

Title	セキュリティシステムに適した顔認証技術 : FACELOCKの開発をめぐる
Author(s)	平山, 高嗣
Citation	大阪大学, 2005, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/347
rights	
Note	

Osaka University Knowledge Archive : OUKA

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

Osaka University

平成 16 年度 博士学位論文

セキュリティシステムに適した顔認証技術
—FACELOCK の開発をめぐる—

平山 高嗣

内容梗概

近年、情報化社会の発展に伴い、個人情報の流出とそれを利用した犯罪が増加している。その原因として、機密情報の保護のための入退室管理や情報システムへのアクセス制御に、流出や看破の危険性がある暗証番号やパスワードを利用してきたことが挙げられる。そこで、偽造、盗難、改ざん、紛失といった危険性が低い生体情報に注目が集まっている。生体情報の分析や認識に関する研究は、情報技術の発達に伴い活性化している。その生体情報を個人認証システムに適用することで高セキュリティを実現できる可能性がある。

生体情報の中でも、顔の認証システムへの適用は、特に高セキュリティ、高ユーザビリティを実現できるため注目されている。顔画像を用いた個人認証は、顔の撮影条件と経年変化によって、画像に写った顔の見え方が様々に変化するため、精度を高く安定させることが難しい。これまでに数多くの顔画像照合研究が行われてきたが、従来手法のほとんどは、顔認証とは異なるタスクである顔認識を目的に提案されたものである。ゆえに、従来手法を応用した個人認証は高く安定した精度を実現できていない。本研究では、顔認証の問題設定を分析し、その解決に適した顔認証技術を提案する。そして、それらの技術に基づく顔認証セキュリティシステム FACELOCK を開発する。

顔認証において、顔位置の推定は最も重要な過程である。顔の形状モデルを用いて顔位置を推定する場合、最適なモデルが用いられなければ、その精度は低下する。この問題を解決するにあたり、推定精度と処理時間のトレードオフが生じる。本研究では、顔の位置と大きさと個人モデルの推定を統合した手法を提案する。顔の大きさと個人性の変化は、セキュリティシステムの運用時に最も頻繁に起こりうる。位置と大きさは、粗密法とスケール変換という統計的手法に基づく繰り返し処理によって推定する。個人モデルの推定は、クラスタリングによって階層化したモデルデータベースを利用する。この推定は、位置と大きさの推定が進むにつれ、その階層構造に従って最適なモデルを含むクラスタを選択して

いく。最終的に選択された個人モデルは個人識別結果とみなすことができる。つまり、提案手法は、顔位置推定と個人識別を同時に行う手法であり、これは従来提案されなかった新しい方法論である。提案手法が市販セキュリティシステムに導入されている顔認識手法より、高精度かつ高効率であることを実験により確かめた。

従来の顔認証セキュリティシステムが持つ問題としては、顔写真による詐称に未対応であることも挙げられる。本研究では、赤外光を利用して高効率に顔領域の検出と、人物と顔写真の判別を行う技術を提案する。赤外光は、背景成分の変化にロバストな顔領域検出を可能にする。人物と顔写真の判別には、認証物体面の凹凸に注目し、照度差ステレオ法によって観測できる凹凸の情報を利用する。提案技術は、詐称者が写真を湾曲させた顔マスクを装着するような状況にも対応できる。この状況は従来研究では考慮されなかった。提案技術の有効性は、FACELOCKを用いた実環境での実験によって確認した。

目次

第1章	序論	1
1.1	顔認証とその問題設定	2
1.2	本研究の位置づけ	3
第2章	関連研究と提案技術	5
2.1	顔画像照合	5
2.2	顔の大きさ変化に対する特徴量の正規化	8
2.3	人物写真判別	9
2.4	赤外光を利用した顔検出	10
第3章	顔の位置と大きさと個人性を考慮した顔照合システム	11
3.1	顔グラフによる顔の表現	11
3.1.1	ガボール特徴量	12
3.1.2	グラフ特徴量類似度	15
3.2	個人モデルデータベースの階層的クラスタリング	15
3.3	顔位置と大きさの推定と個人識別の統合アルゴリズム	17
3.3.1	粗密法による顔位置推定	18
3.3.2	大きさ推定のためのスケール変換	20
3.3.3	個人モデル推定	22
3.3.4	状態遷移	22
3.4	標準顔画像データベースによる評価実験	23
3.4.1	実験環境	23
3.4.2	構成された階層モデルデータベース	25
3.4.3	スケール変換の性能検証	26
3.4.4	設定パラメータに対する性能分析	28
3.4.5	個人モデル推定の有効性の検証	32
3.4.6	従来研究との性能比較	33
3.5	まとめ	38

第4章	赤外面像を用いた人物写真判別システム	41
4.1	赤外光の特性と赤外面像の特徴	41
4.2	システム構成	44
4.3	顔領域検出アルゴリズム	45
4.4	人物写真判別アルゴリズム	49
4.5	まとめ	53
第5章	顔認証システム FACELOCK	55
5.1	システム概要	56
5.2	操作部と制御部	56
5.3	顔認証部	57
5.4	通信部	58
5.5	評価実験	58
5.5.1	実験環境	58
5.5.2	人物写真判別用の顔パラメータの導出	60
5.5.3	顔領域検出の性能検証	61
5.5.4	人物写真判別の性能検証	63
5.5.5	1対多照合の性能検証	63
5.5.6	1対1照合の性能検証	65
5.6	まとめ	68
第6章	結論	69
付録A	顔の大きさの変化にロバストな個人識別システム	87
A.1	システム概要	87
A.2	顔位置推定	87
A.3	フレキシブル特徴照合法	88
A.4	個人識別	89
A.5	システム性能	90
A.5.1	特徴点抽出の精度	91
A.5.2	個人識別の精度	93
付録B	顔認証システム用通信プロトコル SFAP	97
B.1	プロトコルスタック	97
B.2	データ表現	98
B.3	プロトコルの概略	100
B.4	プロトコルの実装	100

B.5	プロトコル評価のためのモデル化	101
B.6	メッセージ最悪応答時間	102
B.7	平均スループット	103
B.8	SFAP の性能	103
B.8.1	実験環境	103
B.8.2	データ伝送平均時間	104
B.8.3	スループット性能	105

目次

3.1	顔の表現	12
3.2	ガボールウェーブレットの形状	13
3.3	改良したガボールウェーブレットの形状	14
3.4	階層モデルデータベースの例	16
3.5	顔画像	17
3.6	顔画像 (図 3.5) と (a) 平均モデルグラフ, (b) 個人モデルグラフの相関値の分布	18
3.7	提案手法の状態遷移	19
3.8	顔位置推定	19
3.9	スケール変換	21
3.10	構成された階層モデルデータベース	26
3.11	最適な比率 r_{new} の候補が持つ確率の累積	27
3.12	$\Delta R_{th}, d_{smp}, n_w, c$ の変化に対する顔位置推定の精度と計算コスト	29
3.13	$\Delta R_{th}, d_{smp}, n_w, c$ の変化に対する大きさ推定の精度と計算コスト	30
3.14	$\Delta R_{th}, d_{smp}, n_w, c$ の変化に対する個人識別の精度と計算コスト	31
3.15	顔の位置と大きさと個人モデル推定の推移例	34
3.16	推定された顔位置の収束性	35
3.17	顔の位置と大きさと個人モデルの推定結果画像	36
3.18	個人識別精度の従来システムとの比較	39
3.19	再照合後の個人識別精度	40
4.1	可視光画像と赤外画像の違い	42
4.2	可視光と赤外光の波長域	42
4.3	光沢紙に印刷された写真を撮影した可視光画像と赤外画像	43
4.4	普通紙に印刷された写真を撮影した可視光画像と赤外画像	43
4.5	カメラ, 照明, 認証物体の位置関係	44
4.6	左右照射画像と左右照度差画像	45

4.7	人物写真判別システムの処理の流れ	46
4.8	2値化後の画像とノイズ除去後の画像	47
4.9	頭部領域の推定	48
4.10	顔領域と目領域の推定	50
4.11	光沢紙に印刷された顔写真の照度差画像	51
4.12	普通紙に印刷された顔写真の照度差画像	51
4.13	凹凸度	52
5.1	FACELOCK の概略	56
5.2	操作部と制御部	57
5.3	赤外カメラ KP-M2R の分光感度特性	58
5.4	赤外光パスフィルタの光透過率	59
5.5	顔認証部の処理の流れ	60
5.6	左右照射画像と左右照度和画像	61
5.7	顔領域の推定結果画像	62
5.8	顔グラフ照合の結果画像	65
5.9	1対多照合モードの個人識別精度	66
5.10	1対多照合モードのテストセットごとの個人識別精度	66
5.11	1対1照合モードの個人識別精度	67
5.12	1対1照合モードのテストセットごとの個人識別精度	68
A.1	個人識別システムの概略	88
A.2	ランダム拡散によるシャドウグラフの生成	90
A.3	顔特徴点抽出結果	92
A.4	個人識別精度	95
B.1	プロトコルスタック	97
B.2	メッセージフォーマット	98
B.3	応答メッセージのフォーマット	99
B.4	基本プロトコルのやりとり	101
B.5	SFAP の状態遷移	102
B.6	認証要求フェーズのやりとり	103
B.7	結果照会フェーズのやりとり	104
B.8	画像データ転送フェーズのやりとり	105
B.9	サーバスループット	106

表目次

3.1	個人モデル推定を行うシステムと行わないシステムの顔位置推定の誤差 (画素)	33
3.2	顔の位置と大きさの推定精度と計算コストの従来システムとの比較	37
3.3	登録人物と未登録人物の個人識別率のクロスポイントによる従来システムとの比較	38
5.1	人物写真判別部における顔領域の推定精度	63
5.2	人物写真判別の精度	63
5.3	1対多照合モードの顔の位置と大きさの推定精度	65
5.4	1対1照合モードの顔の位置と大きさの推定精度	67
A.1	モデルと入力画像のガボール特徴量の差	91
A.2	特徴点抽出の精度 (画素)	93
B.1	データ伝送平均時間 (ms)	104

第1章

序論

顔は、個人を特定するための生体情報の1つである。顔パスという言葉が存在することや、免許証やパスポートに顔写真が貼付されることから、我々にとって顔は最も馴染みある個人特定のための媒体といえる。顔をはじめとする生体情報は、偽造、盗難、改ざん、紛失といった危険性が低く、高セキュリティを実現できる可能性がある。このため、近年、生体情報、特に顔画像を利用したセキュリティシステムに関する研究が盛んに行われている。

顔画像を用いた個人認証は、システムが個人情報を非接触で取得できる点で有効である。また、人と人は普段、対面コミュニケーションを行っているため、顔画像の利用はユーザの心理的抵抗を軽減できるという利点がある。さらには、システムが顔画像を記録するため、不正に対する心理的抑止効果がある。顔以外の指紋、虹彩、静脈などの生体情報では、これらの利点を実現することができない。一方で、欠点としては、顔の撮影条件と経年変化によって、画像に写った顔の見え方が様々に変化する点が挙げられる。従来、多数の顔画像照合技術が提案されてきた。しかし、それらを応用した個人認証の精度が高く安定していないのが現状である。

顔画像による個人認証を利用したセキュリティシステムに対するニーズは、入退室や入出国管理、コンピュータシステムへのアクセス制御などの面で高まっている。今後の応用のためには、より高精度な認証技術が求められている。

1.1 顔認証とその問題設定

顔画像を用いた個人認証，いわゆる顔認証 (face authentication) は，撮影された人物がデータベースに登録されているかどうかを判定し，登録されていれば，それが誰であるかを識別する．これは，データベース中のどのモデルに類似しているかを求める顔認識 (face recognition) とは異なるタスクである [Kotropoulos 2000]．解錠制御やアクセス制御などのユーザの意識下で動作するセキュリティシステムの顔認証タスクには，

- 1対1照合
- 1対多照合

がある．1対1照合は，ユーザがシステムに顔画像とユーザ名などの個人情報を提供し，顔画像とその情報に対応したモデルとが同一人物であることを判定する．これは，顔検証 (face verification) とも呼ばれる．一方，1対多照合は，ユーザがシステムに提供するものは顔画像のみで，それと全登録ユーザのモデルとを照合し，認証する．これは顔識別 (face identification) とも呼ばれる．1対多照合は，1対1照合よりも認証精度を高めることが困難であるが，高いユーザビリティを持つ．

これらの顔認証タスクは，一般的には以下の処理を順に行う．

1. 顔検出 (顔位置推定)
2. 顔特徴抽出
3. 個人識別

まず，顔検出では，1人のユーザの顔位置を正確に推定する．画像中に存在する複数の顔を検出し，それらの領域を大まかに特定する一般的な顔検出 [Yang 2002] とは異なる．次に，顔特徴抽出では，個人識別に用いる画像特徴量を抽出する．また，顔の表情と向きの変化に伴う顔の見え方の変化を正規化するために，顔器官の位置を抽出する場合もある．ただし，顔認証の場合は，顔の表情と向きを容易にほぼ統制することができるため，それらの大きな変化に対応する必要はない．そして，個人識別では，顔画像とモデルの類似度を求め，閾値処理によってユーザとそのモデルが同一人物かどうかを判定する．

顔認証において解決が要求される技術課題としては，

- 顔の位置，大きさの変化への対応

- 顔の表情，向きの変化への対応
- 撮影時の照明，背景の変化への対応
- 経年変化への対応
- 写真詐称への対応
- 実時間認証

が挙げられる。顔の位置と大きさも表情と向きと同様に統制することは可能であるが，認証のつど，ユーザが立ち位置を合わせることは一般的に難しいタスクとされている [土橋 2002]。照明変化は，顔認証システムを屋外で使用する場合や，モバイル端末に搭載する場合を想定して考慮しなければならない。また，背景にユーザ以外の顔や顔に類似する物体が写りこむ状況は多々考えられる。経年変化に関しては，定期的に登録画像を変更することによって対応することができる。ゆえに，顔認証アルゴリズムを強化することで解決する必要はない。そして，顔は偽造される危険性が低いが，顔写真によるなりすまし行為を考慮する必要がある。つまり，人物と写真を判別する技術が求められる。それから，顔認証の利点である高いユーザビリティを低下させないために，実時間認証を実現しなければならない。

1.2 本研究の位置づけ

従来の顔認証セキュリティシステムに導入されている技術は，顔認証と顔認識の問題設定を明確に分け隔てることなく提案されたものである。この問題設定の違いは，従来のほとんどの顔画像照合研究で意識されていない。本研究では，顔認証にフォーカスを合わせ，それに適した顔画像照合手法を提案する。特に，1対多照合の高精度化を目指し，その新しい方法論を展開する。そして，将来的にそれが確立されることを目標とする。この手法の詳細については第3章で述べ，その有効性を検証する。また，従来システムには，顔写真による詐称に対応できる技術は導入されていない。そこで本研究では，人物写真判別技術も開発する。本論文では，第4章においてこの技術を提案する。そして，本研究の提案技術を集結し，それらに基づく顔認証セキュリティシステム FACELOCK の構築も行う。第5章でその詳細を述べ，実環境での FACELOCK の運用性の検証も行う。

第2章

関連研究と提案技術

1970年代にコンピュータビジョン分野で初めて顔画像が扱われてから、今日までに数多くの顔画像処理研究が行われてきた。いまや、顔画像処理研究はコンピュータビジョン分野の中心課題の1つである。本章では、本研究で提案する顔画像照合手法と人物写真判別技術に関連する従来研究について取り上げる。また、それらの提案技術についても簡単に述べる。

2.1 顔画像照合

顔認証セキュリティシステムは主に、顔検出、顔特徴抽出、そして個人識別から構成される。この構成の中で、最も重要とされるのが顔検出である。なぜなら、その検出精度が個人識別に影響を及ぼすからである。従来の顔検出手法は、以下のように大別される。

- 顔特徴に基づく手法 (feature based method)
顔の局所的もしくは大域的な画像特徴量を抽出し、顔検出に利用する。また、顔器官を抽出し、その位置関係を利用する手法もこれに含まれる。
- アピアランスに基づく手法 (appearance based method)
顔領域全体の画像輝度値を1つの特徴量とする。一般的には、矩形窓を画像上で走査することによって顔を検出する。

顔特徴に基づく手法で使用される特徴量は以下のようなものが代表的である。

- 肌色
RGB[Jebara 1997], HSV[Kjeldsen 1996], YCrCb[Chai 1998], YIQ [Dai 1996], CIE XYZ[呉 1997] などの様々な色空間で肌色のモデル化が試みられている。環境中に肌色に似た色を持つ物体が存在しうることで、人種、化粧、日焼けによって色の変動することがモデル化を困難にしている。
- エッジ
顔輪郭や眉毛、目、口に現れるエッジを利用する。顔輪郭を楕円でモデル化したり [Wang 2000]、顔器官の幾何構造の知識を利用する [Yow 1997] 必要がある。形状の変化に敏感に反応する特徴量である。
- 周波数成分
フーリエ変換 [Terrillon 1998] やウェーブレット変換 [Wang 2001] によって抽出される周波数成分を利用する。円形フィルタリング [福井 1997] の応答もこの一種である。大きさの変化により特徴量の変動する。

また、それぞれの手法に適用される代表的な識別器を以下に挙げる。ここでいう識別器とは、入力データが顔であるか顔以外であるかを判別する手法のことである。それらは部分空間を利用する識別器と確率を利用する識別器に大別される。

- 部分空間法
入力データを元の特徴量空間からその部分空間へ投影し、部分空間内で入力データが顔クラスと顔以外クラスのどちらに含まれるかを求める。従来研究で用いられた部分空間法は次のようなものがある。
 - 主成分分析 (PCA)[Turk 1991]
 - 線形判別分析 (LDA)[Yang 2000]
 - 独立成分分析 (ICA)[Antonini 2003]
 - サポートベクタマシン (SVM)[Osuna 1997, Y. Li 2004]
- 確率的手法
モデルの顔画像から抽出された特徴量の確率分布に従い、入力データが顔である確率を求める。従来研究で用いられた確率的手法を以下に挙げる。

- ガウス分布 [Sung 1998, McKenna 1999]
- ベイズ識別 [Schneiderman 1998, Liu 2003]
- ニューラルネットワーク (NN)[Rowley 1996, Féraud 2001]
- 隠れマルコフモデル (HMM)[Samaria 1994, Meng 2000]

顔検出手法は一般に、画像中に存在する複数の顔を検出し、それらの領域を大まかに特定することを目的に提案される [Yang 2002]. しかし、顔認証に要求される顔検出タスクは、1人のユーザの顔位置を正確に推定することである。従来の個人識別手法は、顔位置が正確に推定されるという仮定のもとで提案されたものが多い。代表的な手法に、PCA[Penev 1996], LDA[Belhumeur 1997, Chen 2000], ICA[Yuen 2002], SVM[Guo 2001], ベイズ識別[Z. Li 2004], NN[Ranganath 1997], HMM[Othman 2003]を用いたものがある。つまり、これらの個人識別手法の前処理に従来の顔検出手法を適用しても、高い個人識別精度が保たれる保証はない。

従来研究では、顔検出から個人識別までの一連の処理を行う手法も提案されている [Lanitis 1997, Moghaddam 1997, Wiskott 1999]. その代表が Elastic Bunch Graph Matching[Wiskott 1999]で、現在最も高い認識性能を示す [Phillips 2000, Ayinde 2002]. この手法は顔グラフ照合に基づく顔認識手法である。顔グラフ照合は、顔器官の配置を考慮した特徴ベースの手法といえる。顔グラフのノードは、顔器官の特徴点に配置され、その位置のガボール特徴量 [Daugman 1985, Wiskott 1999, 中野 1999]を持つ。Elastic Bunch Graph Matchingの問題点は、平均モデルグラフを利用したラスタスキャンによって顔位置推定を行うことにある。つまり、平均モデルと顔画像の類似度に差があるほど位置推定の精度が低下する。一般的に、アピアランスベースの手法や顔器官の配置を考慮した手法で顔位置を推定する場合、顔の見え方が正確に推定されなければ、位置も正確に推定されない。また、この逆も成り立つ。この依存性は、あらゆる見え方のモデルを用いてラスタスキャンすることによって解決できる。しかし、多数のモデルを照合に用いれば、計算コストが増大する。Krügerらもこれを問題視している [Krüger 1997]. セキュリティシステムはリアルタイム性が要求されるため、この推定精度と計算コスト(処理時間)のトレードオフを解決しなければならない。

本研究では、顔の見え方として主に大きさと個人性を扱う。これらの変化は、セキュリティシステムの運用時に最も頻繁に起こりうる。Kotropoulosらは、Elastic Bunch Graph Matchingが顔の大きさの変化に弱いことを

実証している [Kotropoulos 2000]. また, この手法は顔の特徴点の検出時に個人性などのモデルを推定することで, その検出精度と個人識別精度を向上させている. しかし, 上述のように顔位置推定の精度がボトルネックになっている [Phillips 2000]. 本研究では, 顔位置の推定時に個人モデルを推定することが有効であると考ええる.

そこで本研究では, 顔の位置と大きさと個人モデルの推定を統合した手法を提案する. 位置の推定には, 粗密法に基づく計算効率の良い探索手法を提案し, 大きさの推定には, スケール変換という高効率な統計的手法を提案する. この2つの手法を交互に繰り返すことで, 双方の推定精度を高めていく. さらに, この処理に個人モデルの推定を加える. この推定が常に全ての個人モデルから行われることは非効率である. Stengerらは手の形状推定という課題に, クラスタリングによって階層化された形状モデルデータベースを用いることでモデル推定の効率を向上させた [Stenger 2003]. 本研究では, 位置と大きさの推定が進むにつれ, 階層化された個人モデルデータベースの構造に従って最適なモデルを選択していく手法を提案する. 最終的に選択された個人モデルは個人識別結果とみなすことができる.

2.2 顔の大きさ変化に対する特徴量の正規化

顔の大きさの変化に対応するためには, 画像の空間周波数の変動に対応した特徴量の抽出も必要である. 従来研究 [Lades 1993, Konen 1994, Wiskott 1995, Wiskott 1997, Krüger 1997, Würtz 1997, Pramadihanto 1998, Lyons 2000, Pramadihanto 2001, プラマディハント 1997] では, あらかじめ多段階の大きさのガボールウェーブレットを用いてガボール特徴量を抽出することにより, 顔の大きさの変化に対応させることが主であった. しかし, この対応は最適な形状のガボールウェーブレットを利用していないため, 顔認識の精度の低下を招いている [Kotropoulos 2000]. これに対し, Ladesらは推定した顔の大きさを利用して, 多段階の特徴量からの補間によって, 顔の大きさに最適な特徴量を求める手法を提案している [Lades 1995]. しかし, 補間の精度が高くないので, これに代わる手法として, 画像の大きさを正規化する手法が提案されている [Okada 1998]. この手法の問題点としては, 正規化に伴う画像の情報落ちが顔認識の精度の維持を妨げることである. また, 画像を正規化するための処理時間が必要になる. 本研究では, スケール変換で推定した顔の大きさの情報を

利用してガボールウェーブレットの形状を変更し、顔の大きさの変化に対して正規化した特徴量を抽出する手法を提案する。この手法は、画像を正規化するよりも効率良く正規化した特徴量を抽出できる。さらに、この手法により、ガボールウェーブレットの大きさの段階を削減しても顔認証の精度が維持されると考えられる。この削減が顔認証手法の計算効率の面で非常に有効と思われる。また、この正規化の処理を効率良く行うために、一般的なガボールウェーブレットの生成式を改良した式を提案する。

2.3 人物写真判別

従来の顔認証セキュリティシステムは顔写真による詐称に未対応という問題を抱えている。そのため、人物と顔写真を判別する技術が求められ、近年になって研究され始めた。

人間の顔には凹凸が存在する。それに対して、写真は平坦である。ゆえに、認証物体の3次元形状を観測すれば、人物と顔写真を判別できる。この観測を行う手法の1つとして、複数の照明を切り換えて撮影した複数枚の画像を用いる照度差ステレオ法[新本 1999]がある。津田らはこの方法を応用し、赤外光照度差画像を用いた人物写真判別手法[津田 2004]を提案した。それは左右の赤外光照射画像から認証物体の面の方向を求め、人物か顔写真かを判別する手法である。可視光照明で照度差ステレオ法を適応することも可能であるが、人間が可視光に反応するため、照度差を得るために十分に強い光量を利用するとユーザのストレスが大きくなる。また、ユーザに左右の照明の切り換えを感知されることによって、顔認証システムの構成が見破られ、詐称される可能性もある。これらの問題を解決するために、可視光照明のかわりに赤外光照明を用いている。津田らの手法の問題点としては、詐称者が写真を湾曲させた顔マスクを装着するような状況を想定していないことが挙げられる。また、顔全体の面の方向を詳細に推定しているため、計算効率が低いことも挙げられる。

2.4 赤外光を利用した顔検出

赤外光照明を用いることは、顔検出に対しても有効である。従来の顔検出手法 [Rowley 1996, Viola 2001] や、顔検出の精度と計算効率を向上させるために導入される背景差分法 [Mitsuyoshi 2003] は照明変化に弱い。従来、この欠点を補うために様々なモデルが用意されてきたが、多数のモデルを持つことは処理時間が増大するという問題を生じる。つまり、顔検出の精度と計算効率の間にはトレードオフが生じる。屋内での顔認証システムの運用に限れば、赤外光照明を利用することで、このトレードオフを解決することができる。なぜなら、屋内には強い赤外光が存在しないため、認証物体まで赤外光が照射されるように光量を調整すれば、認証物体と背景のコントラストを大きくすることができ、効率的な顔検出が可能となる。また、背景付近の物体が移動しても、画像の変化はほとんど生じないため、一般的な背景差分法よりも赤外光照明の利用は有効である。

赤外光照明を用いた顔検出手法がこれまでにいくつか提案されてきた。Dowdallらは、低域波長と高域波長の2種類の近赤外光を用いて撮影した画像から肌領域を検出し、ついで眉と目を検出する手法を提案した [Dowdall 2003]。肌、眉、目の検出には、Pavlidisらによって分析された近赤外光に対する物質の反射特性 [Pavlidis 2000] を利用している。この手法の問題点は、動画像のフレーム間差分を利用するため、システムが顔認証用の正面顔を獲得するためには、ユーザがカメラに正面を向いて平行移動しなければならないという煩わしさが挙げられる。また、遠赤外光を利用した顔検出手法も提案されている [Eveland 2003]。遠赤外画像は物質の熱放射光を計測したものであるため、詳細な顔のテクスチャを得ることができない。ゆえに、顔を検出した遠赤外画像を用いて、顔認証を行うのは困難である。なお、先に述べた津田らの研究では、近赤外光が利用されている。近赤外画像には、遠赤外画像より詳細に顔のテクスチャが現れるため、これを顔認証に用いることができる可能性がある。

本研究では、上述の近赤外光を利用した顔検出手法と人物写真判別法が持つ問題点の解決を目的とし、より安定に写真詐称に対応し、なおかつ計算効率の良い手法を提案する。

第3章

顔の位置と大きさと個人性を考慮した顔照合システム

本章では，顔の位置と大きさと個人モデルの推定を統合した手法を提案する．従来，顔の位置と大きさを同時に推定する手法は提案されている [Kato 2002, Liu 2002] が，顔位置推定と個人識別を同時に行う手法は存在しないため，本手法は顔画像照合研究の新たな方法論に位置づけられる．

本章では，3.1 で顔グラフによる顔の表現方法について述べ，3.2 では階層構造を持つ個人モデルデータベースの生成手法を説明する．3.3 で，粗密法を用いた顔位置推定，大きさ推定のためのスケール変換，個人モデル推定についてと，それらを統合したアルゴリズムを説明する．そして，このアルゴリズムを実装し，3.4 で標準顔画像データベースを用いた実験を示し，その有効性を検証する．

3.1 顔グラフによる顔の表現

本研究では，顔位置推定のための特徴量表現と個人識別のための特徴量表現を兼ね備えた顔グラフによって顔を表現する．図 3.1 に顔グラフの例を示す．顔グラフのノードは，目，唇，鼻といった顔の 30 個の特徴点上に配置する．図 3.1 中の灰色のノードは顔の輪郭に対応したノード，黒色のノードは輪郭以外の特徴点に対応したノードを表す．各ノードは対応する特徴点の位置におけるガボール特徴量を保持する．この特徴量はガボールウェーブレット変換によって求められる．本研究では，ガボール特徴量を顔の大きさに対して高効率に正規化できるガボールウェーブ

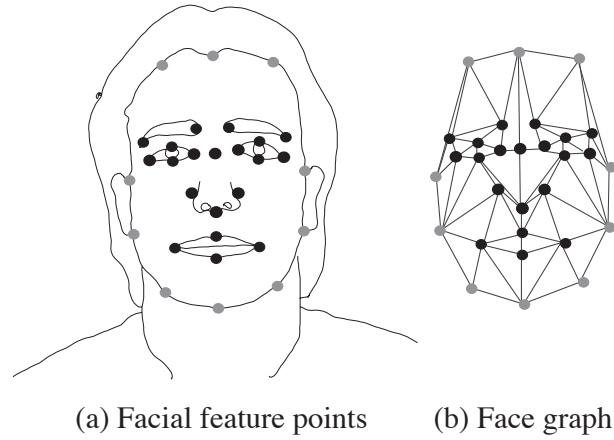


図 3.1: 顔の表現

レット変換を提案する.

