なものとすることが難しい半面，筋肉の配置を変更することでさまざまな顔の造作に対応することが出来る．しかし，その計算量が膨大になることも問題とされている．幾何学的モデルでは，意図する表情を表出させる規則を作成することは簡単であるが，逆に個人性の吸収という問題が表面化する．

モデル変形のための表情の記述法に関しては，Ekmanらが開発したFACS(Facial Action Coding System)を利用することが一般的である[P. Ekman 78]．崔らはFACSの記述単位AU(Action Unit)を36種類コンピュータにインプリメントし表情画像を作成することを実現している[崔 89]．

上記の技術を利用して，顔ロボットを試作した研究例もあった．[原 97]．

• 顔認識システム

顔認識に関する研究は，顔の位置と動きの検出，顔器官およびその特徴点の抽出，個人の識別，表情の認識の各領域に分けることができる．ここで．個人識別と表情認識の技術動向を概説する．

(a) **個人識別：**

個人識別には指紋やパスワードなどを利用することが一般的であるが，顔画像による個人識別は，被験者に負担をかけない手法として古くから注目されてきた[南 86]．萩原らは，人相学的知見をもとに36個の特徴パラメータを求め，日本人男性を用いた個人識別実験を行った[萩原 88]．Sakaguchiらは，三次元情報を反映している等濃線を用い，等濃線が個人により大きく異なることを利用して，パターン照合法，等濃線の方向情報による方法，等濃線の局所密度による方法等を用いて異なる人物像に対して個人識別を行った[T. Sakaguchi 89]．宋らは，自然に撮影された時系列顔画像から，まばたきを利用して目の位置を求め，それに基づき顔領域を抽出し，空間低周波数成分からなるモザイク画像を利用してユークリッド距離による個人識別を行っている[宋 93]．赤松らは，正面顔の濃淡パターンを次元圧縮（フーリエスペクトルのＫＬ展開）し，それを用いたパターン照合により個人識別を行っている[赤松 93]．

被験者の顔の姿勢，サイズなどの制約条件を解消するため，阿部らは，三次元データから顔の向きを推定し，横顔の輪郭線を三次のＢ－スプライン曲線を用いて最小自乗法で近似し，得られる制御点を特徴量として識別を行う方法を提案した[阿部 90]．Ladesらは，入力画像を二次元空間内で変形させてサンプルビューモデルと照合した柔軟な照合法を提案し，表情変化，上下，左右の回転を有する顔画像に対して有効であると報告した[M. Lades 93]．Dadetらは，階層的なガボールウェーブレット変換(GWT)よって局所特徴を抽出し，ダイナミックリンクアーキテクチャ(DLA)による柔軟な照合に基づて，個人別に一枚の正面顔画像をモデルとして登録するだけで，位置，表情，姿勢並びに照明条件などの様々な変化のある入力顔画像からの個人識別方法を提案している[Dadet 97]．

(b) **表情認識：**

顔表情の工学的認識に関する研究は，以下の二つに大別できる．

• 表情筋の動きを推定して表情の分析と認識を行うもの

間瀬らは，オプティカルフロー(時空間画像の輝度の傾きから速度場を求める手法)により表情筋の動きを検出し，ｋ－最近傍法という方法を用い，被験者

の各表情の典型的な顔動画像から得られた特徴ベクトルと，ある顔の動画像から得られたベクトルとの距離により表情を識別している [間瀬 91]．崔らは，AU の動きを推定して表情認識の方法を提案した [崔 91]．AU を用いる方法は，現在のところ顔画像の再合成のための表情の分析に留まる．また，AU による方法において，AU を求めるための特徴点の抽出は人間が行っている．

- **顔の特徴点や特徴量によるもの**

  そのほとんどが特徴点の抽出ができたものと仮定して行われている．Suwa らは，一人の被験者について喜・怒・怖・悲・妬の5種類の感情の認識を行い，平均７９％の正解率を得ている．この研究では，特徴量として，眉間の近傍の三角形の面積，口の開き具合を示す長さ，鼻の中心と上唇の左右両端の上方の点を結ぶ直線が水平方向に対して持つ角度の四つを取り上げた．標準顔におけるこれら四つの特徴量と，ある表情におけるものとのユークリッド距離を用いて表情の識別を行っているが，識別のためのユークリッド距離の判断基準値が明確でない．小林らは，顔画像から 30 特徴点の座標値を人間が入力し，その座標値の「中立」からの移動量を情報として，NN(Neural Network) を用いて 6 基本表情（驚き，恐怖，嫌悪，怒り，幸福，悲しみ）及びその強さの認識認識テストを行った [小林 93a][小林 93b]．松野らは，ポテンシャルネットモデル (二次元グリッド状に相互にバネで連結されたもの) を用い，人間が指定された眉から口を含む顔領域の表情の特徴に合わせて変形させる．求められた変位ベクトルをＫＬ展開することによって次元圧縮し，表情を解析する方法を提案している [松野 94]．坂口らは，HMM(Hidden Markov Model) を用いて 6 基本表情の認識法を提案している [坂口 95]．

1.3 本論文の構成

背景に依存せず，どんな画像からでも姿勢，大きさが未知である顔を発見すると共に，表情認識，個人識別を行う際に望まれる姿勢，大きさの顔画像を，自然な状態でいるユーザに対して積極的に獲得することは，マン・マシン・インタフェースの向上だけでなく，顔画像処理技術の実用化にも貢献できるものである．

本研究では，色情報(主に，人間の肌及び髪の色)を重視する上に処理を行う．色情報をコンピュータ内でうまく表現するために，心理学者Fransworthが提案した均等知覚色空間を (Perceptually uniform color space) 用いることにする．

本論文は七つの章から構成される．次章では顔と顔器官の検出について，従来提案されてきた手法の有効性や問題点を述べ，また本研究が提案する手法について概説する．

第3章では，ファジイパターン照合を用いた色彩画像からの顔検出について論じる．まず，肌色/髪色分布モデルの構築法や，それらを用いた肌色/髪色らしい領域の抽出手法を明らかにする．それから，頭部形状モデルの構築法や，得られた肌色/髪色類似度マップに，提案したファジイパターン照合法を適用し,入力画像から頭部らしい部分を検出する手法について説明する．最後に，実画像を用いた顔の検出実験を通じてその有効性や問題点を考察する．

第4章では，色彩画像からの人物の頭部の三次元姿勢推定について論じる．マン・マシン・インタフェースへの応用のために，マシン前に操作しているユーザの顔を検出してから，その姿勢を高速かつ安定に推定できる方法について説明する．また，実画像を用いた実験結果を示し，考察を行う．

第5章では，顔器官による顔候補の検証について論じる．正面を向いた顔画像を用いて,顔器官の同定アルゴリズムや，人間の顔器官幾何関係の知識を利用した顔候補の検証方法について説明する．実画像を用いた実験結果に対して考察を行う．

第6章では，動的輪廓モデルを用いた顔器官の輪廓と特徴点の抽出について論じる．まず，器官別における複数本のスネークの定義，求められた顔器官を囲む長方形領域よりスネークの初期位置の設定について説明する．それから，スネークの収縮結果による顔器官の特徴点の決定について説明する．また，実画像を用いた実験結果を示し，考察を行う．

第7章では，本論文で述べた手法のまとめ及び全体に対する考察を行う．

第2章　顔と顔器官の検出の従来技術と提案手法

顔情報の処理に関する研究は，60年代からセキュリティーのための個人識別技術を中心に始めたが，90年代から，コンピュータ，通信技術の急速な発展に伴い，世界的にその重要性が改めて認識され，様々な面から再び盛んに行われている．

顔認識の前処理として，ビデオカメラ等で取り込んだ画像から人間の顔を自動的に探索し，認識に耐えうる一定の姿勢，大きさの顔画像を獲得することがある．しかしながら，これまでの研究は表情認識，個人識別，顔特徴の抽出などに集中しており，認識のための顔の探索・位置決めについての検討がまだ不十分である．できるだけ制約条件が少なく，自然な状態で得られる顔画像を対象とした顔の検出から認識までの一貫処理はこれからの研究方向であろう．

2.1　顔検出の現状と問題点

顔の探索・位置決めについては画像認識の一分野として世界的な範囲で研究が行われ，様々な手法が提案されているが，処理対象は主に単純背景あるいは背景既知の濃淡画像であった．ここではそのいくつかの例の概略を述べ．有効な点や問題点などについて概説する．

- **多重解像度画像とのテンプレート照合に基づく方法**

  Burtは，coarse to fineアプローチを利用して，テンプレート照合法を用いて顔を検出した[P. J. Burt 89]．塚本らは，画像中の「暗い」「平坦な」画素の割合により顔領域を探索し，その顔領域(QMF)とテクスチャマッピングにより作成したモデルとの照合を行い，顔の動きを実時間で推定した[塚本94]．Crawらは，Sobelオペレータを使って濃淡画像からエッジの強度と方向を計算し，5階層の画像スケールとテンプレートスケールを利用して顔輪郭を抽出した[I. Craw 87]．小杉は、個人識別への適用を目的とした顔の探索・位置決めの方法として，連続的な多重解像度モザイクパタンを用いて頭部を抽出し，さらにモ

ザイクパタンにより頭部領域内から顔の中心部を検出し，目や鼻部分のヒストグラムからそれらの正確な位置を決定する方法を提案した[小杉 94]．これらの方法は，大きさが未知の顔を探索するのに有効であるが，顔の個人差や姿勢の違いによって生じる画像中の顔の見え方の変化を吸収するために，顔モデルを平均化し，あるいは解像度を落すような処理が必要である．これらの処理によって，顔が本来持っている特徴の大部分は失ってしまうので、背景が複雑である場合，顔が正しく検出できなかったり，背景の一部が顔として検出されたりすることが多かった．

- **顔の輪郭モデルと輪郭周辺の明るさの変化の規則に基づく方法**

  Sakaiらは入力画像から抽出されたエッジマップに色々なサイズの楕円形のテンプレートを照合することによって顔を検出して，予測される目と口のところにエッジが存在するか否かによって，その顔の仮説を確認する方法を提案した[T. Sakai 69]．Kelly氏はプランニングのようなtop-down画像分析法を用いて，エッジ情報及びヒューリステックな経験によって頭部輪郭を抽出する方法を提案した[M. D. Kelly 70]．Govindarajuらは、頭部の右側，左側，髪の生えぎわのところの3本の不連続な曲線よりdeformable templateを構成し，各曲線における四次元のコスト関数を定義する．エッジ情報にモデルを当て嵌めることによって，複雑な背景から顔位置を検出する方法も提案された[V. Govindaraju 90]．これらの方法は，照明状況や画質に直接依存するので画像中の顔のサイズが小さい場合や，画像が不鮮明である場合，正しく処理することは困難である．又，顔の大まかな位置と大きさが既知である場合，その位置や大きさ等の正確な情報を得る処理に向いているが，顔を自動的に検出する処理には不向きである．以上の方法には，明示，あるいは暗黙で画像内の顔がほぼ正面に向いていることを前提にしているので，非正面の顔の検出には不向きであった．

- **ニューラルネットワークによる方法**

  顔画像をニューラルネットワークに学習させ，その出力により顔らしいものの切りだしを行う手法も提案されている[J. J. Weng 93]．この手法は，処理時間がかかり，未学習の顔に対する探索性能には問題があると思われる．

- **色情報を利用する方法**

  肌色の部分を顔領域として抽出する方法も提案された．福田らは，自然画像からの顔領域の抽出の手法として色情報を利用し，まず肌色，黒色の抽出を行って顔領域の候補を抽出し，ファジィ推論で「顔の向き」を推定し，さらにNNで「顔領域」の抽出を行った。室内での10枚，室外での10枚の顔写真を用

いて「顔領域」の抽出を行った結果，室内90％，室外70％の認識正解率を得ている[福田 93]．彼等は，顔画像上の皮膚領域内の1000点のサンプルを選択し，それらの点のRGB情報をR-G, R-Bという二次元空間に投影して，最小自乗法でそれぞれの空間内に1本の直線で近似する．その2本の線上の付近に投影される色は肌色と判定する．この方法は肌色である所が肌色でないと判断され，逆に，肌色ではない所が肌色であると判断される可能性が大きい．また，宮脇ら[宮脇 89]，Imaiら[M. Imai 91]は，肌色部分の抽出が領域分割に基づいて行なうので，しきい値等のパラメータに対して非常に敏感で，安定に抽出できない場合が多かった．

- **その他**

  シーン中から人物動画像の実時間抽出を行い，コンピュータとのインタラクションを実現した例として，間瀬らは髪と顔のコントラストに基づいて，頭部と顔部の領域の変化を画像処理によって抽出し，頭の動作を検出する方法を提案している．そして,「ハイ・イイエ」等のコマンドを頭部動作により入力することによりコンピュータとのインタラクションを実現している[間瀬　91]．長谷川らは，人間の初期視覚機能を模擬したモデルを並列計算機上に実装し，複数の人物が随時出入りする動画像中からカメラの方を向いている人物のみを安定して実時間認識･抽出し，その結果に応じてコンピュータ内の擬人化エージェントを駆動するシステムを開発している[長谷川　94]．複雑な背景を持つ画像から顔を検出する方法がいくつか提案されたが，検出能力は顔の姿勢や背景に依存する場合が多く，安定した検出はできていない[G. Yang 93][X. Song94][A. Tsukamoto 94][H. Zabrodsky 92].

2.2 顔器官の同定の現状と問題点

目，口，眉，鼻などの顔器官は，個人差があり，更に表情の変化や発話等によってその形状が大きく変化する．従って，顔器官を精密に抽出することは困難である．大まかな顔器官の抽出について，今までさまざまな角度から研究されて，テンプレートマッチング，エッジ加算投影，デフォーマブルモデルなどを用いる方法が報告されている．

- **テンプレートマッチングに基づく方法**

  画像処理の主要な手法であるテンプレートマッチングを用いた顔器官の認識

では，顔の向きと表情の変化や，個人差による差異などに対して弱い．個人差を吸収するために，小さな，部分的特徴を記述したテンプレートを使うと [宋 94]，しわやほくろなど雑音となる部分に影響されやすい．一方，それを避けるために大きなテンプレートを使うと，個人差，表情の変化に対応しきれなくなる．

- **デフォーマブルモデルに基づく方法**

  顔器官の全体形状を認識する方法として，Yuilleらは，目，口のそれぞれについてパラメータ化されたテンプレートを用意し，エッジ情報などからマッチング度をエネルギー関数として表し，テンプレートを動的に変形させることで目，口の抽出を行っている [A. L. Yuille 92]．この手法で得られた認識結果は各パラメータの初期値に影響されやすい．また，表情や動きによって顔器官の形状は大きく変わるので，顔器官の全体の形状を一つの処理単位として扱っている限り，各器官の形状を表すパラメータの探索空間が大きくなる可能性がある．

- **顔器官の分布の対称性に基づく方法**

  Reisfeldらは，小規模なローカルな特徴を用いて顔器官を抽出する方法を提案している [D. Reisfeld 92]．この方法は，顔構造の特徴の一つ，つまり顔器官の分布が対称性を持つことに基づいて，ローカルな勾配情報を取り出し，統計的な判別によって顔器官を抽出するものである．この方法によれば，対称となっている顔特徴は抽出できるが，その特徴点がどの器官のものであるかは他の情報を用いなければ判断できない．また，データの候補も多いので，処理時間が長くなる．

- **加算投影に基づく方法**

  坂井ら，Brunelliらは，Integral Projectionという方法を用いている．これは，水平と垂直方向の濃淡またはエッジ画像の累積分布を計算し，それに顔器官の対称性を加え，ヒューリスティックな手法で顔と各顔器官の位置の同定を行う方法も提案されている [坂井 73][R. Brunelli 93]．この手法では，比較的安定した結果が得られるが，画像の濃淡情報だけを使うので，顔の向き，照明条件，背景に影響されやすい．

2.3 提案手法の着目点と概略

複雑な背景を持つ画像から，サイズ，位置，姿勢が未知である顔の検出は，顔認識の分野における重要な研究課題であるが，非常に困難な研究課題である．人間がシーンの中から顔を認識する場合，まず身体の一部として頭部の位置を見極め，正面や側面の顔をとらえ，更に詳細な顔つきを見て判断する．本研究はそのような人間の顔認識の処理過程を模倣し，テレビカメラで取り込んだシーンの中から顔を発見・認識する場合，まず，顔候補を背景から切り出す．次に，見つかった顔候補の姿勢を推定し，顔の認識処理のための正面顔画像を獲得する．更に，顔器官の位置を同定し，顔候補に対して検証を行う．最後に，顔とその器官の輪廓と特徴点を獲得する．

本研究の一つの重要な部分は，色彩入力画像から肌色らしい/髪色らしい領域を抽出することである．肌/髪の色領域の抽出をより安定に行うために，均等知覚色空間を用いて色情報を表現する．多種多様なサンプル顔画像より肌の色，髪の色の均等知覚色空間上での頻度を求める．その統計データに基づいて，肌色分布モデル/髪色分布モデルを構築する．これらの分布モデルを用いて，入力画像中の各点の色 (R,G,B) を均等知覚色空間に変換し，肌色らしさと髪色らしさを求める．その結果を肌色/髪色類似度マップにそれぞれ保存する．このように処理の早い段階で，肌色/髪色と決定的に決めるのではなく，肌色らしさ/髪色らしさということを保存しておくことで雑音に対してロバストになる．

肌色/髪色類似度マップを用いて入力画像から顔を検出するのに，複数の男女の頭部形状の情報を測定し，一般化した頭部形状モデルを構築して利用する．このモデルにも曖昧さを導入し，肌色，髪色，背景，それぞれの確率を保存している．

本研究のもう一つの重要な部分は，従来のパターン照合法の中にファジイ理論を適用することである．2項ファジイ関係を導入したパターン照合法を用いて，種類や意味の異なるもの(頭部形状モデルと抽出された肌色/髪色類似度マップ)を照合させる．この照合度マップのピークを検出し，その値がしきい値より高い場合その位置に顔が存在する確率が高い．この手法は従来のパターン照合法にあったテンプレートと入力データが同種のデータでなければいけないという制限を破り，抽出したいパターンの本質的な特徴を表すモデルを採用できるため，画像のノイズ等の影響を受けにくく，パターンを安定して抽出できる．

複雑な背景の場合，肌色らしい領域の上に髪色らしい領域が存在する可能性がある．この時，顔候補として検出されてしまう領域内で，エッジ情報と色彩情報を加算投影し，その累積分布によって，真の顔か偽の顔候補か初期検証をする．

表情認識，個人識別に耐え得る顔画像を獲得するために，色情報に基づいてカメラに対する頭部の三次元姿勢を推定する．

ほぼ正面を向いた顔画像を対象として，顔とその器官の位置を正確に決定するために，色の空間的な1次微分の大きさマップ，明るさの空間的な1次微分の符合マップ，肌色らしさマップおよび，髪らしさマップを求め，4種類の情報を個々に加算投影し，総合的に考慮する．複数種の情報を融合して利用することによって，異なる照明条件と複雑な背景を持つ画像に対しても安定に処理できるようになる．顔器官が存在するか否かを検証することにより，偽の顔候補を除外することができる．同時に，真の顔に対しては，目，鼻，口などの顔器官の位置を求めることができる．

CGへの入力や個人識別のためのデータベース作りなどのためには，顔とその器官の正確な輪郭を求める必要がある．正しく，速く輪廓を求めるために，既に求められた各顔器官を囲む長方形領域の情報を用いてスネークの初期位置を設定する．又，顔器官の端が尖っているところをシャープに検出するために，複数本のひも状のスネークを収縮させ，二本のスネークの交点を求めることにより，端を正確に求めることができる．

本研究は，自然画像より顔の検出から認識までの一貫する処理の基礎と考えられる．

第3章　ファジイパターン照合を用いた色彩画像からの顔検出

本章は，肌と髪の色情報を用いて色彩画像から顔の自動検出を行うシステムを紹介する．本研究には次の特徴がある．

(1) 色情報を，心理学者Farnsworthによって提案され，今まで画像処理分野に使われていなかった均等知覚色空間(Perceptually uniform color space, 以下F氏均等知覚色空間と呼ぶ)で表現する．人間の目は非常に優れた視覚システムである．人間の色彩感覚と一致する色空間で色を表現することにより，シーンの中の色がより均質かつコンパクトに表現できる．

(2) ファジイ理論を導入することによって，肌色および髪色というような主観的な概念をうまく表現できるファジイ色分布モデルを構築する．顔画像中の各画素の色は，肌色(髪色)か否かという従来の単純な2値分割の代わりに，肌色(髪色)らしさで表現する．このことより，色彩画像から肌色(髪色)らしい領域をしきい値によらず安定に抽出できる．

(3) 2項ファジイ関係を利用するパターン照合法を提案する．従来のパターン照合に基づく顔検出法では，顔のサンプル画像を平均化したものを頭部パターンとして用いられた．しかしながら，平均化処理によって個人差が吸収されると同時に，頭部パターンの重要な特徴も失なってしまうので，顔検出の性能を向上させることが困難であった．本研究では，頭部パターンの特徴をうまく表現するために，頭部の肌の部分と髪の部分の形状を用いて，頭部の形状を表すモデルを構築した．一方，色彩画像から抽出されるのは，各画素の肌色らしさと髪色らしさの情報である．この場合，パターンのモデルと入力画像から得られたものとは異種情報となるため，従来のテンプレート照合法は利用できない．我々は，異種情報を比較するために，新たにファジイパターン照合法を提案し，顔の検出に利用することにした．この手法は従来のパターン照合法にあったテンプレートと入力データが同種のデータでなければいけないという制限を破り，抽出したいパターンの本質的な特徴を表すモデルが採用できるため，画像のノイズ等の影響を受けにくく，パター

ンを安定して抽出できる．

本手法に基づいて，実際に動作する高性能の顔検出システムを構築した．このシステムを用いて，色彩動画像から顔候補を安定に抽出できる．これは，本手法の有効性を示すだけでなく，他の画像処理に利用できる可能性も示唆する．

以降では，3.1で色情報による肌色領域と髪色領域の抽出について，3.2で頭部形状モデルの構造と生成について述べ，3.3でファジイパターン照合による顔候補の抽出について，3.4で顔候補の初期検証について説明し，また，3.5で実験結果を通じて本手法の有効性を示す．そして，最後に，討論とまとめを行う．

3.1 色情報による肌色領域と髪色領域の抽出

3.1.1 F氏均等知覚色空間

本研究において，色情報を利用する目的は，画像中の肌および髪の領域を抽出することである．“肌の色”あるいは“髪の色”という概念は，人間が観察を通じて得た主観的なものである．従って，肌および髪の領域を抽出する場合，人間の色彩感覚と一致する尺度で色を表現すれば，人間の視覚システムのように安定な出力を得ることができると考えられる．人間の目の色彩感覚と一致する色表現の尺度である均等知覚色空間とは，その空間上の二つの色を表す2点のユークリッド距離が，人間の目が感じる色の差と比例する色空間のことをいう．

CIE(Commission Intermationale de l'Éclairage：国際照明委員会)は，X，Y，Zという三つの原刺激を定義した．RGB色空間からXYZ色空間への変換は次の式で行うことができる．

$$\begin{cases} X &= 0.619R + 0.177G + 0.204B \\ Y &= 0.299R + 0.586G + 0.115B \\ Z &= 0.000R + 0.056G + 0.944B \end{cases} \tag{3.1}$$

$$\begin{cases} x &= \dfrac{X}{X+Y+Z} \\ \\ y &= \dfrac{Y}{X+Y+Z} \end{cases}$$

上記の式では，Yは明るさの情報を持ち，xとyは明るさと独立した色の情報を持つ．全ての色がx-y平面では

$$\begin{cases} x &\to [0,1] \\ y &\to [0,1] \end{cases} \tag{3.2}$$

の範囲に入る．CIEのXYZ色表現には，色度差は人間が感じる色の差と比例しない問題点がある．これは，人間が見分けられる最小の色度差は光の波長によって変化するからである．すべての可視光の色をx–y平面に描けば，CIE色度図ができる．同じ明るさを有する刺激の色度を肉眼で見る場合に，見分けることができる最小の色度差の測定結果をCIE色度図上に表示すると，図3.1に示すように，多数の大きさや方向の違う楕円ができる．それらの楕円はMacAdamの偏差楕円と呼び，楕円の中心は標準色であり，楕円上の点は標準色からの人間が見分けられる最小の色度差を持つ色である．図3.1から，標準色の位置によって，最小色度差を表すMacAdamの偏差楕円が大きく変化するので，色度図上で表される二つの色度点間

の距離は，人間の感覚的な色の差に比例しないと言える．現在，画像処理，コンピュータグラフィクス，テレビ放送，および出版等の分野でよく利用されるRGB，YUV，YIQおよびCMY色空間はすべてXYZ色空間からの線形変換なので，同じ問題点が存在する．

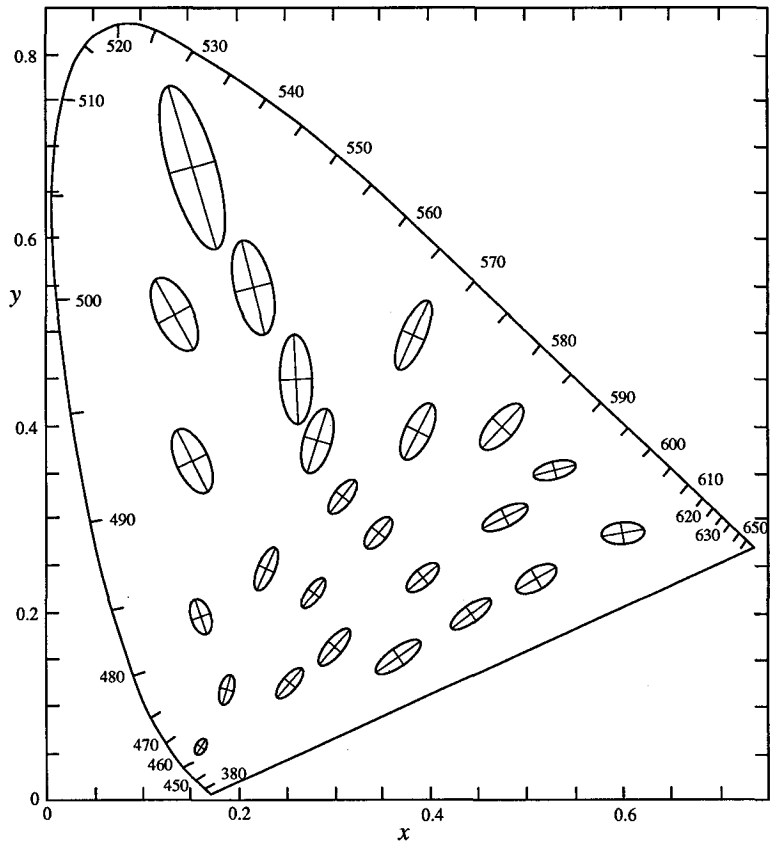


図 3.1: CIE色度図上に表示するMacAdamの偏差楕円.

