

博士論文

拡張固有顔による
実時間顔追跡・認識融合系の高性能化

平成 28 年 3 月

中岸久佳

岡山大学大学院
自然科学研究科

要約

本論文では、拡張固有顔による実時間顔追跡・認識融合系の高性能化について述べる。固有顔は、照明変動などに起因する顔の見え方の変化を直接取り扱うことができるという利点から顔画像認識に広く用いられてきた。しかし、姿勢変動や表情変化などにより発生する顔の見え方の変化は顔の3次元形状に起因したものであるため、画像情報のみから構成される従来の固有顔では、これらの変化に対して頑健な顔認識系を構成することは容易でなかった。

この問題に対する1つの方法として、拡張固有顔と3次元疎固有テンプレート追跡系を組み合わせた実時間顔追跡・認識融合系が岡-尺長によって提案されている。ここで、拡張固有顔は、正規化固有顔に形状推定機能を持たせたものであり、形式的には正規化固有顔に画像情報から形状情報への連想写像を追加することにより定式化される。連想写像は、加重方程式と名付けた線形連立方程式に帰着できることが示されている。加重方程式は、入力画像を登録画像集合の最適な加重平均で表現するための線形連立方程式であり、この方程式を解くことで、形状推定や人物識別の処理が一元的に可能となる。また、3次元疎固有テンプレート追跡は、3次元疎テンプレート追跡系で用いるテンプレートを正規化固有顔に拡張した手法であり、固定形状を対象として見え方の変化にロバストな実時間顔追跡系を実現できる。さらに、正規化固有顔を、形状推定を可能とした拡張固有顔に置き換えることで、照明変動、表情変化、姿勢変動に頑健な系へ拡張することができる。

この方法において、加重方程式を用いた人物識別や形状推定は高速かつ頑健な系を構成するために重要な役割を果たす。ここで、岡-尺長の実時間顔追跡・認識融合系では、加重方程式を過剰決定系で解くことができるように系を構成していた。しかし、過剰決定系では、加重方程式を構成する際に登録者数よりも正規化固有顔の次元数を多く設定する必要があるため、登録者数が増大すると計算コストが大きくなり実時間性を維持することは容易でなくなるという問題があった。本論文では、この問題を解決するために加重方程式をベースとした新たな2つの加重計算法(1, 2)を提案する。また、これらの方法をもとに、顔追跡・認識融合系のさらなる高性能化を目指すための基礎検討(3, 4)を行った。以下にこれらの概要と関係をまとめる。

1. 階層化加重方程式による加重計算法の提案

岡-尺長の加重方程式では、登録処理を個人ごとに行っているため、登録者数が増大した場合に実時間性を維持することは容易でない。そこで、登録処理を登録人物集合のサブセットを用いて階層的に行うことで、登録者数が増加した場合においても、小規模な加重方程式を構成することが可能な方法を提案する。この方法では、サブセットの構成人物を入れ替えることで異なる階層化加重方程式を複数作成できる。本論文では、識別性能を改善するため複数の階層化加重方程式を組み合わせた加重計算法を提案する。

289人登録のデータベースを用いた静止画像実験と、100パターン登録(10人×10表情)のデータベースを用いた動画実験を行った結果から、この方法は、登録者数が増加した場合の実時間顔追跡・認識融合系の高性能化に有効な手法であることを確認した。

2. 並列不足決定系による加重計算法の提案

前項で述べた加重方程式の解法では、過剰決定系の範囲内で加重を求め、人物識別や形状推定に用いている。ここで、登録者数の増加を考えた場合、過剰決定系においては固有顔の次元数が系の構成にとって大きな制約となる。一方、不足決定系においてはこの制約がないため、原理的により柔軟に系を構成できると思われる。そこで、本論文では、もう一つの異なるアプローチとして、不足決定系の枠組みにおいて大規模化によって生じる問題を回避する方法を検討した。具体的には、単一の不足決定系加重方程式を構成して加重を求めるのではなく、与えられた固有顔の中に複数の不足決定系加重方程式を構成し、個々の不足決定系における最適解から平均値を求め、元の加重方程式の解の代わりとして用いる加重計算法を提案する。

289人登録のデータベースを用いた静止画像実験と、100パターン登録(10人×10表情)のデータベースを用いた動画実験を行った結果から、この方法は、登録者数が増加した場合の実時間顔追跡・認識融合系の高性能化に有効な手法であることを確認した。

3. 階層化加重方程式と並列不足決定系を用いた大規模識別の検討

本論文では、大量のシミュレーション画像で構成されるデータベースを用いて、加重方程式による識別系における登録者数を大規模化することを検討する。まず、2つの加重計算法を単独に用いて大規模識別系を構成し、基礎的な性能を評価する。次に、各大規模識別系の特性を分析した経緯から、2つの加重計算法の組み合わせによって大規模識別系を構成する。

シミュレーションデータベース(2197人登録)を用いた実験の結果から、並列不足決定系による候補選択の後に階層化加重方程式による識別を行う方法が最も有効であることを確認した。

4. 加重方程式を用いた顔モデリングの検討

顔追跡・認識融合系では、照明変動や形状変化に対応するために、あらかじめ登録者ごとに照明変動を含む学習画像や形状データを特定の環境下で収集しておく必要があり、データベースの拡張が容易でないという問題があった。この問題に対して、本論文では、不足決定系加重方程式をベースとした顔モデル生成法を検討した。これにより、正面向き画像1枚から必要なデータを推定し、顔モデルを生成できることを確認した。

以上の結果から，階層化加重方程式と並列不足決定系は実時間顔追跡・認識融合系の高性能化に有効であることが確認できた．また，大規模識別系の構成や登録処理の簡略化へも応用可能であり，今後の顔追跡・認識融合系のさらなる高性能化が期待できる．

目次

| | | |
|-------|---------------------------|----|
| 第1章 | 序論 | 1 |
| 1.1 | 研究背景 | 1 |
| 1.2 | 関連研究 | 1 |
| 1.3 | 本論文の構成 | 5 |
| 第2章 | 拡張固有顔を用いた3次元顔追跡・認識融合系 | 6 |
| 2.1 | 拡張固有顔による3次元顔情報処理 | 6 |
| 2.1.1 | 正規化顔モデル | 6 |
| 2.1.2 | 正規化固有顔 | 8 |
| 2.1.3 | 正規化固有顔への部分s-射影 | 9 |
| 2.1.4 | 画像から形状への連想写像 | 9 |
| 2.1.5 | 拡張固有顔 | 10 |
| 2.1.6 | 加重方程式 | 10 |
| 2.1.7 | 加重方程式による人物識別 | 11 |
| 2.1.8 | 照明適応を伴う加重方程式の実現 | 11 |
| 2.1.9 | 拡張固有顔と個人固有空間の例 | 14 |
| 2.2 | 3次元疎固有テンプレート追跡 | 17 |
| 2.2.1 | 3次元物体モデル | 17 |
| 2.2.2 | 姿勢空間と変換行列 | 17 |
| 2.2.3 | 画像生成過程とモデルベースレンダリング | 19 |
| 2.2.4 | 姿勢追跡問題 | 19 |
| 2.2.5 | 基準姿勢における固有テンプレートの作成 | 20 |
| 2.2.6 | 3次元疎固有テンプレートの作成 | 20 |
| 2.2.7 | 6次元姿勢空間における3次元疎固有テンプレート追跡 | 20 |
| 2.2.8 | 3次元疎固有テンプレートマッチング | 22 |
| 2.3 | 拡張固有顔による3次元疎固有テンプレート追跡 | 23 |
| 2.3.1 | 拡張固有顔による3次元疎固有テンプレート追跡の概要 | 23 |
| 2.3.2 | 3次元顔追跡・認識融合系の動作例 | 24 |
| 2.4 | まとめと問題点 | 26 |
| 第3章 | 階層化加重方程式による顔追跡・認識融合系の高性能化 | 28 |
| 3.1 | 階層化加重方程式による解法 | 28 |

| | | |
|-------|--------------------------|----|
| 3.1.1 | サブセット固有空間 | 28 |
| 3.1.2 | サブセット間加重方程式 | 29 |
| 3.1.3 | 個人固有空間 | 30 |
| 3.1.4 | サブセット内加重方程式 | 31 |
| 3.1.5 | 人物識別 | 32 |
| 3.1.6 | 階層化照明適応と階層化加重方程式 | 32 |
| 3.2 | 直交サブセットを用いた階層化加重方程式の並列化 | 32 |
| 3.2.1 | 直交サブセット分割の定義と解法 | 32 |
| 3.2.2 | 直交サブセット分割を用いた人物識別 | 33 |
| 3.2.3 | 直交サブセットと並列部分射影を用いた人物識別 | 35 |
| 3.2.4 | 計算コストの比較 | 35 |
| 3.3 | 静止画像を用いた性能評価 | 36 |
| 3.3.1 | データセットの仕様 | 36 |
| 3.3.2 | テストデータからの位置ずれデータの作成 | 38 |
| 3.3.3 | 比較した識別方法 | 38 |
| 3.3.4 | 階層化加重方程式による静止画像実験の結果 | 38 |
| 3.3.5 | 処理時間の比較 | 39 |
| 3.4 | 実動画像上での性能評価 | 40 |
| 3.4.1 | 学習データセットおよび拡張固有顔 | 40 |
| 3.4.2 | テストシーケンス | 42 |
| 3.4.3 | 比較した識別方法 | 42 |
| 3.4.4 | 動画像に対する識別判定 | 42 |
| 3.4.5 | 動画像実験の結果 | 43 |
| 3.5 | まとめ | 44 |
| 第4章 | 並列不足決定系による顔追跡・認識融合系の高性能化 | 46 |
| 4.1 | 並列不足決定系 | 46 |
| 4.1.1 | 不足決定系における加重方程式の解 | 46 |
| 4.1.2 | バイアス付き加重方程式 | 47 |
| 4.1.3 | 最近傍法との関係 | 47 |
| 4.1.4 | 並列不足決定系の構成 | 48 |
| 4.1.5 | 並列不足決定系における人物識別 | 49 |
| 4.2 | 静止画像を用いた基礎性能評価 | 49 |
| 4.2.1 | 実験条件 | 49 |
| 4.2.2 | 比較した識別方法 | 50 |
| 4.2.3 | 並列不足決定系による静止画像実験の結果 | 52 |
| 4.2.4 | 次元数と並列度の変化による識別率と処理時間の変化 | 52 |
| 4.3 | 実動画像上での性能評価 | 54 |
| 4.3.1 | 実験条件 | 54 |

| | | |
|------------|--|-----------|
| 4.3.2 | 比較した識別方法 | 55 |
| 4.3.3 | 動画像に対する識別結果 | 55 |
| 4.3.4 | 処理時間の比較 | 57 |
| 4.4 | まとめ | 58 |
| 第5章 | 階層化加重方程式と並列不足決定系による登録者数の大規模化 | 60 |
| 5.1 | 階層化加重方程式と並列不足決定系による大規模識別系の構成 | 60 |
| 5.2 | 大規模問題におけるシミュレーション実験系の構成 | 61 |
| 5.2.1 | 合成顔画像データベースの作成 | 62 |
| 5.2.2 | 正規化固有顔の構成 | 63 |
| 5.3 | 階層化加重方程式による大規模化 | 64 |
| 5.3.1 | 2階層加重方程式を用いた事前実験 | 64 |
| 5.3.2 | 3階層以上の加重方程式による大規模化 | 65 |
| 5.3.3 | 2階層加重方程式の2段構成による大規模化 | 68 |
| 5.4 | 並列不足決定系における大規模識別系の構成 | 68 |
| 5.4.1 | 直交サブスペース分割による並列化 | 68 |
| 5.4.2 | 直交サブスペース分割を用いた人物識別 | 68 |
| 5.4.3 | 並列不足決定系を用いた基礎性能比較 | 70 |
| 5.5 | シミュレーション画像を用いた大規模識別実験 | 71 |
| 5.5.1 | 実験の概要および実験条件 | 71 |
| 5.5.2 | 実験条件と比較した手法 | 72 |
| 5.5.3 | K^3 人登録における性能比較 | 72 |
| 5.5.4 | 計算量の比較 | 74 |
| 5.5.5 | さらに大規模な問題での絞り込み性能 | 76 |
| 5.5.6 | 2階層加重方程式と並列不足決定系の融合 | 77 |
| 5.6 | まとめ | 78 |
| 第6章 | 加重方程式を用いた顔モデリングによる登録処理の簡略化 | 80 |
| 6.1 | 加重方程式を用いた登録処理の簡略化 | 80 |
| 6.2 | 拡張基準固有顔 | 81 |
| 6.2.1 | 加重方程式による顔画像処理 | 81 |
| 6.2.2 | 基準固有顔 | 81 |
| 6.2.3 | 多重登録の場合の加重方程式 | 81 |
| 6.2.4 | 多重登録のバイアス付き加重方程式 | 82 |
| 6.2.5 | 拡張基準固有顔 | 82 |
| 6.3 | 拡張基準固有顔を用いた顔モデル生成法 | 83 |
| 6.3.1 | 線形変換を用いた照明変換 | 83 |
| 6.3.2 | 加重方程式を用いた照明変換 | 83 |
| 6.3.3 | 残差画像の利用 | 84 |

| | | |
|------------|-------------------------------|-----------|
| 6.3.4 | 複数枚を登録とした場合の作成方法 | 85 |
| 6.3.5 | 複数枚を登録とした場合の残差の取り扱い | 85 |
| 6.4 | 加重方程式による顔モデリング | 85 |
| 6.5 | 静止画像実験 | 86 |
| 6.5.1 | データセット | 86 |
| 6.5.2 | 実験条件 | 86 |
| 6.5.3 | 比較した手法と評価方法 | 86 |
| 6.5.4 | 静止画像実験結果 | 87 |
| 6.5.5 | 顔モデリングの結果 | 89 |
| 6.6 | 顔追跡・認識融合系への適用 | 91 |
| 6.7 | まとめ | 92 |
| 第7章 | 結論 | 95 |

目 次

| | | |
|------|-----------------------------------|----|
| 2.1 | 立体顔モデルの例 [17] | 7 |
| 2.2 | 個人固有空間と照明適応の例 | 12 |
| 2.3 | 並列部分射影に使用した部分領域 | 14 |
| 2.4 | 立体顔モデルの登録例 | 14 |
| 2.5 | AEF25x24 の作成に用いた登録顔画像と形状の例 | 15 |
| 2.6 | 拡張固有顔 AEF25x24 | 16 |
| 2.7 | 個人固有空間の例 | 16 |
| 2.8 | 姿勢パラメータ | 17 |
| 2.9 | 物体中心座標系とカメラ中心座標系 | 18 |
| 2.10 | 疎テンプレート追跡系の構成 [17] | 22 |
| 2.11 | 拡張固有顔による 3 次元疎固有テンプレート追跡系の構成 [17] | 23 |
| 2.12 | 3 次元疎固有テンプレートの初期値 [17] | 24 |
| 2.13 | 初期姿勢検出からの追跡・人物認識の例 [17] | 25 |
| 2.14 | 表情認識への適用例 [17] | 26 |
| 3.1 | サブセット固有空間とサブセット間加重方程式の例 | 29 |
| 3.2 | 個人固有空間とサブセット内加重方程式の例 | 31 |
| 3.3 | 直交するサブセット分割の例 (K=25 の場合) | 34 |
| 3.4 | CMU Multi-PIE(ID-002) の原画像と切り出し画像 | 36 |
| 3.5 | 正規化固有空間 EF289 | 37 |
| 3.6 | 位置ずれ画像の例 | 37 |
| 3.7 | Data10x10 の例 | 41 |
| 3.8 | 拡張固有顔 AEF10x10 の中心および最初の 4 軸 | 41 |
| 3.9 | 階層化加重方程式を用いた顔追跡・認識の例 | 44 |
| 4.1 | 次元数と並列度を変化させた場合の識別率の変化 | 53 |
| 4.2 | 次元数と並列度を変化させた場合の処理時間の変化 | 54 |
| 4.3 | 人物識別時間の変化 | 56 |
| 4.4 | 並列不足決定系による顔追跡・認識の例 | 57 |
| 4.5 | 登録者数が増加した場合の処理時間の変化 | 58 |
| 5.1 | Data100 の例 | 62 |
| 5.2 | 合成顔画像の作成手順 | 63 |

| | | |
|------|--|----|
| 5.3 | ある人物における学習データとテストデータ | 64 |
| 5.4 | 2階層加重方程式の2段構成 | 67 |
| 5.5 | 25次元正規化固有顔を用いた場合の直交サブスペース分割の例 . . . | 69 |
| 5.6 | 並列度の変化に対する2197人登録時の169位累積分類率 | 74 |
| 6.1 | 未登録人物の画像復元例と残差画像 | 84 |
| 6.2 | 入力画像の例 | 87 |
| 6.3 | テスト人物1の仮想照明画像の比較 | 88 |
| 6.4 | テスト人物5の仮想照明画像の比較 | 89 |
| 6.5 | 人物ごとの評価値 | 90 |
| 6.6 | 全ての照明条件における評価値の分布 | 91 |
| 6.7 | 各画像集合から作成された正規化固有顔の比較 | 92 |
| 6.8 | 実形状データ, 推定形状, 平均形状の比較 | 93 |
| 6.9 | 人物ごとのRMS残差 | 93 |
| 6.10 | 顔追跡・認識融合系への適用例 | 94 |

表 目 次

| | | |
|------|--|----|
| 3.1 | CMU Multi-PIE データベースでを使用した ID | 36 |
| 3.2 | 階層化加重方程式と基本加重方程式の識別率 (%) の比較 | 39 |
| 4.1 | 静止画像実験結果 (%) | 51 |
| 5.1 | 階層化による K^2 人登録時の識別率 (%) | 65 |
| 5.2 | 階層化による K^2 人登録時の K 位累積分類率 (%) | 66 |
| 5.3 | 階層化による K^3 人登録時の識別率 (%) | 67 |
| 5.4 | 階層化による K^3 人登録時の K 位累積分類率 (%) | 67 |
| 5.5 | 並列不足決定系による K^2 人登録時の識別率 (%) | 71 |
| 5.6 | 並列不足決定系による K^2 人登録時の K 位累積分類率 (%) | 71 |
| 5.7 | 大規模識別系 (K^3 人登録時) における識別率 (%) | 73 |
| 5.8 | 大規模識別系 (K^3 人登録時) における K 位累積分類率 (%) | 73 |
| 5.9 | 大規模識別系 (K^3 人登録時) における K^2 位累積分類率 (%) | 74 |
| 5.10 | 2 階層加重方程式と並列不足決定系の計算量の比較 | 75 |
| 5.11 | K^3 人登録時の処理時間 (msec) | 76 |
| 5.12 | 並列不足決定系による 169 位累積分類率 (%) と正解者の最低順位 | 76 |
| 5.13 | 大規模識別における単独系と融合系の計算量の比較 | 77 |
| 5.14 | 融合系での 2197 人登録時の識別率 (%) と 13 位累積分類率 (%) と 処理時間 (msec) | 78 |

第1章 序論

1.1 研究背景

顔追跡および顔認識はヒューマンインターフェースやセキュリティシステムへの需要から広く研究が行われてきた。この中で、見え方に基づく方法は、照明変動などによる顔の見え方の変化を直接取り扱うことができるという利点から広く用いられてきた。代表的な方法として固有空間を利用する方法があり、特に、顔画像集合から作成される固有空間は固有顔 [1] と呼ばれ、様々な応用手法が考案されてきた [2, 3, 4, 5, 6, 7, 8]。

固有顔は、顔画像集合を主成分分析することにより作成される。このとき、固有顔を照明条件など様々な条件下で撮影した顔画像集合から作成することにより、顔の見え方の変化に対して比較的安定した系を実現できる。一方、従来研究の多くでは2次元画像情報のみを用いて処理を行っているが、表情変化や姿勢変動などの顔の見え方の変化の多くは、顔の3次元形状に起因したものであるため、画像情報のみから構成される固有顔を用いるのではこれらの変化を適切に処理することが容易でない。

近年では、顔に関する様々な処理(追跡, 認識, 照明推定, 形状推定)を統一かつ高速に実現できる系が、実応用(特に, セキュリティシステム)の観点から需要が増している。このため、顔の様々な変化に対応するために、画像情報だけでなく形状情報も取り扱うことのできる実時間系を構成する必要がある。

1.2 関連研究

本節では、関連研究についてまとめる。

A. 固有顔とその応用

以下では、固有顔 [1] と固有顔をベースとした関連研究 [2, 3, 4, 5, 6, 7, 8] についてまとめる。

Turk-Pentland[1]により提案された固有顔は、固有空間をベースとした代表的な方法である。固有顔は、予め位置合わせを行った顔画像集合を用いて構成される。具体的には、顔画像集合から作成される共分散行列を主成分分析(PCA)もしくは特異値分解(SVD)することで得られ、このとき得られる固有軸を固有顔と呼び、

顔認識などに用いる．固有顔は，見え方に基づく手法であり，画像の変化を直接取り扱うことができるという利点から広く研究に用いられてきた．

Belhumeur et al.[2]では，固有顔と競合する手法として Fisherfaces が提案されている．この方法では，固有顔において PCA や SVD で求めていた軸を線形判別分析により求め，これらを顔識別に用いている．実験において，固有顔よりも照明変動や表情の変化に対して頑健であることが示されている．Moghaddam-Pentland [3]は，固有空間法に確率的な解釈を与えた方法であり，文献中には，この方法を用いた顔検出や顔認識などの応用例が報告されている．

Georghiades et al.[4]では，照明変動を取り扱う手法として Illumination cone が提案されている．この方法では，照明変動による見え方の変化は照明変動を含む顔画像集合から構成される凸錐で表現できることが示されている．Illumination cone は顔認識にも応用されており，照明変動に対して頑健であることが示されているが，学習には複数枚の照明条件の異なる画像が必要となる．Shakunaga-Shigenari [7]は，固有顔の直交分解と仮想固有投影により照明変動に頑健な顔認識を行う方法を提案している．この方法では，登録画像集合から構成される固有顔（基準固有顔）を用いて，顔画像を射影成分と個性およびノイズ成分に分解し，顔画像の標準的な情報を含む固有空間と個性およびノイズを含む固有空間を作成し，両者を顔認識に用いる．さらに，仮想固有投影により異なる照明条件の画像を仮想的に学習する方法も示されている．これにより，与えられた学習画像が少ない場合でも，照明変動に対して頑健な識別系を構成できることが示されている．

Vasilescu-Terzopoulos [5]では，テンソルを用いた顔認識法 (Tensorfaces) が提案されている．一般的な顔認識手法では，行列化した顔集合から固有顔を構成していたが，この方法では，テンソルを用いて人物，照明条件，表情，姿勢などの様々な顔の変化を解析している．Tensorfaces により，様々な変化を取り扱うことのできる顔認識系が構成できるが，学習量が多く，姿勢変動が微小な場合以外は挙動が安定しない．Wright et al. [6]では，スパース表現を用いた顔認識系が提案されている．スパース表現とは少数の基底の組み合わせで信号を表現する信号処理の考え方であり，文献中では，固有顔と組み合わせた顔認識系も報告されている．この方法は，表情，隠れやノイズなどに対して頑健な識別が可能であるが，安定動作のためには正確な顔画像切り出しが重要となる．

Matsubara-Shakunaga [8]では，固有顔をベースとした顔追跡系が提案されている．この方法は，疎テンプレート追跡法 (Matsubara-Shakunaga[9]) で用いられる疎テンプレートを照明変動を含む画像集合から作成される固有顔 (疎固有テンプレート) に置き換えた追跡手法である．疎固有テンプレートにより，物体の見え方の変化に頑健な顔追跡系を構成できるが，2次元画像をベースとした追跡手法であるため追跡中の顔の正確な姿勢情報を得ることはできない．

前述のとおり，これらの系は，あらかじめ位置決めされた2次元画像をベースとして構成されたものであるため表情や姿勢の変化など3次元形状に起因する見え方の変化に十分な精度で対応するのは容易でない．以下(節BおよびC)では，こ

これらの問題に対応するために、画像情報だけでなく形状情報も取り扱った関連研究についてまとめる。

B. 画像情報と形状情報を取り扱った関連研究 (静止画像ベース)

以下では、静止画像を対象とした関連研究についてまとめる。Blaž-Vetter[10]は画像情報と形状情報から独立に作成した固有空間を用いて、1枚の画像から形状と照明条件(反射特性)を同時に推定する方法を提案している。この方法は、画像情報と形状情報を独立に取り扱うことによって、形状と照明条件を推定できることを示している点が興味深い。この方法では、両者の間に独立性を仮定しているため、照明条件が異なる同一人物の画像間で同じ形状が推定される保証が無いという問題がある。さらに、推定に非線形最小二乗法を用いているため、計算量が膨大になり実時間処理には向いていない。

Castelan et al. [11]は画像情報と形状情報のそれぞれで作成した固有空間の間で、最適な線形写像を統計的に求め計算量を削減する方法を提案している。この方法では、線形写像による推定の効率化が可能であるが、照明条件を限定できない場合には画像と形状の間に線形写像を仮定できないという問題がある。

C. 画像情報と形状情報を取り扱った関連研究 (動画ベース)

ここで、動画を対象とした場合、系の安定動作のためには顔の位置情報の抽出(位置や姿勢)が問題となる。位置情報の推定手法としては大きく2つに分類される。1つ目は、顔検出器を用いた方法である。この方法では、顔の位置情報をフレームごとに呼び出される顔検出器の結果により顔切り出しを行い、得られた画像(列)を顔画像処理に用いる。代表的な顔検出器として、Viola-Jones [12]がよく知られている。この方法では、高速に計算可能な Haar-like 特徴を用いて学習された分類器で顔検出を行っており、この特徴を用いた顔認識系(Vaishak et al.[13])も報告されている。さらに、検出と動画(画像列)を組み合わせた顔画像処理系としては相互部分空間法(西山ら [14])がある。相互部分空間法では、フレームごとの特徴点検出器から得られる画像列から部分空間を作成し、作成された部分空間と登録者の部分空間との類似度を計算することで顔認識を行う。

一方で、顔検出器は2次元画像を用いて学習されている。このため、顔検出では3次元情報を取り扱うことができず、荒い位置推定しか行えないため(姿勢までは推定できない)、顔画像の切り出しが十分な精度で行えないという欠点がある。

この問題に対する2つ目の方法としては、顔追跡を用いた方法がある。この方法では、あらかじめ学習しておいたテンプレートと前フレームの情報を利用して顔の位置を推定する(初期フレームは検出器(Viola-Jones [12]など)を使う必要がある)。このアプローチは、特に、3次元姿勢追跡系を用いた場合に、より詳細な姿勢推定および画像切り出しを行うことができるという利点がある。以下では、追跡をベースとした関連研究についてまとめる。

Cootes et al.[15] は、画像と特徴点の 3 次元座標を組み合わせた固有空間により、形状の変形を含む顔追跡 (顔の特徴点追跡) を実現している (Active appearance models) . Active appearance models と認識系を組み合わせた融合系も報告されているが、この方法は特徴点の安定的な追跡を前提としているため、照明変動や姿勢変動が激しい場合に動作が安定しないという問題がある .

Xu - Roy-Chowdhury[16] は、様々な要因 (照明条件、姿勢変動、表情変化) による見え方の変化を局所多重線形モデルにより取り扱う方法を提案している . しかし、局所多重線形モデルにより顔の表情変化、姿勢変動や照明変動などへの対応が可能となる一方で、最適化の段階で全ての変動要因を同時に推定する必要があり、実時間処理には向いていないと考えられる .

ここまで述べてきた関連手法の中には、追跡、認識、照明推定や形状推定などの処理を統一的に行える応用例も報告されているが、高速処理を同時に実現した系は少ない . さらに、認識系において登録者数が大規模になった場合についての議論 (拡張性や実時間性) も十分に行われているとは言えない . この問題に対して、Oka-Shakunaga [17, 18] は、固有顔をベースとした 2 つの手法 (拡張固有顔、3 次元疎固有テンプレート追跡) を組み合わせることにより見え方や形状の変化に対応できるロバストな実時間顔追跡・認識融合系を実現している .

拡張固有顔は、L1 正規化を施した顔画像集合から構成される正規化固有顔 (坂上-尺長 [7, 19]) に形状推定機能を持たせたものである . 形式的には正規化固有顔に画像から形状への連想写像を追加することにより定式化されるが、連想写像は加重方程式と呼ぶ線形連立方程式に帰着でき、この方程式を解くことにより人物識別や形状推定などの処理を行う . さらに、登録者ごとに照明変動に対応する部分空間を正規化固有顔内に作成しておき、照明適応と呼ぶ操作を施すことで照明変動を考慮した系へ拡張することもできる .

3 次元疎固有テンプレート追跡は、パーティクルフィルタを用いて姿勢空間内で追跡対象の 3 次元姿勢を推定する 3 次元疎テンプレート追跡 (Oka et al.[20]) で用いるテンプレートを正規化固有顔に拡張した手法であり、固定形状を対象として照明変動にロバストな実時間顔追跡系を実現できる . さらに、この正規化固有顔を拡張固有顔に置き換えることで、未知形状を対象とした顔追跡・認識融合系へ拡張することができる .

この融合系において、加重方程式は高速な人物識別と形状推定を実現するために重要な役割を果たす . 加重方程式は、画像から形状への連想写像から導出される線形連立方程式であり、最適解 (加重) により未知画像は登録画像集合の最適な加重平均で表現される . ここで、加重方程式では加重和制約 (加重合計が 1 になるという制約) を加えて加重を求めている . 加重和制約は正規化画像空間・正規化固有顔の枠組みから自然に導入されたものであり、平均操作に閉じた議論を行うことができる .

