

前腕回旋を含んだ手指動作の3次元形状推定

玉城 絵美

筑波大学 システム情報工学研究科

知能機能システム 博士前期課程2年

概要

手指に何も装着してない状態から，カメラを用いて手指の3次元形状を推定する手法の一つに，データベースから類似データを高速検索する手法がある．著者らの研究グループは，対にした手指の輪郭線特徴量および関節角度データを類似度に応じて多階層にクラスタリングすることにより類似画像を高速に検査するヒト手指形状システムを構築した．しかし，手の表裏情報を十分に利用していなかったため，前腕回旋の動作を含んだ3次元手指の推定精度に限界があった．

そこで本研究では，第一に，ヒトが表裏を含んだ手指の形状を理解する場合に用いる視覚的手掛かりを，手の皺，爪，皮膚の色を用いて心理知覚実験により検討した．その結果，視覚的手掛かりにおける関与度の比率は，爪が25%，皺が20%，皮膚の色が15%であった．第二に，その知見に基づき，手指の輪郭線の特徴量に爪の位置情報を追加することにより，指先の位置および表裏の情報も併用する手指形状推定システムを提案した．

提案するシステムでは，データグローブを装着した状態で手画像を取得し，手指の輪郭線特徴量，爪の位置情報及び関節角度データを持ったデータセットを作成した．次に，データセットを前腕回旋角度と手指関節角度によってクラスタリングし，データベースを構築した．最後に，手指画像がシステムへ入力されると，その手指画像の爪の位置情報と輪郭線特徴量を元にデータベース内の類似データセットを検索し，そのデータセットが持つ手指関節角度を出力する手指形状推定システムを構築した．また，このとき，肌領域と爪領域を高速に抽出できる個人別表色系を提案した．

データグローブを装着した状態で手指形状推定を行い，データグローブから出力された関節角度と，本システムから出力された関節角度を比較して行った評価実験の結果，前腕回旋動作に対応しながら，推定誤差の標準偏差7.23度での手指関節角度の推定を実現した．一方，爪の位置情報を持たない推定システムの標準偏差は22.0度であった．また，CPUがPentium IV，クロック周波数2.8GHzのPCを用いたとき，本システムの平均処理速度は100fpsの高速処理であった．一連の結果から本システムの有用性が確認できた．

目次

第1章	序論	1
第2章	心理知覚実験による検討	4
2.1	目的および方法	4
2.2	結果と考察	5
第3章	システムの構成	7
3.1	目的	7
3.2	データベース作成	7
3.2.1	処理の流れ	7
3.2.2	個人別表色系	9
3.2.3	原理	10
3.2.4	個人別表色系の作成	12
3.2.5	肌色領域の抽出	14
3.2.6	爪領域の抽出	15
3.2.7	輪郭線の特徴量化	16
3.2.8	爪の位置データ取得	16
3.2.9	データセット作成	18
3.2.10	データベース構築	19
3.3	形状推定	20
3.3.1	処理の流れ	20
3.3.2	データ検索範囲指定	21
3.3.3	類似度の計算	21
第4章	評価実験	23
4.1	個人別表色系	23
4.1.1	方法	23
4.1.2	速度	24
4.1.3	背景別	25
4.1.4	照明別	26
4.1.5	肌色分布の切り出し精度についての楕円体モデルとの比較	28
4.1.6	計算量についての楕円体モデルとの比較	29

4.2	形状推定	29
4.2.1	実験手法	29
4.2.2	結果と考察	30
4.2.3	爪の位置情報を持たないシステムとの比較	33
第 5 章	結論	35
	謝辞	37
	参考文献	38
付録 A	カラーリストバンドによる前腕回旋角度の検出方法	40
付録 B	形状推定の結果	43

目次

1.1	手指動作による3次元の情報入力ユーザインタフェースの例	1
1.2	片手の各PIP関節角度が10度違う手指CGモデルの例	2
2.1	手の要素を表裏逆にした画像.(a)通常の手.(b)皺.(c)爪.(d)皮膚の色	4
2.2	視覚的手掛かりとその強度ごとの関与度.(a)皺.(b)爪.(c)皮膚の色	6
3.1	データベース作成のフローチャート	8
3.2	色差空間上の肌色領域切り出しの比較.(a)通常の場合.(b)個人の表色系の例	11
3.3	個人別表色系での肌色分布.(a)蛍光灯環境下.(b)白熱灯環境下	13
3.4	肌色抽出処理の流れ	14
3.5	爪領域の抽出	15
3.6	高次局所自己相関の局所パターン	16
3.7	誤って形状推定される例	17
3.8	爪の位置情報による手指の表裏と指先位置認識の例.(a)爪なし.(b)爪あり	17
3.9	データセット作成概要図	18
3.10	データベース内の階層概略図	19
3.11	推定処理のフローチャート	20
3.12	形状推定の結果	21
3.13	検索範囲の例	22
4.1	速度評価実験	24
4.2	背景別誤認識率	25
4.3	照明別誤認識率.(a)誤認識率.(b)肌色未検出率.(c)背景部分誤認識率	27
4.4	色差空間上の肌色領域の切り出しの比較.(a)楕円体モデルの内外判定の例.(b)個人別表色系の例	28
4.5	前腕回旋時の中指第2関節角度.(a)推定精度.(b)推定誤差	31
4.6	前腕回旋時の親指第1関節角度.(a)推定結果.(b)推定誤差	32
4.7	爪の位置情報を持たないシステムでの人差し第2関節角度の推定精度.(a)推定精度.(b)推定誤差	34
5.1	カラーリストバンドによる前腕回旋角度の検出の様子	40
5.2	カラーリストバンドによる前腕回旋角度の検出のフローチャート	41
5.3	カラーリストバンドのペーパークラフト例	42

5.4	手指の各関節を動かさない状態.	43
5.5	手指の各関節を自由に動かした状態.	46

第1章 序論

カメラ動画入力を中心とした、手指に何も装着しない状態での、3次元手指動作による情報入力ユーザインタフェースを実現するためには、前腕回旋動作を含んだ3次元の手指形状推定を高速かつ高精度に行わなければならない。

図 1.1 に、手指動作による3次元の情報入力ユーザインタフェースの例を示す。この例では、ビデオカメラつきHMD(Head Mounted Display)を装着し、作業者自身の手動作を撮影して手指形状推定を行っている。推定された手指形状のCGモデルと仮想物体同士に当たり判定をつけることによって、3次元物体を操作することが可能となる。

3次元手指動作による情報入力ユーザインタフェースが実現すると、マウスやキーボードに比べ、入力できる情報量が格段に増えるため、作業効率が向上する。また、操作自体が実世界に近くなるため、仮想空間内の操作がより直感的になる。