3.1.1 ガボール特徴量

一般的なガボールウェーブレット変換は、式 (3.1) で定義されるガボールウェーブレットと画像の畳み込みによって求められ、画像を周波数領域での表現に変換するものである.

$$\begin{aligned}
 \psi_{j,\theta} &= \frac{1}{4\pi a^{2j}\sigma} \exp\left(-\frac{x^2 + y^2}{(2a^j\sigma)^2}\right) \cdot \\
 &\quad \left[\exp\left(\frac{i u (x \cos \theta + y \sin \theta)}{a^j}\right) - \exp(-(u\sigma)^2) \right], \\
 j &= 1, \dots, N_j, \\
 \theta &= \frac{\kappa\pi}{N_\theta} \text{ with } \kappa = 0, \dots, N_\theta - 1
 \end{aligned} \tag{3.1}$$

ガボールウェーブレットは、その特性が大きさや周波数のファクタ j と方向ファクタ θ によって制御されるため、画像の局所的な濃度変化を検出するフィルタとしてよく用いられる。画像の特徴を十分に解析するためには、異なった j と θ を持ったウェーブレットを用意すれば良い。なお、式 (3.1) 中の a と σ はウェーブレットの基底サイズ、 u は基底周波数、 N_j は大きさの段階数、 N_θ は方向数を表す。ガボールウェーブレットの特性としては、式 (3.1) の第2項中の減算により、関数の直流成分を補正してい

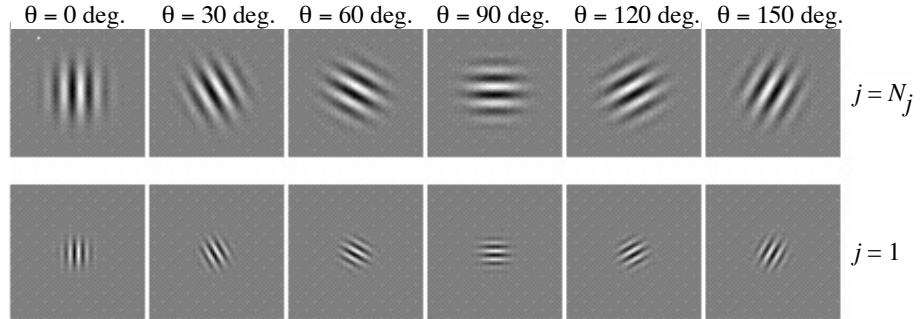


図 3.2: ガボールウェーブレットの形状

るため、均一な照明変動に対してロバストになることも挙げられる。また、この変換による応答は複素数となるが、その振幅成分は畳み込みの位置の変化に沿ってゆっくりと変化し、位相成分は激しく変化するため、一般的には振幅成分のみを用いることが多い。図 3.2 にウェーブレットの形状の例を示す。

本研究では、顔の大きさの変化に対応した特徴量を容易に抽出するために、上述のガボールウェーブレット変換を式 (3.2) のように改良する。

$$\psi_{j\theta} = \frac{1}{4\pi a^2 j^2 \sigma} \exp\left(-\frac{x^2 + y^2}{(2aj\sigma)^2}\right) \cdot \left[\exp\left(\frac{i u (x \cos \theta + y \sin \theta)}{aj}\right) - \exp(-(u\sigma)^2) \right] \quad (3.2)$$

式 (3.1) と式 (3.2) における a と j の関係に注目すると、式 (3.1) で定義されるガボールウェーブレットは、 j によって大きさと周波数が非線形に制御される。このため、顔の大きさの変化に対して自在にガボールウェーブレットの大きさと周波数を操作するには、 j の設定が複雑となる。一方、本研究で提案するものは、その関係が線形である。ゆえに、参照モデルグラフに対する入力画像の顔の大きさの比率 s_{opt} を j にかけて ($j \leftarrow js_{opt}$)、改良したウェーブレットを拡大縮小することで、顔の大きさの変化による空間周波数の変動を吸収したガボール特徴量が容易に求められる。つまり、顔の大きさの変化と正比例してウェーブレットの大きさを変えることが容易にできる。参照モデルグラフとは、顔の位置と大きさと個人モデルの推定に利用する顔グラフのことである。 a と j の関係が非線形で

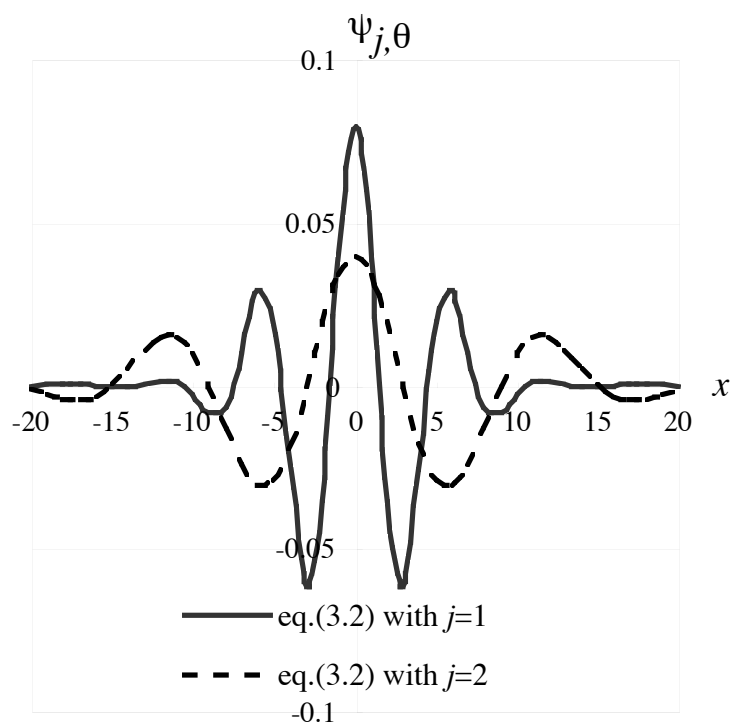


図 3.3: 改良したガボールウェーブレットの形状

ある式 (3.1) においても, $j \leftarrow \log(ajs_{opt})/\log a$ とすることで, s_{opt} の変化と正比例してウェーブレットの大きさを変えることができる. しかし, これは式 (3.2) に比べ, 約 20 倍の計算コスト (計算機による時間計測の結果) を必要とする. $j = 1, 2$ に対応する式 (3.2) を 2 次元表示したものを図 3.3 に示す. この図から, $j = 2$ のウェーブレットのピークは, $j = 1$ に対して x 軸方向に 2 倍された位置に現れていることが確認できる. つまり, 改良したガボールウェーブレットは j に正比例してウェーブレットの大きさが変化することが確認できる.

3.1.2 グラフ特徴量類似度

次に、顔グラフと画像の照合に用いるグラフ特徴量類似度について述べる。本研究では、この特徴量を式 (3.3) に示す正規化相関によって求める。

$$R(G(x, y, s_{opt})) = \frac{1}{30} \cdot \frac{\sum_{m=1}^{30} V_m W_m(G(x, y, s_{opt}))}{\sqrt{\sum_{m=1}^{30} V_m^2 \sum_{m=1}^{30} W_m(G(x, y, s_{opt}))^2}} \quad (3.3)$$

ここで、 V_m は参照モデルグラフのノード m が持つガボール特徴量、 W_m はシャドウグラフ G のそれである。シャドウグラフとは、参照モデルグラフから生成され、推定した位置 (x, y) と大きさ (s_{opt}) を持つグラフである。

3.2 個人モデルデータベースの階層的クラスタリング

本研究では、個人モデル推定の効率化を図るために、個人モデルデータベースを階層的クラスタリング [Yager 2000, Duda 2001] によって階層構造化する。階層化されたモデルデータベースは図 3.4 のような構造となる。階層的クラスタリングは、ボトムアップ手法とトップダウン手法がある。ボトムアップ手法は、 k 個の 1 要素クラスタから連続的にクラスタを融合し、各階層でのクラスタを構成していく。一方、トップダウン手法は、 k 個全ての要素を含む 1 個のクラスタから連続的にクラスタを分割していく。ボトムアップ手法で全ての要素を完全に融合するまでクラスタリングを続けた場合と、トップダウン手法で全ての要素を完全に分割するまでクラスタリングを続けた場合に構成される階層的クラスタは、融合と分割の評価が同じであれば同等となる。 k 個のクラスタを l ($l \neq 1$) 個のクラスタへと段階的に融合する場合にはボトムアップ手法が、1 個のクラスタを l ($l \neq k$) 個のクラスタに段階的に分割する場合にはトップダウン手法が適用される。つまり、クラスタリングを途中で打ち切る場合には適用する手法が異なる。本研究では、最上位階層が 1 個、最下位階層が k 個のクラスタから構成される階層構造を必要とするので、どちらの手法も適用することができるが、アルゴリズムがより簡単なボトムアップ手法を適用することにする。

次に、クラスタの融合の評価方法について述べる。まず、評価に用いる尺度についてである。一般的には、特徴空間のユークリッド距離を用いるが、本研究では顔グラフと画像の照合に用いるグラフ特徴量類似度

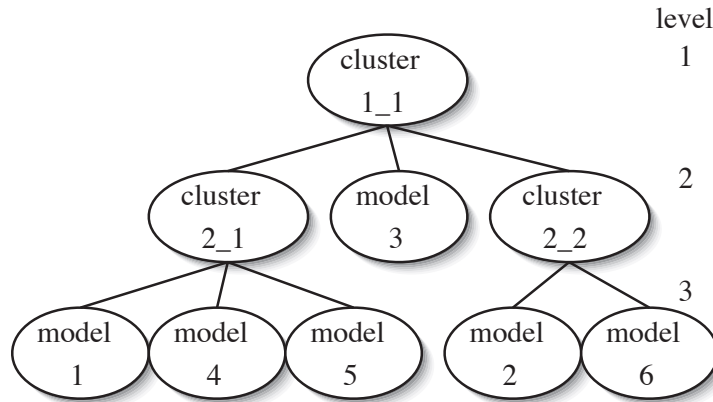


図 3.4: 階層モデルデータベースの例

R とする。クラスタリングの場合には、式 (3.3) の V_m と W_m はそれぞれ個人モデルグラフが持つガボール特徴量である (この場合、 x , y , s_{opt} は既知データである)。評価のアルゴリズムは、最近傍アルゴリズムと最遠近傍アルゴリズムが代表的である。最近傍アルゴリズムは、クラスタ間の最大相関値がある閾値 R_{th} 以上であればクラスタを融合する。これは単連結アルゴリズムとも呼ばれ、クラスタ内の要素が互いに必ずしも類似しない細長いクラスタを構成する可能性がある。一方、最遠近傍アルゴリズムは、クラスタ間の最小相関値がある閾値 R_{th} 以上であればクラスタを融合する。これは完全結合アルゴリズムとも呼ばれ、構成されたクラスタは、クラスタ内の要素が互いに閾値 R_{th} 以上の類似度を持つことになる。本研究では、クラスタ内の要素が互いに類似していることが望ましいので、最遠近傍アルゴリズムを利用してクラスタを構成する。階層構造は、ボトムアップ手法を利用する場合、閾値 R_{th} を大きな値から徐々に小さくすることで構成できる。本研究では、この R_{th} を ΔR_{th} ずつ減少させる。階層モデルデータベースの構成は、この ΔR_{th} に依存する。

個人モデルの推定は、このように構成された階層構造を上位階層からたどることで行う。そのためには、各クラスタの代表データを定義しなければならない。本研究では、クラスタ内の要素が互いに類似していることから、代表データをクラスタ中心とする。これは、個人モデルグラフが持つガボール特徴量の平均と、ノード間距離の平均から求められ、これをクラスタ平均モデルグラフと呼ぶことにする。このグラフが参照モデルグラフとなる。



図 3.5: 顔画像

3.3 顔位置と大きさの推定と個人識別の統合アルゴリズム

探索問題の解決法を決定する上で、探索空間の特徴量の分布を調べることは重要である。本研究の顔位置推定の場合、その分布は画像と参照モデルグラフのグラフ特徴量類似度 R から求まる。図 3.5 の顔画像と平均モデルグラフおよび個人モデルグラフをラスタスキャンで照合して求めたグラフ特徴量類似度の分布を図 3.6 に示す。この図から分布が完全な単峰性ではないことがわかる。ゆえに、多峰性の分布に対して探索結果が局所解に陥る可能性がある降下法を用いることができない。そこで、本研究では常に複数の解候補を推定しながら大局解を求める手法を提案する。

提案手法の概略を図 3.7 に示す。本手法は、顔位置推定、顔の大きさ推定、個人モデル推定、推定データの検証の 4 つの状態から構成される。推定したデータはシャドウグラフという形で優先度付き待ち行列に格納する。そして、この待ち行列の上位に存在するシャドウグラフの位置と大きさを更新し、そのグラフが参照するクラスタ平均モデルグラフから階層構造をたどって、最適なモデルを選択していくという処理を繰り返す。複数の解候補を推定するために、この処理にビーム探索 [Barr 1981] を導入する。検証状態は、シャドウグラフの妥当性をグラフ特徴量類似度によって検証する。3 つの推定状態のそれぞれについてと、各状態間の遷移について 3.3.1 より詳しく順に述べる。

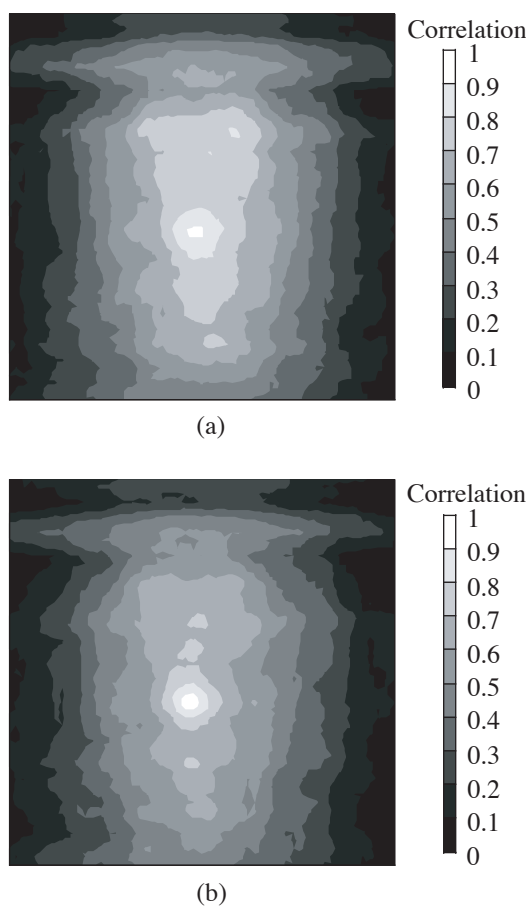


図 3.6: 顔画像 (図 3.5) と (a) 平均モデルグラフ, (b) 個人モデルグラフの相関値の分布

3.3.1 粗密法による顔位置推定

顔位置推定への1回目の遷移が起きた時には, まず, $d_{img} \times d_{img}$ 画素の画像全体に, d_{smp} 画素間隔で矩形格子状に $n_{smp} \times n_{smp}$ ($n_{smp} = d_{img}/d_{smp}$) 個のサンプリング点を配置し, それらの点上からガボール特徴量を抽出する. そして, 参照モデルグラフの各ノード m に対して次の特徴量類似度 ϕ_{mn} を計算する.

$$\begin{aligned} \phi_{mn} &= 1 - \|V_m - D_n(s_{opt})\|, & (3.4) \\ m &= 1, \dots, 20, \\ n &= 1, \dots, n_{smp}^2 \end{aligned}$$

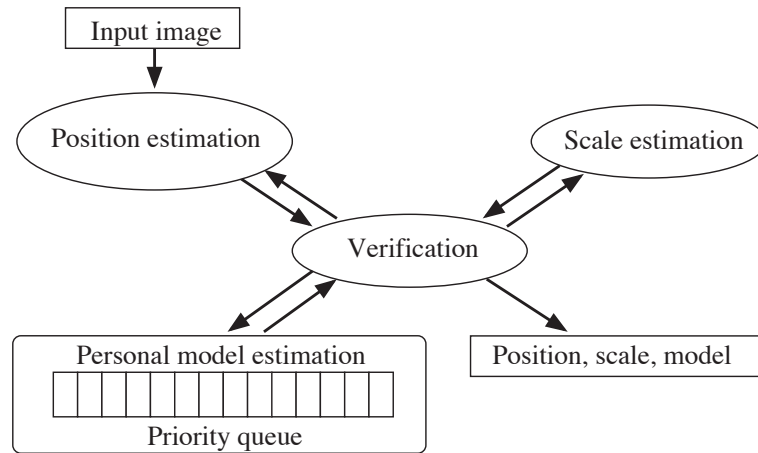


図 3.7: 提案手法の状態遷移

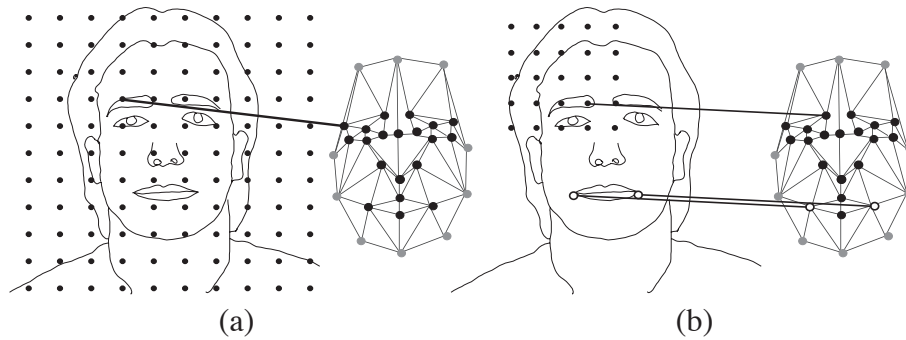


図 3.8: 顔位置推定

これを局所特徴量類似度と呼ぶことにする．ここで，この参照モデルグラフは最上位階層のクラスタ平均モデルグラフ，つまり，全ての個人モデルグラフの平均グラフ G_{ave} である．また， $D_n(s_{opt})$ はサンプリング点 n のガボール特徴量を，推定した顔の大きさ s_{opt} によって正規化したものを表す．顔の大きさ推定については 3.3.2 で示す．なお，1 回目の遷移では $s_{opt} = 1$ である．位置の推定は，ノード m に対して最も局所特徴量類似度が高いサンプリング点 n を探索し (図 3.8 (a))，ノード m がその点 n と一致するように参照モデルグラフを配置することによって行う．この一連の処理を，参照モデルグラフの顔輪郭部以外のノードに対して行い，20 個の位置を推定する．

2回目以降の遷移が起きた場合には、以下のように位置を推定する。まず、検証状態から遷移してきたシャドウグラフが持つ推定位置のまわりに、 $d_{smp} \times s_{opt}$ 画素間隔で矩形格子状に $n_{smp} \times n_{smp}$ ($n_{smp} = d_{face}/d_{smp} + 1$) 個のサンプリング点を配置する。ここで、 d_{face} を平均グラフ $G_{ave}(s_{opt} = 1)$ の横方向の長さとする。なお、状態遷移の過程でサンプリング点が同じ位置に配置されることを防ぐために、 d_{smp} を単調減少させる。次に、シャドウグラフが参照するクラスタ平均モデルグラフのノード m に対して、局所特徴量類似度 ϕ_{mn} を計算する。そして、最も高い局所特徴量類似度を持つ位置へ参照モデルグラフを移動させる。それから、図 3.8 (b) のように、そのグラフの顔輪郭部およびノード m 以外から n_w (< 20) 個のガボール特徴量 (W_q) を抽出する。図中の白色ノードは n_w 個のガボール特徴量を抽出するためのサンプリング点の例である。そして、次の式で表される局所特徴量類似度の合計 ψ_m を計算する。

$$\psi_m = \max_n \phi_{mn} + \sum_{q=1}^{n_w} [1 - \|V_q - W_q(s_{opt})\|] \quad (3.5)$$

この一連の処理を、参照モデルグラフの顔輪郭部以外のノードに適用し、最も高い ψ_m を持つグラフの位置を推定結果とする。

3.3.2 大きさ推定のためのスケール変換

顔の大きさを高効率に推定し、シャドウグラフを適切な大きさに伸縮させるために、スケール変換という手法を提案する。スケール変換は、スケール辞書というデータベースを利用する。スケール辞書はあらかじめシステム構築時に作成しておく。

まず、スケール辞書の生成について説明する。顔の大きさが平均グラフ G_{ave} とほぼ等しい N_t 枚の教師顔画像を用意する。次に、平均グラフを伸縮率 s_T で T 段階に伸縮させたグラフを、各教師画像の顔位置 (x_{true}, y_{true}) に配置する。そして、このグラフの妥当性をグラフ特徴量類似度 R によって求める。つまり、式 (3.3) の V_m を平均グラフのガボール特徴量、 W_m を伸縮グラフのそれとして求める。このグラフ特徴量類似度をスケール辞書と呼ばれるデータベースに学習データとして登録する。

入力画像の顔の大きさの推定は、この辞書を用いて以下のように統計的に行う。まず、検証状態から遷移してきたシャドウグラフ G のグラフ特徴量類似度をスケール辞書の学習データと比較する。比較は教師画像 t

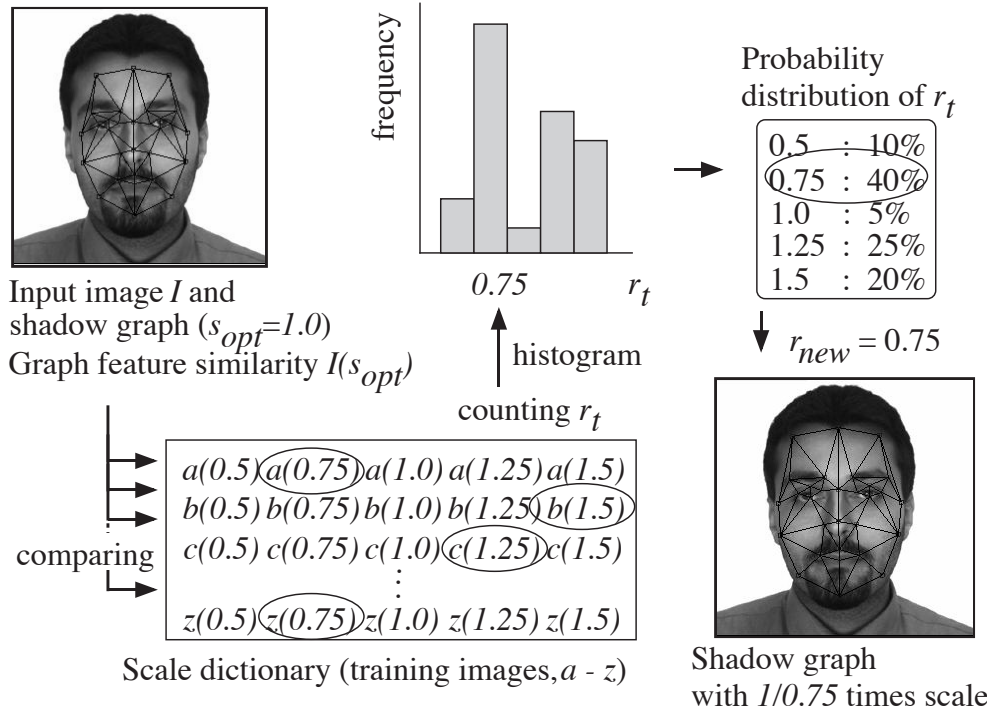


図 3.9: スケール変換

ごとに行い，最も誤差の小さい学習データをそれぞれ検出し，そのデータに対応した伸縮グラフの伸縮率 r_t を導き出す。

$$r_t = \arg \min_{s_T} [R(G(x, y, s_{opt})) - R(G_{ave}(x_{true}, y_{true}, s_T))] \quad (3.6)$$

for each training image,

$$t = 1, \dots, N_t, \quad T = 1, \dots, T$$

ここで，式 (3.6) の $R(G_{ave}(x_{true}, y_{true}, s_T))$ はスケール辞書の学習データである．次に，全ての教師画像から求めた r_t のヒストグラムを生成する．このヒストグラムから求まる r_t の確率分布に従い，入力画像に対するグラフの大きさの比率 r_{new} を推定する．シャドウグラフの適切な大きさへの伸縮は， r_{new} によって更新された伸縮率 s_{opt} を求めることで行う ($s_{opt} \leftarrow s_{opt} \times s_{new}$ ($s_{new} = 1/r_{new}$)).