Oka-Shakunaga[17, 18] では、過剰決定系の加重方程式を用いて、登録者 25 人の顔追跡・認識融合系を実時間動作させていた . 過剰決定系で加重方程式を構成

するためには，登録者数よりも正規化固有顔の次元数を多く設定する必要がある．しかし，従来の枠組みをそのまま用いて加重方程式を構成していたのでは，登録者数が増大した場合に加重方程式の規模が大きくなり，実時間性を維持できなくなる．すなわち，融合系の高性能化のためには，加重方程式の解法に改良を施す必要がある．

1.3 本論文の構成

本論文では，加重方程式をベースとした新たな加重の算出方法について検討し，実時間顔追跡・認識融合系（もしくは，単に顔認識系のみ）の高性能化を検討した．以下では，本論文の構成について述べる．まず，2章で拡張固有顔と3次元疎固有テンプレート追跡の基本的な枠組みと，これらを組み合わせた顔追跡・認識融合系の構成について述べる．

3章では，階層化加重方程式により，過剰決定系の枠組みの中で顔追跡・認識融合系を改良する方法について述べ，289人規模の静止画像データベースと100種類（10人×10表情）のパターンが登録された実動画データベースを用いて実験を行った結果について報告する．4章では，並列不足決定系を用いて，不足決定系の枠組みの中で顔追跡・認識融合系を改良する方法について述べ，3章と同様のデータベースを用いて実験を行った結果について報告する．

5章では，階層化加重方程式と並列不足決定系を用いた大規模識別系の構成について述べ，最大2197人登録のシミュレーションデータベースを用いた実験の結果を報告する．6章では，加重方程式をベースとした顔モデル生成法について述べ，生成した顔モデルの評価と顔追跡・認識融合系に適用した結果を報告する．7章では，これまで述べてきた手法や結果についてまとめ，加重方程式を用いた顔追跡・認識融合系および顔認識系の課題や将来的な展望について議論する．

第2章 拡張固有顔を用いた3次元顔追跡・認識融合系

固有顔 [1] は，顔の見え方の変化を効率よく扱うことができるため顔追跡や顔認識に広く用いられてきた．一方，表情変化や姿勢変動などの顔の見え方の変化は，顔の3次元形状に起因する．このため，様々な顔の変化に対応するためには，画像と3次元形状を組み合わせた顔画像処理系が必要となる．この中で，Oka-Shakunaga[17, 18] は，固有顔をベースとする2つの手法を組み合わせることにより実時間で動作可能な3次元顔追跡・認識融合系を提案している．

本章では，この系を構成する2つの手法である拡張固有顔および3次元疎固有テンプレート追跡の枠組みと，両者を組み合わせた実時間顔追跡・認識融合系についてまとめる．まず，拡張固有顔の構成とこれを用いた人物識別や形状推定について述べ，次に，3次元疎固有テンプレートを用いた3次元顔追跡法について述べる．最後に，両者を組み合わせた顔追跡・認識融合系とその動作例を示す．

2.1 拡張固有顔による3次元顔情報処理

2.1.1 正規化顔モデル

(1) 入力データ

本論文では，人物顔の3次元物体モデルをサーフェスモデルで取り扱い，以降では立体顔モデルと呼ぶ．サーフェスモデルでは，物体の表面は小さなパッチの集合で表現され，各パッチは物体中心座標系の点によって定義される．

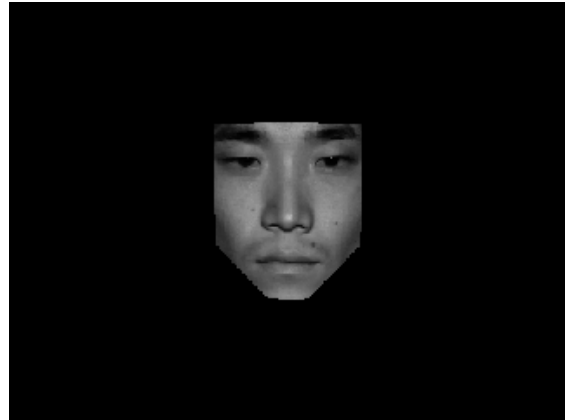
それぞれのパッチにはテクスチャパターンと面法線が定義されており，これらはレンダリングを行う際に物体中心座標からカメラ中心座標へと変換される．

(2) 幾何学的正規化

顔の3次元形状は人物ごとに異なる．そのため，顔画像認識に有効な固有顔を作成するためには，各人物の立体顔モデルに対して幾何学的な形状正規化が必要である．本論文では，顔の表面における右眼および左眼の中心と，唇の中心の3点を用いて形状正規化を行う．



幾何学的正規化前 (正面向き)



幾何学的正規化後 (正面向き)



幾何学的正規化前 (横向き)



幾何学的正規化後 (横向き)

図 2.1: 立体顔モデルの例 [17]

ここで、3点の物体中心座標をそれぞれ $M_{\text{reye}} = [X_{\text{reye}} \ Y_{\text{reye}} \ Z_{\text{reye}}]^T$ $M_{\text{leye}} = [X_{\text{leye}} \ Y_{\text{leye}} \ Z_{\text{leye}}]^T$ $M_{\text{lip}} = [X_{\text{lip}} \ Y_{\text{lip}} \ Z_{\text{lip}}]^T$ とする。

基準座標における3点をそれぞれ $[X_{\text{EYE}} \ 0 \ 0]^T$, $[-X_{\text{EYE}} \ 0 \ 0]^T$, $[0 \ Y_{\text{LIP}} \ 0]^T$ とする。このとき、与えられたモデルの基準座標への幾何学的正規化は次式で表される。

$$\mathbf{H} = \begin{bmatrix} X_{\text{EYE}} & -X_{\text{EYE}} & 0 \\ 0 & 0 & Y_{\text{LIP}} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_{\text{reye}} & X_{\text{leye}} & X_{\text{lip}} \\ Y_{\text{reye}} & Y_{\text{leye}} & Y_{\text{lip}} \\ Z_{\text{reye}} & Z_{\text{leye}} & Z_{\text{lip}} \end{bmatrix}^{-1} \quad (2.1)$$

幾何学的正規化を行った後に、基準座標から顔領域を切り出したものを正規化顔モデルとする。本論文では、眉から口にかけての領域を切り出し領域としている。ただし、顔の頬に相当する部分は、Z値の変化が最も激しくなる領域であり、各人物の顔の大きさや骨格によりZ値に大きな差が出るため、切り出し領域には含まない。図 2.1 に幾何学的正規化前後の立体顔モデルを示す。

(3) 正規化ベクトル表現

基準座標において作成された正規化顔モデルは，画像と奥行きの2つのベクトルで表すことができる．正規化顔モデルが N 組の画素値と Z 値から構成されている場合，画像ベクトル \mathbf{V} は N 次元ベクトルであり， $n(= 1, \dots, N)$ 番目の要素は n 番目の画素の輝度値を表す．また，奥行きベクトル \mathbf{Z} は N 次元ベクトルであり， n 番目の要素は n 番目の画素の Z 値を表す．

幾何学的正規化を行った後，輝度の変化に対応するため，画像ベクトルの総和が1となるように正規化する．画像ベクトル \mathbf{V} に正規化を施した正規化画像ベクトル \mathbf{v} は次式で定義される．

$$\mathbf{v} = \mathbf{V} / (\mathbf{1}^\top \mathbf{V}) \quad (2.2)$$

ここで， $\mathbf{1}$ は全ての要素が1のベクトルである．

2.1.2 正規化固有顔

登録者数を K とする．このとき， k 番目の人物の正規化画像ベクトルと奥行きベクトルはそれぞれ $\mathbf{v}_k, \mathbf{z}_k$ で表される．正規化画像ベクトルと形状ベクトルの平均ベクトルをそれぞれ $\bar{\mathbf{v}}, \bar{\mathbf{z}}$ とする．これらは次式で求められる．

$$\bar{\mathbf{v}} = \frac{1}{K} \sum_{k=1}^K \mathbf{v}_k \quad (2.3)$$

$$\bar{\mathbf{z}} = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \quad (2.4)$$

正規化画像ベクトル集合の分散共分散行列 \mathbf{C}_v は次式により求められる．

$$\mathbf{C}_v = \frac{1}{K} \sum_{k=1}^K (\mathbf{v}_k - \bar{\mathbf{v}})(\mathbf{v}_k - \bar{\mathbf{v}})^\top \quad (2.5)$$

ここで，対角行列 Λ は \mathbf{C}_v の固有値を降順に並べたものを示し， Φ は第 i 列が \mathbf{C}_v の固有ベクトルに対応する行列とする．このとき，主成分分析は $\Lambda = \Phi^\top \mathbf{C}_v \Phi$ で表される．

Φ の左 m 列からなる部分行列を Φ_m と表す．また， $\bar{\mathbf{v}}$ を中心とし，固有ベクトル Φ_m を固有軸として構成される m 次元固有顔を正規化固有顔と呼び $\langle \bar{\mathbf{v}}, \Phi_m \rangle$ で表す．

正規化画像 \mathbf{v} の正規化固有顔 $\langle \bar{\mathbf{v}}, \Phi_m \rangle$ への射影 (画像の復元) は次式で計算される．

$$\mathbf{v}^{\$} = \Phi_m \mathbf{s} + \bar{\mathbf{v}} \quad (2.6)$$

$$\text{where } \mathbf{s} = \tilde{\Phi}_m^\top (\mathbf{v} - \bar{\mathbf{v}}) \quad (2.7)$$

ここで, \mathbf{s} は一般的に射影係数と呼ばれ, 入力画像 \mathbf{v} の正規化固有顔内の座標を m 次元で表したものである.

数学の概念では式 (2.6) の方がよく知られているが, 本論文においては, \mathbf{s} が重要な役割を果たす. このため, 以降では射影係数 \mathbf{s} を求める操作 (式 (2.7)) を \mathbf{s} -射影と呼ぶ.

2.1.3 正規化固有顔への部分 \mathbf{s} -射影

部分画像の正規化固有顔への部分 \mathbf{s} -射影を考える. 画像 \mathbf{V} の部分画像を \mathbf{PV} で表す. ここで, \mathbf{P} は $N \times N$ 対角行列である. \mathbf{P} の対角要素は 0 または 1 であり, 正規化固有顔への部分 \mathbf{s} -射影が有効である画素に対応する対角要素は 1 となる. 部分画像 \mathbf{PV} の正規化固有顔への部分 \mathbf{s} -射影は $\bar{\mathbf{v}}$ と $\tilde{\Phi}_m$ の同次表現を用いて次式で表される [21].

$$\tilde{\mathbf{s}} = (\mathbf{P}\tilde{\Phi}_m)^\dagger (\mathbf{PV}) \quad (2.8)$$

$$\text{where } \tilde{\Phi}_m = [\tilde{\Phi}_m \quad \bar{\mathbf{v}}] \quad (2.9)$$

$$\tilde{\mathbf{s}} = [\alpha \mathbf{s}^\top \quad \alpha]^\top \quad (2.10)$$

ここで, α は入力画像の明るさに対応する係数であり, α で除すことにより \mathbf{s} -射影の正規化表現が得られる.

$$\hat{\mathbf{s}} = [\mathbf{s}^\top \quad 1]^\top \quad (2.11)$$

$\hat{\mathbf{s}}$ を用いた場合, 正規化画像の復元は次式で行われる.

$$\mathbf{v}^\S = \tilde{\Phi}_m \hat{\mathbf{s}} = [\tilde{\Phi}_m \quad \bar{\mathbf{v}}] \hat{\mathbf{s}} \quad (2.12)$$

また, 正規化された入力画像 \mathbf{v} と復元された画像 \mathbf{v}^\S の残差画像は次式で計算される.

$$\mathbf{v}^\# = \mathbf{v} - \mathbf{v}^\S = \mathbf{v} - \tilde{\Phi}_m \hat{\mathbf{s}} \quad (2.13)$$

2.1.4 画像から形状への連想写像

正規化固有顔 $\langle \bar{\mathbf{v}}, \tilde{\Phi}_m \rangle$ に K 人の顔画像を登録する場合を考える. ここで, \mathbf{v}_k の \mathbf{s} -射影の正規化表現を $\hat{\mathbf{s}}_k$ で表す. このとき, K 人全てに対して $\hat{\mathbf{s}}_k$ から形状 \mathbf{Z}_k が得られるような写像を定義できれば, この写像を形状推定に用いることができる. この問題には様々な解が考えられるが, Oka-Shakunaga [17, 18] では次式で定義される最適線形連想写像を取り上げている.

$$\Psi_m = \zeta_K \hat{\mathbf{S}}_K^+ \quad (2.14)$$

ここで， ζ_K と \hat{S}_K は，それぞれ K 人の形状ベクトルと s -射影を並べたものであり，次式で定義される．

$$\zeta_K = [\mathbf{Z}_1 \cdots \mathbf{Z}_K] \quad (2.15)$$

$$\hat{S}_K = [\hat{s}_1 \cdots \hat{s}_K] = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_K \\ 1 & \cdots & 1 \end{bmatrix} \quad (2.16)$$

なお，式 (2.14) 中の \hat{S}_K^+ は \hat{S}_K の疑似逆行列 (Moore-Penrose 逆行列) であり， $m+1 > K$ の場合は

$$\hat{S}_K^+ = (\hat{S}_K^T \hat{S}_K)^{-1} \hat{S}_K^T \quad (2.17)$$

となり， $m+1 < K$ の場合には

$$\hat{S}_K^+ = \hat{S}_K^T (\hat{S}_K \hat{S}_K^T)^{-1} \quad (2.18)$$

となる．

写像 Ψ_m を求めておけば，任意の s -射影 に対する形状ベクトルを次式により推定できる．

$$\mathbf{Z} = \Psi_m \hat{\mathbf{s}} \quad (2.19)$$

また，平均画像 $\bar{\mathbf{v}}$ が入力として与えられた場合，平均形状ベクトル $\bar{\mathbf{Z}}$ は次式によって求められる．

$$\Psi_m \tilde{\Phi}_m^+ \bar{\mathbf{v}} = \Psi_m \hat{\mathbf{0}} = \bar{\mathbf{Z}} \quad (2.20)$$

2.1.5 拡張固有顔

画像から形状の推定には，正規化固有顔 $\langle \bar{\mathbf{v}}, \Phi_m \rangle$ の同次表現 $\tilde{\Phi}_m$ と連想写像 Ψ_m を利用できる．ここで， $\tilde{\Phi}_m$ は固有顔であるが， Ψ_m は連想写像である．一方，両者を組み合わせることによって，画像情報と形状情報を一体として扱えるため，本論文では，両者をまとめて拡張固有顔 ($\tilde{\Phi}_m \Psi_m$) と呼ぶ．

2.1.6 加重方程式

ここで，

$$\mathbf{w} = \hat{S}_K^+ \hat{\mathbf{s}} \quad (2.21)$$

とおくと，式 (2.19) を次のように書き直せる．

$$\mathbf{Z} = \zeta_K \mathbf{w} \quad (2.22)$$

一方，式 (2.21) は次の線形連立方程式を \mathbf{w} について解くことに帰着できる．

$$\hat{S}_K \mathbf{w} = \hat{\mathbf{s}} \quad (2.23)$$

この線形連立方程式を加重方程式と呼ぶ。なお，加重方程式が過剰決定系 ($m+1 > K$) の場合，最適解は式 (2.23) において最小二乗誤差を与える \mathbf{w} となる。一方，不足決定系 ($m+1 < K$) の場合は，加重方程式を満たす \mathbf{w} の中から $\mathbf{w}^\top \mathbf{w}$ を最小とするものが最適解として得られる。

拡張固有顔においては，正規化固有顔によって画像情報を，連想写像によって形状情報を処理しているが，加重方程式を用いることにより拡張固有顔内で両者を一体として取り扱うことが可能となる。

2.1.7 加重方程式による人物識別

加重方程式の最適解 $\mathbf{w} = [w_1 \cdots w_K]^\top$ は制約条件 $\mathbf{1}^\top \mathbf{w} = 1$ のもとで最適化されることから，最適解は \mathbf{s} を $\{s_k\}$ の加重平均で表すことになる。従って， \mathbf{s} に最も類似した人物は加重が最も大きい人物と考えることができる。

$$k_{max} = \underset{k}{\operatorname{argmax}} w_k \quad (2.24)$$

このように， \mathbf{w} を用いることにより形状推定だけでなく人物識別も統一的に行うことができる。

2.1.8 照明適応を伴う加重方程式の実現

(1) 多重登録への拡張

2.1.6 節の議論では，各登録人物について 1 枚の画像を登録する場合についての加重方程式の構成法を述べた。この方法を，各人物の同じ形状に対して複数の画像を登録するように拡張することができる。これを多重登録と呼ぶ。多重登録は，照明変動の影響で登録人物の見え方が大きく変化する場合に有効である。

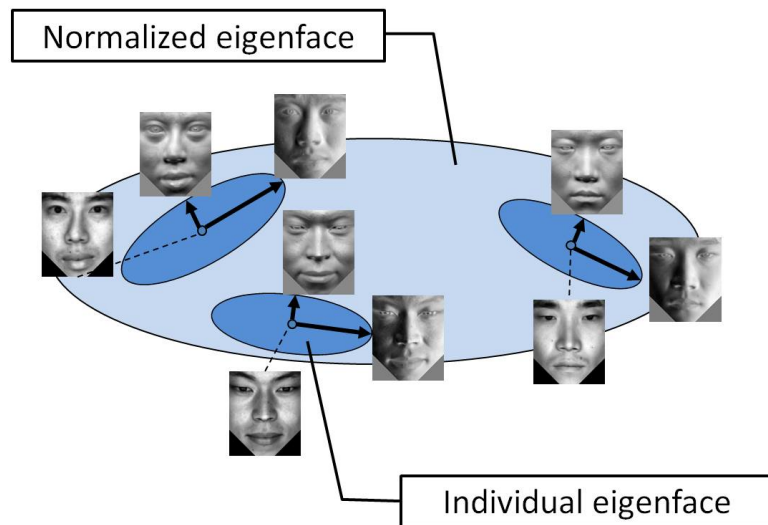
例えば，人物 $k (= 1, \dots, K)$ を照明条件 $l (= 1, \dots, L)$ で撮影した画像を \mathbf{V}_{kl} とし， \mathbf{V}_{kl} の s -射影の正規化表現を \hat{s}_{kl} とした場合，各 \hat{s}_{kl} を人物 k の形状 \mathbf{Z}_k に多重登録すればよい。 $K \times L$ 枚の画像が登録された場合，加重方程式は $K \times L$ 変数の連立方程式

$$\hat{\mathbf{S}}_{KL} \mathbf{w} = \hat{\mathbf{s}} \quad (2.25)$$

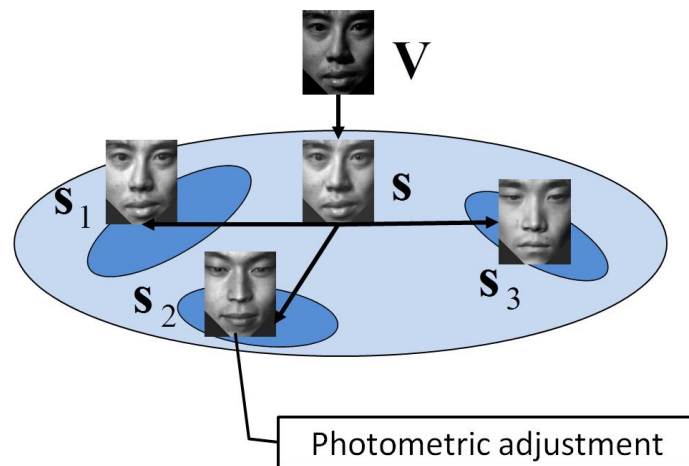
となり，登録 s_{kl} ごとに加重 w_{kl} が求められる。多重登録の場合，形状推定や人物識別に用いられる加重は同一人物に関する登録画像の加重の合計となる。

$$k_{max} = \underset{k}{\operatorname{argmax}} \sum_l w_{kl} \quad (2.26)$$

(2) 個人固有空間



(a) 個人固有空間の例



(b) 照明適応の例

図 2.2: 個人固有空間と照明適応の例

画像生成過程を考えると，画像は物体形状・表面反射特性・照明条件から生成される．特に，完全拡散反射面については無限遠点光源でカメラ位置固定と仮定した場合，対象物体の見え方は3自由度しか持たない．この仮定が厳密には成り立たない場合についても，低次元の部分空間を用いることで照明変動に伴う画像の見え方の変化を精度よく近似できる．そこで，事前に多様な照明条件で撮影した顔画像から各人物の個人固有空間を正規化固有顔内に作成し，加重方程式に適用することを考える．

以下の議論では，正規化固有顔 $\langle \bar{\mathbf{v}}, \Phi_m \rangle$ が複数の照明条件からなる画像集合 $\{\mathbf{V}_{kl} | k = 1, \dots, K, l = 1, \dots, L\}$ を主成分分析することにより作成された場合を考える．このとき，画像集合内の各画像の正規化固有顔への (部分)s-射影は次式で与えられる．

$$\tilde{\mathbf{s}}_{kl} = (\mathbf{P}\tilde{\Phi}_m)^+(\mathbf{P}\mathbf{V}_{kl}) \quad (2.27)$$

$$\text{and } \tilde{\mathbf{s}}_{kl} = [\alpha \mathbf{s}_{kl}^\top \alpha]^\top \quad (2.28)$$

これから，s-射影の正規化表現 $\hat{\mathbf{s}}_{kl} = [\mathbf{s}_{kl}^\top \ 1]^\top$ が得られる．

ここで，各人物 k について， $\mathcal{S}_k = \{\mathbf{s}_{kl} | l = 1, \dots, L\}$ を主成分分析することにより固有空間 $\langle \bar{\mathbf{s}}_k, \eta_k \rangle$ を作成できる．この空間を (人物 k の) 個人固有空間と呼ぶ．図 2.2(a) に個人固有空間の例を示す．この例では，3 人の登録画像集合から構成される正規化固有顔内に 3 つの 2 次元個人固有空間が作成されている．

任意の入力画像 \mathbf{V} の正規化固有顔への s-射影 \mathbf{s} が与えられたとき， \mathbf{s} の人物 k の個人固有空間 $\langle \bar{\mathbf{s}}_k, \eta_k \rangle$ への射影は次式で与えられる．

$$\mathbf{s}_k = \eta_k \eta_k^\top (\mathbf{s} - \bar{\mathbf{s}}_k) + \bar{\mathbf{s}}_k \quad (2.29)$$

(3) 照明適応を伴う加重方程式

入力画像 \mathbf{V} の正規化固有顔への s-射影 \mathbf{s} が与えられたとき， K 人の個人固有空間に対する射影 \mathbf{s}_k を式 (2.29) によってすべて求めることにより，各個人固有空間を代表する 1 点を選択できる．ここで，個人固有空間が多様な照明条件で撮影された画像集合から作成されていることから，個人固有空間への射影は，入力画像に最も類似した画像を個人固有空間内で求めることで照明条件を特定していることに相当する．本論文では，この操作を照明適応 (photometric adjustment) と呼ぶ．図 2.2(b) に照明適応の例を示す．この例では，3 人の登録人物について照明適応が行われており，入力画像 \mathbf{V} の s-射影 \mathbf{s} から，3 人の代表点 $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$ がそれぞれ計算されている．

照明適応により，各個人固有空間を代表する 1 点を自動的に生成できるため，各登録人物について 1 つの射影係数 s_k だけを考慮して，加重方程式を解けばよいことになる．即ち，照明適応の結果得られる \mathbf{s}_k を用いて，式 (2.23) を解くことにより，多重登録を行う必要なく，入力画像の照明条件を考慮した人物識別・形状推定を実現できる．

(4) 並列部分射影との併用

2 次元の顔認識におけるロバストな人物認識方法に並列部分射影は有効であることが知られている [21]．この手法は拡張固有顔による顔認識や形状推定にも応用できる．ここで，第 $q = (1, \dots, Q)$ 部分画像の射影の領域を指定する行列を \mathbf{P}_q とする．この時， \mathbf{P}_q に対する射影 $\tilde{\mathbf{s}}_q$ は次式で求められる．

$$\tilde{\mathbf{s}}_q = (\mathbf{P}_q \tilde{\Phi}_m)^+(\mathbf{P}_q \mathbf{V}) \quad (2.30)$$

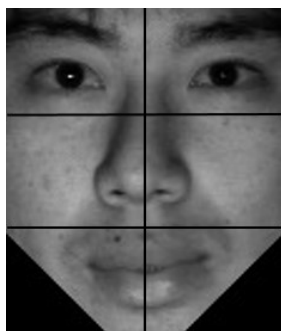


図 2.3: 並列部分射影に使用した部分領域

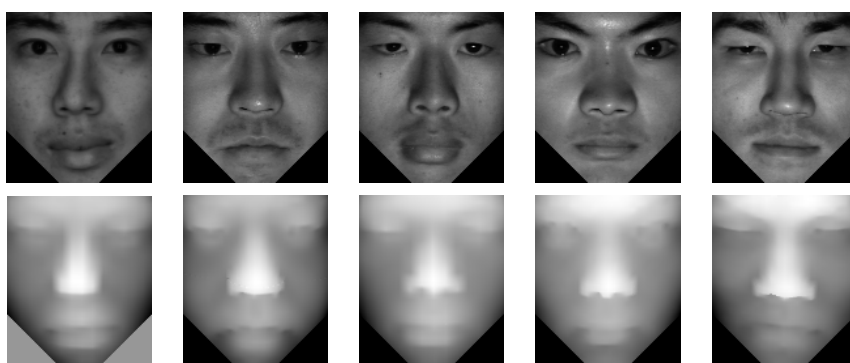


図 2.4: 立体顔モデルの登録例

次に, \hat{s}_q を式 (2.29) によって個人固有空間に射影した結果を s_{qk} とし, これらから構成される加重方程式 (2.23) を解くことで, 第 q 部分画像についての最適加重 w_q を求めることができる. このとき, Q 個の最適加重を平均することにより, 人物識別および形状推定のための加重を生成できる. この場合, 人物識別は次式で実現できる.

$$k_{max} = \operatorname{argmax} \frac{1}{Q} \sum_q w_{qk} \quad (2.31)$$

一方, w_q は, 部分毎の人物識別や形状推定に適用することも可能である.

図 2.3 に並列部分射影に用いた部分画像の例を示す. 本論文では, 図のように均等に分割された 6 個 (3 行 \times 2 列) の部分画像を使用した. なお, 最終的な加重の算出には, これらの部分画像と全体画像の計 7 領域 (すなわち, $Q = 7$) を使用した.

2.1.9 拡張固有顔と個人固有空間の例

ここでは, 25 人から作成した拡張固有顔の例を簡単に示す. 図 2.4 に 25 人の登録人物のうち 5 人の顔を示す. 各 2 段組中の上段と下段はそれぞれ, 顔の見えと形



図 2.5: AEF25x24 の作成に用いた登録顔画像と形状の例

状（奥行き画像）である．奥行き画像では手前にあるほど白くなり，奥に行くほど黒くなるように表示している．このため，鼻の付近は白く，両頬は黒くなっている（以降の奥行き画像についても同じ方法で表示する）．

各人物について照明条件の異なる 24 枚を登録画像とした．24 枚の登録画像の例を図 2.5 に示す．合計 600 枚の画像から 60 主成分からなる正規化固有顔を作成し，顔形状と組み合わせて拡張固有顔を作成した．この拡張固有顔を AEF25x24 と呼ぶ．

図 2.6 に AEF25x24 を示す．左端は平均画像 \bar{v} と平均形状ベクトル \bar{Z} のペアであり，2 列目以降に正規化固有顔の第 1～4 軸（上段），連想写像の第 2～5 軸（下段）を示す．また，図 2.7 に AEF25x24 中に作成された個人固有空間の例を示す．この

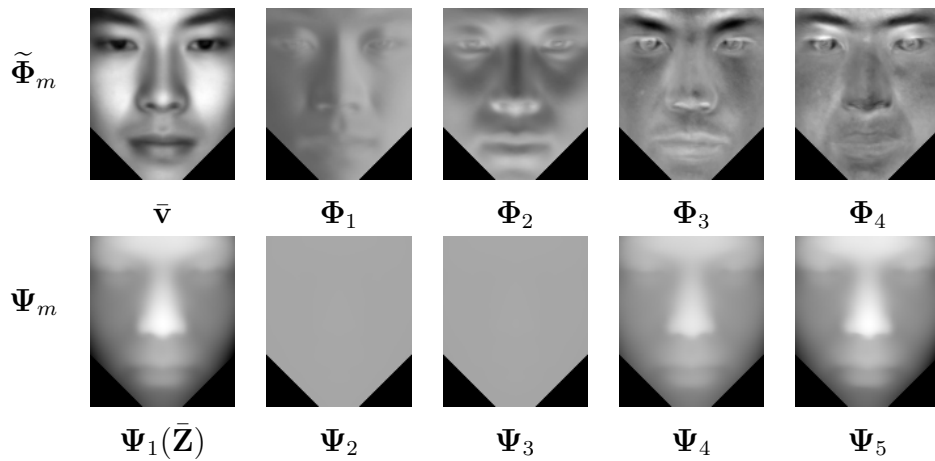


図 2.6: 拡張固有顔 AEF25x24

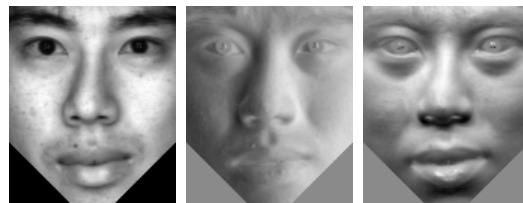


図 2.7: 個人固有空間の例

個人固有空間は，図 2.5 の人物の (2次元) 個人固有空間であり，図 2.5 に示した 24 枚から作成されたものである (なお，他の人物の個人固有空間も同様に作成されている)。

AEF25x24 中の正規化固有顔は，各人物について照明変動を施した場合に得られる 24 枚の画像を原画像として作成されているため，第 1 軸と第 2 軸には照明の変化を表す成分が現れている (正規化画像空間を用いているため，形状固定の完全拡散反射モデルでは 2 軸しか得られない。なお，この例では 25 人の顔画像を主成分分析しているため，厳密には 3 軸目以降にも照明の影響が残る)。この結果，第 1 軸と第 2 軸に対応する連想写像の軸 (形状ベクトル) は，形状の変化に寄与しないほぼ平らな面になっており，他の軸に対応する形状ベクトルと異なった特性を示していることが判る。

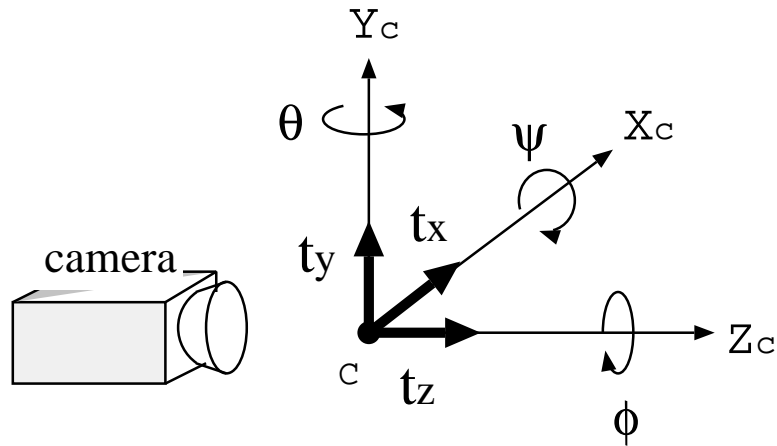


図 2.8: 姿勢パラメータ

2.2 3次元疎固有テンプレート追跡

ここでは，追跡物体形状が3次元サーフィスモデルとして与えられた場合における，単眼カメラによる剛体の3次元姿勢追跡について議論する．ここで，Oka et al.[20]では表面テクスチャを拡散反射（ランバート面モデル）で取り扱っているが，3次元疎固有テンプレート追跡では正規化固有顔で取り扱う．以下では，議論を簡単化するために物体の初期姿勢は既知であると仮定する．

2.2.1 3次元物体モデル

本論文では，追跡対象である3次元物体はサーフィスモデルで取り扱う．サーフィスモデルでは，物体の表面を小さなパッチの集合として扱うが，3次元疎固有テンプレート追跡では，このパッチの集合は正規化固有顔を用いて表現される．ここで，物体モデルの表面を構成する点を物体中心座標を用いて， $M_o = [X_o Y_o Z_o]^T$ で表す．

一方，画像生成過程はカメラ中心座標で記述する．ある姿勢の物体モデルから画像を生成するためには，物体中心座標からカメラ中心座標 $M_c = [X_c Y_c Z_c]^T$ に変換する必要がある．表記の簡単化のために， M_o と M_c の同次座標表現 $\tilde{M}_o = [M_o^T \ 1]^T$ ， $\tilde{M}_c = [M_c^T \ 1]^T$ を用いる．

2.2.2 姿勢空間と変換行列

対象物体を剛体とした場合，対象物体の3次元姿勢は並進3自由度，回転3自由度の計6自由度となる（図 2.8）．ここで 3×3 の回転行列 R ，3次元の並進ベクト

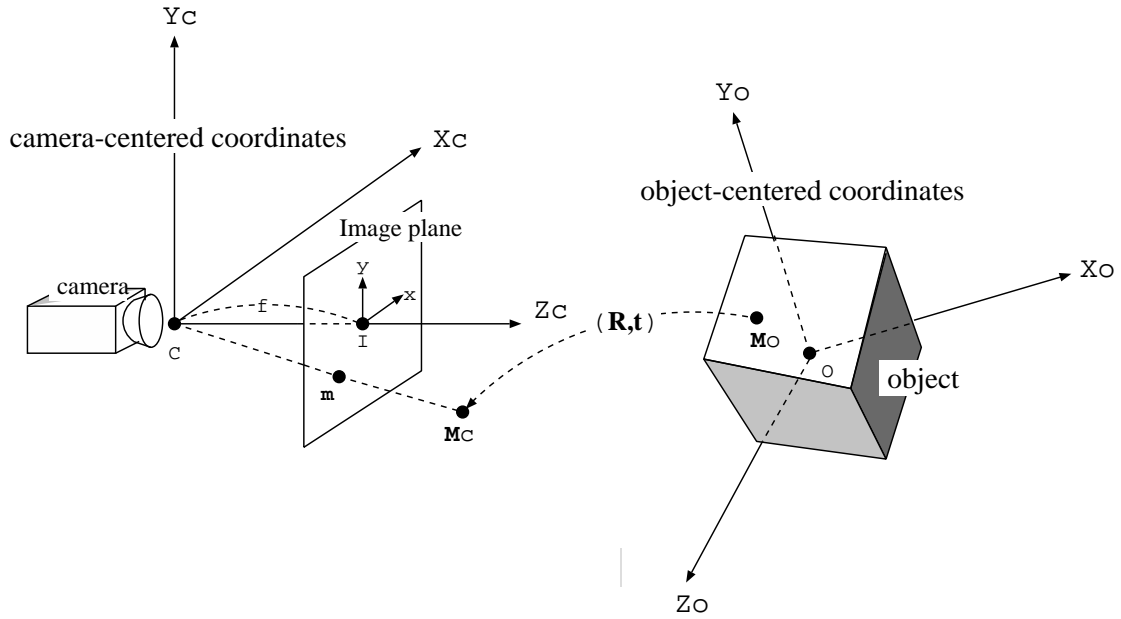


図 2.9: 物体中心座標系とカメラ中心座標系

ル t を以下のように定義する .

$$\mathbf{R} = \mathbf{R}_\psi^x \mathbf{R}_\theta^y \mathbf{R}_\phi^z, \quad \mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

ただし ,

$$\mathbf{R}_\psi^x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\psi & -s_\psi \\ 0 & s_\psi & c_\psi \end{bmatrix}, \quad \mathbf{R}_\theta^y = \begin{bmatrix} c_\theta & 0 & s_\theta \\ 0 & 1 & 0 \\ -s_\theta & 0 & c_\theta \end{bmatrix},$$

$$\mathbf{R}_\phi^z = \begin{bmatrix} c_\phi & -s_\phi & 0 \\ s_\phi & c_\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

ここで , $c_\psi = \cos \psi$, $s_\psi = \sin \psi$ であり , 他も同様である . このとき , 物体中心座標 $\tilde{\mathbf{M}}_o$ からカメラ中心座標 $\tilde{\mathbf{M}}_c$ への変換は次式で表される . また , この関係は図 2.9 のように示される .

$$\tilde{\mathbf{M}}_c = \mathbf{D} \tilde{\mathbf{M}}_o \quad (2.32)$$

$$\text{ただし } \mathbf{D} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

この定義より , 姿勢行列 \mathbf{D} は姿勢パラメータ $\mathbf{b} = [t_x \ t_y \ t_z \ \psi \ \theta \ \phi]^\top$ により表される .