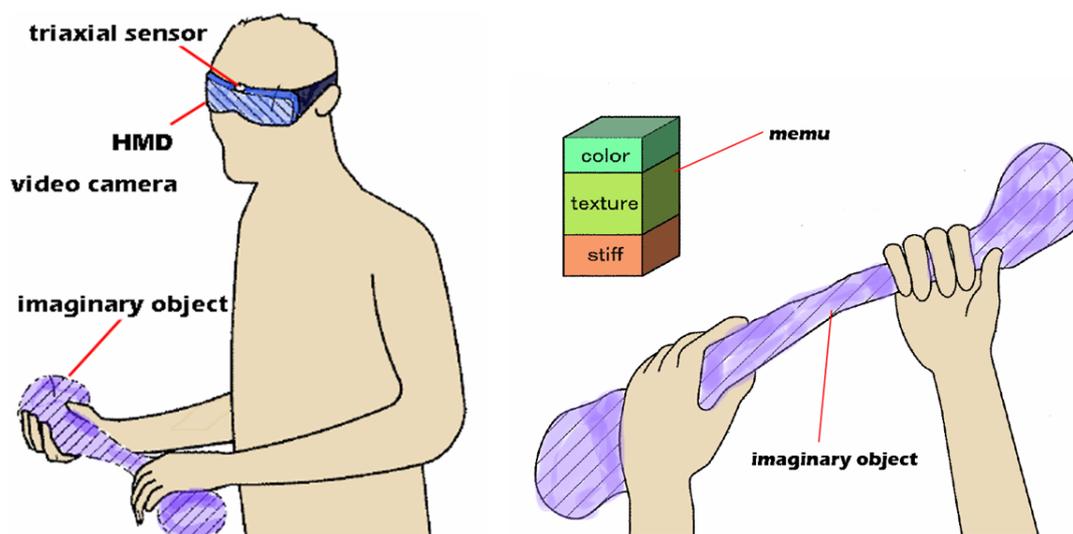


図 1.1: 手指動作による3次元の情報入力ユーザインタフェースの例

手指形状推定システムには、形状推定の誤差の標準偏差 10 度以内の推定精度と 70fps 以上の処理速度が必要である。手指ジェスチャ認識では、図 1.2 に示す様に、関節角度の違いが 10 度以内であれば、視覚上の差異をほとんど認められないため、システムの誤差出力として許容範囲内といえる。従って、形状推定誤差の平均値もしくは標準偏差 10 度以内の推定精度が必要である。また、ディスプレイ描画の際に画面のちらつきが気にならない 70Hz の垂直操作周波数と同期できる処理速度、つまり 70fps 以上の速度で形状推定処理を行う必要がある。

しかしながら、人間の手指は多関節構造を持ち、さらに前腕回旋動作により手指全体が大きく位置と姿勢を変えるため、単眼カメラによる形状推定は非常に困難である。物体の位置と姿勢を測定するための一般的な手法として多用されるステレオカメラなどの両眼視による認識は、処理時間が増大する問題があり、また以下に示す他の手法も推定精度と速度に問題をかかえている。

島田らはモデルマッチングを用いて単眼カメラのみの実時間の形状推定を行った [1]。画面上の突起領域が指先に対応するという対象物体に関する知見を取り入れている。しかし、マッチング処理にモデルとの対応付けをすることは処理の効率化には有利であるが、対応付けできない時は被覆領域を使ったモデルマッチングであり、処理速度や精度に問題が生じる。つまり、指先の位置情報をマッチング以外の手法で取得する必要があると考えられる。

一方、J.Cui ら [2]、今井ら [3]、K.Hoshino ら [4] は、あらかじめ CG 画像を作成しておくことで手指形状の輪郭線特徴量をデータベースマッチングする手法を考案した。しかし、これらの手法は前腕回旋動作への対応について明確に述べられておらず、手指を自由に動かした状態でのユーザビリティインタフェースに適した高速性が実現できるかどうかは不明である。

また、谷本らはデータベースの自己組織化において、自己増幅と自己消滅のアルゴリズムを付加して、高速な手指形状推定を達成した [5]。さらに、データベースの内部構造を自己組織化させることで、データ検索時間を大幅に短縮した [6]。データベース量を増やすことにより手の回転時の形状推定を実現したが、回転自体の安定した精度と速度は保証されておらず、定量的な評価も行われていない。他の非線形マップを利用できる可能性も残されている [7][8]。

単に輪郭線特徴のデータ量を増やすだけでは、手の表裏情報が含まれておらず、誤って推定される可能性がある。輪郭線のみデータの場、表裏逆の左右の手指形状を同一と推定

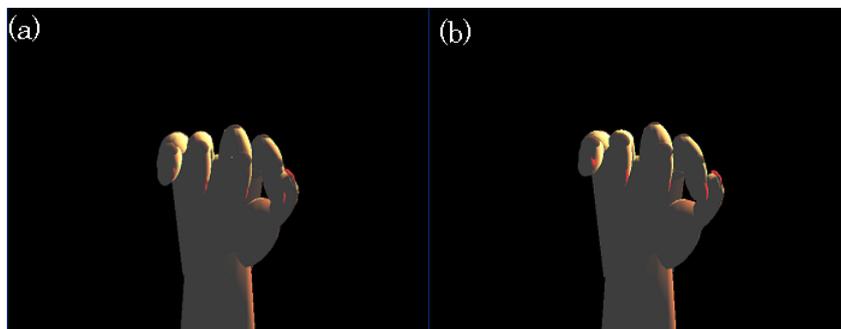


図 1.2: 片手の各 PIP 関節角度が 10 度違う手指 CG モデルの例.

する可能性が高い。また、このときの輪郭線は手指の外郭を示しており、手指が互いに遮蔽したり、手の甲により遮蔽が発生する手指の自己遮蔽が起きる。自己遮蔽時の精度も低い。前腕回旋動作を含んだ手指形状推定には、データベース量の増加による処理速度の低下問題と前腕回旋動作時の手指表裏誤認識と自己遮蔽による推定精度低下の問題がある。この問題を解決するためには、手指の表裏情報と指先の位置情報が必要と考えられる。

そこで本研究では、第一に、ヒトが表裏を含んだ手指の形状を理解する場合に用いる視覚的手掛かりを、手の皺、爪、皮膚の色を用いて心理知覚実験により検討した。第二に、その知見に基づき、手指の輪郭線の特徴量に爪の位置情報を追加することにより、指先の位置および表裏の情報も併用する手指形状推定システムを提案した。

第2章 心理知覚実験による検討

2.1 目的および方法

本章では，ヒトが手の表裏情報を含んだ手指形状を理解する際の視覚的特徴を明らかにすることにより，カメラで手指形状推定を行う際に最適な要素を検討する．通常の手の画像上に，手を構成する視覚的手掛かりの内の一つを表裏逆にした画像を提示し，表裏逆と認識した場合には，関与度が高いとした．また，視覚的手掛かりを強調した実験も行った．

図 2.1 に手の要素を表裏逆にした画像の一部を示す．ただし，画像の加工は，各視覚的手掛かり周辺を元の画像に馴染ませる程度にしか行っていない．

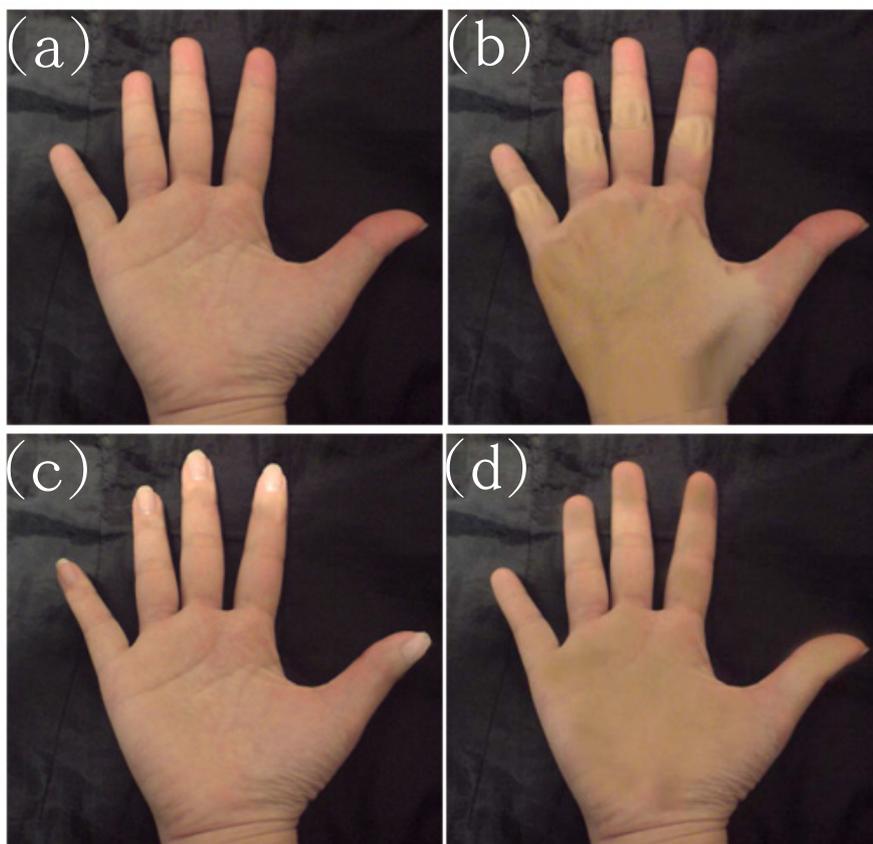


図 2.1: 手の要素を表裏逆にした画像. (a) 通常の手. (b) 皺. (c) 爪. (d) 皮膚の色.