XYZ色空間から均等知覚色空間を求めようとする試みは，多くの研究者によって発表，提案された．その中のG. Wyszeckiが提案した$L^*u^*v^*$色空間と$L^*a^*b^*$色空間は変換式が簡単で，手軽に利用できるという特徴があり，CIEも推奨している．しかし，その両空間とも近似的な均等知覚色空間であるため，変換後の色空間におけるMacAdamの偏差楕円は均一にならず，ばらつきがかなり残る．一方，心理学者のFarnsworthは一連の心理実験によって得られた結果から一つの均等知覚色空間を1957年に提案した[Wyszecki 67]．この均等知覚色空間の色度図を図3.2に示す．

図3.2から分かるように，この色空間では，MacAdamの偏差楕円は半径がほぼ同じ円になっている．これは，この色度図上で表現した二つの色の距離が人間の感じる色の差に比例していることを意味する．本研究では，F氏均等知覚色空間を採用

して色情報を表現するにする.

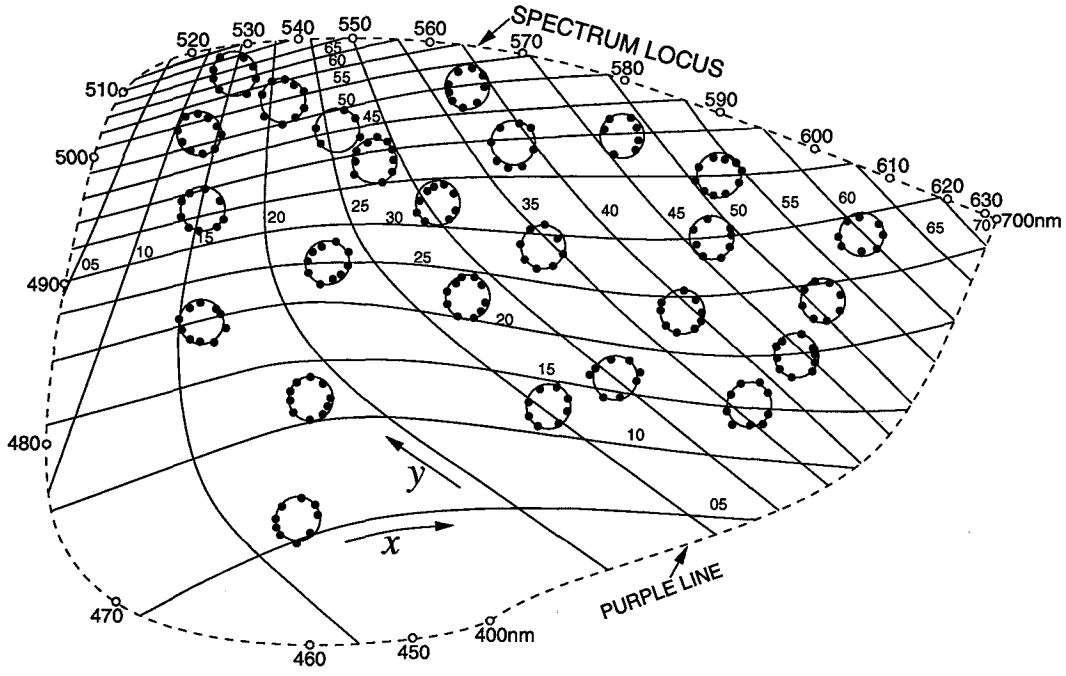


図 3.2: F氏均等知覚色空間の色度図.

我々は，F氏均等知覚色空間をu_f–v_f平面と名付けて，その平面上の点の座標を(u_f, v_f)で表す．XYZ色空間からF氏均等知覚色空間への変換は複雑な非線形変換を必要とする．これでは計算時間がかかるので，F氏均等色度図表より，XYZ色空間とFarnsworth色空間との対応点を全色空間にわたって約1500点について事前に変換を行い，変換テーブルを持っておく．なお，サンプル点以外の色変換はサンプル点からの線形補間により計算する．全ての色はXYZ色空間では式(2)で示す範囲に入るが，F氏均等知覚色空間では

$$\begin{cases} u_f & \rightarrow [0, 91] \\ v_f & \rightarrow [0, 139] \end{cases} \tag{3.3}$$

の範囲に入る.

3.1.2 肌の色分布モデル

肌色は個人差だけでなく，性別，年齢，部位，季節によって変化し，千差万別に見える．しかし，同じ人種の肌の色が色空間上で非常に狭い範囲に集中していることは，種々の研究調査によって明らかになっている．従来の研究では，肌の色の

分布に基づいて色空間上のある範囲内の色を“肌色”と規定し，それ以外は“肌色”でないと分類する．しかし，このような肌色の表現法には，境界上に隣合う二つの色の差がごくわずかにも関わらず，一方は肌色に分類され，もう一方は肌色でないものと分類されてしまうことに，不自然さがある．このことが従来の方法で安定に肌色領域を抽出することが実現できなかった一つの原因である．本研究では，全ての色を肌色とそうでないものに分類する代わりに，各色がどのぐらい肌色に似ているかという度合で表現する．言い代えると，多くの肌色のサンプルから均等知覚色空間上での頻度を求める．頻度の高い色は肌色らしい色で，頻度の低いところは肌色らしくない色であり，これを肌色分布モデル(Skin Color Distribution ModelあるいはSCDM)と名付ける．頻度(すなわち肌色らしさ)は0から1の値に正規化する．

様々な環境で撮影された顔画像を収集し，それらを用いて肌色分布モデルを次のような手順で構築する．

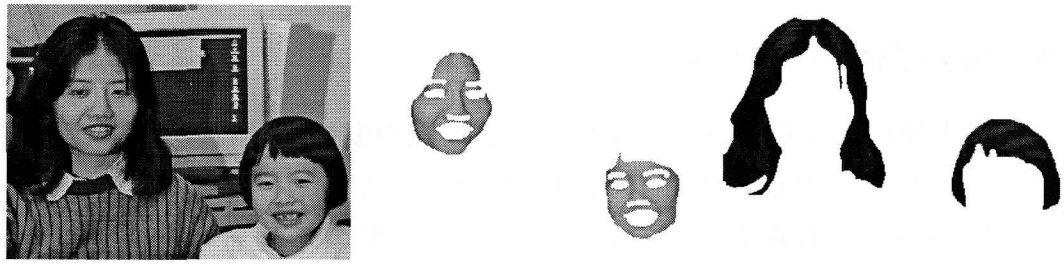
(1) ペイントソフトを利用して顔画像中の肌以外の部分を白く塗りつぶす．(図3.3(b)参照)

(2) F氏均等知覚色空間の中に肌色類似度を記録する表(SCDM，92×140項目)を用意し，0.0で初期化する．

(3) (1)で処理した画像に対して，色が“白”でなければ(肌画素であれば)，その色をF氏均等知覚色空間に変換し，肌色類似度表の対応の項目に1を加算する．

(4) 肌色類似度表の中の最大値を見つけ，全ての項目の数値をその最大値で割ることにより，肌色類似度表の正規化を行う．

こうして構成された肌の色分布モデルは，92×140個の項目を持つテーブルで，F氏均等知覚色空間の各色(u_f, v_f)がサンプル顔画像の中の肌領域に現れる頻度を0.0から1.0の間の値で表現するものである．もし，肌の色分布モデル内のある色の値が0.0であれば，その色はサンプル顔画像の肌領域の中に一回も現れなかったことを意味し，肌色ではないと考えられる．逆に，もし，肌の色分布モデル内のある色の値が1.0であれば，その色はサンプル顔画像の肌領域の中に出現する頻度が最も高いことを意味し，肌色である可能性も高い．従って，肌の色分布モデルを利用して，ある色の肌色らしさを評価することができる．



(a) The original image (b) The skin region selected by viewer. (c) The hair region selected by viewer.

図 3.3: 肌の色分布モデルおよび髪の色分布モデルを構築するためのサンプル顔画像.

本研究の現段階では，東洋人のサンプル顔画像を利用し，東洋人向きの肌の色分布モデルを構築した．図3.3は肌の色分布モデルを構築する際に用いられたサンプル顔画像の一例 (図3.3(a)) と指定された肌の領域 (図3.3(b)) を示す.

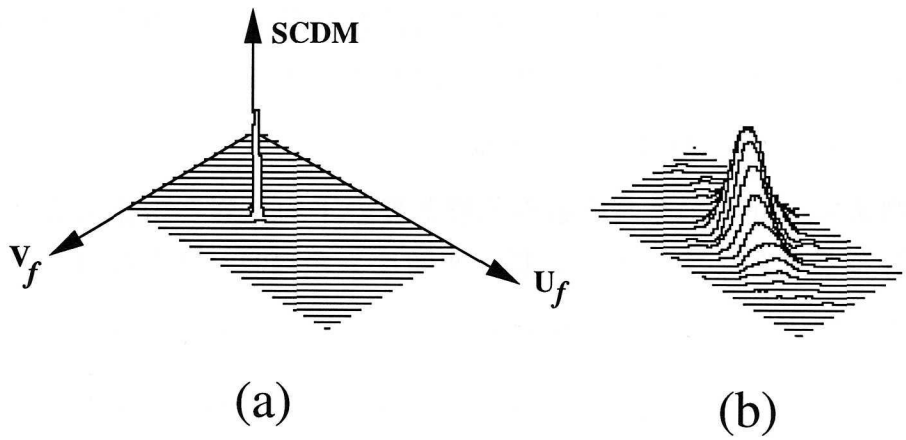


図 3.4: 肌の色分布モデル.

図3.4は，F氏均等知覚色空間に構築された肌の色分布モデル (SCDM) を示している．このモデルを構築するのに，ビデオカメラで撮影した研究室内や，大学構内のシーンの画像と，テレビ放送番組 (ニュース，ドラマなど) から録画した画像を，サンプル顔として男女計38人分を使った．図3.4(a)から分かるように同じ人種であれば，肌色はかなり狭い範囲に集中しているといえる．なお，同図 (b) は同図 (a) の“山”の部分 (0でない部分) を8倍に拡大したものである.

3.1.3　髪の色分布モデル

髪の色は人種によって違いがあるが，同じ人種の髪の色はほぼ同じである．東洋人の場合，髪の色は黒色あるいは黒褐色が多い．このような髪の色と対応する画素は，明度が低く，色成分が不安定であるので，髪の色分布モデルを構築するとき，明度の成分 (y) と色彩の成分 (u_f, v_f) の両方を考慮する必要がある．本研究では，髪の色分布モデルは三つのパラメータ，二つの分布頻度表から次のように構成される．

(1) 髪の平均色彩を計算するための変数(二つの色彩成分 u_f, v_f，髪の画素の数 n)，及び髪の色の明るさの分布頻度表 ($HCDM_y$，256項目) と髪の色彩成分の分布頻度表 ($HCDM_c$，166項目) を用意し，0.0で初期化する．

(2) ペイントソフトを利用して処理した画像(図3.3(c)参照)に対して，色が“白”でなければ(髪の画素であれば)，その色をF氏均等知覚色空間に変換し，明るさを求める．そして，髪の色彩の合計及び髪の毛の画素の数を計算し，髪の明るさ分布頻度表の対応の項目に1を加算する．

(3) 髪の平均色彩 $(\overline{hc_u}, \overline{hc_v})$ を計算する．

(4) 髪の画素であれば，その画素の色彩と髪の平均色彩との距離を計算し，その色彩特徴量に対応する髪の色彩分布頻度表の項目に1を加算する．

(5) 二つの頻度表の中の最大値をそれぞれ見つけ，正規化を行う．

構築された髪の色分布モデル (Hair Color Distribution Model あるいは HCDM) は，次の式で表現できる．

$$\begin{cases} HCDM(y, u_f, v_f) = HCDM_y(y) \times HCDM_c(c) \\ c = \sqrt{(u_f - \overline{hc_u})^2 + (v_f - \overline{hc_v})^2} \end{cases} \tag{3.4}$$

ここで，yは明るさを表し，(u_f, v_f)はF氏均等知覚色空間での色を表す．cは(u_f, v_f)で表す色彩と髪の平均色彩との距離である．

髪の色分布モデルは全ての色の髪の色との類似度を0.0から1.0の数値で表すもので，それを利用して，入力画像内の各画素の髪色らしさを推定することができる．

現段階では，髪の色分布モデルを構築する際に，東洋人向きの肌の色分布モデルに合わせて，同じ東洋人のサンプル顔画像を利用した．図3.5の実線は構築された髪の色の明るさの分布頻度表 $HCDM_y(y)$ を表す．実線に対して，横軸は明るさ (y) を表示し，縦軸は頻度 $(HCDM_y(y))$ を示す．図3.5の点線は髪の色彩と髪の平均

色彩との距離の分布頻度表 $HCDM_c(c)$ を表す．点線に対して，横軸は色差の距離 (c) を表示し，縦軸は頻度 $(HCDM_c(c))$ を表示する．表示のために，最大頻度の1.0を255に対応させている．

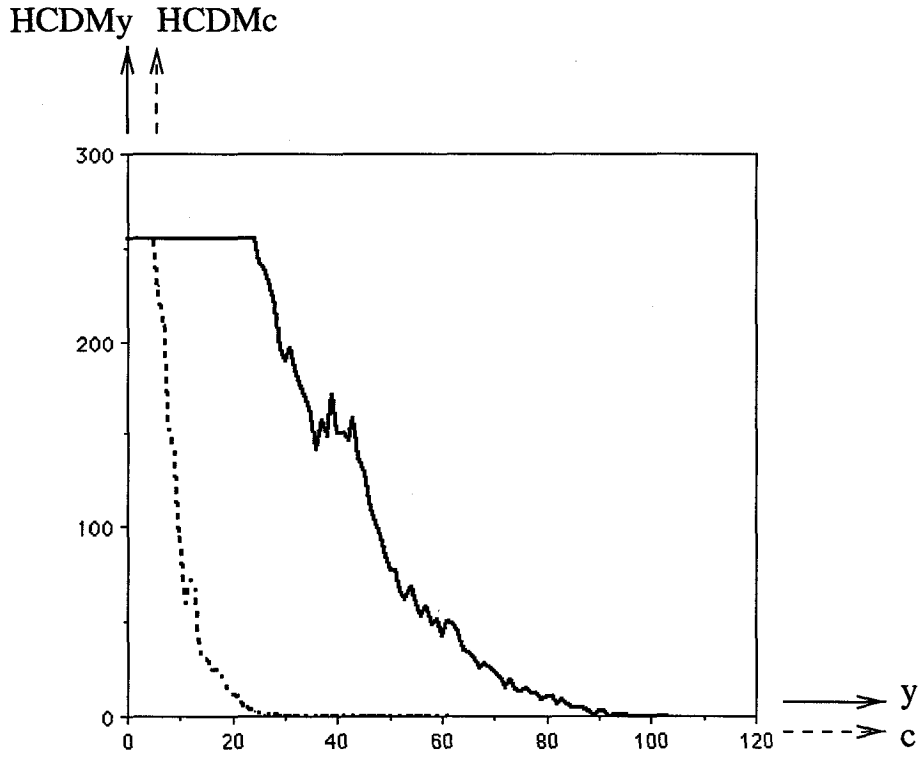


図 3.5: 髪の色の明るさのヒストグラムと髪の色と無彩色の距離のヒストグラム．

3.1.4 肌色領域と髪色領域の抽出

複雑な背景を持つ入力画像から安定に顔を抽出するには，肌色情報と髪色情報をともに利用する必要がある．本研究では，まず，入力画像の色を3.1.1で述べた方法でF氏均等知覚色空間に変換する．そして，3.1.2と3.1.3で述べた方法で構築された肌の色分布モデルと髪の色分布モデルに基づいて，入力画像の各画素の色の肌色らしさと髪色らしさを評価することによって，肌である可能性の高い肌色の部分と，髪である可能性の高い黒い部分を抽出する．

入力画像から肌色領域/髪色領域を安定に抽出するために，各画素の色が肌色/髪色であるか否かという，いわゆる単純な2値化の代わりに，その色がどのぐらい肌色/髪色に似ているかを評価し，0.0から1.0までの連続な値で肌色らしさ/髪色らしさを表現する．この値を肌色類似度(Skin Color Similarity あるいは SCS)/髪色類

似度 (Hair Color Similarity あるいは HCS) と名付ける．入力画像より評価された各画素の肌色/髪色類似度は，肌色類似度マップ (Skin Color Similarity Map あるいは SCSM)/髪色類似度マップ (Hair Color Similarity Map あるいは HCSM) という二次元配列を用いて表現する．入力画像の色を肌色/髪色類似度で表現することによって，処理の早い段階で，髪色，肌色と決定的に決めるのではなく，肌色らしさ，髪色らしさということを保存しておくことで雑音に対してロバストになる．又，色のセグメンテーションのためのしきい値問題がなくなる．

3.1.5　肌色類似度マップと髪色類似度マップ

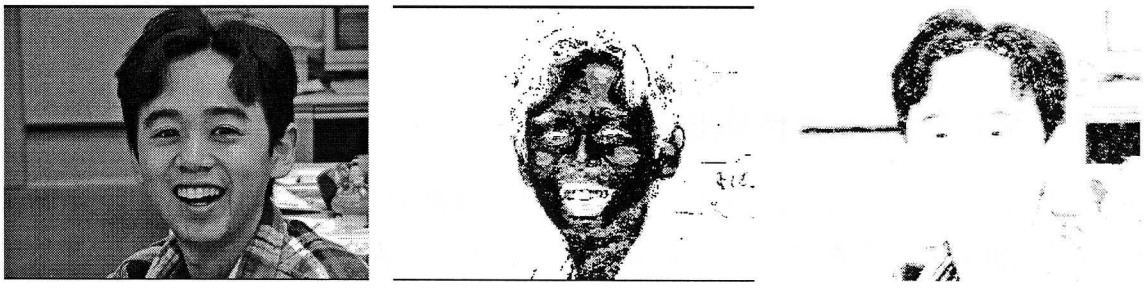
入力画像中のある画素 p の肌色類似度 ($SCS(p)$) を求めるには，その色と対応する肌の色分布モデルの値 (SCDM($u_f(p), v_f(p)$)) を調べればよい．

$$SCS(p) = SCDM(u_f(p),\ v_f(p)) \tag{3.5}$$

又，入力画像中のある画素 p の髪色類似度 ($HCS(p)$) を求めるには，定義した髪の色分布モデル ($HCDM(y(p), u_f(p), v_f(p))$) に基づいて次のように推定する．

$$HCS(p) = HCDM(y(p),\ u_f(p),\ v_f(p)) \tag{3.6}$$

ここで，$y(p)$ は画素 p の明るさであるが，$(u_f(p),\ v_f(p))$ はF氏均等知覚色空間における画素 p の色である．



(a) Original image.　(b) Estimated the skin color similarity map.　(c) Estimated the hair color similarity map.