2.2.3 画像生成過程とモデルベースレンダリング

物体の姿勢が D により与えられた時，物体上の点 M_o は画像平面上の点 $m_c = [x \ y]^T$ に射影される．これは次式で表される．

$$\tilde{m}_c = AD\tilde{M}_o \quad (2.33)$$

ここで， $\tilde{m}_c = [sx \ sy \ s]^T$ は， m_c の同次座標表現である．また， A は 3×4 の射影行列であり，実際のカメラでは内部パラメータを求めることにより得られる．

本論文では，既知物体モデルの見えマッチングによる姿勢追跡を議論することから，さまざまな姿勢における画像を生成するためのモデルベースレンダリング技術が必要となる．レンダリング法は基本的には，Z-buffer 法に従う．ここで，Z-buffer 法では式 (2.33) と同時に深さ計算を行っている．これを明示的に表現すると，Z-buffer 法は次式で表現される．

$$\begin{bmatrix} \tilde{m}_c \\ Z_c \end{bmatrix} = \begin{bmatrix} A \\ 0 \ 0 \ 1 \ 0 \end{bmatrix} D\tilde{M}_o \quad (2.34)$$

ここで，式 (2.34) の 4 行目は奥行きを表している．Z-buffer 法では，Z 値 $Z(x, y)$ が更新されると共に，画素値 $V(x, y)$ が更新される．

Oka et al.[20] は， $V(x, y)$ をランバート面モデルに従って求めているが，本論文では，固有テンプレートを用いるため，ランバート面モデルによる画素値計算は不要になる．その代わりに，固有テンプレートから各画素に対応する情報を取り出す操作が必要になる．ここで，固有テンプレートを $\tilde{\Phi}_m = [\Phi_m \ \bar{v}]$ で表現すると，画素 (x, y) に対応する $(m + 1)$ 次元ベクトル $\tilde{\Phi}_m(x, y)$ が $V(x, y)$ の代わりに用意されることになる．

2.2.4 姿勢追跡問題

姿勢追跡を，初期姿勢が与えられた場合に連続的に姿勢推定を行う問題と定義する． b_t は時刻 t における姿勢パラメータ b の値を示し， D_t は b_t から求められる姿勢行列を表す．このとき，姿勢追跡とは姿勢更新行列 $\delta D = D_t D_{t-1}^{-1}$ を逐次推定することに帰着できる．

ここで， $\delta b = [\delta x \ \delta y \ \delta z \ \delta \psi \ \delta \theta \ \delta \phi]^T$ は δD の 6 パラメータを表す．すなわち， δD は次のように表される．

$$\delta D = D_t D_{t-1}^{-1} = \begin{bmatrix} R_{\delta \psi}^x & R_{\delta \theta}^y & R_{\delta \phi}^z & [\delta x \ \delta y \ \delta z]^T \\ \mathbf{0}^T & & & 1 \end{bmatrix} \quad (2.35)$$

対象物体の姿勢を正確に追跡し続けるためには，姿勢更新行列を推定した後に， D_t を次式で正しく計算する必要がある．

$$D_t = \delta D D_{t-1} \quad (2.36)$$

2.2.5 基準姿勢における固有テンプレートの作成

前時刻 $t - 1$ の推定姿勢 \mathbf{b}_{t-1} が与えられた場合，この姿勢から固有テンプレートを作成できる．以下では，前時刻の推定姿勢から予想される姿勢を基準姿勢と呼び， $\hat{\mathbf{b}} = [\hat{t}_x \ \hat{t}_y \ \hat{t}_z \ \hat{\psi} \ \hat{\theta} \ \hat{\phi}]^\top$ と表す．

Z-buffer 法ではある画像上の画素に値を書き込む際に，画素値 $V(x, y)$ だけでなく，その画素に対応する3次元座標のZ値 $Z(x, y)$ をZ-bufferに保存しておく．Oka et al.[20]では，Z-buffer法の拡張として， $V(x, y)$ ， $Z(x, y)$ を更新する際に，同時に $X(x, y)$ ， $Y(x, y)$ の値も更新する拡張Z-buffer法を提案している．これは式(2.33)による画像生成時に，物体のカメラ中心座標を計算することに相当し，形式的には次の計算を行うことと等価である．

$$\begin{bmatrix} \tilde{\mathbf{m}}_c \\ \mathbf{M}_c \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{I} \ \mathbf{0}_3 \end{bmatrix} \mathbf{D}\tilde{\mathbf{M}}_o \quad (2.37)$$

拡張Z-buffer法により，各画素 (x, y) において画素値 $V(x, y)$ が与えられるとともに，3次元座標 $[X(x, y) \ Y(x, y) \ Z(x, y)]$ が計算される．本論文では，固有テンプレートを使用することから， $V(x, y)$ を計算する代わりに， $\tilde{\Phi}_m(x, y)$ が計算されることになる．

2.2.6 3次元疎固有テンプレートの作成

基準姿勢 $\hat{\mathbf{b}}$ において固有テンプレート $\tilde{\Phi}_m(x, y)$ が作成されると，このとき， $\tilde{\Phi}_m(x, y)$ から3次元疎固有テンプレートが作成できる．ここで，点選択方法としてMatsubara-Shakunaga[9]が提案している局所領域最大最小点選択法(テンプレートを $h \times h$ の局所領域に分け，各局所領域内で輝度値が最大・最小となる点を選択する)を用いる場合には，輝度値 $V(x, y)$ を与える必要がある．本論文では，初期値として正規化固有顔の平均ベクトル $\bar{\mathbf{v}}(x, y)$ を用いて3次元疎固有テンプレートを作成する．一方，追跡開始後は観測される画像 $V(x, y)$ が得られることから，これを用いて3次元疎固有テンプレートの更新を行う．選択された点集合を $\{(x_j, y_j)\}$ とすると，3次元疎固有テンプレートは次のように記述できる．

$$\{(\mathbf{M}_j; \tilde{\Phi}_{mj})\} = \{(X_j \ Y_j \ Z_j; \tilde{\Phi}_{mj})\} \quad (2.38)$$

ここで， $X_j = X(x_j, y_j)$ ， $Y_j = Y(x_j, y_j)$ ， $Z_j = Z(x_j, y_j)$ ， $\tilde{\Phi}_{mj} = \tilde{\Phi}_m(x_j, y_j)$ である．なお， m 次元の正規化固有顔から3次元疎固有テンプレートが構成される場合，固有顔の次元数 $(m + 1)$ 個のテンプレートが生成されることになる．

2.2.7 6次元姿勢空間における3次元疎固有テンプレート追跡

追跡対象の物体は6自由度を持つため，3次元物体の追跡は6次元姿勢空間において取り扱われる．しかし，各フレームにおいて6姿勢パラメータを直接推定す

るのは効率的ではない．ここで，3次元疎固有テンプレートが基準姿勢 $\hat{\mathbf{b}}$ から作成されることから，基準姿勢からの姿勢変動を推定するのが有効である．

すなわち，サンプルを基準姿勢 $\hat{\mathbf{b}}$ の近傍で定義される6次元姿勢空間の小領域内に限定して生成する．ここで，時刻 t における G 個のサンプル集合を $\{\delta\mathbf{b}_t^{(g)}\}$ ，サンプルを

$$\delta\mathbf{b}^{(g)} = [\delta x^{(g)} \ \delta y^{(g)} \ \delta z^{(g)} \ \delta\psi^{(g)} \ \delta\theta^{(g)} \ \delta\phi^{(g)}]^\top \quad (2.39)$$

で表す．

(1) 疎テンプレート追跡系

図 2.10 に疎テンプレート追跡系の構成を示す．追跡法は疎テンプレートマッチングとコンデンセーション法を組み合わせた方法である．追跡法では，サンプル集合の一部をフレームからフレームへと伝搬させる．

ある時刻のサンプル集合を $\{\delta\mathbf{b}_t^{(g)}, g = 1, \dots, G\}$ ，サンプル $\delta\mathbf{b}_t^{(g)}$ の重みを $\pi_t^{(g)}$ とする．ここで $\pi_t^{(g)}$ は時刻 t における確率密度を表し，drift，diffuse，measureの3段階によって推定される．このアルゴリズムの概要を以下で述べる．

drift ステップ:

時刻 $t-1$ におけるサンプルセット $\{\delta\mathbf{b}_{t-1}^{(g)}, \pi_{t-1}^{(g)}\}$ ， $g = \{1, \dots, 10\}$ を時刻 $t-1$ における G 個のサンプルから選択する．

diffuse ステップ:

選択された10個のサンプルに対して，正規分布に従って次時刻のサンプルを G 個生成する．標準偏差 $\sigma = [\sigma_x \ \sigma_y \ \sigma_z \ \sigma_\psi \ \sigma_\theta \ \sigma_\phi]^\top$ のガウスノイズが各サンプルのすべての姿勢パラメータに対して付加される．

measure ステップ:

サンプル $\delta\mathbf{b}_t^{(g)}$ の評価値 d_g を疎固有テンプレートマッチングにより算出する．その後， G 個のサンプルの中から評価値の上位10個を選択し，サンプル $\delta\mathbf{b}_t^{(g)}$ の重みを次式で計算する．

$$\pi_t^{(g)} = \frac{1/d_g}{\sum_{j=1}^{10} (1/d_j)} \quad (2.40)$$

10個のサンプルとその重みから時刻 $t-1$ からの姿勢変動を次のように推定する．

$$\delta\hat{\mathbf{b}}_t = \sum_{g=1}^{10} \pi_t^{(g)} \delta\mathbf{b}_t^{(g)} \quad (2.41)$$

これにより， $\delta\hat{\mathbf{D}}$ が計算される．最後に， $\delta\hat{\mathbf{D}}$ から時刻 t における推定姿勢が次式で求める．

$$\hat{\mathbf{D}}_t = \delta\hat{\mathbf{D}}_t \hat{\mathbf{D}}_{t-1} \quad (2.42)$$

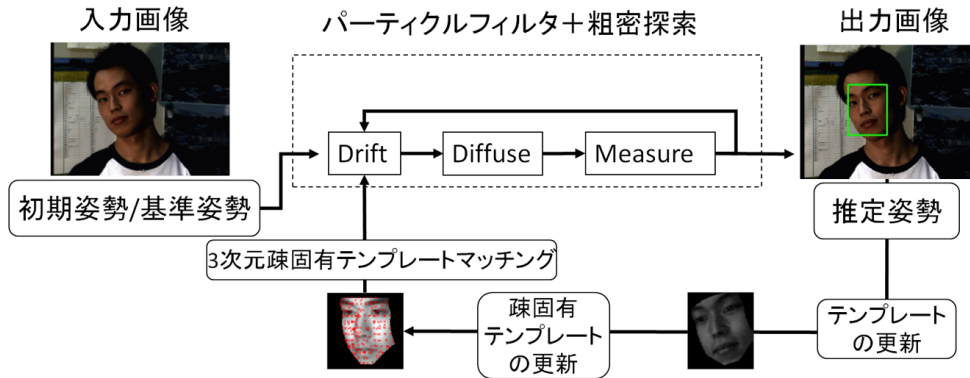


図 2.10: 疎テンプレート追跡系の構成 [17]

(2) 粗密追跡法

粗密追跡法は，drift，diffuse，measure から構成される追跡法を， σ を変化させながら繰り返し行うことにより追跡性能を改善できることが確認されており，本論文もこれに従う．なお，本論文の実装では i 段目の標準偏差 σ_i は $i-1$ 段目の標準偏差を半分にした $\sigma_{i-1}/2$ としている．これにより，diffuse では，drift で選択された各サンプルに対し標準偏差を σ_i とするランダムノイズを付加してサンプル集合を生成することとなる．

2.2.8 3次元疎固有テンプレートマッチング

2.2.6 節で述べたように，3次元疎固有テンプレートをいったん作成すると，姿勢毎に Z-buffer 法を適用し 3次元疎固有テンプレートマッチングに用いる必要はなくなり，3次元疎固有テンプレートを各サンプル姿勢に変換するだけでよい．

3次元疎固有テンプレートマッチングでは，3次元疎固有テンプレートを各サンプル姿勢に変換し，疎な点の画像平面座標を求めることにより入力画像中の疎な画素集合を得る．この画素集合を正規化固有顔 $(\tilde{\mathbf{v}}, \tilde{\Phi}_m)$ に部分射影し，得られる射影残差中の選択された画素に対応するものに，ロバスト関数 (Geman-McClure 関数) を適用して積算したものを評価値とする．以下では，評価値の計算方法を具体的に示す．

各候補姿勢 g から得られる画像を \mathbf{V}_g とした場合，射影残差は次式で計算される．

$$\mathbf{V}_g^\# = (\mathbf{P}_{temp} \tilde{\Phi}_m)(\mathbf{P}_{temp} \tilde{\Phi}_m)^+ (\mathbf{P}_{temp} \mathbf{V}_g) - (\mathbf{P}_{temp} \mathbf{V}_g) \quad (2.43)$$

ここで， \mathbf{P}_{temp} は，固有テンプレートを構成する画素のうち疎点に対応する対角要素が 1 である領域指定行列である．

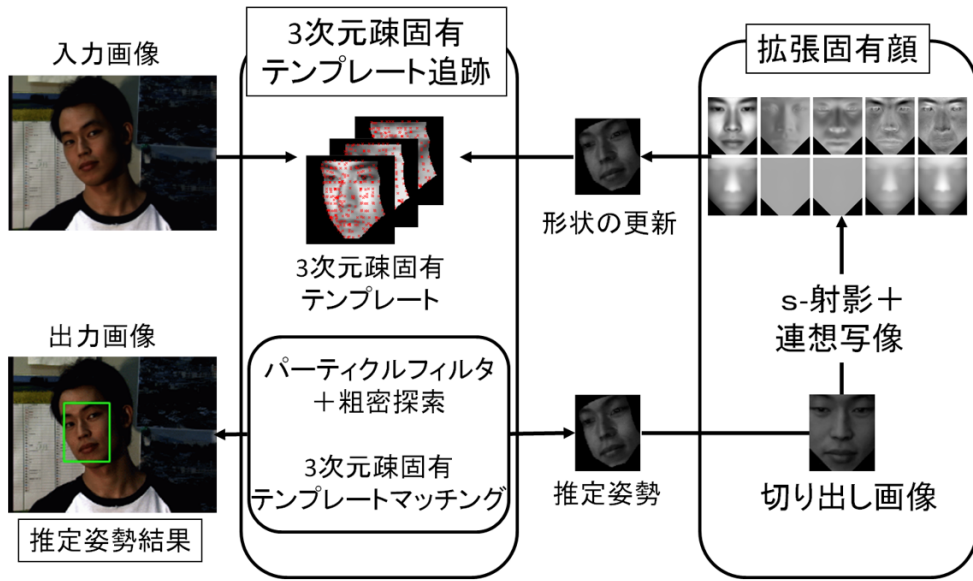


図 2.11: 拡張固有顔による 3 次元疎固有テンプレート追跡系の構成 [17]

$V_g^\#$ の第 j 画素に対応する残差を $v_{gj}^\#$ とすると, 候補姿勢 g の評価値 d_g はロバスト関数 (Geman-McClure 関数) を用いて次式で表される.

$$d_g = \sum_{j=1}^{j_{max}} \frac{v_{gj}^{\# 2}}{\alpha^2 c^2 + v_{gj}^{\# 2}} \quad (2.44)$$

式 (2.44) が式 (2.40) の重み計算に用いられる. ここで α は $(P\tilde{\Phi}_m)^+(PV_k)$ の最後の要素であり, d_{gj}/α は正規化後の残差を表す. また, c は Geman-McClure 関数のパラメータ (定数) であり, 適切な値を設定することでノイズの影響を低減することができる (本論文では $c = 0.4$ としている).

2.3 拡張固有顔による 3 次元疎固有テンプレート追跡

2.3.1 拡張固有顔による 3 次元疎固有テンプレート追跡の概要

顔追跡・認識融合系は, 図 2.11 に示すように, 3 次元疎固有テンプレート追跡系と拡張固有顔を組み合わせる構成される. 3 次元疎固有テンプレート追跡系は, 固定形状を対象とした 3 次元疎固有テンプレートマッチングを用いて, 6 次元姿勢空間において姿勢追跡を行う. これによって, 各フレームにおいて推定姿勢が求まることになる. このとき, 推定姿勢で切り出された画像を拡張固有顔に射影することによって画像情報からの形状復元が可能であるため, 3 次元形状モデルの更新が可能となる. また, 拡張固有顔を使用することで, 人物識別も同時に行えるた

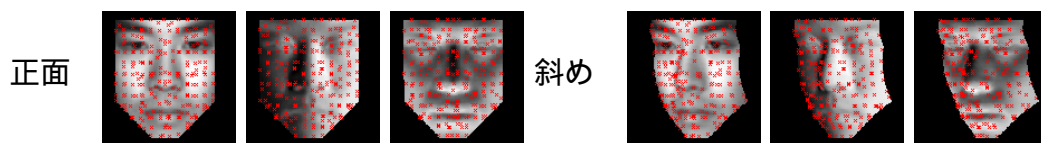


図 2.12: 3次元疎固有テンプレートの初期値 [17]

め、この系全体としては、3次元を対象とした顔追跡・認識系を実現できることになる。

(1) 追跡開始時における見えと形状

本論文の実装では、追跡開始時の立体顔モデル(3次元疎固有テンプレートの初期値)の見えは正規化固有顔の平均画像ベクトルとし、形状は正規化固有顔の平均画像ベクトルに対応付けられる平均奥行きベクトルとする。つまり、 \bar{v} 、 \bar{Z} を用いて初期の3次元疎固有テンプレートを作成する。図 2.12 に初期テンプレートから作成された3次元疎固有テンプレート(平均画像と正規化固有顔の第1軸、第2軸)を示す。上段は正面から見たものであり、下段は斜めから見たものである。なお、追跡の開始前に、適当な顔検出器を用いて顔の位置を自動的に推定しておく必要があるが、本論文では、Viola-Jones[12]を用いて学習された検出器の結果を初期値として利用した。

(2) 3次元形状の更新と人物識別

前節で述べたように、初期モデルは平均形状と平均画像から作成されている。すなわち、初期状態では被写体とは異なる3次元モデルを追跡していることになる。そこで、追跡性能を高めるために、姿勢推定によって得られる(密な)テンプレートを拡張固有顔に射影することによって3次元モデルを各フレームの各段で更新する。これにより、この系では初期モデルが被写体とは異なっても、拡張固有顔によって徐々に被写体のモデルへ更新していくことが可能となっている。なお、照明適応によって照明変動や、姿勢変動による照明の相対的な変動に対応することもできる。また、人物識別は拡張固有顔へのs-射影を用いて行うことができる。本論文では、2.1.8節で述べた、7部分領域(図 2.3)の並列部分射影による識別系を採用している。

2.3.2 3次元顔追跡・認識融合系の動作例

(1) 顔追跡・認識の動作例

ここでは、拡張固有顔による顔追跡・認識融合系の動作例について簡単に示す。図 2.13 に拡張固有顔 AEF25x24 による実画像の追跡・認識の例を示す。この例で

は、比較的穏やかな照明条件下で対象人物の顔がゆっくりと動いている 600 フレームの動画を使用した。

追跡開始時に用いる顔モデルの見えと形状は AEF-25x24 の平均とし、初期姿勢はカメラに対して正面を向いているものとした。また、拡張固有顔とパーティクルフィルタの姿勢推定結果を用いて、顔の見えと形状の更新を各フレームの格段で行った。図 2.13 は初期姿勢と実姿勢が異なる例の追跡開始から 2 フレームの結果を左から右へ順に並べたものである。フレーム 1 では姿勢推定誤差が大きいため、立体顔モデルに大きな変化がない。フレーム 2 では推定姿勢が実姿勢に近づくことで、立体顔モデルが対象人物へと変化しており、人物識別に用いる最大加重も正しく得られていた。

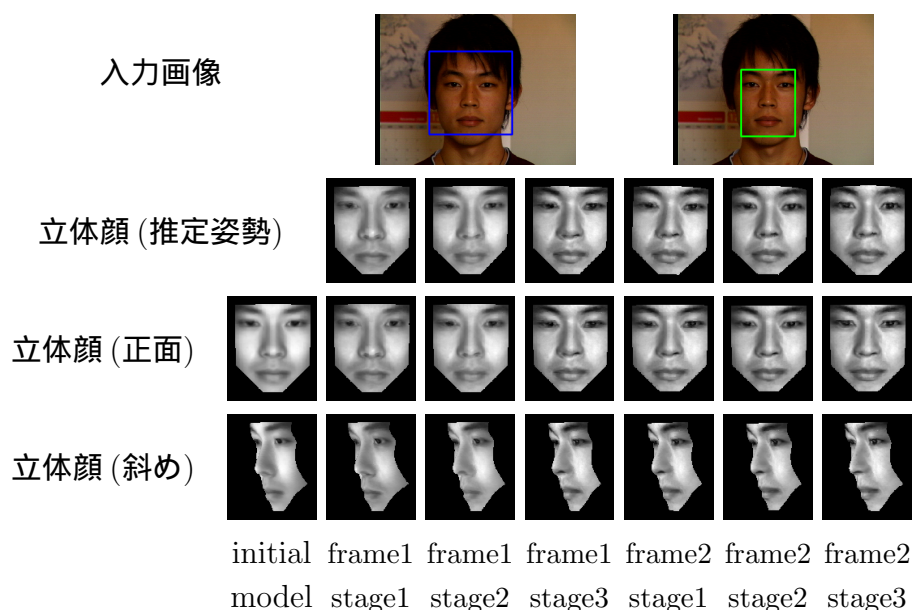


図 2.13: 初期姿勢検出からの追跡・人物認識の例 [17]

(2) 表情変化を含む場合の動作例

ここでは、拡張固有顔による表情追跡の実験を行った。1 人の人物の 9 つの表情から、人物識別の場合と同様の方法で拡張固有顔を作成し追跡に用いた。図 2.14 に表情変化への対応例を示す。図 2.14 中の各画像の上段は入力画像と出力画像であり、各画像の下段は追跡により得られた画像から推定されるモデルをそれぞれ正面と斜め方向から見た画像である。図 2.14 を見ると、人物顔の姿勢と照明条件が異なっている入力画像のそれぞれで表情変化に対応できていることがわかる。

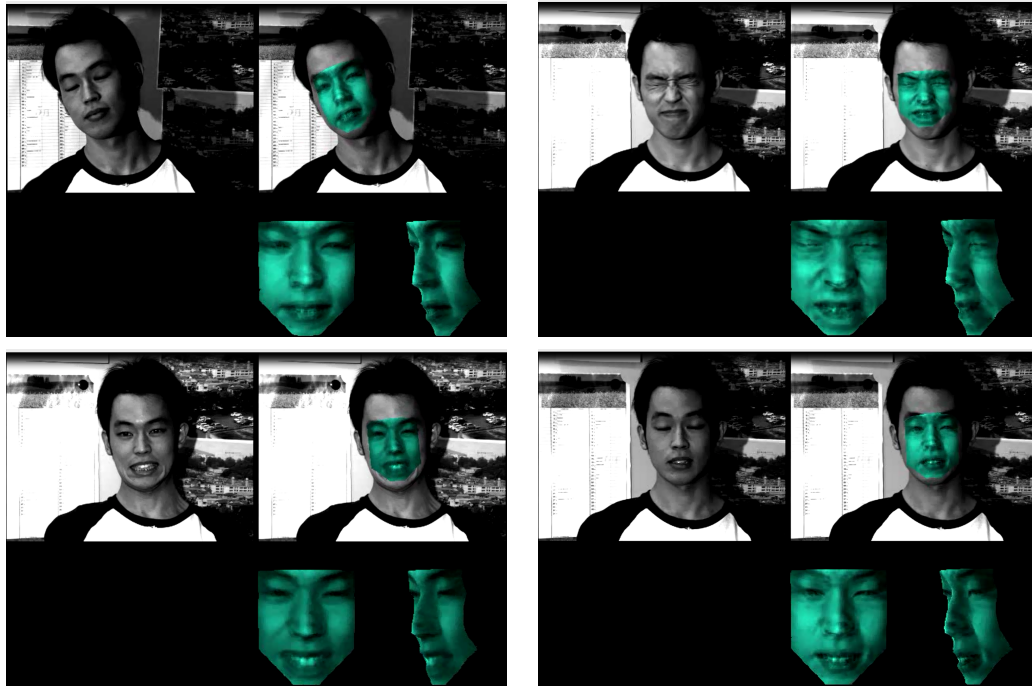


図 2.14: 表情認識への適用例 [17]

2.4 まとめと問題点

本章では，拡張固有顔と3次元疎固有テンプレート追跡を組み合わせた実時間顔追跡・認識融合系について述べ，その動作例を示した．

まず，拡張固有顔の構成方法について論じた．拡張固有顔は，正規化固有顔と画像情報への連想写像により定式化され，この連想写像によって画像情報（射影係数）から形状の推定が可能となる．ここで，連想写像は正規化固有顔上の登録射影係数の集合から各形状に対する写像で定義されるため，疑似逆行列と入力の射影係数ベクトルの内積を求めることで形状が決まる．一方，式変形を行うことにより，この操作は入力の射影係数を登録の射影係数の最適な加重平均で表現する線形連立方程式（加重方程式）を解く問題へと帰着できる．また，あらかじめ各登録人物の照明変動に対応する部分空間を正規化固有顔内に作成しておき照明適応と呼ぶ操作と組み合わせることで，照明変動に頑健な加重方程式を構成できることを示した．

次に，正規化固有顔をベースとした追跡法である3次元疎固有テンプレート追跡について述べた．また，正規化固有顔を拡張固有顔で置き換えることにより，見え方の変化だけでなく，形状の取り扱いも可能となった実時間顔追跡・認識融合系を実現した．

実時間顔追跡・認識融合系において，画像情報と形状情報の取り扱いは連想写像により定式化されており，連想写像から導出される加重方程式は，高速な顔認識や

形状推定の実現のために重要な役割を果たす．ここで，連想写像は擬似逆行列により定式化(式(2.14))されているが，これは従来のパターン認識においても最適線形連想写像として登場する．例えば，ニューラルネットワークの教科書[22]にも連想写像の導入として最適線形連想写像が記載されている．ここで，ニューラルネットワークでは線形性を排除して理論を発展させている．一方，Oka-Shakunagaは，和が1であるという制約を加え，加重平均を求める問題として連想写像を定式化している．ここで，線形連立方程式の枠組みで解くことが本質的に重要であり，実時間系を実現できるだけでなく，学習に必要な計算コストを削減できる．即ち，ニューラルネットワークにおいて重要な役割を果たす逆伝播(back-propagation)による学習を考えることなく，パターン空間に存在する点集合の加重配分を線形連立方程式の範囲内で解決することができる．

Oka-Shakunaga[17, 18]の系では，最大25人登録の実時間顔追跡・認識融合系を60次元の正規化固有顔を用いて実現している．このとき，加重方程式のサイズは 61×25 (過剰決定系)である．Oka-Shakunaga[17, 18]では，加重方程式を過剰決定系で解くように系を構成しているが，過剰決定系の加重方程式を構成するためには，登録者数よりも正規化固有顔の次元数を多く設定する必要がある．ここで，登録者数を大規模化し，顔追跡・認識融合系を高性能化する場合，元の加重方程式の枠組みをそのまま用いていたのでは，登録者数が増大した場合に加重方程式の規模が大きくなり，実時間性を維持できなくなる．

この問題に対し，3章，4章では，元の加重方程式をベースとした新たな加重の算出方法について検討し，顔追跡・認識融合系の高性能化(登録者数の大規模化，実時間性の維持，識別性能の改善)を実現する．

第3章 階層化加重方程式による顔追跡・認識融合系の高性能化

3.1 階層化加重方程式による解法

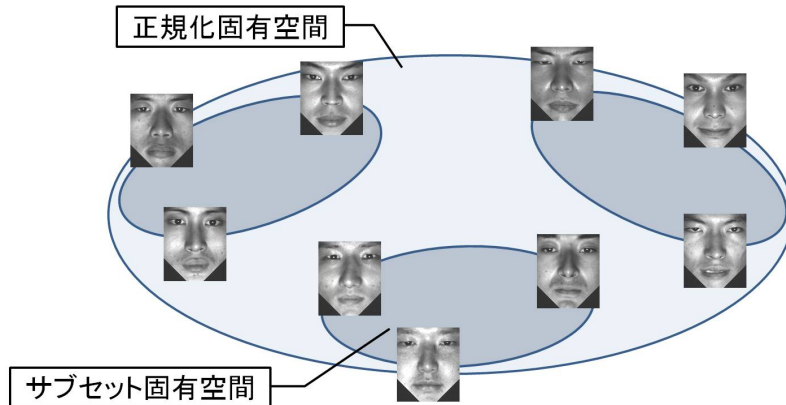
加重方程式では、正規化固有顔内に登録人物全員の個人固有空間を作成することにより、照明適応を実現し、照明条件を考慮した加重方程式を構成できる。一方、過剰決定系の加重方程式では、正規化固有顔の次元数を登録者数より大きくする必要があり、過剰決定系においては、登録者数が増加すると、それに伴い必要な次元数も増加するため、計算コストが大きくなり実時間性を保つことは容易でなくなる。

本章では、元の加重方程式(用語の混同を避けるため式(2.23)を基本加重方程式と呼ぶ)において個人単位で行っていた操作を登録者集合のサブセットを用いて階層的に行うことにより、この問題を回避する。具体的には、個人固有空間を正規化固有顔内に直接作るのではなく、登録人物集合のサブセットから作成される部分空間内に作成し、サブセット間とサブセット内の2階層で小規模な加重方程式を構成する[27]ことにより登録者の増大に対処する。

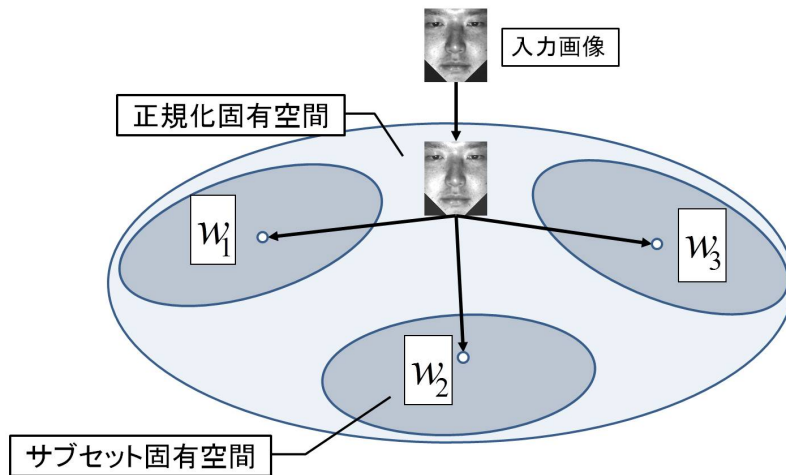
なお、この方法では、サブセットの構成人物を入れ替えることで、複数パターンの加重方程式が作成可能である。本章では、これらを用いて加重方程式の並列度を高め、大数の法則により最終的な識別性能を改善する方法も合わせて提案する。

3.1.1 サブセット固有空間

人物 k に対して照明条件 l で撮影された画像の正規化固有顔への s -射影の正規化表現 $\hat{s}_{kl} = [s_{kl}^T \ 1]^T$ が与えられているとする。 K_1 個のサブセットを K_2 人 ($K = K_1 K_2$ とする) で構成する場合、第 $k_1 (= 1, \dots, K_1)$ サブセットについて、 $S_{k_1} = \{s_{(k_1, k_2)l} | k_2 = 1, \dots, K_2, l = 1, \dots, L\}$ を主成分分析することにより m' 次元の固有空間 $\langle \bar{s}_{k_1}, \eta_{k_1} \rangle$ を作成できる。この空間をサブセット固有空間と呼ぶ。なお、 (k_1, k_2) には任意の登録人物 k が割り当てられるが、サブセット内外で重複がないように選択する。図3.1の例では、9人の登録者集合は3人の登録者からなる3つのサブセットに分割され、3つのサブセット固有空間が正規化固有顔内に作成されている。



(a) サブセット固有空間



(b) サブセット間加重方程式

図 3.1: サブセット固有空間とサブセット間加重方程式の例

任意の入力画像 V の正規化固有顔への s -射影 s が与えられたとき, s の k_1 番目のサブセット固有空間 $\langle \bar{s}_{k_1}, \eta_{k_1} \rangle$ への射影は次式で与えられる.

$$s_{k_1} = \eta_{k_1} \eta_{k_1}^T (s - \bar{s}_{k_1}) + \bar{s}_{k_1} \quad (3.1)$$

3.1.2 サブセット間加重方程式

K_1 個のサブセット固有空間に対する射影 s_{k_1} を式 (3.1) によってすべて求めることにより, 各サブセット固有空間を代表する 1 点を決定できる. これらを用いて, 基本加重方程式と同様に, 各サブセットについて 1 つの点 s_{k_1} のみを考えた加

重方程式を以下のように構成できる．

$$\widehat{\mathbf{S}}_{K_1} \mathbf{w}_{K_1} = \widehat{\mathbf{s}} \quad (3.2)$$

$$\text{ただし, } \widehat{\mathbf{S}}_{K_1} = [\widehat{\mathbf{s}}_1 \cdots \widehat{\mathbf{s}}_{K_1}] \quad (3.3)$$

$$\mathbf{w}_{K_1} = [w_1 \cdots w_{K_1}]^\top \quad (3.4)$$

本論文では，式 (3.2) をサブセット間加重方程式と呼ぶ．解は次式で与えられる．

$$\mathbf{w}_{K_1} = \widehat{\mathbf{S}}_{K_1}^+ \widehat{\mathbf{s}} \quad (3.5)$$

サブセット間加重方程式を解くことによって，サブセット間の重みが算出される様子を図 3.1(b) に示す．この例では，3 つのサブセット固有空間についてそれぞれ加重が算出されている．このとき得られる解は，入力画像をサブセットの加重平均で表すための重みである．これは各個人の重みを直接表現するものではないため，人物識別や形状推定に直接用いることはできない．

なお， $K_2 = 1$ の場合，サブセット間加重方程式は基本加重方程式と等価である．

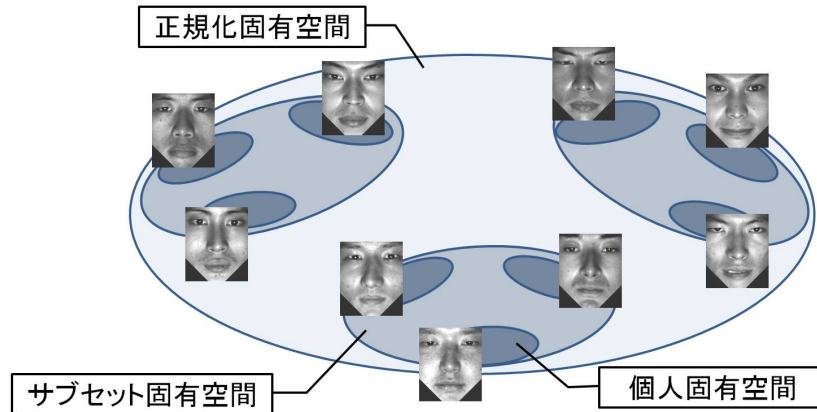
3.1.3 個人固有空間

前節で述べたように，式 (3.2) を解くことで得られる重みは各登録人物の重みを直接表現していない．しかし，サブセット内においても，構成人物毎に重みを求めることができれば，これとサブセット間に割り当てられた重みから各登録人物の重みを算出することが可能である．