この推定をより精度良く行うには，以下の処理が考えられる．まず，高い確率を持つ r_t を e 個選び， r_{new} の候補 r_e とする．次に，それらによって更新された伸縮率 (s_{opt}/r_e) で伸縮させたシャドウグラフと参照モデル

グラフのグラフ特徴量類似度を求める。そして、グラフ特徴量類似度が最も大きくなった r_e を r_{new} として採用する。 e の設定については3.4.3で述べる。図3.9にスケール変換の処理の流れを示す。この図は、参照モデルグラフの大きさを1として、入力画像の大きさが $1/0.75$ である場合のスケール変換を表している。

3.3.3 個人モデル推定

個人モデルの推定は、検証状態から遷移してきたシャドウグラフ G がどのクラスタに属するかを判定する。シャドウグラフが参照するクラスタ平均モデルグラフが、レベル L_v の階層のあるクラスタの代表とする。まず、このクラスタに属する下位階層 ($L_v + 1$) の全てのクラスタ平均モデルグラフとシャドウグラフの照合を行う。照合にはグラフ特徴量類似度 R を用いる。つまり、式(3.3)の V_m をクラスタ平均モデルグラフのガボール特徴量とする。次に、照合後の全てのシャドウグラフを、そのグラフ特徴量類似度を優先度として優先度付き待ち行列に格納する。この際、シャドウグラフの参照モデルグラフを、照合に用いたそれぞれのクラスタ平均モデルグラフに変更する。そして、優先度付き待ち行列の上位に存在するシャドウグラフの参照モデルグラフがモデル推定の結果となる。

3.3.4 状態遷移

前節までに示したそれぞれの処理の遷移について述べる。本手法は、顔位置推定と個人モデル推定を検証状態を介して交互に行う。顔の大きさ推定への遷移は、顔位置推定から検証へ遷移した時に求めるグラフ特徴量類似度 R がスケール辞書の学習データの範囲内となった時に行う。状態遷移の過程で、優先度付き待ち行列に残されたシャドウグラフを除去しないので、推定データの妥当性が低下すれば、同階層か上位階層の参照モデルグラフを用いた位置と大きさの推定が繰り返される。なお、終了状態への遷移は、モデルの推定が階層データベースの最下位層まで進んだ時に起こる。その時のシャドウグラフの参照モデルグラフが個人モデルグラフとなっているので、それを個人識別の結果とみなす。状態遷移の例は3.4.5で示す。

本手法では、推定データが局所解に陥ることを防止し、推定精度を向上させるために、ビーム探索を状態遷移に組み込む。本研究では、優先

度付き待ち行列を利用してこれを実現する。つまり、この待ち行列から優先度の高いシャドウグラフを c 個選択し、それぞれからさらなる推定を行う。本手法では、同じ推定データを持つシャドウグラフが偶然に生成される可能性があるが、これらの重複するデータは c 個を選択する時に削除する。

3.4 標準顔画像データベースによる評価実験

3.4.1 実験環境

提案手法の有効性を検証するために実験システムを構築し、評価実験を行った。性能比較に用いる従来システムは、現在最も高い顔認識性能を持つとされる Elastic Bunch Graph Matching と、著者が従来提案した顔の大きさの変化にロバストな個人識別システム [平山 2003] である。本論文では、前者を従来システム 1、後者を従来システム 2 と呼ぶ。従来システム 1 は、Beveridge らによる公開プログラム [Beveridge] と文献 [Wiskott 1999] に基づいて構築した。従来システム 2 は、平均グラフを利用したラスタスキャンによって顔位置を推定した後、スケール変換によって顔の大きさを推定し、グラフが持つ特徴量の正規化を行い、フレキシブル特徴照合法 [Pramadihanto 2001] によって顔特徴点を抽出するシステムである。従来システム 2 の詳細は付録 A で示す。それぞれのシステムの実装は SGI 製ワークステーション Origin300 (CPU R14000, 500 MHz, memory 4 GB, OS IRIX6.5.16m) 上で行った。画像は Purdue 大学の AR Face データベース [Martinez 1998] (256 × 256 画素, 8bit 濃淡) の一部を利用した。個人モデルデータベースには、ほぼ等しい大きさの 50 人のモデル画像とモデルグラフを登録した。モデル画像とグラフは 1 人につき 1 対で、モデル画像は全て無表情の正面顔である。ただし、眼鏡をかけた画像や照明変動のある画像も含まれた。モデルグラフの特徴点は著者が目視によって画像中で定義した。テスト画像は、モデルデータベースに登録された人物のモデル画像、登録されていない人物の画像がそれぞれ 50 枚 (テストセット 100%) と、それらを 80% と 120% に伸縮した画像 (テストセット 80%, 120%) 200 枚の計 300 枚とした。全てのテスト画像に複雑な背景は含まれない。本実験は、前処理によって人物領域がおおまかに検出された後の顔位置推定と個人識別を想定している。なお、画像の伸縮は MacOS X 上で Adobe Photoshop7.0 を用いて bi-cubic 内挿法で行った。

本実験で用いた提案システムのパラメータを以下に示す.

- $\sigma = 1, a = 2.5, u = \pi$
:ガボールウェーブレットのパラメータ
- $N_j = 2, N_\theta = 8$
:ガボールウェーブレットの大きさの段階数と方向数
- R_{th} の初期値 = 0.9
:階層データベース構成用の相関値の閾値
- $\Delta R_{th} = 0.02, 0.03, \text{ or } 0.04$
: R_{th} の減少幅
- $d_{img} = 256$
:画像の一辺の画素数
- $d_{face} = 120$
:平均グラフ G_{ave} の横方向の長さ (画素)
- $d_{smp} = 15, 17, 20, 24, 30, \text{ or } 40$
:顔位置推定のサンプリング点の間隔 (画素)
- $n_w = 2, 4, 6, \text{ or } 19$
:顔位置推定の式 (3.5) 用のサンプリング点数
- $N_t = 20$
:スケール辞書の教師顔画像数
- $s_T = 0.1$ 刻みで $0.5 \sim 1.5$ ($T = 1 \sim 11$)
:スケール辞書生成用の顔グラフ伸縮率
- $c = 1, 3, 5, \text{ or } 7$
:ビーム探索のビーム幅

$\Delta R_{th}, d_{smp}, n_w, c$ を上記のように変化させて 288 通りの実験を行った。なお、 ΔR_{th} は使用したデータベースから異なった階層構造が構成されるように設定した。 d_{smp} は、適当な値から位置の推定精度に差が生じなくなる値まで減少させ、また、同様に計算コストに差が生じなくなる値まで増加させた。実際には、 $d_{face}/d_{smp} + 1 (= n_{smp} : \text{サンプリング点数})$ が 9, 8, 7, 6, 5, 4 となる d_{smp} に設定した。また、 n_w に関しては、 $n_w = 2$ の特徴点を左右の目、 $n_w = 4$ を左右の目と口の左右の端点、 $n_w = 6$ を左右の目と鼻および口の左右の端点に設定した。目、鼻、口は顔認識に重要な部位とされ、左右とした理由は顔の対称性を考慮したためである。そして、 c に関しては、ビーム探索を導入しない状態である $c = 1$ から推定精度が向上しなくなる値まで、システム性能を十分に示すことができる間隔で設定した。

3.4.2 より、まず、構成された階層データベースを示し、続いて、スケール変換の性能を検証する。そして、位置と大きさの推定実験と個人識別実験の結果から、設定パラメータに対する提案システムの性能を分析する。さらに、顔位置推定と個人モデル推定の統合の有効性を検証し、最後に提案システムと従来システムの性能比較を行う。

3.4.2 構成された階層モデルデータベース

全ての個人モデルグラフ間のグラフ特徴量類似度 R を求めたところ、最大が 0.92、最小が 0.78 であった。そこで、閾値 R_{th} の初期値と閾値の減少幅 ΔR_{th} を上述のように設定し、3 種類の階層データベースを構成した。それぞれを図 3.10 に示す。図中の楕円と円はクラスタを表す。数字はクラスタ内の要素 (個人モデルグラフ) の数である。 ΔR_{th} を 0.02 として構成した階層データベース (図 3.10 (a)) はレベル数が 8、最大分岐数が 8 である (図中では最下位層を省略している)。また、 ΔR_{th} が 0.03 に対しては、レベル数が 6、最大分岐数が 11 の階層データベース (図 3.10 (b))、0.04 に対しては、レベル数が 5、最大分岐数が 13 の階層データベース (図 3.10 (c)) が構成された。レベル数が大きいほど、顔位置推定と個人識別の計算コストが高くなる。逆に、レベル数が小さいほど、分岐数が平均的に大きくなり、それらの精度が低くなる可能性がある。これらの階層データベースから得る見解は、本実験で用いた個人モデルデータベースが、ガボール特徴量空間でいくつかの大きなクラスタに分割できるような分布ではなく、単峰性の分布をしているということである。

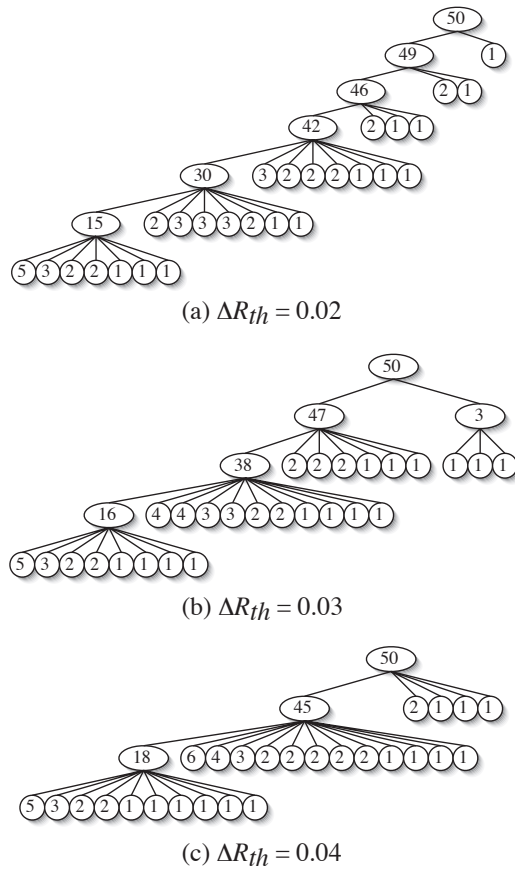


図 3.10: 構成された階層モデルデータベース

3.4.3 スケール変換の性能検証

グラフと入力画像の顔の大きさを合わせるための一般的な方法に対するスケール変換の有効性を、比率 r_{new} の候補 r_e の数 e の設定と絡めて検証する．一般的な方法とは、 T 段階の大きさの参照モデルグラフと入力画像を照合し、最も適合する参照モデルグラフを選ぶことで顔の大きさの変化に対応するというものである．まず、スケール変換の開始から r_{new} の候補となる e 個の r_e を求めるまでの処理時間を t_e 、1つのグラフと1枚の画像からグラフ特徴量類似度 R を求める時間を t_R とすると、スケール変換と一般的な方法にかかる処理時間はそれぞれ、 $t_e + e \times t_R$ と $T \times t_R$ で表される． t_e はスケール辞書の教師データ数にほぼ比例して大きくなるが、本実験の設定 ($N_t = 20, T = 11$) で約 1 ミリ秒と微小である．これ

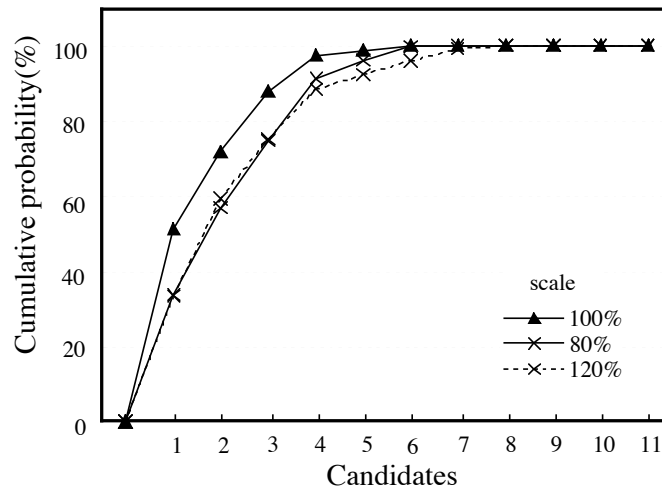


図 3.11: 最適な比率 r_{new} の候補が持つ確率の累積

に対し, t_R はガボールウェーブレット変換に要する時間を含むため, 約 60 ミリ秒と t_e より大きい. すなわち, t_e は t_R に対して無視でき, 両手法の処理時間は e と T に依存することになる. そこで, e についてであるが, 本システムでは図 3.11 に示す統計を利用して設定する. 図 3.11 は入力画像にスケール変換を適用した時に求まる r_i の確率を降順に累積したものである. この図から, 第 5 候補までの累積確率が 90% を越えていることがわかる. また, 5 つの候補の中に r_{new} として最適なものが必ず含まれることが実験的に確認されている. これらのことから, 本実験における s_T と $T (= 11)$ の設定でのスケール変換では, e を 5 ($< T$) とすることで, 一般的な方法と同精度で r_{new} を求めることができる. 同精度と言える理由は, 候補を選択した後の処理が一般的な方法に基づくためである. ゆえに, スケール変換の方が計算時間が短いため有効であると言える.

スケール変換の最も有効な点は, 上述のように t_e が微小であるため, 計算効率がスケール辞書の大きさには依存しないところにある. さらに, あらゆる顔の大きさを確率の形で瞬時に求められる点も有効である. ゆえに, 顔の大きさの推定処理を行う大きさの範囲を限定でき, その処理を効率良く進めることができる. これはより細かな精度で顔の大きさを推定する必要がある場合に, 特に有効であると考えられる. これに対し一般的な方法では, 大きさの推定精度は参照グラフの数 T に依存するので, 計算効率とトレードオフの関係にある.

3.4.4 設定パラメータに対する性能分析

顔の位置と大きさの推定実験

図 3.12 に上述の 288 通りの顔位置推定に関する実験結果を示す。この図は d_{smp} の変化に対する顔位置推定の誤差と計算コストを ROC 曲線で表したものである。また、全曲線に関する包絡線も示す。顔位置推定の誤差は鼻の頭に配置されるノードの推定位置と真値の誤差である。計算コストは処理時間ではなく、ガボールウェーブレット変換を行ったノード数とした。つまり、ガボールウェーブレット変換の計算回数である。処理時間は、計算機と実装スキルに依存するため、計算コストの指標として相応しくないと考えている。図中のデータは登録人物の全てのテスト画像に対する実験結果の平均である。この曲線は、グラフの左下に位置するほど、設定したパラメータに対するシステム性能が良いことを示す。各曲線の 6 点のプロットは、計算コストが低い方から $d_{smp} = 40, 30, 24, 20, 17, 15$ に対する結果である。 d_{smp} は画像中のサンプリング間隔であるから、大きくするほどサンプリング点の数が減り、計算コストが低くなる。この図から、 $d_{smp} = 40$ の推定誤差が、他の d_{smp} に比べて非常に大きくなっていることがわかる。また、 ΔR_{th} より c の変化に対して推定誤差が大きく影響を受けている。なお、 $c = 1$ より $c = 3, 5, 7$ の推定誤差が小さいので、ビーム探索の効果は明らかである。 n_w に関しては、それを大きくするほど計算コストが増大するが、推定誤差は大きな影響を受けないことがわかる。

次に、顔位置の推定結果と同様に、図 3.13 に顔の大きさの推定に関する実験結果を示す。大きさ推定の誤差は、推定した伸縮率 s_{opt} と真値との誤差である。この図から、 d_{smp} を小さくしても、推定誤差が小さくなるとは限らないことがわかる。また、 $\Delta R_{th} = 0.02$ より、 $\Delta R_{th} = 0.03$ の推定誤差が小さくなる傾向がある。 c と n_w に関しては、顔位置の推定結果に対する考察と同様である。

個人識別実験

図 3.14 に、顔位置の推定結果と同様に 288 通りの個人識別に関する実験結果を示す。個人識別精度は登録人物の個人識別率を指標とする。これは登録人物のテスト画像が入力された場合に、その人物の個人モデルが正しく推定された割合である。この図から、個人識別の精度と、位置と

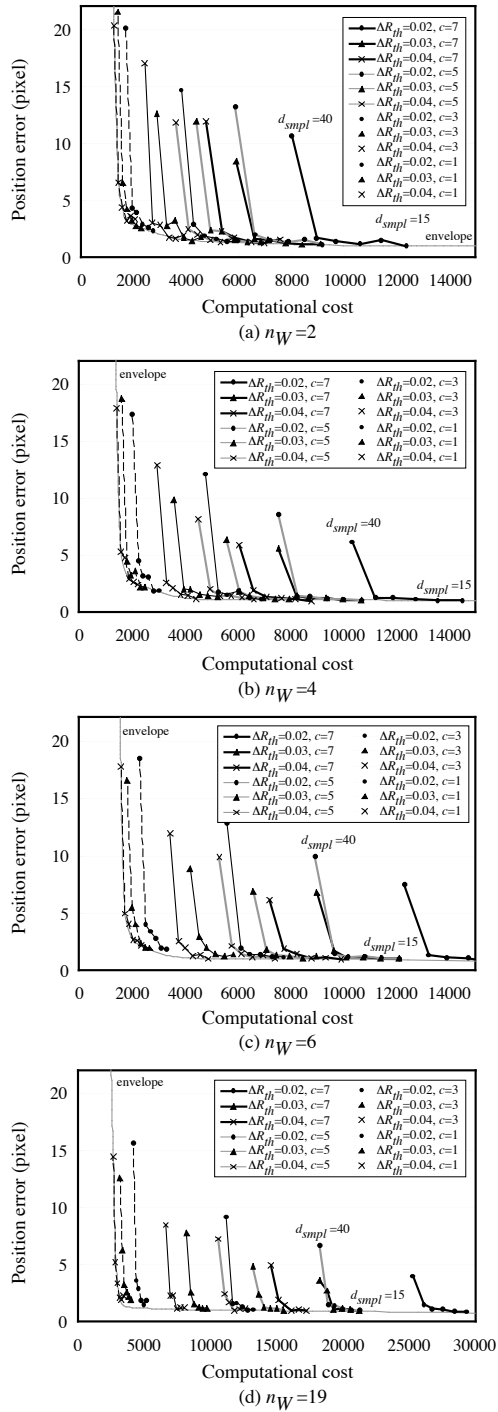
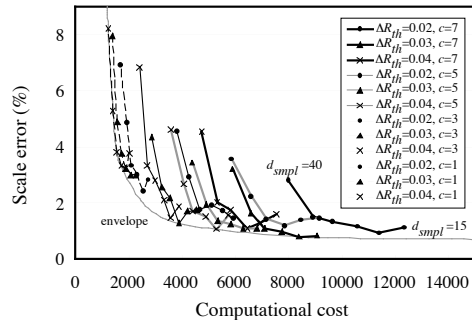
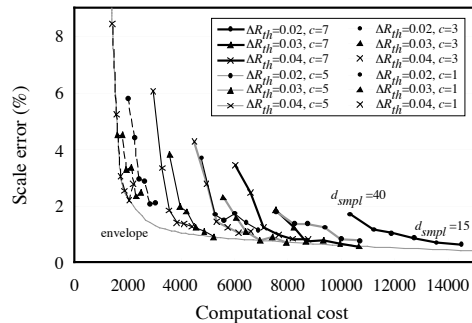


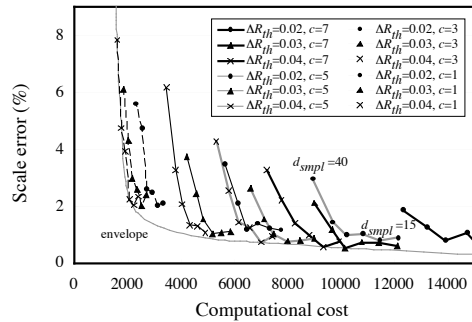
図 3.12: $\Delta R_{th}, d_{smp1}, n_W, c$ の変化に対する顔位置推定の精度と計算コスト



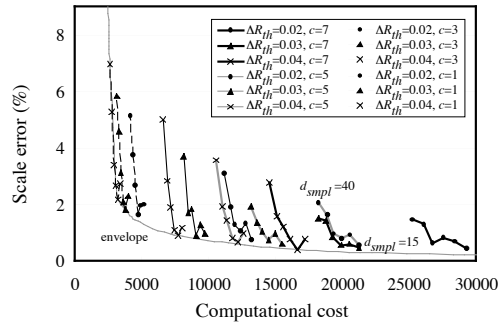
(a) $n_W=2$



(b) $n_W=4$



(c) $n_W=6$



(d) $n_W=19$

図 3.13: ΔR_{th} , d_{smp1} , n_W , c の変化に対する大きさ推定の精度と計算コスト

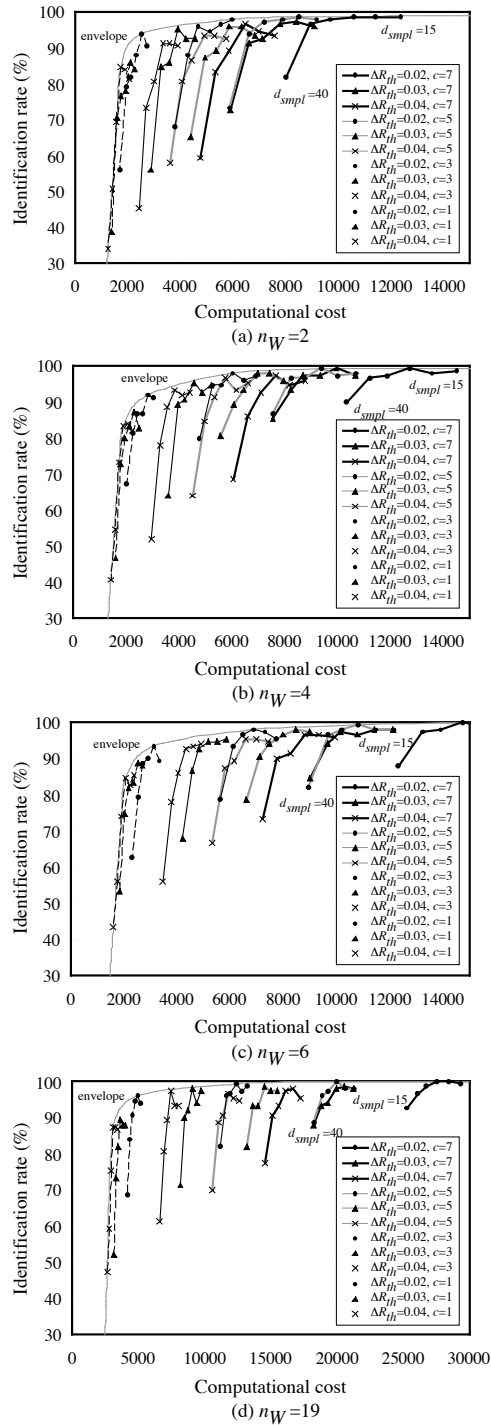


図 3.14: $\Delta R_{th}, d_{smpl}, n_W, c$ の変化に対する個人識別の精度と計算コスト

大きさの推定精度 (図 3.12, 図 3.13) が密接に関わっていることがわかる。

提案システムを運用する場合に適用するパラメータは, その状況で要求されるセキュリティレベルもしくは処理時間を満たすものを, 図 3.12, 図 3.13, 図 3.14 の包絡線付近から選べば良い. 本研究では個人識別精度を重視し, それが 99% を越えたパラメータの中から, 位置と大きさの推定精度が最も良く, 計算コストが最も小さくなった $\Delta R_{th} = 0.03, d_{smp} = 17, n_w = 4, c = 7$ を選んだ. 以降の実験では, このパラメータを用いる.

3.4.5 個人モデル推定の有効性の検証

まず, 提案システムによる顔の位置と大きさと個人モデル推定の推移の例を図 3.15 に示す. 図中の四角い枠に隣接する数値は推定した順番, Lv. は階層データベースのレベル, Cluster と Pos. と Scale はそれぞれ推定したクラスタと位置と大きさ, Sim. はグラフ特徴量類似度を表す. この例に用いた入力画像の顔位置の真値は $(x_{true}, y_{true}) = (122, 155)$, 大きさの真値は 1.2, モデルは Lv.=1 で Cluster=1 に含まれる. この図より, 推移が進むにつれて推定値が真値に向って収束し, グラフ特徴量類似度が高くなっていくことがわかる. 真値から遠ざかる方向への推定も多少確認できるが, これは局所解への収束を防ぐ役割を果たす. また, 重複した推定を行わなかったことも確認できる.

次に, 個人モデル推定が顔位置推定の性能を向上させることを, 状態遷移の繰り返しに伴う顔位置の推定誤差の収束性から検証する. 提案システムとの収束性の比較には, 提案システムから個人モデル推定を省いたシステムとした. この比較システムでは参照モデルグラフに常に平均グラフ G_{ave} を適用した. 図 3.16 に, それぞれのシステムの計算コストの推移に対する顔位置の推定誤差の変化をテストセットごとに示す. この図から両システムのグラフに周期性が見られ, 周期ごとに誤差が収束している様子が見られる. これは優先度付き待ち行列を用いたビーム探索を導入したためである. 各周期が c 回の探索を行う 1 組のビーム探索に相当する. 例えば, 図 3.16 (a) の 5 周期目は, 図 3.15 の 78 番目の推定から 92 番目までに当たる. そして, この図から提案システムの収束性の方が良く, 顔位置推定と個人モデル推定を統合することによって, 推定位置の収束性を向上させられることが明らかである. 計算コストが 1000 を示すあたりまでは両システムの収束性に差がないが, それは, この時点までは両システムで全く同一の処理を適用するためである. なお, 推定処理

表 3.1: 個人モデル推定を行うシステムと行わないシステムの顔位置推定の誤差 (画素)

System	Test set 100%	80%	120%
Model estimation	0.60	1.05	1.67
Non model estimation	3.45	2.50	3.93

終了時の両システムの推定誤差は表 3.1 の通りである。この結果からも、個人モデル推定が有効であるといえる。

3.4.6 従来研究との性能比較

提案システムと2つの従来システムの比較を行う。テスト画像には、テストセット 100%の画像から顔の大きさは変化せずに表情が変化した画像 100 枚 (テストセット expression) と不均一照明下で撮影された画像 100 枚 (テストセット light) を含めた。提案システムには表情変化と不均一な照明変化に対応できるアルゴリズムを導入していないが、現状でどの程度対応できるかを検証する。

図 3.17 に提案システムの処理結果の画像を示す。この結果から、顔の大きさが変化した画像だけでなく、照明が眼鏡に反射し、目が隠れている顔画像や表情、照明条件の異なる顔画像に対しても位置と大きさの推定が正しく行われていることがわかる。また、個人モデルの推定によって最適なモデルグラフを選択しているため、それぞれの人物の顔画像に照合されている顔グラフの形状が異なり、表情変化のない顔画像に対しては特徴点も正確に抽出できていることもわかる。

顔の位置と大きさの推定性能

表 3.2 に、テストセット 100%, 80%, 120%, expression, light ごとの登録人物のテスト画像に対して得られた推定位置の平均誤差、推定した大きさの平均値、平均計算コストを示す。この表から、提案システムの推定精度が従来システムより向上していることがわかる。そして、その計算コストは従来システムの 1/5 程度に軽減されている。また、従来システム 1 が顔の特徴点の検出時に、常に 50 個のモデルとシャドウグラフを