初めに、任意の物体画像を様々な時間間隔で提示し、形状がその物体を理解できる間隔測定し、時間間隔を各被験者ごとに設定した。次に、手の要素を表裏逆にした画像(図 2.1)を通常の画像とランダムに提示し、被験者自身の手や言葉により詳しい手指形状の説明を行ってもらった。ただし、被験者には実験目的を説明せず、手以外に気づいた点も述べてもらった。また、被験者が実験途中に実験目的や趣旨を察してしまった場合は実験を中止した。

必要最低限の手の要素を計測するため、物体形状を理解する最小の時間で画像を提示し、被験者自身の手や言葉により手指形状の説明を行ってもらった。

2.2 結果と考察

18 歳から 23 歳の学生 48 人に実験を行った結果を図 2.2 に示す。横軸の濃さは、視覚的手掛かりの強度である。手の要素の一つを表裏逆にした画像を提示し、表裏逆の要素を手掛かりとして、手指の形状を理解した場合には、その要素の得点を 1 ポイント増加し、この得点を全回答数で除した値を関与度とした。元の画像の手指形状ではなく、表裏逆にした要素によった手指形状と理解した場合には関与度 1 ポイントとする。表裏逆と理解した場合でも元画像との手指形状と関連性が無い場合は、関与していないとした。全ての回答数に対する関与の割合の平均は、爪が 25%、手の皺が 20%、皮膚の色が 15%であった。すなわち、爪、皺、皮膚の色の順で関与度が高いことがわかった。

また、図 2.2 より皺と皮膚の色は関与度が全体的に低く、爪の関与度は強度に依存することがわかる。よって、爪と認識できる状態であれば、爪の関与度は高いと言える。したがって、爪の情報は、表裏を含んだ手指形状理解において重要である。これは、爪は他要素に比べ個人差が少なく、指先の表裏と位置情報を確実に判断できるためだと考えられる。

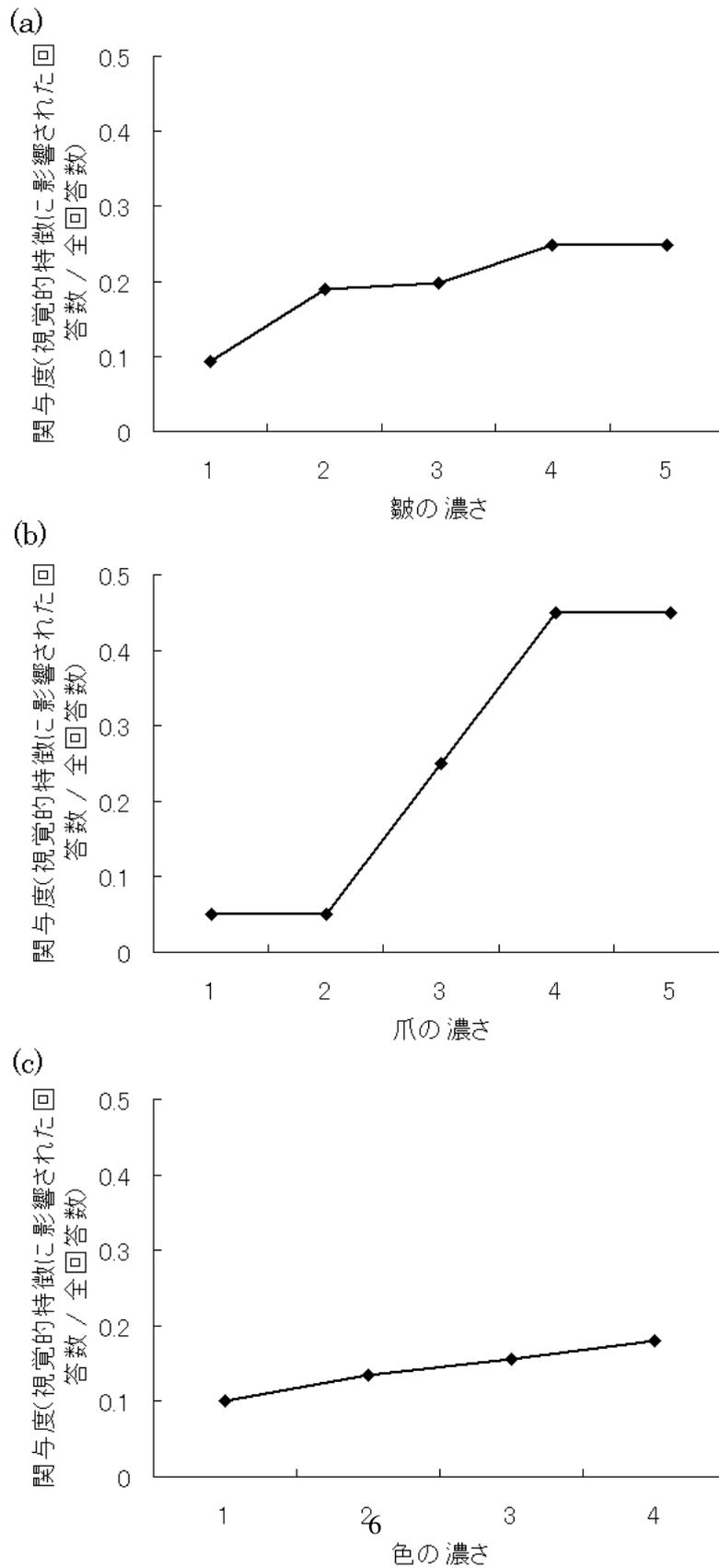


図 2.2: 視覚的手掛かりとその強度ごとの関与度.(a) 皺. (b) 爪. (c) 皮膚の色.

第3章 システムの構成

3.1 目的

前章で得られた知見に基づき、手指形状推定システムを構築する。システム全体の処理はデータベース作成と実際的手指形状の推定に分かれる。データベースは手指の輪郭特徴量、爪位置と関節角度の3つをデータセットにして作成される。そして、手指形状推定を行うときには、手指の輪郭特徴量と爪の位置情報を基にデータベース中で類似した手指形状をもつデータセットを見つけ出し、そのデータセットが持つ関節角度データを推定結果とする。

手指を撮影するカメラは、PointGray社製のカラー高速カメラを使用した。解像度は320x240画素、画像取得測度は30fpsで撮像した。データベース作成時には、Virtual Technologies社製のデータグローブCyberGloveを用いて手指関節角度を取得し、カラーリストバンド[13]を用いて前腕回旋角度の取得を行った。

3.2 データベース作成

3.2.1 処理の流れ

データベース作成処理の大まかな流れを図3.1に示す。はじめに、手指の画像を取得する。腕の位置を固定した状態で、カメラの前で自由に手指を動かして、データ取得を行う。手指画像の取得時は、照明環境を特定せず、カメラから1m程度離れた位置の手指を撮影する。このときの取得画像の解像度は320x240である。手には関節角度取得のためにデータグローブを装着し、前腕回旋角度を測定するためにカラーリストバンド[13]を装着する。データグローブは、Virtual Technologies社製のデータグローブCyberGloveを装着する。このとき、複雑なテクスチャを含んだデータグローブが手指全体を覆ってしまうため、白い手袋と赤い付け爪を装着した。また、付け爪は、装着後に大きさを切りそろえ、余計な装飾はとりはずしている。