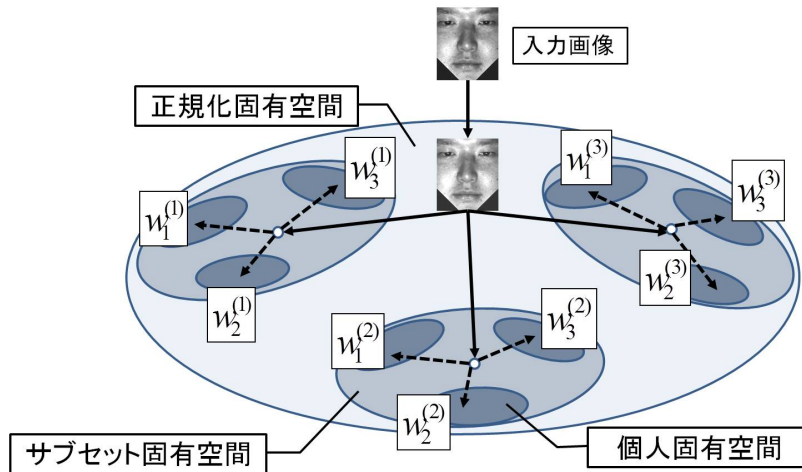
以下では，各サブセット固有空間内に構成人物毎の固有空間を作成する．第 k_1 サブセットのある人物 k_2 に対して照明条件 l で撮影された画像のサブセット固有空間への射影係数 $s_{k_2 l}^{(k_1)}$ が与えられているとする．このとき， $\mathcal{S}_{k_2}^{(k_1)} = \{s_{k_2 l}^{(k_1)} | l = 1, \dots, L\}$ を主成分分析することにより m'' 次元の固有空間 $\langle \bar{s}_{k_2}^{(k_1)}, \eta_{k_2}^{(k_1)} \rangle$ を作成できる．この空間を (サブセット内の) 個人固有空間と呼ぶ．図 3.2 の例では，個人固有空間は，9 人の登録者集合から作成された 3 つのサブセット固有空間内にそれぞれ 3 つずつ作成されている．

入力画像から第 k_1 サブセット固有空間への射影 s_{k_1} が式 (3.1) により与えられたとき，射影係数は $s^{(k_1)} = \eta_{k_1}^\top (s - \bar{s}_{k_1})$ で与えられる．このとき， $s^{(k_1)}$ の個人固有空間 $\langle \bar{s}_{k_2}^{(k_1)}, \eta_{k_2}^{(k_1)} \rangle$ への射影は次式で与えられる．

$$\mathbf{s}_{k_2}^{(k_1)} = \eta_{k_2}^{(k_1)} \eta_{k_2}^{(k_1)\top} (\mathbf{s}^{(k_1)} - \bar{\mathbf{s}}_{k_2}^{(k_1)}) + \bar{\mathbf{s}}_{k_2}^{(k_1)} \quad (3.6)$$



(a) 個人固有空間



(b) サブセット内加重方程式

図 3.2: 個人固有空間とサブセット内加重方程式の例

3.1.4 サブセット内加重方程式

式 (3.6) から個人固有空間に対する射影 $s_{k_2}^{(k_1)}$ をすべて求めることにより、サブセット内の各構成人物を代表する 1 点を決定できる。これらを用いると、各サブセット内の人物について 1 つの代表点 $s_{k_2}^{(k_1)}$ のみを登録とした加重方程式を以下のように構成できる。

$$\widehat{\mathbf{S}}_{K_2}^{(k_1)} \mathbf{w}_{K_2}^{(k_1)} = \widehat{\mathbf{s}}^{(k_1)} \quad (3.7)$$

$$\text{ただし, } \widehat{\mathbf{S}}_{K_2}^{(k_1)} = [\widehat{\mathbf{s}}_1^{(k_1)} \ \dots \ \widehat{\mathbf{s}}_{K_2}^{(k_1)}] \quad (3.8)$$

$$\mathbf{w}_{K_2}^{(k_1)} = [w_1^{(k_1)} \ \dots \ w_{K_2}^{(k_1)}]^\top \quad (3.9)$$

本論文では，式 (3.7) をサブセット内加重方程式と呼ぶ．解は次式で与えられる．

$$\mathbf{w}_{K_2}^{(k_1)} = \widehat{\mathbf{S}}_{K_2}^{(k_1)+} \widehat{\mathbf{S}}^{(k_1)} \quad (3.10)$$

サブセット内加重方程式を解くことによって，サブセット内の各個人の重みが算出される様子を図 3.2(b) に示す．この例では，3つのサブセット固有空間内において，それぞれ3つずつ重みが算出されている．このとき得られる解は，サブセット内部の人物の重みであり，サブセット間の重みと同様に単独で識別や形状推定に用いることはできない．

3.1.5 人物識別

3.1.2 節と 3.1.4 節で求めた 2 種類の重みを組み合わせることで，登録人物全体の重みを算出する方法を述べる．ここで，加重方程式の構成から，サブセット間の重み \mathbf{w}_{K_1} とサブセット内の重み $\mathbf{w}_{K_2}^{(k_1)}$ はどちらも総和が 1 となる．このため，第 k_1 サブセットに割り当てられた重み w_{k_1} を第 k_1 サブセットの内得られた重み $\mathbf{w}_{K_2}^{(k_1)}$ に掛けることで，総和が 1 となる重み \mathbf{w}_K を次式で計算できる．

$$\mathbf{w}_K = [w_1 \cdots w_K]^\top \quad (3.11)$$

$$w_k = w_{k_1} w_{k_2}^{(k_1)} \quad (3.12)$$

ここで，人物 k はサブセットのインデックス k_1 とサブセット内のインデックス k_2 で一意に表されるものとする．

式 (3.12) によって，各登録人物の重みを算出することが可能となった．人物識別は基本加重方程式と同様に式 (2.24) で行う．

3.1.6 階層化照明適応と階層化加重方程式

ここまで述べてきたように，登録人物集合を何人かのサブセットに分割し，2 階層 (サブセット固有空間，個人固有空間) にすることにより，登録を少数に抑えた小規模な加重方程式の構成が可能となった．式 (3.1) と式 (3.6) をあわせて階層化照明適応と呼び，サブセット間加重方程式とサブセット内加重方程式をまとめて階層化加重方程式と呼ぶ．

3.2 直交サブセットを用いた階層化加重方程式の並列化

3.2.1 直交サブセット分割の定義と解法

階層化加重方程式による解法では，登録人物全員を排他的なサブセットに分割する．以下では，この際に作成される分割をサブセット分割と呼ぶ．

サブセット分割は，サブセットを構成する人物を入れ替えることにより複数作成可能である．以下では，複数のサブセット分割を用いて，識別性能を改善する方法を検討する．ここで，サブセット分割をランダムに作成するだけでも，大数の法則によりある程度の識別率向上を期待できるが，ランダムな分割では異なるサブセット分割間の2つのサブセットに同一人物が多数含まれる可能性がある．そこで，このような偶然性を排除するため，直交サブセット分割を次の直交条件により定義する．

直交条件: 異なるサブセット分割に属するサブセット間には共通のメンバーは高々1人しか存在しない．

2つのサブセット分割が直交サブセット分割であることは，以下に定義するサブセットベクトルとサブセット分割行列から確認することができる．

サブセットベクトルは，登録者数を K とした場合， ${}_K C_2$ 次元ベクトルで定義される．サブセットベクトルの各次元は登録者のペアに対応し，2値(当該ペアを含む場合に1，含まない場合に0)を取るものとする．また，サブセット分割内の全てのサブセットから求められるサブセットベクトルを並べた行列をサブセット分割行列と定義する．このとき，2つのサブセット分割行列の内積がゼロ行列であることが2つのサブセット分割が直交サブセット分割であるための必要十分条件になる．

直交サブセット分割の例を図3.3に示す．この例は，25人の人物を5行5列に配置し，各行をサブセットとしたサブセット分割と各列をサブセットとしたサブセット分割である．

なお，登録人物数 K が素数の二乗である場合，ラテン方阵 [23] をベースとした直交サブセット分割を $\sqrt{K} - 1$ 個作成できる．さらに，列および行ごとの分割によっても直交条件を満たすサブセット分割を作成できるため，直交サブセット分割は最大で $\sqrt{K} + 1$ 個構成できる．本論文では，このようにして得られる直交サブセット分割により階層加重方程式を並列させて解くことで，ロバストな顔認識系の構成を目指す．

3.2.2 直交サブセット分割を用いた人物識別

ここでは，直交サブセット分割を用いた人物識別法について述べる．直交サブセット分割を C 個作成した場合を考える．第 $e(= 1, \dots, E)$ サブセット分割から構成される階層化加重方程式の最適解を $\mathbf{w}^{(e)}$ で表す．このとき， E 個の最適解を平均することにより最終的な解が得られる．

$$\mathbf{w} = \frac{1}{E} \sum_{e=1}^E \mathbf{w}^{(e)} \quad (3.13)$$



図 3.3: 直交するサブセット分割の例 ($K=25$ の場合)

ここで、平均操作によって得られた加重ベクトルは元の階層化加重方程式を満たす保証はない。

通常，階層化加重方程式から得られる各最適解は，基本加重方程式とは異なり他人の情報を含む部分空間を加重の算出に使用する．このため，階層化加重方程式の最適解は，基本加重方程式の最適解と比較して認識性能は劣化していると考えられる．一方，従来からパターン認識においては相補性の高い情報を組み合わせることにより最終的な認識性能が向上することが知られている [24]．階層化加重方程式では，各サブセットは直交条件を満たすように作成されるため相補性は高いと考えられ，この効果と大数の法則により最終的な認識性能を改善できると考えられる．

3.2.3 直交サブセットと並列部分射影を用いた人物識別

並列部分射影 [25] を用いた顔認識は，基本加重方程式へも適用でき，併用により識別性能が向上することが確認されている．このことから，階層化加重方程式と直交サブセット分割の並列化を組み合わせた解法においても有効であると考えられる．

ここで，第 $q(= 1, \dots, Q)$ 領域の画像を用いて構成した，第 e サブセット分割の階層化加重方程式で得られる最適な加重を $w_{(e,q)}$ で示す．このとき， EQ 個の解を平均することによって最終的な解が得られる．

$$\mathbf{w} = \frac{1}{EQ} \sum_{e=1}^E \sum_{q=1}^Q \mathbf{w}_{(e,q)} \quad (3.14)$$

ここで， $Q = 7$ (全体と部分を用いた計 7 領域) である，式 (5.5) も式 (5.4) と同様に，元の加重方程式を満たす保証はない．しかし，並列部分射影による解法も，各部分領域をそれぞれ相補性の高い部分領域とすることで認識性能を改善できると考えられる．

3.2.4 計算コストの比較

基本加重方程式の係数行列のサイズが $m+1 \times K$ ($m+1 > K$) である場合，サブセット間加重方程式とサブセット内加重方程式の係数行列のサイズはそれぞれ $m+1 \times K_1$ ， $m'+1 \times K_2$ となる．ここで，直交サブセット分割を用いた解法では， $\sqrt{K} = K_1 = K_2$ となる．また，簡単化のため次元数を $m' = m$ とすると，基本加重方程式，サブセット間加重方程式，サブセット内加重方程式の係数行列のサイズはそれぞれ， $(m+1) \times K$ ， $(m+1) \times \sqrt{K}$ ， $(m+1) \times \sqrt{K}$ となる．

最適解を求める際に QR 分解を用いた場合，基本加重方程式，サブセット間加重方程式，サブセット内加重方程式の計算量はそれぞれ， $O(mK^2)$ ， $O(mK)$ ， $O(mK)$ となるが，階層化加重方程式においてサブセット内加重方程式は \sqrt{K} 個構成されることから総計算量は $O(mK^{3/2})$ となる．さらに，直交サブセット分割を用いて

表 3.1: CMU Multi-PIE データベースで使した ID

| | |
|----------|-------------------------------------|
| Session1 | 1~212, 214~250 |
| Session2 | 251~269, 271~274, 276~280, 284~292, |
| Session3 | 293~295 |



図 3.4: CMU Multi-PIE(ID-002) の原画像と切り出し画像

階層化加重方程式を並列に解く場合，並列数を上限 $(\sqrt{K} + 1)$ まで上げると総計算量は $O(mK^2)$ となり，計算量は基本加重方程式と同じになる．しかし，この場合の厳密な総計算量は $O(mK^2) + O(mK^{3/2})$ であるため，実際の処理時間は基本加重方程式のものより遅くなる．すなわち，実時間性維持のためには並列度の調整が重要となる．

3.3 静止画像を用いた性能評価

ここでは，階層化加重方程式の基本性能を検証するため，CMU Multi-PIE データベース [26] を用いた基礎実験の結果について述べる．

3.3.1 データセットの仕様

以下に，実験に使用するデータセットの概要を示す．

Multi-PIE

公開データベースである CMU Multi-PIE データベースから作成したデータセットである．CMU Multi-PIE [26](Session1~3) に含まれる $289 (= 17^2)$ 人の正面顔を

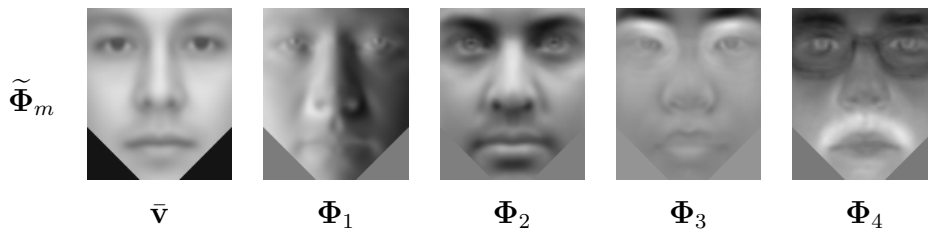


図 3.5: 正規化固有空間 EF289

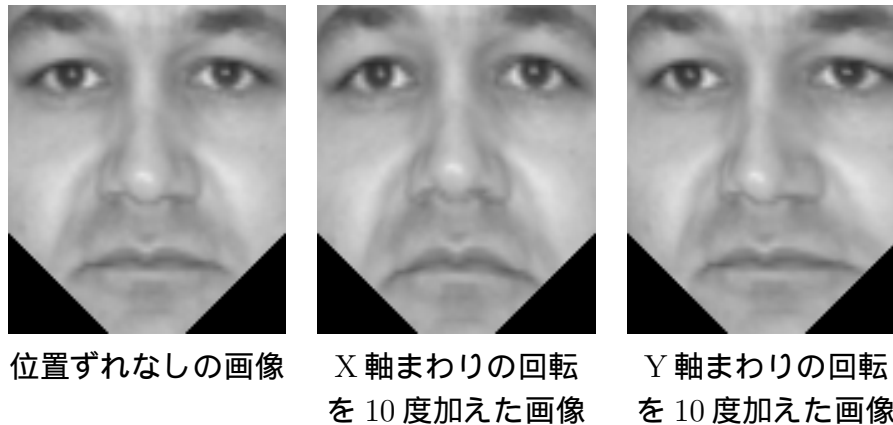


図 3.6: 位置ずれ画像の例

20 種類の照明条件で撮影した画像集合から構成され、以降では Multi-PIE と呼ぶ。また、Session ごとに使用した ID をまとめたものを表 3.1 に示す。

実験用画像の作成方法を述べる。まず、原画像において両眼と口の中央の位置を手動で求め、位置の正規化を 2 次元で行った画像集合を作成する。次に、この画像集合から作成される拡張固有顔を用いて、原画像に対する顔追跡を人物ごとに行い切り出される画像を実験用画像とした。なお、Multi-PIE 中には各登録人物の形状情報は存在しない。このため、5 章、6 章で用いられるデータベース（後述）の登録者の形状を擬似的に使用し、追跡パラメータを人物ごとに手動で調整することにより切り出しを行った。図 3.4 に原画像と切り出し画像の例を示す。

Multi-PIE の各人物について 4 枚の画像を学習データセットとした。学習データセットから 130 次元の正規化固有空間（以下では、EF289 と呼ぶ）を作成した。ここで、固有空間の次元数は累積寄与率が 95% を越える次元数から設定した。EF289 の平均と上位 4 軸を、図 3.5 に示す。EF289 の中にサブセット固有空間、サブセット内の個人固有空間を作成した。ここで、個人固有空間の次元数は 2 とし、サブセット固有空間の次元数は累積寄与率が 99% を越える次元数とした。

3.3.2 テストデータからの位置ずれデータの作成

データセット中の各人物について，学習データとして使用した画像以外をテストデータとした．すなわち，本実験では，Multi-PIE では $16 \times 289 = 4624$ 枚を実験に使用した．

本実験では，3.3.1 節で切り出したデータを位置ずれのない場合とした．ここで，各画像は 3 次元姿勢追跡により切り出されているため，切り出しの際に用いた推定姿勢にノイズを付加することにより位置ずれが発生した画像を生成することができる．本実験では，主に回転誤差に対してどのように性能が変化するかを確かめた．位置ずれ画像は X 軸 (画像の横軸) まわりの回転を 5, 10 度加えたものと Y 軸 (画像の縦軸) まわりの回転を 5, 10 度加えたものを位置ずれ画像として用意した．図 3.6 に位置ずれ画像の例を示す．

3.3.3 比較した識別方法

本実験では，次の 2 種類の識別方法を比較した．

1. 基本加重方程式：289 個の変数の基本加重方程式 (式 (2.23)) を解いた．人物識別は式 (2.24) に従った．
2. 階層化加重方程式：正規化固有顔内にサブセット固有空間を 17 個，各サブセット固有空間内に 17 個の個人固有空間を作成した．これらの空間を用いてサブセット間加重方程式およびサブセット内加重方程式の構成し，加重を算出する．また，直交サブセットの並列数を 1, 5, 10, 18 と変化させながら識別率の変化を調べた．人物識別は式 (2.24) に従った．

3.3.4 階層化加重方程式による静止画像実験の結果

各方法について，識別率を比較すると表 3.2 のようになる．位置ずれがない場合，基本加重方程式では識別率が 99.96% となった．対して，階層化加重方程式では，単独のサブセット分割を用いた場合に 99.81% であり，基本加重方程式よりも低い識別率となった．しかし，直交サブセット分割を用いて並列化した場合，識別率は 100% を達成できており，並列化による識別性能の改善を確認できた．

また，位置ずれが発生した場合は，基本加重方程式と階層化加重方程式のどちらにおいても性能が劣化していることが判る．ここで，直交サブセット分割を用いて並列化を行った場合，階層化加重方程式は基本加重方程式と比較して識別率の低下を抑えることができた．特に，Y 軸まわりの回転に 5 度加えた位置ずれ画像を用いた実験の場合，階層化加重方程式では，並列化により識別率 100% を達成できた．

表 3.2: 階層化加重方程式と基本加重方程式の識別率 (%) の比較

| (a) X 軸まわりの回転にノイズを与えた場合 | | | |
|-------------------------|----------------|-------|-------|
| Method | 回転ずれ: X 軸まわり回転 | | |
| | なし | 5 度 | 10 度 |
| 基本加重方程式 | 99.96 | 99.59 | 96.13 |
| 階層化加重方程式 | 99.81 | 98.92 | 86.91 |
| 5 並列 " | 100.00 | 99.96 | 97.90 |
| 10 並列 " | 100.00 | 99.98 | 98.59 |
| 18 並列 " | 100.00 | 99.98 | 98.79 |

| (b) Y 軸まわりの回転にノイズを与えた場合 | | | |
|-------------------------|----------------|--------|-------|
| Method | 回転ずれ: Y 軸まわり回転 | | |
| | なし | 5 度 | 10 度 |
| 基本加重方程式 | 99.96 | 99.78 | 98.66 |
| 階層化加重方程式 | 99.81 | 99.55 | 95.72 |
| 5 並列 " | 100.00 | 100.00 | 99.35 |
| 10 並列 " | 100.00 | 100.00 | 99.78 |
| 18 並列 " | 100.00 | 100.00 | 99.76 |

表 3.2 節に示す結果から階層化加重方程式は、画像の劣化のあるなしに関わらず、並列化により識別性能を改善でき、3.2.3 節の操作が妥当であることを確認できた。なお、階層化加重方程式では、最大 126(部分領域数 7×直交サブセット分割数 18)まで並列数を増やすことができるが、実験結果から 35 並列(部分領域数 7×直交サブセット分割数 5)もあれば充分識別性能を改善できることが確認できた。

本データセットでは、次元数と登録者数の関係から基本加重方程式は不足決定系に陥ってしまう。一方、階層化加重方程式では、サブセットを用いて加重方程式のサイズを小さくできるため不足決定系に陥ること無く識別系を構成できた。このことから、本手法は、登録者が増大した場合においても過剰決定系の枠組みを維持したまま識別を行うための有効な手法であることが確認できた。

なお、4 章では過剰決定だけでなく不足決定系の枠組みを用いて、性能の改善を目指した手法も提案している(詳細については、4 章を参照されたい)。

3.3.5 処理時間の比較

ここでは、処理時間の比較を行った(CPU は Core i7 CPU 980X(3.33GHz)を使用した)。登録者 289 人、正規化固有顔の次元数が 130 次元である場合、基本加重

方程式のサイズは 131×289 となる。このとき、基本加重方程式の処理時間は約 4.8 ミリ秒であった。これに対して、階層化加重方程式においては、サブセット間加重方程式のサイズが 131×17 、サブセット内加重方程式のサイズが $37 \sim 42 \times 17$ となる。このとき、直交サブセット分割数を 1, 5, 10, 18 とした場合の処理時間はそれぞれ、約 3.3 ミリ秒、約 7.5 ミリ秒、約 14.7 ミリ秒、約 27.3 ミリ秒であり、並列度が増した場合の階層化加重方程式の処理時間は基本加重方程式よりも遅かった。

このことから、階層化加重方程式を用いた顔追跡・認識融合系において実時間性を保つためには、直交サブセット分割の並列数を適切に調整する必要がある。

3.4 実動画像上での性能評価

ここでは、実動画像を用いた顔追跡・認識実験を行った。追跡アルゴリズムは基本的に [17, 18] に従い、加重方程式の解法の違いによる性能の比較を行った結果について述べる。

3.4.1 学習データセットおよび拡張固有顔

以下に、実験に使用するデータセットの概要を示す。

Data10x10

実動画像実験のために作成した独自データセットである。10 人の 10 表情の正面顔を 24 種類の照明条件で撮影した画像集合と 10 人の 10 表情の立体顔モデル集合から構成され、以降では Data10x10 と呼ぶ。

実験用画像の作成方法を述べる。Data10x10 においては、各登録について形状データが存在するため、まず、各立体顔モデルにおいて両眼と口の中央の位置を手動で求め、正規化顔モデルを作成した。次に、各正規化顔モデルに対応する人物および表情の原画像(集合)に対して顔追跡を行い、切り出される画像を拡張固有顔の作成に用いた。図 3.7 に、Data10x10 の例を示す。

Data10x10 から 140 次元の正規化固有顔(以下では、EF10x10 と呼ぶ)を作成した。ここで、固有空間の次元数は累積寄与率が 95% を越える次元数から設定した。EF10x10 の中に正規化固有顔内の個人固有空間、サブセット固有空間、サブセット固有空間内の個人固有空間を作成した。ここで、個人固有空間の次元数は 4 とし、サブセット固有空間の次元数は累積寄与率が 99% を越える次元数とした。

これらと各登録の形状データから拡張固有顔を作成し(以降、AEF10x10 と呼ぶ)実験に使用した。図 3.8 に AEF10x10 の中心と最初の 4 軸を示す。

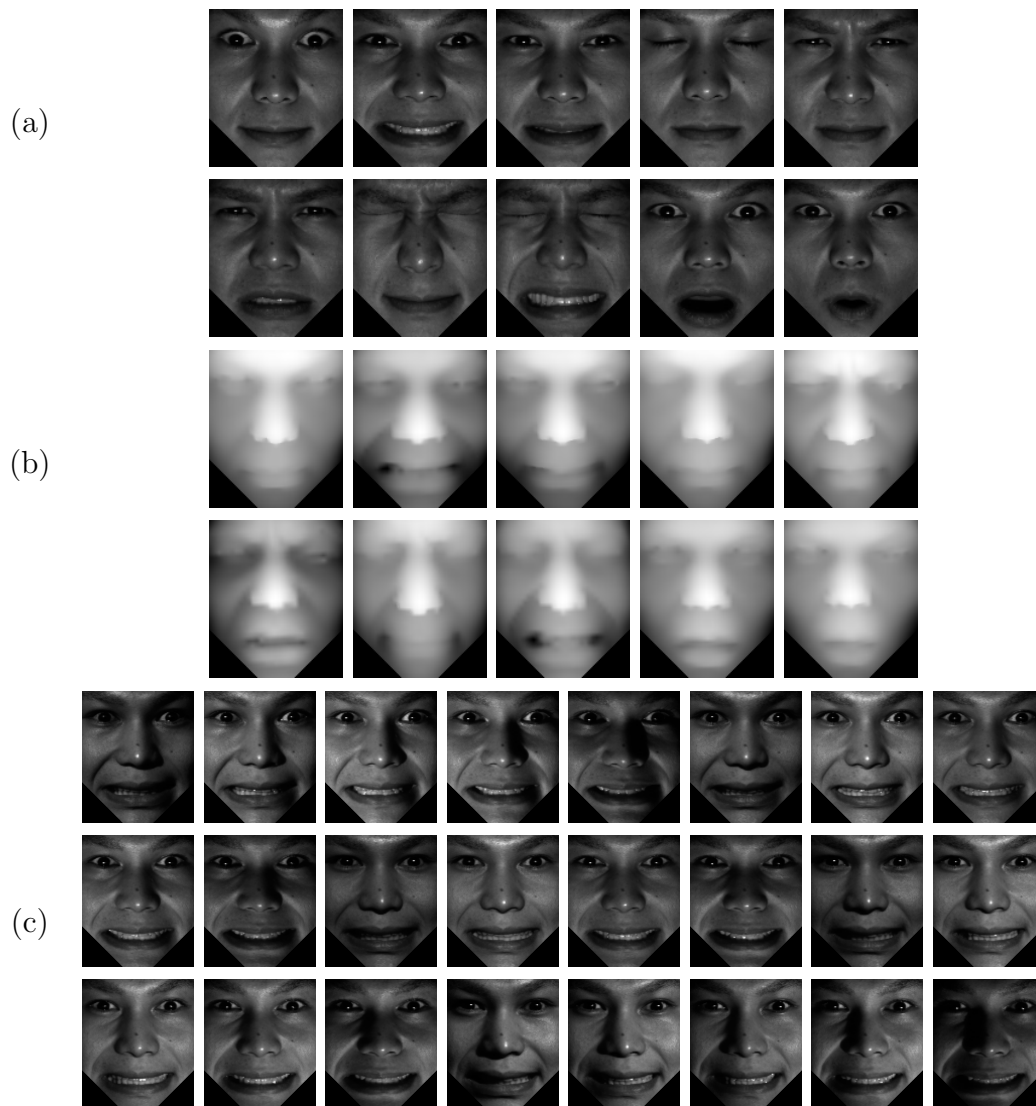


図 3.7: Data10x10 の例

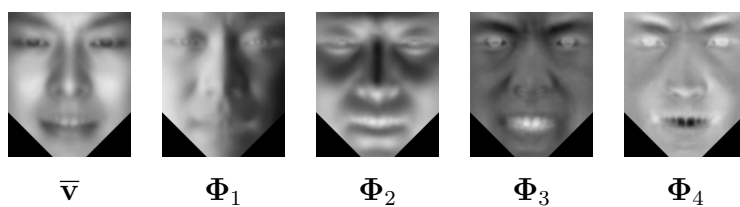


図 3.8: 拡張固有顔 AEF10x10 の中心および最初の 4 軸

3.4.2 テストシーケンス

実験には、天井照明(蛍光灯)のみが存在する窓のない部屋で撮影した10人(1200フレーム, 30fps)の動画像を用いた。動画像中では、人物は自由に動き、様々な表情変化をするようにした。また、天井照明以外の点光源(ハロゲンランプ)を人物の前で移動させているため、動画像中の照明条件は変化している。

各動画像に対し、追跡開始フレームを第1フレーム, 第10フレーム... , 第892フレームと変化させることで各動画像から100種類のテストシーケンスを作成した。従って、10人分の動画像から作成したテストシーケンスの総数は1000である。

3.4.3 比較した識別方法

静止画像実験を元に、3.3.3節で示した方法の中から、次の2種類の識別方法を比較した。

1. 基本加重方程式(過剰決定系): 141(次元) \times 100(変数)のサイズの基本加重方程式を解く。また、Data10 \times 10を用いた実験では、人物ごとに表情に関する加重和を求めてから、式(2.24)で人物識別を行った。
2. 5並列階層化加重方程式: 正規化固有顔内にサブセット固有空間を10個、各サブセット固有空間内に10個の個人固有空間を作成し、サブセット間加重方程式、サブセット内加重方程式の作成に用いた。また、直交サブセットの並列数は5としている、すなわち5つの階層化加重方程式の最適解を平均したものが識別に用いられる。なお、人物識別は基本加重方程式と同様の方法で行った。

3.4.4 動画像に対する識別判定

顔追跡・認識融合系では、フレーム間で姿勢追跡精度が一定しないため、位置ずれが不規則に発生する。一方、静止画像を対象とした識別実験結果から、どの方法においても位置ずれが大きい場合には識別性能が低下することが判っている。これを前提としてできるだけ短時間で正確な識別を行う必要がある。

姿勢推定精度を算出するのは簡単ではないため、本論文では位置ずれの指標として正規化相関を用いる。本論文では、次のように正規化相関を計算している。まず、入力画像から得られる推定姿勢と推定形状に基づいて、顔の向きを正面方向に変換した場合の画像(\mathbf{V}_{crop})とこの画像の正規化固有顔への射影($\mathbf{V}_{crop}^{\$}$)を取得する。次に、 \mathbf{V}_{crop} と $\mathbf{V}_{crop}^{\$}$ を用いて、正規化相関 $C(\mathbf{V}_{crop}, \mathbf{V}_{crop}^{\$})$ を次式で計算する。

$$C(\mathbf{V}_{crop}, \mathbf{V}_{crop}^{\$}) = \frac{\mathbf{V}_{crop}^{\top} \mathbf{V}_{crop}^{\$}}{\|\mathbf{V}_{crop}\| \|\mathbf{V}_{crop}^{\$}\|} \quad (3.15)$$

また、この系では最大加重を与える人物を識別に用いるが、位置ずれが大きい場合には加重が複数人物に分散する傾向がある。従って、最大加重の値が小さい場合には信頼性の低い判定しか行えない。一方、最大加重の値が大きい場合には信頼性の高い判定が1フレームからでも可能である。さらに、この系では正面向きの顔画像で識別処理を行うため、推定姿勢が正面向きから離れている場合には、信頼性の高い識別を行える可能性は小さい。これらの特性を反映して、各フレームを次の3種類に分類する。

1. 決定フレーム: 相関が0.992以上、正面姿勢からの角度が30度以下でかつ、最大加重が決定閾値以上であり、最大加重と第2位の加重との比が閾値以下
2. 有効フレーム: 決定フレーム以外で、相関が0.992以上、角度が30度以下でかつ、最大加重が有効閾値以上
3. 棄却フレーム: 上記以外のフレーム

なお、本論文では誤識別率の発生を抑えるために予備実験から上記のパラメータを設定した。基本加重方程式および階層化加重方程式に用いる決定閾値、比率、有効閾値は、それぞれ0.70, 0.50, 0.65と0.35, 0.33, 0.30とした。

これらの2フレームを用いて、動画像に対する識別は次の2つの規則に従って行う。

- (I) 決定フレームにおいては、最大加重を与える人物を選択する。
- (II) 有効フレームにおいては、直前の2有効フレームと同じ人物が最大加重を与える場合についてのみ、その人物を選択する。

3.4.5 動画像実験の結果

実験では、各テストシーケンスについて、比較手法のそれぞれについて1回ずつの追跡・認識を行った。

追跡開始から識別完了後までの状況を確認したところ、基本加重方程式においては、1000シーケンス中の982シーケンスで正解、残りの18シーケンスで棄却であった。これに対し、5並列階層化加重方程式においては、1000シーケンス中の976シーケンスで正解、残りの24シーケンスで棄却であった。

また、追跡開始から300フレームまでの推移を詳しく見ると、基本加重方程式については、30フレーム後までに約50%、90フレーム後までに約78.0%のシーケンスで正解が得られた。これに対し、階層化加重方程式については、30フレーム後までに約79%、90フレーム後までに約93%のシーケンスで正解が得られた。これらの結果から、階層化加重方程式は棄却率の上昇は見られたものの、基本加重方程式と比べて早い段階で識別を完了可能であることが確認できた。これは、階層