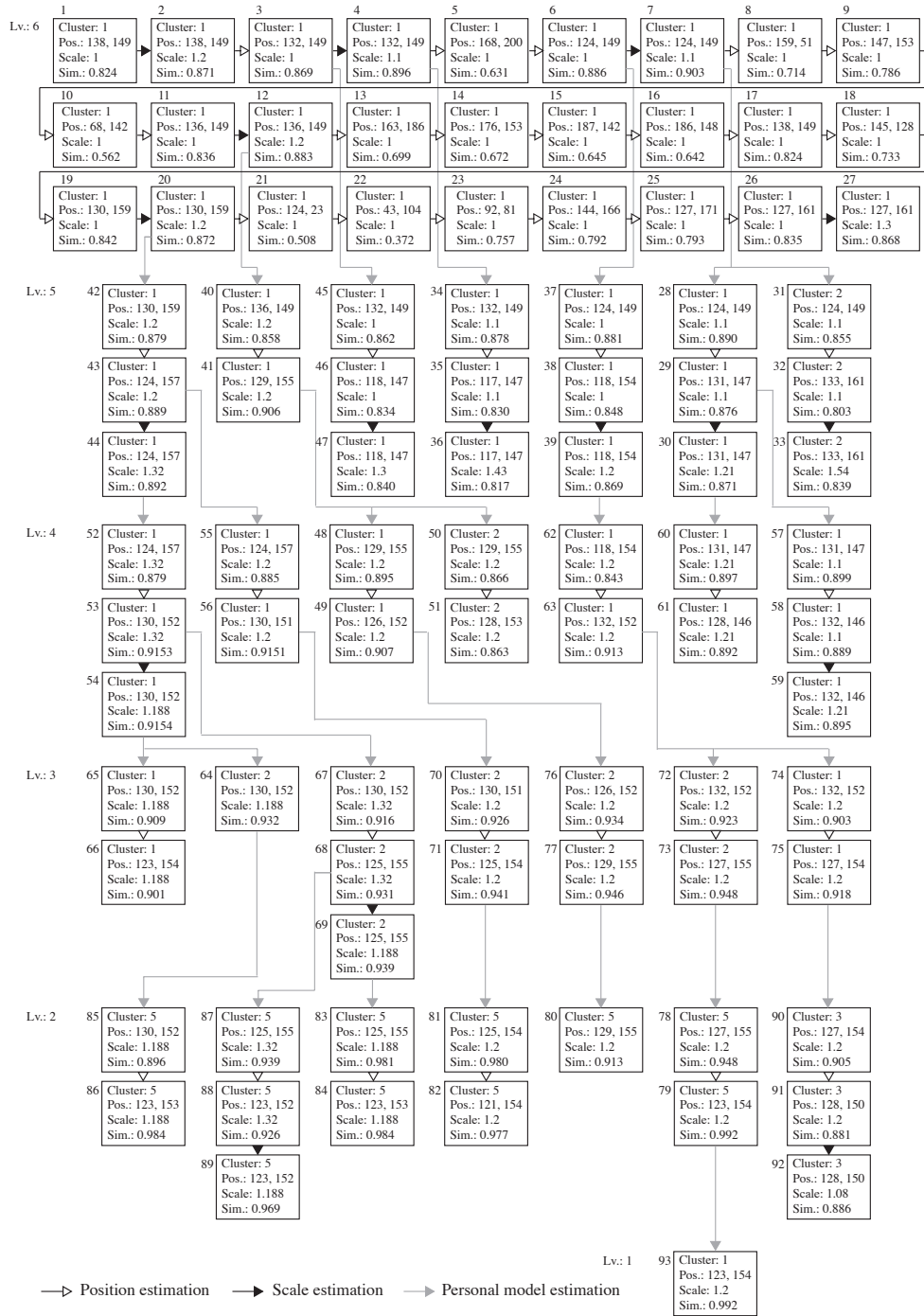


図 3.15: 顔の位置と大きさと個人モデル推定の推移例

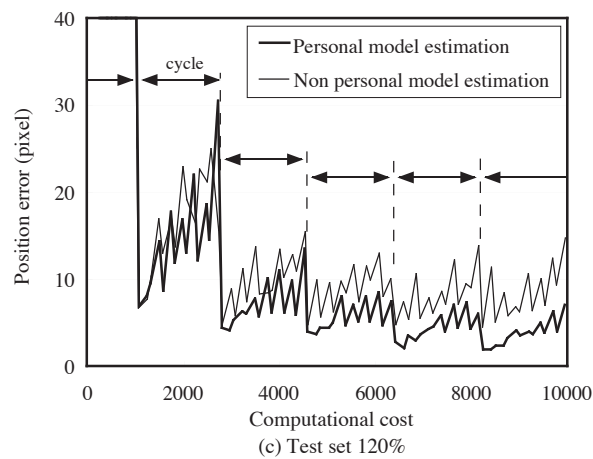
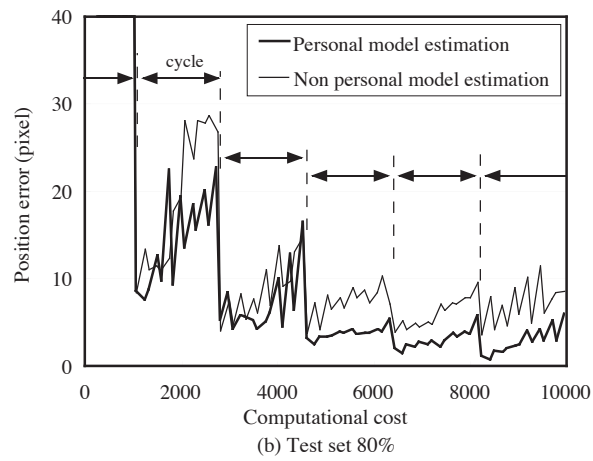
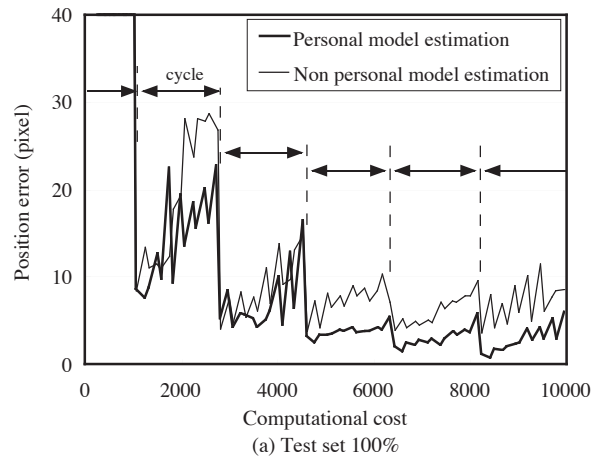


図 3.16: 推定された顔位置の収束性

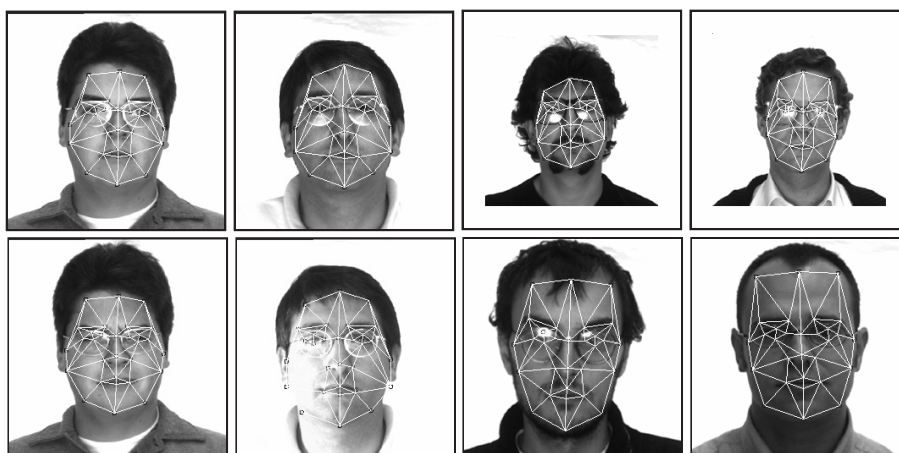


図 3.17: 顔の位置と大きさと個人モデルの推定結果画像

照合して最適なモデルを推定することに対し、提案システムは階層データベースの最大分岐数である 11 個以下のモデルから最適なものを推定する。実際には、モデル推定のための照合回数は、従来システム 1 が 300 回、提案システムが 275 回で、この点からも提案システムの計算効率は向上している。ゆえに、提案システムは位置と大きさの推定精度と計算コストのトレードオフを解決できるといえる。従来システム 1 を導入したオムロン製 FaceKey[大久保 2000] は 1 秒程度で認証処理を行うため、提案システムが十分な計算効率を持つといえる。なお、本実験では、1 個のノードに対してガボールウェーブレット変換を施すのにかかる処理時間は約 2 ミリ秒であった。この処理時間は高速フーリエ変換を導入することによって大幅に軽減させることが可能である。

個人識別の精度

個人識別の精度は watch list と呼ばれる評価方法 [Grother 2003] に基づいて求める。watch list 評価では、登録人物の個人識別率 TAR (True Acceptance Rate) は以下のように求まる。

$$TAR(s) = \frac{|\{p_g : R \geq s, id(p_g) = g\}|}{|\mathcal{P}_G|} \quad \forall p_g \in \mathcal{P}_G \quad (3.7)$$

\mathcal{P}_G は登録人物のテストセット、 p_g は登録モデル g のテスト画像、 $id()$ は推定した個人モデルを返す関数、 R は終了状態に遷移したシャドウグラ

表 3.2: 顔の位置と大きさの推定精度と計算コストの従来システムとの比較

System	Test set	Position error (pixel)	Estimated scale (%)	Cost
proposed	100%	0.60	100 ± 0.2	10022
	80%	1.05	80 ± 0.9	9820
	120%	1.67	120 ± 0.9	10058
	expression	3.42	100 ± 1.1	9979
	light	2.99	100 ± 1.8	9446
previous 1	100%	2.74	100 ± 2.0	46595
	80%	8.78	80 ± 9.2	46595
	120%	14.7	120 ± 13	46597
	expression	4.21	100 ± 4.4	46596
	light	3.30	100 ± 4.0	46595
previous 2	100%	2.62	100 ± 1.6	46355
	80%	9.05	80 ± 3.2	46355
	120%	15.3	120 ± 11	46355
	expression	4.36	100 ± 3.4	46355
	light	3.39	100 ± 2.8	46355

フのグラフ特徴量類似度, s は未登録人物の判定用の閾値を表す. また, 未登録人物の識別率 TRR (True Rejection Rate) は以下のように求まる.

$$TRR(s) = \frac{|\{p_n : R < s\}|}{|\mathcal{P}_N|} \quad \forall p_n \in \mathcal{P}_N \quad (3.8)$$

\mathcal{P}_N は未登録人物のテストセット, p_n は未登録人物のテスト画像である.

図 3.18 にそれぞれの実験システムの個人識別精度を示す. この図は, 未登録人物の判定用の閾値 s を変化させ, それに対する TAR と TRR を ROC 曲線の形で表したものである. この曲線はグラフの右上に位置するほど, その精度が良いといえる. また, 表 3.3 に TAR と TRR が等しいクロスポイントの値を示す. 図 3.18 と表 3.3 より, 全てのテストセットに対して, 提案システムの精度向上が確認できる. 従来システムは, 表情変化や照明の不均一な変化に対応するための特徴点抽出アルゴリズムを導入しているが, 顔位置の推定精度の悪さが特徴点抽出と個人識別処理に影響を及ぼしている.

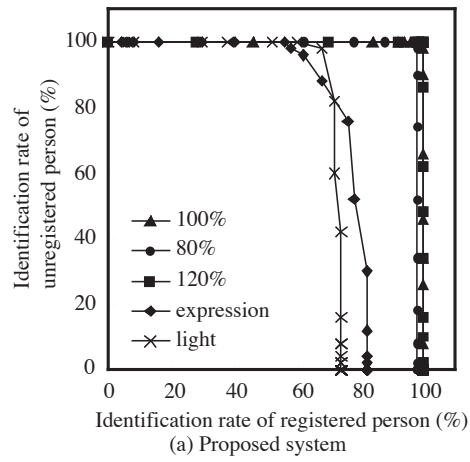
表 3.3: 登録人物と未登録人物の個人識別率のクロスポイントによる従来システムとの比較

System	Test set				
	100%	80%	120%	expression	light
proposed	100%	99%	100%	76%	73%
previous 1	81%	59%	54%	56%	66%
previous 2	84%	73%	65%	54%	49%

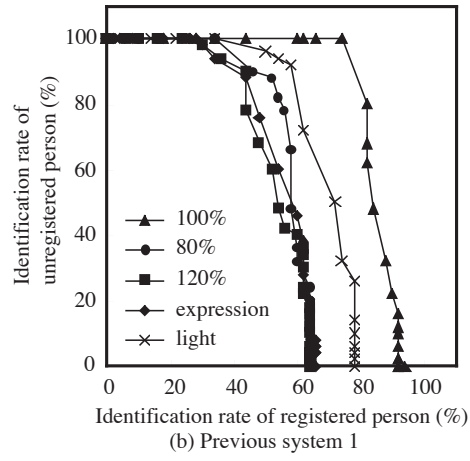
提案システムに関して、テストセット `expression` と `light` に対する精度が他より著しく低下していることがわかる。提案システムは表情変化や照明の不均一な変化に未対応であるため、本人のモデルが含まれるクラスタ以外のクラスタ平均モデルとそれらの変化を持つ顔画像の類似度が高くなると考えられる。ゆえに、個人モデルの推定は失敗しているが、表情や照明が変化した顔画像に最適なモデルを推定している可能性があると考えられる。そこで、終了状態に遷移したシャドウグラフを全ての個人モデルとグラフ特徴量類似度 R を用いて再び照合させた。その結果を図 3.19 に示す。この結果から、再照合によりテストセット `expression` と `light` の個人識別精度が向上したことがわかる。それでも、それらの精度はテストセット 80%, 100%, 120% に比べ著しく低下している。セキュリティシステムの運用時の表情変化や照明の不均一な変化は、顔の大きさや個人性の変化より抑えやすいが、起こりうる変化である。今後、それらに対応できるアルゴリズムを提案システムに組み込む必要がある。

3.5 まとめ

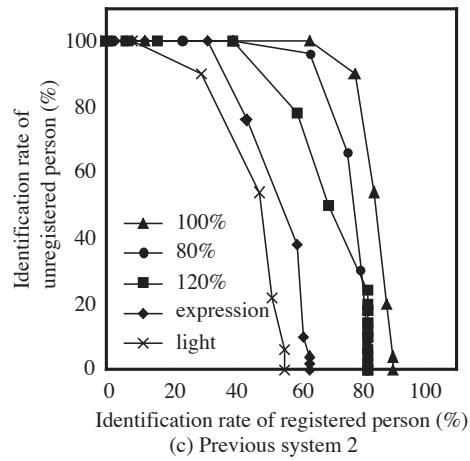
顔の位置と大きさと個人モデルの推定の統合手法を提案した。顔認証実験によって、提案システムが `Elastic Bunch Graph Matching` より高性能であることを検証し、位置と大きさと個人モデル推定の統合が顔認証の精度と計算コストのトレードオフを解決できるという着目点の正しさも実証した。本手法に導入したビーム探索は、 c 個からの更新をそれぞれ独立に行うことができるため、 c 個のプロセスに負荷分散させるという並列処理に展開することが可能である。すなわち、この並列処理を導入すれば、顔認証の精度と処理時間のトレードオフのさらなる解決が期待



(a) Proposed system



(b) Previous system 1



(c) Previous system 2

図 3.18: 個人識別精度の従来システムとの比較

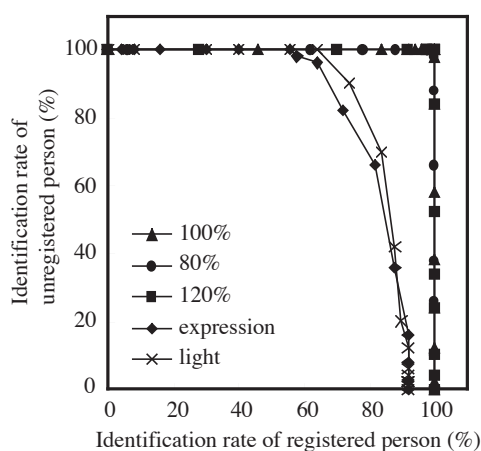


図 3.19: 再照合後の個人識別精度

できる。

提案手法の問題点としては、データベースに登録される個人モデルが増加した場合に、1つのクラスタから分岐するクラスタ数も増加するということである。この場合、本実験結果の精度を維持するためには、ビーム探索の c を大きくしなければならない可能性がある。それに伴い計算コストが増加する。そこで、個人モデルが膨大となった場合には、類似する個人モデルも多数存在すると考え、それらを主成分分析などを利用して圧縮し、階層モデルデータベースを構成する手法を考えている。

また、顔の位置と大きさの推定精度が画像中のサンプリング方法に左右されるので、その方法として有効とされる retina 法 [Smeraldi 2000] を提案手法に導入すれば、さらに推定精度が向上する可能性がある。今後は表情変化や照明の不均一な変化への対応にも取り組む必要がある。

第4章

赤外画像を用いた人物写真判別システム

顔写真による詐称への対応は、従来の顔認証セキュリティシステムが持つ課題である。本研究では、赤外画像によって人物と顔写真を判別するシステムを提案する。本システムは、加算投影法などの計算効率の良い画像処理手法に基づいて顔領域を検出する。赤外画像を利用することで、背景成分の変化にロバストな顔領域検出が可能になる。人物と顔写真の判別には、照度差ステレオ法に基づく効率的な手法を提案する。本手法は、従来研究では考慮されなかった湾曲写真にも対応できる。本章では、まず赤外光の特性と赤外画像の特徴について述べる。そして、提案システムの構成、顔領域検出アルゴリズム、人物写真判別アルゴリズムについて説明する。

4.1 赤外光の特性と赤外画像の特徴

赤外光は電磁波であり、その発見が熱効果によることで熱線ともいわれる [赤外線技術研究会]。可視光の波長が 300nm~700nm に対し、赤外光の波長は可視光に比べて、帯域幅が広く、700nm~1mm にも及ぶ。図 4.1 に可視光画像と赤外画像の違いを示す。可視光画像とは、可視光に対して最も高い感度を持つセンサで撮影した画像のことである。また、赤外画像とは、赤外光照明を被写物体に照射し、赤外カメラを用いて撮影したものである。赤外カメラは赤外光に対して感度のピークを持つセンサから構成される。可視光画像は人物と背景がはっきり写っている。一



(a) Visible image

(b) Infrared image

図 4.1: 可視光画像と赤外画像の違い

Visible	Near infrared		Middle infrared	Far infrared	
	Low band	High band			
0.3 μm	0.7 μm	1.5 μm	2.5 μm	25 μm	1mm

図 4.2: 可視光と赤外光の波長域

方、赤外光の光量にもよるが、赤外画像には背景がほとんど写っていない。これは屋内環境には強い赤外光が存在しないためである。従って、赤外光の光量を調整することで、画像中の人物と背景のコントラストを大きくすることができる。これにより、背景成分の変化にロバストかつ効率良く顔を検出することができる。

赤外光はいくつかの種類に分類される。分光学上は次のように定義される。700nm \sim 2.5 μm が近赤外光、2.5 μm \sim 25 μm が中赤外光、25 μm 以上のものが遠赤外光と分類される。また、近赤外光は1.5 μm を境界に低域波長と高域波長に分類される。図 4.2 は赤外光の波長域を図示したものである。それぞれの波長に対する物質の反射特性が大きく異なることが Pavlidis らによって示されている [Pavlidis 2000]。人間の顔に限って言うと、その反射特性が異なる部位は肌と毛髪とそれ以外に分類される。本研究では、近赤外光の低域波長を用いる。なぜなら、その波長を用いて撮影された画像がアピアランスの点で可視光画像と最も類似し、顔検出後に行われる顔認証に可視光画像用の従来手法が適応できるためである。なお、本論文では、これ以降で赤外という言葉が使われているところは全て近赤外光の低域波長を意味する。

赤外光の特性として、その反射特性が人間の肌と写真で異なることも

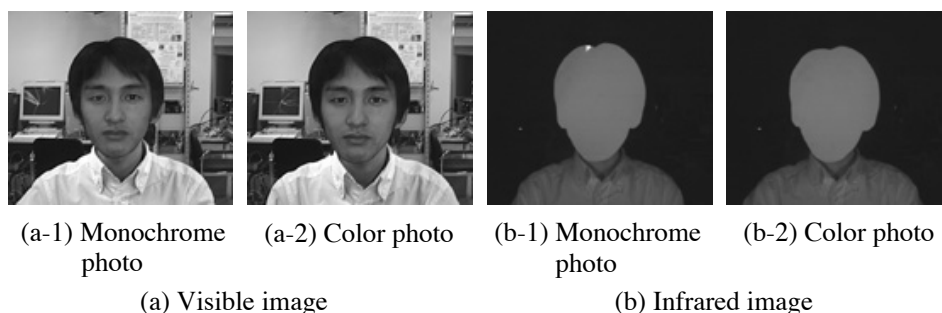


図 4.3: 光沢紙に印刷された写真を撮影した可視光画像と赤外画像

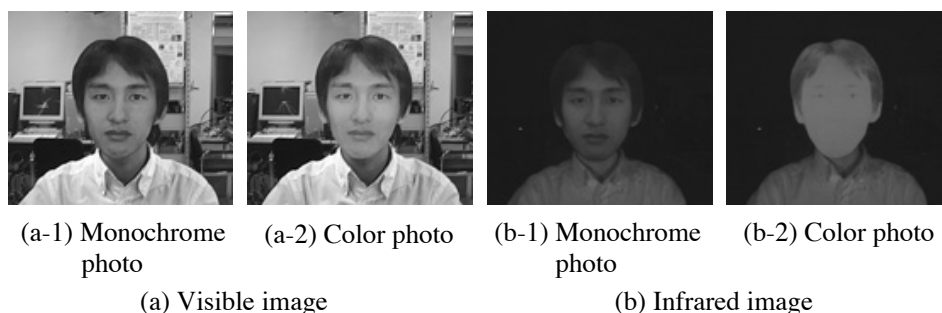


図 4.4: 普通紙に印刷された写真を撮影した可視光画像と赤外画像

挙げられる．図 4.3 に光沢紙に印刷された顔写真を撮影した可視光画像と赤外画像を示す．光沢紙に印刷された顔写真を撮影した赤外画像は白黒写真やカラー写真に関係なく，顔のテクスチャを失っている．次に，普通紙に印刷された顔写真を撮影した可視光画像と赤外画像を図 4.4 に示す．普通紙に印刷されたカラー写真は光沢紙と同様に，図 4.4(b-2)のように顔のテクスチャを失う場合が多いが，カラーの色彩によっては，顔のテクスチャがわずかに現れる場合もある．一方，普通紙に印刷された白黒写真は人間の肌が撮影された赤外画像 (図 4.1(b)) と類似していることが分かる．ゆえに，人物と顔写真を赤外光の反射特性の違いだけから完全に判別することは困難である．

そこで，本研究では，人間の顔と写真の違いとして挙げられる凹凸の有無に注目し，被写物体の 3 次元形状を求める方法である照度差ステレオ法を利用する．この手法は，複数の照明を切り換えて，複数の画像を撮影する必要がある．赤外光照明を利用すれば，人間が赤外光を可視できないため，ユーザは照明の切り換えによるストレスを受けない．さら

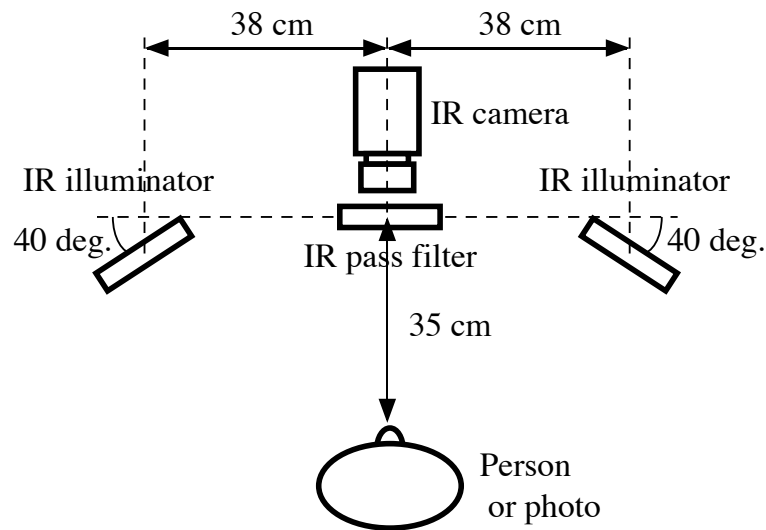


図 4.5: カメラ, 照明, 認証物体の位置関係

に, 人物写真判別のアルゴリズムが見破られ, 詐称される可能性が低い.

4.2 システム構成

顔認証システムにおいて, 顔画像の撮影時に次の想定が考えられる.

1. ユーザは無表情で, 故意に動かず静止する.
2. ユーザはカメラに対して正面を向く.
3. ユーザの立ち位置はカメラのほぼ正面で, カメラからの距離に大きな変化がない.

ただし, ユーザに毎回同じ状態での撮影を要求することは困難であるため, 顔の見え方の多少の変動を考慮する必要がある. 特に, 立ち位置の変化は最も頻繁に起こり, 他の変化よりも大きくなりやすい. 本研究では, これらの想定を考慮して, 人物写真判別システムを開発する.

本システムは, その構成を簡素化するために, 照度差ステレオ法に最低限必要な2台の照明を利用する. 2台の赤外光照明は, 図4.5のように赤外カメラを中心に左右に配置する. 左右に配置する理由は, 顔の凹凸が最も大きな部位が鼻筋と頬であり, 左右照度差画像によってこの凹凸

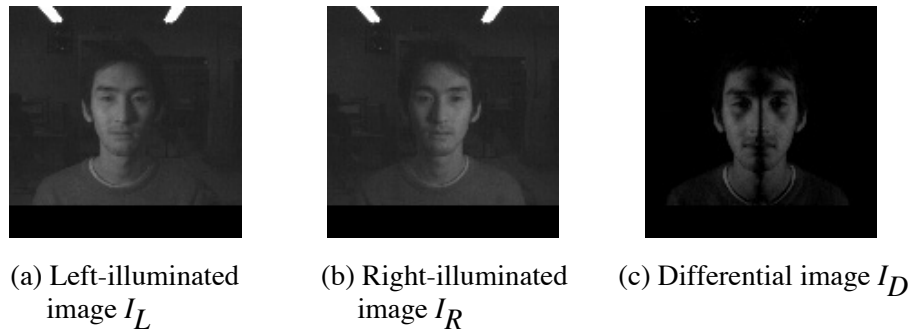


図 4.6: 左右照射画像と左右照度差画像

を観測できるためである。照度差画像 I_D は、左右の照明を切り換えて認証物体を 2 回撮影し、以下のように左照射画像 I_L と右照射画像 I_R の差分を取ることで生成する (図 4.6)。

$$I_D(x, y) = |I_L(x, y) - I_R(x, y)| \quad (4.1)$$

なお、左右照度差画像の生成には、全く同一のシーンを撮影した左右の照射画像を用いることが理想である。つまり、2 枚の画像の撮影中は、人物が完全に静止し、環境が全く変化しないという状態である。しかし、ユーザに完全な静止を求めることが困難であるため、シャッタースピードの高い赤外カメラと画像取り込み速度の高い画像取り込みボードで連続撮影することが必要である。左右の照明の切り換えのタイミングは、カメラのシャッターを切るタイミングと同期をとる。タイミング信号はデジタル I/O ボードによって、カメラと照明コントローラに入力される。

本システムの処理の流れを図 4.7 に示す。照度差画像に対して、顔領域検出と人物写真判別を順に施す。

4.3 顔領域検出アルゴリズム

まず、照度差画像 I_D に対して 2 値化を行う。顔認証システムの想定より、顔の大きさの変化はほとんどなく、顔領域の面積がほぼ一定であることに基づき、p-tile 法を利用して 2 値化の閾値 B_{th} を決める。照度差画

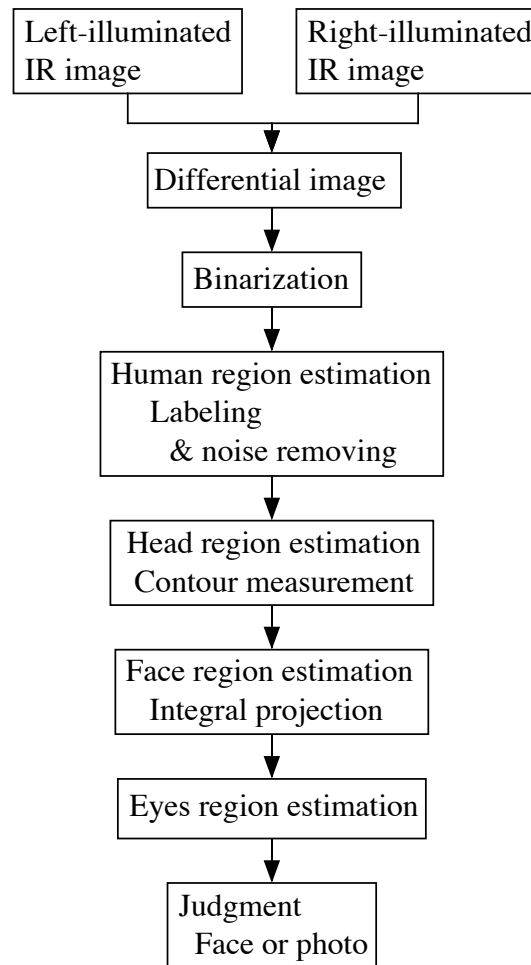


図 4.7: 人物写真判別システムの処理の流れ

像の2値化画像は以下によって生成する.

$$I_B(x, y) = \begin{cases} 1 & \text{if } I_D(x, y) \geq B_{th} \\ 0 & \text{otherwise} \end{cases} \quad (4.2)$$

2値化された照度差画像は、顔領域とそれ以外が分離されるべきである。しかし、髪や服の領域の輝度分布が顔領域と類似する場合、人物領域が全体的に抽出される。抽出された人物領域はいくつかの領域に分かれ、背景領域には雑音成分が存在する。この雑音を除去するために、各領域の面積を求める必要があるため、ラベリングを行う。ラベリングには浸水法 [萬上 2001] を用いる。ラベリングを高速に行うために、間隔 d 画素の