次に、肌領域と爪領域を抽出するために、取得画像の表色系を個人別表色系[14]に変換する。その後、個人別表色系を基に肌と爪の領域を抽出し、手指の輪郭線と爪の位置を検出する。そして、輪郭線情報は特徴量化を行い、数値に変換する。このとき、手指に装着したデータグローブより手指の関節角度を取得し、同時に前腕回旋角度も取得する。カラーリストバンドによる前腕回旋角度の検出方法は付録Aに詳細を記載する。手指の輪郭線特徴量データ、爪の位置データ、関節角度データの3つが揃うと、それを一つのデータセットとして保存する。データセットが十分な数に達すると、画像取得を終了する。最後に、関節角度によってデータをクラスタリングし、データベースを構成する。

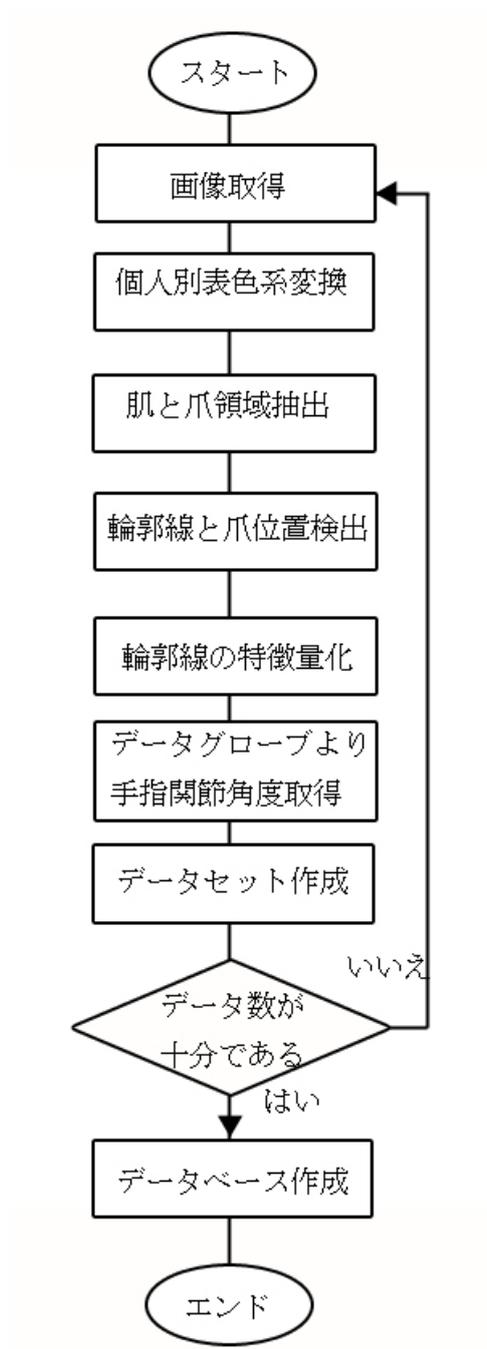


図 3.1: データベース作成のフローチャート.

3.2.2 個人別表色系

身体動作による入力インタフェースでは、人肌の色を高速にかつ高精度に出す必要がある。ところが、肌色抽出の主流となっている個人データベース探索や肌色モデルによる抽出 [16][17] は、計算量が多く実時間処理には向かない。また、照明の微弱な変化に対応できず、抽出率が 10%以下になる事もある。照明環境変化に対応するためには、個人と各環境ごとのデータを照合する処理時間が必要となってくるので、高速処理の弊害となり、これらの手法は最適とはいえない。

一方、予めデータを準備せずに肌色抽出する手法も提案されている。S.Matsushashi ら [18] は背景からの肌色分離に優れた HQV 表色系を提案した。しかし、RGB からの表色系変換時と閾値設定時に、取得画像全体の平均値と分散値を求めなければならず、計算量が多いことが難点である。また照明変化に対する定量的な評価実験も行なわれていない。

対照的に、川戸ら [19] は肌色情報そのものを使わずに抽出する手法を提案している。同手法では、まず取得画像をブロックに分割し、ブロック単位の赤み成分の密度で肌色判定する。同時に、赤み成分の抽出量に従って、より細かいブロック分割を行っている。しかし、ブロックごとの色の空間密度を評価するので、手指などの細い部分の抽出は困難とされている。また、これらの手法は、照明光の色分布変化について考慮されていない。肌色は照明光が変化すると、常に赤み成分が特徴的に残存するとは限らない上に、画面全体の色も変わってしまう。これらは高精度な抽出を困難とさせる原因となる。

現在のところ、表色系を用いた肌色領域の抽出手法の中で、色差空間上の肌色分布の楕円体内外判定による肌色領域の抽出 [20],[21] は、最も高精度な抽出手法の一つである。しかし、この手法は色差空間 X,Y と楕円体との関連性が無いために、1 画素ごとに楕円体内外判定を行う計算量の多い関数型で抽出しなければならず、実時間処理には向かない。

楕円体の内外判定以外の方法では、色差空間 X,Y 上の閾値のみでの抽出が最も高速な抽出手法である。しかし、色差空間上の肌色ではない部分を多く含むため抽出精度が低下する。

予め表色系が楕円体の傾き情報を含んでいるならば、閾値のみでの抽出で楕円体の内外判定と殆ど同じ領域を切り出すことが出来るため、楕円体の内外判定による抽出法と同じ精度が得られる。さらに、色差空間上ならば、肌色分布は照明の色に応じて色差空間上の任意の位置に写像されるので、その特徴を利用し照明環境変化に対応可能だと考えられる。

そこで、本研究では、個人の肌色分布の楕円体の長軸に極力一致した個人別の表色系を作成することで、閾値のみで高精度な抽出を可能とし、背景変化と特に照明環境変化に対応する手法を提案する。具体的には、予め求めておいた色差空間上の各個人の肌色分布図の重心から、肌色軸とそれに直交する肌色直交軸、つまり個人別の表色系を求め、取得画像から照明環境に応じた閾値設定を行い、肌色抽出を行う。

3.2.3 原理

本節では、肌色の分布特性と個人の表色系を作成する原理の説明を行う。肌色の分布は、ある一定色の濃淡の広がりであるために、主に色差空間上の原点近くから、個人ごとの赤みの強い点までの分布となっている。逆に、XY軸の色差空間上ならば、方向への広がりが少ないという特徴がある。つまり、原点方向からの長軸をもった楕円状となる場合が多い。これを任意の表色系で抽出した場合、閾値設定を正に行っても、ある程度の肌色類似した非肌色領域も拾ってしまう。個人の肌色領域の広がりが表色系の軸と45度ずれている時に領域の切出しを行った様子を図3.2-(a)に表す。このとき、分布の楕円は切り出し領域の四角形の対角線上にあるので、非肌色領域が最も多くなる。特に、白熱灯などの赤みがかかった照明環境の場合、取得画面上の非肌色領域の割合は大きくなる。しかし、図3.2-(b)に示すように楕円の中心方向の軸Aとその直交軸Bで閾値を設定した場合、分布の傾きと軸の傾きが近似なため、非肌色領域が極力少ない状態で抽出できる。また、照明が変化した場合も肌色分布上のすべての色が照明光によって同じく写像されるため、傾きは変わらないので、抽出の時点で非肌色領域を低減することができる。従って、原点と肌色分布の中心を結ぶ個人の肌色軸とそれに直交する軸で構成される個人の表色系を基準にして、抽出を行うのが効率的と言える。

背景が茶色がかかった乳白色で照明光が白熱灯のときの取得画面上の肌色を、値の最小最大値による閾値設定のみで抽出した。表色系ごとの肌色抽出画素数を表3.1に示す。ただし、取得画面上の肌色部分の画素は手動で判別をしたものである。左端の欄に示すように、個人の表色系で抽出を行った画素数が他表色系より少ないことから、非肌色領域を拾いにくいことがわかる。