図 3.6: 入力画像から推定された肌色類似度マップと髪色類似度マップ．

図3.6には入力画像から推定された肌色類似度マップと髪色類似度マップの一例を示す．図3.6(a)は色彩入力画像である．図3.6(b)，(c)はそれぞれ肌色類似度マップ，髪色類似度マップであるが，ここでは各画素の肌色/髪色類似度を濃淡で表しており，黒いほど肌色/髪色類似度が高い．

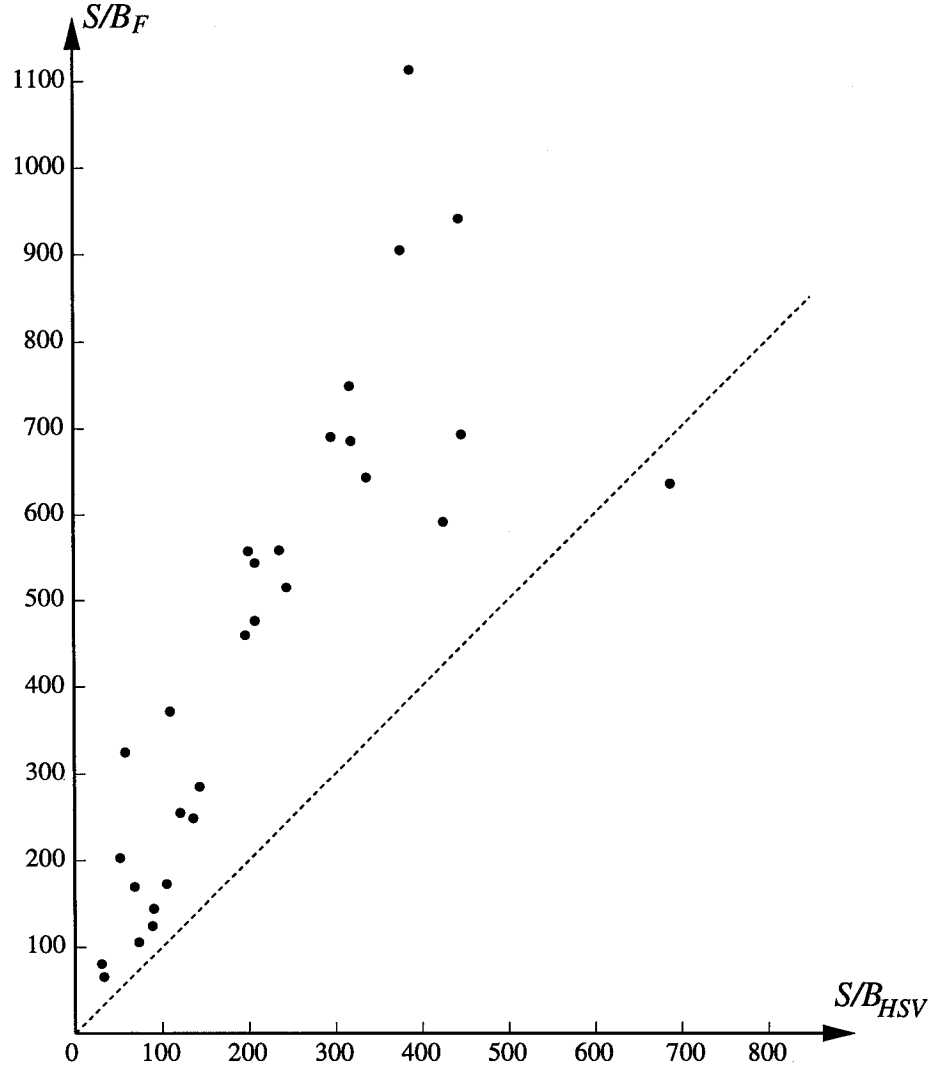
色情報を表現する色空間として，人間の目の特性と一致するものがより適切であるということを示すために，我々は色情報処理に広く利用されているSmith氏が提案したHSV色空間を比較対象とし，その色空間においても，同じサンプル顔画像を利用して，同様に肌の色分布モデルを構築した．F氏均等知覚色空間，HSV色空間を利用して肌色領域を抽出する性能を比較するために，まず，色空間別に構築された肌の色分布モデルを用いて，色彩入力画像から色空間別の肌色類似度マップを推定する．また，画像中の肌領域は人間が目視で抽出しておく．そして，得られた肌色類似度マップより，肌領域内の肌色類似度の平均値と領域外の肌色類似度の平均値を色空間別に計算して，比較実験を行った．比較実験に用いた29枚の入力画像には，実験室内で蛍光灯照明と白熱灯の照明で撮影したもの，屋外で撮影したもの，そしてテレビ放送から録画したニュース番組，ドラマ，コマーシャル等のシーンを含む．

色空間別に推定された肌色類似度マップの比較結果は図3.7に示す．図3.7(a)には，3枚の画像の比較結果を表で示し，各画像に対して，推定された肌の部分の肌色類似度の平均値(Skin)とそれ以外の部分の肌色類似度の平均値(Back)，そして，その比(S/B)を利用した色空間別に示す．図3.7(b)には，全ての画像においてのF氏均等知覚色空間内のS/BとHSV空間内のS/Bとの関係を点で示す．図3.7(b)の横軸はHSV空間のS/Bであり，縦軸はF氏均等知覚色空間のS/Bである．図3.7(b)では，一点を除いて，全ての点が原点を通る45度の直線の左上にある．これで，F氏均等知覚色空間を利用する方が，HSV色空間より，人間の肌である部分の肌色類似度の評価が高く，そうでない部分の肌色類似度の評価が低いという傾向が明示された．

この原因は次のように考えられる．HSV色空間の色度図上で表される二つの色度点間の距離は，人間の感覚的な色の差に比例しない．従って，人間がほぼ同じように見える二つの色が，HSV空間においては，その距離が大きくなることがある．その結果，人間がほぼ同じ色に見える大多数の肌領域の画素は，HSV空間において広範囲な分布になり(頻度が低くなる)，肌である部分の肌色類似度は低くなるという望ましくない現象が現れる．以上のことより，F氏均等知覚色空間はCIE非線型色空間より，シーンの中の肌色がより均質かつコンパクトに表現できることがわかった．これで，F氏均等知覚色空間を利用することは，肌色の処理にとって有効である言える．

		Skin	Back	S/B
Image1	F	0.5668	0.0023	248.5
	HSV	0.2287	0.0016	136.3
Image2	F	0.6224	0.0078	80.0
	HSV	0.3247	0.011	30.5
Image3	F	0.6263	0.0031	203.1
	HSV	0.3450	0.0066	52.1

(a)



(b)

図 3.7: 異なる肌の色分布モデルを用いて推定された肌色類似度マップの比較結果.

3.2 頭部形状モデル

入力画像中の人間の頭部形状は，個人差及びカメラに対する姿勢によって様々に変化する．しかし，顔の詳細部分を省略すれば，画像中の頭部形状は次のようなパターンになる: 中央には肌の部分，上には髪の部分があり，両側に髪あるいは背景の部分がある．これによって本研究では，画像中の人間の頭部形状を，その頭部の肌の部分と髪の部分の相対形状情報を用いて表現する．その形状情報は，いくつかの2次元モデルによって表現される．

頭部の回転は，左右に振る，上下に振る，そして左右に傾けるという三つの成分があるが，ヒューマンインターフェースのための研究や応用において，水平方向の回転についてのみ考えればほぼ十分であると思われる．我々は画像に現れる頭部の形状を，正面，左斜め向き，左向き，右斜め向き及び右向きの5種類に分類する．それに応じて5つの$m \times n$のセルを持つ基本頭部形状モデルを構築する．正面頭部形状モデルには10×13個のセルがあり，その他の側面頭部形状モデルには10×12個のセルがある．モデル中の各セルは二つの属性を持つ．

1. M_F：セル全体に占める肌の部分の面積の割合．
2. M_H：セル全体に占める髪の部分の面積の割合．

セル中の肌ではなく，髪でもない部分は，背景と定義する．背景の割合が下記の式で計算できる：

$$M_B = 1.0 - M_F - M_H$$

頭部形状の個人差を吸収するために，二つの工夫をした．まず，モデルの中のセル数を少なめに設定し，モデルの解像度を低くする．そうすると，顔の形と髪のスタイルの小さな変化が吸収できる．更に，複数の男女の頭部形状の情報を測定し，平均化することにより一般化したモデルを構築し，顔の形や髪のスタイルなどの大きな変化を吸収する．基本頭部形状モデルは次のような手順で構築する．

(1) 画像から頭部を含む矩形領域および，その中の肌領域と髪領域を指定する．

(2) 指定された頭部矩形領域に対して，$m \times n$に等分する．各セルに含まれる肌の画素の割合と髪の画素の割合を調べる．

(3) 集めた複数人分の正面，左向き15度と右向き15度の顔画像に対して，(2)で求めたデータを用いて，各対応するセル内の肌の画素の割合および髪の画素の割合の平均値を計算して，正面の基本頭部形状モデルを生成する．(図3.8(c)参照)

(4) 同様に，集めた複数人分の左(あるいは右)向き30，45度の顔画像のデータを用いて，左(あるいは右)斜め向きの基本頭部形状モデルを生成する(図3.8(b)，図3.8(d)参照)．又，左(あるいは右)向き60，75，90度の顔画像のデータを用いて，左(あるいは右)向きの基本頭部形状モデルを生成する(図3.8(a)，図3.8(e)参照)．

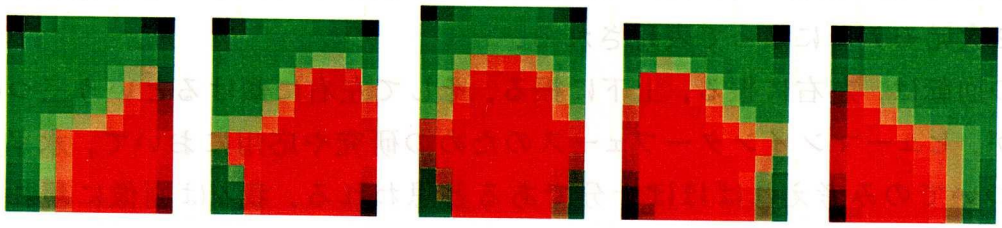


図 3.8: 基本頭部形状モデル．

図3.8は，男性5人，女性2人の顔画像系列を使って生成した5種類の基本頭部形状モデルを示している．基本頭部形状モデルのセルの大きさを変えられる．同図に示したように，セル全体に占める肌(髪)の部分の面積の割合を赤(緑)成分の大きさで表す．

3.3 ファジイパターン照合による顔候補の抽出

本研究では，画像から顔を検出するために，パターン照合法を利用する．顔パターンは，3.2に述べた頭部形状モデルを採用する．一方，照合対象は，画像から推定した肌色類似度マップと髪色類似度マップである．頭部形状モデルと肌色類似度マップ及び髪色類似度マップというような異種情報をうまく比較するために，我々は，ファジイ理論に基づいたパターン照合法を提案する．本章では，2項ファジイ関係を導入したパターン照合法(以下ファジイパターン照合法と呼ぶ)を紹介し，それを用いて入力画像から顔候補を切り出す方法を説明する．

3.3.1 顔候補の抽出

入力画像中の顔の個数，大きさ，位置，向きが未知なので，顔候補を検出する際，まず，肌色類似度マップと髪色類似度マップを推定する．そして，それらの情報に，

5種類の頭部形状モデルの大きさを変化させながらマッチングさせ，画像中の頭部らしい矩形領域を見つけ，顔候補を発見する．異なるサイズの顔を抽出する処理は互いに独立で処理できるので，並列マシンを使えば効率良く行うことができる．

3.3.2 肌色/髪色面積の割合のファジイ集合

頭部形状モデルには，頭部の形状情報が肌部分と髪部分の面積の割合としてセル単位で保存される．従って，頭部形状モデルを用いて入力画像とパターン照合を行うために，入力画像内の矩形領域中のある正方形領域内の肌色/髪色面積の割合を評価する必要がある．本研究では，得られた肌色/髪色類似度マップより，正方形領域内の平均肌色/髪色類似度を計算して，肌色/髪色面積の割合を推定する時，人間の主観的な曖昧さが存在するために，ファジイ理論を導入している．ファジイ集合は次のように定義される．

ある集合Aにおけるファジイ集合Xとは，次のようなファジイメンバーシップ関数μ_Xによって特性づけられた集合である．即ち：

$$\mu_X : A \to [0,1] \tag{3.7}$$

要素$a \in A$に対して，値$\mu_X(a) \in [0,1]$は，aがファジイ集合Xに属する度合を表す．

我々は，集合Aを肌色/髪色類似度マップの中の正方形領域内の平均肌色/髪色類似度の集合A_{scs}，A_{hcs}と定義し，集合Xを肌色/髪色面積の割合のファジイ集合F，Hと定義する．そうすると，肌色/髪色面積の割合のファジイ集合は次のようなファジイメンバーシップ関数によって特性づけられた集合となる．

$$\mu_F : A_{scs} \to [0,1]$$
$$\mu_H : A_{hcs} \to [0,1]$$

メンバーシップ関数μ_Xの決め方は，ファジイ理論における重要な課題であるが，一般的には，人間が経験的に決めれば良い．直観的に，肌色/髪色類似度マップの中の正方形領域内の平均肌色/髪色類似度が高いほど，それにおける肌色/髪色面積の割合が高い．更に，肌色/髪色類似度マップの中の正方形領域内の平均肌色/髪色類似度がほぼa以上の時，それにおける肌色/髪色面積の割合が1になると考えら

れる．従って，ファジイメンバーシップ関数μ_Xは次のようなS型標準関数を用いる：

$$S(x;\ a,\ b) = \begin{cases} 0 & x \le a; \\ \dfrac{2(x-a)^2}{(b-a)^2} & a < x \le \dfrac{(a+b)}{2}; \\ 1 - \dfrac{2(x-b)^2}{(b-a)^2} & \dfrac{(a+b)}{2} < x \le b; \\ 1 & b < x \end{cases} \tag{3.8}$$

ここで，係数aとbは0から1までの実数である．aとbのいろいろな組合せを用いて，顔候補を抽出する実験をし，その結果，肌に関する系数a,bをそれぞれ0と0.6に設定し，髪に関する系数a,bをそれぞれ0と0.75に設定する．これで，肌色/髪色面積の割合がそれぞれ次ぎのように得られる．

$$\begin{aligned} \text{肌色面積の割合} &= \mu_F(a_{scs}) = S(a_{scs};\ 0,\ 0.6) \\ \text{髪色面積の割合} &= \mu_H(a_{hcs}) = S(a_{hcs};\ 0,\ 0.75) \end{aligned}$$

ある小さい正方形領域内の平均肌色/髪色類似度はその領域内の全画素の肌色/ 髪色類似度 (SCS(p), HCS(p)) の平均値と定義する．

$$a_{scs} = \frac{\sum_{p \in area} SCS(p)}{N} \tag{3.9}$$

$$a_{hcs} = \frac{\sum_{p \in area} HCS(p)}{N} \tag{3.10}$$

ここで，Nは正方形領域内の画素数である．

上に定義したファジイメンバーシップ関数によって，この領域内の肌色/髪色面積の割合が次のように推定できる:

$$\begin{aligned} D_F &= \mu_F(a_{scs}) \\ D_H &= \mu_H(a_{hcs}) \end{aligned}$$

3.3.3　2項ファジイ関係に基づくファジイパターン照合法

画像中のある正方形領域と頭部形状モデル中の対応するセルとが“類似している”とは，画像から推定されたその正方形領域内における肌色面積の割合および髪色面積の割合 ($\mathbf{D} = (D_F, D_H)$) と，それぞれ頭部形状モデル内の対応するセルにおける肌色画素の割合および髪色画素の割合 ($\mathbf{M} = (M_F, M_H)$) が“ほぼ等しい”ことを意味する．

ファジイ理論では，2つの実数値の集合x_1とx_2が“類似している”度合を，2項ファジイ関係で記述している．2項ファジイ関係AEは例えば次のようなメンバーシップ関数と考えられる．

$$AE(x_1,x_2) = e^{-a|x_1 - x_2|^b} \tag{3.11}$$

上式におけるa,bは適当な定数である．このメンバーシップ関数の値は，x_1とx_2がほぼ等しい場合1に近付き，逆に，差が大きくなるにつれ0に近付く．

x_1:$\mathbf{D}$および，x_2:$\mathbf{M}$と定義すれば，上記のファジイメンバーシップ関数AEを利用して，入力画像中のある正方形領域($\mathbf{D} = (D_F,D_H)$)と頭部形状モデル中の対応するセル($\mathbf{M} = (M_F,M_H)$)との照合度を下記のように評価できる．

$$\begin{aligned} match(cell) &= AE(\mathbf{D},M) \\ &= e^{-a\|\mathbf{D} - \mathbf{M}\|^b} \\ &\triangleq e^{-a(|D_F - M_F| + |D_H - M_H|)^b} \end{aligned} \tag{3.12}$$

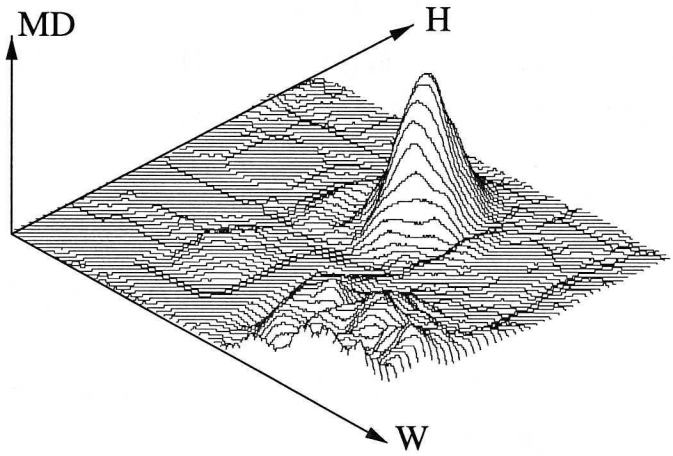
我々は，aとbのいろいろな組合せを用いて，顔候補の抽出実験を行い，その結果，aとbをそれぞれ0.5と1と決めた．

入力画像中の矩形領域と頭部形状モデルとの照合度は，画像中の矩形領域内の各正方形領域と頭部形状モデル内の各対応するセルとの照合度の合計である．

$$Match(reg) = \frac{\sum_{cell \in reg} match(cell)}{m \times n} \tag{3.13}$$

矩形領域と頭部形状モデルとが完全に照合する場合，照合度は1.0になる．顔を含む矩形領域を抽出するために，我々は，頭部形状モデル系列を用いて，入力画像上を走査しながら，画像中の矩形領域と頭部形状モデル系列の中の各モデルとの照合度を評価する．画像の同じ位置において，異なるサイズや向きのモデルを用いて評価された照合度の最大値を選びだし，それをこの位置における頭部らしさとして，矩形領域とモデルとの照合度マップ(Map of Matching Degreeあるいは MMD)を作る．その中には，入力画像における各位置の最も顔らしい矩形領域と対応する頭部らしさ，サイズおよび向きの情報が保存されている．

(a) Original image. (b) Extracted SCSM (c) Extracted HCSM



(d) The MMD obtained by comparing the extracted SCSM & HCSM.

図 3.9: 肌色類似度マップと髪色類似度マップより評価された照合度マップ.

図 3.9(a) は色彩入力画像である. 図 3.9(b) と図 3.9(c) は濃淡画像で, それぞれ入力画像より評価された肌色/髪色類似度マップを表す. 図 3.9(d) は顔候補を抽出する時に評価された照合度 (頭部らしさ) マップ (MMD) で, 見やすくするために, 入力画像に相対して 45 度回転させた. 座標軸の W 軸と H 軸はそれぞれ入力画像の幅と高さで, MD 軸は照合度である (山が高いほど照合度が高い). 図 3.9(d) に示すように, 入力画像中の顔のある所に高い山が見られる.

照合度マップの局所最大値がしきい値より高い場合, その局所最大値と対応する位置における矩形の中に顔が存在する可能性が高く, その矩形領域を顔候補とする.

3.4 顔候補の初期検証

本研究では, 顔候補を検出するとき, 肌・髪の色情報のみ利用している. 背景が複雑である場合, 肌色らしい領域の上に髪色らしい領域が存在することがある. こ

の時，顔候補として検出されてしまったものの中に顔でないものが存在する可能性もある．一般的に，顔領域の中に目，鼻，口などの顔の器官が存在するかどうかを調べることにより，本当の顔か偽の顔候補かを検証できる．しかし，現段階では検出された顔候補の大きさと姿勢が様々なので，顔器官の抽出による精密な検証は困難であると思われる．その代わりに，顔領域内のエッジ分布状況を分かれば，概略な検証が可能と考えられる．

顔器官には水平方向に長いものが多いので，まず，色々な向き，大きさの本当の顔を対象として，顔領域内に垂直エッジを抽出し，垂直方向に加算投影をする．その累積分布を調べることによって，次のような結果を明らかにした：メガネを掛けているか否かと関係なく，眉，目，鼻，口の所，場合によって前髪，あごの所に分布の山があり，分布の平均値の大きさが一定範囲内である．

顔ではない領域に対しても調べた．その結果は，分布の平均値が極端に大きくなったり，小さくなったりとか，上，下半分累積分布のバランスが本当の顔の場合と合わないとかの傾向が顕著している．

図 3.10: 顔候補の検証.

以上の調査結果を基づいて，本研究では，肌色と髪色の相対関係を改めてチェッ

クしたのち，顔候補領域内に求められた垂直エッジを垂直方向へ加算投影し，その累積分布状況という大雑把な情報より顔候補の初期検証を行う．検証結果の二つの例は図3.10に示す．同図には，赤色矩形は検出された顔候補の領域，青色矩形は真顔と判断された領域を示す．又，白色の点線は垂直方向のエッジ加算投影分布，黄色の直線は全顔候補領域の分布平均値，緑色の2本直線は領域の上半分，下半分の分布平均値を表す．

3.5 顔の検出実験

提案した方法の有効性を検証するために，ビデオカメラで撮影した研究室内や，大学構内のシーンの画像と，テレビ放送から録画した画像を対象とし，顔検出の実験を行った．肌の色分布モデルおよび髪の色分布モデルを構築するために，男女38人分のサンプル顔画像を使った．正面，右斜め向き，右向き，左斜め向き，左向きというような基本頭部形状モデルを構築するのに，7人(男性5人と女性2人)の顔画像系列を使った．

3.5.1 顔候補の検出実験

まず，合計47枚(その中に女性34人と男性33人が含まれる)の色彩静止画像に対して顔検出の実験を行った．実験の結果は図3.11にまとめてある．入力画像の明るさが極端に明るい場合，色が"白"になり，また，極端に低い場合，色成分が不安定になり，肌色が正確に抽出できないので，顔が抽出できない時がある．また，小さ過ぎる顔(20×24未満)はノイズの影響を受けやすく，顔が正確に抽出できないことがある．入力画像内のサイズが20×24以上の顔の検出率は，95.5%であった．一般的に，エッジ検出などの手法で顔の特徴を抽出して顔を検出する方法は，画像中の顔のサイズが70×70画素以下の場合，安定な処理結果を期待できないといわれている[Song 94]．従って，本手法は，画像中の顔のサイズにおける変化に強いという利点がある．

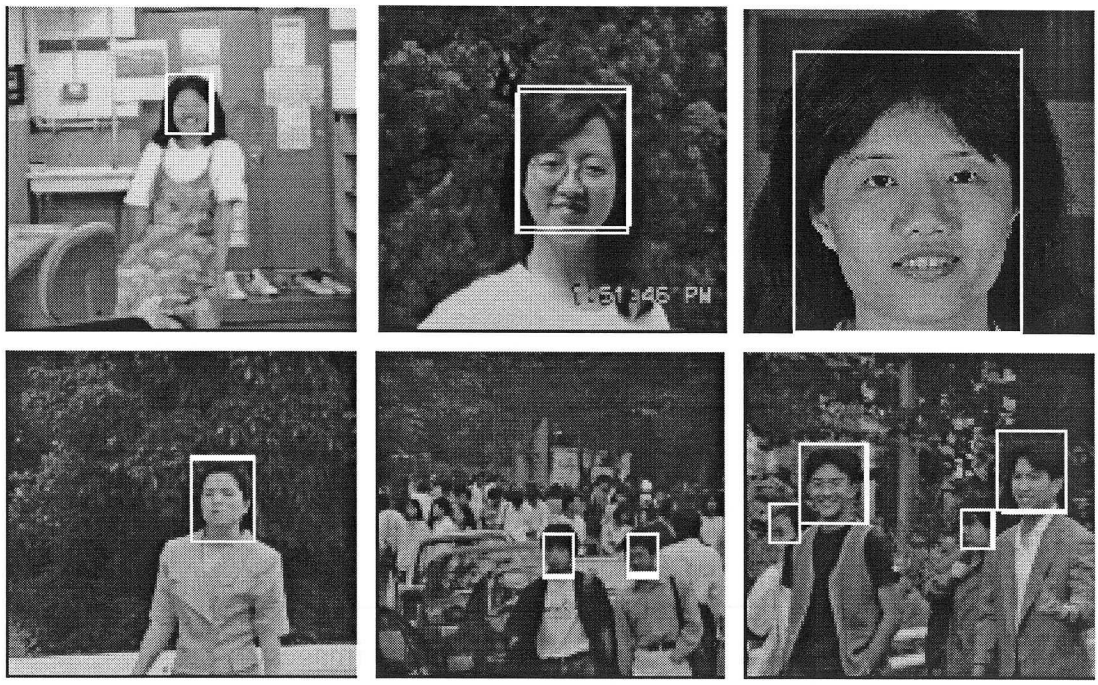
現在，本アルゴリズムに基づいて，ImpactビデオとIndigo^2の上に1フレームにつき2秒の速さで実際に動作する顔検出システムを構築した．

実験結果の例は図3.12に示す．図3.12(a)には，静止画像から検出した顔を示し，図3.12(b)には，実際に動作する顔検出システムの出力結果を示す．図3.12(b)の上には抽出された肌領域と髪領域を示し，下には検出された顔を示す．

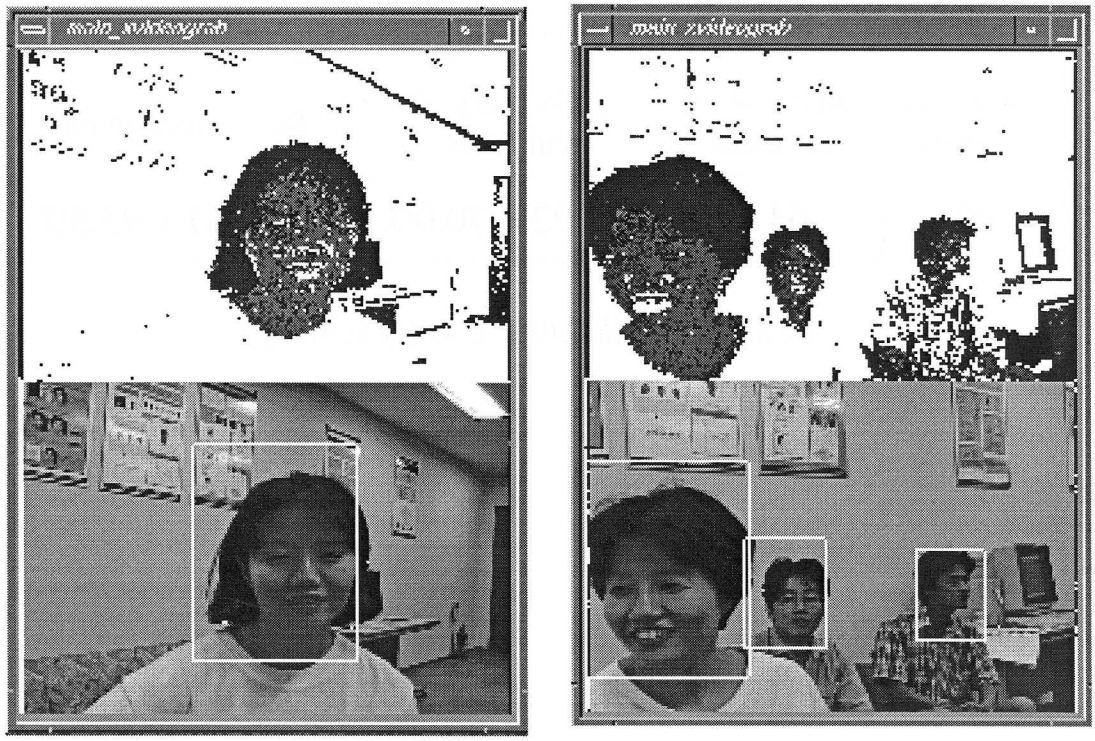
total images	total faces in the images
47	67

candidate of faces	candidate of real faces	probability of real faces	face detection rate
92	64	64 / 92 = 70.0%	64 / 67 = 95.5%

図 3.11: 実験結果のまとめ (静止画像).