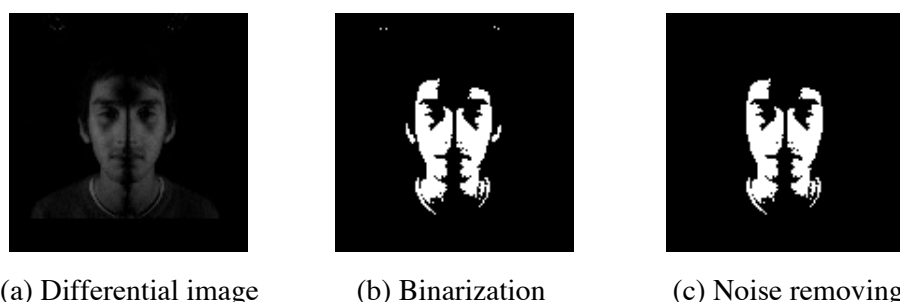


図 4.8: 2 値化後の画像とノイズ除去後の画像

格子状のサンプリングを浸水法に導入する．雑音成分を排除した人物領域は式 (4.3) によって得る．ラベル画像 I_{label} 中のラベル L の領域 R_L とその面積 S_L を求め，面積 S_{th} を超えた領域のラベルを A_{person} に更新する． A_{person} は人物領域 R_{person} を表すラベルである (式 (4.4))．2 値化後の画像とノイズ除去後の画像を図 4.8 に示す．

$$I_{label}(x, y) = A_{person} \quad \text{if } S_L \geq S_{th}, \quad (x, y) \in R_L \quad (4.3)$$

$$R_{person} \triangleq \{(x, y) | I_{label}(x, y) = A_{person}\} \quad (4.4)$$

次に，人物領域から頭部領域，顔領域，目領域を順に推定する．頭部領域 R_{head} は，人物領域の最上部である頭頂 y_{top} から首の位置 y_{neck} までの領域と定義する．ただし，本システムでは頭部領域の幅は設定しない (式 (4.5))．首の位置は，人物領域の幅 W_{person} が耳の位置 y_{ear} から人物領域の最下部 y_{bottom} までで最小値を取る位置として推定する (式 (4.6))．耳の位置は，画像を垂直方向に順次走査し，各列における人物領域の最上部位位置の変動が最大となる位置とする (式 (4.7))．頭部領域の推定の様子を図 4.9 に示す．

$$R_{head} \triangleq \{(x, y) | y_{top} \leq y \leq y_{neck}, (x, y) \in R_{person}\} \quad (4.5)$$

$$y_{neck} = \arg \min_{y_{ear} \leq y \leq y_{bottom}} W_{person}(y), \quad (4.6)$$

$$W_{person}(y) = x_R(y) - x_L(y),$$

$$x_L(y) = \min\{x | (x, y) \in R_{person}\},$$

$$x_R(y) = \max\{x | (x, y) \in R_{person}\}$$

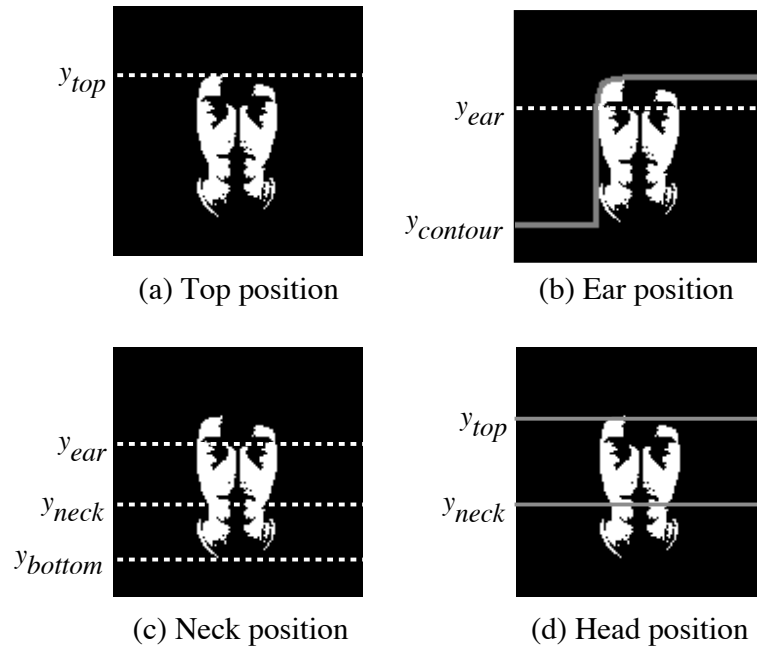


図 4.9: 頭部領域の推定

$$\begin{aligned}
 y_{ear} &= y_{contour}(x_{ear}), \\
 y_{contour}(x) &= \min\{y \mid (x, y) \in R_{person}\}, \\
 x_{ear} &= \arg \max_x \{y_{contour}(x) - y_{contour}(x-1) - y_{bottom}\}
 \end{aligned} \tag{4.7}$$

顔領域 R_{face} は眉の位置 $y_{eyebrow}$ から、すでに推定された首の位置までの領域と定義する。顔領域の幅 W_{face} は眉から首の距離と同じとし、幅の2分点の x 成分は頭部領域の水平方向の重心 x_{center} とする(式(4.8))。まず、眉の位置を推定する前に、鼻の位置 y_{nose} を頭部領域の水平加算投影 $P(y)$ の最大位置 [リュウ 2004] から求める(式(4.9))。そして、眉の位置は、頭頂から鼻の位置までの水平加算投影の極小値 $Y_{localmin}$ の中で最小を取る位

置 [リユー 2004] として推定する (式 (4.10)).

$$R_{face} \triangleq \left\{ (x, y) \mid y_{eyebrow} \leq y \leq y_{neck} \right. \\ \left. x_{center} - \frac{W_{face}}{2} \leq x \leq x_{center} + \frac{W_{face}}{2} \right\}, \quad (4.8)$$

$$W_{face} = y_{neck} - y_{eyebrow},$$

$$x_{center} = \frac{\sum_{(x,y) \in R_{head}} x}{\sum_{(x,y) \in R_{head}} P(y)},$$

$$P(y) = \sum_{(x,y) \in R_{head}} I_B(x, y)$$

$$y_{nose} = \arg \max_{y_{top} \leq y \leq y_{neck}} P(y) \quad (4.9)$$

$$y_{eyebrow} = \arg \min_{y \in Y_{localmin}, y_{top} \leq y \leq y_{nose}} P(y) \quad (4.10)$$

目領域 R_{eyes} は、すでに推定された眉の位置から鼻の位置までの領域とする。目領域の幅 W_{eyes} は顔領域の幅と等しく、幅の2分点の x 成分は頭部領域の水平方向の重心 x_{center} とする (式 (4.11))。目領域は、顔の凹凸が最も大きな部位である。この目領域を人物写真判別に用いる。顔領域と目領域の推定の様子を図 4.10 に示す。

$$R_{eyes} \triangleq \left\{ (x, y) \mid y_{eyebrow} \leq y \leq y_{nose} \right. \\ \left. x_{center} - \frac{W_{eyes}}{2} \leq x \leq x_{center} + \frac{W_{eyes}}{2} \right\}, \quad (4.11)$$

$$W_{eyes} = W_{face}$$

4.4 人物写真判別アルゴリズム

次に、人物写真判別処理について述べる。顔写真を用いてユーザが詐称するパターンは2つ存在すると考えられる。通常、平坦な顔写真がそのまま赤外カメラに向けて差し出されることがほとんどであると思われる。もう1つのケースとしては、写真を湾曲させた顔マスクをユーザが装着することも考えられる。顔写真は、紙の湾曲度 (平坦, 湾曲) の他に

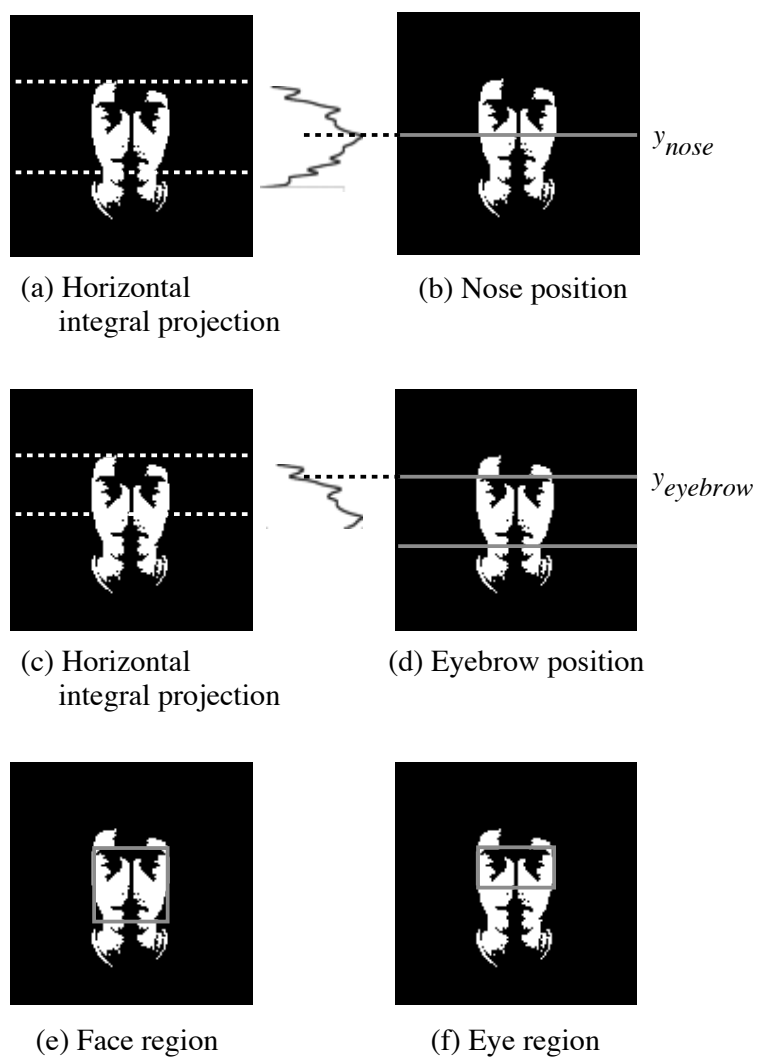


図 4.10: 顔領域と目領域の推定

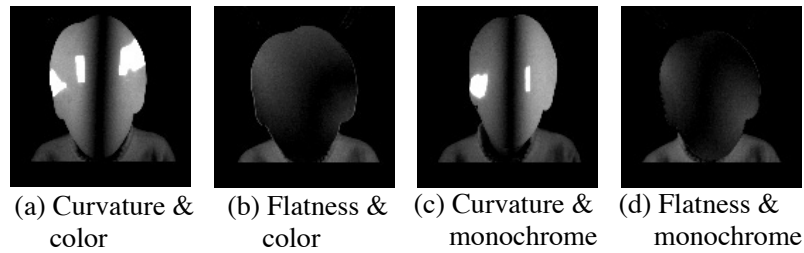


図 4.11: 光沢紙に印刷された顔写真の照度差画像

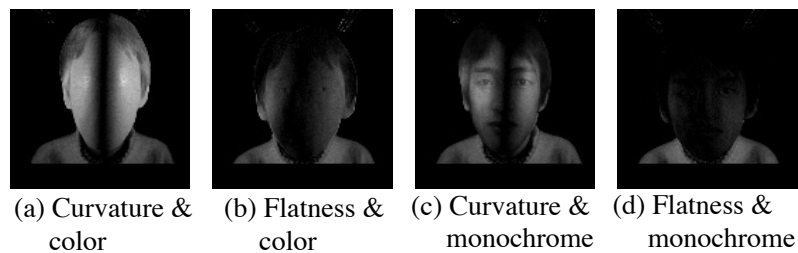


図 4.12: 普通紙に印刷された顔写真の照度差画像

も、紙質 (光沢紙, 普通紙), 色 (カラー, 白黒) によって照度差画像への写りが変化する. その特徴としては, 平坦写真ではほとんど照度差が得られないこと, 普通紙に印刷された白黒写真以外では顔のテクスチャが現れないことが挙げられる. 図 4.11 と図 4.12 に顔写真を撮影した照度差画像を示す. それらと人物の照度差画像 (図 4.6 (c)) の違いを分析すると次のようになる.

1. 顔写真の照度差画像は, 頭部領域がほとんど存在しないか, 大きく隔たった左右の領域に分かれるが, 人物の照度差画像の頭部領域は, 左右に分かれたとしても隔たりが小さい. これは被写物体面の湾曲度に依存する.
2. 人物の顔には凹凸があるため, 人物の照度差画像の目領域に筋状のテクスチャが現れるが, 顔写真の照度差画像にはそれが現れない.

本システムでは, この 2 つの違いを反映した 2 つの特徴量

1. 湾曲度
2. 凹凸度

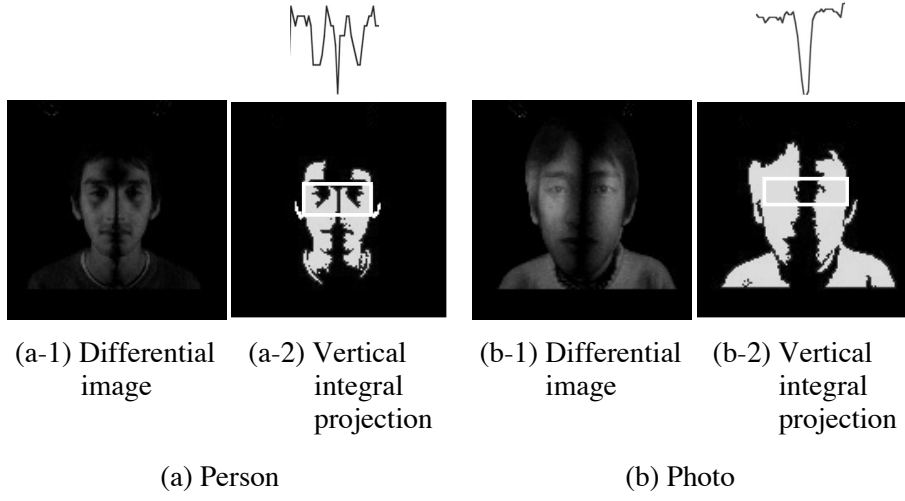


図 4.13: 凹凸度

を導入し，人物と顔写真を判別する．湾曲度 D_{curve} は，鼻の位置 y_{nose} における人物領域の左右領域間の距離 (式 (4.12)) とし，凹凸度 D_{bamp} は，目領域の垂直加算投影の値 $P(x)$ から求める (式 (4.13))．凹凸度を求める様子を図 4.13 に示す．図 4.13 (a) の凹凸度が 6，(b) の凹凸度が 2 である．認証物体 Z が人物か顔写真であるかの判別は，閾値処理によって行う．それは，湾曲度の閾値 c_{th} と凹凸度の閾値 b_{th} を用いた式 (4.14) に従う．

$$D_{curve} = v_R(y_{nose}) - v_L(y_{nose}), \quad (4.12)$$

$$v_L(y_{nose}) = \max_{x \leq x_{center}} \{x | (x, y_{nose}) \in R_{person}\},$$

$$v_R(y_{nose}) = \min_{x \geq x_{center}} \{x | (x, y_{nose}) \in R_{person}\}$$

$$D_{bamp} = \sum_{x \in R_{eyes}} 1 \quad \text{if } P(x) = P_{th} \cap P(x-1) \neq P_{th}, \quad (4.13)$$

$$P(x) = \sum_{(x,y) \in R_{eyes}} I_B(x, y),$$

$$P_{th} = \frac{\max_{x \in R_{eyes}} P(x) - \min_{x \in R_{eyes}} P(x)}{2}$$

$$Z = \begin{cases} 'photo' & \text{if } D_{curve} > c_{th} \cap D_{bamp} < b_{th} \\ 'person' & \text{otherwise} \end{cases} \quad (4.14)$$

4.5 まとめ

赤外光の特性を利用した人物写真判別システムを提案した。このシステムは加算投影法などの計算効率の良いアルゴリズムで構成されるため、従来システムより高効率なシステムとなっていることが期待できる。システム性能の検証については、第5章で提案する顔認証セキュリティシステム FACELOCK の性能を検証する際に行う。

第5章

顔認証システム FACELOCK

本研究に先行して開発された顔認証セキュリティシステムの代表として、土居らによる PASSFACE[土居 1997]，東芝製 FacePassTM[土橋 2002]，Visionics 製 FaceIt^(R)，オムロン製 FaceKey[大久保 2000] が挙げられる。PASSFACE はテンプレート照合，FacePassTM は制約相互部分空間法[福井 1999]，FaceIt^(R) は Local Feature Analysis[Penev 1996] を利用して顔認証を行っている。FaceKey では，顔グラフを利用した Elastic Bunch Graph Matching が顔認証に利用されている。Elastic Bunch Graph Matching は，数ある顔認識手法の中で最も精度が高いとされている [Phillips 2000, Ayinde 2002]。本研究では，顔の見え方の変化に対する認証精度が Elastic Bunch Graph Matching を上回る顔認証手法を第3章で提案した。本章では，この手法を導入した顔認証セキュリティシステムを提案する。その名は FACELOCK である。本システムは，第4章で提案した人物写真判別機能を持つ。従来システムはこの機能を備えていない。

本システムは，運用性とセキュリティ性を考慮して，LAN を介したクライアントサーバ構成とする。つまり，LAN のある所であればどこにでもクライアントを設置できる利点がある。それから，計算コストが高いプロセスはサーバに集められるので，クライアントが高パフォーマンスを持つ必要がない。つまり，本システムを多数導入する場合，クライアントを安価で構築できるため，全体的なコストを抑えることができる。この構成は FacePassTM，そして FaceKey と同様である。さらに，顔認証に使われるユーザの顔画像は暗号化の機構を持つネットワークを通してサーバに送られるので，ユーザのプライバシーをクラッキングから守ることができる。また，顔画像はサーバに記録されるので，クライアントが盗難されてもユーザのデータを損失することはない。なお，クライアントとサー

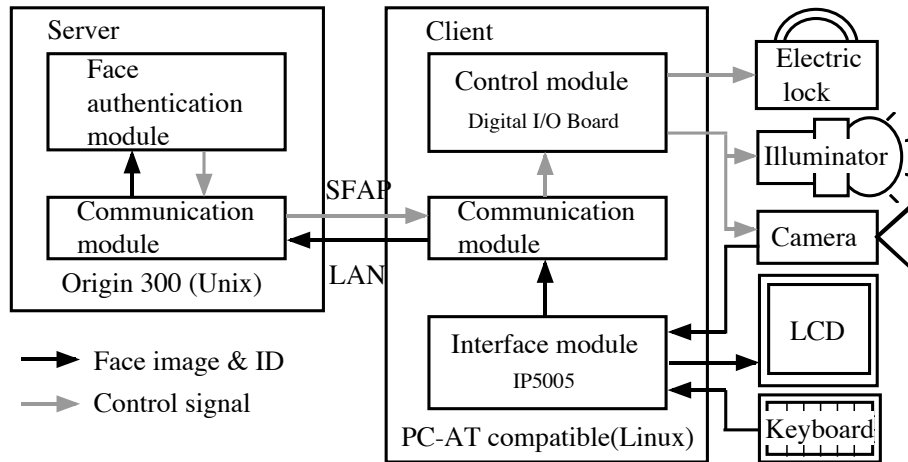


図 5.1: FACELOCK の概略

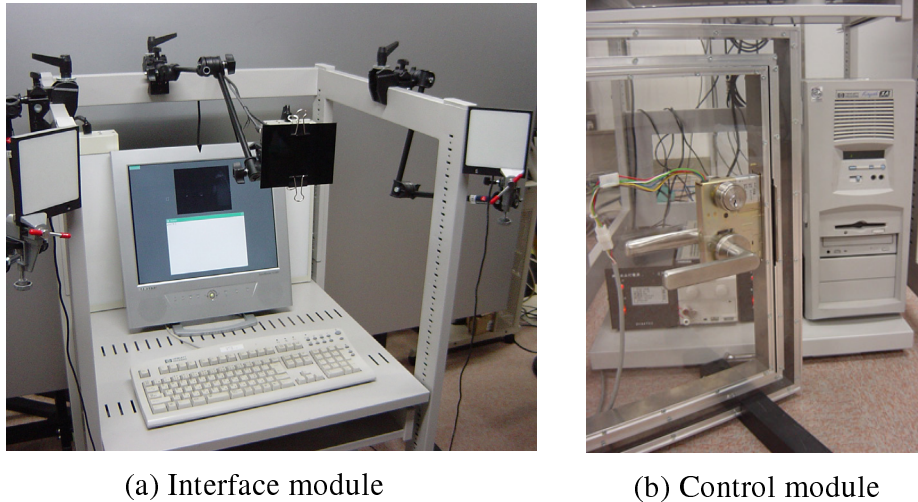
バ間の通信には、一般的な RADIUS[Rigney 1997] や Kerberos [Kohl 1993] という通信プロトコルではなく、独自に開発したものを導入する。

5.1 システム概要

FACELOCK の概略を図 5.1 に示す。クライアントは PC を用いて構築し、登録や認証の際にユーザとのインターフェイスとなる操作部と、電子錠などの制御を行う制御部からなる。サーバは認証処理と撮影画像などの管理を行う顔認証部で、SGI 製ワークステーション Origin300 (8CPU R14000, 500 MHz, memory 4 GB, OS IRIX6.5.16m) 上で構築される。認証処理には、2つのモードが用意されている。全登録モデルと撮影画像を照合する 1 対多照合モードと、ユーザが ID 番号をシステムに入力し、その番号に対応したモデルと撮影画像を照合する 1 対 1 照合モードである。クライアントとサーバ間の通信は、独自に開発した通信プロトコル SFAP を利用して通信部と LAN を介して行う。

5.2 操作部と制御部

操作部はモニタ、日立製赤外カメラ KP-M2R, ID 入力用キーボード、日立製画像処理ボード IP5005 から構成される (図 5.2 (a))。KP-M2R は図 5.3



(a) Interface module

(b) Control module

図 5.2: 操作部と制御部

に示すように 600nm 付近の波長に感度のピークを持つ。可視光の波長帯にも感度を持つため、カメラの前に赤外光パスフィルタを取り付ける。このフィルタの光透過率を図 5.4 に示す。制御部は電子銃、赤外光照明、それらを制御するためのデジタル I/O ボード、赤外カメラから構成される (図 5.2 (b))。赤外光照明は 880nm の赤外光を放射する。

5.3 顔認証部

図 5.5 に、顔認証部の処理の流れを示す。顔認証部は人物写真判別部と個人識別部から構成される。人物写真判別部は第 4 章で提案した人物写真判別システムに基づく。一方、個人識別部は第 3 章で提案した顔の位置と大きさの推定と個人識別の統合手法に基づく。まず、人物写真判別部が実行され、照度差画像に対して、顔領域検出と人物写真判別が順に施される。人物写真判別において認証物体が人物であると判定された場合は、左右照射画像から式 (5.1) に従って照度和画像 I_A が生成される。照度和画像を図 5.6 に示す。

$$I_A(x, y) = I_L(x, y) + I_R(x, y) \quad (5.1)$$

そして、照度和画像に人物写真判別部で推定された顔領域 R_{face} を射影し、その領域に対して個人識別部の顔認証を適用する。

個人識別部は、1 対多照合モードと 1 対 1 照合モードによって個人モデ

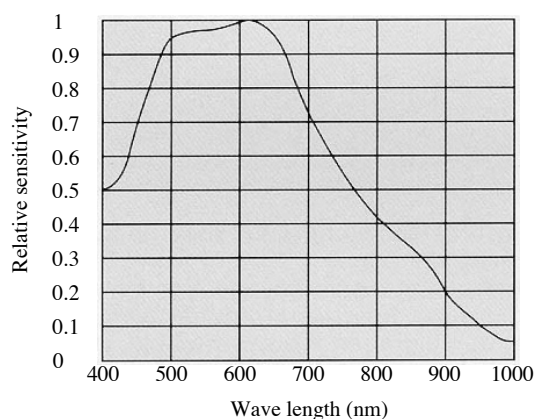


図 5.3: 赤外カメラ KP-M2R の分光感度特性

ルデータベースの利用方法が異なる。1対多照合モードにおいては階層構造を持つ個人モデルデータベースが用いられ、1対1照合モードにおいてはID番号に対応した人物のモデルグラフだけを参照モデルグラフとして利用する。

5.4 通信部

通信部は、認証サーバとクライアントとがネットワーク透過的に通信するためのモジュールである。顔認証のための通信プロトコルはこのモジュールで完全に隠蔽され、アプリケーションは、ネットワークの存在を考慮する必要がない。顔認証のための通信プロトコル SFAP は付録 B で詳しく述べる。

5.5 評価実験

5.5.1 実験環境

FACELOCK の性能を評価実験により検証する。システムは室内の乱雑な場所に設置した。人物の撮影は 50 人に対して行った。まず、カメラとユーザ間の距離を 35cm としてモデル画像を撮影した。次に、システムの運用時にその距離を完全に統制することは困難であるため、28cm, 42cm とわずかに変えてテスト画像を 2 枚撮影した。そして、再びその距離を

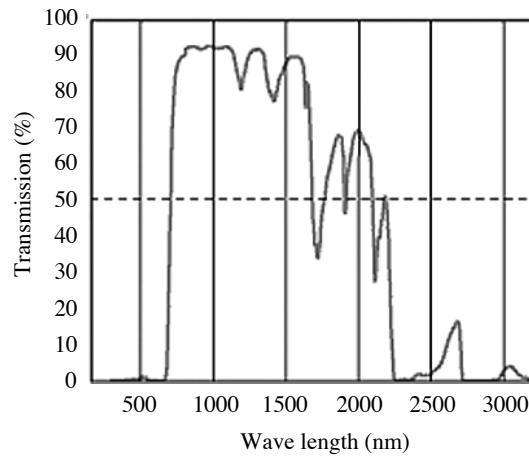


図 5.4: 赤外光パスフィルタの光透過率

35cm としてテスト画像を撮影した。距離 35cm における顔画像の大きさを標準顔 (100%) とすると、28cm では標準顔の 120%，42cm では標準顔の 80% の大きさとなる。なお、標準顔の顔領域の幅は平均で 138 画素であった。写真詐称に対する撮影は、6 人の顔写真をそれぞれ 8 パターン ((平坦写真 or 湾曲写真) and (光沢紙 or 普通紙) and (白黒写真 or カラー写真)) 用意し、人物の撮影と同様に行い、モデル画像 48 枚、テスト画像 144 枚を撮影した。顔認証システムの想定に基づき、撮影された全ての画像は無表情正面顔であった。また、画像は 8 ビット濃淡で、 512×512 画素の大きさであった。人物写真判別部では、その処理の精度が画像の大きさにほとんど影響を受けないので、処理を高速化するために原画像を 128×128 画素の大きさへと縮小したものを用いる。個人識別部では原画像を用いる。なお、左右の赤外光照明の切り換えと、それと同期をとっての撮影は現状で自動化していないため、手動によって行った。人物写真判別部で設定する閾値等のパラメータは 5.5.2 で導出する。また、個人識別部の設定パラメータは 3.4 で示したものと同様とした。本論文では他の顔認証システムとの比較を行わず、本システムがセキュリティシステムとして要求される認証精度を満たすかどうかの検証を行う。その理由は、他システムの関連文献に報告されている実験環境は詳細なものではなく、また、開発者独自の判定基準で認証精度を求めているため、比較するのが困難なためである [瀬戸 1999]。