表 3.1: 閾値のみでの抽出結果

肌色部分	HSV	Luv	YIQ	個人表色系
435	647	697	599	498

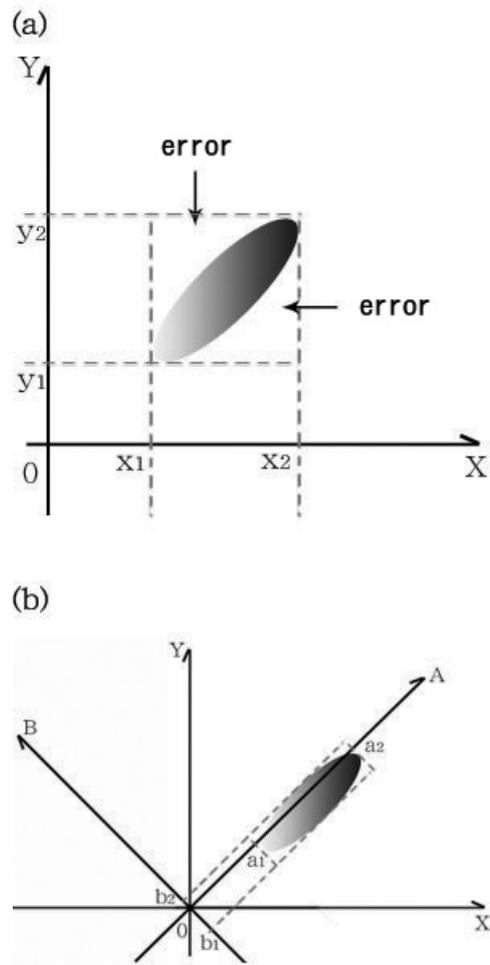


図 3.2: 色差空間上の肌色領域切り出しの比較.(a) 通常の例.(b) 個人の表色系の例.

3.2.4 個人別表色系の作成

動作環境として，機材はCPUがPentiumIV，2.8GHz，主記憶512MBのPCと，PointGray社製のRGB入力高速カメラを使用した．解像度は320x240画素，画像取得速度は150fpsで撮像した．カメラから被験者手までの距離は1から2mであり，肘から指先までの間を抽出対象とした．個人の表色系作成の流れを以下に示す．

- 処理1：取得画像の平滑化処理．
- 2：輝度と色差空間の分離．
- 3：肌色データ取得．
- 4：個人の肌色分布図を作成．
- 5：個人の表色系を作成．

処理1では，カメラからの取得画面上の細かなノイズを除去するために，平滑化を行った．

処理2では，肌色分布図を取得する前に，照明の強さの影響を省くためYIQ表色系を利用して，RGB入力のカメラ画像を輝度(Luminance)と色差や色度(Chrominance)に分離した．次式にRGBからYIQ表色系への変換式を示す．

$$\begin{aligned} R' &= \begin{cases} 4.5 \times R & R < 0.018 \\ 1.099 \times (R^{0.45}) - 0.099 & R \geq 0.018 \end{cases} \\ G' &= \begin{cases} 4.5 \times G & G < 0.018 \\ 1.099 \times (G^{0.45}) - 0.099 & G \geq 0.018 \end{cases} \\ B' &= \begin{cases} 4.5 \times B & B < 0.018 \\ 1.099 \times (B^{0.45}) - 0.099 & B \geq 0.018 \end{cases} \end{aligned} \quad (3.1)$$

$$\begin{aligned} Y &= 0.299 \times R' + 0.587 \times G' + 0.144 \times B' \\ I &= 0.596 \times R' - 0.274 \times G' - 0.322 \times B' \\ Q &= 0.211 \times R' - 0.522 \times G' + 0.311 \times B' \end{aligned} \quad (3.2)$$

ただし，式中のRとGとBはRGB表色系のそれぞれの値であり，YとIとQはYIQ表色系のそれぞれの値である．YIQ表色系は北アメリカのテレビ放送に使用され，肌色表現に優れている表色系である．肌色が分布すると思われる領域を，補正により伸張して粒度が高く表現できるため分離に使用した．しかし，厳密には，後に個人の表色系を作成するので，輝度と色差空間の分離が行えるのならば，YUV(YCrCb)などの表色系による色差空間でも良い．

処理3では，100フレーム分の肌色データを取得する際には，取得画像の中心部に掌より小さい枠を作り，1秒弱ほど枠内に手をおくことによりデータ取得を行った．

処理4では，個人の表色系の肌色軸の傾きを肌色モデルの重心と原点を結ぶ線から求めた．

なお，重心に対する重みは色の出現回数 m とした．重心点 (Q', I') を求める式を以下に示す．

$$\begin{aligned} Q' &= \frac{\sum m_n^q \times q_n}{M} \\ I' &= \frac{\sum m_n^i \times i_n}{M} \end{aligned} \quad (3.3)$$

ただし， m_n^q は， Q 軸上の位置 n の色の出現回数， m_n^i は I 軸の位置 n の色の出現回数， q_n は Q 軸上の位置， i_n は I 軸上の位置， M は取得画像の画素数を示している．式 (3.3) より求めた重心点を単位ベクトル化し， $\overrightarrow{\alpha(Q', I')}$ とするとき，式 (2) の I を I' だけ傾きをずらすことにより，肌色軸 S_a の軸成分 S'_a を求める．次式に S'_a を示す．

$$\begin{aligned} S'_a &= (0.596 \times I' + 0.211 \times Q') \\ &+ (-0.274 \times I' - 0.522 \times Q') \\ &+ (-0.322 \times I' + 0.311 \times Q') \end{aligned} \quad (3.4)$$

同様に肌色直交軸 S_b の軸成分 S'_b も単位ベクトル， $\overrightarrow{\alpha(Q', I')}$ に直交する単位ベクトルから計算することができる．図 3.3 に個人の表色系に変換後の肌色分布の一例を示す．照明が環境した場合も，肌色軸に沿って分布領域が変化しているのが確認できる．

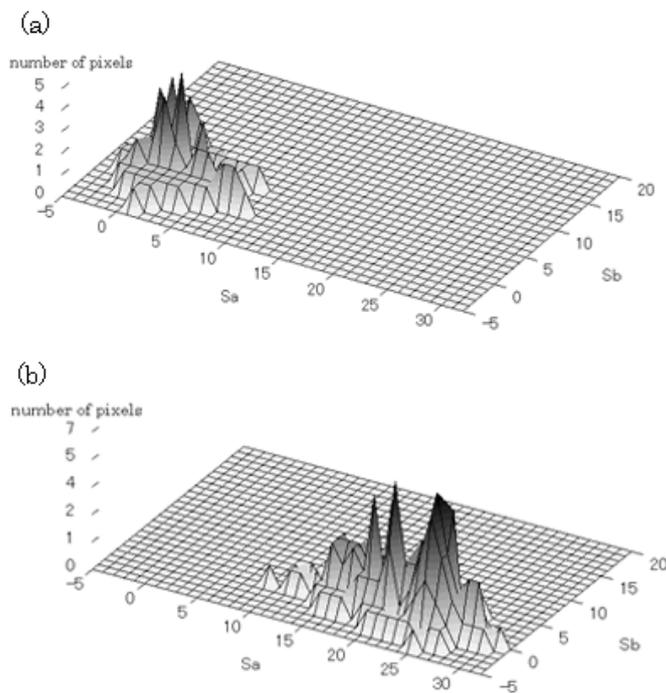


図 3.3: 個人別表色系での肌色分布.(a) 蛍光灯環境下.(b) 白熱灯環境下.