(a) Extracted faces candidate in static images.



(b) Extracted faces candidate in live video sequence. The upper part shows the extracted region of the skin color and the hair color. The lower part shows the face detection result.

図 3.12: 顔候補の抽出の実験結果.

3.5.2 顔候補の検証実験

検出された顔候補に対して，初期検証の実験を行った．その一部の結果を図3.13に表す．図に示すように，大きさ，向きを変わっても，眼鏡を掛けても，本当の顔であればその顔候補は真顔であると判断し，その矩形領域は青色で再表示する．一方，ある顔候補が偽顔であると判断されると，そのものは候補から削除し，青色で再表示しない(図3.10参照)．この概略な検証を行った後，偽の顔候補がほとんど除外できるが，真顔として残される可能性がまだある．

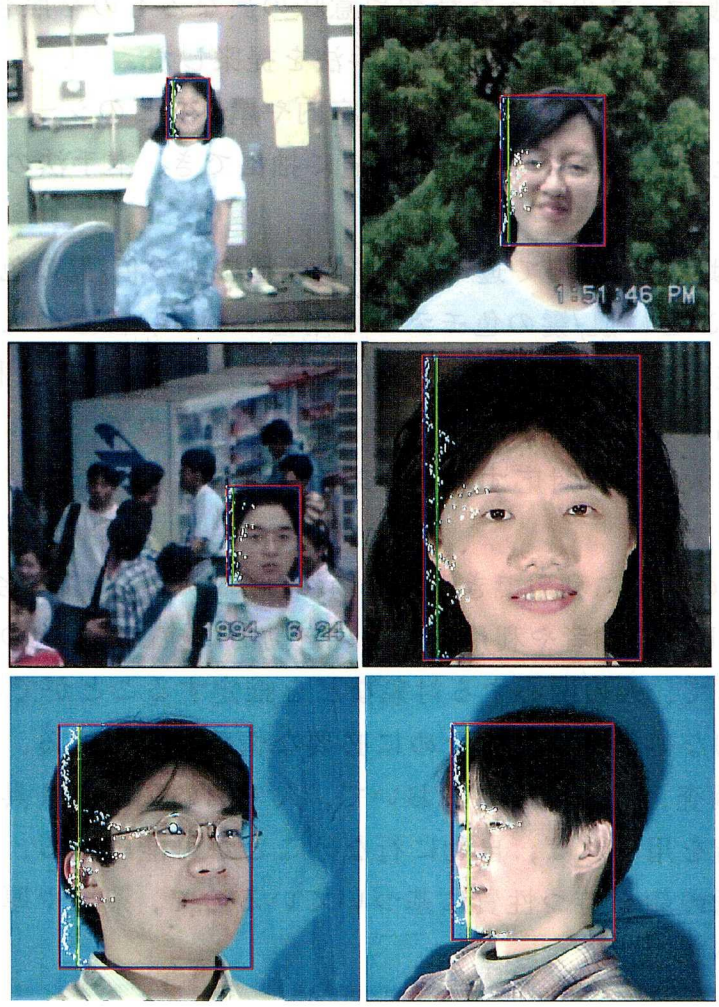


図 3.13: 顔候補の検証結果.

3.6 まとめ

本章は，色彩自然画像から個数，位置，大きさ，姿勢が未知である顔を検出する新しい手法を提案した．肌色/髪色を検出する際に，F氏均等知覚色空間を利用して，肌/髪の色分布モデルを構築し，それによって，肌色/髪色を“らしさ”で表現するという新しい手法を提案した．これより，肌色領域/髪色領域を安定に抽出することができた．更に，顔の検出を安定に行うために，顔パターンはその形状特徴に基づくモデルを構築して表現し，そして，顔形状モデルと入力画像から抽出した肌色/髪色類似度マップのような種類や意味の異なるものを照合させるために，ファジイ理論に基づいたパターン照合の手法を提案した．

多数の実験結果より，複雑な背景を持つ室内や屋外などのシーンでも，また画像中に顔が二つ以上ある場合でも，顔が安定に抽出でき，本手法の有効性が実証できた．

肌の色/髪の色は人種によって違うが，同じ人種の肌の色/髪の色はほぼ同じである．従って，人種ごとに肌/髪の色モデルを用意すれば，任意の人種，あるいは複数人種の人が混在する画像からも顔が検出できると思われる．人種別の肌/髪の色モデルで顔を検出する処理は独立に行うことが可能なので，それらの処理を並列に行うと，一つの人種のみ対応する本システムと同じ速さで顔が検出できると考えられる．

顔の検出は多くの分野に利用できると思われる．例えば，テレビ会議や銀行キャッシュコーナに設置し，参加者や利用者の顔を自動的に検出し，次の章で述べる顔の姿勢推定と併用して，人の関心する対象などを推定することができるようになり，それに応じて適切な処置を行なうために必要な情報を提供する．またスーパーや百貨店の顧客の調査，美術館などにおいて，人の関心が集まり具合の調査などに顔の自動検出技術を用いれば，自動に行なうことが可能になる．

本研究では普通の顔を持つ性質に基づいているので，茶髪や厚化粧の顔については検出できない．

第4章　色彩画像からの人物の頭部の三次元姿勢推定

顔認識の研究としては，顔の動き検出 [A. Tsukamoto 94][間瀬 91][P. Ballard 95][S. Basu 96]，顔の表情認識 [K. Matsuno 95]，個人識別などが挙げられるが，これらの研究において，カメラに対する顔の三次元的な姿勢を求めることは特に重要な課題である．

顔の動き検出に関して考えてみると，これまでの研究と同じくおおまかな顔の姿勢を検出することで計算機にジェスチャ認識させ，これにより計算機に意志表示をすることが可能である．さらに詳細な顔の姿勢推定ができれば，計算機にユーザがどこを注視しているのかを判定させ，ユーザが何に興味を持っているのかを認識させることが可能となり，人と計算機とがよりインタラクティブなものになると期待できる．

一方，計算機による表情認識や個人識別に関して，入力顔画像の向きに対してより柔軟性の高いものを考えたとき，顔の姿勢推定は重要なものであると言うことができる．これまでの研究では，表情モデルや個人モデルが正面顔画像に限られていた場合，顔が横を向いた入力画像に対して判定できなかった．また，回転に対して複数のモデルを持っている場合でも，やはり実際の入力の顔の向きとモデルの顔の向きの間にずれが存在し，精度が良くないなどの問題点があった．しかし，入力画像中の顔の姿勢を求めることができれば，あらかじめ保持しておいた三次元的な顔モデルを，推定された姿勢をもとに回転させることにより，入力画像とモデルの精度の高い比較が可能となる．

また他方では，正面を向いた顔は利用できる情報が一番多いので，モデルを回転させるのではなく，正面顔画像を常に獲得すれば良いという考え方がある．そのため，能動的カメラシステムや複数台のカメラを用いた協調システムなどを用いて，正面顔をカメラで追跡することが考えられているが，この場合でも顔の姿勢推定は重要な役割を果たすと考えられる．これらのカメラシステムにおいては，顔を追跡するカメラの制御をどうするかが重要な鍵となる．顔の姿勢を推定する

ことは，能動的カメラをどこに移動させれば良いか，または，複数台あるカメラのうち次にどのカメラを選択するのが最適かといったシステムの制御にとってとても重要な情報となる．

本研究では，自然な状態でいるユーザに対しても応用できる顔認識システムの開発を目指し，ロバストかつ高速な頭部姿勢の推定方法を開発する．

間瀬らは，顔と髪のコントラストに基づいて，抽出された頭部と顔の領域，ならびにその重心の情報をもとにした頭部の三次元動作の検出法を提案した[間瀬 91]．この方法では，重心のずれから回転角度への変換を，頭部の三次元モデルを使って推定した．しかし、この手法を適用する際に，撮影条件が(照明，レンズ焦点距離，カメラ位置，方向)一定であり，背景画像が既知であるなどの制約条件は必要であった．また，頭部の傾きの角度を推定する際に，首の中点の位置を利用しているので，その位置を精度良く求めることも必要であった．従って，この方法は画質や画像内の顔の大きさなどの変動による影響を受けやすいと思われる．さらに，この方法には，各回転軸に関する情報が独立に求められないという短所もあった．

これに対して，本研究では，マシンの前に操作をしているユーザの顔を検出したのち，その顔領域内で，抽出された肌色領域と髪色領域の重心という一次モーメントを利用する上に，さらに頭部領域(肌色領域と髪色領域を併せた領域)の慣性主軸など二次モーメントを用いて，カメラに対する頭部の姿勢(単軸回転および3軸複合回転)を推定する方法を提案する．

本手法は，目，鼻，口などの顔の特徴[P. Ballard 95]や首などの位置[間瀬 91]の情報の代わりに，肌色・髪色領域の一次，二次モーメントといった安定に求められる情報のみを用いて姿勢を推定することにより，画質や画像内の顔の大きさなどの変動による影響を受けにくくなり，眼鏡や髭がある場合にも効果的かつ柔軟的に対応できる．

頭部の姿勢推定の手法をまず4.1では，単軸回転の場合について述べ，次に4.2では，3軸複合回転の場合について述べる．4.3に実際に人物の頭部の姿勢推定を行った結果と能動的カメラシステムのシミュレーションの結果を示す．最後にまとめを行う．

4.1 単軸回転における姿勢推定

カメラに対する頭部の三次元姿勢は，頭部座標系のX，Y，Z軸の3軸の回転により表現できる（図4.1参照）．X軸は頭部を上下に振る方向を，Y軸は頭部を左

右に振る方向を，そしてZ軸は頭部を左右に傾ける方向に当たる．

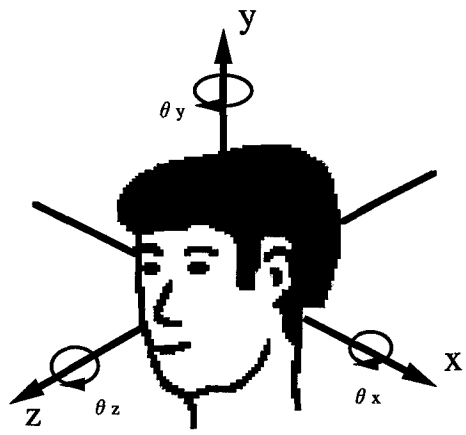


図 4.1: 顔の姿勢の回転要素

以下，姿勢推定の処理は検出された顔領域の範囲内で行う．頭部の姿勢を推定するのに，肌色領域と髪色領域の情報を用いる．まず各軸が単独回転をした場合の姿勢推定法について述べる．

4.1.1 Z軸回りの回転角度の推定

頭部の形状を考えてみると，頭は中心軸に関してほぼ左右対称であり，またほぼ楕円体として見なすことができる．このことから，頭がZ軸回りに回転した場合，画像面に投影された頭部を楕円と見なしたときの長軸の方向は，頭の傾きを表している．この長軸の方向を調べれば，Z軸回りの回転角度を推定できることが予想される．この長軸方向の傾きを，頭部領域の慣性主軸を求めることにより計算する [Klaus 86][谷内田 86].

実際に頭部がZ軸回りに回転をした入力顔画像から，慣性主軸を計算することによって推定されたθ_zの結果を図4.2と表4.1に示す．図4.2において，頭部領域上に示される直線が求められた慣性主軸を表す．この図から，求められた慣性主軸がZ軸回りの回転結果とほぼ一致していることが分かる．表4.1からも推定誤差が許容範囲内であると言える．

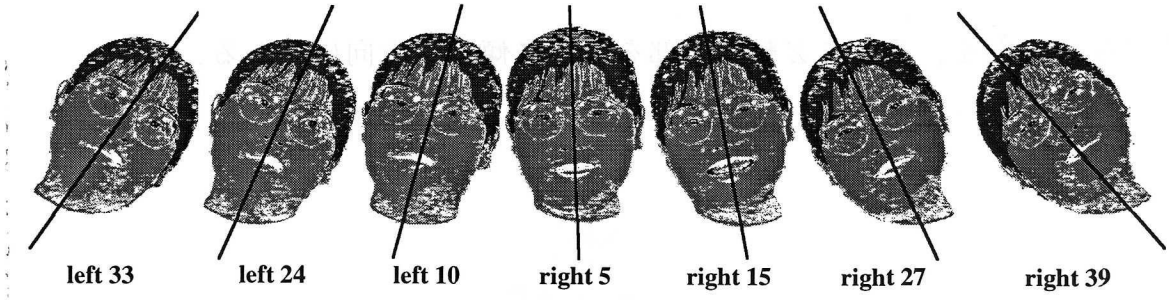


図 4.2: 推定された慣性主軸の例

実際の角度 [degree]	-33	-24	-10	5	15	27	39
推定された角度 [degree]	-35	-24	-15	-3	11	25	41
誤差 [degree]	-2	0	-5	2	-4	-2	2

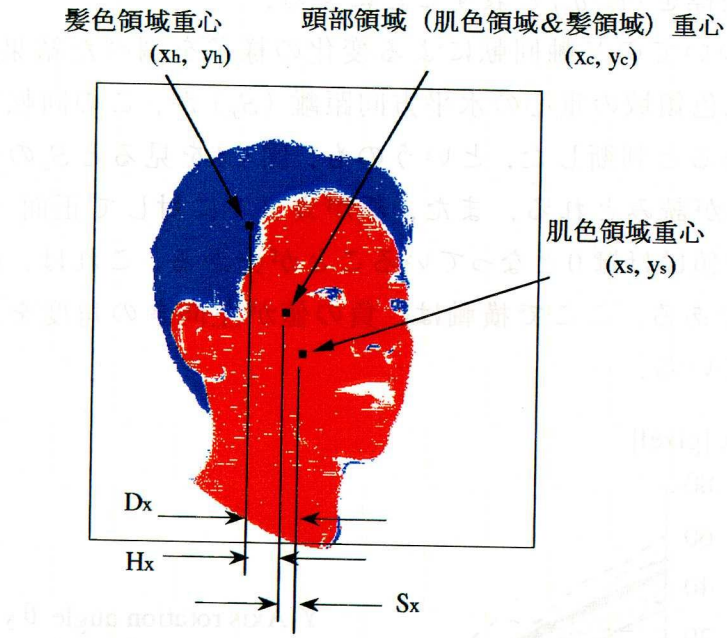
表 4.1: 推定された慣性主軸の結果

4.1.2 Y 軸回りの回転角度の推定

顔がカメラに対して正面を向いているとき，得られた顔画像中では，肌色領域が頭部の中心部にあり，髪色領域が頭部の周辺部分にあるであろう．しかも，それらの領域は左右対称に見えるであろう．顔が左(あるいは右)を向いているときは，画像中では肌色領域は頭部の右(あるいは左)に現れ，髪色領域は左(あるいは右)に現れるだろう．このことから，肌色領域，髪色領域，頭部領域の間の関係を調べれば，Y 軸回りの回転角度を推定できることが予想される．

そこで，男女複数の被験者（男性6名＋女性1名）に対して，正面向きを基準として左向き90度から右向き90度まで，Y 軸回りに頭部を15度ずつ回転させて，被験者一人あたり13枚（総計91枚）の入力画像を取り込んだ．そして，抽出された肌色，髪色，頭部の三つの領域に関してそれぞれ面積と重心を求めた．これより図4.3に示すように，頭部領域の重心に対する肌色領域の重心の水平方向距離（S_x），頭部領域の重心に対する髪色領域の重心の水平方向距離（H_x），髪色領域の重心に対する肌色領域の重心の水平方向距離（D_x）を考える．

ただし，得られた画像中の顔の大きさは一定ではないので，各水平方向距離は



Sx ：頭部領域重心に対する肌色領域重心の水平方向距離
Hx ：頭部領域重心に対する髪色領域重心の水平方向距離
Dx ：髪色領域重心に対する肌色領域重心の水平方向距離

図 4.3: 各重心間の水平方向距離

顔の面積の平方根に比例すると仮定して次式を用いて正規化する．

$$S_x = (x_s - x_c)\sqrt{S_0/S} \tag{4.1}$$

$$H_x = (x_h - x_c)\sqrt{S_0/S} \tag{4.2}$$

$$D_x = (x_s - x_h)\sqrt{S_0/S} \tag{4.3}$$

ここで，S_0は人間の頭部領域の基準面積，Sは実際の画像から計算された頭部領域の面積を表す．また，肌色領域の重心座標を(x_s, y_s)，髪色領域の重心座標を(x_h, y_h)，頭部領域の重心座標を(x_c, y_c)と表すことにする．

S_x，H_x，D_xについてのY軸回転による変化の様子を調べた結果より，頭部領域の重心に対する肌色領域の重心の水平方向距離（S_x）が，この回転要素を推定するのに最も有用であると判断した．というのも，図4.4を見るとS_xの変化がほぼ直線的となっているのが読みとれる．また，顔がカメラに対して正面（$\theta_y = 0$）を向いているとき，S_xの値はほぼ0となっていることが分かる．これは，顔の対称性から予想された通りである．ここで横軸は，負の値が左向きの角度を，正の値が右向きの角度を表している．

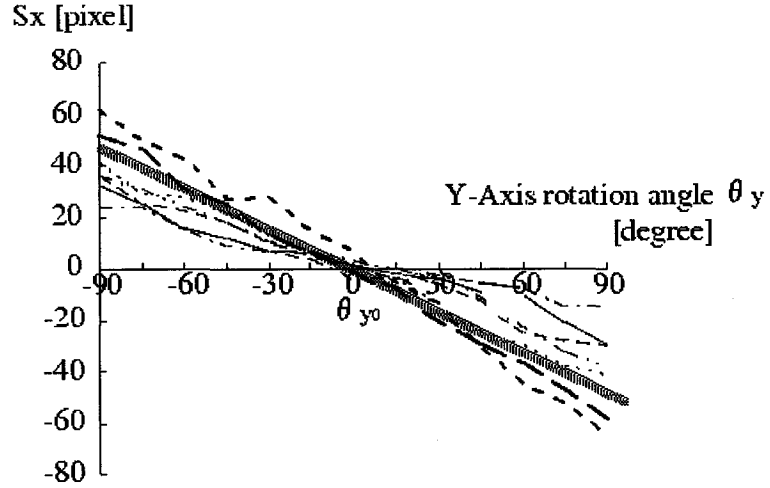


図 4.4: Y軸回転におけるθ_yの推定式

図4.4より，S_xからY軸回りの顔の回転角度θ_yを推定する一次近似式を立てると，次式のようになる．

$$\theta_y = \theta_{y0} + a_y S_x \tag{4.4}$$

ここで，θ_{y0}は$S_x = 0$のときのY軸回転の値を，a_yは近似直線の勾配を表している．

4.1.3 X軸回りの回転角度の推定

X軸回りの回転角度θ_xの推定方法は，4.1.2で記述した方法と類似している．ただし，θ_yの推定では，重心間の水平方向距離S_xを利用したが，θ_xの推定には，頭部領域に対する肌色領域の垂直方向距離S_yを用いる(図4.5).

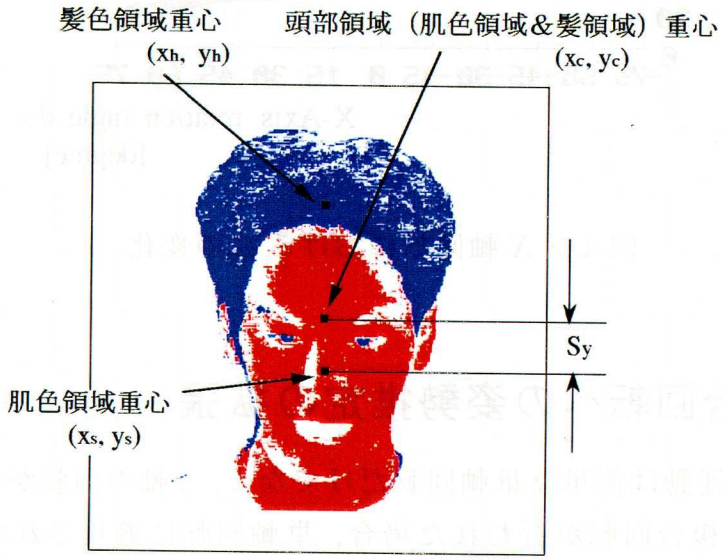


図 4.5: 重心間の垂直方向距離

実際に，5人の被験者に対して下向き60度から上向き60度まで，頭部を回転させたときの画像を取り込み，S_yとX軸回りの回転角度の関係を図4.6に示す.

図4.6より，$\theta_x = 0$を境とし，$\theta_x < 0$側と$\theta_x > 0$側で別々にS_yの値がほぼθ_xの値に比例していることが分かる．そこで，図4.6から場合分けした一次近似式を立てると，次式のようになる.

$$\theta_x = \begin{cases} \theta_{x_1} + a_{x_1} S_y & (S_y \leq S_{y0}) \\ \theta_{x_2} + a_{x_2} S_y & (S_y > S_{y0}) \end{cases} \tag{4.5}$$

ここで，$\theta_{x_1}, \theta_{x_2}$はそれぞれ$S_y = 0$のときの$X$軸回転の値を，$a_{x_1}, a_{x_2}$はそれぞれの近似直線の勾配を表している.