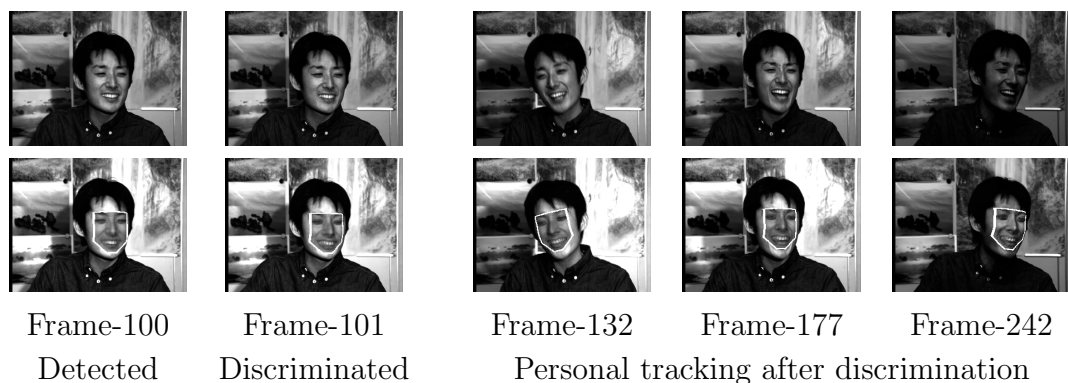


図 3.9: 階層化加重方程式を用いた顔追跡・認識の例

化加重方程式は基本加重方程式よりも位置ずれに対してロバストであるため、位置ずれノイズを含む画像に対しても正確な加重を算出できたからであると考えられる。

図 3.9 に、階層化加重方程式による顔追跡・認識の結果を例示する。図の左から 1 列目は初期フレームの顔の位置，図の左から 2 列目は識別完了時の顔の位置，3 列目以降は識別後の追跡結果を示している。これらの結果から本人に近い顔モデルが顔追跡・認識融合系により与えられていることが判る。

1 フレームあたりの 3 次元疎固有テンプレート追跡系と加重方程式による識別系の合計処理時間は、基本加重方程式を用いた場合に 28.2 ミリ秒，階層化加重方程式を用いた場合に 30.6 ミリ秒であった。これらの結果から、階層化加重方程式は、基本加重方程式に比べ処理時間が若干大きくなっているが、100 人規模の顔追跡・認識融合系において実時間を達成可能であることが確認できた。

3.5 まとめ

本章では、顔認識・融合系において、登録者数が増大した場合に発生する問題を回避するための方法として階層化加重方程式を提案し、その有効性を 289 人規模の静止画像データベースおよび 100 登録規模 (10 人 × 10 表情) の実動画像データベースで確認した。

この方法では、正規化固有顔内に直接作成していた個人固有空間を登録者集合のサブセットから作成される部分空間内に作成することで、登録者数を小規模に抑えた加重方程式構成できる。また、複数のサブセット (直交サブセット分割) を構成できることを利用した並列実装も可能である。

静止画像データベースおよび動画像データベースを用いた実験の結果から、複数の直交サブセットを用いた階層化加重方程式は、登録者数の増加にともなう問題を回避するとともに画像の位置ずれなどのノイズに対しても有効であった。こ

これらの結果は、階層化加重方程式が、過剰決定系の枠組みで登録者数を大規模化するのに適した手法であることを示している。

階層化加重方程式を単独で用いた場合は基本加重方程式よりも処理時間が小さかったが、直交サブセットの並列数を上げた場合には、基本加重方程式よりも処理時間が大きくなった。一方、この問題は、適切な並列数(本実験では、並列数5で実時間性の維持と識別性能の改善の両方を達成)を選択することにより回避可能であった。現在の処理時間の比較結果から、階層化加重方程式では、100~289人程度が実時間性を維持できる最大の登録者数となっており、さらなる大規模登録での実時間系の構成は今後の課題である。

なお、階層化加重方程式は、原理的に3以上の階層を持つことも可能であるが、これに関する検討は本章では行わず5章において行う。

第4章 並列不足決定系による顔追跡・認識融合系の高性能化

4.1 並列不足決定系

Oka-Shakunaga[17, 18] は、基本加重方程式を過剰決定系で構成し、人物識別や形状推定に用いている。一方、基本加重方程式を過剰決定系で精度良く解くには、登録者数に比べて大きな次元数を持つ正規化固有顔を準備する必要がある。このため、過剰決定系では、登録者数が増大すると必要な次元数も増大し、計算コストが増加するため系の実時間性を保つことは容易でなくなる。また、過剰決定系を維持するために必要な次元数を確保できず、基本加重方程式が不足決定系となってしまう場合もある。

この問題に対して、3章では、登録者集合のサブセットを用いて小規模な加重方程式を構成することができる階層化加重方程式を提案している。この方法により、過剰決定系を維持したまま登録者数の増加に対応できることが示されている。しかし、階層化加重方程式による解法では、識別性能を改善するための並列処理によって、計算量が大きくなってしまう。

さらに、過剰決定系においては、登録者が増加した場合に固有顔の次元数が系の構成にとって大きな制約となる。一方、不足決定系においてはこの制約がないため、原理的にはより柔軟に系を構成できると思われる。

本章では、不足決定系の枠組みにおいて登録者数の増加に対応し、より柔軟に識別系を構成するための方法 [28, 29] を検討する。具体的には、単一の不足決定系の加重方程式を用いるのではなく、与えられた正規化固有顔の中で複数の不足決定系を構成し、個々の不足決定系における最適解から平均値を求め、元の加重方程式の解の代わりとして用いる方法を提案する。

4.1.1 不足決定系における加重方程式の解

$m+1 < K$ のとき、基本加重方程式(式(2.23))は不足決定系となる。このとき、(w に関する)解空間は $(K-m-1)$ 次元となり、疑似逆行列による解は $1^T w = 1$ と $s = S_K w$ を満たす解空間内で $w^T w$ を最小とする w を与える。

4.1.2 バイアス付き加重方程式

基本加重方程式が不足決定系である場合，疑似逆行列による解は $\mathbf{w}^\top \mathbf{w}$ を最小化するが，この計算には入力画像 \mathbf{v} と登録画像 \mathbf{v}_k の類似性が考慮されていない．ここで，入力画像と登録画像の類似性を適切に考慮した最適化を実現できれば，基本加重方程式よりも優れた解が得られると考えられる．

本論文では， \mathbf{s} と \mathbf{s}_k の距離の逆数を対角要素とする K 行 K 列の対角行列 \mathbf{B} を導入し，入力画像 \mathbf{v} と登録画像 \mathbf{v}_k の類似性を考慮した加重の調整を行う．

$$\mathbf{B} = \text{diag}(d^{-1}(\mathbf{s}, \mathbf{s}_1) \cdots d^{-1}(\mathbf{s}, \mathbf{s}_K)) \quad (4.1)$$

$$\text{where } d(\mathbf{s}, \mathbf{s}_k) = \sqrt{(\mathbf{s} - \mathbf{s}_k)^\top (\mathbf{s} - \mathbf{s}_k)} \quad (4.2)$$

ここで， $\mathbf{s}_k = \mathbf{s}$ である場合には，適当な大数を \mathbf{B} の第 k 対角要素に設定するものとする．

$\mathbf{w} = \mathbf{B}\mathbf{w}'$ を式 (2.23) に代入すると次式が得られる．

$$\hat{\mathbf{S}}_K \mathbf{B}\mathbf{w}' = \hat{\mathbf{s}} \quad (4.3)$$

式 (4.3) をバイアス付き加重方程式と呼ぶ．式 (4.3) の最適解から加重ベクトルは次式で与えられる．

$$\mathbf{w} = \mathbf{B}[\hat{\mathbf{S}}_K \mathbf{B}]^+ \hat{\mathbf{s}} \quad (4.4)$$

式 (4.4) で示される最適解は基本加重方程式の解空間内にあり， $\mathbf{w}'^\top \mathbf{w}' = \mathbf{w}^\top \mathbf{B}^{-2} \mathbf{w}$ を最小化することになるため，入力画像と各登録画像の距離を考慮に入れた最適化を実現できる．

4.1.3 最近傍法との関係

バイアス付き加重方程式は m 次元固有顔において定義され，バイアスとして同じ空間での距離を用いている．ここで，距離をそのままにしておいて加重方程式の次元数を減少させていくと，その極限状態 ($m = 0$ の場合) では， $\tilde{\Phi}_0 = [\Phi_0 \bar{\mathbf{v}}] = [\bar{\mathbf{v}}]$ となる．この場合，式 (2.7) と (2.29) はスカラーとなるため，正規化を施すことにより $\hat{\mathbf{S}}_K = \mathbf{1}$ および $\hat{\mathbf{s}} = \mathbf{1}$ となる．このとき，式 (4.4) は次式のように変形される．

$$\begin{aligned} \mathbf{w} &= \mathbf{B}[\mathbf{1}^\top \mathbf{B}]^+ = \mathbf{B}(\mathbf{1}^\top \mathbf{B})^\top [(\mathbf{1}^\top \mathbf{B})(\mathbf{1}^\top \mathbf{B})^\top]^{-1} \\ &= \mathbf{B}\mathbf{B}^\top \mathbf{1}[\mathbf{1}^\top \mathbf{B}\mathbf{B}^\top \mathbf{1}]^{-1} = \mathbf{B}^2 \mathbf{1}[\mathbf{1}^\top \mathbf{B}^2 \mathbf{1}]^{-1} \\ &= \frac{1}{\sum_k d^{-2}(\mathbf{s}, \mathbf{s}_k)} \begin{bmatrix} d^{-2}(\mathbf{s}, \mathbf{s}_1) \\ \vdots \\ d^{-2}(\mathbf{s}, \mathbf{s}_K) \end{bmatrix} \end{aligned} \quad (4.5)$$

この場合，加重方程式による人物識別 (式 (2.24)) は最近傍法の結果と等価になる．

この結果から，バイアス付き加重方程式は最近傍法の拡張の一つであると解釈できる．すなわち，式 (4.5) において次元数を増加させることにより，最近傍法を入力画像と未知画像間の距離だけではなく固有顔の情報を考慮した識別系に拡張できることが判る．

4.1.4 並列不足決定系の構成

バイアス付き加重方程式では，未知サンプルと登録サンプル間の距離 (式 (4.2)) は加重の調整において重要な役割を果たすため，適切な次元の固有顔で距離を正確に測定するのが望ましい．また，処理時間の観点から，不足決定系は距離を計測した次元数よりも低い次元数で構成することが望ましい．従って，適切な次元の固有顔中に独立な部分空間を複数取って独立に加重方程式を構成し，各解の平均を取ることが処理時間や識別性能の観点から望ましいと考えられる．

ここで， m 次元固有顔の中に独立な J 個の m' 次元部分空間を考える ($m = Jm'$ とする) と， J 並列の不足決定系を構成できる．具体的には，基本加重方程式 (式 (2.23)) を次式のように書き換えた場合を考える．

$$\begin{bmatrix} \mathbf{S}_K^{(1)} \\ \vdots \\ \mathbf{S}_K^{(J)} \\ \mathbf{1}^\top \end{bmatrix} \mathbf{w} = \begin{bmatrix} \mathbf{s}^{(1)} \\ \vdots \\ \mathbf{s}^{(J)} \\ 1 \end{bmatrix} \quad (4.6)$$

ここで， $\mathbf{S}_K^{(j)}$ と $\mathbf{s}^{(j)}$ はそれぞれ \mathbf{S}_K と \mathbf{s} の第 j 番目の m' 行の部分行列，および部分ベクトルを表す．この時，第 j 不足決定系は次式で与えられる．

$$\widehat{\mathbf{S}}_K^{(j)} \mathbf{w}^{(j)} = \widehat{\mathbf{s}}^{(j)} \quad (4.7)$$

$$\text{where } \widehat{\mathbf{S}}_K^{(j)} = \begin{bmatrix} \mathbf{S}_K^{(j)} \\ \mathbf{1}^\top \end{bmatrix} \quad (4.8)$$

$$\text{and } \widehat{\mathbf{s}}^{(j)} = \begin{bmatrix} \mathbf{s}^{(j)} \\ 1 \end{bmatrix} \quad (4.9)$$

ここで， $\mathbf{w}^{(j)}$ は第 j 不足決定系で求められる最適解を表し， $\mathbf{w}^{(j)} = [w_1^{(j)} \cdots w_K^{(j)}]^\top$ である．

対角行列 \mathbf{B} を導入すると，バイアス付きの第 j 不足決定系が次式で得られる．

$$\begin{bmatrix} \mathbf{S}_K^{(j)} \\ \mathbf{1}^\top \end{bmatrix} \mathbf{B} \mathbf{w}^{(j)} = \begin{bmatrix} \mathbf{s}^{(j)} \\ 1 \end{bmatrix} \quad (4.10)$$

なお，本論文の並列実装においては，各不足決定系は同じ距離 (m 次元固有顔内の距離) を用いてバイアス付き加重方程式を解くこととする．

最終的な解 (加重ベクトル) は， J 個の不足決定系の解の平均により求められる．

$$\mathbf{w} = \frac{1}{J} \sum_{j=1}^J \mathbf{w}^{(j)} \quad (4.11)$$

通常，不足決定系の (バイアス付き) 加重方程式から得られる各最適解は，使用することのできる次元数が少ない．また，不足決定系では最適解は $\mathbf{w}^{(j)\top} \mathbf{w}^{(j)}$ を最小化するため，基本加重方程式 (過剰決定系) の最適解と比較して認識性能は劣化していると考えられる．一方，従来からパターン認識においては相補性の高い情報を組み合わせることにより最終的な認識性能が向上することが知られている [24]．並列不足決定系では，各部分空間は正規直交であるため相補性は高いと考えられ，この効果により認識性能を改善できると考えられる．

4.1.5 並列不足決定系における人物識別

並列不足決定系においても，並列部分射影 [21] を上述の議論と組み合わせることにより，さらなる識別性能の向上を期待できる．ここで，第 q 領域 (ただし， $q = 1 \dots Q$) の画像の部分 s -射影の正規化表現を \hat{s}_q で表す．

次に，各 s_q を式 (2.30) により個人固有空間に射影した結果を s_{qk} とし，加重方程式を各不足決定系で解くことで，第 q 部分画像についての最適加重を求めることができる．ここで，第 q 部分画像の第 j 並列目の最適加重を $\mathbf{w}_q^{(j)}$ で示すことにすると，最終的な解は， QJ 個の不足決定系の解の平均により求められる．

$$\mathbf{w} = \frac{1}{QJ} \sum_{q=1}^Q \sum_{j=1}^J \mathbf{w}_q^{(j)} \quad (4.12)$$

ここで，並列不足決定系においても $Q = 7$ とした．

並列部分射影による解法も，各部分領域をそれぞれ相補性の高い部分領域とすることで認識性能を改善できると考えられる．

4.2 静止画像を用いた基礎性能評価

並列不足決定系の基本性能を検証するために静止画像を用いた実験を行った．

4.2.1 実験条件

実験条件は，3.3 節に従う．以下に簡単にまとめる．

使用したデータベース

本実験では，3.3 節に示した使用した静止画像実験用データベース (Multi-PIE) を使用した．

Multi-PIE は，公開データベースである CMU Multi-PIE [26](Session1~3) に含まれる $289 (= 17^2)$ 人の正面顔を 20 種類の照明条件で撮影した画像集合から構成される．Multi-PIE は，あらかじめ人物ごとに顔追跡系をベースとした位置の正規化が行われており，位置正規化が施された画像集合を実験に使用した．Multi-PIE の各人物について，4 枚を登録画像，16 枚をテスト画像とした．

学習画像集合 ($289 \times 4 = 1156$ 枚) から 130 次元の正規化固有顔 EF289 を，また，EF289 内に各人物について 2 次元の個人固有空間を作成し，基本加重方程式や並列不足決定系の構成に使用した．

テスト画像

本実験では，位置ずれのないテスト画像集合 ($289 \times 16 = 4624$ 枚) と，位置ずれの発生したテスト画像集合を使用した．階層化加重方程式との比較のため，本実験でも回転誤差に対してどのように性能が変化するかを確かめた．3.3 節同様に，位置ずれ画像は X 軸 (画像の横軸) まわりの回転を 5, 10 度加えたものと Y 軸 (画像の縦軸) まわりの回転を 5, 10 度加えたものを位置ずれ画像として用意した．

4.2.2 比較した識別方法

本実験では，次の 4 種類の識別方法を比較した．

1. 基本加重方程式: 289 個の変数の基本加重方程式 (式 (2.23)) を解いた．人物識別は式 (2.24) に従った．
2. 並列不足決定系 (バイアスなし): EF289 (130 次元) を 5 つの部分空間 (各 26 次元) に分割し，各部分空間で加重方程式 (サイズ: $(27) \times 289$) を構成した．人物識別 (式 (2.24)) には，各不足決定系から得られる最適加重を平均したものをを用いた．また，不足決定系の並列数を 1 ~ 5 と変化させながら識別率の変化も調べた．
3. 並列不足決定系 (バイアスあり): EF289 (130 次元) を 5 つの部分空間 (各 26 次元) に分割し，各部分空間でバイアス付き加重方程式 (サイズ: $(27) \times 289$) を構成した．ここで，バイアスの計算には 130 次元の正規化固有顔内で求められる s と s_k を使用した．人物識別 (式 (2.24)) には，各バイアス付き加重方程式から得られる最適加重を平均したものをを用いた．また，バイアス付き加重方程式の並列数を 1 ~ 5 と変化させながら識別率の変化も調べた．

表 4.1: 静止画像実験結果 (%)

(a) X 軸まわりの回転にノイズを与えた場合 (括弧内はバイアス付き加重方程式の識別率)

| Method | 回転ずれ : X 軸まわり回転 | | |
|--------------|-----------------|--------------|--------------|
| | なし | 5 度 | 10 度 |
| 基本加重方程式 | 99.96 | 99.59 | 96.13 |
| 5 並列階層化加重方程式 | 100.00 | 99.96 | 97.90 |
| 単独不足決定系 | 94.70(98.94) | 86.25(94.92) | 64.53(80.95) |
| 2 並列 " | 99.70(99.98) | 97.51(99.13) | 85.01(91.85) |
| 3 並列 " | 99.91(100.00) | 99.39(99.76) | 91.83(95.37) |
| 4 並列 " | 100.00(100.00) | 99.65(99.85) | 93.86(96.65) |
| 5 並列 " | 99.98(100.00) | 99.81(99.94) | 95.76(97.40) |
| 最近傍法 | 99.61 | 97.49 | 84.69 |

(b) Y 軸まわりの回転にノイズを与えた場合 (括弧内はバイアス付き加重方程式の識別率)

| Method | 回転ずれ : Y 軸まわり回転 | | |
|--------------|-----------------|---------------|--------------|
| | なし | 5 度 | 10 度 |
| 基本加重方程式 | 99.96 | 99.78 | 98.66 |
| 5 並列階層化加重方程式 | 100.00 | 100.00 | 99.35 |
| 単独不足決定系 | 94.70(98.94) | 90.64(97.60) | 79.22(90.94) |
| 2 並列 " | 99.70(99.98) | 99.05(99.76) | 95.16(97.60) |
| 3 並列 " | 99.91(100.00) | 99.85(99.98) | 97.62(98.70) |
| 4 並列 " | 100.00(100.00) | 99.96(100.00) | 98.51(99.35) |
| 5 並列 " | 99.98(100.00) | 99.96(100.00) | 99.03(99.44) |
| 最近傍法 | 99.61 | 98.88 | 93.30 |

4. 5 並列階層化加重方程式: EF289 内にサブセット固有空間を 17 個, 各サブセット固有空間内に 17 個の個人固有空間を作成した. これらを用いて 5 並列の階層化加重方程式を構成し加重を計算した. 人物識別 (式 (2.24)) には, 各階層化加重方程式から得られる最適加重を平均したものをを用いた.
5. 最近傍識別: EF289 において, s と s_k の距離を求め, 最も小さいものが属する人物を選ぶ.

4.2.3 並列不足決定系による静止画像実験の結果

上述の5種類の方法について、識別率を比較すると表 4.1 のようになる。位置ずれがない場合は、最近傍法を除いてどの方法についても良い結果が得られている。なお、位置ずれが発生した場合は、基本加重方程式やバイアスなしの並列不足決定系では誤識別が生じていることが判る。対して、バイアス付き並列不足決定系については、並列数が少ない場合に誤識別が生じているが、誤識別率は低減しており、特に、5並列の系については、Y軸まわりの回転に5度のノイズが加わった場合でも5並列の階層化加重方程式と同様に、誤識別が発生することはなかった。

表 4.1 に示すとおり、並列不足決定系では、並列度が上がるのに伴って識別性能が改善していることが分かる。単独系では従来法と同等、もしくはそれ以下の識別率しか得られない場合においても、並列度が増えるに伴って識別率が高まっていることが判る。

バイアスなしの並列不足決定系とバイアス付き並列不足決定系の結果からバイアスの効果が確認できた(すなわち、バイアスを用いることで不足決定系において解を修正できることが判った)。また単独系と並列系の識別結果から並列度を高めることの効果を確認できた(なお、この方法では、基本加重方程式ではただか7並列の認識系を最大35並列(5並列×7並列)まで増加させている)。これらの識別結果は、4.1.4節と4.1.5節の操作が妥当であることを示している。

4.2.4 次元数と並列度の変化による識別率と処理時間の変化

Multi-PIEにおいて、並列不足決定系の次元数 m' と並列度 J を変化させた場合の識別率の変化を図 4.1 に、処理時間の変化を図 4.2 に示す。ここで、これらの図の横軸は総次元数(部分空間の次元数×並列数)である。図 4.1 から、どの次元数を用いた場合でも並列度を増加させることで識別率が向上していることを確認できた。一方で、次元数の変化はほとんど識別率に影響を及ぼさなかった(単独系および少数並列時において若干の差がでるが、並列度を上げることで性能の差はほとんどなくなった)。また、ノイズがある場合とない場合でも同様の傾向が確認できた。これらの結果から、識別率の向上に関しては次元数よりも並列度の調整が重要であることがわかった。

処理時間(登録者数289人)は、基本加重方程式を用いた場合に約4.8msであるのに対して、単独系の不足決定系を用いた場合で約2.2ms~3.0ms、最大並列の並列不足決定系を用いた場合では2.9~3.5msであった。これらの結果から、並列不足決定系は基本加重方程式や階層化加重方程式よりも高速に動作することが確認できた。

前節と同様に、実験結果は並列不足決定系の有効性と4.1.4節のねらいが妥当であることを示していると考えられる。

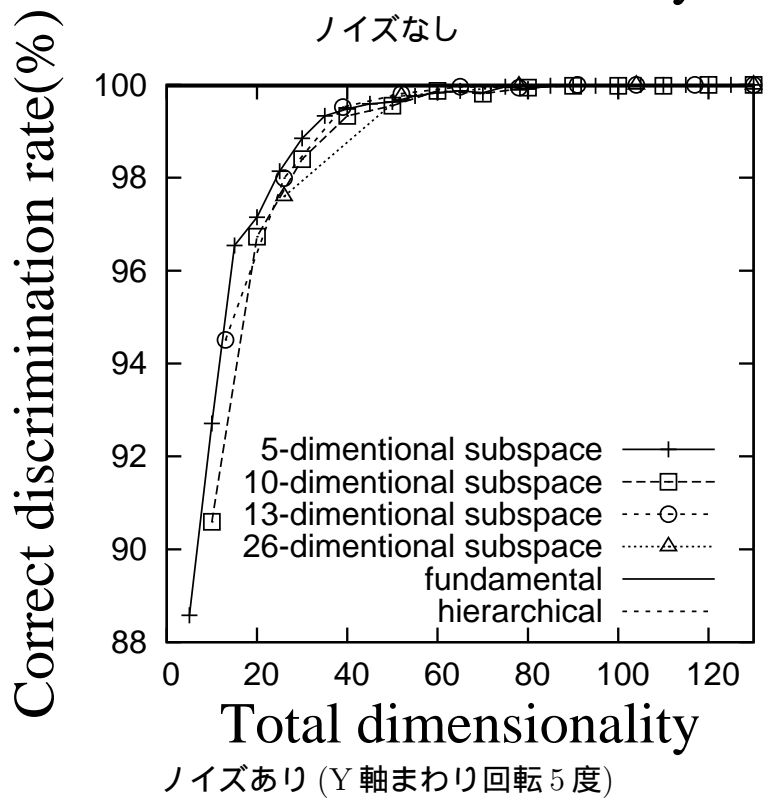
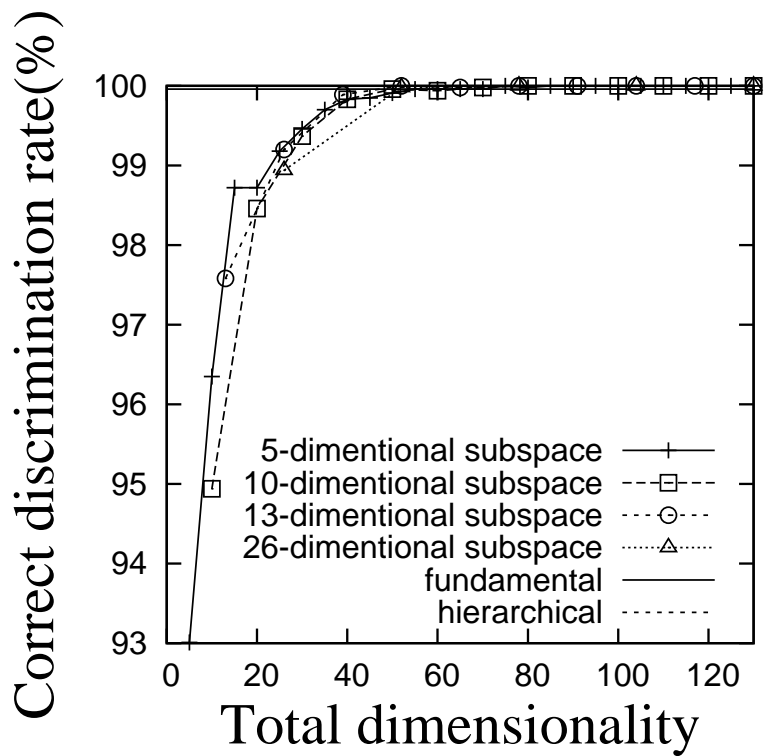


図 4.1: 次元数と並列度を变化させた場合の識別率の変化

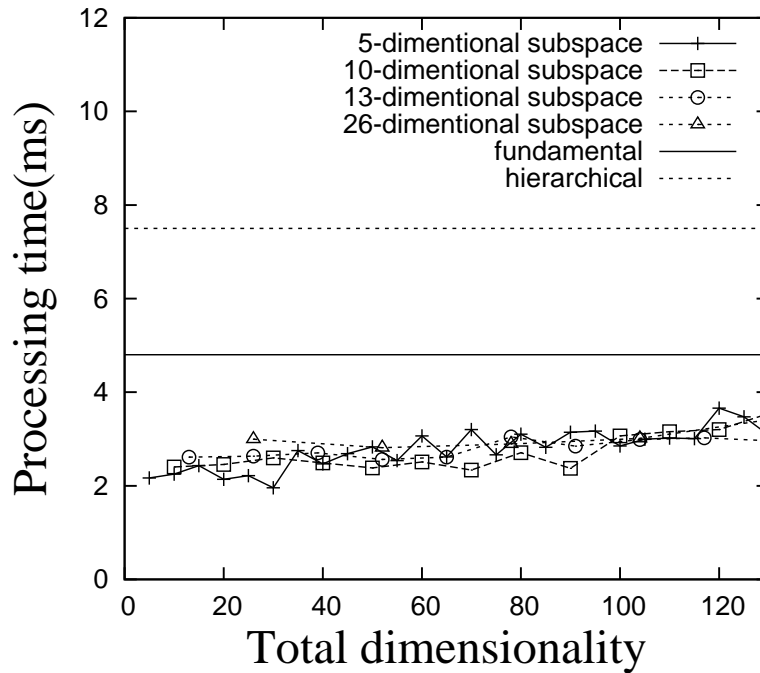


図 4.2: 次元数と並列度を変化させた場合の処理時間の変化

4.3 実動画像上での性能評価

ここでは、実動画像を用いて顔追跡・認識融合系を行い、加重方程式の解法による性能の比較を行った（追跡アルゴリズムは [17, 18] に従う）。

4.3.1 実験条件

実験条件は、3.4 節に従う。以下に簡単にまとめる。

使用したデータベースと拡張固有顔

本実験においても、独自データベース (Data10x10) を使用した。Data10x10 は、10 人の 10 表情の正面顔を 24 種類の照明条件で撮影した画像集合と 10 人の 10 表情の立体顔モデル集合から構成される。Data10x10 の登録画像集合および形状モデルは、あらかじめ位置正規化が施されており、これらを拡張固有顔の作成に用いた。

Data10x10 から 140 次元の正規化固有顔 EF10x10 を作成した。また、EF10x10 の中に正規化固有顔内の個人固有空間を作成した。また階層化加重方程式との比較のため、EF10x10 の中には、サブセット固有空間、サブセット固有空間内の個人固有空間も作成した。ここで、個人固有空間の次元数は 4 とし、サブセット固有空間の次元数は累積寄与率が 99% を越える次元数とした。

テストシーケンス

実験には、3章と同じ動画像を使用した。各人物の1200フレーム(30fps)の動画像から、開始フレームが9フレームずつ異なる100種類のテストシーケンスを作成した。総数1000のテストシーケンスを用いて、顔追跡・認識を行った。

4.3.2 比較した識別方法

静止画像実験を元に、4.2.2節で示した方法の中から、次の3種類の識別方法を比較した。

1. 過剰決定系基本加重方程式 (OD)
2. 5 並列階層化加重方程式 (HE)
3. 5 並列バイアス付き不足決定系 (PUD)

4.3.3 動画像に対する識別結果

追跡・認識融合系では、フレーム間で姿勢追跡精度が一定しないため、位置ずれが不規則に発生する。一方、静止画像を対象とした識別実験結果から、階層化加重方程式だけでなく並列不足決定系においても位置ずれが大きい場合には識別性能が低下することが判っている。このため本実験においても、短時間で正確な識別を行うための判定を行った。

ここで、並列不足決定系においても3章と同様の評価値(正規化相関、加重の最大値、最大加重と第2位の加重の比、正面姿勢からの角度)を用いて各フレームを次の3種類に分類し、動画像に対する識別を行った。

1. 決定フレーム：相関が0.992以上、正面姿勢からの角度が30度以下でかつ、最大加重が決定閾値以上であり、最大加重と第2位の加重との比が閾値以下
2. 有効フレーム：決定フレーム以外で、相関が0.992以上、角度が30度以下でかつ、最大加重が有効閾値以上
3. 棄却フレーム：上記以外のフレーム

なお、本実験においても誤識別率の発生を抑えるために予備実験から上記のパラメータを設定した。決定閾値は基本加重方程式では0.70、並列不足決定系では0.35とし、このときの比率は基本加重方程式では0.50、並列不足決定系では0.50とした。また、有効閾値は基本加重方程式では0.65、並列不足決定系では0.30とした。

ここで、動画像に対する識別は次の2つの規則に従って行う。

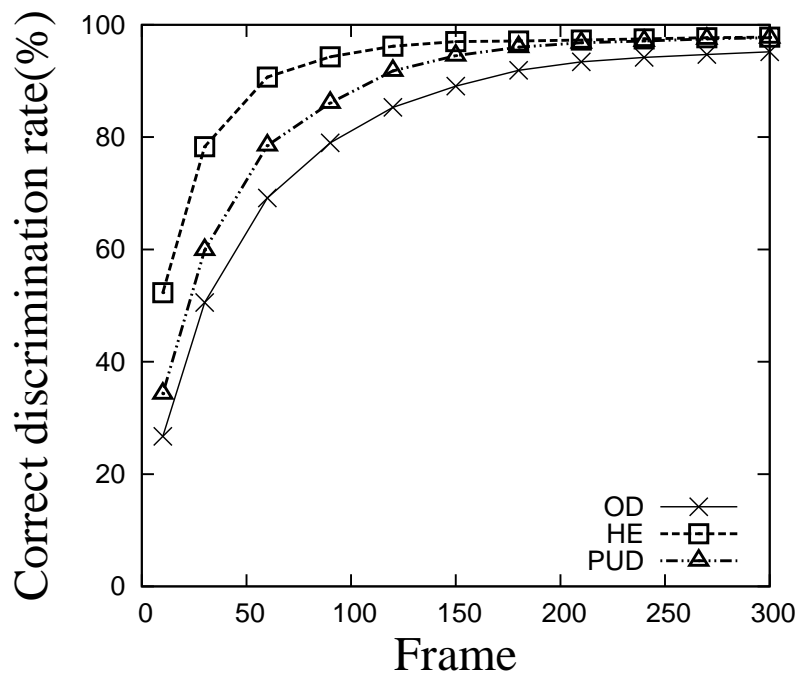


図 4.3: 人物識別時間の変化

- (I) 決定フレームにおいては，最大加重を与える人物を選択する．
- (II) 有効フレームにおいては，直前の 2 有効フレームと同じ人物が最大加重を与える場合についてのみ，その人物を選択する．

実験では，各テストシーケンスについて，比較手法のそれぞれについて 1 回ずつの追跡・認識を行った．追跡開始からのフレーム数を横軸に取ったときの識別成功率の推移をグラフで表すと図 4.3 が得られる．

追跡開始から 300 フレーム後の状況を詳しく見ると，基本加重方程式においては，1000 シーケンス中の 982 シーケンスで正解，残りの 18 シーケンスで棄却であった．これに対し，5 並列階層化加重方程式においては，1000 シーケンス中の 976 シーケンスで正解，残りの 24 シーケンスで棄却，並列不足決定系（バイアスあり）においては，1000 シーケンス中の 981 シーケンスで正解，残りの 19 シーケンスで棄却であった．

また，追跡開始からの推移を詳しく見ると，基本加重方程式 (OD) については，30 フレーム後までに約 50%，90 フレーム後までに約 78.0% の確率で正解が得られている．これに対し，階層化加重方程式については，30 フレーム後までに約 79%，90 フレーム後までに約 93% のシーケンスで正解が得られ，並列不足決定系 (PUD) については，30 フレーム後までに約 60%，90 フレーム後までに約 86% のシーケンスで正解が得られた．これらの結果から，並列不足決定系は，階層化加重方程式と同様に基本加重方程式よりも表情や照明の変動が発生するシーケンスに対して