3.2.5 肌色領域の抽出

個人の表色系を作成した後、実時間での肌色抽出を行う。図 3.4 に処理の流れ図を示す。細かいノイズを除去するために、平滑化処理を行った。さらに、身体動作による入力インタフェースでは画面上にある程度の面積の身体部位が映っていることが前提にあるため、ラベリング処理を行い、面積値が極端に低いものに関しては除外した。肌色抽出を行う際は、処理中に照明環境が変化してしまうことも考えられるので、一定時間ごとに閾値の再設定を行った。本

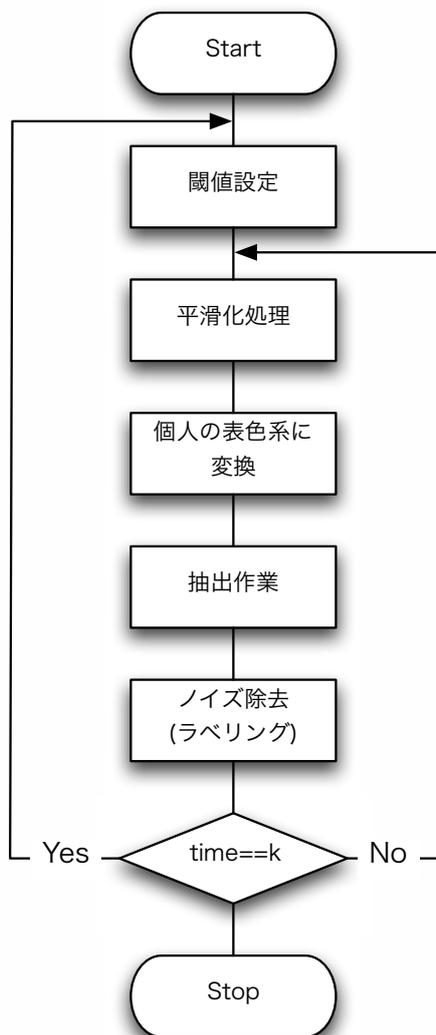


図 3.4: 肌色抽出処理の流れ

研究では、次式によって閾値を求めた。

$$Threshold = \overline{S_0} + (\overline{G_n} - \overline{G_0}) \times h \pm k \quad (3.5)$$

ただし、 $\overline{S_0}$ は肌色データ取得時の肌色の平均値、 $\overline{G_0}$ は肌色データ取得時の画像全体の色の平均値である。 $\overline{G_n}$ は現在の取得画像の全体の色平均値、 h は、照明環境が閾値変化にもたらす割合であり、 k は閾値の最大値と最小値の幅を表す定数である。

また、この閾値決定手法は、照明環境が一定の割合で肌色分布が、照明の色分だけ写像される性質を利用して、逐次閾値を更新し環境変化に対応するための手法である。ただし、取得画像中の肌色領域が画像中の1割以下や9割以上を占めるなど、極端に増減した場合には対応できない。

3.2.6 爪領域の抽出

個人別表色系の肌色直行軸 S_b を使用して、手領域抽出後に爪領域の抽出を行った。その様子を図 3.5 に示す。図の淡色が肌色、濃色が爪領域を示しており、爪領域を抽出可能であることがわかる。画像から人の肌色領域を抽出した際に、さらにその領域内から、関節角度などの特定要素を抽出することにより身体動作計測を行う可能性がある。身体の特異要素では皮膚の厚さが変化している場合があり、それによって特異要素の色を抽出することが可能である。皮膚下の構造がある程度同一の場合、皮膚の厚さによって光の反射量が変化する。人の肌色に似た色は、反射光の他にも血流など皮膚下の身体的構造に左右されるので、個人別表色系の肌色軸 S_a による特定部位の抽出は困難である。逆に、肌色とはかけ離れた色は左右されないといえる。つまり、肌色直行軸 S_b は身体の特異要素を抽出することにも、特に有効であると考えられる。

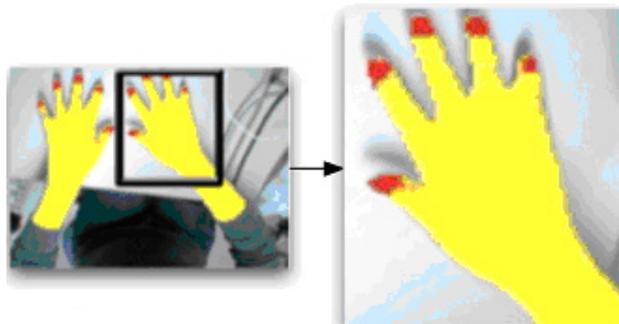


図 3.5: 爪領域の抽出.

3.2.7 輪郭線の特徴量化

本研究では，Otsu ら [15] が提案している高次局所自己相関関数を，輪郭の特徴量の抽出に用いた．高次局所自己相関関数特徴は画像の認識や計測のために有用な特徴であり，次式で定義される高次相関関数について，参照点とその近傍についての計算をしたものである．

$$X^N(a_1, \dots, a_N) = \int f(r + a_1) \dots f(r + a_N) dr \quad (3.6)$$

ここで， X^N は N 次元での r 点近傍の相関関数である．一般に自然画像を処理対象とする場合，対象点まわりの画素が重要なため，係数 N を 2 次までとした．また，平行移動による等価なものを除いて，図 3.6 で示す 25 の特徴量で表現できる．ただし，No.1 から No.5 までは，No.6 以降の特徴量に比べスケールが小さいので，スケール差を合わせる必要がある．No.2 から No.5 までは参照点の画素をさらに掛け合わせ，No.1 は参照点の画素の 2 乗を掛けることで他の特徴量に一致させた．手指の画像は，全体の 320x240 画素からさらに手指部分だけ切り出し，輪郭抽出を行った後に 18x21 画素に縮小し，縦 6，横 7 ブロック分割を行い，そのすべてに特徴量の次元低減を行った．

3.2.8 爪の位置データ取得

手指の輪郭特徴量だけでは，手の表裏情報が含まれておらず誤って推定される可能性がある．図 3.7 にその例を示す．同図の場合，まったく違う手の形状にもかかわらず，手指の輪郭線特徴量は同等となってしまう．しかし，爪の位置情報を持った場合，同図は，まったく違うデータとして認識される．また，手内部にある指の輪郭線を取得した場合は，自己遮蔽問題が解決されるが，手指の表裏を認識することはできない．爪の位置情報を持った場合は，手指の表裏を認識するとともに，指先位置がわかることから自己遮蔽問題も解決可能である．

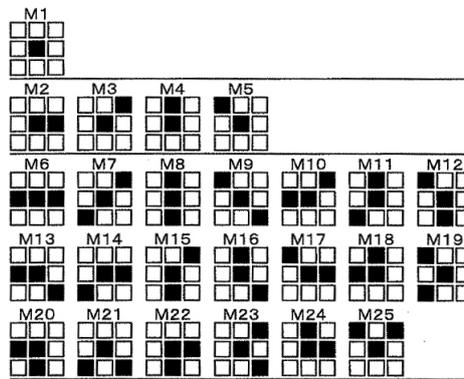


図 3.6: 高次局所自己相関の局所パターン.

図 3.8 にその例を示す．同図-(a) の場合，左右の手があるのか，左手が二つあるのかは判断できない．また，詳しい指先の位置情報も認識できない．しかし，爪の位置情報を加えた同図-(b) では，左右の手があると判断でき，指先位置を認識することが可能である．

本研究では，手指の領域内から爪の領域を抽出し，ラベリング処理を行った結果，得られた爪の領域の重心点を爪の位置とした．ただし，どの指の爪であるかは推定しない．なぜならば，手指の輪郭特徴量と爪の位置情報には，爪がどの指に対応しているかの情報が含まれているからである．今回は，爪の領域の抽出精度を上げるため，データベース作成時に一般的な赤い付け爪を使用した．



図 3.7: 誤って形状推定される例.

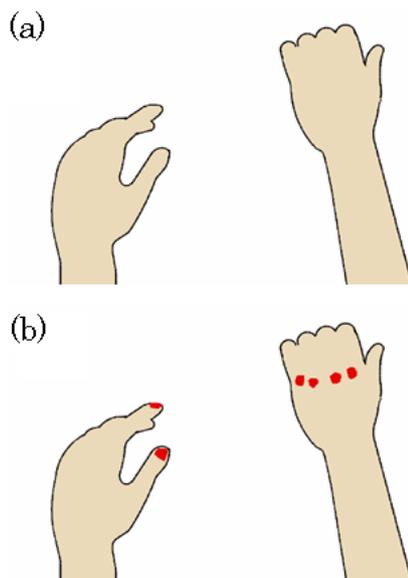


図 3.8: 爪の位置情報による手指の表裏と指先位置認識の例.(a) 爪なし.(b) 爪あり.