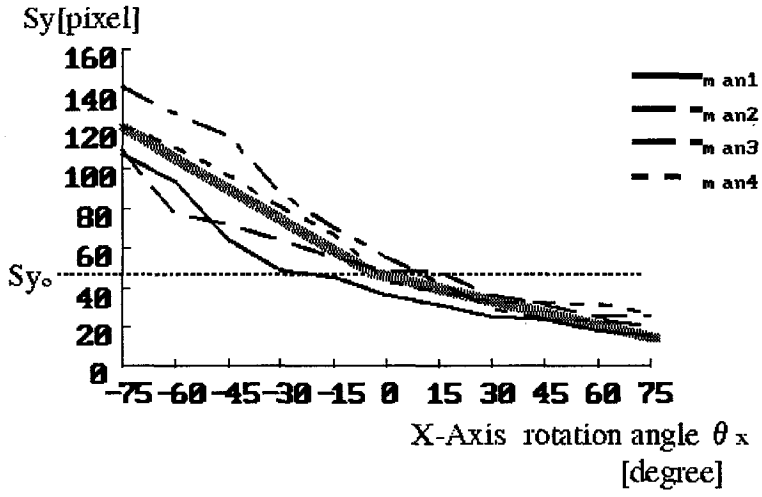


図 4.6: X 軸回転における S_y の変化

4.2 3軸複合回転への姿勢推定の拡張

実際に，頭部の運動は簡単な単軸回転だけでなく，3軸の回転が組合わさった複雑なものである．複合回転が行われた場合，単軸回転に適用された推定方法に他の回転の影響が入り込んで，そのままでは使用できない．そこで，単軸回転の推定から3軸複合回転における各軸の回転の推定へ拡張する．

4.2.1 Z 軸回転の推定に対する3軸複合の影響

まず，Y 軸回転による慣性主軸への影響を考える．ある Z 軸回転において，画像面で観測される慣性主軸の傾きは，カメラに対して顔が正面を向いているときに値が最大となり（これが真値である），Y 軸が左右に回転していくにつれて，傾きは $atan$ 関数に従って減少し，左90度あるいは右90度において値が理論上0となる．その様子を図4.7に示す．また，Y 軸回転の値が大きい場合には，観測される慣性主軸の傾きは Y 軸回転の影響だけでなく，X 軸回転の影響によっても左右されることになる．特に左(あるいは右)90度の場合は，X 軸の回転角度がそのまま慣性主軸の傾きとなる．

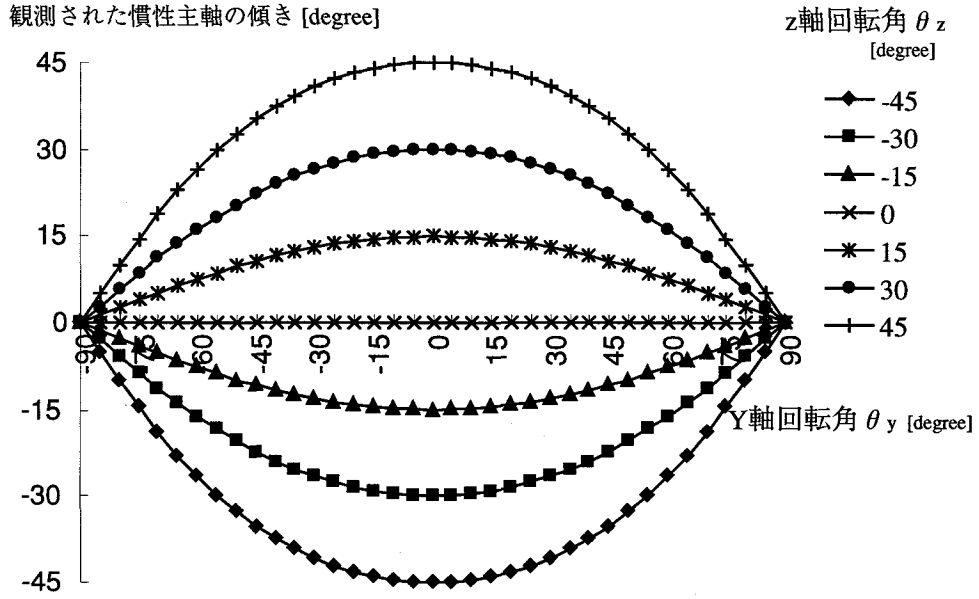


図 4.7: Y 軸回転による観測される慣性主軸の変化

Z 軸回転から X,Y 軸回転の影響を分離するのは，単眼のみの観測では難しい．ただし，Y 軸の回転角度が小さい場合，Y 軸回転と X 軸回転よりの影響は小さいので，無視できる．

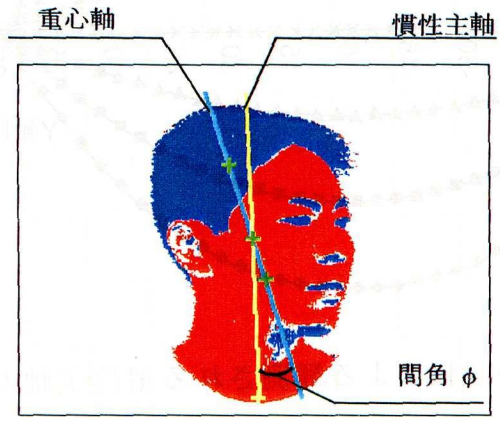
4.2.2 Y 軸回転の推定に対する3軸複合の影響

Z 軸回転による Y 軸回転角度の推定への影響であるが，顔が左右に傾けば，画像の水平方向を基準としている頭部領域と肌色領域の重心間水平距離（S_x）も当然変化するので，推定されたθ_yは真値からはなれてしまう．しかし少し考えれば，各重心が頭部領域の重心を中心とした回転移動を行ったと見なせることが分かる．そこで，頭部領域の重心を原点とし，顔の中心軸である慣性主軸を縦軸とするような座標空間（x'-y'空間)を考え，その空間上に各重心を座標変換し，そのときの重心間の水平方向距離を求めれば，Z 軸回転の影響を取り除くことができる(図4.8)．つまり，次式により S_xを求め，式4.4に代入すれば良い．

$$S_x = r\sin\phi \tag{4.6}$$

ここで，rは頭部領域重心と肌色領域重心を結ぶ線分を，ϕは慣性主軸と重心軸(各重心を通る直線)の間角を示す．

図4.8を見ると，(b)の左側の図においてて$\theta_y = 0$で Z 軸回転を行っても，間角ϕはほぼ0度であることが分かる．それに対して，右側の図は，θ_yを変化させた場合



(a) 間角 ϕ

(b) 回転による間角 ϕ の変化

図 4.8: 慣性主軸と重心の間角ϕ

であるが，間角ϕが大きくなり慣性主軸に対する水平方向距離が大きくなっているのが分かる．

4.2.3 X軸回転の推定に対する3軸複合の影響

まず，Z軸回転による影響であるが，これもY軸回転の推定に対する影響を考える場合と同じく，各重心が回転運動を行っただけであるので，慣性主軸を縦軸と取る座標空間に変換し，その空間での頭部領域–肌色領域の重心間垂直方向距離を求めれば良い．S_yは次式で求められる．

$$S_y = r\cos\phi \tag{4.7}$$

さらに，Y軸回転による影響であるが，推定式4.5はX軸の単独回転($\theta_y = 0$の場合)においてのみ，θ_xの値を推定する式である．$\theta_y \neq 0$の場合は，そのままのS_yの値からθ_xを決定できない(図4.9)．つまり，あるθ_yにおけるS_yの値を，$\theta_y = 0$におけるS_yの値に補正する必要がある．そこで，各θ_yにおけるX軸回転によるS_yの変化を調べてみた．その結果を図4.10に示す．

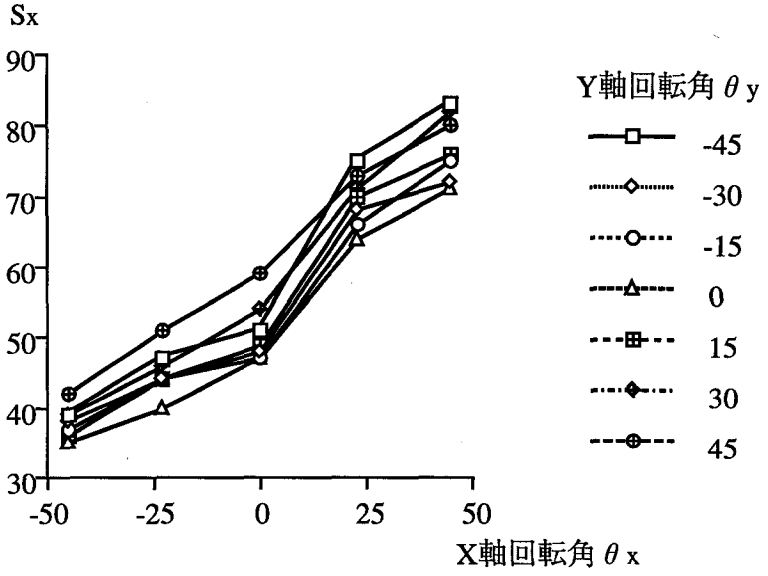


図 4.9: 各θ_yのX軸回転によるS_xの変化

図4.10を見ると，各S_yの値が$\theta_x = 0$において最小となるような，緩やかな二次曲線を描いているのが分かる．また，各θ_yにおける二次曲線の変化率もほぼ同じであると言える．これより，観測されたS_yの値とそのときのθ_yの値から，$\theta_x = 0$におけ

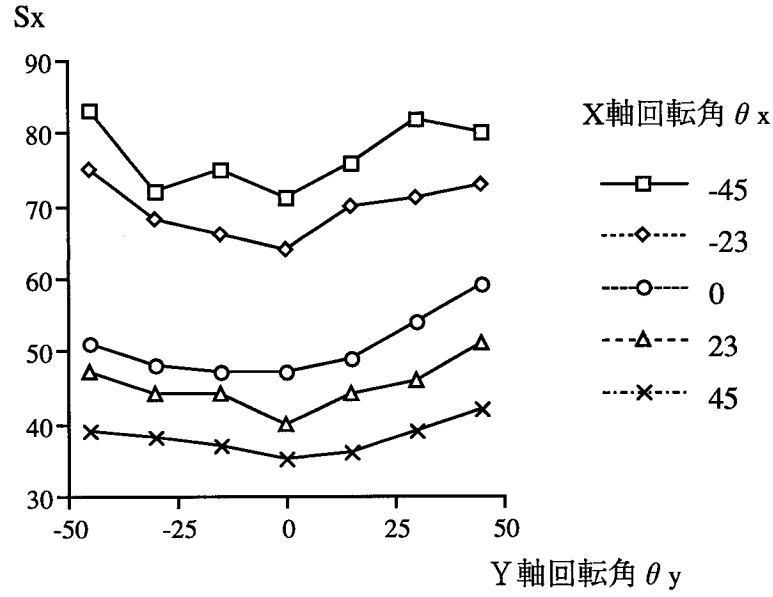


図 4.10: 各θ_xのY軸回転によるS_xの変化

るS_yの値S_{y_0}を次式から推定することができる.

$$S_y = b\theta_y^2 + S_{y_0} \tag{4.8}$$

ただし，bは二次曲線の二次係数を表す. これより，S_yを求め推定式4.5よりθ_xの値を推定することができる.

4.3 実験及び考察

上述したX, Y, Zの3軸による頭部の姿勢推定の手法を用いて実際に実験を行った. 実験は2種類行った. 一つは，カメラの前で被験者に自由に頭を動かしてもらい，頭部の三次元姿勢を連続的に推定していく実験で，推定手法の精度を調べた. もう一つは，能動的カメラシステムによる正面顔追跡のシミュレーションの実験で，顔追跡のシステムへの本手法の有効性を確かめた.

4.3.1 頭部の三次元姿勢推定実験

実験では，被験者に最初Z, Y, X軸の順番に単軸回転を行わせている. 図4.11中で網線の掛かった部分A, B, Cがそれぞれ，Z軸,Y軸,X軸の単軸回転を行っている

フレームである．この部分を見ると動かしていない他の軸について推定された回転角に多少の誤差も見られるが，単軸回転をうまく捉えていることが分かる．

図4.11後半のフレームでは被験者に自由に顔を動かしてもらった．図4.12に入力画像と推定された姿勢をもとに，別に用意した顔の三次元モデルにより合成された結果を示す．合成された顔モデルの結果を見たところ，おおまかには被験者の姿勢をうまく推定できている．細かい部分をみるとやはり関係ない軸が回転していたりと誤差もある程度出ているが，単眼のみの情報により推定したことを考えれば，全体的に見て良い結果が得られたと思われる．

なお，本実験では，入力装置としては，SONYのDigital Handycam DCR-VX1000を使用し，姿勢推定の計算処理はSilicon Graphics社のIndigo^{2}とImpactビデオを用いた．実時間処理における処理スピードは，入力画像のサイズが320×240pixelにおいて，約3frame/secであった．

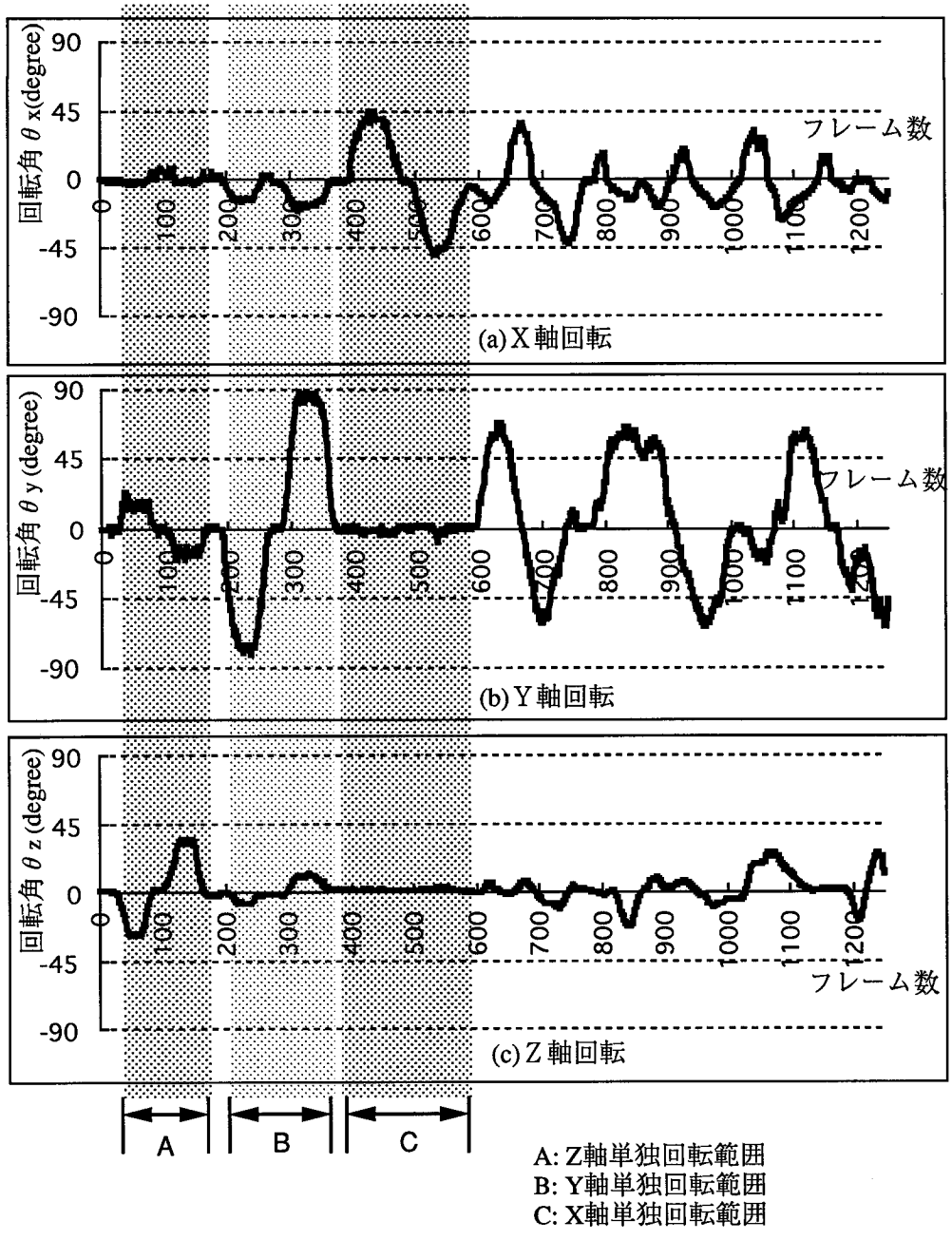


図 4.11: 姿勢推定の結果

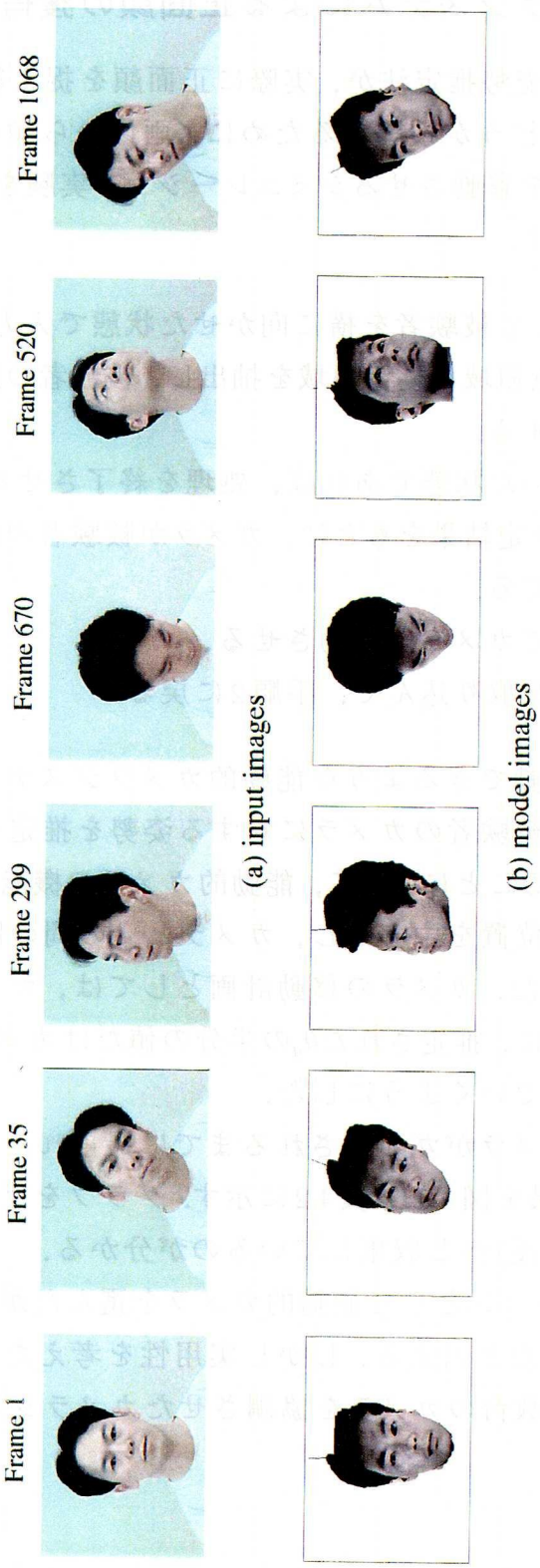


図 4.12: 入力画像系列と合成された顔モデル結果

4.3.2 能動的カメラシステムによる正面顔の獲得シミュレーション

本章で述べた頭部の姿勢推定法が，実際に正面顔を捉え続けるような能動的カメラシステムに有用かどうか検証するために，画像から顔の姿勢（向き）を推定し，顔の正面にカメラを移動させるシミュレーション実験を行った．

実験手順を以下に示す．

1. まず，カメラに対して被験者を横に向かせた状態で入力画像を取り込む．
2. 入力画像から，肌色領域と髪色領域を抽出し，被験者の頭部のカメラに対する三次元姿勢を推定する．
3. 顔が正面を向いていた状態であれば，処理を終了させる．
4. そうでなければ，推定結果をもとに，カメラが被験者の顔の正面に回り込むような移動計画を立てる．
5. 移動計画に基づいてカメラを移動させる．
6. 新しい視点で画像を取り込んで，手順2に戻る．

現段階では実際に移動できるような能動的カメラシステムはないので，カメラから得られた画像から被験者のカメラに対する姿勢を推定し，その結果をもとに人がカメラを移動させることによって，能動的カメラの機能の肩代りをした．またカメラの軌道は，頭の位置を中心とし，カメラと人の間の距離を半径とするような仮想円の円周上とした．カメラの移動計画としては，カメラが顔の正面の位置に安定に収束するように，推定されたθ_yの半分の値だけカメラを回転させて，徐々に顔の正面に回り込んでいくようにした．

正面の顔に向けてカメラがガイドされるまでに撮られた画像系列一例を図4.13に示す．また，その結果を図4.14，表4.2に示す．グラフを見ると，カメラの位置が数回の処理で顔正面(0度)へと収束しているのが分かる．

今回は，シミュレーションとして能動的カメラを選んだが，結果はうまく正面の顔を捉えることができたと言える．しかし実用性を考えたとき，正面顔を追跡するシステムとして，複数台のカメラを協調させたカメラシステムの方がより有用であろう．

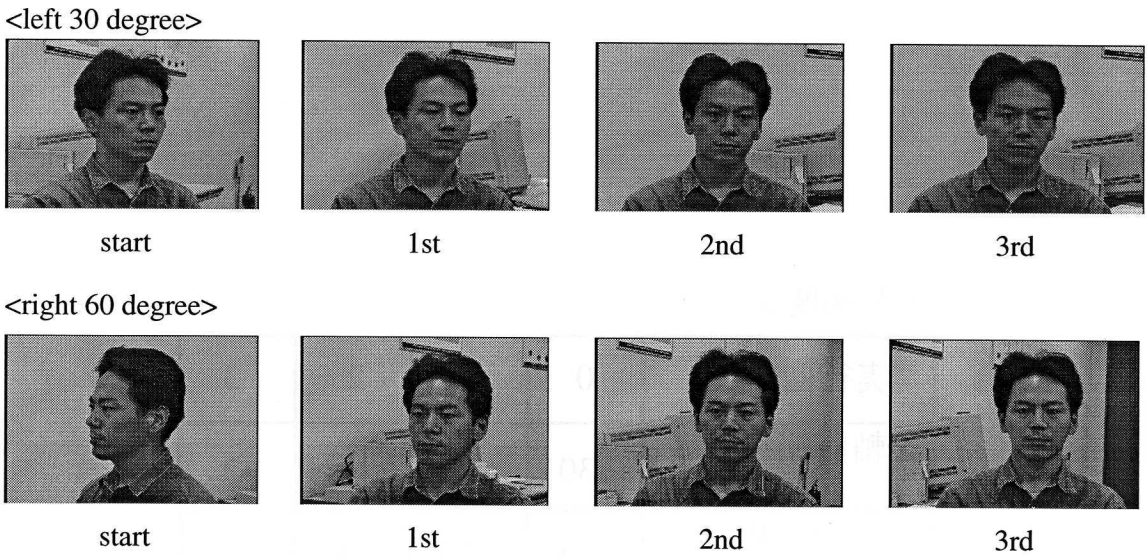


図 4.13: ガイドされたカメラからの画像系列

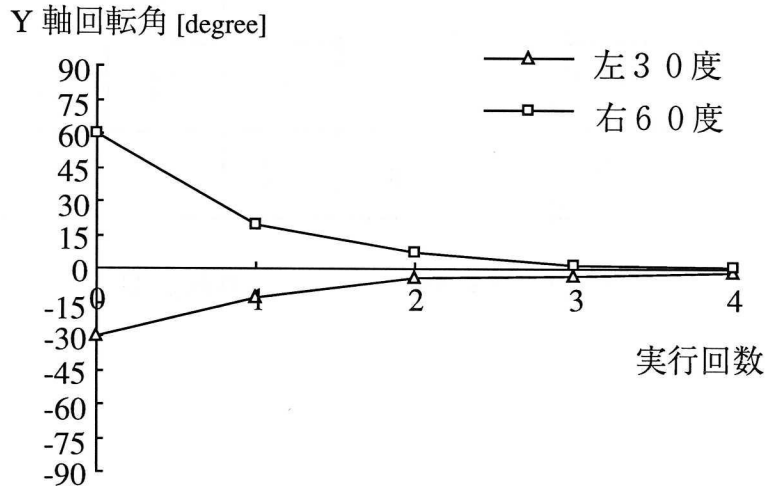


図 4.14: カメラの移動

＜左30度＞

実行回数	0	1	2	3
Y軸回転角度 [degree]	-30	-13	-4	-3
推定結果 [degree]	-33	-18	-1	-3
カメラの移動角度 [degree]	-17	-9	-1	-2

＜右60度＞

実行回数	0	1	2	3
Y軸回転角度 [degree]	60	20	8	2
推定結果 [degree]	79	24	11	6
カメラの移動角度 [degree]	40	12	6	3