ロバストかつ高速に識別を完了できることが確認できた。

図 4.4 に、バイアス付き並列不足決定系による追跡・認識結果を例示する。図の左から 1 列目は追跡開始時点の顔検出位置 (初期モデルを描画, テクスチャは平均画像, 形状は平均形状) を示す。一方, 図の左から 2 列目は識別完了時の顔の位置 (顔モデルの更新結果を描画) を示し, 本人に近い顔モデルが追跡・認識融合系により与えられていることが判る。



図 4.4: 並列不足決定系による顔追跡・認識の例

4.3.4 処理時間の比較

QR 分解を用いて, 過剰決定系基本加重方程式, $\sqrt{K} + 1$ 並列の階層化加重方程式 (各階層の次元を $m + 1$ とする) および J 並列不足決定系 (各不足決定系の次元を $m' + 1$ とする) を解く場合を考える。このとき, 各解法の総計算量は, それぞれ $O(mK^2)$, $O(mK^2)(+O(mK^{3/2}))$, $O(Jm'^2K)$ となる。ここで, $Jm' = m$ であることから, J 並列不足決定系の計算量は $O(mm'K)$ となる。なお, 不足決定系では $m' < K$ となることから, 並列不足決定系による解法はその他の解法に比べて計算コストが小さいことが分かる。

以下では, 登録者数の増加に伴う処理時間の変化を各手法ごとに調べた。登録者数は 50 人から 50 人刻みにとり, 最大 289 人登録とした。実験結果を図 4.5 に示す。図 4.5 から, 並列不足決定系は登録人数が増えた場合においても, 処理時間の増加が緩やかに抑えられていることが確認でき, 高速な人物識別に有効であるこ

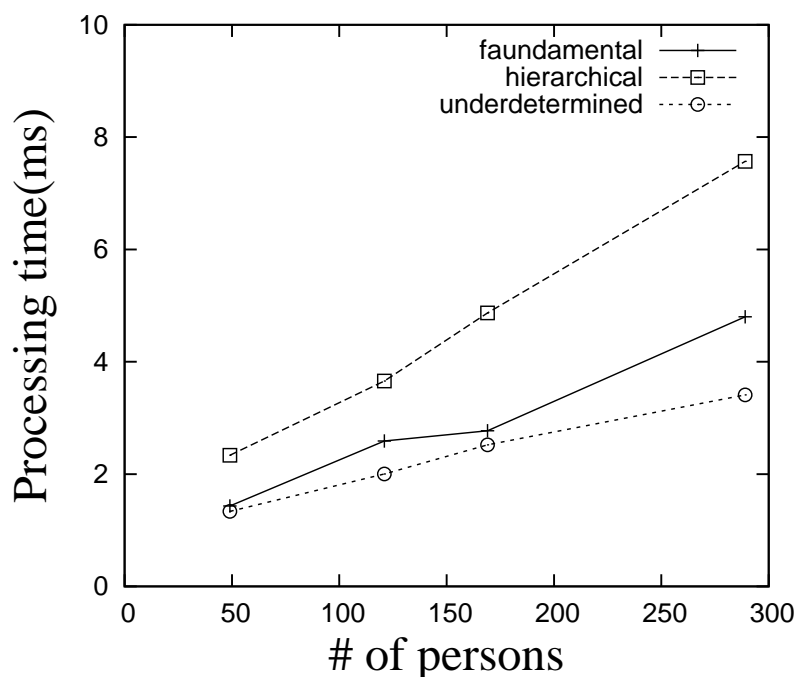


図 4.5: 登録者数が増加した場合の処理時間の変化

とが確認できた。これは、並列不足決定系は、登録者数が増加した場合でも計算量が増大しないように、加重方程式を効率的に構成することができるためであると考えられる。

ここで、1フレームあたりの追跡・認識融合系の処理時間はバイアス付き並列不足決定系を用いた場合に 26.6 ミリ秒であり、他の加重方程式の解法と比べ最も高速な系を実現できた。また、他手法では、Xu[16]らの方法では1フレームあたり約 15 秒(追跡のみ)、Wright[6]らの方法では約数秒(認識のみ)となっており、本手法が他手法に比べて高速であることが判る。

4.4 まとめ

本章では、顔追跡・認識融合系において、登録者数が増大した場合に発生する問題を不足決定系の枠組みで回避するための方法として並列不足決定系を提案し、その有効性を 289 人規模の静止画像データベースおよび 100 登録規模 (10 人 × 10 表情) の実動画データベースで確認した。

この方法では、与えられた正規化固有顔の中で複数の不足決定系を構成し、個々の不足決定系における最適解から平均値を求め、基本加重方程式の解の代わりとして用いる。

静止画像データベースおよび動画データベースを用いた実験の結果から、並列不足決定系は階層化加重方程式と同様に、登録者数の増加にともなう問題を回

避するとともに画像の位置ずれなどのノイズに対しても有効であった。これらの結果は、並列不足決定系が、不足決定系の枠組みで登録者数を大規模化するのに適した手法であることを示している。

また、並列不足決定系と階層化加重方程式の結果から、加重方程式をベースとした加重の算出方法は、過剰決定系と不足決定系のどちらを用いた場合でも改良できることが示せた。一方、並列不足決定系の処理時間は、階層化加重方程式の処理時間と比較して小さかった。現在の処理時間の比較結果から、さらなる大規模登録での実時間系の構成も可能であると考えられ、性能限界を調べることは今後の課題である。

第5章 階層化加重方程式と並列不足決定系による登録者数の大規模化

5.1 階層化加重方程式と並列不足決定系による大規模識別系の構成

Oka-Shakunaga [17, 18] は、顔の追跡・認識融合系において、正規化固有顔から形状への連想写像と、照明変動への適応を同時に実現する方法を提案し、25人系において実時間処理が可能であることを示している。この方法では、全登録者から構成される正規化固有顔内(同次固有空間)に各登録者の照明変動に対応する個人固有空間を作成しておき、顔画像が入力された時点で、正規化固有顔への s -射影および s -射影の各個人固有空間への射影を求める。ここで、入力画像の s -射影を未知サンプル、 s -射影の各個人固有空間への射影を登録サンプルと呼び、未知サンプルを登録サンプルの最適加重平均で表す問題を解く。この問題は、各登録者の加重を変数とし、固有空間の各次元を1つの式で表した線形連立方程式(加重方程式)で表現される。加重方程式を解くことにより、人物識別と形状復元を統一的に論じることができるが、Oka-Shakunaga [17, 18] の系では、登録者数増加への対応が課題であった。本論文では、この問題に対して2つのアプローチを提案し対策を行ってきた。まず3章では、加重方程式を過剰決定系の範囲内で取り扱う階層化加重方程式を提案し、登録者数100の実時間顔追跡・認識融合系を実現した。これに対して、4章では、加重方程式を不足決定系の範囲内で取り扱う並列不足決定系を提案し、登録者数100の場合において、一定の識別性能と高速処理を可能とした実時間顔追跡・認識融合系を実現した。

本章では、階層化加重方程式 [27] および並列不足決定系 [28, 29] をベースとして、登録者数の大規模な識別系(以降では、単に大規模識別系と呼ぶ)を構成する。それぞれの加重方程式を用いて構成された大規模識別系は次の3点にまとめられる。

1. 2階層加重方程式の多段化方式 [30]

この方法では、まず、大規模な登録者集合を2階層の階層化加重方程式(以降では、2階層加重方程式と呼ぶ)で取り扱える規模の登録者集合に分割し、各集合内で2階層加重方程式を構成した後、得られる最適解を用いて候補者を絞り込む。次

に，得られた候補者を集めて再び2階層加重方程式を用いて候補者を絞り込むという操作を多段に繰り返し，候補者数が十分に少なくなった時点で，最後の2階層加重方程式によって大規模識別を行う．

2. 直交サブセット分割を用いた並列不足決定系 [31, 32]

並列不足決定系を用いた方法では，不足決定系で用いる部分空間を固有顔の第1軸から順に一定数ずつ選択して作成しているが，この方法では各軸を1回ずつしか使わないため並列度を高く取れない．そこで本章では，階層化加重方程式において登録者集合を分割するために使用している直交サブセット分割を，正規化固有空間から部分固有空間への分割（以降，直交サブスペース分割と呼ぶ）に利用する．最終的な識別は，直交サブスペース分割により，並列度が高められた並列不足決定系を用いて大規模識別を行う．

なお，本論文では，バイアス付き加重方程式を用いて全ての並列不足決定系を構成した．

3. 並列不足決定系と2階層加重方程式を組み合わせた大規模識別 [31]

不足決定系では，過剰決定系のように正規化固有顔の次元数は識別系の構成の際に制約とならない．このため，不足決定系は原理的に次元数をはるかに上回る登録者数の場合でも識別系を構成できる．また，並列不足決定系は2階層加重方程式に対して，識別性能は若干劣るが処理時間は高速であることが実験から確認されている．これらのことから，並列不足決定系において，2階層加重方程式の多段化方式のような候補選択が精度よくできていれば，候補選択を並列不足決定系に置き換えることで，より高速で精度の高い大規模識別系を構成できると思われる．このため，本実験では，各手法の最終的な識別性能だけでなく，絞り込み性能や処理時間の観点からも評価を行う．

なお，本実験は，加重方程式をベースとした顔認識系を登録者数のキャパシティの観点から評価する基礎実験であるため，予め位置決めされている静止画像を対象とした．また，現時点では，調整された照明環境下で収集された1000人以上の規模のデータベースは存在しない．このため，本論文では，大量のシミュレーション顔画像を用いて大規模識別系を構成した．

5.2 大規模問題におけるシミュレーション実験系の構成

従来の顔追跡・認識融合系 [18]，3章および4章では，並列部分射影 [21] を使用することで顔認識性能を向上させている．本章の議論においても並列部分射影は有効と思われるが，実験を簡潔にすることを目的とするため，本論文では，並列部分射影については考えないで，顔全体を用いた識別問題に議論を限定する．



図 5.1: Data100 の例

5.2.1 合成顔画像データベースの作成

以下に、合成顔画像の作成に用いた原データセットの概要を示す。

Data10x10

顔画像合成のための独自データセットである。100 人の無表情の正面顔を 24 種類の照明条件で撮影した画像集合から構成され、以降では Data100 と呼ぶ。なお、各顔画像は、3.4.1 節と同様の方法で切り出されている。図 5.1 に Data100 の例を示す。この例では 100 人の無表情をある照明条件から撮影した画像を並べたものである。

Data100 中の 100 人の実画像から選んだ 4 人の人物 q_1, q_2, r_1, r_2 から 1 人の (仮想人物の) 画像を合成した。ここで、人物 q_1, q_2 は顔全体 (両眼以外の部分) を合成す

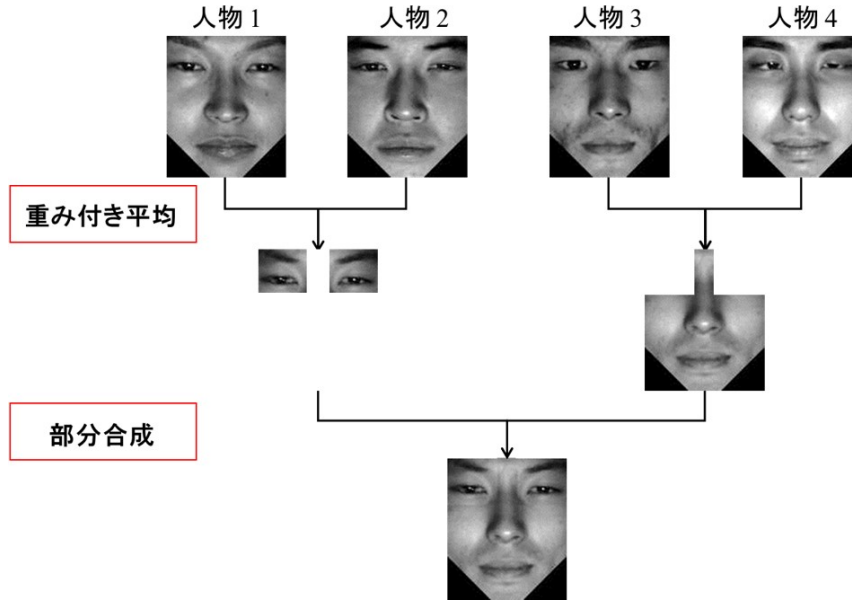


図 5.2: 合成顔画像の作成手順

るのに用い，パラメータ β_1 により合成画像を求めた．一方，人物 r_1, r_2 は両眼部分を合成するのに用い，パラメータ β_2 を用いて合成画像を求めた．

ある照明条件 l で撮影された人物 q_1, q_2, r_1, r_2 の正規化画像を $\mathbf{v}_{q_1}, \mathbf{v}_{q_2}, \mathbf{v}_{r_1}, \mathbf{v}_{r_2}$ とし，各部分領域で生成される合成画像をそれぞれ $\mathbf{v}'_{kl}, \mathbf{v}''_{kl}$ とすると，仮想人物 k の照明条件 l の合成画像 \mathbf{v}_{kl} は次式で求められる．

$$\mathbf{v}_{kl} = (\mathbf{I} - \mathbf{P}_e)\mathbf{v}'_{kl} + \frac{\mathbf{1}^\top \mathbf{P}_e \mathbf{v}'_{kl}}{\mathbf{1}^\top \mathbf{P}_e \mathbf{v}''_{kl}} \mathbf{P}_e \mathbf{v}''_{kl} \quad (5.1)$$

$$\text{where} \quad \mathbf{v}'_{kl} = \beta_1 \mathbf{v}_{q_1 l} + (1 - \beta_1) \mathbf{v}_{q_2 l}, \quad (5.2)$$

$$\text{and} \quad \mathbf{v}''_{kl} = \beta_2 \mathbf{v}_{q_1 l} + (1 - \beta_2) \mathbf{v}_{r_1 l}. \quad (5.3)$$

ここで， \mathbf{P}_e は両眼部分の領域指定行列である． q_1, q_2, r_1, r_2 を異なる人物としてランダムに選択し，その際に， $0 < \beta_1 < 1, 0 < \beta_2 < 1$ の範囲で一様乱数を発生させることで，任意枚数の合成画像を生成できる．図 5.2 に合成画像の作成手順をまとめたものを示す．

5.2.2 正規化固有顔の構成

上記の方法で生成した画像集合を用いて，静止識別実験を行った．この実験では，各人物について 4 枚の照明画像 (図 5.3(a)) を学習データとして正規化固有顔

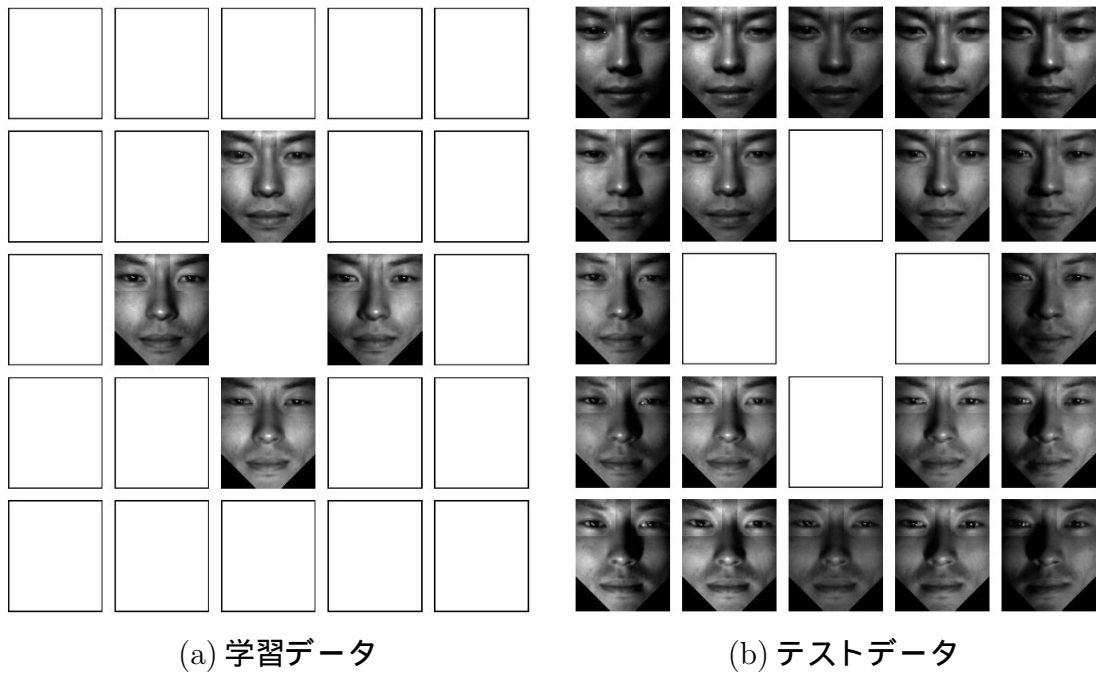


図 5.3: ある人物における学習データとテストデータ

(次元数は登録者数などの実験条件によって異なるため、その都度明記する)を作成し、その中に2次元の個人固有空間を作成した。識別実験では、各人物について、残りの20枚をテスト画像(図5.3(b))とし、登録者全員について識別率を求めた。また、従来の顔追跡・認識融合系では並列部分射影を使用することで顔認識性能を向上させている。しかし、本実験では評価を単純化するために並列部分射影を用いず、顔全体を用いた識別問題に限定した。

5.3 階層化加重方程式による大規模化

5.3.1 2階層加重方程式を用いた事前実験

まず、2階層加重方程式において、シミュレーション画像を用いた基礎性能評価実験を行った。実験では、サブセットサイズとサブセット数が同じ K である場合

について、 K を変化させることによる識別性能 (識別率、および、 K 位累積分類率) の変化を調べた。

ここで、 K が素数である場合、互いに直交するサブセット分割を $K + 1$ 組作成できる。本実験では、並列数が 1, 4, $K + 1$ である場合について識別性能を比較した。また、本識別実験には、 $K = 17$ について作成した 140 次元の正規化固有空間を使用した (この中に識別に用いる人数分の個人固有空間を作成した)。表 5.1 に識別率、表 5.2 に K 位累積分類率の変化を示す。比較のため、それぞれを距離の比較で識別した場合の値を「最近傍法」の欄に標記している。

表 5.1 より、最近傍法に比べ、2 階層加重方程式では、登録者数の増加に対する識別率の低下が少ない。また、並列度を上げることで得られる識別率向上効果は並列度 4 程度でなくなることが判る。一方、表 5.2 を見ると、最近傍法では登録者数が増加するのに伴い K 位累積分類率が低下していくのに対し、2 階層加重方程式では高い K 位累積分類率を維持している。特に、並列度 4 以上の場合には、 K 位累積分類率が 100 % を維持している。ここで、 $K = 17$ においては、登録者数 289 人を 17 人に無損失で絞り込めることを意味している。一方、並列度 1 の場合についても、 K が 13 以下で K 位累積分類率が 100 % である。これは、登録者数 169 人を 13 人に無損失で絞り込めることを意味している。

以上の結果は、登録者数がさほど大きくない場合に、2 階層加重方程式を用いた候補の絞り込みが有効であることを示している。

表 5.1: 階層化による K^2 人登録時の識別率 (%)

| 識別手法 | | 識別率 (%) | | | | |
|------|---------|---------|---------|----------|----------|----------|
| | | $K = 5$ | $K = 7$ | $K = 11$ | $K = 13$ | $K = 17$ |
| 2 階層 | 並列度 1 | 100 | 99.90 | 99.34 | 99.17 | 98.27 |
| | 4 | 100 | 100 | 99.96 | 99.91 | 99.83 |
| | $K + 1$ | 100 | 100 | 99.96 | 99.91 | 99.79 |
| 最近傍法 | | 99.6 | 98.67 | 98.14 | 97.69 | 96.78 |

5.3.2 3 階層以上の加重方程式による大規模化

階層化加重方程式において登録者数が増大した場合、単純に階層化加重方程式における階層を 3 以上にすることが考えられる。この方法は原理的には妥当であり、 K^n の登録者を $n - 1$ 階層のサブセット (それぞれの階層のサブセットは、 K^{n-1} 人、 K^{n-1} 人、...、 K 人で構成される) を用いて構成することにより、 $n - 1$ 階層の加重方程式で求められる各加重の積によって、各個人の加重を求めることができる。

表 5.2: 階層化による K^2 人登録時の K 位累積分類率 (%)

| 識別手法 | | K 位累積分類率 (%) | | | | |
|------|---------|----------------|---------|----------|----------|----------|
| | | $K = 5$ | $K = 7$ | $K = 11$ | $K = 13$ | $K = 17$ |
| 2 階層 | 並列度 1 | 100 | 100 | 100 | 100 | 99.88 |
| | 4 | 100 | 100 | 100 | 100 | 100 |
| | $K + 1$ | 100 | 100 | 100 | 100 | 100 |
| 最近傍法 | | 99.68 | 99.26 | 98.32 | 98.01 | 97.37 |

以下では、3 階層加重方程式の性能評価実験を行った。実験では 3 階層加重方程式において、2 階層目と 3 階層目のサブセット数を同じ K とする場合（すなわち、2 階層目のサブセットサイズは K^2 、3 階層目のサブセットサイズは K 、登録者数は K^3 となる）について、 K を変化させることによる識別性能（識別率および K 位累積分類率）の変化を調べた。また、本識別実験には、各 $K = 17$ について作成した 140 次元の正規化固有空間を使用した（この中に識別に用いる人数分の個人固有空間を作成した）。

K が素数である場合について識別性能を調べた。なお、この実験では直交サブセット分割を用いた並列系は構成していない。表 5.3 に識別率、表 5.4 に K 位累積分類率の変化を示す。

表 5.3 を見ると、 K が 7 以上になると識別率が急速に低下し、最近傍法よりも低い値を示しており、特に、 K が 17 の場合には識別率が 50 % を切る状態が出現している。また、表 5.4 が示すように、 K 位累積分類率の値も低迷したままであり、最近傍法よりも悪い結果である。これは、表 5.2 と大きく異なる結果を示しており、3 階層加重方程式には問題があることが判る。

3 階層加重方程式における識別率低下の主な原因としては、サブセットサイズが大きくなり過ぎることが考えられる。本論文で用いた固有空間は 140 次元であるが、3 階層目のサブセットサイズは K^2 であるため、 $K = 17$ の場合には表現能力の限界に達していると考えられる。

パターン認識の古典的アプローチとしてクラスタリングがあり、妥当なクラスタリングが可能な場合には、サブセットサイズをこれよりも大きく取れる可能性は残されているが、現状では妥当なクラスタリング法がないため、本論文では 3 階層以上の階層化加重方程式を用いた大規模化は行わないこととする。

また、2 階層加重方程式に関しては、前節で述べたように直交サブセット分割による並列化が有効と考えられる。一方、3 階層以上の階層化加重方程式における並列化は、2 階層の場合と比べ複雑であり、現状では妥当な方法が見つからない。

表 5.3: 階層化による K^3 人登録時の識別率 (%)

| 識別手法 | 識別率 (%) | | | | |
|-----------|---------|---------|----------|----------|----------|
| | $K = 5$ | $K = 7$ | $K = 11$ | $K = 13$ | $K = 17$ |
| 3 階層加重方程式 | 99.40 | 95.65 | 81.47 | 69.83 | 47.19 |
| 最近傍法 | 98.20 | 96.59 | 93.59 | 92.53 | 90.55 |

表 5.4: 階層化による K^3 人登録時の K 位累積分類率 (%)

| 識別手法 | K 位累積分類率 (%) | | | | |
|-----------|----------------|---------|----------|----------|----------|
| | $K = 5$ | $K = 7$ | $K = 11$ | $K = 13$ | $K = 17$ |
| 3 階層加重方程式 | 99.92 | 99.18 | 92.54 | 85.77 | 69.37 |
| 最近傍法 | 99.68 | 99.26 | 98.32 | 98.01 | 97.37 |

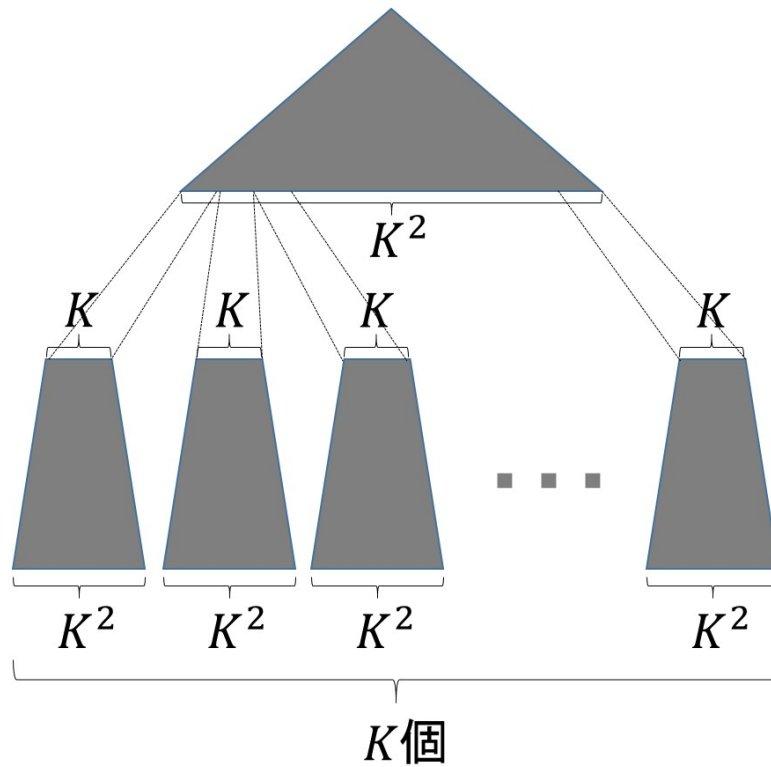


図 5.4: 2 階層加重方程式の 2 段構成

5.3.3 2階層加重方程式の2段構成による大規模化

前2節の検討を受けて、階層化加重方程式を用いた大規模識別系の構成には、基本モジュールとして2階層加重方程式を用い、かつ、サブセットサイズを一定に保ちつつ、登録者数の増加に対応する方法を検討する必要がある。ここで、累積分類率を100%近くに保ちつつ、登録者数の絞り込みを行えば、2階層加重方程式を多段に繋ぐ構成で登録者数を拡大できると考えられる。一方、表5.2は、 K^2 人で構成した2階層加重方程式において、 K 位累積分類率を100%近くに保つことができることを示しており、2階層加重方程式を2段に繋ぐことで K^3 人に対応できる可能性がある。この場合、1段目では K 個の K^2 人組において K 人への絞り込みを行い、2段目では絞り込まれた合計 K^2 人に対して2階層加重方程式が解かれることになる。

図5.4に2階層加重方程式の2段構成のイメージ図を示す。この図では、下段の K 個の台形が1段目の2階層加重方程式(K^2 人登録)に、上段の三角形が2段目の2階層加重方程式(K^2 人登録)に相当する。1段目では、各2階層加重方程式から上位 K 人が抽出され、2段目において、絞り込まれた K^2 人で構成される2階層加重方程式を解くことにより最終的な人物識別を行う。

5.4 並列不足決定系における大規模識別系の構成

5.4.1 直交サブスペース分割による並列化

4章では、不足決定系で用いる部分空間を正規化固有顔の第1軸から順に一定数ずつ選択している。しかし、この方法では各軸を1回ずつしか使わないため、最大並列度は部分空間の数以上にはできない。そこで本論文では、階層化加重方程式において並列度を高めるために用いている直交サブセット分割を、正規化固有顔から部分空間への分割に適用する。

直交サブスペース分割の例を図5.5に示す。この例では、25次元正規化固有顔の固有軸を5行5列に配置し、各行を部分空間とした分割と各列を部分空間とした分割である。次元数は素数の2乗となるように設定されているため、この例では、列および行ごとの分割とラテン方阵[23]をベースとした分割が4個、すなわち、直交条件を満たす分割は合計6個作成できる。

5.4.2 直交サブスペース分割を用いた人物識別

直交サブスペース分割を用いることで、並列度を(直交サブスペース分割数) \times (部分空間の次元数)まで高めることができる。

互いに直交するサブスペース分割を E 個、各分割において不足決定系の加重方程式を J 作成した場合を考える。サブスペース分割 e の第 j 不足決定系において



Partition 1



Partition 2



Partition 3



Partition 4



Partition 5



Partition 6

図 5.5: 25 次元正規化固有顔を用いた場合の直交サブスペース分割の例

求められる最適加重を $w_{(e,j)}$ で表す．このとき， EJ 個の最適解を平均することにより最終的な解が得られる．

$$\mathbf{w} = \frac{1}{EJ} \sum_{e=1}^E \sum_{j=1}^J \mathbf{w}_{(e,j)} \quad (5.4)$$

なお，本実験では使用していないが，並列部分射影との組み合わせも有効であると考えられる．ここで，第 $q(= 1, \dots, Q)$ 領域の画像を用いて構成した，第 e サブスペース分割の第 j 不足決定系において求められる最適加重を $w_{(e,j,q)}$ で表す．このとき， EJQ 個の解を平均することによって最終的な解が得られる．

$$\mathbf{w} = \frac{1}{EJQ} \sum_{e=1}^E \sum_{j=1}^J \sum_{q=1}^Q \mathbf{w}_{(e,j,q)} \quad (5.5)$$

ここで，平均操作によって得られる両者の加重ベクトルは元の加重方程式を満たす保証はないが，有効性はこれまでの実験により確かめられている．

なお本論文の実装では，次元数が m の正規化固有空間から \sqrt{m} 次元ずつ用いて不足決定系を構成していく．ここで， \sqrt{m} が素数である場合，上述のように互いに直交するサブスペース分割を最大で $\sqrt{m} + 1$ 個構成できる．従って，この場合，各サブスペース分割で構成される \sqrt{m} 並列不足決定系を $\sqrt{m} + 1$ 個並列させて解くことによって，最大で $(\sqrt{m} + 1)\sqrt{m}$ 並列不足決定系を構成できる．

5.4.3 並列不足決定系を用いた基礎性能比較

5.3.1 節では，2 階層加重方程式において，サブセットサイズとサブセット数が同じ K である場合について， K を変化させることによる登録者数 K^2 の識別性能（識別率および K 位累積分類率）の評価を行った．本節では，同様の実験を並列不足決定系で行い，識別性能の比較を行う．ここで，2 階層加重方程式の識別では並列度（直交サブセット分割数）が $1, 4, K + 1$ である場合についての識別性能を示す．対して，並列不足決定系では並列度を変化させるため，各部分空間を K 次元とし，直交サブセット分割数を $1, 4, K + 1$ と変化させた場合の識別性能を求めた．

また，正規化固有空間の次元数は 2 階層加重方程式の識別では全て 140 次元（5.3.1 節と同様），並列不足決定系では直交サブセット分割を行うため， K^2 次元とした．表 1 に識別率，表 2 に K 位累積分類率の変化を示す．比較のため，それぞれを距離の比較で識別した最近傍法の性能も載せる．

表 5.5 より，並列不足決定系は K が 7 以上の場合に常に最近傍法に比べて，識別率が高いことが分かる． $K = 5$ の場合は，固有空間の次元数が 25 次元（累積寄与率 92 %）と少なかったために，識別率が低くなったと考えられる．また，全ての K に対して，並列度の増加による識別率向上効果は見られなかった．2 階層加重方程式との比較を行うと，並列度 1 の部分では並列不足決定系の識別率が勝つ

表 5.5: 並列不足決定系による K^2 人登録時の識別率 (%)

| | | K=5 | K=7 | K=11 | K=13 |
|---------|--------|------|-------|-------|-------|
| 最近傍法 | | 99.6 | 98.67 | 98.14 | 97.69 |
| 2 階層 | 並列度 1 | 100 | 99.90 | 99.34 | 99.17 |
| | 4 | 100 | 100 | 99.96 | 99.91 |
| | K+1 | 100 | 100 | 99.96 | 99.91 |
| 並列不足決定系 | 並列度 K | 98.6 | 99.49 | 99.71 | 99.79 |
| | 4K | 99.4 | 100 | 99.67 | 99.76 |
| | (K+1)K | 99.4 | 100 | 99.67 | 99.79 |

表 5.6: 並列不足決定系による K^2 人登録時の K 位累積分類率 (%)

| | | K=5 | K=7 | K=11 | K=13 |
|---------|--------|-------|-------|-------|-------|
| 最近傍法 | | 99.68 | 99.26 | 98.32 | 98.01 |
| 2 階層 | 並列度 1 | 100 | 100 | 100 | 100 |
| | 4 | 100 | 100 | 100 | 100 |
| | K+1 | 100 | 100 | 100 | 100 |
| 並列不足決定系 | 並列度 K | 100 | 100 | 100 | 100 |
| | 4K | 100 | 100 | 100 | 100 |
| | (K+1)K | 100 | 100 | 100 | 100 |

ている。しかし、並列度を上げた場合には全体的に 2 階層加重方程式の識別率が高いことが分かる。

一方、表 5.6 を見ると、最近傍法では登録者数が増加するのに伴い K 位累積分類率が低下している。それに対し、2 階層加重方程式、並列不足決定系はともに K 位累積分類率 100 % を維持している。これは $K = 13$ の場合、登録者数 169 を 13 人に無損失で絞り込めることを意味している。以上の結果から、並列不足決定系を用いた候補の絞り込みも大規模識別系の構成に有効であると考えられる。

5.5 シミュレーション画像を用いた大規模識別実験

5.5.1 実験の概要および実験条件

ここでは、前述のシミュレーション顔画像を用いた大規模識別実験を行い、2 階層加重方程式の多段構成と直交サブスペース分割を用いた並列不足決定系の性能評価および比較を行った。ここで、登録者数が大規模な場合、2 階層加重方程式では候補者を一定数まで絞り込む操作も重要となってくる。そのため本実験では、識別率だけでなく任意の順位までの累積分類率も求めることで、それぞれの手法

が登録者数が大規模な場合に候補者の絞り込みにどれだけ有効かの評価も行った。また、識別や絞り込みを可能な限り高速に行うため、各手法の計算量の比較を行った。さらに、登録者の増加に対する絞り込み性能の限界を確認するために、登録者をより増加させた実験も行った。

5.5.2 実験条件と比較した手法

本実験に用いる学習画像やテスト画像および正規化固有顔の構成などの基本的な実験条件は5.2.2節にならう。登録者数は、前述の実験よりもさらに増やして K^3 人とし識別実験を行った。

比較した手法を以下に示す。

1. 2階層加重方程式の2段構成： K^3 人を K^2 人の K 個のサブセットに分割し、1段目の2階層加重方程式を構成する。各2階層加重方程式の結果から上位 K 人を抽出し、2段目の2階層加重方程式(K^2 人登録)を構成する。なお、正規化固有顔の次元数は140とした。最終的な人物識別は2段目の加重を用いて行った(式(2.24))。また、直交サブセット分割の並列数を変化させて識別性能の変化も調べた。
2. 直交サブスペース分割を用いた並列不足決定系： K^2 次元の正規化固有顔を K 個の部分空間(各 K 次元)に分割し、各部分空間でバイアス付きの不足決定系加重方程式(サイズ: $(K+1) \times K^3$)を構成した。ここで、バイアスの計算には K^2 次元の正規化固有顔内で求められる s と s_k を使用した。人物識別(式(2.24))には、各バイアス付き加重方程式から得られる最適加重を平均したものを用いた。また、直交サブスペース分割の並列数を変化させながら識別率の変化も調べた。
3. 最近傍法： K^3 人について、 s と s_k の距離を求め、最も小さいものが属する人物を選ぶ。

5.5.3 K^3 人登録における性能比較

1. 識別率の比較

ここでは、比較した手法の識別率と K 位累積分類率、 K^2 位累積分類率を求め比較した。表5.7に識別率、表5.8に K 位累積分類率、表5.9に K^2 位累積分類率を示す。

表5.7より、識別率は K^2 人登録時(表5.5)とほぼ同じ傾向が見られた。しかし、表5.5では並列度を増やすことによる効果が見られなかったのに対して、 K^3 人登録時には性能の向上が確認できた。また、表5.8より K 位累積分類

表 5.7: 大規模識別系 (K^3 人登録時) における識別率 (%)

| | | K=5 | K=7 | K=11 | K=13 |
|----------|--------|-------|-------|-------|-------|
| 最近傍法 | | 98.40 | 96.57 | 93.66 | 93.29 |
| 2 階層 2 段 | 並列度 1 | 98.96 | 98.35 | 96.70 | 94.62 |
| | 4 | 99.88 | 99.71 | 99.47 | 98.98 |
| | K+1 | 99.96 | 99.85 | 99.78 | 99.62 |
| 並列不足決定系 | 並列度 K | 98.12 | 98.40 | 98.16 | 98.48 |
| | 4K | 98.44 | 98.38 | 98.27 | 98.52 |
| | (K+1)K | 98.60 | 98.44 | 98.29 | 98.53 |

表 5.8: 大規模識別系 (K^3 人登録時) における K 位累積分類率 (%)

| | | K=5 | K=7 | K=11 | K=13 |
|----------|--------|-------|-------|-------|-------|
| 最近傍法 | | 99.72 | 99.29 | 98.35 | 98.27 |
| 2 階層 2 段 | 並列度 1 | 99.96 | 99.90 | 98.32 | 99.27 |
| | 4 | 100 | 99.98 | 99.95 | 99.90 |
| | K+1 | 100 | 100 | 99.98 | 99.96 |
| 並列不足決定系 | 並列度 K | 99.96 | 99.91 | 99.91 | 99.94 |
| | 4K | 99.96 | 99.93 | 99.93 | 99.97 |
| | (K+1)K | 99.96 | 99.93 | 99.93 | 99.97 |

率を見ると, $K = 11$ までは 2 階層加重方程式の 2 段構成の方が性能が高い. 一方, $K = 13$ になると並列不足決定系の性能が上回っていることが確認できた. しかし, とともに K の値によっては K 位累積分類率が 100 % に達しないことがあることが分かった.