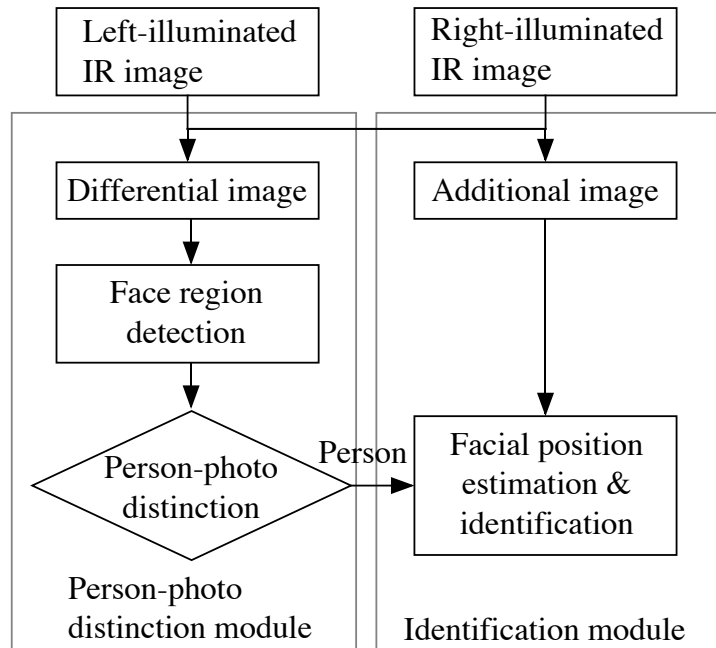


図 5.5: 顔認証部の処理の流れ

5.5.2 人物写真判別用の顔パラメータの導出

人物写真判別アルゴリズムで用いられる顔パラメータを以下のように導出した。

1. 2値化の閾値 B_{th}
 顔領域を手作業によって求め、その面積と画像全体の面積の割合の平均を計算し、p-tile法で画像濃淡値の閾値 B_{th} を求めた。25人分の標準顔の人物モデル画像からそれを求めたところ、28という濃淡値が得られた。
2. ラベリングのサンプリング間隔 d とノイズ除去の面積 S_{th}
 ラベリングを高速に行うために、サンプリング間隔 d を5画素とした。このとき、 5×5 画素の正方領域が画像中のどこに存在しても、正方領域の中には必ず1つのサンプリング点が入ることになる。従って、25画素 (5×5 画素) の面積をもつ正方領域が必ずサンプリングされるラベリング法であるから、ノイズ除去の閾値 S_{th} は25と決定される。

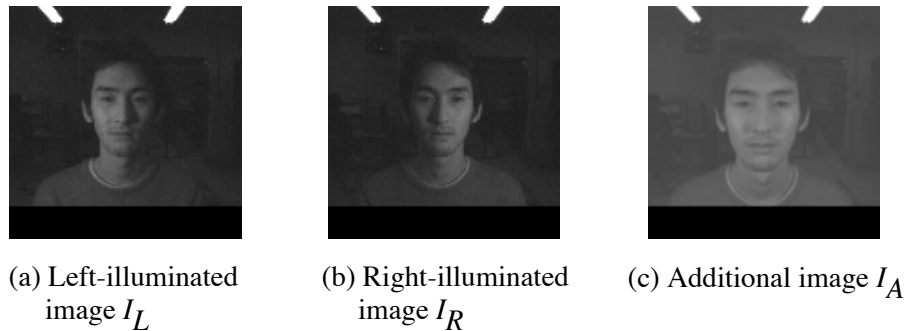


図 5.6: 左右照射画像と左右照度和画像

3. 湾曲度の閾値 c_{th}

湾曲度は左右に分かれた顔領域の間隔である。平坦写真を撮影したモデル画像3人分12枚に対して、2値化を行ったところ、左右の顔領域は少なくとも15画素離れていた。それに対し、顔領域の大きさが標準顔の1.2倍の人物モデル画像25人分を分析したところ、顔領域の左右の隔りは最大で5画素であった。従って、 c_{th} を5と設定した。

4. 凹凸度の閾値 b_{th}

凹凸度は加算投影 $P(x)$ から得られるグラフの谷の数に基づく。写真を湾曲させて撮影したモデル画像3人分12枚を分析したところ、全ての画像に対し、凹凸度は2であった。一方、25人分の標準顔の人物モデル画像を分析したところ、凹凸度は5以上であった。従って、 b_{th} を5と設定した。

5.5.3 顔領域検出の性能検証

顔領域の推定結果を図5.7に示す。人物、顔写真に関わらず、ほぼ正確に顔領域が推定されていることがわかる。次に、人物画像の顔領域の推定精度を表5.1に示す。標準顔、標準顔の80%の大きさ、標準顔の120%の大きさのテスト画像それぞれ50枚(テストセット100%, 80%, 120%)に対して実験を行った。顔領域の推定精度の検証には、3つの指標を用意した。それらは、位置誤差、大きさ誤差、顔検出率である。位置誤差は、手作業で抽出した顔領域の中心と推定された顔領域の中心の平均距離であ

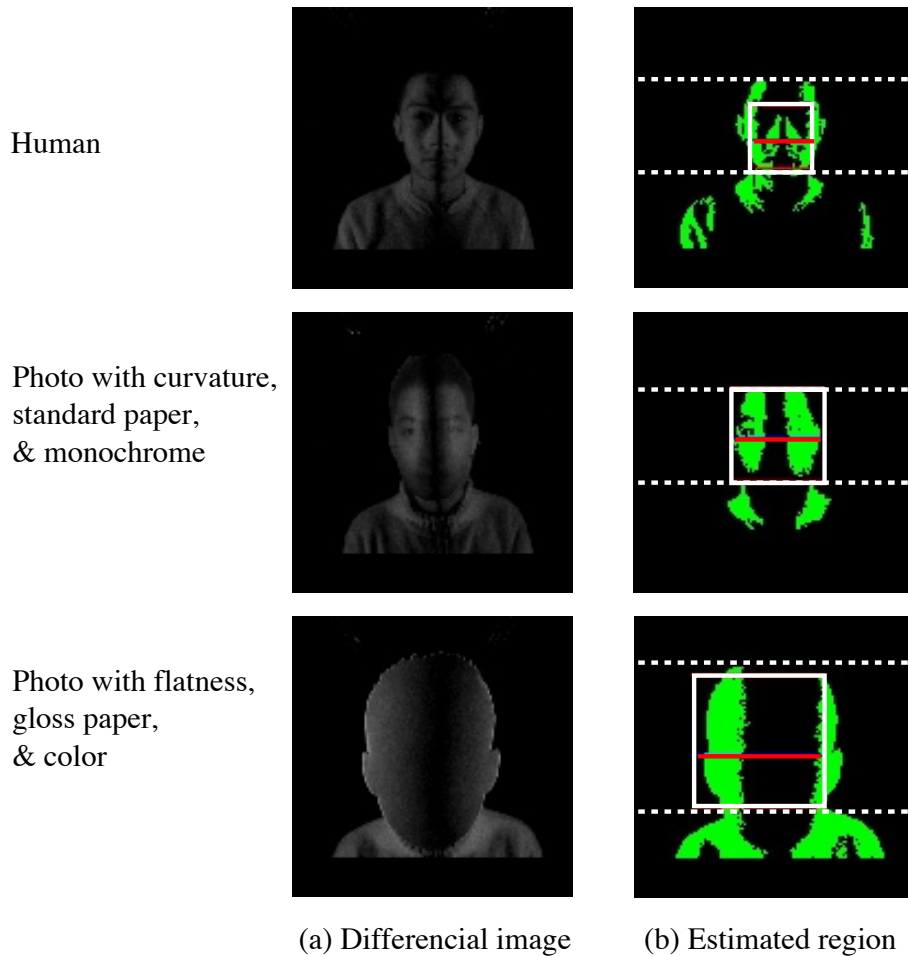


図 5.7: 顔領域の推定結果画像

る。大きさ誤差は、手作業で抽出した顔領域の幅と推定された顔領域の幅の比率の平均である。そして、顔検出率は、推定された顔領域内に目と口が完全に含まれる場合を検出成功として求めた平均成功率 [Yang 2002] である。表 5.1 より全ての指標において推定精度が良くないことがわかる。原因としては、髪の毛や眼鏡の影響が挙げられる。また、本システムが口の位置を検出せずに顔領域を推定するため、顔検出率が低下したと考えられる。

表 5.1: 人物写真判別部における顔領域の推定精度

Test set	Position error (pixel)	Scale error (%)	Detection rate (%)
80%	18.3	100±17.9	78.0
100%	17.6	100±20.6	82.0
120%	16.8	100±17.8	70.0

表 5.2: 人物写真判別の精度

Test set	FRR (%)	FAR (%)
80%	6.0	0
100%	6.0	0
120%	6.0	0

5.5.4 人物写真判別の性能検証

次に、人物写真判別の精度を、人物を顔写真であると判定したエラー率 FRR (False Rejection Rate) と、顔写真を人物であると判定したエラー率 FAR (False Acceptance Rate) から求めた。これらを表 5.2 に示す。一般的なセキュリティシステムでは、FAR が重視されるため、FAR が 0% である本システムの有効性は高く評価できる。FRR の対象となるエラーの原因としては、人物が眼鏡をかけた場合に起こる可能性が高い。人物写真判別部の処理時間は、全体で平均 27 ミリ秒であった。人物写真判別部の顔領域検出アルゴリズムは、非常に高い計算効率で顔領域を推定できるため、詳細な顔位置推定の前処理用として十分な性能を持つと言える。

5.5.5 1対多照合の性能検証

撮影された 50 人のうちの 25 人を登録人物とし、個人モデルデータベースにそれらのモデル画像とモデルグラフを登録した。モデルグラフの特徴点は著者が手作業で抽出した。残りの 25 人は未登録人物である。実験は人物を撮影した全てのテスト画像に対して行った。図 5.8 に個人識別部

による顔グラフ照合の結果画像を示す。この結果から、顔の位置と大きさがほぼ正確に推定されていることがわかる。また、個人モデルの推定によって最適なモデルグラフを選択しているため、顔の特徴点も正確に検出されていることがわかる。次に、顔の位置と大きさの推定精度および計算コストを表5.3に示す。位置の推定精度は、手作業で抽出した顔グラフの鼻の位置と推定された鼻の位置の平均誤差である。大きさの推定精度は、推定された顔の大きさの平均である。また、計算コストは、ガボールウェーブレット変換を行ったノード数である。顔の大きさの変化に伴って推定精度が低下していることがわかる。

そして、個人識別の精度を watch list 評価によって検証する。図5.9に watch list 評価の閾値 s に対する本人拒否率 (FRR) と他人受入率 (FAR) の変化を示す。なお、 $FRR(s) = 1 - TAR(s)$ 、 $FAR(s) = 1 - TRR(s)$ である。この曲線は、全てのテスト画像に対する結果の平均を示す。FARが0%の時にFRRは64%、FRRが12%の時にFARは44%であった。また、FRRとFARが等しくなるクロスポイント (EER: Equal Error Rate) は17%であった。アクセス制御などのセキュリティシステムは、一般的にユーザに3度の再試行を許可する。この場合のFRRは本実験で得られた値を3乗し、FARはそれを3倍することで推定される [瀬戸 1999]。これに従えば、EERは6.6%となる。顔認証を用いたセキュリティシステムに要求される認証精度がEERで1%程度とされている [JAISA]。本実験は一般的に運用実験と言われているものに比べ、撮影条件の変化が大きいと考えられるため、この実験結果が1%という精度に達しなかった可能性もある。さらに、運用実験ではシステム操作に対するユーザの慣れによって精度が向上することも考えられる。

さらに、個人識別精度をテストセットごとに求め、顔の大きさの変化への対応を検証する。図5.10にFRRとFARのROC曲線を示す。この曲線はグラフの左下に位置するほど、個人識別精度が良いことを示す。この結果から、本システムの個人識別精度が顔の大きさの変化に伴って低下していることがわかる。EERはテストセット80%で26%、テストセット100%で12%、テストセット120%で16%であった。この原因としては、カメラと人物の距離の変化に伴い、撮影された顔画像に多少の歪みが生じたことと、水平軸方向の顔の回転が多少生じたことが挙げられる。人物の表情が変化することも考慮して、顔の変形に対してロバストな認証処理を加えなければならない。

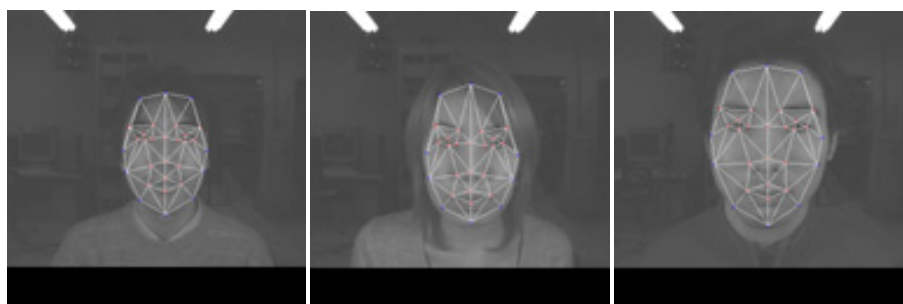


図 5.8: 顔グラフ照合の結果画像

表 5.3: 1 対多照合モードの顔の位置と大きさの推定精度

Test set	Position error (pixel)	Estimated scale (%)	Cost
80%	5.6	80±11.2	7203
100%	2.7	100±3.2	7443
120%	4.2	120±4.5	7429

5.5.6 1 対 1 照合の性能検証

1 対多照合実験で用いた 50 人の顔画像 150 枚で 1 対 1 照合実験を行った。照合の試行数は、FRR 用が登録人物の画像と本人のモデルのみとの照合を行うので 25×3 、FAR 用が未登録人物の画像と全ての登録人物のモデルとの照合を行うので $25 \times 25 \times 3$ であった。顔の位置と大きさの推定精度および計算コストを表 5.4 に示す。1 対多照合の結果と同様に、顔の大きさの変化に伴って推定精度が低下していることがわかる。

個人識別の精度は verification と呼ばれる評価方法 [Grother 2003] によって検証する。verification 評価では、FRR は watch list 評価と同様であるが、FAR は以下のように求める。

$$FAR(s) = \frac{|\{p_n : R < s\}|}{|\mathcal{P}_N \setminus \mathcal{G}|} \quad \forall p_n \in \mathcal{P}_N \quad (5.2)$$

\mathcal{G} は個人モデルデータベースである。図 5.11 に閾値 s に対する FRR と FAR の変化を示す。FAR が 0% の時に FRR は 43%、FRR が 0% の時に FAR は 97% であった。また、EER は 16% であった。ユーザに 3 度の再試行を許可する場合の EER は 2.9% となった。

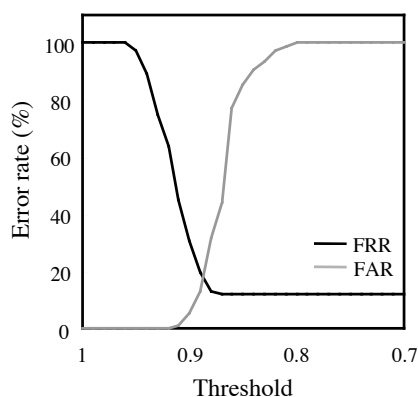


図 5.9: 1 対多照合モードの個人識別精度

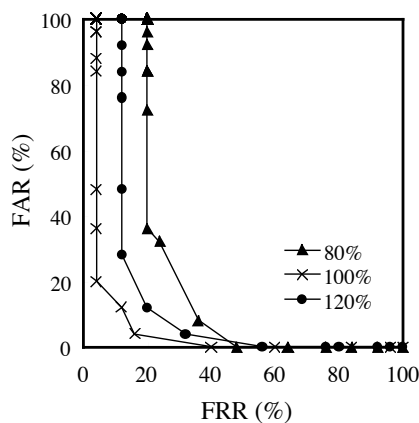


図 5.10: 1 対多照合モードのテストセットごとの個人識別精度

次に、個人識別精度をテストセットごとに求め、顔の大きさの変化への対応を検証する。図 5.12 に FRR と FAR の ROC 曲線を示す。この結果も 1 対多照合と同様に、個人識別精度が顔の大きさの変化に伴って低下していることがわかる。EER はテストセット 80% で 23%、テストセット 100% で 5%、テストセット 120% で 20% であった。

高い認証精度が期待できる 1 対 1 照合でも高い精度を得ることができなかったことから、パラメータの設定方法に問題がある可能性が考えられる。本実験のパラメータは、AR Face データベースを利用した実験結果 (第 3 章) から決定している。このデータベースのモデル顔画像と本実験で撮影したモデル顔画像の大きさが異なるため、ガボールウェーブレッ

表 5.4: 1 対 1 照合モードの顔の位置と大きさの推定精度

Test set	Position error (pixel)	Estimated scale (%)	Cost
80%	4.7	80±10.6	8204
100%	2.0	100±1.2	8654
120%	5.5	120±8.3	8077

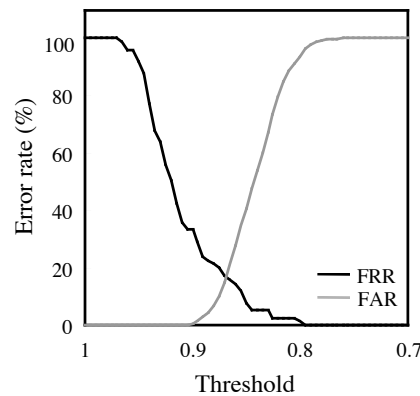


図 5.11: 1 対 1 照合モードの個人識別精度

トの基底サイズと基底周波数のパラメータ (a, σ, u) と顔位置推定のサンプリング点の間隔 d_{smp} が適当ではないと考えられる。また、赤外画像による個人識別が可視光画像より困難である可能性が挙げられる。ゆえに、個人識別に最適なガボールウェーブレットの基底サイズと基底周波数や赤外画像と可視光画像から抽出されるガボール特徴量の違いを分析する必要がある。さらに、個人識別に用いる照度和画像は、顔のテクスチャが劣化する可能性があり、それが個人識別精度を低下させていることが考えられる。これは、左右両方の照明を点灯させた状態で、もう 1 枚顔画像を撮影し、それを照度和画像の代わりに用いることで解決できる。なお、個人識別部の処理時間は平均で 33 秒で、実時間処理の実現というにはほど遠い結果となった。個人識別部のアルゴリズムが高効率であることは実証しているので、実装プログラムを改良しなければならない。

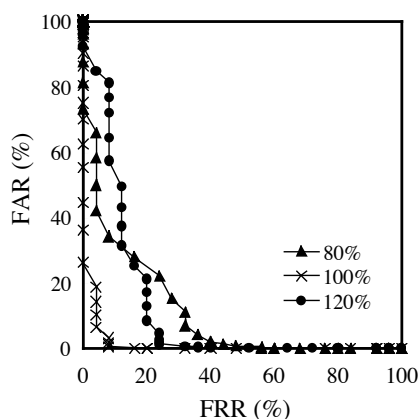


図 5.12: 1 対 1 照合モードのテストセットごとの個人識別精度

5.6 まとめ

セキュリティシステムに適した顔認証技術を導入したシステム FACELOCK を提案した。その技術は従来技術に比べ、顔の大きさの変化に対して認証精度を高く維持し、計算効率を向上させている点が特徴として挙げられる。ゆえに、本システムはユーザに顔画像撮影時の立ち位置を気にさせたくない状況や立ち位置の制約が困難な状況で最大限に性能が生かされる。また、従来システムには備えられていない人物と写真の判別機能を備える。しかし、性能評価実験の結果から、FACELOCK の有効性を十分に示すことができなかった。現状では、個人の入退室情報を管理せず、単に解錠することが求められるような状況 (negative identification) に適用できる可能性がある。この場合、認証順位が 1 位の人物がユーザであるかを判定する必要がなく、例えば、3 位までにユーザが認証されれば認証成功とすることができる。つまり、他人受入の判定を変化させずに本人拒否の判定を緩和させるため、EER を低下させることができる。今後、このような状況への適用を考慮し、認証判定に用いる順位と EER の関係を分析する必要がある。

第6章

結論

本論文では、セキュリティシステムへの応用を目的として、それに適した顔認証技術を提案した。そして、提案技術に基づく顔認証セキュリティシステム FACELOCK を開発した。

従来の顔画像照合研究のほとんどが、顔認識というタスクを意識したものであり、顔認証に関わる問題設定を十分に分析しなかった。本論文では、その問題設定を明確にし、最も重要でかつ基本的な問題である顔の位置と大きさの変化に注目した。顔の形状モデルを用いて顔位置を推定する場合、最適なモデルが用いられなければ、その精度は低下する。そこで、顔の位置と大きさと個人モデルの推定を統合した手法を第3章で提案した。本手法は、顔位置推定と個人識別を同時に行う手法とみなせ、このような手法は従来提案されなかった。標準顔画像データベースを用いた実験では、本手法が従来手法より性能が良く、顔認証の精度と計算コストのトレードオフを解決できることを示した。

また、従来の顔認証セキュリティシステムは顔写真による詐称に未対応であったため、人物と顔写真を判別する技術を第4章で提案した。この提案技術は、人物の顔と顔写真の違いが凹凸であることに注目し、2台の赤外光照明を用いた照度差ステレオ法によってその凹凸を観測する。従来研究では、詐称者が写真を湾曲させた顔マスクを装着するような状況が想定されなかったが、これにも対応できる。赤外光を利用することによって、環境の変化にロバストで計算効率の高い顔領域検出も実現できる。

そして、提案技術の実環境での運用性を検証するために、顔認証セキュリティシステム FACELOCK を構築し、第5章でその詳細を明らかにした。FACELOCK は、クライアントサーバ構成を適用したり、赤外光照明を利用するなど、運用性とセキュリティとユーザビリティを十分に考慮した

システムとなっている。評価実験によって、人物写真判別技術の有効性を高く評価できたものの、顔画像照合手法の有効性を十分に示すことができなかった。しかし、問題点を明らかにすることに至ったため、それを解決すれば実用化の段階まで発展する可能性が十分にある。顔認証システムの評価に関する実験環境や認証精度の定義等は、近年中に標準化される見込みである [坂野 2003]。今後、これに従って本システムの精度評価と他システムとの比較を行う必要がある。

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，素晴らしい研究環境を用意して頂きました大阪大学大学院基礎工学研究科谷内田正彦教授に心より深く感謝致します。全盛期が過ぎた顔研究に携わりたいという私の意志を快く受けとって頂きましたことに感謝致します。

本研究をまとめるにあたり，親切丁寧なご指導と数々のご教示を賜り，学位審査を行って頂きました大阪大学大学院基礎工学研究科西田正吾教授，同佐藤宏介教授，大阪大学産業科学研究所八木康史教授に心より深く感謝致します。

私の研究活動全般において，直接的で細部にいたる御指導と様々な御助言を賜りました岩井儀雄助教授に心より深く感謝致します。岩井助教授からは論文の添削や研究効率の大幅な向上につながる研究環境の構築など，非常に多くの御助力を賜りました。重ねてお礼申し上げます。

本研究を進めるにあたり様々な御指導と，研究活動の励みとなる御助言を賜りました砂山渡助手（現広島市立大学助教授），長原一助手に心より深く感謝致します。大西彩子事務補佐，中谷愛子事務補佐，沖合美紀事務補佐には出張の手配など，様々な御助力を賜りました。厚くお礼申し上げます。

Bruno Jean-Michel Lenski 氏，King-Foong Lew 氏は，私の未熟な指導に耐え，本論文に関わる成果をあげてくれました。ならびに，山添大丈氏，間下以大氏はじめ谷内田研究室で共に研究を行い，公私共にお世話になった方々に深くお礼申し上げます。皆様の研究に取り組む姿勢が励みになりました。そして，顔画像を撮影させて頂いた方々にお礼申し上げます。

最後に，精神的にも経済的にも支えてくれた両親に感謝致します。何も言わず温かく見守ってくれた姿が何よりの力となりました。それから，研究活動や日々の生活を支えて頂いた方々に心より深く感謝致します。博士学位を取得することは最終目標ではなく，通過点にすぎません。今後，よりいっそう研究活動に精進致しますので，御指導御鞭撻のほどよろしくお願い致します。

参考文献

- [Antonini 2003] G. Antonini, V. Popovici, and J. -P. Thiran,, “Independent component analysis and support vector machine for face feature extraction,” *Proceedings of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication*, pp.111–118, 2003.
- [Ayinde 2002] O. Ayinde and Y.-H. Yang, “Face recognition approach based on rank correlation of Gabor-filtered images,” *Pattern Recognition*, vol.35, no.6, pp.1275–1289, 2002.
- [Barr 1981] A. Barr and E. A. Feigenbaum: *The handbook of artificial intelligence*, Addison-Wesley, Menlo Park, CA, 1981.
- [Belhumeur 1997] P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.711–720, 1997.
- [Beveridge] Evaluation of face recognition algorithms,
<http://www.cs.colostate.edu/evalfacerec/index.html>
- [Chai 1998] D. Chai and K. N. Ngan, “Locating facial region of a head-and-shoulders color image,” *Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition*, pp.124–129, 1998.
- [Chen 2000] L. F. Chen, H. Y. M. Liao, M. T. Ko, J. C. Lin, and G. J. Yu, “A new LDA-based face recognition system which can solve the small sample size problem,” *Pattern Recognition*, vol.33, no.10, pp.1713–1726, 2000.
- [Dai 1996] Y. Dai and Y. Nakano, “Face-Texture model based on SGLD and its application in face detection in a color scene,” *Pattern Recognition*, vol.29, no.6, pp.1007–1017, 1996.

- [Daugman 1985] J. G. Daugman, “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters,” *Journal of the Optical Society of America A*, vol.2, pp.1160–1169, 1985.
- [Dowdall 2003] J. Dowdall, I. Pavlidis, and G. Bebis, “Face detection in the near-IR spectrum,” *Image and Vision Computing*, vol.21, no.7, pp.565–578, 2003.
- [Duda 2001] R. O. Duda, P. E. Hart, and D. G. Stork: *Pattern classification*, Wiley-interscience, 2001.
- [Eveland 2003] C. K. Eveland, D. A. Socolinsky, and L. B. Wolff, “Tracking human faces in infrared video,” *Image and Vision Computing*, vol.21, no.7, pp.579–590, 2003.
- [Féraud 2001] R. Féraud, O. J. Bernier, J. E. Viallet, and M. Collobert, “A fast and accurate face detector based on neural networks,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.23, no.1, pp.42–53, 2001.
- [Grother 2003] P. Grother, R. J. Micheals, and P. J. Phillips, “Face recognition vendor test 2002 performance metrics,” *Proceedings of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication*, pp.937–945, 2003.
- [Guo 2001] G. Guo, S. Z. Li, and K. L. Chan, “Support vector machines for face recognition,” *Image and Vision Computing*, vol.19, no.9–10, pp.631–638, 2001.
- [Jebara 1997] T. S. Jebara and A. Pentland, “Parameterized structure from motion for 3D adaptive feedback tracking of faces,” *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp.144–150, 1997.
- [Kato 2002] T. Kato, T. Kurata, and K. Sakaue, “VizWear-Active: distributed Monte Carlo face tracking for wearable active cameras,” *Proceedings of the 16th International Conference on Pattern Recognition*, vol.1, pp.395–400, 2002.

- [Kjeldsen 1996] R. Kjeldsen and J. Kender, "Finding skin in color images," *Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition*, pp.312–317, 1996.
- [Kohl 1993] J. Kohl, "The Kerberos network authentication service (V5)," *Internet Engineering Task Force (IETF)*, RFC, no.1510, 1993.
- [Konen 1994] W. Konen, T. Maurer, and C. von der Malsburg, "A fast dynamic link matching algorithm for invariant pattern recognition," *Neural Network*, vol.7, pp.1019–1030, 1994.
- [Kotropoulos 2000] C. Kotropoulos, A. Tefas, and I. Pitas, "Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions," *Pattern Recognition*, vol.33, no.12, pp.1935–1947, 2000.
- [Krüger 1997] N. Krüger, M. Pöttsch, and C. von der Malsburg, "Determination of face position and pose with a learned representation based on labelled graphs," *Image and Vision Computing*, vol.15, no.8, pp.665–673, 1997.
- [Lades 1993] M. Lades, J. C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Würtz, and W. Konen, "Distortion invariant object recognition in the Dynamic Link Architecture," *IEEE Transactions on Computers*, vol.42, pp.300–311, 1993.
- [Lades 1995] M. Lades: *Invariant object recognition based on dynamical links, robust to scaling, rotation and variation of illumination*, PhD. Thesis, Ruhr Universität, Bochum, Germany, 1995.
- [Lanitis 1997] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.743–756, 1997.
- [Y. Li 2004] Y. Li, S. G. Gong, J. Sherrah, and H. Liddell, "Support vector machine based multi-view face detection and recognition," *Image and Vision Computing*, vol.22, no.5, pp.413–427, 2004.