3.2.9 データセット作成

手指の輪郭特徴量，爪の位置，データグローブとカラーリストバンド [13] より取得した各関節角度からデータセットを作成した．図 3.9 にデータセットの作成概要図を示す．

爪の位置は，画像を 18×21 画素に縮小処理を行った後の x と y 座標で表し，爪がない場合は $(-1, -1)$ 座標とした．手指関節角度は前述のデータグローブを使用し，18 自由度の取得を行った．前腕回旋角度は，30 度ごとの 6 段階で計測を行った．

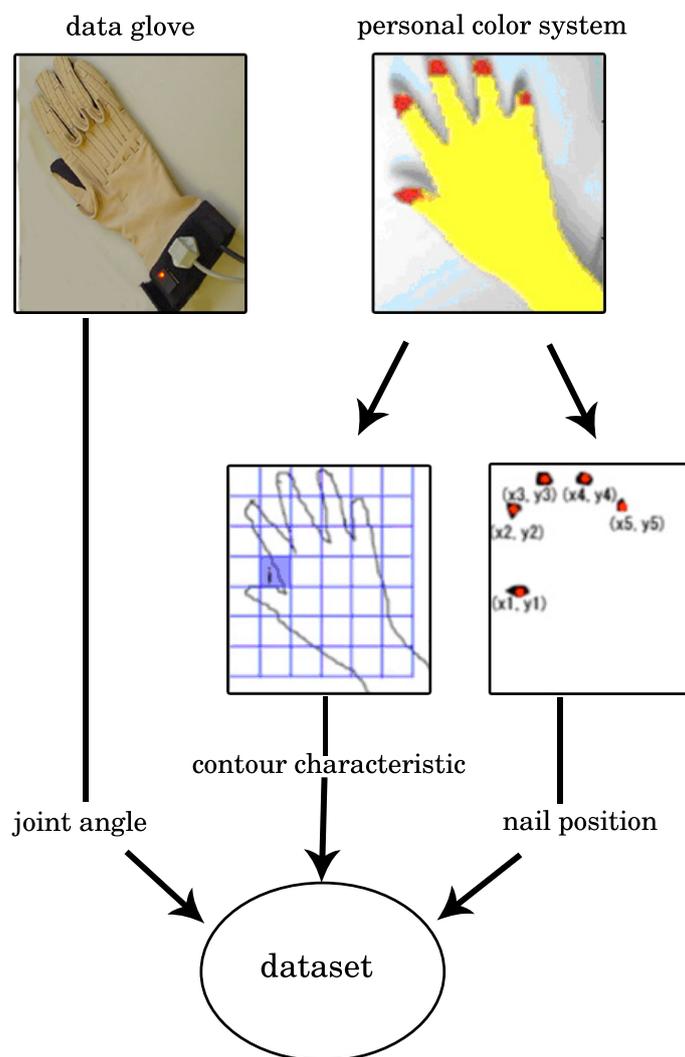


図 3.9: データセット作成概要図.

3.2.10 データベース構築

前腕回旋角度を基にして、3層以上に階層化されたデータベースを構築した。図3.10にデータベースの概略図を示す。第一階層では、前腕回旋角度を基にしている。第二階層以降では、データ収集時に、指の関節の中で最も角度の分散が少ないものを基にする。データ量に応じて、この第二階層目と同じような階層を作成する。さらに、最下位であるデータセットは、指の関節角度順に並べられている。データベース作成前には、重複したデータセットが存在しないように、手指の輪郭特徴量と爪の位置が極端に近い値をもっているデータセットを削除した。

実験時のデータベース内のデータ数は合計1万個弱程度で、第二層の一つのクラスが持つデータベースは平均して1500個であった。このときの第二層以下のデータベース階層は5階層である。なぜならば、前腕回旋が固定された状態での手指形状推定のためのデータベース階層は5階層が最も計算量が少ないと報告がなされているからである [6]。

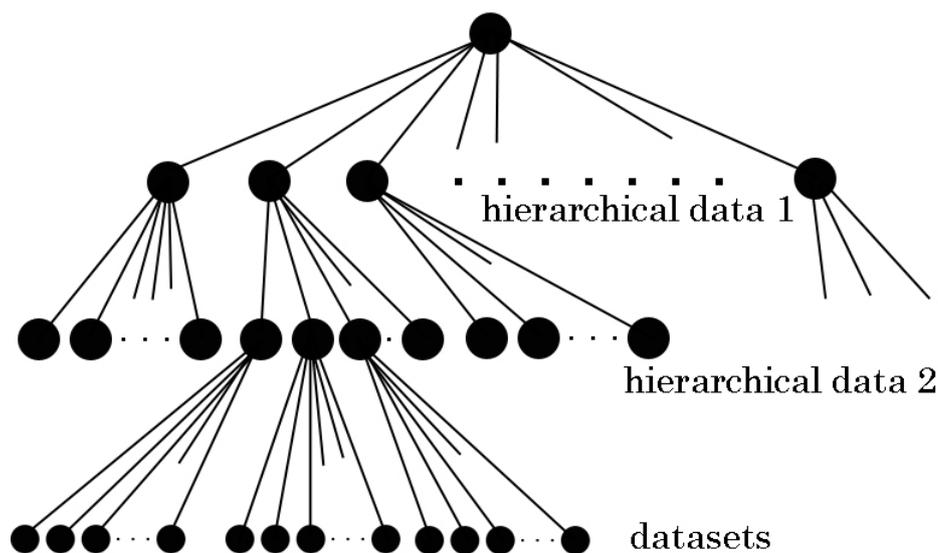


図 3.10: データベース内の階層概略図.

3.3 形状推定

3.3.1 処理の流れ

手指画像がシステムへ入力されると、爪の位置情報と輪郭線特徴量を検出し、データベース内の類似データセットを検索し、関節角度を出力することで推定を行う。手指形状推定手法の大まかな流れを図 3.11 に示す。

はじめに、手指の画像を取得する。腕の位置を固定した状態で、カメラの前で自由に手指を動かして、データ取得を行う。手指画像の取得時は、背景と照明環境を特定しない。また、このときのカメラとカメラからの手指の距離、取得画像の解像度は、データベース作成時と同じである。ただし、手には何も装着していない。

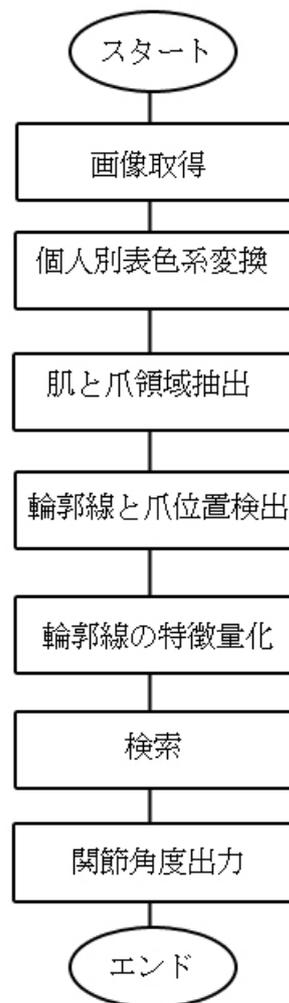


図 3.11: 推定処理のフローチャート.

次に，肌領域と爪領域の抽出を行うために，取得画像の表色系を個人別の表色系 [9] に変換する．その後，個人別表色系を基に肌と爪の領域を抽出し，輪郭線と爪の位置を検出する．そして，手指の輪郭線情報を数値化するために，特徴量化を行う．

最後に，取得画像より検出した手指の輪郭線特徴量データと爪の位置データを用いて，データベース内の検索を行う．データベース中で，最も類似した手指形状のデータセットを見つけ出し，そのデータセットが持つ関節角度データを推定結果とする．

また，形状推定の様子を視覚的に理解しやすくするため，手の CG モデルを作成し，本システムから出力された関節角度によって動作させた．その様子を図 3.12 と付録 B に記載する．

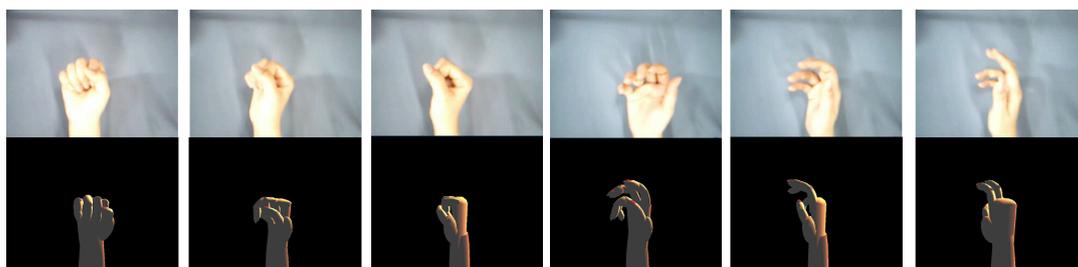


図 3.12: 形状推定の結果.