表 4.2: 推定結果とカメラの移動量

4.4 まとめ

本章において，色彩画像から頭部の三次元姿勢をロバストかつ高速に推定する方法を提案した．提案された肌色，髪色領域の重心間距離や頭部領域の慣性主軸などの情報を利用した姿勢推定手法が，単眼からの情報だけで，頭部の三次元姿勢を求められることを実験により示した．

本手法はあらかじめ用意された三次元頭部モデルや，実行中に頭部の三次元形状復元を必要とせず，それらの手法に比べてはるかに容易で，しかも計算コストにおいて効果的な手法であると言える．

現在は1台のカメラによる情報のみで姿勢推定を行っているが，複数台のカメラによる協調観測を行えば，より精度の高い姿勢推定が行えると思われる．又，姿勢推定の結果をもとに能動的カメラや複数台のカメラを用いたシステムを制御すれば，個人識別，表情認識のためのユーザの正面顔を常に捉えることが可能となり，実環境への応用が広がると思われる．

第5章　顔器官による顔候補の検証

第3章に述べた顔検出システムには，色情報のみを利用しているので，検出された顔候補の中に偽の顔候補が存在する可能性がある．同章の最後に，顔候補に対して，エッジを求め，垂直方向への加算投影することによって，初期検証を行ったが，偽の顔候補が存在する可能性がまだ残っている．

入力画像から頭部の三次元姿勢を推定し，その結果をもとに能動的カメラを適切に誘導することによって，望ましい大きさの正面顔画像を獲得できることは第4章で明らかにした．このような顔画像を用いれば，顔候補の中に目，口等の顔器官を抽出することができるようになる．

本章では，顔器官による顔候補の検証について論じる．ここで，顔器官の位置を正確に検出するために，色情報(肌と髪の色)および，エッジ情報(強度と符合)を総合的に考慮した加算投影法を提案する．顔器官を抽出できるか否か，同定された顔器官と人間の顔器官の幾何関係モデルとの一致度を計算することによって，顔候補が本当の顔であるか否かを最終判定する．

以降では，5.1で加算投影法の改良について，5.2で色情報を用いた顔位置と大きさの評価について，5.3で色とエッジ情報を用いた顔器官位置の決定について，5.4で顔器官幾何関係モデルによる顔候補の検証について述べる．5.5で顔器官の決定と顔候補の検証の実験結果を示す．最後に討論とまとめを行う．

5.1　加算投影法の改良

本論文では，顔とその器官の位置を決定するのに加算投影法を用いることにする．異なる照明条件と複雑な背景を持つ顔画像にも安定に適用できるように，エッジ情報のほかに，色情報も重視して利用する．つまり，エッジ強度，エッジ符合，肌色らしさ，髪色らしさという4種類の情報を併用する．さらに，各領域を精密に決定する前に，色情報，エッジ情報，ならびに前段階まで決定されてしまった顔器官の位置情報を全面的に考慮し，次の段階の大まかな探索範囲を予測する．そして，予測された範囲内で水平あるいは垂直方向に有用な情報を加算投影する．なお，位

置を決定するための閾値が各領域内の累積分布の平均値によって動的に計算して決定される．

5.1.1 肌色類似度マップと髪色類似度マップ

第3章に述べたように，肌色分布モデル，髪色分布モデルを用いて，色彩入力画像の各画素の色においての肌の色らしさ，髪の色らしさを評価できる．この結果によって，肌色類似度マップ，髪色類似度マップが求められる．そのマップ内の各画素に0から255までの値が保存される．

5.1.2 エッジ強度マップとエッジ符合マップ

顔器官には水平方向に長いものが多い．それをうまく抽出するために，本研究では，3×3，5×3，7×3，9×3画素の垂直微分用のマスクを定義し，エッジ抽出の比較実験を行った．比べた結果より大きさ512×480以上の入力画像に対して，7×3画素という大きめなマスクを用いると，顔器官に関するエッジ情報が一番よく得られることがわかった．7×3画素のマスクは次のように定義する．

a_{-3}	a_{-2}	a_{-1}	a_0	a_1	a_2	a_3
b_3	b_2	b_1	b_0	b_1	b_2	b_3
c_{-3}	c_{-2}	c_{-1}	c_0	c_1	c_2	c_3

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

図 5.1: 定義された7×3のマスク．

色彩情報を充分に利用すれば，処理の安定性を向上できると考えられるので，エッジの強度を計算するのに，入力画像内の明るさと色の変化とに基づいて，空間的な1次微分を求めることにする．すらわち，R,G,Bの値より次式でエッジの強度$Strength$を求める．この結果を，0から255までの値でエッジ強度マップと名付ける画像に保存する．

$$Strength(p) = \triangle I_{p-3} + 2\triangle I_{p-2} + 2\triangle I_{p-1} + 3\triangle I_p + 2\triangle I_{p+1} + 2\triangle I_{p+2} + \triangle I_{p+3} \quad (5.1)$$

ここで

$$\triangle I_i = \sqrt{(R_{ci} - R_{ai})^2 + (G_{ci} - G_{ai})^2 + (B_{ci} - B_{ai})^2}$$

更に，明るさ変化の方向における情報をうまく利用すれば，前髪などの影響を受けにくくなり，目，眉などの器官に対してより正確な位置付けができると考えられる．従って，本論文ではエッジの強度だけでなく，エッジの符号も求めて利用する．エッジの符号 $Sign$ は次式で明るさ (Y) より求める．この結果を，-1，0，1 という3値データでエッジ符合マップと名付ける画像に保存する．

$$Sign(Yedge,\ a) = \begin{cases} -1 & Yedge \leq -a; \\ 0 & -a < Yedge \leq a; \\ 1 & a < Yedge; \end{cases} \tag{5.2}$$

ここで，a は閾値となる定数である．また，明るさが暗から明に変化する場合は"+"，逆に，明から暗に変化する場合は"−"と定義する．なお，$Yedge$ は次式で計算する．

$$\begin{aligned} Yedge \ &= \ Y_{c1} - Y_{a1} + 2(Y_{c2} - Y_{a2}) + 2(Y_{c3} - Y_{a3}) \\ &\quad + 3(Y_{c4} - Y_{a4}) + 2(Y_{c5} - Y_{a5}) + 2(Y_{c6} - Y_{a6}) + Y_{c7} - Y_{a7} \end{aligned}$$

5.1.3 動的閾値

加算投影法という簡単であり，しかも有効な画像処理技術を用いて，安定に顔器官を同定するために，上に述べた4種類のマップの情報を融合して利用する．更に，できるだけ照明状況や色，画質などの変動に影響されないように二つの工夫をした．一つは，その4種類の情報を加算投影するとき，各画素のデータを2値化せずそのまま累積をする．もう一つは，位置を特定するための各閾値は固定せず，当前処理範囲内の累積分布の平均値によって，動的決定する．

$f(x,y)$ がある画像マップとすると，矩形範囲内 $((x_0,\ y_0),\ (x_m,\ y_n))$ に水平あるいは垂直方向への累積データより平均値を求め，動的閾値を次式で計算する．

$$THX = \frac{1}{a(x_m - x_0)} \sum_{x=x_0}^{x_m} \sum_{y=y_0}^{y_n} f(x,y) \tag{5.3}$$

$$THY = \frac{1}{b(y_n - y_0)} \sum_{y=y_0}^{y_n} \sum_{x=x_0}^{x_m} f(x,y) \tag{5.4}$$

ここで，a と b は定数であるが，マップ別，器官別によって異なる．

5.2 色情報を用いた顔位置と大きさの評価

入力画像から得られた顔候補という矩形領域より，その内の頭部の位置と大きさを精密に求める必要がある．その理由が二つ挙げられる．一つは，顔候補を検出する時，使われた頭部形状モデルがかなり粗いので，検出された顔候補の位置と大きさに誤差が存在する．もう一つは，各顔器官の位置と大きさを予測するのに，その頭部の位置と大きさの情報が必要である．頭部の位置と大きさを示す矩形を図5.2のように定義する．

図 5.2: 定義された頭部の位置と大きさ．

頭部の位置と大きさは顔候補領域内の色情報によって決定する．すなわち，矩形の上部は髪色の上部，矩形の両側は肌色の両側によって決定する．しかし，顔と首の間は色も同じであり，エッジの抽出も一般的に難しいので，矩形の下部の位置を正確に決定するのは困難である．一方，人間の頭部は幅と高さの比率がほぼ一定なので，既に得られた矩形の上部位置と幅の情報に基づいて，矩形の下部位置を近似的に推定できる．

肌色類似度マップより肌領域の左(右)位置を求めるために，顔候補を囲む矩形領域$((x_0, y_0), (w, h))$内で各列上の各画素の肌色らしさを水平(X)方向へ加算投影する．この加算投影の平均値によって，肌領域の両側のX座標を決定する閾値を求めておく．そして，$((x_0, y_0), (w, h))$の中央から左(右)側へ探索していき，X方向の加算投影値が初めて閾値より低い所のX座標$(xl\ or\ xr)$を肌領域の左(右)側位置とする．

髪色類似度マップより頭部の上位置を求める時，既に求められた肌領域の左(右)側X座標$(xl\ or\ xr)$を用いて，$((xl, y_0), (xr, h))$範囲内の各行上の各画素の髪色ら

しさを垂直(Y)方向に加算投影する．この矩形範囲の上から下へ探索することによって，Y方向の加算投影値が初めて閾値より高い所のY座標(yt)を発見し，その位置を頭部領域の上位置とする．頭部領域の下位置の座標(yb)が次のように推定して得られる．

$$yb = yt + (xr - xl) * 1.65 \tag{5.5}$$

これで，頭部の位置と大きさが決定され，白い枠の矩形((xl, yt), (xr, yb))で表示される(図5.3参照).

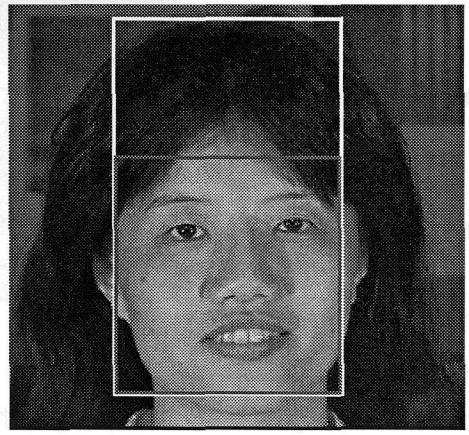


図 5.3: 肌色類似度マップと髪色類似度マップより決定された頭部と顔の位置と大きさ．

更に，この白い枠の矩形内で肌色類似度マップの各行上の肌色らしさを垂直(Y)方向に加算投影して，上から下へ探索すれば，肌色領域の上座標(yt_f)を決定できる．これで，顔領域を含む矩形((xl, yt_f), (xr, yb))が得られグレー枠で表示される(図5.3参照).

5.3 色とエッジ情報を用いた顔器官位置の決定

特定された顔領域(グレー枠)に，5.1.2で述べた方法で，色情報(RGB)よりエッジの強度を求め，エッジ強度マップを獲得する(図5.4(a)参照)．又，輝度情報よりエッジの符号を求め，エッジ符号マップを獲得する(図5.4(b)参照)．これらの情報と色情報(肌と髪の色らしさ)をもとに，顔器官の位置を決定していく．

5.3.1 顔器官の存在しうる領域

個人差や前髪など影響で，得られた顔領域(グレー枠)がかなりばらついている．また，前髪が眉の一部にかかると，顔器官の位置特定の段階ではその影響がしばしばノイズになってくる．従って，顔器官の位置をより安定に特定するために，両眉のすぐ上と両側，ならびに口を含む下という顔器官の存在しうる矩形領域を獲得することが望ましい．

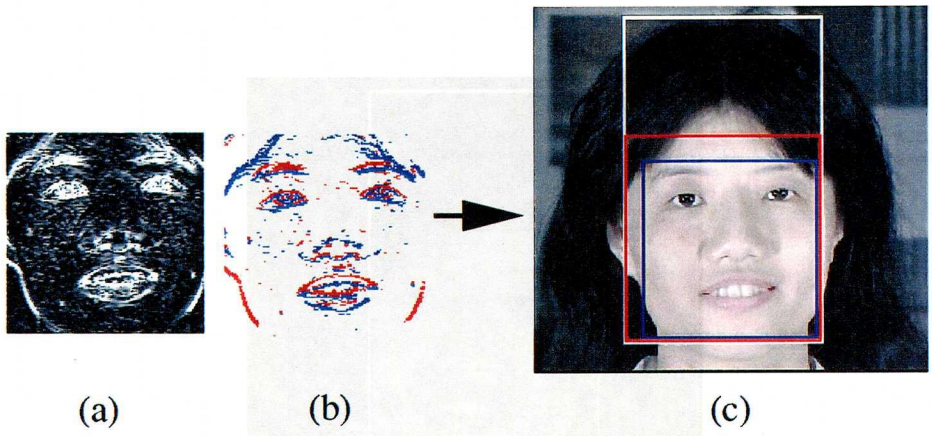


図 5.4: エッジ強度マップとエッジ符号マップより決定された顔器官の存在しうる領域．

このために，顔領域の上半分にエッジ符合マップ (図5.4(b)) の各行上の"-1" の画素数を，Y 方向に加算投影し，この累積分布の頂点 (y_{top}) を見つける．同時に，同じ範囲内にエッジ強度マップ (図5.4(a)) の各行上のエッジの強度も Y 方向に加算投影する．上から下へ探索することによって，頂点 (y_{top}) の近くかつエッジ強度の加算投影値が閾値より高い所を見つけ，その位置の Y 座標 (yt_o) を顔器官の存在しうる上部とする．

そして，y_{top}点の上下 N 画素程度の範囲内で，エッジ符合マップの各列上の"−1" の符合の画素数および，エッジ強度マップの各列上のエッジの強度を X 方向に加算投影し，左(右)から中央へ探索することによって，2種類の加算投影値が同時にそれぞれの閾値より高い所の X 座標 (xl_o or xr_o) を求める．(xl_o or xr_o) を顔器官が存在しうる領域の左(右)の X 座標とする．この領域の下座標 (yb_o) は，次のように推定する．

$$yb_o = yt_o + (xr_o - xl_o) * 1.2 \tag{5.6}$$

これで，顔器官の存在しうる領域$((xl_o, yt_o), (xr_o, yb_o))$を決定でき，青い枠で表示する(図5.4(c)参照).

5.3.2 眉と目

顔器官の存在しうる領域$((xl_o, yt_o), (xr_o, yb_o))$の左上1/4領域内に，エッジ強度マップとエッジ符号マップの情報をY方向に加算投影する．上から下へ探索すると，符合の累積は"-1"の方が多く，強度の累積分布は動的閾値より大きい所を左眉毛の上Y座標$(yl_{eb}[0])$とする．次に，累積分布の頂点を越えてから，符合の累積は"$+1$"の方が多く，強度の累積分布は動的閾値より小さい所を左眉毛の下Y座標$(yl_{eb}[1])$とする．更に，符合の累積は"-1"の方がまた多くなり，強度の累積分布は閾値より大きい所を左目の上Y座標$(yl_e[0])$とする．もう一回累積の頂点を越えて，符合の累積は"$+1$"の方が多くなり，強度の累積分布は閾値より小さい所を左目の下Y座標$(yl_e[1])$とする．

左眉毛の上下座標範囲$(yl_{eb}[0], yl_{eb}[1])$，左目の上下座標範囲$(yl_e[0], yl_e[1])$内でそれぞれ，エッジ強度マップ上の情報をX方向に加算投影することによって，左眉毛の左右X座標範囲$(xl_{eb}[0], xl_{eb}[1])$，左目の左右X座標範囲$(xl_e[0], xl_e[1])$を求められる．

同様な処理を，顔器官の存在しうる領域$((xl_o, yt_o), (xr_o, yb_o))$の右上1/4領域内にすれば，右眉の領域上下座標範囲$(yr_{eb}[0], yr_{eb}[1])$，左右座標範囲$(xr_{eb}[0], xr_{eb}[1])$，ならびに右目の領域上下座標範囲$(yr_e[0], yr_e[1])$，左右座標範囲$(xr_e[0], xr_e[1])$がそれぞれ特定できる．

5.3.3 鼻と口

本研究では，特定された両目の下位置$yl_e[1]$と$yr_e[1]$のうち，より下にあるものを鼻の上Y座標$(y_n[0])$とする．

ここで，今までの処理結果により，$((xl_e[0], y_n[0]), (xr_e[1], yb_o))$という範囲を鼻と口の存在しうる領域とする．この範囲内で，エッジ強度マップ上の情報をX方向へ加算投影し，左(右)から中央へ探索する．累積値が動的閾値より高い所のX座標$(x_m[0], or x_m[1])$を，口の左(右)位置とする．

続いて，$((x_m[0], y_n[0]), (x_m[1], yb_o))$の範囲内でエッジ強度マップと肌色類似度マップの情報を，それぞれY方向へ加算投影して，肌色類似度マップの累積分布の谷

の位置によって，口の中部と鼻の下部の位置を推定する．それらの位置に基づいて上から下へ探索すると，強度の累積分布状況を用いて，鼻の下端Y座標($y_n[1]$)，口の上下Y座標($y_m[0]$と$y_m[1]$)を求める．

最後に，(($x_m[0]$, $y_n[0]$), ($x_m[1]$, $y_n[1]$)) の領域内でエッジ強度マップの情報をX方向に加算投影して，鼻の左右X座標($x_n[0]$と$x_n[1]$)を求める．

以上の処理により，眉を大きく囲む矩形二つ，目に関して二つ，鼻，口に関して一つずつの矩形が各顔器官領域として得られた(図5.5参照)．

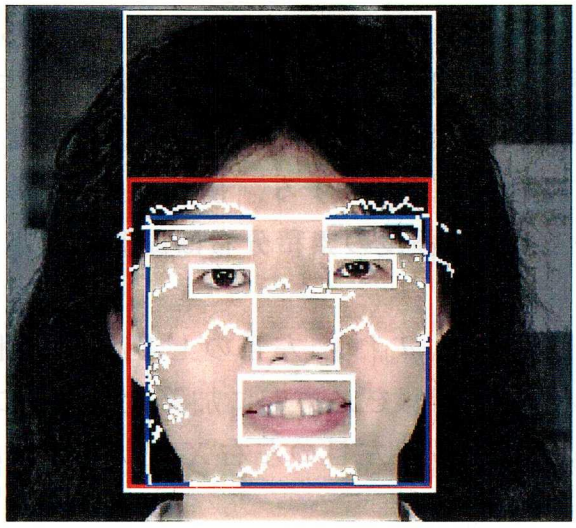


図 5.5: 加算投影による顔器官位置の同定．

5.4 顔器官幾何関係モデルによる顔候補の検証

人間の目や口などの顔器官の相対的な位置関係はほぼ一定である．本論文では，図5.6に示されるモデルで人間の両目と口の幾何関係を表現する．このモデルは顔器官幾何関係モデル (Relational Face Model あるいは RFM) と名づけ，並行移動，スケールの変化及び回転に依存しない．

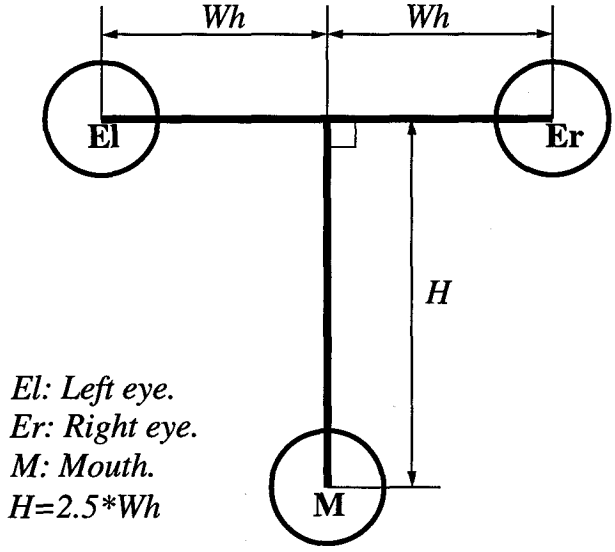


図 5.6: 顔器官幾何関係モデル．

顔候補の真偽は，抽出された顔特徴とそれに合わせた顔器官幾何関係モデルとの間の距離差によって判断する．定義された顔器官幾何関係モデルを抽出された顔特徴に合わせる処理は次の手順で行う．(図5.7参照)

1. 左右の目領域の中心点 (e_l) と (e_r)，口領域の中心点 (m) を求める．
2. 両目の真中の点 (M_i) を求める．
3. 点 (M_i) と点 (m) の間に直線 (L_1) を描く．
4. 点 (M_i) を通って，直線 (L_1) に直交する直線 (L_2) を描く．
5. 点 (e_l) を通って，直線 (L_2) に垂直する直線 (L_3) を描き，直線 (L_3) と (L_2) の交点 (e'_l) を求める．
6. 点 (e_r) を通って，直線 (L_2) に垂直する直線 (L_4) を描き，直線 (L_4) と (L_2) の交点 (e'_r) を求める．
7. 顔器官関係モデル上の点 (E_l) と点 (E_r) 間の距離は，点 (e'_l) と点 (e'_r) 間の距離と一致させるように，顔器官関係モデルを拡大（縮小）する．

8. 顔器官関係モデル上の$(E_l)/(E_r)$を$(e'_l)/(e'_r)$にセットして，モデルと抽出された顔特徴にフィットさせる．

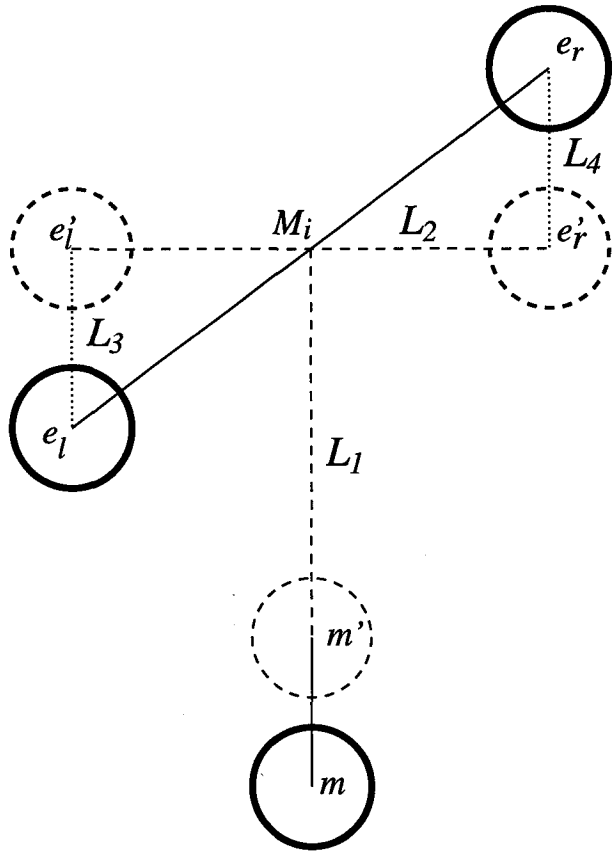


図 5.7: 抽出された顔器官と顔器官幾何関係モデルとの照合．

顔器官幾何関係モデルと抽出された顔特徴の間の距離差$D(face)$は，次式で求める．

$$D(face) = \frac{|e'_l - e_l| + |e'_r - e_r| + |m' - m|}{|e'_l - e'_r|} \tag{5.7}$$

この式には，分子は変換された顔器官幾何関係モデル$(e'_l,\ e'_r,\ m')$と抽出された顔特徴$(e_l,\ e_r,\ \mathrm{m})$の間の距離和であり，分母は顔器官幾何関係モデルの両目の間の距離である．経験的に，$D(face)$の値が0．4以下であれば，この顔候補は本当の顔と認める．