次に表 5.9 より, 2 階層加重方程式の 2 段構成では並列度を $K + 1$ とした場合に, すべての K で K^2 位累積分類率が 100 % となった. 一方, 並列不足決定系では $K = 7$ 以上の場合において, 並列度 K の時点ですでに 100 % となっていることを確認できた. この結果は重要であり, $K = 13$ とした場合, 登録者数 2197 人を 169 人に無損失で絞り込めることを意味している. このことから, 2 階層加重方程式の 2 段構成と並列不足決定系は共に登録者数が大規模な場合において, 候補者の絞り込みに有効であることが確認できた.

2. 並列度の変化による並列不足決定系の性能比較

表 5.9 より, 並列不足決定系を用いて登録者数 K^3 を K^2 人に絞り込むのに, $K = 13$ の場合では並列度が K , つまり 13 次元の部分固有空間を 13 並列させることで達成していた. ここでは, さらに並列度を少なく変化させることで, 累積分類率がどのように変化するかを調べた. ただし, 直交サブセット分割の並びによって選択する次元が変化するので, 複数のサブセット分割ご

表 5.9: 大規模識別系 (K^3 人登録時) における K^2 位累積分類率 (%)

| | | K=5 | K=7 | K=11 | K=13 |
|----------|--------|-------|-------|--------|--------|
| 最近傍法 | | 99.92 | 99.88 | 99.76 | 99.72 |
| 2 階層 2 段 | 並列度 1 | 100 | 100 | 99.99 | 99.98 |
| | 4 | 100 | 100 | 99.996 | 99.997 |
| | K+1 | 100 | 100 | 100 | 100 |
| 並列不足決定系 | 並列度 K | 99.58 | 100 | 100 | 100 |
| | 4K | 99.92 | 100 | 100 | 100 |
| | (K+1)K | 100 | 100 | 100 | 100 |

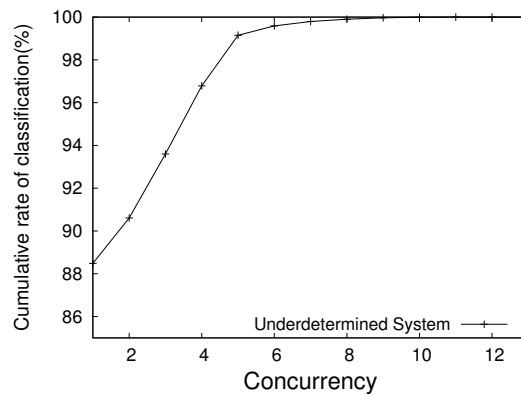


図 5.6: 並列度の変化に対する 2197 人登録時の 169 位累積分類率

とに評価を行い，平均をとった．図 5.6 に並列度の変化に対する 2197 人登録時の 169 位累積分類率の変化を示す．

図 5.6 より，並列度 1 では低かった累積分類率が並列度を増やす毎に上昇しており，並列度 11 の段階ですでに 100 % に達していることが分かる．これにより並列不足決定系は，個々の部分空間毎のみでは低性能な識別系を，並列を繰り返すことで補いながら高性能化していることが確認できた．ここで，並列度を最低限に抑えることで処理速度を上げることができる．一方，大規模な問題では並列度を増やすことで識別性能を維持できると考えられる．

5.5.4 計算量の比較

本節では，2 階層加重方程式の 2 段構成と並列不足決定系の計算量を比較する．

1. 2 階層加重方程式の 2 段構成

まず，過剰決定系加重方程式の計算量を考える．登録者を K 人，同次固有顔の次元数を $m + 1$ とした場合，1 つの加重方程式 (線形連立方程式) を解くのに必要な計算量は，QR 分解を用いた場合に $O(mK^2)$ となる．

表 5.10: 2 階層加重方程式と並列不足決定系の計算量の比較

| | 加重方程式の計算 | 固有空間作成 | 総計算量 |
|----------|----------|----------|----------|
| 2 階層 2 段 | $O(K^7)$ | $O(K^8)$ | $O(K^8)$ |
| 並列不足決定系 | $O(K^7)$ | 0 | $O(K^7)$ |

登録者 K^3 人の 2 階層加重方程式の 2 段構成を $K+1$ 並列で解いた場合の計算量を考える。1 つの 2 階層加重方程式の計算量は $(K+1)O(mK^2) = O(mK^3)$ であり、2 段構成の解法においてはこの計算を $(K+1)^2$ 回 (2 階層加重方程式の計算回数 $(K+1) \times$ 並列度 $(K+1)$) 行うため、計算量は $O(mK^5)$ となる。ここで、本実験では $m = K^2$ としているため、実際の計算量は $O(K^7)$ となる。

ここで、2 階層加重方程式の 2 段構成においては、1 段目の結果からサブセットを作成するため、入力ごとに固有空間 (サブセット固有空間と個人固有空間) を動的に作る必要があり、解法全体の計算量にはこれらの計算量も含めなければならない。動的に生成される固有空間は、 $(K+1) \times K$ (並列数 \times サブセットの数) 個のサブセット固有空間と、 $(K+1) \times K^2$ (並列数 \times 2 階層加重方程式の登録者数) 個の個人固有空間である。 $n \times n$ 行列の固有値分解の計算量は一般的に $O(n^3)$ である。本論文の実装では、1 つのサブセット固有空間の計算量は、 $m (= K^2) \times K$ 行列から $m \times m$ 行列を準備してから作成しているため、 $K(K+1)O(m^3) = K(K+1)O(K^6) = O(K^8)$ となる。一方で、個人固有空間の計算量は、個人固有空間の作成に用いる行列のサイズが次元圧縮により K と m のサイズがほぼ同等となるとすると、 $K^2(K+1)O(K^3) = O(K^6)$ 程度ですむ。よって、すべての固有空間作成のための計算量は $O(K^8)$ となり、解法全体の計算量も $O(K^8)$ となる。

2. 並列不足決定系

登録者数を K^3 、各不足決定系の次元数を m' 、並列度を J とする。この場合、直交サブスペースを用いた並列不足決定系の計算量は $O(Jm'^2K^3)$ となる。ここで、 $m' = K, J = K(K+1)$ となるので、計算量は $O(K^7)$ となる。

また、並列不足決定系の場合、最初に作成した固有顔を分割しながら計算を行うため、入力ごとに固有空間を作成する必要はなく、解法全体の計算量も $O(K^7)$ となる。

2 手法の計算量をまとめると表 5.10 が得られた。直交サブスペースを用いた並列不足決定系の加重算出のための計算量は、2 階層化加重方程式の 2 段構成と変わらない。しかし、並列不足決定系では、認識ごとの動的な固有空間の作成が不要であるため、総計算量は 2 階層加重方程式の 2 段構成と比較して小さくなる。ゆえに、大規模識別系における候補者の絞り込みには並列不足決定系の方が適しているといえる。

表 5.11: K^3 人登録時の処理時間 (msec)

| | | K=5 | K=7 | K=11 | K=13 |
|----------|--------|------|-------|--------|--------|
| 2 階層 2 段 | 並列度 1 | 17.9 | 38.1 | 131.5 | 212.1 |
| | 4 | 45.3 | 122.2 | 485.4 | 800.5 |
| | K+1 | 63.9 | 234.5 | 1429.5 | 2883.5 |
| 並列不足決定系 | 並列度 K | 5.1 | 13.1 | 56.5 | 112.7 |
| | 4K | 7.0 | 21.1 | 111.4 | 251.6 |
| | (K+1)K | 8.2 | 29.9 | 244.5 | 688.5 |

表 5.12: 並列不足決定系による 169 位累積分類率 (%) と正解者の最低順位

| 登録者数 | 2500 人 | 5000 人 | 7500 人 | 10000 人 |
|----------------|--------|--------|--------|---------|
| 169 位累積分類率 (%) | 100 | 99.988 | 99.983 | 99.972 |
| 正解者の最低順位 (位) | 132 | 423 | 668 | 1231 |

3. 処理時間の比較

K^3 人登録時の各手法の処理時間を表 5.11 にまとめる。これらの結果から、並列不足決定系は 2 階層加重方程式の 2 段構成と比較して高速であることが分かる。これは前述のとおり、並列不足決定系では、認識ごとの動的な固有空間の作成が不要であるためである。さらに絞り込み性能の結果からも、並列不足決定系は、大規模識別系における候補者の絞り込みに適していると言える。

5.5.5 さらに大規模な問題での絞り込み性能

ここでは登録者をさらに大規模化した場合の絞り込み性能を評価する。大規模識別系を構成する際、並列不足決定系は、次元数 > 登録者数の制約を満たす必要が無い場合 2 階層加重方程式の 2 段構成よりも柔軟に系を構成できる。

本実験では、並列不足決定系において、登録者数を 2500 人ずつ 10000 人まで増やし、169 位累積分類率を求めていく。また、累積分類率が 100 % でない場合でも、最低でも何人にまで絞り込んでいるかを確認するため、各テストの正解人物の順位も求めるようにする。

登録者数を大きく増やした場合、固有顔を作成するのに長い時間を必要とする。そのため、本実験では 169 次元固有顔を原画像集合から作成し、その固有顔内に個人固有空間を作成するようにする。また、本実験では、処理時間をできるだけ減らすため、使用する並列度は 13 とした。表 7 に各登録者数における 169 位累積分類率と正解人物の最低順位を示す。表より、2500 人の累積分類率が 100 % となることから、データ I を用いた場合でも、高い絞り込み性能を維持できていること

表 5.13: 大規模識別における単独系と融合系の計算量の比較

| | 加重方程式の計算 | 固有空間作成 | 総計算量 |
|--------|----------|----------|----------|
| 手法 (1) | $O(K^7)$ | 0 | $O(K^7)$ |
| 手法 (2) | $O(K^7)$ | $O(K^8)$ | $O(K^8)$ |
| 手法 (3) | $O(K^7)$ | 0 | $O(K^7)$ |
| 手法 (4) | $O(K^7)$ | $O(K^8)$ | $O(K^8)$ |

が確認できた．次に 5000 人から 10000 人の結果を見ると，人数が増えるごとに累積分類率は下がっているが，いずれも 99.97 % を越える性能は保てていることが確認できた．さらに，並列度を 13×14 まで増やしての 10000 人の実験を行ったが，累積分類率は向上しないことを確認した．

5.5.6 2 階層加重方程式と並列不足決定系の融合

前節までの実験結果から，並列不足決定系よりも 2 階層加重方程式の 2 段構成の識別性能の方が高いことが確認できた．一方，候補者の絞り込み性能については 2 つの方法はほぼ同等であり，並列不足決定系の方が計算量は少ないことが確認できた．これらの結果を総合すると，絞り込みには高速な並列不足決定系を用い，最終的な識別には 2 階層加重方程式を用いることで，高速かつ高性能な識別を行うことができると考えられる．

この方法の有効性を確かめるため，以下では，登録者を 2197 人 ($K = 13$ 時の K^3) とした場合について，次の 4 つの方法でそれぞれ識別実験を行い，計算量と識別性能の比較を行う．

1. 候補者 2197 人からの絞り込みを行わず，並列不足決定系によって識別を行う．
2. 候補者 2197 人からの絞り込みを行わず，2 階層加重方程式の 2 段構成で識別を行う．
3. 並列不足決定系を用いて候補者を 169 人に絞り込んだ後に，再び並列不足決定系を用いて識別を行う．
4. 並列不足決定系を用いて候補者を 169 人に絞り込んだ後に，2 階層加重方程式 (1 段構成) を用いて識別を行う．

まず表 5.13 にそれぞれの手法での計算量を示す．次に，表 5.14 にそれぞれの手法よる識別率と 13 位累積分類率を示す．ただし，表 5.14 中の並列度は最終的な識別を行うときのものであり，絞り込みを行うときの並列不足決定系の並列度は 13 とした．

表 5.14: 融合系での 2197 人登録時の識別率 (%) と 13 位累積分類率 (%) と処理時間 (msec)

| 並列度 | 手法 | 識別率 | 13 位累積分類率 | 処理時間 |
|--------------------------------|-----|-------|-----------|--------|
| 2 階層 2 段 1 並列不足決定系 13 | (1) | 98.48 | 99.94 | 112.7 |
| | (2) | 94.62 | 99.27 | 212.1 |
| | (3) | 98.08 | 99.94 | 115.4 |
| | (4) | 97.20 | 99.50 | 128.2 |
| 2 階層 2 段 4 並列不足決定系 4 × 13 | (1) | 98.52 | 99.97 | 251.6 |
| | (2) | 98.98 | 99.90 | 800.5 |
| | (3) | 98.98 | 99.97 | 294.4 |
| | (4) | 99.30 | 99.92 | 436.3 |
| 2 階層 2 段 14 並列不足決定系 14 × 13 | (1) | 98.53 | 99.97 | 688.5 |
| | (2) | 99.62 | 99.96 | 2883.5 |
| | (3) | 99.20 | 99.98 | 774.1 |
| | (4) | 99.70 | 99.99 | 1736.2 |

表 5.14 より，共に並列不足決定系しか用いない手法 (1) と (3) を見ると，並列度が少ない場合には手法 (1) の方が，並列度が高くなると手法 (3) の方が性能が高くなることが確認できた．一方，表 5.13 より，処理時間は手法 (1) の方が高速ではあるが，並列不足決定系だけを用いて精度良く識別を行うには，絞り込みを行った後で並列度を増加させる方法が有効であることが分かった．

次に，2 階層加重方程式を用いる手法 (2) と (4) を見ると，全ての並列度で手法 (4) の方が高い識別率を示している．表 5.13 から，計算量は両手法で差がない．しかし，解かなければならない加重方程式の数が階層化加重方程式では $(K + 1)^2$ ，並列不足決定系では $K(K + 1)$ であるため，実際の処理時間は手法 (4) の方が少なかった

手法 (3) と (4) を比較すると，識別性能において手法 (4) の方が優れており，最終的に並列不足決定系と 2 階層加重方程式を融合した手法 (4) が比較した 4 つの手法の中で最も良い手法であることが確認できた．

5.6 まとめ

本章では，階層化加重方程式と並列不足決定系を用いてさらなる大規模識別系を構成し，実験により性能を評価した．まず，階層化加重方程式においては，2 階層加重方程式を多段化することにより登録者数の大規模化を図った．最大 2197 人登録のデータベースを用いた実験では，高い識別性能を得られた．並列不足決定系では，階層化加重方程式で用いられている直交サブセット分割の考え方を正規化固有顔の分割に適用した直交サブ空間分割を用いた大規模識別系を提案し

た．実験結果から識別性能は少し劣るものの高速な候補選択 (絞り込み) が可能な方法であることが分かった．

最後に，両者を組み合わせた大規模識別系を構成し実験を行ったところ，並列不足決定系による候補選択ののちに 2 階層加重方程式を用いて最終的な識別を行う系が最も高い識別率を示した．絞り込みの基礎性能評価から数万人登録を目指した拡張も可能であると思われ，さらなる大規模識別系の構成が今後の課題として挙げられる．

第6章 加重方程式を用いた顔モデリングによる登録処理の簡略化

6.1 加重方程式を用いた登録処理の簡略化

顔画像処理は、セキュリティ応用の観点から近年急速に重要性を増している。また、インターネット上に氾濫する画像を対象とした顔画像処理は、今後も応用分野の広がりが期待される分野である。一方、このようなビッグデータを対象とする場合に要求されるのは、少数枚（できれば1枚）の写真しか利用できない場合にも、照明条件や姿勢の変動にある程度耐えられるモデルを作成する方法を確立することであると考えられる。

ここで、Blanz-Vetter[10]は、1枚の画像から形状と照明条件を同時に推定する方法を提案している。しかし、この方法では形状や反射特性をパラメータ化して取り扱っており、モデルに基づく最適化を行っているため、人物固有の特徴を精確に取り扱うには限界があると思われる。モデルでは表現しきれない精確さを目指すためには、画像データに基づく手法の確立が待たれる。そこで、本論文では、2つの顔認識手法をベースとして、少数枚画像から顔モデルを生成する方法を検討した。

従来の実時間顔追跡・認識融合系 [17, 18] では、照明変動や形状変化に対応するために、あらかじめ登録者ごとに照明変動を含む学習画像や形状データを特定の環境下で撮影しておく必要があるため、登録者数の拡張が容易でないという問題があった。一方、2次元画像から作成される固有顔をベースとして、1枚の画像から異なる照明条件の画像を線形変換により近似的に作成する方法 [33, 7] を提案している。しかし、この方法では画像の照明条件を推定する必要があり、推定が正確でない場合には、作成される画像の質が劣化するという問題があった。また、この方法は、学習画像集合から得られる統計的な変換(モデル)を用いるため、人物固有の特徴を取り扱えない点では [10] と同様の限界を持っている。このため、本章では、加重方程式 [18] をベースとした照明条件の推定や線形変換を必要とせず、人物固有の特徴を復元できる新しい画像作成方法 [34] を提案する。また、実時間顔追跡・認識融合系に適用した結果についても簡単に報告する。

6.2 拡張基準固有顔

6.2.1 加重方程式による顔画像処理

加重方程式 [17, 18] は，入力画像の射影係数を基準画像の射影係数の集合の最適な加重平均で表現する線形連立方程式である．顔追跡・認識融合系においては，あらかじめ各個人ごとの照明変動に対応する部分空間を正規化固有空間内に作成しておき，照明適応と呼ぶ操作により，各人物 1 枚登録の場合と同じサイズの加重方程式を解く．これに対して本論文では，仮想照明画像作成時に各照明条件ごとに登録された射影係数の集合を使用するため部分空間化を行わず，多重登録の場合の加重方程式を構成する．

6.2.2 基準固有顔

k 番目の登録人物を l 番目の照明条件で撮影した画像に対応する正規化画像ベクトルを v_{kl} で表す．以下では， K は登録者数， L は照明条件数を示すものとする．

正規化画像ベクトル集合 $\{v_{kl}\}$ の平均を \bar{v} ，上位 m 個の主成分を Φ_m で表す．このとき，中心を \bar{v} とし，固有軸を Φ_m とする m 次元の正規化固有顔 [7] が作成できる．

入力画像 V の正規化固有顔への s -射影は同次固有顔 $\tilde{\Phi}_m = [\Phi_m \ \bar{v}]$ を用いて次式で計算できる [21]．

$$\tilde{s} = [\alpha s^\top \ \alpha]^\top = \tilde{\Phi}_m^+ V \quad (6.1)$$

ここで， $\tilde{\Phi}_m^+$ は $\tilde{\Phi}_m$ の擬似逆行列を表す．

このとき，同次射影係数 \tilde{s} の正規化表現は， $\hat{s} = [s^\top \ 1]^\top$ で与えられる．射影係数 s が得られると，これに対応する正規化画像 v^s は次式で与えられる (v^s は普通射影と呼ばれる)．

$$v^s = \tilde{\Phi}_m \hat{s} = \Phi_m s + \bar{v} \quad (6.2)$$

正規化固有顔において，顔画像は複数人の学習画像集合から作られる固有空間の少数の固有ベクトルの線形結合で効率良く表現される．すなわち，正規化固有顔は顔画像の標準的な情報を表現する空間であると見なせる．この意味から，本論文では正規化固有顔を基準固有顔と呼ぶ．また，正規化固有顔の作成に用いる学習画像集合を基準画像集合と呼び，各基準画像に対応する形状情報の集合を基準形状集合と呼ぶ．

6.2.3 多重登録の場合の加重方程式

各登録人物について複数枚の画像が登録されているとき，多重登録の加重方程式が構成される．多重登録の加重方程式は，入力画像の射影係数 s を基準画像の

射影係数の集合 $\{s_{kl}\}$ の最適な加重平均で表現する線形連立方程式となり，次式で表現される．

$$\widehat{\mathbf{S}}_{KL}\mathbf{w} = \widehat{\mathbf{s}} \quad (6.3)$$

ここで， $\widehat{\mathbf{S}}_{KL} = [\widehat{s}_{11} \cdots \widehat{s}_{KL}]$ ， $\mathbf{w} = [w_{11} \cdots w_{KL}]^\top$ であり，最適解は次式で求められる．

$$\mathbf{w} = \widehat{\mathbf{S}}_{KL}^+ \widehat{\mathbf{s}} \quad (6.4)$$

ここで， $\widehat{\mathbf{S}}_{KL}^+$ は $\widehat{\mathbf{S}}_{KL}$ の擬似逆行列である．本章では，登録者数と次元数の関係は $m+1 < KL$ (不足決定系) となるため， $\widehat{\mathbf{S}}_{KL}^+ = \widehat{\mathbf{S}}_{KL}^\top (\widehat{\mathbf{S}}_{KL} \widehat{\mathbf{S}}_{KL}^\top)^{-1}$ となる．不足決定系では加重方程式を満たす \mathbf{w} の中から $\mathbf{w}^\top \mathbf{w}$ を最小化するものが得られる．

ここで， $\mathbf{w} = [w_{11} \cdots w_{KL}]^\top$ は制約条件 $\mathbf{1}^\top \mathbf{w} = 1$ のもとで最適化されることから， $\widehat{\mathbf{S}}_{KL}\mathbf{w}$ は， \mathbf{s} を $\{s_{kl}\}$ の最適な加重平均で表現したものと見なせる．

$$\mathbf{s} = \sum_{k=1}^K \sum_{l=1}^L w_{kl} \mathbf{s}_{kl} \quad (6.5)$$

6.2.4 多重登録のバイアス付き加重方程式

基本加重方程式が不足決定系である場合， \mathbf{s} と \mathbf{s}_{kl} の距離の逆数を対角要素とする \mathbf{B} を導入し，入力と登録の類似性を考慮した加重の調整を行うことができる．

$$\mathbf{B} = \text{diag}(d^{-1}(\mathbf{s}, \mathbf{s}_{11}) \cdots d^{-1}(\mathbf{s}, \mathbf{s}_{KL})) \quad (6.6)$$

$$\text{where } d(\mathbf{s}, \mathbf{s}_{kl}) = \sqrt{(\mathbf{s} - \mathbf{s}_{kl})^\top (\mathbf{s} - \mathbf{s}_{kl})} \quad (6.7)$$

ここで， $\mathbf{s}_{kl} = \mathbf{s}$ である場合には，適当な大数を \mathbf{B} の第 kl 対角要素に設定するものとする． $\mathbf{w} = \mathbf{B}\mathbf{w}'$ 基本加重方程式に代入すると次式が得られる．

$$\widehat{\mathbf{S}}_{KL}\mathbf{B}\mathbf{w}' = \widehat{\mathbf{s}} \quad (6.8)$$

式 (6.8) を多重登録のバイアス付き加重方程式と呼ぶ．式 (6.8) の最適解は $\mathbf{w} = \mathbf{B}[\widehat{\mathbf{S}}_{KL}\mathbf{B}]^+ \widehat{\mathbf{s}}$ となる．この最適解は基本加重方程式の解空間内にあり， $\mathbf{w}'^\top \mathbf{w}' = \mathbf{w}^\top \mathbf{B}^{-2} \mathbf{w}$ を最小化するため，入力画像と各登録画像の距離を考慮に入れた最適化を実現できる．

6.2.5 拡張基準固有顔

Oka-Shakunaga[17, 18] では，拡張固有顔は正規化固有顔と射影係数から形状への連想写像 Ψ_m により定式化されている．一方，本論文では基準固有顔をベースとして系を構成するため，以降では基準固有顔と射影係数から形状への連想写像をまとめたものを拡張基準固有顔と呼ぶ．

6.3 拡張基準固有顔を用いた顔モデル生成法

6.3.1 線形変換を用いた照明変換

山元ら [33], Shakunaga-Shigenari[7] は, 線形変換により画像の照明変換を行う方法を提案している. この方法では, 基準固有顔内においてある照明条件 l の射影係数を異なる照明条件 l' の射影係数へ変換する. これにより, 画像が 1 枚しか得られない場合でも, 照明変動を含む画像集合を近似的に作成できる (このとき作成される画像を, 以下では仮想照明画像と呼ぶ).

ある照明条件 l から異なる照明条件 l' への線形変換を $T_{l'l}$ と表す. このとき, ある照明条件 l で得られた射影係数 \hat{s}_l は $T_{l'l}$ により異なる照明条件 l' の射影係数 $\hat{s}_{l'}$ へ変換される.

$$\hat{s}_{l'} = T_{l'l} \hat{s}_l \quad (6.9)$$

ここで, 照明変換 $T_{l'l}$ は, 照明条件 l の射影係数の集合 $\hat{S}_{Kl} = [\hat{s}_{1l} \cdots \hat{s}_{Kl}]$ と照明条件 l' の射影係数の集合 $\hat{S}_{Kl'} = [\hat{s}_{1l'} \cdots \hat{s}_{Kl'}]$ を用いて定義でされる.

$$\hat{S}_{Kl'} = T_{l'l} \hat{S}_{Kl} \quad (6.10)$$

また, $T_{l'l}$ は次式で計算される.

$$T_{l'l} = \hat{S}_{Kl'} \hat{S}_{Kl}^+ \quad (6.11)$$

線形変換 $T_{l'l}$ により異なる照明条件の射影係数が得られると, 式 (6.2) により仮想照明画像を作成できる.

6.3.2 加重方程式を用いた照明変換

線形変換による方法では, 入力画像の照明条件を正確に推定する必要がある. また, 線形変換を用いた方法は統計的な意味で妥当であるが射影係数の最適線形変換にすぎず, 個人ごとに異なる変換を表してはいない. さらに, この方法は特定の照明条件における画像変換を対象としているため, 不特定の照明条件下で撮影された画像に適用するのは問題である. これらの問題を解決するため, 本論文では加重方程式をベースとし, 照明条件の推定や統計的変換を必要としない仮想照明画像の作成方法を検討した.

ここで, 式 (6.10) において共通の $T_{l'l}$ を推定することやめ, $T_{l'l}$ が集合 \hat{S}_{Kl} を集合 $\hat{S}_{Kl'}$ に変換する作用素であると考え, K 人について基準画像上での変換が (外延的に) 与えられていることが判る. 以下では, この変換を用いて最適線形変換における問題を回避する.

登録者の画像から式 (6.4,6.8) で求められる最適解を用いて別の照明条件で撮影された同じ登録者の画像を作成することを考える. この場合, 個人ごとに計算さ

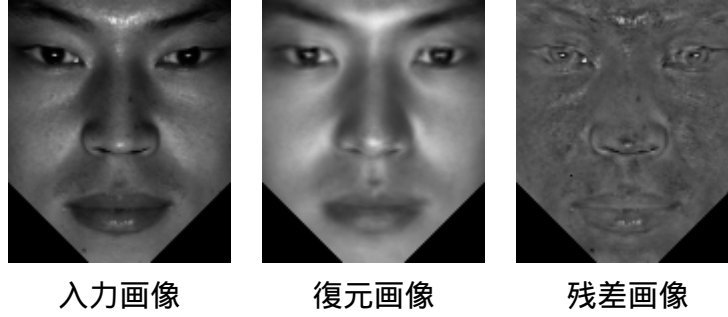


図 6.1: 未登録人物の画像復元例と残差画像

れる加重の和 ($w_k = \sum_{l=1}^L w_{kl}$) は変わらないと仮定する．また，この操作により得られる加重を人数分並べたベクトルを $\mathbf{w}_K = [w_1 \cdots w_K]^\top$ で表す．上記の仮定を利用すると， \mathbf{w}_K と各照明条件の射影係数の集合を用いて，次式のように $\hat{\mathbf{s}}$ から照明条件 l への変換を照明推定なしに行うことができる．

$$\mathbf{s}_l = \sum_{k=1}^K \left(\sum_{l'=1}^L w_{kl'} \right) \mathbf{s}_{kl} \quad (6.12)$$

式 (6.12) により \mathbf{s}_l が得られると，式 (6.2) により照明条件 l の仮想照明画像を作成できる．

6.3.3 残差画像の利用

顔には皺や黒子などその人物特有の情報が存在する．しかし，その人物が基準固有顔に登録されていない，もしくは，同様の特徴を持つ人物が基準固有顔内に存在しない場合，それらの情報は画像の復元時に失われてしまう．例えば，図 6.1 の例では，入力画像に存在する皺や黒子が復元画像では失われている．一方で，これらの失われた情報はその人物らしい仮想照明画像を作成する上で重要な情報である．ここで，これらの情報は基準固有顔で表現することができないため入力画像と復元画像の残差として現れる．ここで残差画像は次式で計算される．

$$\mathbf{v}^\# = \mathbf{v} - \tilde{\Phi}_m \hat{\mathbf{s}} = \mathbf{v} - \Phi_m \mathbf{s} - \bar{\mathbf{v}} \quad (6.13)$$

本論文では，残差画像 $\mathbf{v}^\#$ を利用することにより復元時に失われた人物特有の情報含む仮想照明画像を生成する．照明変換によって得られた射影係数を \mathbf{s}_l とすると残差画像 $\mathbf{v}^\#$ を用いて，人物特有の情報を含む仮想照明画像は次式で得られる．

$$\mathbf{v}_l^* = \tilde{\Phi}_m \hat{\mathbf{s}}_l + \mathbf{v}^\# = \Phi_m \mathbf{s}_l + \bar{\mathbf{v}} + \mathbf{v}^\# \quad (6.14)$$

6.3.4 複数枚を登録とした場合の作成方法

山元ら [33] の方法では，登録が 2 枚以上得られた場合についても議論しており，同様の拡張がこの系についても行える．

加重方程式を用いた方法では，入力画像の照明条件が異なっても個人ごとに計算される加重の和 (w_K) は変わらないという仮定を導入しているため，複数登録時はこの制約を満たす加重方程式を構成する必要がある，

登録画像が 2 枚与えられた場合を考える．それぞれの射影係数を $\hat{s}^{(1)}$ ， $\hat{s}^{(2)}$ とした場合，制約を満たす加重方程式は次式で構成される．

$$\begin{aligned} \begin{bmatrix} \hat{s}^{(1)} \\ \hat{s}^{(2)} \\ \mathbf{0}_K \end{bmatrix} &= \begin{bmatrix} \hat{\mathbf{S}}_{KL} & \\ & \hat{\mathbf{S}}_{KL} \\ \mathbf{\Omega}_{KL} & -\mathbf{\Omega}_{KL} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{w}}_{KL}^{(1)} \\ \hat{\mathbf{w}}_{KL}^{(2)} \end{bmatrix}, \\ \mathbf{\Omega}_{KL} &= \left. \begin{bmatrix} \mathbf{1}_L^\top & & \\ & \ddots & \\ & & \mathbf{1}_L^\top \end{bmatrix} \right\} K \text{ 行} \end{aligned} \quad (6.15)$$

ここで， $\mathbf{w}_{KL}^{(1)} = [w_{11}^{(1)} \cdots w_{KL}^{(1)}]^\top$ ， $\mathbf{w}_{KL}^{(2)} = [w_{11}^{(2)} \cdots w_{KL}^{(2)}]^\top$ は各入力画像に対応する最適解を示し， $\mathbf{1}_L$ は要素が全て 1 の L 次元ベクトルを示す．

式 (14) は，制約条件 $\sum_l w_{kl}^{(1)} - \sum_l w_{kl}^{(2)} = 0$ のもとで最適化されるため，人物間のバランスを変えることなく仮想照明画像の作成が可能となる．

6.3.5 複数枚を登録とした場合の残差の取り扱い

なお，入力画像が複数ある場合，残差画像は各入力画像に対して計算される．このため，両者の情報を利用することで仮想照明画像の質が向上すると思われる．ここで，各画像の残差を比較すると，皺や黒子などの情報は影や鏡面反射などにより顔のテクスチャが見えなくなっている場合をのぞき，これらの情報は両方の残差に同様に現れると思われる．しかし，影や鏡面反射が入力画像に存在する場合，それらが存在する部分では残差の絶対値が大きくなる．これらのことから，本論文では，複数枚の残差画像を比較し入力画像との残差の絶対値が最も小さな画素値を選択して作成される残差画像を足し合わせて仮想照明画像を作成した．

6.4 加重方程式による顔モデリング

本論文では，形状復元に関しては，Oka-Shakunaga [17, 18] に従うこととした．Oka-Shakunaga [17, 18] では，入力画像から射影係数 \hat{s} が与えられると，連想写像 Ψ_m により入力画像の形状 Z が生成される．

$$Z = \Psi_m \hat{s} \quad (6.16)$$

ここで, $\Psi_m = \zeta_{KL} \hat{S}_{KL}^+$ であり, $\zeta_{KL} (= [Z_{11} \cdots Z_{KL}])$ は各登録者の奥行きベクトルを並べた行列である.