- [Z. Li 2004] Z. Li and X. Tang, "Bayesian face recognition using support vector machine and face clustering," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.374–380, 2004.
- [Liu 2002] C. Liu, H. -Y. Shum, and C. Zhang, "Hierarchical shape modeling for automatic face localization," *Proceedings of the 7th European Conference on Computer Vision*, pp.687–703, 2002.
- [Liu 2003] C. Liu, "A Bayesian discriminating features method for face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.25, no.6, pp.725–740, 2003.
- [Lyons 2000] M. J. Lyons, J. Budynek, A. Plante, and S. Akamatsu, "Classifying facial attributes using a 2-D Gabor wavelet representation and discriminant analysis," *Proceedings of the 4th International Conference on Automatic Face and Gesture Recognition*, pp.202–207, 2000.
- [Martinez 1998] A. M. Martinez and R. Benavente, "The AR face database," CVC Technical Report24, 1998.
- [McKenna 1999] S. J. McKenna, Y. Raja, and S. Gong, "Tracking colour objects using adaptive mixture models," *Image and Vision Computing*, vol.17, no.3–4, pp.225–231, 1999.
- [Meng 2000] L. Meng, T. Q. Nguyen, and D. A. Castan, "An image-based Bayesian framework for face detection," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.302–307, 2000.
- [Mitsuyoshi 2003] T. Mitsuyoshi, Y. Yagi, and M. Yachida, "Real-time human feature acquisition and human tracking by omnidirectional image sensor," *Proceedings of IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp.258–263, 2003.
- [Moghaddam 1997] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.17, no.7, pp.696–710, 1997.
- [Okada 1998] K. Okada, J. Steffens, T. Maurer, H. Hong, E. Elagin, H. Neven, and C. von der Malsburg: *The Bochum/USC face recognition system:*

and how it fared in the FERET phaseIII test, In Face Recognition: From Theory to Applications, Springer-Verlag, Sterling, UK, 1998.

[Openssl] OpenSSL, <http://www.openssl.org/>

[Osuna 1997] E. Osuna, R. Freund, and F. Girosi, “Training support vector machines: an application to face detection,” *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.130–136, 1997.

[Othman 2003] H. Othman and T. Aboulnasr, “A separable low complexity 2D HMM with application to face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.25, no.10, pp.1229–1238, 2003.

[Pavlidis 2000] I. Pavlidis and P. Symosek, “The imaging issue in an automatic face disguise detection system,” *Proceedings of IEEE Workshop on Computer Vision beyond the Visible Spectrum: Methods and Applications*, pp.15–24, 2000.

[Penev 1996] P. Penev and J. Atick, “Local feature analysis: a general statistical theory for object representation,” *Neural Systems*, vol.7, no.3, pp.477–500, 1996.

[Phillips 2000] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, “The FERET evaluation methodology for face-recognition algorithms,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.22, no.10, pp.1090–1104, 2000.

[Pramadihanto 1998] D. Pramadihanto, Y. Iwai, and M. Yachida, “A flexible feature matching for automatic face and facial points detection,” *Proceedings of the 14th International Conference on Pattern Recognition*, pp.324–329, 1998.

[Pramadihanto 2001] D. Pramadihanto, Y. Iwai, and M. Yachida, “Integrated person identification and expression recognition from facial images,” *IEICE Transactions on Information and System*, vol.E84-D, no.7, pp.856–866, 2001.

- [Ranganath 1997] S. Ranganath and K. Arun, "Face recognition using transform features and neural networks," *Pattern Recognition*, vol.30, no.10, pp.1615–1622, 1997.
- [Rigney 1997] C. Rigney, A. Rubens, W. Simpson, and S. Willens, "Remote Authentication Dial In User Service (RADIUS)," *Internet Engineering Task Force (IETF)*, RFC, no.2138, 1997.
- [Rowley 1996] H. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.203–208, 1996.
- [Samaria 1994] F. Samaria and S. Young, "HMM-based architecture for face identification," *Image and Vision Computing*, vol.12, no.8, pp.475–544, 1994.
- [Schneiderman 1998] H. Schneiderman, "Probabilistic modeling of local appearance and spatial relationships for object recognition," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp.45–51, 1998.
- [Smeraldi 2000] F. Smeraldi, O. Carmona, and J. Bigün, "Saccadic search with Gabor features applied to eye detection and real-time head tracking," *Image and Vision Computing*, vol.18, no.4, pp.323–329, 2000.
- [Stenger 2003] B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla, "Filtering using a tree-based estimator," *Proc. of the 9th International Conference on Computer Vision*, vol.2, pp.1063–1070, 2003.
- [Sung 1998] K. -K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.20, no.1, pp.39–51, 1998.
- [Terrillon 1998] J. C. Terrillon, M. David, and S. Akamatsu, "Automatic detection of human faces in natural scene images by use of a skin color model and invariant moments," *Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition*, pp.112–117, 1998.
- [Turk 1991] M. Turk and A. Pentland, "Eigenface for recognition," *Journal of Cognitive Neuroscience*, vol.3, no.1, pp.71–86, 1991.

- [Viola 2001] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.511–518, 2001.
- [Wang 2000] J. Wang and T. Tan, "A new face detection method based on shape information," *Pattern Recognition Letters*, vol.21, no.6–7, pp.463–471, 2000.
- [Wang 2001] Y. Wang and B. Yuan, "A novel approach for face detection from color images under complex background," *Pattern Recognition*, vol.34, no.10, pp.1983–1992, 2001.
- [Wiskott 1995] L. Wiskott, J. M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition and gender determination," *Proceedings of the International Workshop on Automatic Face- and Gesture-Recognition*, pp.92–97, 1995.
- [Wiskott 1997] L. Wiskott, J. M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition by Elastic Bunch Graph Matching," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.775–779, 1997.
- [Wiskott 1999] L. Wiskott, J. M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, CRC Press, pp.355–396, 1999.
- [Würtz 1997] R. P. Würtz, "Object recognition robust under translation, deformation, and changes in background," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.769–774, 1997.
- [Yager 2000] R. Yager, "Intelligent control of the hierarchical agglomerative clustering process," *IEEE Transactions on System, Man, and Cybernetics - Part B: Cybernetics*, vol.30, no.6, pp.835–845, 2000.
- [Yang 2000] M. -H. Yang, N. Ahuja, and D. Kriegman, "Mixtures of linear subspace for face detection," *Proceedings of the 4th International Conference on Automatic Face and Gesture Recognition*, pp.70–76, 2000.

- [Yang 2002] M. -H. Yang, D. Kriegman, and N. Ahuja, “Detecting faces in images: a survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.24, no.1, pp.34–58, 2002.
- [Yow 1997] K. C. Yow and R. Cipolla, “Feature-based human face detection,” *Image and Vision Computing*, vol.15, no.9, pp.713–735, 1997.
- [Yuen 2002] P. C. Yuen and J. H. Lai, “Face representation using independent component analysis,” *Pattern Recognition*, vol.35, no.6, pp.1247–1257, 2002.
- [大久保 2000] 大久保 竜也, 安達 澄昭, 岩尾 博之: 顔認識技術を用いた入退室管理システム, *OMRON TECHNICS*, vol.40, no.3, pp.202–207, 2000.
- [呉 1997] 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 D-II*, vol.J80-D-II, no.7, pp.1774–1785, 1997.
- [坂野 2003] 坂野 鋭: 顔画像認証精度評価方法の標準化—世界の動向—, *情報処理学会研究報告 CVIM*, vol.2003, no.66, pp.41–44, 2003.
- [新本 1999] 新本 浩一, 本多 庸悟, 金子 俊一: 2光源照度差ステレオ法による肌表面形状の再構成, *電子情報通信学会論文誌 D-II*, vol.J82-D-II, no.1, pp.10–18, 1999.
- [赤外線技術研究会] 赤外線技術研究会 編: 赤外線工学—基礎と応用—, オーム社, 1991.
- [瀬戸 1999] 瀬戸 洋一, 三村 昌弘: バイオメトリックス認証技術における精度評価の動向, *情報処理*, vol.40, no.11, pp.1099–1103, 1999.
- [プラマディハント 1997] プラマディハント ダデット, 呉 海元, 谷内田 正彦: 多様な姿勢を有する入力顔画像からの個人識別, *電子情報通信学会論文誌 D-II*, vol.J80-D-II, no.8, pp.2232–2238, 1997.
- [津田 2004] 津田 剛宏, 山本 和彦, 加藤 邦人: 近赤外光による人物と写真との判別手法の提案, *電気学会論文誌 C*, vol.124, no.12, pp.2446–2453, 2004.

- [土居 1997] 土居 元紀, 陳 謙, 眞溪 歩, 大城 理, 佐藤 宏介, 千原 國宏: 顔画像照合による解錠制御システム, 電子情報通信学会論文誌 D-II, vol.J80-D-II, no.8, pp.2203–2208, 1997.
- [土橋 2002] 土橋 浩慶, 岡崎 彰夫, 高木 一圭: 顔照合セキュリティシステム FacePass™, 東芝レビュー, vol.57, no.8, pp.48–51, 2002.
- [中野 1999] 中野 宏毅, 山本 鎮男, 吉田 靖夫: ウェブレットによる信号処理と画像処理, 共立出版株式会社, 1999.
- [JAISA] 日本自動認識システム協会: <http://www.jaisa.or.jp/>
- [平山 2003] 平山 高嗣, 岩井 儀雄, 谷内田 正彦: 顔の大きさの変化にロバストな個人識別システム, 情報処理学会論文誌, vol.44, no.6, pp.1625–1634, 2003.
- [福井 1997] 福井 和広, 山口 修: 形状抽出とパターン照合の組合せによる顔特徴点抽出, 電子情報通信学会論文誌 D-II, vol.J80-D-II, no.8, pp.2170–2177, 1997.
- [福井 1999] 福井 和広, 山口 修, 鈴木 薫, 前田 賢一: 制約相互部分空間法を用いた環境変動にロバストな顔画像認識—照明変動の影響を抑える制約部分空間の学習—, 電子情報通信学会論文誌 D-II, vol.J82-D-II, no.4, pp.613–620, 1999.
- [藤田 1979] 藤田恒太郎: 人体解剖学, 南江堂, 1979.
- [萬上 2001] 萬上 圭太: 統計的な動きモデルを用いた人の姿勢推定, 大阪大学大学院基礎工学研究科修士学位論文, 2001.
- [森 1965] 森 於菟, 小川 県三, 大内 弘, 森 富: 解剖学, 金原出版, 1965.
- [リユー 2004] リユー キンフォン: 写真詐称に対応した赤外画像による顔検出システム, 大阪大学基礎工学部特別研究報告, 2004.

本研究に関する発表

A. 学術論文

1. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔の大きさの変化にロバストな個人識別システム, 情報処理学会論文誌, vol.44, no.6, pp.1625–1634, 2003.
2. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔画像認識を用いた施錠セキュリティシステム FACELOCK の開発, 電気学会論文誌 C, vol.124, no.3, pp.784–797, 2004.
3. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔認証のための顔位置推定と個人識別の統合, 電子情報通信学会論文誌 D-II, vol.J88-D-II, no.2, pp.1–15, 2005.

B. 国際会議

1. Takatsugu Hirayama, Yoshio Iwai, Masahiko Yachida, Person identification system robust for facial scale variants, IEEE Workshop on Cues in Communication, Kauai, USA, 2001.
2. Takatsugu Hirayama, Yoshio Iwai, Masahiko Yachida, Face recognition based on efficient facial scale estimation, 2nd International Workshop on Articulated Motion and Deformable Objects, pp.201–212, Palma de Mallorca, Spain, 2002.

3. Takatsugu Hirayama, Yoshio Iwai, Masahiko Yachida, Face recognition system using accurate and rapid estimation of facial position and scale, 4th International Conference on Audio- and Video-Based Biometric Person Authentication, Guildford, UK, pp.154–163, 2003.
4. Takatsugu Hirayama, Yoshio Iwai, Masahiko Yachida, Integrated estimation of facial scale and position, IEEE Conf. on Multisensor Fusion and Integration for Intelligent Systems, pp.222–227, Tokyo, 2003.
5. Takatsugu Hirayama, Yoshio Iwai, Masahiko Yachida, Parallelization between face localization and person identification, 6th Int'l Conf. on Automatic Face and Gesture Recognition, pp.183–188, Seoul, 2004.
6. Takatsugu Hirayama, Yoshio Iwai, Masahiko Yachida, Face authentication system with facial personation prevention by using infrared images, 2nd International Symposium on Systems and Human Science “Complex Systems for Safety, Security and Reliability”, San Francisco, 2005.

C. 国内会議 (研究会・シンポジウム)

1. 平山高嗣, 岩井儀雄, 谷内田正彦, スケール変化にロバストな個人識別システム, 電子情報通信学会技術研究報告 PRMU, vol.100, no.566, pp.15–22, 2001.
2. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔認識のためのスケール変化に応じた特徴量抽出, 情報処理学会研究報告 CVIM, vol.2001, no.66, pp.33–40, 2001.
3. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔位置と大きさの高速同時推定, 電子情報通信学会技術研究報告 PRMU, vol.102, no.318, pp.61–66, 2002.

4. 平山高嗣, 岩井儀雄, 谷内田正彦, ヒューマンコンピュータコミュニケーションのための顔認識システムの構築, 第3回 SICE システムインテグレーション部門講演会, pp.191-192, 2002.
5. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔画像認識を用いた施錠セキュリティシステム FACELOCK の開発, 電子情報通信学会技術研究報告 PRMU, vol.102, no.651, pp.7-12, 2003.
6. Takatsugu Hirayama, Yoshio Iwai, Masahiko Yachida, Face recognition system using efficient methods for facial scale variations, SICE Annual Conference 2003, pp.1360-1365, 2003.
7. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔位置検出と個人識別の並列化, 情報処理学会研究報告 CVIM, vol.2003, no.66, pp.1-8, 2003.
8. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔画像の大きさと個人性を考慮した顔位置推定, 情報処理学会研究報告 CVIM, vol.2003, no.88, pp.111-118, 2003.
9. 平山高嗣, 岩井儀雄, 谷内田正彦, 顔認証セキュリティシステム FACE-LOCK, 第4回 SICE システムインテグレーション部門講演会, pp.77-78, 2003.
10. 平山高嗣, リュー キンフォン, 岩井儀雄, 谷内田正彦, 写真詐称に対応した赤外画像による顔認証, 画像の認識・理解シンポジウム (MIRU2004), pp.703-708, 2004.
11. Takatsugu Hirayama, King-Foong Lew, Yoshio Iwai, Masahiko Yachida, Face detection and facial personation prevention by using infrared images, SICE Annual Conference 2004, pp.461-465, 2004.

-
12. 岩井儀雄, 平山高嗣, 谷内田正彦, 顔画像認識を利用した生体認証 API の開発, 第 5 回 SICE システムインテグレーション部門講演会, 2004.

付録A

顔の大きさの変化にロバストな個人識別システム

A.1 システム概要

システムの概略を図 A.1 に示す。本システムは、顔位置推定、顔特徴点抽出、個人識別から構成される。顔特徴点抽出は、スケール変換とフレキシブル特徴照合を順に行う。また、顔モデルデータベースには、顔の大きさがほぼ等しい人物のモデル画像を登録する。システム構築時には、モデル画像からモデルグラフと平均モデルグラフ、そしてスケール変換に用いるスケール辞書を作成する。なお、モデルグラフの各ノード位置は、鼻の頭に位置するノードを基準とした相対位置へと変換する。

A.2 顔位置推定

顔位置の推定には、参照モデルグラフとして平均モデルグラフを用いる。平均モデルグラフは、モデルグラフの各ノードが保持するガボール特徴量の平均と、ノード間距離の平均から作成される。顔位置の推定はまず、入力画像中を粗くラスタスキャンし、各位置で参照モデルグラフをマッピングしてシャドウグラフ G を生成する。シャドウグラフの妥当性は式 (A.1) に示すネットワーク応答によって求める。

$$O(G) = \sum_{m \in \mathbf{G}} (\|V_{m,j} - W_{m,j}\|) \quad (\text{A.1})$$

ここで、式 (A.1) の \mathbf{G} はグラフ中のノードを表す。 $V_{m,j}$ は参照モデルグラフのノード m において、大きさ j のガボールウェーブレットで抽出

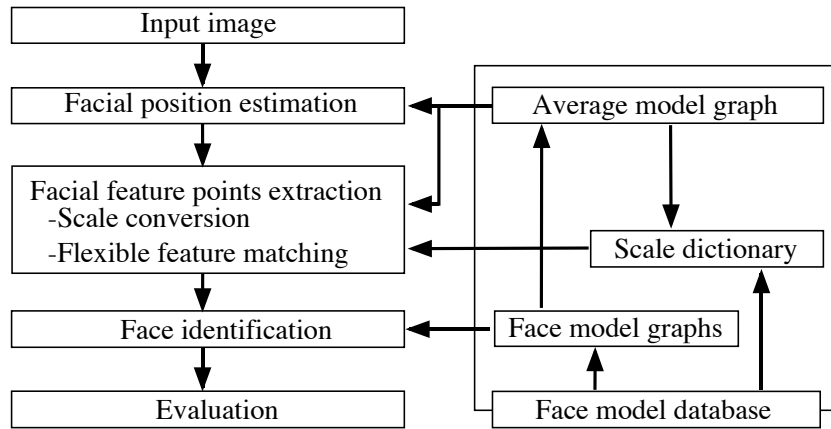


図 A.1: 個人識別システムの概略

されたガボール特徴量を要素とするベクトルであり， $W_{m,j}$ はシャドウグラフのそのベクトルである．この応答が小さいほど，シャドウグラフは参照モデルグラフとの全体的な調和を高める．このネットワーク応答が最小となるシャドウグラフの位置を顔の位置とみなす．この照合を，参照モデルグラフのノードの空間位置関係を固定して順次行うことから，固定照合と呼ぶ．顔位置推定により，顔特徴点の大まかな位置も得ることができる．これらの位置は，正確な特徴点の位置を求めるために，顔特徴点抽出の初期位置として利用する．つまり，推定した顔位置のシャドウグラフにスケール変換を適用する．

A.3 フレキシブル特徴照合法

フレキシブル特徴照合はまず，スケール変換後のシャドウグラフの各ノードをその近傍にランダムに移動させること（ランダム拡散過程）でシャドウグラフを変形させ，新しいシャドウグラフの候補を生成する．シャドウグラフの1つのノードをその近傍にランダムに移動させている様子を図A.2に示す．新しいシャドウグラフにはこの候補の内，式(A.2)のネットワーク応答が最も小さいものを採用する．この処理をシミュレーテッド・アニーリング法によって繰り返し，シャドウグラフと入力画像の全体的な調和を高めていく．

$$O_{flex}(G) = \sum_{m \in \mathbf{G}} \left[\left(\|V_{m,j} - W_{m,j \times s_{opt}}\| \right) + \lambda \sum_{n \in \mathbf{G}-m} \left(\|v_{mn} - \frac{w_{mn}}{s_{opt}}\| \right) \right] \quad (\text{A.2})$$

ここで、式 (A.2) の $j \times s_{opt}$ は大きさ j のガボールウェーブレットを s_{opt} で拡大縮小させたものを表す。これにより、参照グラフとシャドウグラフが持つガボール特徴量が正規化される。 v_{mn} は参照グラフの、 w_{mn} はシャドウグラフのノード m と n の距離情報である。距離情報 w_{mn} も s_{opt} によって正規化される。また、 λ はガボール特徴量と距離情報のどちらを重視するかを決める係数である。 λ は式 (A.3) に従って、照合の繰り返し処理が進むにつれて単調減少させる。 λ_{ini} は照合の初期段階のノード移動距離、 λ_{fin} は最終段階のノード移動距離、 τ_{max} は最大の繰り返し数である。なお、 τ は現在の繰り返し数である。また、照合のノード移動距離も λ に従って単調減少する。

$$\lambda(\tau) = \lambda_{ini} \left(\frac{\lambda_{fin}}{\lambda_{ini}} \right)^{\frac{\tau}{\tau_{max}}} \quad (\text{A.3})$$

なお、ガボール特徴量は、シャドウグラフの候補が生成される毎にそのノード位置でのみガボールウェーブレット変換が施されて、抽出される。また、ノードが顔の器官の空間位置関係を無視した位置に収束することを防ぐため、ノードの移動に簡単な制約を持たせる。この制約には一般的に知られている顔器官の構成知識 [森 1965, 藤田 1979] を利用している。フレキシブル特徴照合は、空間位置関係を考慮した特徴照合を行うので、ひげや眼鏡等を身に付けていても、特徴点の抽出が可能である。

A.4 個人識別

顔特徴点抽出で得られた最終的なシャドウグラフは、入力画像の顔の表現として利用できる。個人識別は、そのシャドウグラフと顔モデルデータベースとを比較することで行える。

同一人物の顔かどうかの判定には、式 (A.2) で定義されるネットワーク応答を用いる。すなわち、シャドウグラフとモデルグラフのネットワーク応答が最小であるモデルを入力画像の顔の人物であると認識する。ただし、応答があらかじめ設定されている閾値以下でなければ、システムは

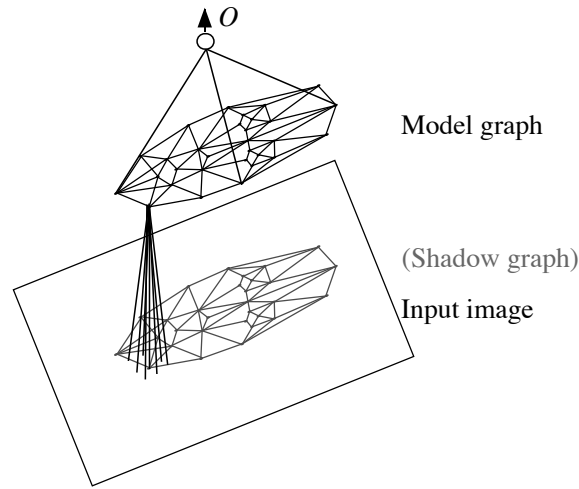


図 A.2: ランダム拡散によるシャドウグラフの生成

データベースに登録されている人物とは判定しないとする。つまり，watch list 評価に基づく。なお，個人識別処理での式 (A.2) の $V_{m,j}$ はモデルグラフのノード m において，大きさ j のガボールウェーブレットにより抽出されたガボール特徴量を要素とするベクトル， v_{mn} はモデルグラフのノード m と n の距離情報である。

A.5 システム性能

提案システムの有効性を検証するための評価実験について述べる。評価実験システムの実装は SGI 製ワークステーション Origin300 (CPU R14000, 500 MHz, memory 4 GB, OS IRIX6.5.16m) 上で行った。画像は Purdue 大学の AR Face データベース [Martinez 1998] のもの (256 × 256 画素, 8bit 濃淡) を利用した。モデルデータベースには，ほぼ等しい大きさの 50 人の顔画像とそれぞれのモデルグラフを登録した。画像は 1 人につき 1 枚で，全てが無表情の正面顔である。モデルグラフの特徴点位置は著者が目視によって顔画像中で定義した。入力画像は，モデルデータベースに登録された人物のモデル画像，登録されていない人物の画像がそれぞれ 50 枚と，それらを 80%, 90%, 110%, 120% に伸縮した画像の計 500 枚とした。画像の伸縮は MacOS X 上で Adobe Photoshop 7.0 を用いて bi-cubic 内挿法で行った。スケール辞書は，モデルデータベース中の 20 枚 ($= N_s$,

表 A.1: モデルと入力画像のガボール特徴量の差

	80%	90%	110%	120%
equation(3.2)	0.53	0.55	0.50	0.48
equation(3.1)	0.74	0.57	0.58	0.66

) の画像と参照モデルグラフを 11 段階 ($= T$) に伸縮させたもの (平均モデルグラフを 1 として $s_T = 1/1.5, 1/1.4, 1/1.3, 1/1.2, 1/1.1, 1, 1/0.9, 1/0.8, 1/0.7, 1/0.6, 1/0.5$ で伸縮) からそれぞれ求めたネットワーク応答より作成した. また, ガボールウェーブレット変換 (式 (3.2)) は, σ を 1, a を 2.5, u を π , ウェーブレットの大きさの段階数 N_j を 2, 方向数 N_θ を 8 とし, 総数 16 個のウェーブレットを用いて行った. 顔位置推定は, x 方向, y 方向ともに 5 画素間隔でラスタスキャンした. 式 (A.3) については, λ_{ini} を 4, λ_{fin} を 1, τ_{max} を 200 とした.

本システムの比較対象として, 2 つのシステムを構築した. 1 つ目のシステムは, 11 段階の大きさの参照モデルグラフ (平均モデルグラフを 50% から 150% まで 10% 刻みで伸縮したもの) と入力画像を照合し, 最も適合する参照モデルグラフを従来のフレキシブル特徴照合法に用いるもので, 従来システム 1 と呼ぶことにする. 2 つ目のシステムは, 利用する参照モデルグラフは 1 つで, 入力画像をこの参照モデルグラフの大きさに正規化するものである. 我々が構築したシステムは, あらかじめ入力画像を 11 段階 (1/50%, 1/60%, 1/70%, 1/80%, 1/90%, 1/100%, 1/110%, 1/120%, 1/130%, 1/140%, 1/150%) に伸縮したものを用意しておき, それらと参照モデルグラフを照合して, 最も適合する伸縮画像を従来のフレキシブル特徴照合法に用いるというもので, 従来システム 2 と呼ぶことにする. このシステムは, 他の 2 つのシステムより顔位置推定の計算コストが約 11 倍必要であり, また, 画像の伸縮を行うための計算コストも必要になる.

A.5.1 特徴点抽出の精度

図 A.3 に本システムを用いて顔特徴点を抽出した結果を示す. この結果から, 特徴量を正規化しているにも関わらず, 抽出された特徴点の位置が入力画像間で一致しないことがわかる. この原因を表 A.1 から考察する.

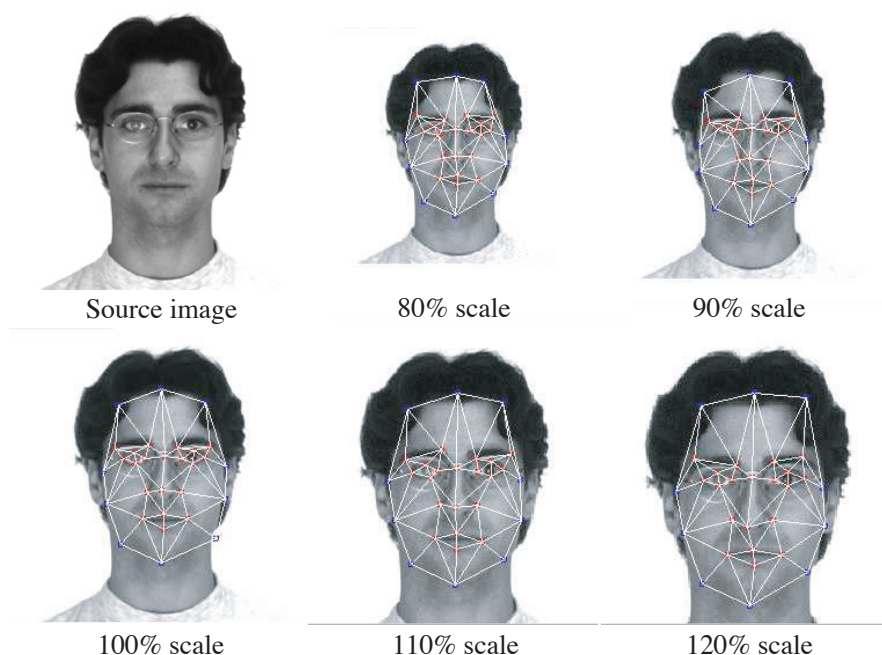


図 A.3: 顔特徴点抽出結果

表 A.1 はモデルグラフが持つガボール特徴量とそのモデルに対応した入力画像 (80, 90, 110, 120%伸縮) から式 (3.1) と式 (3.2) によって抽出される特徴量の差の平均を示す. なお, 式 (3.2) の j には画像伸縮率をかける. 表 A.1 の上段を見ると値がほぼ同じとなっているため, 式 (3.2) を用いることで顔画像の大きさの変化による空間周波数の変動に影響を受けないガボール特徴量が抽出されていることが確認できる. しかし, 値が 0 になっていないことから, 完全には正規化されていないこともわかる. これは画像の伸縮による情報落ちが影響していると考えられる. そして, この情報落ちがそれぞれの大きさの入力画像で抽出される特徴点の位置が一致しない原因となっていると考えられる. また, この情報落ちの影響は, 画像の伸縮によって顔の大きさを正規化する従来システム 2 の特徴点抽出の精度と個人識別率が低下する可能性を示唆する.