5.5 顔器官の決定と顔候補の検証の実験結果

提案した手法の有効性を確認するため，次のような実画像を用いて実験をした．なお，画像の大きさは2種類で，それぞれ 512×480，720×486 である．

- ビデオカメラで撮影した研究室内，大学構内のものや，テレビ放送から録画したもの
- 東大原島研からもらった顔画像データベース

入力画像として次のような制約条件が必要である．

- 画像内の顔は十分大きく，ほぼ正面を向いている
- 画像内の顔は眼鏡，帽子等を身に付けていない
- 画像の色がきれいに取れている

5.5.1 顔位置の決定

この段階では，メガネを掛けても，正確に決定できる．しかし，顔の両側の背景に肌色がある場合，頭部の上背景に黒色がある場合，誤って決定されたケースもあった．

5.5.2 顔器官の位置決定

顔の位置が正しく決定されたのち，メガネを掛けていないものに対して，各器官の位置を決定する処理を行う．正しく処理する可能な入力画像は，その中の顔の大きさは画像一杯から 80×100 画素ぐらいまでである．すべての器官を正しく抽出する確率は約85%であった．

ここで，目，眉のところに明確な輪廓線が多数あるので，ほとんど正確に特定できる．一方，鼻の下と口の上の間にはっきりした輪廓線が少ないので，その位置を特定する時に，照明の変動による影響を比較的に受けやすい．更に，髭あるいはそのあとがある場合，正しく特定することは一層難しくなる．

本アルゴリズムに基づいて，Impactビデオと Indigo^{2}の上に顔器官の輪廓と特徴点の抽出システムを構築した．720×486 の顔画像に対して，顔器官の位置を求める所用時間は約2秒である．

5.5.3 顔候補の検証

各顔器官を正しく同定できた後，その結果を顔器官幾何関係モデルと比較する処理を行った．顔候補が本当の顔であれば，両方の距離差はすべて0.4より小さいことがわかった．一部の例を図5.8，表5.1に示す．

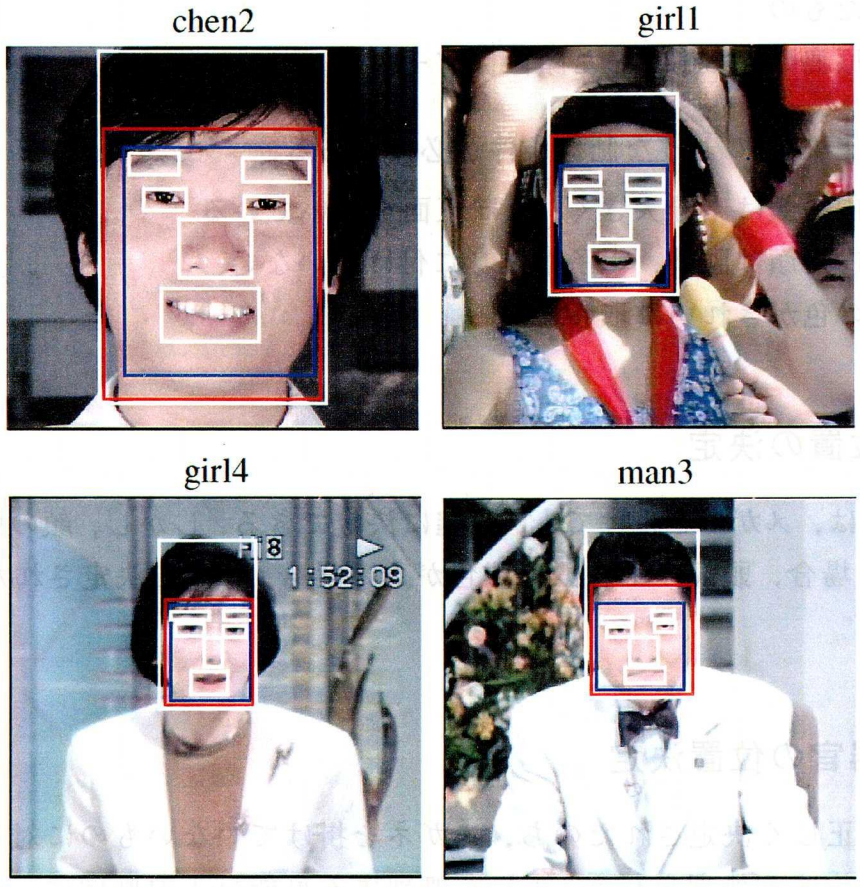


図 5.8: 顔器官の抽出結果．

	girl1	girl4	girl5	chen2	man3	man5	man9	wuhy2
Left eye	(89, 88)	(112, 80)	(118, 78)	(97, 94)	(110, 81)	(118, 75)	(182, 136)	(99, 128)
Right eye	(127, 92)	(141, 80)	(147, 82)	(165, 98)	(138, 78)	(143, 73)	(201, 137)	(163, 125)
Mouth	(107, 136)	(125, 112)	(135, 115)	(122, 162)	(125, 106)	(137, 111)	(191, 159)	(134, 188)
D(face)	0.125	0.191	0.236	0.346	0.375	0.348	0.1	0.29

Estimated difference between the relational face model and the estimated positions of the eyes and the mouth.

表 5.1: 顔候補の検証結果.

5.6 まとめ

本章では，色情報(肌と髪の色)および，エッジ情報(強度と符合)を総合的に考慮した加算投影法を提案した．実画像を対象として顔器官の抽出実験をした．その結果より，提案した手法が顔の大きさの変動に強く，異なる照明条件と複雑な背景を持つ顔画像に対して有効であることを確認した．抽出された各顔器官の位置情報に基づいて，顔候補の最終検証を行った．

求めた顔と顔器官を含む矩形領域は，顔器官の正確な輪郭を抽出する時の初期位置として利用できる．

また，求めた顔と顔器官の位置を利用することによって，より精度の高い姿勢推定が可能になると思われる．

第6章　動的輪廓モデルを用いた顔器官の輪廓と特徴点の抽出

CGへの入力や個人識別のためのデータベース作りなど，個人ごとの顔モデルを必要とする用途が数多くあり，そのためには顔及び顔器官の正確な輪郭，特徴点の位置を求める必要がある．これまで，顔器官の特徴点を得るには，手作業で行われることが多かったが，ここでは，自動的に獲得する方法について検討する．

以降では，6.1で顔器官輪郭の抽出について，6.2で特徴点の決定について述べる．6.3で顔器官の輪廓と特徴点の抽出の実験結果を示す．最後に討論とまとめを行う．

6.1　顔器官輪郭の抽出

物体の輪郭を滑かに求めるのによく用いられる方法として動的輪郭モデル（スネーク）がある．しかし，スネークを一般的な輪郭の検出問題に用いようとすると初期状態の与え方に結果が大きく依存する．また，初期状態の与え方が悪いと処理時間がかかるといった欠点がある．これに対して，5章では各顔器官を囲む矩形領域が既に求められているので，その結果を利用してスネークの初期位置を配置すれば，この問題を解決できる．

スネークを用いる時の他の問題点は，滑らかな曲線を求めるには適しているが，尖っているものの検出には適していないという点である．顔器官には目や口のように尖っているものが多いが，これらの端をシャープに検出するために，ここでは輪状のスネークではなく，複数本ひも状のスネークを用いることにする．

6.1.1　動的輪郭モデルのエネルギ関数

スネーク点列のパラメータを $v(s) = (x(s),\ y(s))$ で表現とすると，スネークの持つエネルギー E_{snake}は次式で表される．

$$E_{snake} = \int (E_{int}(v(s)) + E_{image}(v(s)) + E_{ext}(v(s)))ds \tag{6.1}$$

ここで，E_{snake}を最小化することにより輪郭が得られる．一般的に，内部エネルギーE_{int}は輪郭の滑らかさ，画像エネルギーE_{image}は輝度勾配，外部エネルギーE_{ext}はスネークに外部から強制的な力を表す．

顔器官の輪郭を正しく抽出するのに，スネークに適切な初期位置を与えるだけでは不十分であり，各顔器官の特徴に合わせたエネルギー関数の設定も重要である．本研究では，エネルギー関数を設定する時，一般的な画像エネルギー項の他に，個人性や表情変化，撮影条件などに左右されにくい色情報に関するエネルギー項を追加する．

$$E_{image} = w_{edge}E_{edge} + w_{skin}E_{skin} + w_{red}E_{red} \tag{6.2}$$

ここで，wは重みであるが，器官別に異なる値を持つ．E_{skin}は肌色らしさによる，E_{red}は赤ぽい色によるエネルギー項である．画像中ある画素(x,y)に対して，それぞれ次のように求める．

$$E_{skin} = SCDM(u_f(x,y),\ v_f(x,y)) \tag{6.3}$$

$$E_{red} = 2 \times R(x,y) - G(x,y) - B(x,y) \quad (R > G\ and\ R > B) \tag{6.4}$$

6.1.2 顔輪郭について

顔全体の輪郭を求めるためのスネークの初期位置としては，顔を囲む長方形領域をそのまま用いると複雑な背景の影響を受ける可能性があるので，図6.1の黒線に示すように，眉，目，唇の外側を連結したものを初期位置とした．そこから肌色領域を追跡することによって，スネークを広がるように設定する．前髪が一部眉毛にかかっている場合，この方法では，初期位置が取れないので，左目の脇から肌色領域の追跡をして，右目の脇までを初期位置とする開いたスネークを配置する．残りの部分は，髪部分から，肌色領域に向かう，開いたスネークを設定する．

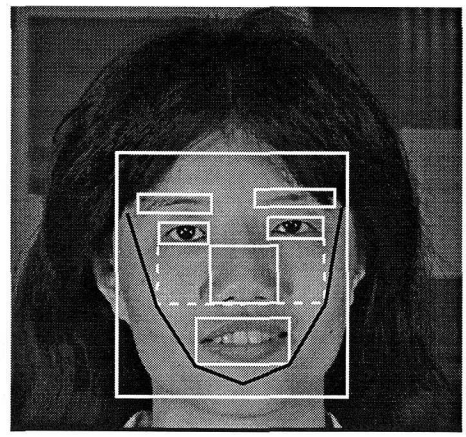


図 6.1: スネークの初期位置（顔）.

これを初期条件とすると，顔の輪郭まであまり雑音がないと思われるが，実際はほくろやしわといった雑音成分があり，また輪郭線も背景との関係で，鮮明ではなく曖昧な所もあるため，必ずしも正確な輪郭線が得られるとは限らない．このため顔輪郭の軸対称性に着目し，エネルギー関数に軸対称性を付加することにより，正確でロバストに顔輪郭を求めることが考えられる [横山 97].

6.1.3 眉について

眉の形状は個人性がある．さらに，眉の内側であっても，低いながらも肌色らしいと認識される場合もある．そこで，慎重な操作が必要となる．

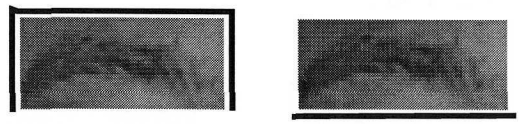


図 6.2: スネークの初期位置（眉）.

眉の初期位置は（図6.2参照），大まかな位置として与えられた長方形の上に二本のスネークを，上から'コ'の字に下から直線で配置する．上から内側に向かって収束させ，エネルギー関数に使う性質としては，肌色領域上であれば動きやすく，非肌色の領域では動きにくい，輝度の勾配を下らないなどである．下のものは上に向かうようにコントロールエネルギを加える．端点は左右に広がるようにする．

6.1.4 目について

目は，左右の端がとがっている上にそこでのエッジが一般的に弱い．それゆえ外側に囲むように普通のスネークを配置して，収束させても左右の端がうまく取れない．そこで，上下から'コ'の字型に開いた二本のスネークを配置し（図6.3参照）収束させる．エネルギー関数については，上まぶたの場合，肌色領域内で，輝度の勾配を上り詰めるように設定し，下まぶたの場合，肌色領域に於いて，輝度の勾配を下るように設定する．端点は，二本とも左右に広がるように設定する．

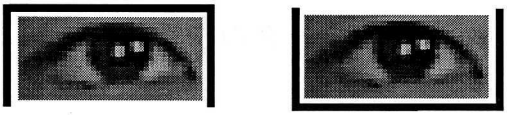


図 6.3: スネークの初期位置（目）．

6.1.5 鼻について

鼻には，左，右からと，下から，三本のスネークを適用することを考える（図6.4参照）．そのエネルギー関数を輝度の勾配に従って落ちていくような性質とする．鼻の穴の部分は，肌色領域として検出されないことが多いので，肌色でないところに入らないような性質もエネルギー関数に加える．それぞれ，端点は，広がる方向に少しづつ動くものとする．

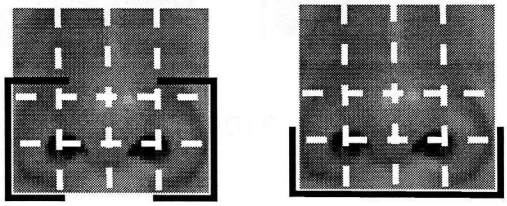


図 6.4: スネークの初期位置（鼻）．

6.1.6 口について

くちびるは，開いているときと閉じているときでは，形状が大いに違う．外側の輪郭の検出のために，まず大まかな位置を示す長方形から，上下から変形した'コ'の字型に開いた二本のスネークを配置し（図6.5参照）収束させる．外部エネルギーとして，赤色成分の強いところに入らないよう，肌色領域内では，輝度値の高い方に向かうようにエネルギー関数を設定する．次に，外の輪郭から少しうちに入った部分から，輝度値の低い方にいくようにエネルギー関数を設定する．そうすると，上唇の下輪郭，下唇の上輪郭は検出できる．

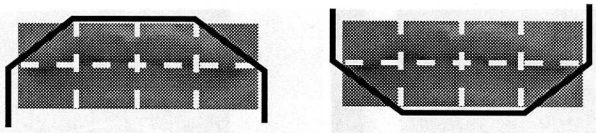


図 6.5: スネークの初期位置（口）．

6.2 特徴点の決定

個人モデルを作る時，標準的な三次元の顔モデルにその人の二次元顔画像をフィッティングするために最低限必要となる特徴点は以下に示すものである．

- 眉：右，左，上，下
- 目：右，左，上，下，瞳中心
- 鼻：頂点，右，左，下
- 口：右，左，上唇上，上唇下，下唇上，下唇下
- 他：顎下端，顔輪廓20点程度（髪型）

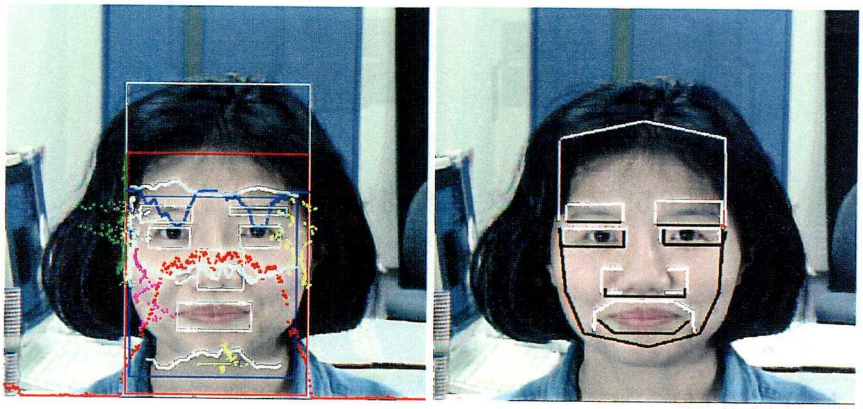
本研究では，眉に関して輪郭線及び上下左右の端点を，目については上下の輪郭線及び左右上下の端点．また，鼻の輪郭線，左右の端点，頂点を，くちびるは，左右の端点，上下唇のそれぞれ上下の端点を求める．

抽出された各顔器官の輪郭線より，その特徴点を求める．それぞれの顔器官の左右の端点は，開いた二本のスネークの収束後の二つの交点として明らかになる．上下の点は，それぞれの開いたスネークの端点を結び，上下それぞれのスネークの点列上の点の中央付近でその線から最も離れた点となる．

6.3　顔器官の輪廓と特徴点の抽出の実験結果

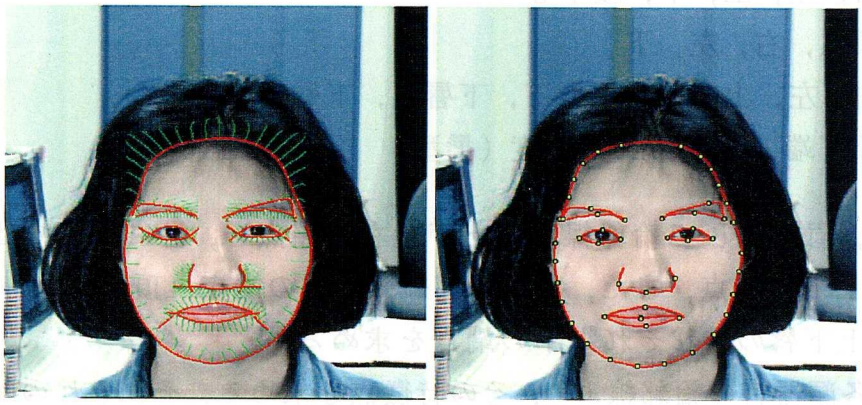
本研究で提案した方法の有効性を検証するために，実験室内で3CCDビデオカメラで撮影した顔画像と，原島研からもらった顔画像を入力画像として処理を行った．

各顔器官を囲む矩形領域という大まかな位置(図6.6(a)参照)を初期位置として，上に述べた手法のように，滑らかな輪郭を検出する．スネークの全ての器官の初期位置(図6.6(b)参照)からの収束過程状態の一例は図6.7(a)の様になる．その処理結果(求められた輪郭線と特徴点)を，図6.7(b)に示す．また，一部の実験結果を，図6.8に示す．



(a)　　(b)

図 6.6: (a)顔器官を囲む矩形領域，(b)スネークの全ての器官の初期位置．



(a)　　(b)

図 6.7: (a)スネークの収束過程，(b)求められた輪郭線と特徴点．

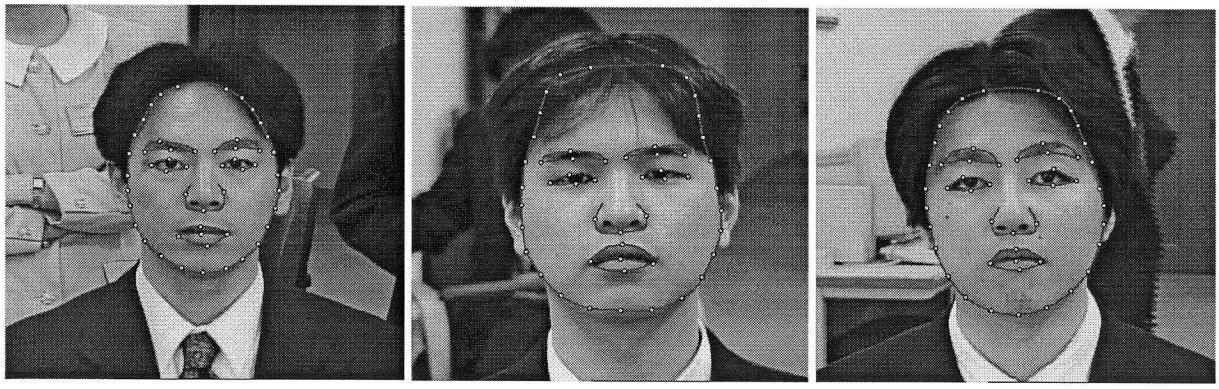


図 6.8: 輪郭線と特徴点の実験結果.

本アルゴリズムに基づいて，Impactビデオと Indigo^{2}の上に顔器官の輪廓と特徴点の抽出システムを構築した．720×486の顔画像に対して，顔器官の位置を正確に求めた後，その輪廓を抽出する所用時間は約5秒である．

6.4 まとめ

本章では，顔器官の輪郭の抽出，特徴点の獲得について述べた．顔器官の輪廓線の抽出には，複数本のスネークを用いた輪廓抽出法を提案した．本手法を用いることにより，眉，目，口など，尖った部分を持つ顔器官に対して有効であることが示された．さらに安定な結果を得るために，ポテンシャル場を導入したエネルギー項の再検討や，顔器官の対称性などを考える必要があると思われる．

顔および顔器官を自動的に抽出する技術を確立することによって，顔画像の入力から合成顔画像の生成までの一連の手続きを自動的に行うことができる．また，顔の構成要素である眉や目などの動きを分析することが可能となり，その分析結果を自然で感性豊かな顔表情の表出に役立てることができる．さらに，眉や目などの動きの分析は，人間の感情の変化を情報システムが検知するための手がかりにもなり，心理的側面を考慮した人間と情報システムとのコミュニケーションの実現に大いに貢献するものと考えられる．

第7章　おわりに

本論文では，人工現実感やヒューマンインターフェースなどのシステムへの応用を目的として，肌と髪の色情報を重視し，カメラから入力された色彩画像からの顔情報処理の方法を提案した．

複雑な背景を持つ画像から，大きさ，位置，姿勢，個数等がともに未知である顔の発見法を第3章で提案した．本手法では，色情報を均等知覚色空間で表現し，肌の色/髪の色を表現するモデルを構築した．これによって，肌色/髪色を“らしさ”で表現するという新しい方法を提案し，肌色領域/髪色領域を安定に抽出することを実現した．又，曖昧さを導入した頭部形状モデルも前もって構築した．異種情報を比較するために，新たに2項ファジイ関係を利用したパターン照合法を提案した．本手法に基づいた顔検出システムを構築し，それを用いて，色彩動画像から顔候補を安定に抽出できた．これは，本手法の有効性を示すだけでなく，他の画像処理に利用できる可能性も示唆している．

マシン前に操作しているユーザの顔を検出したのち，カメラに対する頭部の三次元姿勢の推定法を第4章で提案した．肌色・髪色領域の一次，二次モーメントといった安定に求められる情報を用いて姿勢を推定することによって，システムの信頼性を格段に向上した．この結果をもとに能動的カメラや複数台のカメラを用いた協調システムを制御すれば，自然な状態でいる人に対する正面顔を積極的に捉えることが可能であることを実験から示した．

ほぼ正面を向いた顔画像を対象として，顔器官による顔候補の検証方法を第5章で提案した．顔器官の位置を正確に検出するために，色情報(肌と髪の色)および，エッジ情報(強度と符合)を総合的に考慮した加算投影法を提案した．実験結果より，提案した手法が顔の大きさの変動に強く，異なる照明条件と複雑な背景を持つ顔画像に対して有効であることを確認した．又，顔器官が存在するか否かを検証することにより，偽の顔候補を除外することが可能であることを示した．

さらに，顔及び顔器官の正確な輪郭を求める方法を第6章で提案した．求められた顔器官の矩形領域よりスネークの初期位置の設定法を検討した．顔器官の端が尖っているところをシャープに検出するために，複数本のスネークを用いた輪廓抽

出法を提案した．実験結果より，二本のスネークの交点を求めることにより，端が正確に求められることを実証した．

以上，4点は顔情報処理の研究の基礎的発展に寄与するものであると確信する．

謝 辞

本研究を行うにあたり，常に適切な御指導御鞭撻を賜わり，研究に関する様々な御助言を頂きました指導教官の大阪大学基礎工学部谷内田正彦教授に心から深く感謝致します．また，国内はもとより，国際会議での発表の機会を与えて頂いたことは，本研究を行なう上で得難い経験となりました．さらに，在学期間に身元保証人として大変お世話になっており，重ねてお礼を申し上げます．

本研究をまとめるにあたって，御厚情あふる御指導，数々の有益な御教示を賜わり，学位審査を行なって頂きました大阪大学基礎工学部井口征士教授，同西田正吾教授，大阪大学産業科学研究所北橋忠宏教授に心から深く感謝致します．

なお，研究を進めるにあたり，様々な御助言・御指導を頂きました大阪大学基礎工学部八木康史助教授，山口智浩助手，大澤幸生助手，徐剛講師(現立命館大学助教授)に深く感謝致します．ならびに，様々な御協力を頂きました早瀬敏夫技官，前田祥子女史に感謝致します，さらに，数々の有益なご討論を頂きました馬野元秀助教授(現大阪府立大学教授)に厚くお礼申しあげます．

それから，福本猛氏(現三菱電機)，Dadet Pramadihanto氏(現博士課程1回生)，横山太郎氏(修士課程2回生)は私と共に研究を行ない，この論文の元となった素晴らしい成果をあげてくれました．ここに感謝の意を表します．

ならびに，日頃の討論や，助言等をしてくれました同輩の岩井儀雄氏(現大阪大学基礎工学部助手)，山澤一誠氏(現奈良先端科学技術大学院大学助手)，魏世杰氏に深く感謝致します．そして，日頃から各方面で研究に御協力頂いた(財)イメージ情報科学研究所渡辺孝弘氏、木村聰氏に、谷内田研究室の方々に深くお礼を申し上げます．

そして，本研究を行なう後期にあたり，非常に快適な研究環境を提供して頂いた京都工芸繊維大学工芸学部塩山忠義教授，村田滋助教授に心から感謝の意を表します．

また，一年間の助成・援助金を提供して頂いた(財)国際コミュニケーション基金に深く感謝致します．

最後に，研究生活，家庭生活の両面において全力に支えてくれました夫，そして優しく見守ってくれました両親ならびに娘に心から感謝の意を表します．

参考文献

[1] [R. Chellappa 95] R. Chellappa, C. L.Wilson, and S. Sirohey: "Human and Machine Recognition of Faces: A Survey", Proceeding of the IEEE. Vol.83, No.5 MAY 1995.