ここで, [17, 18] において射影係数から形状への連想写像は加重方程式に帰着できることが示されており, 人物識別や形状推定は w のドメインで統一的に行うことが可能となっている. すなわち, 入力画像の形状は, 射影係数と同様に登録者の奥行きベクトルの最適な加重平均で計算することができる.

$$\mathbf{Z} = \sum_{k=1}^K \sum_{l=1}^L w_{kl} \mathbf{Z}_{kl} \quad (6.17)$$

6.5 静止画像実験

6.5.1 データセット

ここでは, 静止画像を用いた実験を行った. 以下に, 実験に使用するデータセットの概要を示す.

Data100x1 5.2.1 節で述べた独自データセットであり, 100 人の正面顔 (無表情) を 24 種類の照明条件で撮影したもの.

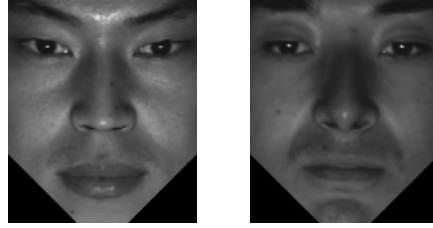
6.5.2 実験条件

各データセットをテスト画像集合と基準画像集合に分割する. Data100x1 において, Data100x1 の最初の 10 人 x24 照明画像をテスト画像集合へ, 各データベースの残りの画像は基準画像集合とした. Data100x1 から 130 次元の基準固有顔 (CS90) を作成し, 仮想照明画像の作成に用いた.

6.5.3 比較した手法と評価方法

本論文では, 以下の 3 通りの方法を比較した.

1. 線形変換による照明変換 (式 (6.9))
2. バイアス付き加重方程式 (多重登録) による照明変換 (式 (6.12))
3. バイアス付き加重方程式 (多重登録) による照明変換と残差画像の併用 (式 (6.12, 6.14))



テスト人物 ID1 テスト人物 ID5

図 6.2: 入力画像の例

なお，簡単な事前実験により加重方程式はバイアス付き加重方程式を採用した（以降では，単に加重方程式と呼ぶ）。

本論文では，顔追跡・認識融合系への適用のため画像集合ではなく画像集合から作成される正規化固有顔を評価に用いた．西山ら [14] は，2つの部分空間の類似度の評価に正準角と呼ばれる2つの部分空間のなす角度を用いている．

実画像集合から作成される M 次元正規化固有顔を $\Phi_M^{(1)}$ ，仮想画像集合から作成される M 次元正規化固有顔を $\Phi_M^{(2)}$ とする（正規化固有顔の平均にはグラムシュミットの直交化を施しておく）．このとき，評価値 c は， $\Phi_M^{(1)} \Phi_M^{(2)\top}$ の特異値 $\lambda_m (m = 1, \dots, M)$ (正準相関) の平均として求まる．

$$c = \frac{1}{M} \sum_{m=1}^M \lambda_m \quad (6.18)$$

また，形状については本人の奥行きベクトルと推定された奥行きベクトル間で RMS 残差を計算し評価に用いた．

6.5.4 静止画像実験結果

登録が1枚の場合の場合（入力の照明条件が比較的良好な条件）について実験を行ったところ，テストに使用した10人の評価値の平均（各個人ごとの評価値を図6.5(a)に示す）は，方法1で0.937，方法2で0.944，方法3で0.945であり，加重方程式と残差画像を併用した方法が最も良かった．図6.2のテスト画像から各手法により作成された仮想照明画像を図6.3，6.4に示す．これらの結果から，方法3により作成される仮想照明画像が最もテスト人物の特徴を復元できていることがわかる．特に，図6.4の例では，方法1の仮想照明画像で一部劣化が見られたが，その他の方法では劣化無く画像を作成できており，さらに方法3では入力人物の特徴も復元できている．これらの結果から，加重方程式による仮想照明変換と残差画像の併用は個人の特徴を反映した仮想照明画像の作成に有効であることが確認できた（評価値で見ると差はわずかであるが，方法3における仮想照明画像中にはその他の方法では復元できない皺などの情報が存在していることが確認できる）．また，登



図 6.3: テスト人物 1 の仮想照明画像の比較

録が 2 枚の場合は，テストに使用した 10 人の評価値の平均（各個人ごとの評価値を図 6.5(b) に示す）は，方法 2 で 0.945，方法 3 で 0.946 であり，登録を増やすことにより精度が上がることを確認できた．

一方で，評価値が悪化した例も存在した．これらの人物（1 枚登録では人物 7，2 枚登録では人物 9）においては，鏡面反射や影などが残差画像からうまく分離されないまま仮想照明画像に加えられたため，評価値が下がったものと思われる．しかし，平均値としては方法 2，3 の方が高く，残差画像の取り扱いを改善することでさらに評価値は上昇すると思われる．

また，全ての照明条件 ($l = 1, \dots, L$) の画像をそれぞれ 1 枚のテスト画像として仮想照明画像を作成した場合の評価値の平均と最大値，最小値をまとめたものを図 6.6 に示す．これらを解析した結果，影などが少ない画像では評価値が高く，影などが多く含まれる画像を入力とした場合には評価値が低いという結果となった．入力に含まれる影が直接照明変換に影響を与えたことによる画像の質の低下のため，もしくは，残差から個人特徴以外を排除できず，そのまま式 (6.14) で用いた



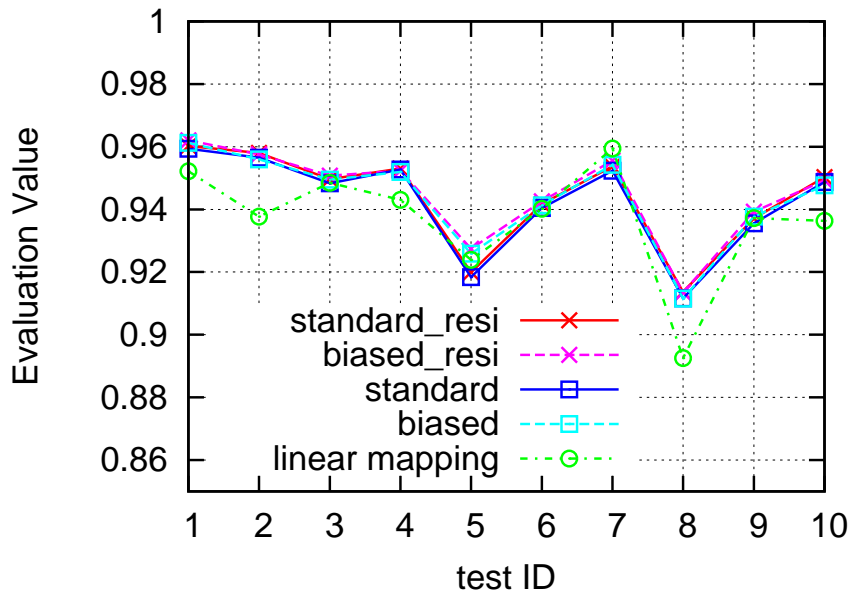
図 6.4: テスト人物 5 の仮想照明画像の比較

ためであると考えられる。現状では、比較的穏やかな照明条件において適用できることが確認できたが、それ以外の画像へ適用する場合は、入力画像に含まれる残差から影や個人特徴以外のノイズを除去する必要があると思われる。

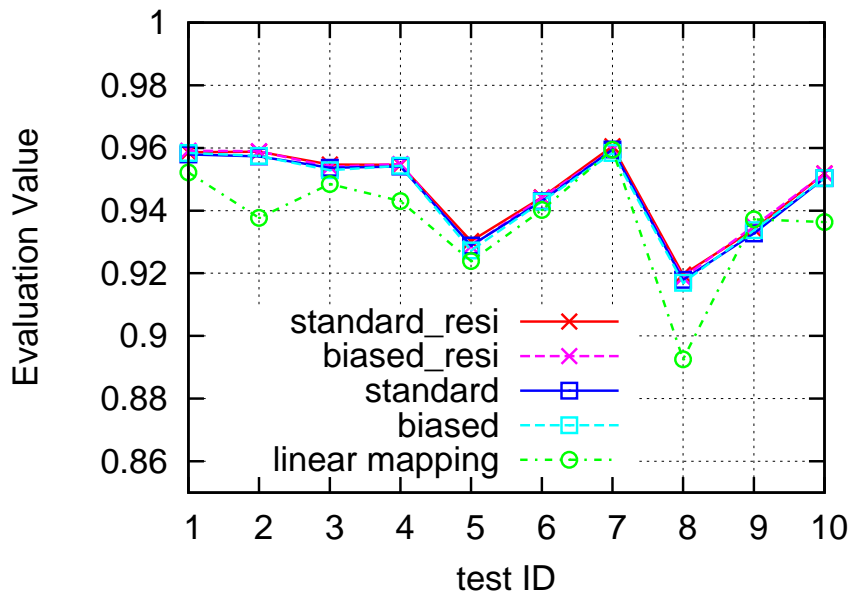
図 6.3, 6.4 の画像から作成される正規化固有顔を画像化したものを図 6.7 に示す。これらの図から、加重方程式と残差の併用による方法は、固有軸に関しては変化を確認しづらいが、平均画像においては個人性を良く復元できていることが確認できた。

6.5.5 顔モデリングの結果

図 6.8 に形状モデリングの結果を示す。図 6.8 から、斜め 45 度から見たモデル (左から 2, 5 列目) や横顔のモデル (左から 3, 6 列目) では鼻や頬の周辺に違いが現れているものの、顔追跡・認識融合系において有効に動作する角度 (30 度, 左から 1, 4 列目) では実物とほぼ変わらない結果が得られた (平均形状に近い人物では



(a) 人物ごとの評価値 (1 枚登録)



(b) 人物ごとの評価値 (2 枚登録)

図 6.5: 人物ごとの評価値

差が確認できないが、頬が痩けている人物では実形状と平均形状間での差が確認できる)。

また、個人ごとに計算された推定形状と実形状の RMS 残差と登録者 90 人の平均形状と実形状の RMS 残差をまとめたものを図 6.9 に示す。実形状データが平均形状に近い人物では、推定形状よりも平均形状との RMS 残差が小さい場合があっ

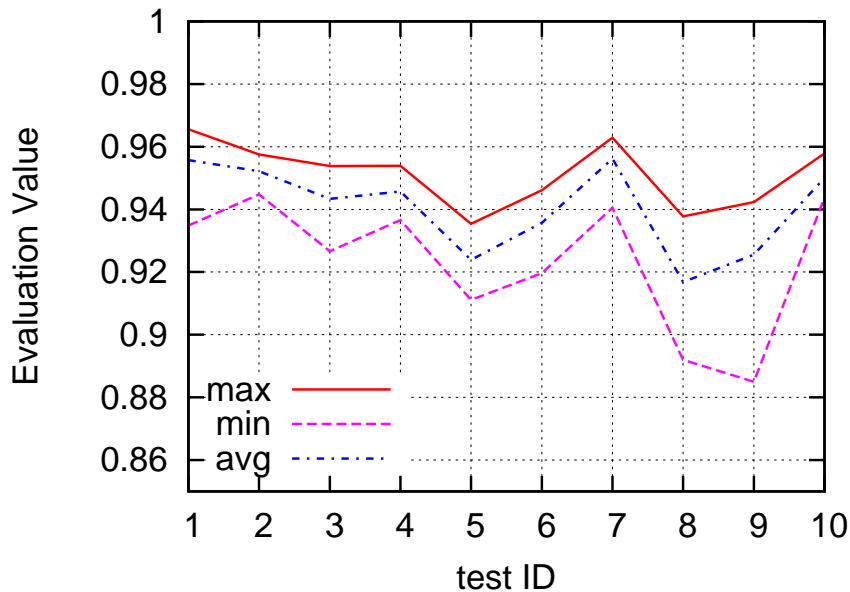


図 6.6: 全ての照明条件における評価値の分布

た．しかし，残差の平均（標準偏差）をとると，推定形状と実形状間における数値は 3.15(1.00)であったのに対して，平均形状と実形状間における数値は 3.47(1.27)であった．これらの結果から，基本的に顔追跡・認識融合系への適用の観点からは，推定を行わずに平均形状を用いることよりも，加重方程式による形状推定を行った方が有効であることが確認できた．

6.6 顔追跡・認識融合系への適用

方法 1,3(線形変換による照明変換，加重方程式による照明変換と残差の併用)により得られた 10 人の仮想照明画像集合と形状データから拡張固有顔を作成し，顔追跡・認識融合系に適用した(人物識別の方法は 3.4.4 節に従う)．適用例を図 6.10 に示す．

この例では，動画像中の人物が 10 人中の誰か判っていない状態(図左列：1 フレーム目)から追跡を開始しており，実画像から作成される拡張固有顔を用いた場合は，5 フレーム目(図右列)において人物が正しく識別された．これに対して，方法 1 で作成された拡張固有顔を用いた場合は，5 フレーム目になってもまだ人物を識別できていないことが判る．しかし，方法 3 では実画像と同様に 5 フレーム目において人物が正しく識別された．(テクスチャや形状を見ても遜色ない追跡が行えていることが判る)．

これらの結果から，加重方程式による照明変換と残差の併用による仮想照明画像の作成方法は，1 枚の画像しか得られない場合においても，実画像を登録した場



図 6.7: 各画像集合から作成された正規化固有顔の比較

合と比較して、遜色なく照明変動に対応しながら人物顔の追跡・認識を行えることを確認できた

6.7 まとめ

本論文では、加重方程式をベースとして、少数の画像から異なる照明条件の画像を近似的に作成する方法を提案した。実験により顔モデル作成が正しく行えることと、顔追跡・認識融合系への適用が可能であることが確認できた。また、人物の細かい特徴など従来の方法では失われていた情報を残差画像の併用によって復元することができ、より本人らしい固有顔を作成できた。今後の課題として、残差画像内の影や鏡面反射などのノイズと個人特徴の適切な分離が挙げられる。



図 6.8: 実形状データ, 推定形状, 平均形状の比較

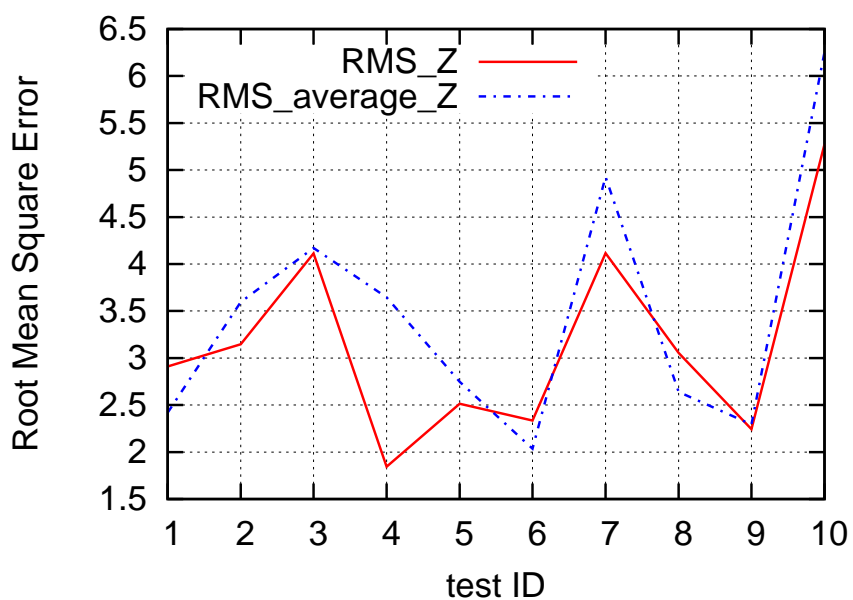
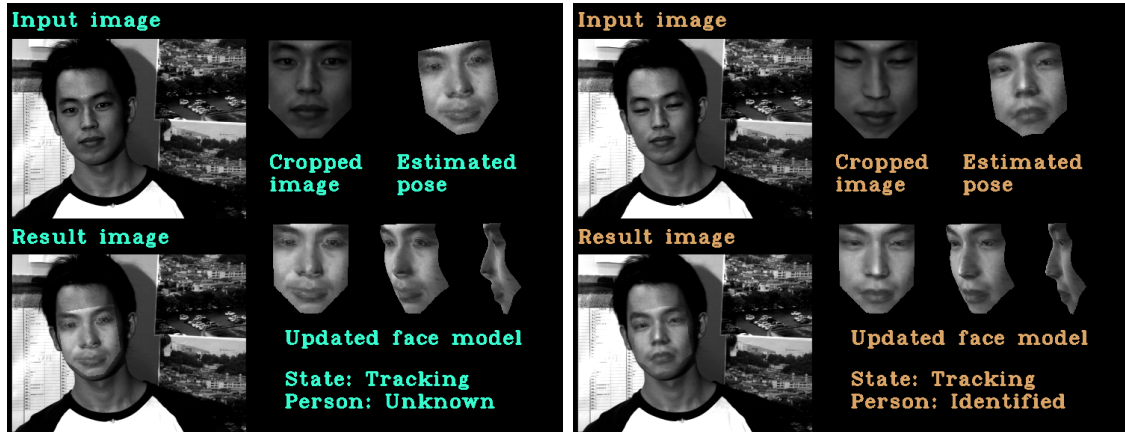
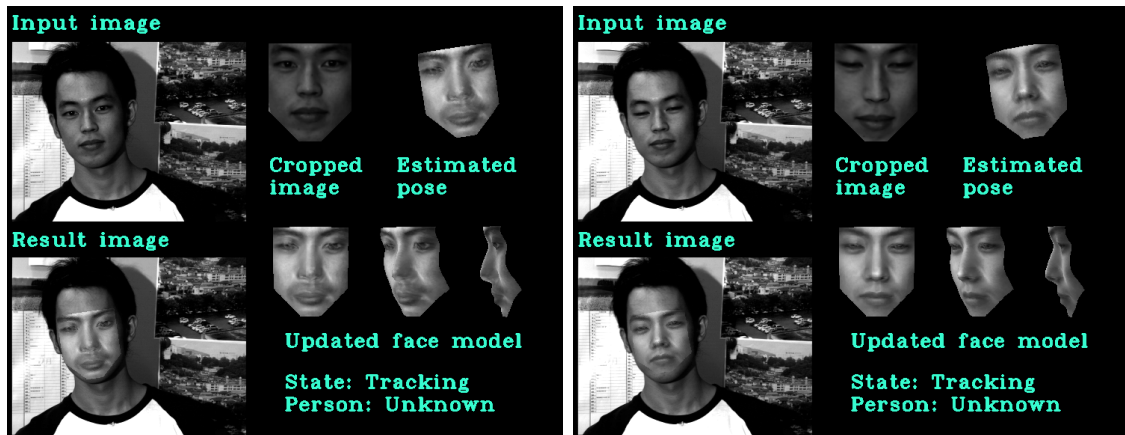


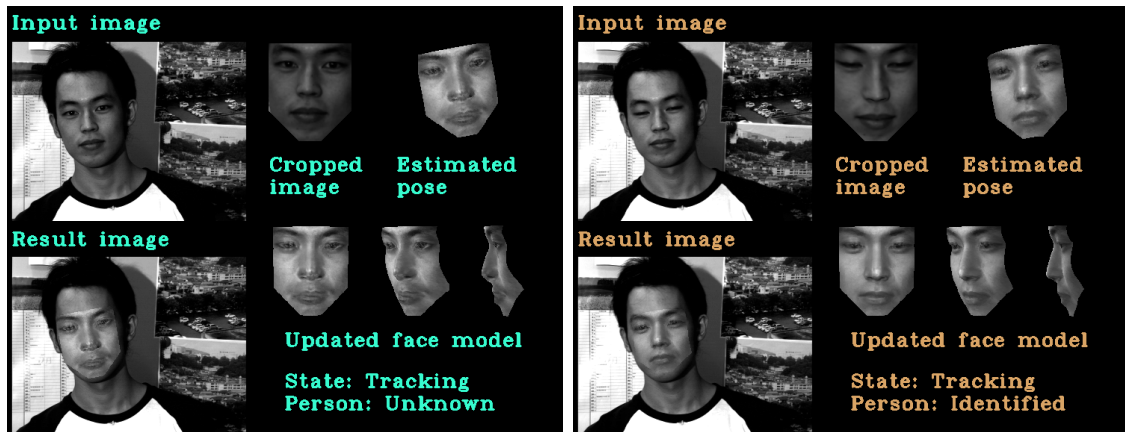
図 6.9: 人物ごとの RMS 残差



実画像から作成された拡張固有顔で追跡した結果



方法 1 から作成された拡張固有顔で追跡した結果



方法 3 から作成された拡張固有顔で追跡した結果

図 6.10: 顔追跡・認識融合系への適用例

第7章 結論

本論文では，Oka-Shakunaga[17, 18]の基本加重方程式をベースとした新たな2つの加重計算法(1, 2)による，実時間顔追跡・認識融合系の高性能化について述べた．また，これらの方法をベースとして，顔追跡・認識融合系のさらなる高性能化を目指すための基礎検討(3, 4)を行った．以下に，本論文で述べた2つの加重計算法および2つの基礎検討についてまとめる．

1. 階層化加重方程式による加重計算法の提案

基本加重方程式において個人ごとに行っていた登録処理を，登録人物集合のサブセットを用いて階層的に行うことで，登録者数が増加した場合においても小規模な加重方程式を構成することが可能な方法を提案した．また，異なるサブセットから得られる複数の階層化加重方程式を組み合わせることで識別性能を改善する方法も提案した．

289人登録のデータベースを用いた静止画像実験と，100パターン登録(10人×10表情)のデータベースを用いた動画像実験の結果から，この方法が，登録者数が増加した場合の実時間顔追跡・認識融合系の高性能化に有効であることを確認した．

2. 並列不足決定系による加重計算法の提案

単一の不足決定系加重方程式を構成するのではなく，与えられた基本加重方程式の中に複数の小規模な不足決定系加重方程式を構成し，個々の不足決定系の最適解を組み合わせることで，大規模化によって生じる問題を回避するとともに識別性能の改善を目指す方法を提案した．

289人登録のデータベースを用いた静止画像実験と，100パターン登録(10人×10表情)のデータベースを用いた動画像実験の結果から，この方法が，登録者数が増加した場合の実時間顔追跡・認識融合系の高性能化に有効であることを確認した．

3. 階層化加重方程式と並列不足決定系を用いた大規模識別の検討

大量のシミュレーション画像で構成されるデータベースを用いて，加重方程式による識別系における登録者数を大規模化した．まず，階層化加重方程式と並列不足決定系を単独に用いて大規模識別系を構成し，基礎性能評価を行った．次に，各大規模識別系の特性を分析した経緯から，2つの加重計算法の組み合わせによって大規模識別系を構成した．

2197 人登録のシミュレーションデータベースを用いた実験の結果から，並列不足決定系による候補選択の後に階層化加重方程式による識別を行う方法が最も良い識別性能を与えることを確認した．

4. 加重方程式を用いた顔モデリングの検討

実時間顔追跡・認識融合系 [17, 18] における登録処理を簡略化する方法として，不足決定系加重方程式をベースとした顔モデル生成法を検討した．実験により，正面向き画像 1 枚から必要なデータを推定し，顔モデルを生成できることを確認した．

以上の実験結果から，階層化加重方程式と並列不足決定系は実時間顔追跡・認識融合系の高性能化に有効であることが確認できた．また，大規模識別系の構成や登録処理の簡略化へも応用できたことから，今後の顔追跡・認識融合系のさらなる高性能化が期待できる．

謝辞

本研究を行なう機会を与えて下さり，御指導・御教示を賜りました尺長健教授に深く感謝致します．本稿の執筆にあたり，御指導・御教示を賜りました太田学教授，竹内孔一講師に深く感謝致します．研究内容について，日頃から適切な御助言を賜りました右田剛史助教に深く感謝致します．また，本研究を進めるに当たり，貴重な御助言を賜りました岡裕希氏，データベース作成および整理に御協力頂いた片山聡一郎氏，大規模シミュレーション実験に御協力頂いた井上航氏，福田剛士氏に深く感謝致します．最後に，画像撮影に御協力頂いた皆様に深く感謝致します．

参考文献

- [1] M. Turk and A. Pentland, Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71–86, 1991.
- [2] P. N. Belhumeur, J. P. Hespanha, and D. J. 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.
- [3] B. Moghaddam and A. Pentland, Probabilistic visual learning for object representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 696–710, 1997.
- [4] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 6, pp. 643–660, 2001.
- [5] M. A. O. Vasilescu and D. Terzopoulos, Multilinear analysis of image ensembles: Tensorfaces, *Proc. The European Conference on Computer Vision*, pp. 447–460, 2002.
- [6] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, Robust face recognition via sparse representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 2, pp. 210–227, 2009.
- [7] T. Shakunaga and K. Shigenari, Decomposed eigenface for face recognition under various lighting conditions, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 864–871, 2001.
- [8] T. Shakunaga, M. Matsubara, and N. Noguchi, Appearance tracker based on sparse eigentemplate, *Proc. International Conference on Machine Vision and Applications*, pp. 13–17, 2005.
- [9] Y. Matsubara and T. Shakunaga, Sparse template matching and its application to real-time object tracking, *IPSJ Transactions on Computer Vision and Image Media*, Vol. 46, No. SIG9, pp. 60–71, 2005.

- [10] V. Blanz and T. Vetter, Face recognition based on fitting a 3d morphable model, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 9, pp. 1063–1074, 2003.
- [11] M. Castelan, W. A. P. Smith, and E. R. Hancock, A coupled statistical model for face shape recovery, *IEEE Transactions on Image Processing*, Vol. 16, No. 4, pp. 1139–1151, 2007.
- [12] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 511–518, 2001.
- [13] V. Belle, T. Deselaers, and S. Schiffer, Randomized trees for real-time one-step face detection and recognition, *Proc. International Conference on Pattern Recognition*, pp. 1–4, 2008.
- [14] 西山正志, 山口修, 福井和広, 多重制約相互部分空間法を用いた顔画像認識, *電子情報通信学会論文誌 (D-II)*, Vol. J88–D-II, No. 8, pp. 1339–1348, 2005.
- [15] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 6, pp. 681–685, 2001.
- [16] Y. Xu and A. K. Roy-Chowdhury, A physics-based analysis of image appearance models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, pp. 1681–1688, 2011.
- [17] Y. Oka and T. Shakunaga, Sparse eigentracker augmented by associative mapping to 3d shape, *Proc. IEEE Conference on Automatic Face and Gesture Recognition*, pp. 649–656, 2011.
- [18] Y. Oka and T. Shakunaga, Real-time face tracking and recognition by sparse eigentracker augmented by associative mapping to 3d shape, *Image and Vision Computing Journal*, Vol. 30, No. 3, pp. 147–158, 2012.
- [19] 坂上文彦, 尺長健, 正規化固有空間への最適部分射影とその応用, *情報処理学会論文誌*, Vol. 44, No. SIG17, pp. 100–108, 2003.
- [20] Y. Oka, T. Kuroda, T. Migita, and T. Shakunaga, Tracking 3d pose of rigid object by 3d sparse template matching, *Proc. International Conference on Image and Graphics*, pp. 390–397, 2009.

- [21] 坂上文彦, 尺長健, 並列部分射影による顔画像認識と decomposed eigenface 法への適用, 電子情報通信学会論文誌 (D), Vol. J92-D, No. 5, pp. 1652–1661, 2008.
- [22] T. Kohonen, Self-Organization and Associative Memory, Third Edition, Springer Verlag, 1989.
- [23] C. F. Laywine and C. M. Mullen, Discrete mathematics using latin squares, Wiley Interscience, 1998.
- [24] 尺長健, 金子博, 淀川英司, 2次統計量の線分スペクトル分解による手書き文字の筆者認識, 電子情報通信学会論文誌 (D), Vol. J67-D, No. 7, pp. 776–783, 1984.
- [25] F. Sakaue and T. Shakunaga, Face recognition by parallel partial projections, Proc. Asian Conference on Computer Vision, Vol. 1, pp. 1440–150, 2004.
- [26] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, Multi-pie, Proc. IEEE Conference on Automatic Face Gesture Recognition, pp. 1–8, 2008.
- [27] H. Chugan and T. Shakunaga, Hierarchical approach to weight equations in face tracking and recognition framework, Proc. IEEE Conference on Automatic Face and Gesture Recognition, pp. 1–8, 2013.
- [28] H. Chugan, Y. Oka and T. Shakunaga, Underdetermined approach to weight equations in tracking and recognition framework, Proc. International Conference on Machine Vision and Applications, pp.242–246, 2013.
- [29] 中岸久佳, 岡裕希, 尺長健, 並列不足決定系による顔追跡認識融合系の大規模化, 電子情報通信学会論文誌 (D), Vol. J97-D, No. 12, pp. 1821–1830, 2014.
- [30] 尺長健, 福田剛士, 中岸久佳, 2階層加重方程式の多段構成による多クラス識別, 電子情報通信学会技術報告, Vol. 114, No. 197, PRMU2014–38, pp. 13–19, 2014.
- [31] 井上航, 福田剛士, 中岸久佳, 尺長健, 並列不足決定系加重方程式による多クラス識別, 電子情報通信学会技術報告, Vol. 114, No. 409, PRMU2014–105, pp. 227–232, 2015.
- [32] H. Chugan, T. Fukuda and T. Shakunaga, Challenge to scalability of face recognition using universal eigenface, Proc. Pacific Rim Symposium on Image and Video Technology, pp. 51–62, 2015.

- [33] 山元宣政, 重成一真, 尺長健, 仮想部分空間法による照明変動に影響されない顔認識, 電子情報通信学会論文誌 (D-II), Vol. J84, No. 8, pp. 1753–1761, 2001.
- [34] 中岸久佳, 福田剛士, 尺長健, 拡張固有顔を活用した少数枚画像からの顔モデル生成法, 電子情報通信学会技術報告, Vol. 115, No. 224, PRMU2015–87, pp. 143–148, 2015.

発表文献リスト

論文誌

1. 中岸久佳, 岡裕希, 尺長健, 並列不足決定系による顔追跡認識融合系の大規模化, 電子情報通信学会論文誌 (D), Vol. J97-D, No. 12, pp. 1821–1830, 2014.

国際会議

1. H. Chugan, T. Fukuda and T. Shakunaga, Challenge to scalability of face recognition using universal eigenface, Proc. Pacific Rim Symposium on Image and Video Technology, pp.51–62, 2015.
2. H. Chugan, Y. Oka and T. Shakunaga, Underdetermined approach to weight equations in tracking and recognition framework, Proc. International Conference on Machine Vision and Applications, pp.242–246, 2013.
3. H. Chugan and T. Shakunaga, Hierarchical approach to weight equations in face tracking and recognition framework, Proc. IEEE Conference on Automatic Face and Gesture Recognition, pp.1–8, 2013.