特徴点抽出の精度は, 顔特徴点の抽出処理が終了した時点のシャドウ

表 A.2: 特徴点抽出の精度 (画素)

	80%	90%	100%	110%	120%
Proposed system	4.95	5.17	4.61	4.66	4.70
Previous system 1	5.94	5.43	4.77	5.20	5.13
Previous system 2	5.69	5.50	5.11	5.64	6.13

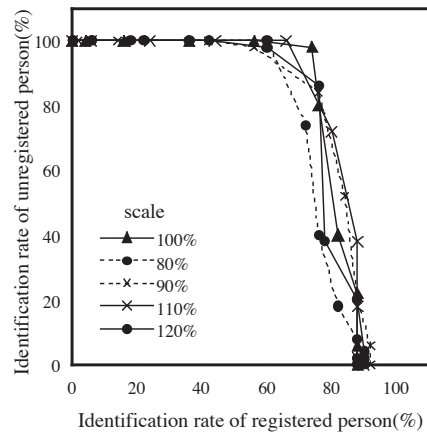
グラフと、入力画像の人物に該当するモデルグラフの特徴点の位置誤差を指標とする。すなわち、入力画像はモデルデータベースに登録されている人物のものである。表A.2に3つのシステムの特徴点抽出の精度を示す。表中の80%、90%、100%、110%、120%は入力画像の伸縮率を、数値は1特徴点あたりの平均誤差距離を表す。なお、精度の計算の際、本システムと従来システム1に関しては式(A.2)と同様の距離情報に対する正規化を行う。この結果から、従来システムでは本システムより、入力画像の伸縮が大きくなるほど抽出誤差も大きくなる傾向があるとわかる。これは、従来システム1ではガボール特徴量の正規化を行っていないこと、従来システム2では画像の伸縮による情報落ちが原因であると考えられる。一方、本システムでは顔の大きさの変化にあまり影響を受けずに特徴点を抽出できていることがわかる。なお、顔の大きさの変化に対する処理が適用されないはずの伸縮率100%の入力画像(つまり伸縮なしの入力画像)に対し、3つのシステムの精度に差があるが、これは伸縮率100%の入力画像の顔の大きさが完全に一様ではないため、これらの画像にも顔の大きさの変化に対する処理が適用されるからである。

A.5.2 個人識別の精度

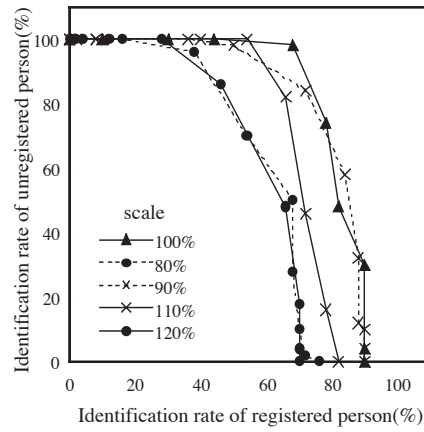
図A.4に3つのシステムの個人識別精度を示す。この図は、watch list 評価で設定する閾値を変化させ、それに対する登録人物の個人識別率と未登録人物の識別率をROC曲線の形で表したものである。この曲線はグラフの右上に位置するほど、個人識別精度が良いといえる。従って、元画像を伸縮させた入力画像(80%、90%、110%、120%)に対して、本システムの精度向上が確認できる。従来システムにおける精度の低下の原因は、A.5.1で述べたものと同様であると考えられる。

本システムによって入力画像の顔の大きさが変化する場合にも効率良

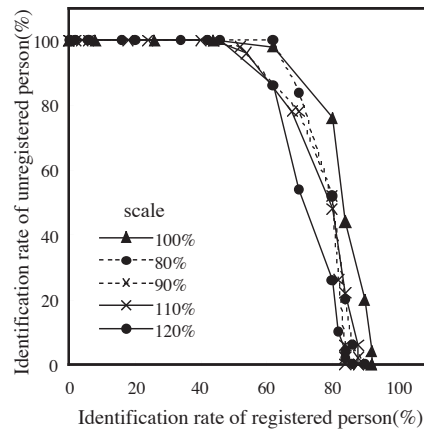
く、かつ精度良く顔特徴点が抽出され、個人識別の精度の低下が抑えられることを確認した。また、実験結果から入力画像の顔の大きさの変化に対して、入力画像自体の大きさを正規化するのではなく、ガボールウェーブレットの大きさを正規化して特徴量を抽出することが有効であると明らかになった。さらに、従来は5段階の大きさのガボールウェーブレットを利用して顔の大きさの変化に対応させるのが一般的であったが、本システムはそれを2段階に削減しても認識精度を維持することができた。ガボールウェーブレット変換の計算コストは非常に高く、ウェーブレットの数に比例するため、この削減が計算効率の面に関して非常に有効と思われる。



(a) Proposed system



(b) Previous system 1



(c) Previous system 2

図 A.4: 個人識別精度

付録B

顔認証システム用通信プロトコル SFAP

B.1 プロトコルスタック

顔認証システム用通信プロトコル SFAP (Simple Face Authentication Protocol) のプロトコルスタックを図 B.1 に示す。SFAP は下位レイヤで、SSL か TLS 相当の暗号機能を備えているものとする。なお、SFAP には DoS 攻撃に対する対策も必要であるが、特に触れないものとする。

Applications
FAP
SSL/TLS
TCP
IP

図 B.1: プロトコルスタック

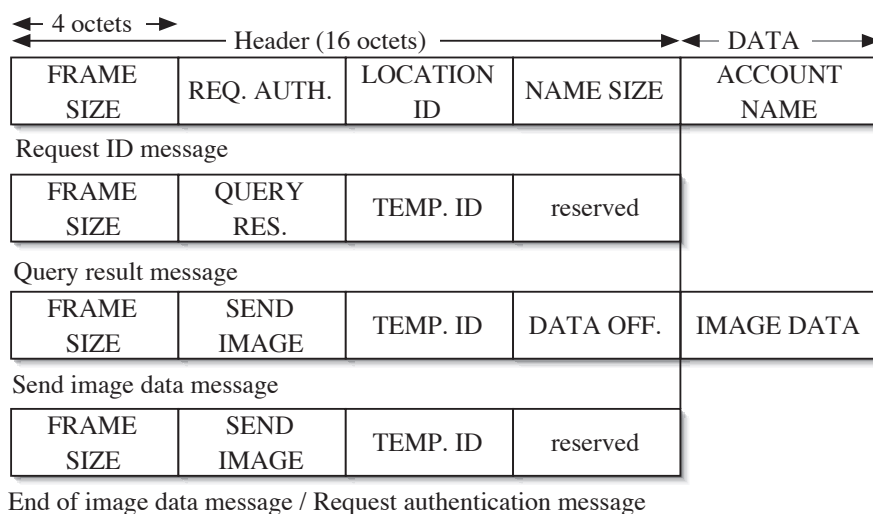


図 B.2: メッセージフォーマット

B.2 データ表現

最近では、相互運用性を考慮して XML 形式でデータが表現されつつあるが、SFAP ではフレーム形式を利用する。図 B.2 にフレームを示す。データ表現は単純で、サーバとのやりとりには固定サイズのヘッダと可変長のサイズのデータからなる 1 つのフレーム形式で行う。

以下、簡単に各フィールドについて説明する。

FRAMESIZE メッセージフレームのサイズ (octets) を示す。

REQ. AUTH. サーバへの接続許可依頼を行う。

LOCATION ID 操作したい鍵の場所を示す。このフィールドにより 1 台のクライアントにつき複数の鍵が制御可能になる。

TEMP. ID 一時利用 ID を示す。サーバへ接続すると一時利用 ID が発行される。以後のメッセージはこの ID を利用する。

NAME SIZE アカウント名の長さ

ACCOUNT NAME アカウント名

QUERY RES. サーバへ認証判定結果を要求する。

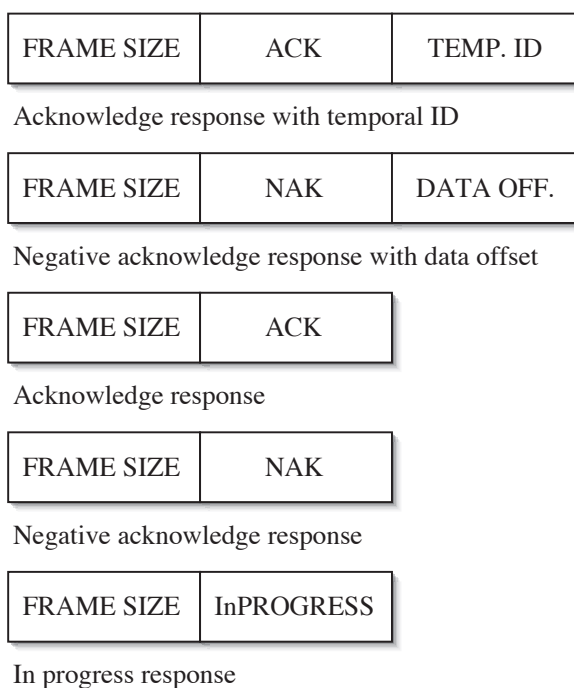


図 B.3: 応答メッセージのフォーマット

SEND IMAGE サーバへ画像を転送する。

DATA OFF. 画像データのオフセット

IMAGE DATA 画像部分データ

次に応答メッセージのフォーマットを図 B.3 に示す。Acknowledge response with temporal ID は、REQ. AUTH の肯定応答として、一時利用 ID を付加したものである。ACK/NAK はそれぞれ、通常の意味である。In-PROGRESS は、顔認証中に QUERY RESULT メッセージが送られてきたときに、返される応答である。これは、処理中であることを示す。この応答を受けたクライアントは、輻輳を避けるため乱数で決定された一定時間の経過後にもう一度 QUERY RESULT を発行する。

B.3 プロトコルの概略

図B.4に基本プロトコルのやりとりを示す。基本的なやりとりは4つの処理に分割することができる。

REQUEST AUTHENTICATION クライアントからサーバへ認証要求を出す。認証要求を受け取ったサーバはクライアントにSSLコネクションを開きサーバ利用者の許可を行う。この時点でサーバはLOCATION ID と接続要求元 IP ADDRESS からサーバ利用者を制限する事ができる。SSLのコネクションが成立した場合、一時利用IDをクライアントに発行し、以後、クライアントはその一時利用IDを使用する。

SEND IMAGE サーバ接続に成功したクライアントはサーバに認証用の画像をバースト転送する。画像は、一定サイズのフレームに分割されて転送される。途中のフレームでエラーが生じた場合、サーバ側はNAKとエラーのデータオフセットをクライアントに返す。

END OF IMAGE DATA クライアントはサーバに画像データ転送が終了したことを知らせる。このメッセージを受け取ったサーバは、画像データが正常に転送終了していればACKを返す。そうでなければ、NAKを返す。その後、サーバはforkし、画像の認証処理を子プロセスで実行する。

QUERY RESULT 転送終了したクライアントは一定時間経過後、認証処理の結果を要求する。認証が完了すればACKを、認証処理中ではInPROGRESSを、認証されなければNAKを返す。サーバが認証処理中であった場合、QUERY RESULTメッセージが大量発行されないように、さらに待ち時間を延ばしてクライアントは待機する。

B.4 プロトコルの実装

上に述べた基本プロトコルをUNIXシステム(Solaris8, IRIX6.5, Debian Linux)に実装した。サーバ、およびクライアントの状態遷移を図B.5に示す。サーバの実装方法にはIterative serverとConcurrent serverの2種類に大別される。Iterative serverはサーバの負荷が上がると、応答が悪くなるという欠点があるが、実装が容易であるという利点がある。そこで、

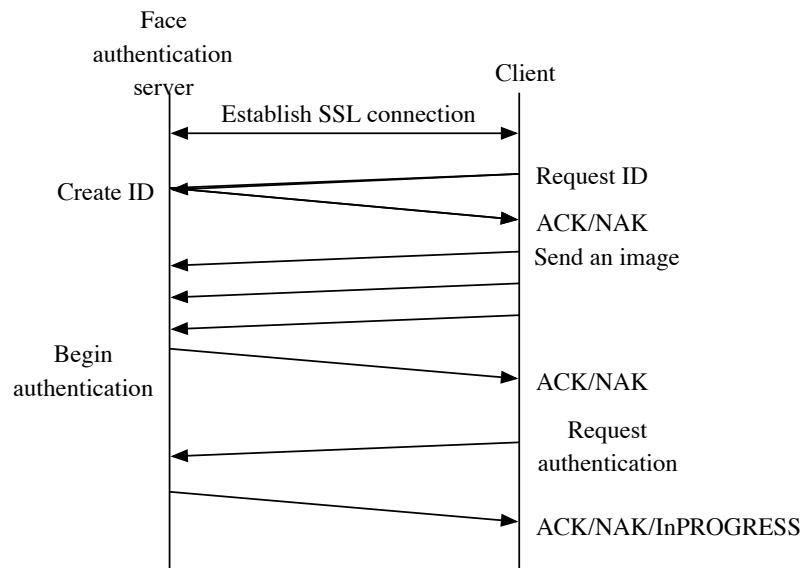


図 B.4: 基本プロトコルのやりとり

主要部分を Iterative server として FACELOCK の顔認証サーバを実装し、負荷の高い顔認証部を fork して並列処理するように実装している。

B.5 プロトコル評価のためのモデル化

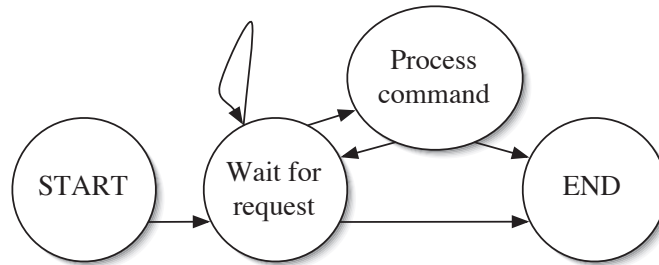
SFAPの性能を検証するために、各処理を次のようにモデル化する。Request authentication phase では、図 B.6 に示すようなやりとりが行われる。この phase の平均処理時間 T_R は、データ転送遅延平均時間 Δt と SSL 接続平均時間 T_{SSL} 、REQUEST 処理平均時間 T_{REQ} 、コマンド転送平均時間 T_{com} 、レスポンス転送平均時間 T_{res} から、以下のように与えられる。

$$T_R \equiv T_{SSL} + T_{REQ} + T_{com} + T_{res} + 2\Delta t \quad (\text{B.1})$$

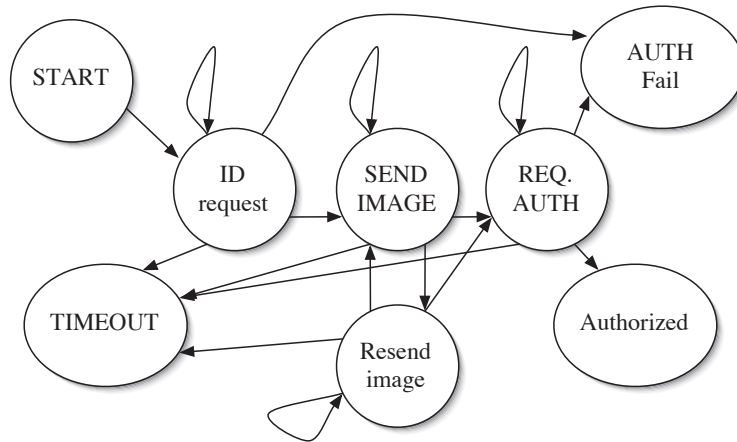
同様に、Query result phase の平均処理時間 T_Q は、QUERY 平均処理時間 T_{QUE} から以下のように与えられる (図 B.7)。

$$T_Q \equiv T_{QUE} + T_{com} + T_{res} + 2\Delta t \quad (\text{B.2})$$

ここで、コマンド転送平均時間、レスポンス転送平均時間は全コマンドで同一であるとしている。Send image phase の平均処理時間 T_I は、画像サイ



(a) サーバの状態遷移



(b) クライアントの状態遷移

図 B.5: SFAP の状態遷移

ズ N_I とフレームサイズ N_{frame} , 1 フレーム毎の平均処理時間 T_{IMG} , 最終フレームの平均処理時間 T_{CLS} により以下の式で与えられる. なお, 最終フレームだけ別にパラメータ化しているのは, サーバプロセスが fork して, 子プロセスに認証処理を実行させる時間が含まれているからである.

$$T_I \equiv \lfloor \frac{N_I}{N_{\text{frame}}} \rfloor T_{\text{IMG}} + T_{\text{CLS}} + 2\Delta t \quad (\text{B.3})$$

B.6 メッセージ最悪応答時間

最悪応答時間 T_{worst} は, 認証サーバが Iterative で実装されているので以下の式で与えられる.

$$T_{\text{worst}} = \max(T_R, T_Q, T_{\text{IMG}}) \quad (\text{B.4})$$

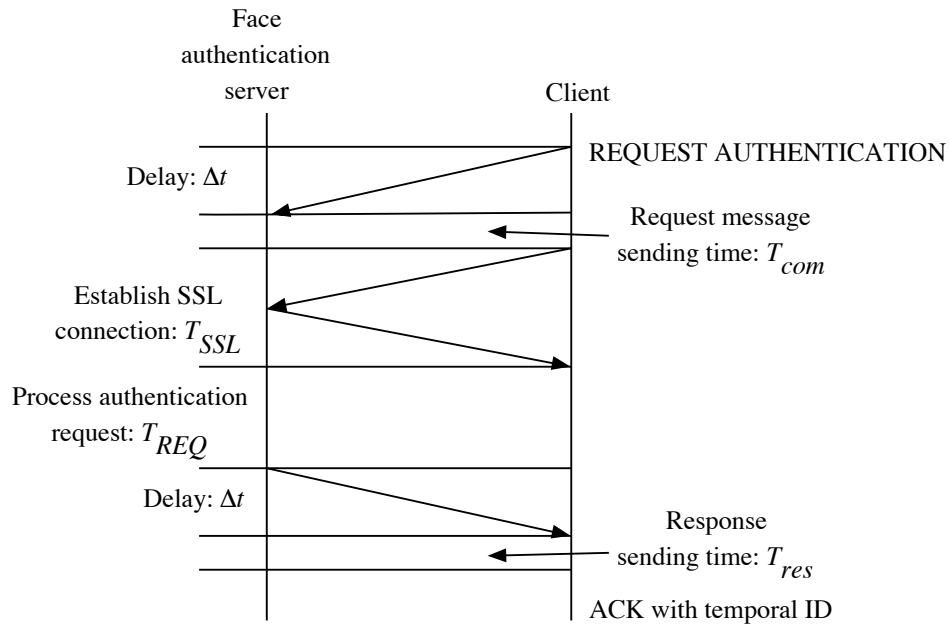


図 B.6: 認証要求フェーズのやりとり

B.7 平均スループット

平均スループット $N_{\text{throughput}}$ (件/秒) は, 平均認証処理時間 T_{rec} と認証処理多重度 M , 平均 QUERY RESULT 発行数 N_Q により以下の式で与えられる.

$$N_{\text{throughput}} = \frac{1}{\max(T_{\text{rec}}/M, T_R + T_I + N_Q T_Q)} \quad (\text{B.5})$$

B.8 SFAP の性能

サーバのパフォーマンスを検証するために, いくつかの実験を行った.

B.8.1 実験環境

実験にはスイッチングハブ (Cabletron 社製 Smart Switch SSR2000) に接続された 2 台のワークステーション (SGI 社製 Origin 300 8CPU 2GB Memory for Server, Origin 300 4CPU 1GB Memory for Client) を利用した.

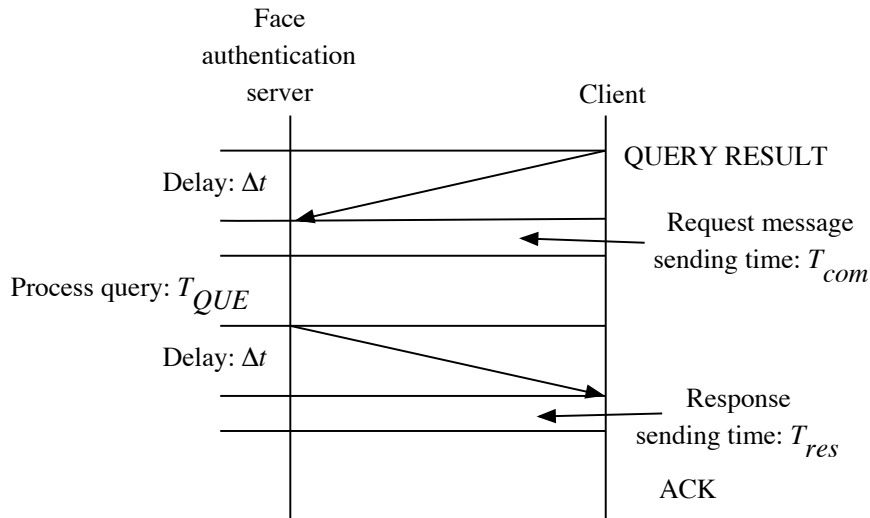


図 B.7: 結果照会フェーズのやりとり

表 B.1: データ伝送平均時間 (ms)

	Δt	T_R	T_Q	T_I	T_{rec}
Encryption	23.5	37.9	6.56	1666.2	26200
No encryption	1.53	32.9	4.57	1108.5	26200

転送に利用した画像のサイズは、512画素×512画素×8ビット、合計256KBである。また、ネットワークにはFast Ethernet (100Mbps)を利用した。

B.8.2 データ伝送平均時間

各メッセージの平均伝送時間を表B.1に示す。各データは10,000試行の平均である。SFAPはアプリケーションレイヤの通信プロトコルであるので伝送遅延は1ms～20ms程度とやや大きい。QueryやRequestのメッセージは、サイズが小さいので暗号化のオーバーヘッドは現れていないが、画像転送の際には500msほどのオーバーヘッドが生じている。画像転送サイズは256KBであるので、2ms/KB程度のオーバーヘッドが生じている。なお、暗号化ありの実験では、OpenSSL [Openssl]を利用した。

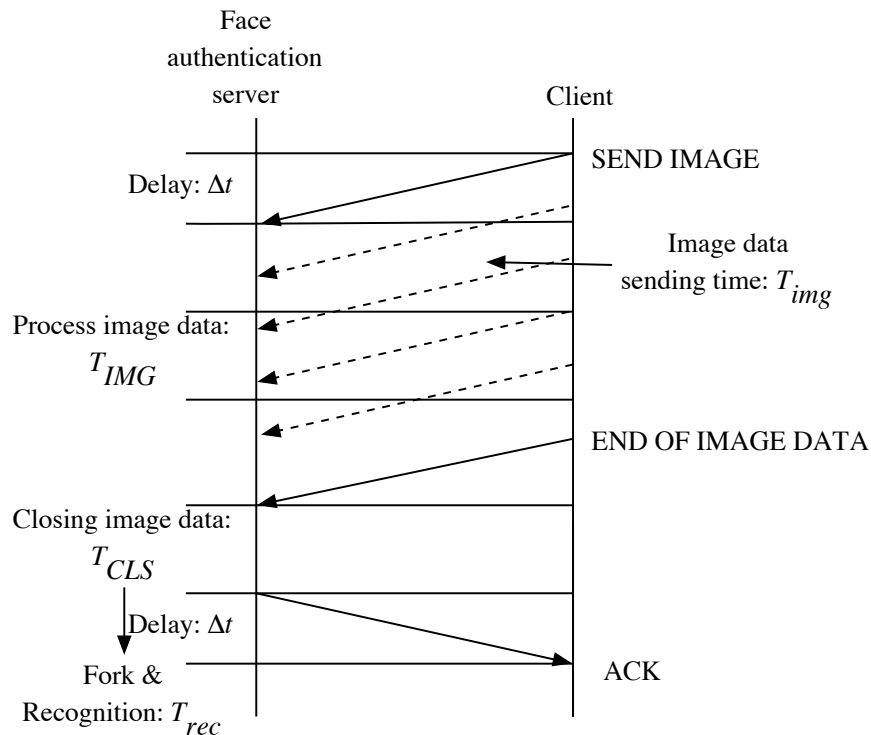


図 B.8: 画像データ転送フェーズのやりとり

B.8.3 スループット性能

サーバの多重度 M とスループットの関係を図 B.9 に示す. $T_{rec} = 26.2$ (ideal) の折れ線グラフは, 前節の実験データから計算された理論値である. $T_{rec} = 26.2$ の折れ線グラフは, 実際のサーバで行った実測値である. $T_{rec} = 26.2$ の理論値と良く一致していることが分かる. ネットワーク性能から最大スループットは 15 clients/min に制限されるが, 実際にはサーバの CPU 数である 8 で性能が頭打ちになる. 将来的には, 認証処理速度の向上が必要で, 改善の余地がある. また, 処理速度が向上した場合には, Fast Ethernet に接続されたサーバ 1 台あたり 15 クライアントを処理できる. 小グループのサーバとしては十分なパフォーマンスがあるが, 大企業などのサーバとしては能力が不足しているため, そのような大規模用途にはサーバをクラスタ化し, 負荷分散を行うか, 画像を圧縮して転送するなどの伝送帯域圧縮の工夫が必要である.

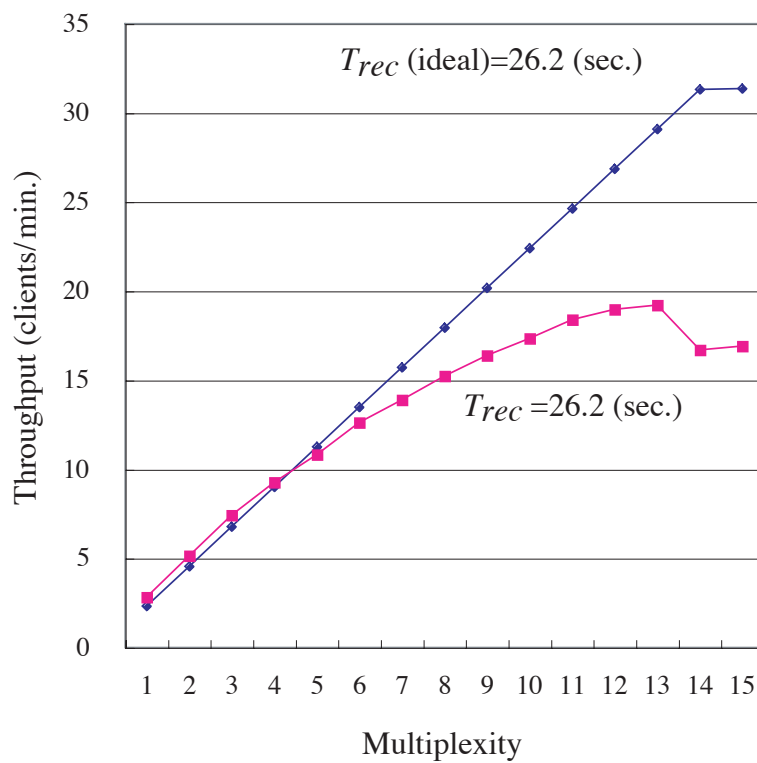


図 B.9: サーバスループット