[2] [D. Terzopoulos 93] D. Terzopoulos and K. Waters: "Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models", IEEE Trans. on PAMI, Vol.15, No.6, pp.569-579, June 1993.

[3] [S. Hashimoto 87] S. Hashimoto, Y. Sato and H. Oda: "Modification of Facial Expression Using Spring Frame Model", Proc. IASTED, pp37-41, 1987.

[4] [H. Kitano 94] H. Kitano: "Artificial Intelligence", Artificial Life & Entertainment: AAAI'94 Workshop Proc. Seattle, 1994.

[5] [上木 94] 上木伸夫, 森島繁生, 山田寛, 原島博: "多層ニューラルネットによって構成された感情空間に基づく表情の分析・合成システムの構築", 電子情報通信学会論文誌, Vol.J77-D-II, No.3, pp.573-1582, 1994.

[6] [崔 91] 崔昌石, 原島博, 武部幹: "顔の 3 次元モデルを用いた顔面表情の分析", 電子情報通信学会論文誌, Vol.J74-D-II, No.6, pp.766-777, 1991.

[7] [金子 89] 金子正秀: "顔画像のモデルベース符号化", 信学会知的コミュニケーションと符号化ワークショップ資料, 2.3, pp.29-36, 1989.

[8] [角 91] 角, 中川, 登: "表情・年齢変化の可能な 3 次元顔画像合成システム", 信学技報, HC91-43, pp.9-16, 1991.

[9] [P. Ekman 78] P. Ekman and W. V. Friesen: "The Facial Action Codeing System", Consulting Psy,hologists Press Inc., San Francisco, CA, 1978.

[10] [崔 89] 崔昌石, 原島博, 武部幹: "分析合成符号化における表情の記述と合成について", 信学技報, IE88-84, 1989.

[11] [原 97] 原文雄: “人工感情をそなえた顔ロボットで機械と人のコミュニケーションを”, ニュートン, pp.40-41, 1997年2月号.

[12] [南 86] 南 敏: “顔の識別技術”, 計測と制御 Vol.25, No.8, pp.25-31, 1986.

[13] [M. Suwa 78] M. Suwa et al.: “A Preliminary Note on Pattern Recognition of Facial Emotional Expression”, 4th Int. Joint Conference on Pattern Recognition, pp.408-410, 1978.

[14] [阿部 90] 阿部享, 阿曽広具, 木村正行: “3次元表面形状による人間の顔の自動識別”, 電子情報通信学会論文誌, Vol.J73-D-II, No.9, pp.1477-1484, 1990.

[15] [萩原 88] 萩原 他: “パタンマッチングを主体にした顔画像による個人ＩＤ”, 電子情報通信学会技術研究報告, Vol.PRU88-46, pp.53-60, 1988.

[16] [T. Sakaguchi 89] T. Sakaguchi et al.: “Personal Identification Through Facial Images using Isodensity Lines”, Proc. Visual Communication and Images Processing IV, Vol.1199, No.2, pp.643-654, 1989.

[17] [宋 93] 宋柄卓, 小沢慎治: “時系列顔画像処理による個人の認識”, 電子情報通信学会技術研究報告, Vol.PRU93-66, pp.29-36, 1993.

[18] [赤松 93] 赤松茂, 佐々木努, 深町映夫, 末永康仁: “濃淡画像マッチングによるロバストな正面顔の識別法-フーリエスペクトルのＫＬ展開の応用-”, 電子情報通信学会論文誌, Vol.J76-D-II, No.7, pp.1363-1373, 1993.

[19] [M. Lades 93] M. Lades et.al.: “Distortion invariant object recognition in the dynamic link architecture”, IEEE Transactions on Computers, 42(3):300-311, 1993.

[20] [Dadet 97] Dadet Pramadihanto, 呉海元, 谷内田正彦: “多様な姿勢を有する入力画像からの個人識別”, 電子情報通信学会論文誌, J80-D-II. (97年8月掲載予定)

[21] [間瀬 90] 間瀬健二: “オプティカルフロー抽出による表情筋の動作検出”, 電子情報通信学会技術研究報告, Vol.PRU89-128, pp.17-24, 1990.

[22] [間瀬 91] 間瀬健二, 末永康仁: “表情動画像からの感情の認識の1手法”, 電子情報通信学会技術研究報告, Vol.PRU91-24, pp.95-101, 1991.

[23] [小林 93a] 小林宏, 原文雄: “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, Vol.29, No.1, pp.112-118, 1993.

[24] [小林 93b] 小林宏, 原文雄: "ニューラルネットによる人の顔の6基本表情の強さの計測に関する研究", 日本機械学会論文集(C編), 59巻567号, pp.177-183, 1993.

[25] [松野 94] 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", 電子情報通信学会論文誌D-II, Vol.J77-D-II, No.8, pp.1591-1600, 1994.

[26] [坂口 95] 坂口竜己, 大谷淳, 岸野文郎: "Wavelet変換による顔動画像の特徴抽出と表情認識への応用", 信学技報, IE94-147, PRU94-147, pp.15-22, 1995.

[27] [山口 95] 山口真美, 加藤 隆, 赤松 茂: "顔と感性情報と物理的特徴との関連について -年齢/性の情報を中心に-", 信学技報, HC94-89, 1995.

[28] [輿水 94] 輿水大和, 村上和人 他: "似顔絵生成システム開発による認知的ビジョンの研究", 重点領域研究平成5年度成果報告書「感性情報処理の情報学・心理学的研究」, pp.103-106, 1994.

[29] [A. Tsukamoto 94] A. Tsukamoto, C. W. Lee and S. Tsuji: "Detection and Tracking of Human Face with Synthesized Templates", Proc. Asian Conf. Computer Vision, pp.183-186, 1994.

[30] [K. Matsuno 95] K. Matsuno, C. W. Lee, S.Kimura and S.Tsuji: "Automatic Recognition of Human Facial Expressions", Proc. 5th ICCV, pp.352-359, 1995.

[31] [A. Lanitis 95] A. Lanitis, C. J. Taylor and T. F. Cootes: "A Unified Approach To Coding and Interpreting Face Images", Proc. 5th ICCV, pp.368-373, 1995.

[32] [T. Sakai 69] T. Sakai, M. Nagao, and S. fujibatashi: "Line extraction and pattern recognition in a photograph", Patt. Recog., Vol.1, pp.233-248, 1969.

[33] [M. D. Kelly 70] M. D. Kelly: "Visual identification of people by computer," Tech. Rep. AI-130, Stanford AI Proj., Stanford, CA, 1970.

[34] [V. Govindaraju 90] V. Govindaraju, S. N. Srihari, and D. B. Sher: "A computational model for face location," Proc. 3rd ICCV, pp.718-721, 1990.

[35] [I. Craw 87] I. Craw, H. Ellis and J. Lishman: "Automatic extraction of face feature," Patt. Recog. Lett., vol.5, pp. 138-187, 1987.

[36] [小杉 94] 小杉 信: "個人識別のための多重ピラミッドを用いたシーン中の顔の探索・位置決め", 電子情報通信学会論文誌, Vol. J77-D-II, No.4, pp.672-681, 1994.

[37] [P. J. Burt 89] P. J. Burt: "Multiresolution techniques for image representation, analysis, and smart transmission", SPIE Proc. Visual Commun. and Image Process. IE, Vol.1199, pp.2-15, 1989.

[38] [塚本 94] 塚本明利, 李七雨, 辻三郎: "複数のモデル画像による顔の動き推定", 電子情報通信学会論文誌 D-II, Vol.J77-D-II, No.8, pp.1582-1590, 1994.

[39] [J. J. Weng 93] J. J. Weng,N. Ahuja,and T. S. Huang: "Learning Recognition and Segmentation of 3-D Objects from 2-D images", 4th ICCV, pp.121-128, 1993.

[40] [福田 93] 福田敏男, 伊藤茂則, 新井史人: "ファジィ推論とニューラルネットワークを用いた人物画像の認識に関する研究", 日本機械学会論文集（C編）, 59巻558号, pp.200-206, 1993.

[41] [宮脇 89] 宮脇隆志, 石橋聡, 岸野文郎: "色彩情報を用いたカラー画像の領域分割", 信学技報, IE89-50, pp.43-48, 1989.

[42] [M. Imai 91] M. Imai, M. Takahata and S. Tsuji: "Detection of Human Faces in Image Sequences", Proc. First Korea-Japan Joint Conference on Computer Vision, pp.174-179, 1991.

[43] [長谷川 94] 長谷川, 横澤, 石塚: "自然感の高いビジュアルヒューマンインタフェースの実現のための人物動画像の実時間並列協調的認識", 電子情報通信学会論文誌 (D-II), Vol. J77-D-II, No.1, pp.108-118, 1994.

[44] [G. Yang 93] G. Yang and T. S. Huang: "Human Face Detection in a Scene", Proc. IEEE conf. on Computer Vision and Pattern Recognitio, pp.453-458, 1993.

[45] [X. Song 94] X. Song, C.W. Lee, G. Xu, and S. Tsuji: "Extracting Facial Features with Prtial Feature Template", Proc. Asian Conference on Computer Vision, pp.751-754, 1994.

[46] [H. Zabrodsky 92] H. Zabrodsky, S. Peleg, and D. Avnir: "Hierachical Symmetory", Proc. IEEE conf. on Computer Vision and Pattern Recognitio, pp.9-12, 1992.

[47] [宋 94] 宋欣光, 李七雨, 徐剛, 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", 電子情報通信学会論文誌,D-II Vol.J77-D-II, No.8, pp.1601-1609, 1994.

[48] [A. L. Yuille 92] A. L. Yuille, P. W. Hallinan and D. S. Cohen: "Feature Extraction from Faces Using Deformable Templates", Int. Journal of Computer Vision, 8:2, pp.99-111, 1992.

[49] [D. Reisfeld 92] D. Reisfeld and Y. Yeshurun: "Robust detection of facial features by generalized symmetry", Proc. 11th IAPR Int. Conf. on Pattern Recognition, pp.117-120, 1992.

[50] [J.M. Vincent 91] J. M. Vincent, J. B. Waite and D. J. Myers: "Precise Location of Facial Features by a Hierarchical Assemly of Neural Nets", 2nd Int. Conf. on Artificial Neural Networks, pp.69-73, 1991.

[51] [M. Kirby 90] M. Kirby and L. Sirobich: "Application of the Karhunen-Loeve procedure for the characterixation of human faces", IEEE Trans. Pattern Anal. & Mach. Intell., Vol.12, No.1, pp.103-108, 1990.

[52] [Z. Q. Hong 91] Z. Q. Hong: "Algebraic Feature Extraction of Image for Recognition", Pattern Recognition, Vol.24, No.3, pp.211-219, 1991.

[53] [坂井 73] 坂井利之, 長尾真, 金出武雄: "計算機による顔写真の解析", 電子通信学会論文誌, Vol.56, D-4, pp.226-233, 1973.

[54] [R. Brunelli 93] R. Brunelli, and T. Poggio: "Face Recognition: Features versus Templates", IEEE. Trans. Pattern Anal. & Mach. Intell. Vol.15, No.10, pp.1042-1052, 1993.

[55] [Wyszecki 67] Günter Wyszecki and W. S. Stiles: "COLOR SCIENCE",John Wiley & Sons, Inc. 1967.

[56] [間瀬 91] 間瀬健二, 渡部保日児, 末永康仁: "ヘッドリーダ：画像による頭部動作の実時間検出", 電子通信学会論文誌, Vol.J74-D-II, No.3, pp398-406, 1991.

[57] [P. Ballard 95] P. Ballard and G. C. Stockman: "Controlling a Computer via Facial Aspect", IEEE Trans. Systems, Man, and Cybernetics, Vol.25, No.4, pp.669-677, Apr. 1995.

[58] [谷内田 86] 谷内田正彦: "ロボットビジョン", 昭晃堂, 1990.

[59] [B. Klaus 86] B. Klaus and P. Horn: "Robot Vision", The MIT Press, 1986.

[60] [S. Basu 96] S. Basu, I. Essa, and A. Pentland: "Motion Regularization for Model-based Head Tracking", 13th Int. Conf. on Pattern Recognition, Aug. 25-30, 1996

[61] [J. Zheng 91a] J. Zheng, Q. Chen, F. Kishino and S. Tsuji: "Active camera controlling for manipulation", Proc. IEEE conf. on Computer Vision and Pattern Recognitio, pp.413-418, 1991.

[62] [J. Zheng 91b] J. Zheng, Q. Chen, F. Kishino and S. Tsuji: "Active camera guided manipulation", IEEE Int. Conf. Robotics and Automation, pp.632-638, 1991.

[63] [M. Kass 87] M. Kass, A. Witkin and D. Terzopoulos: "Snakes: Active Contour Models", Int. J. Comput. vision 1, pp.321-331, 1987.

[64] [G. Xu 94] G. Xu, E. Segawa and S. Tsuji: "Robust Active Contours With Insensitive Parameters", Pattern Recognition, Vol.27, No7, pp.879-884, 1994.

[65] [横山 97] 横山太郎, 八木康史, 谷内田正彦, 呉海元: "顔の軸対称性を考慮した顔輪郭の自動抽出", 電子情報通信学会論文誌, Vol.J80-D-II. (97年8月掲載予定)

[66] [太田 96] 太田肇, 小沢慎治: "顔部品の位置検出に基づく顔の輪郭線画作成 ", 信学技報, PRU95-211, pp.7-12, 1996.

[67] [M. Turk 91] M. Turk and A. Pentland: "Eigenfaces for recognition", Journal of Cognitive Neuroscience, 3(1):pp.71-86, 1991.

[68] [R. Brunelli 93] R. Brunelli and T. Poggio: "Face recognition: Features versus templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):pp.1042-1052, 1993.

[69] [A. Pentland 94] A. Pentland, B. Moghaddam and T. Starner: "View-based and modular eigenspaces for face recognition", Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp.84-91, 1994.

[70] [D. J. Beymer 94] D. J. Beymer: "Face recognition under varying pose", Proc. IEEE conf. on Computer Vision and Pattern Recognition, pp.756-761, 1994.

[71] [T. Maurer 95] T. Maurer and C. von der Malsburg: "Single-view based recognition of faces rotated in depth", Proc. of the International Workshop on Automatic Face- and Gesture Recognition, pp.248-253, Zurich, 1995.

[72] [W. Konen 93] W. Konen and C. von der Malsburg: "Learning to generalize from single examples in the dynamic link architecture", Neural Computation, 5:pp.719-735, 1993.

[73] [W. Konen 94] W. Konen, T. Maurer and C. von der Malsburg: "A fast dynamic link matching algorithm for invariant pattern recognition", Neural Network, 7:pp.1019-1030, 1994.

[74] [I. Her 93] I. Her: "A symmetrical coordinate frame on the hexagonal grid for computer graphics and vision", Transactions of the ASME, J. of Mech. Design, 155:pp.447-449, 1993.

[75] [J. Daugman 88] J. Daugman: "Complete discrete 2-D Gabor transform by neural networks for image analysis and compression", IEEE Transactions on Acoustics, Speech, Signal Proc., 36(7):pp.1169-1179, 1988.

[76] [H. Kobayashi 93] H. Kobayashi and F. Hara: "Dynamic Recognition of Basic Facial Expressions by Discrete-time Recurrent Neural Network", Proc. Int. Joint Conference on Neural Network, pp.155-158, 1993.

研究業績

1. 学術論文

1-1) 呉海元, 陳謙, 谷内田正彦, "ファジイパターン照合を用いた色彩画像からの顔検出システム", 電子情報通信学会論文誌, Vol.J80-D-II, pp.1774-1785, 1997.

1-2) DADET PRAMADIHANTO, 呉海元, 谷内田正彦, "多様な姿勢を有する入力画像からの個人識別", 電子情報通信学会論文誌, Vol.J80-D-II. (97年8月掲載予定)

1-3) 横山太郎, 八木康史, 谷内田正彦, 呉海元, "顔の軸対称性を考慮した顔輪廓の自動抽出", 電子情報通信学会論文誌, Vol.J80-D-II. (97年8月掲載予定)

1-4) Wenbiao JIANG, Haiyuan WU and Tadayoshi SHIOYAMA, "3-D Shape Recovery from Image Brightness for non-Lambertian Surface", Journal of Imaging Science and Tecnology, Vol.41: No.4 (July/August) pp.116-124, 1997.

1-5) Haiyuan WU, Qian CHEN, and Masahiko YACHIDA, "Face Detection from Color Images using Fuzzy Pattern Matching Method", IEEE Transactions on Pattern Analysis and Machine Intelligence. (submitted)

1-6) 呉海元, 福本猛, 陳謙, 谷内田正彦, "色彩画像からの人物の頭部の三次元姿勢推定", 映像情報メディア学会論文誌. (投稿中)

1-7) 呉海元, 江文彪, 塩山忠義, "単眼明度画像からの三次元物体姿勢推定", 日本機械学会論文集, C1編. (投稿中)

1-8) Wenbiao JIANG, Haiyuan WU and Tadayoshi SHIOYAMA, "3-D Shape Recovery from Color Information for non-Lambertian Surface", Journal of Imaging Science and Tecnology. (submitted)

2. 国際会議論文（査読あり）

2-1) Tadayoshi SHIOYAMA, Haiyuan WU, Wenbiao JIANG and Susumu TERAUCHI, "3-D Object Positioning from Monocular Image Brightness", 9th Int. Conf. on Image Analysis and Processing, Florence, 1997-9. (accepted).

2-2) Haiyuan WU, Takesi FUKUMOTO, Qian CHEN and Masahiko YACHIDA, "Face Detection and Rotations Estimation using Color Information", 5th IEEE Int. Workshop on "Robot and Human Communication" pp.341-346, Tsukuba, 1996-11.

2-3) Taro YOKOYAMA, Haiyuan WU, and Masahiko YACHIDA, "Automatic Detection of Contours and Feature Points of Face and Facial Frature", 5th IEEE Int. Workshop on Robot and Human Communication, pp.335-334, Tsukuba, 1996-11.

2-4) DADET PRAMADIHANTO, Haiyuan WU, and Masahiko YACHIDA, "Face Recognition from a Single Example View Based on Flexible Nerual Network Matching", 5th IEEE Int. Workshop on Robot and Human Communication, pp.329-334, Tsukuba, 1996-11.

2-5) Haiyuan WU, Taro YOKOYAMA, DADET PRAMADIHANTO and Masahiko YACHIDA, "Face and Facial Feature Extraction from Color Image" , Proc. 2th Int. Conf. on Automatic Face- and Gesture- Recognition, pp.345-350, Killubgton, vermont, 1996-10.

2-6) Haiyuan WU, Qian CHEN, and Masahiko YACHIDA, "Detecting Human Face in Color Images", 1996 IEEE Int. Conf. on Systems, Man and Cybernetics (SMC), Vol.3, pp.2232-2237, Beiji, 1996-10.

2-7) Dadet Pramadihanto, Haiyuan WU, and Masahiko YACHIDA, "Invariant Face Recognition by Gabor Wavelet and Neural Network Matching", 1996 IEEE Int. Conf. on Systems, Man and Cybernetics (SMC), Vol.1, pp.59-63, Beiji, 1996-10.

2-8) Haiyuan WU, Takesi FUKUMOTO, Qian CHEN, and Masahiko YACHIDA, "Active Face Observation System", Int. Conf. on Pattern Recognition (13th-ICPR), Vol.C, pp.441-445, Vienna, 1996-8.

2-9) Haiyuan WU, Qian CHEN, and Masahiko YACHIDA, "An Fuzzy Theory Based Face Detector", Int. Conf. on Pattern Recognition (13th-ICPR), Vol.C, pp.406-410, Vienna, 1996-8.

2-10) Haiyuan WU, Qian CHEN, and Masahiko YACHIDA, "Facial Feature Extraction and Face Verification", Int. Conf. on Pattern Recognition (13th-ICPR), Vol.C,

pp.484-488, Vienna, 1996-8.

2-11) Haiyuan WU, Takesi FUKUMOTO, Qian CHEN and Masahiko YACHIDA, "Obtaining Optimal Image for Face Recognition Using an Active Camera", Proc. Second Asian Conference on Computer Vision , Vol.3, pp.708–712, Singapore, 1995-12.

2-12) Qian CHEN, Takesi FUKUMOTO, Haiyuan WU and Masahiko YACHIDA, "Face Observation Using an Active Camera", Proc. 6th Int. Conf. Human-Computer Interaction, Vol.B, pp.553-558, Yokohama, 1995-7.

2-13) Qian CHEN, Haiyuan WU and Masahiko YACHIDA, "Face Detection by Fuzzy Pattern Matching", 6th Int. Conf. Human-Computer Interaction, Poster Sesstions: Abridged Proc. p.20, Yokohama, 1995-7.

2-14) Qian CHEN, Haiyuan WU and Masahiko YACHIDA, "Face Detection by Fuzzy Pattern Matching", Proc. 5th Int. Conf. on Computer Vision (5th-ICCV), pp. 591-596, MIT, 1995-6. (採択率25%)

2-15) Haiyuan WU, Qian CHEN and Masahiko YACHIDA, "An Application of Fuzzy Theory: Face Detection", Proc. 1th Int. Workshop on Automatic Face- and Gesture- Recognition, pp. 314-319, Zurich, 1995-6.

2-16) DADET PRAMADIHANTO, Yoshio IWAI, Masahiko YACHIDA and Haiyuan WU, "Identifying Faces under Verying Pose Using a Single Example View", Proc. 3th Asian Conf. on Computer Vision, (submitted).

3. その他

3-1) DADET PRAMADIHANTO, 呉海元, 谷内田正彦, "多様な姿勢を有する顔画像からの個人識別", 情報処理学会研究報告 97-CVIM-105, pp.33-40, 静岡, 1997-5.

3-2) 福本猛, 呉海元, 谷内田正彦, "能動的カメラシステムによる顔認識のための最適顔画像の獲得", 第3回画像の理解・認識シンポジウム (MIRU'96), pp.II-37-42, 奈良, 1996-7.

3-3) 横山太郎, 呉海元, 谷内田正彦, "動的輪郭モデルを用いた色彩画像からの顔器官の同定", 第3回画像の理解・認識シンポジウム (MIRU'96), pp.II-319–324, 奈良, 1996-7.

3-4) 横山太郎, 呉海元, 谷内田正彦, "色彩画像からの顔の発見と顔部品の同定", 情報処理学会研究報告 96-CVIM-100-11, pp.85-92, 東京, 1996-5.

3-5) 福本猛, 呉海元, 谷内田正彦, "能動的カメラによる顔観察", 電子情報通信学会研究報告, 信学技報 Vol.95, PRU95-241, pp.41-46, 千葉, 1996-3.

3-6) 呉海元, 陳謙, 谷内田正彦, "カラー自然画像からの顔の検出", 電子情報通信学会研究報告, 信学技報 Vol.94, PRU94-108, pp.7-14, 竜古大, 1995-1.

4. その他出版物

4-1) 呉海元, 陳謙, 谷内田正彦, "カラー自然画像からの顔の検出", 画像ラボ, Vol.6/No.11, pp.19-22, 1995-11.