3.3.2 データ検索範囲指定

現在のデータセットからみて近傍のデータのみを検索範囲として，手指の輪郭特徴量と爪の位置から，最も類似度の高いデータセットを検索する．図 3.13 に，データベース内の検索範囲の一例を示す．

図 3.13 で白い丸が現在のデータの位置とすると，淡い丸が検索範囲である．次フレームの検索範囲は，現在のデータセットの最下位での近傍データセットと，第二階層での近傍のクラス内の同じ位置のデータセット群と，第一階層での近傍のクラス内の同じ位置の第二階層のクラス内のさらに同じ位置のデータセット群となる．

3.3.3 類似度の計算

類似度を算出する際には，単純なユークリッド距離計算をする．類似度 E_r を次式に示す．

$$E_r = \sum_{i=1}^P (x_i - x'_i)^2 + \sum_{k=1}^N (y_k - y'_k)^2 \quad (3.7)$$

ただし， P は高次局所パターン数， x_i はデータベース内のデータセットの手指輪郭の i 番目の特徴量， x'_i は取得画像の手指輪郭の位置 i 番目の特徴量， N は爪の数， y_k はデータベース内のデータセットの k 番目の爪の位置， y'_k は取得画像の k 番目の爪の位置，25 は図 3.6 で示

す局所パターンの数, 5 は爪の位置情報の数を示している. E_r を最小にするデータセットを検索し, そのデータセットが持つ手指の関節角度を出力することによって, 3次元手指形状推定を行う.

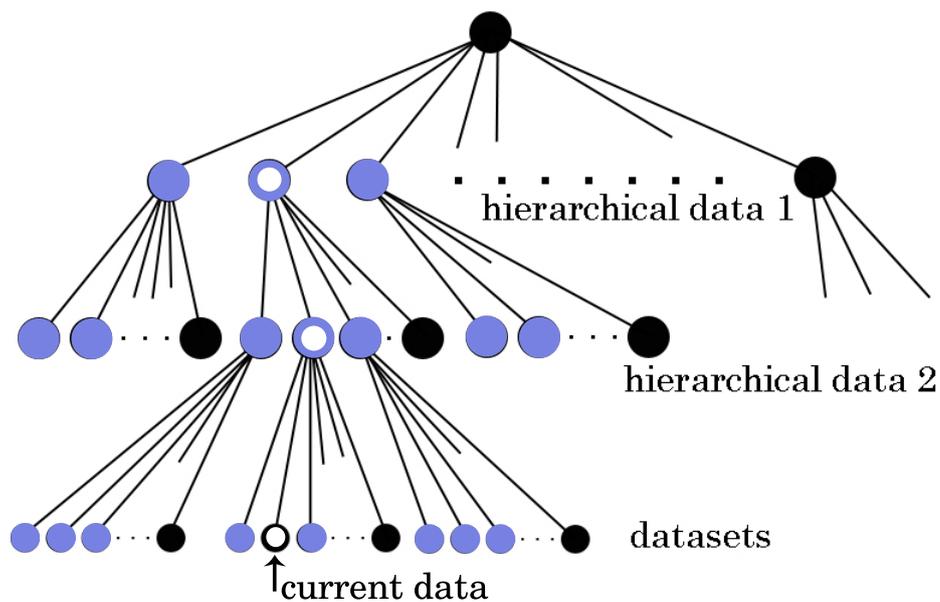


図 3.13: 検索範囲の例.

第4章 評価実験

4.1 個人別表色系

4.1.1 方法

個人の肌色表色系の処理速度と背景と照明環境別の精度を、他の表色系と肌色モデルを用いて抽出を行ったものと比較しながら評価実験を行った。また、各表色系は、RGB 表色系からの変換後の抽出処理速度を最速にした精度を比較するために、閾値のみの簡単な関数型で抽出を行っている。カメラからの入力値を RGB とし、取得動画像は各表色系と肌色モデルとも同じものを使用した。精度評価には次式を用いた。

$$\text{誤認識率 (\%)} = \text{誤認識された画素数} / \text{取得画像の画素数} \times 100 \cdots (a)$$

$$\text{肌色未検出率 (\%)} = \text{肌色部分を認識出来なかった画素数} / \text{肌色部分の画素数} \times 100 \cdots (b)$$

背景部分誤認識率 (\%) = 背景を肌色だと誤認識した画素数 / 背景部分の画素数 $\times 100 \cdots (c)$
ただし、画面上の肌色部分の正値は人が手動で示した指標を用いた。

誤認識率とは、取得画像の画素数に対する、背景と肌を誤って認識した画素数の割合である。肌色未検出率とは、肌部分の画素数に対する、肌を検出できなかった画素数の割合である。背景部分誤認識率とは、背景部分の画素数に対する、背景を肌とを誤認識した画素数の割合である。各率は、低いほど抽出精度が良いと言える。

比較する表色系は、全部で5種類である。第一は、入力値である RGB 表色系。第二は、肌色抽出に最もよく用いられている HSV 表色系の HS 空間。第三は、Luv 表色系の uv 空間。第四は、次式の変換式で示される HQV 表色系の HQ 空間。第五は、個人別表色系 YSaSb の SaSb(Skin-a, Skin-b) 空間による閾値での抽出。また、HS 空間上の特定環境で作成された肌色モデルも比較に加える。

$$\begin{aligned} H &= \text{Arccos} \left(\frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \right) \\ Q &= \sqrt{\left(\frac{(R - G) + (R - B)}{2} \right)^2 + \left(\frac{\sqrt{3}(G - B)}{2} \right)^2} \\ V &= \max(R, G, B) \end{aligned} \quad (4.1)$$

ただし、R, G, B は RGB 表色系のそれぞれの値で、H, Q, V は HQV 表色系のそれぞれの

値を示す。HQV 表色系は、Hue(H) と chroma(Q) と Value(V) の 3 要素で構成され、Q の分散値は人間の肌の色を背景画像から分離するのに良い性質を持っているという報告がなされている [18]。

様々な人種の男女 18 名の腕から手の部位の画像を実験に用いた。被験者の内訳は、中東系人 4 人、東南アジア系人 4 人、極東アジア系人 6 人、白人 4 人である。撮影環境は、背景が 6 種類で照明が 5 種類の計 11 種類である。一つの環境と一人の被験者につき 2 枚の画像を撮影したので、実験に用いた画像の合計は 396 枚である。

比較対象となる全ての表色系および肌色モデルによる肌色抽出手法において、平滑化処理とラベリングによるノイズ除去処理を図 3 に示した順で行った。このときのラベリングによるノイズ除去処理は特定の手法にチューニングされたものではなく、主に平滑化処理でも取り除けなかったごま塩ノイズを除去するために採用した。

また、表色系による肌色領域抽出の際に用いる閾値は、式 (2.5) により自動的に設定した。評価実験では、簡単な関数型、すなわち各表色系上の肌色分布を四角形で囲むことにより肌領域を抽出することが目的である。なぜなら、本研究では各表色系の変換後の処理速度を最速にした場合の抽出精度を比較実験するからである。

4.1.2 速度

速度評価実験では、高速カメラから動画画像を取得し、肌色抽出を行い、ディスプレイに描画するまでを計測した。図 4.1 に評価実験結果を示す。

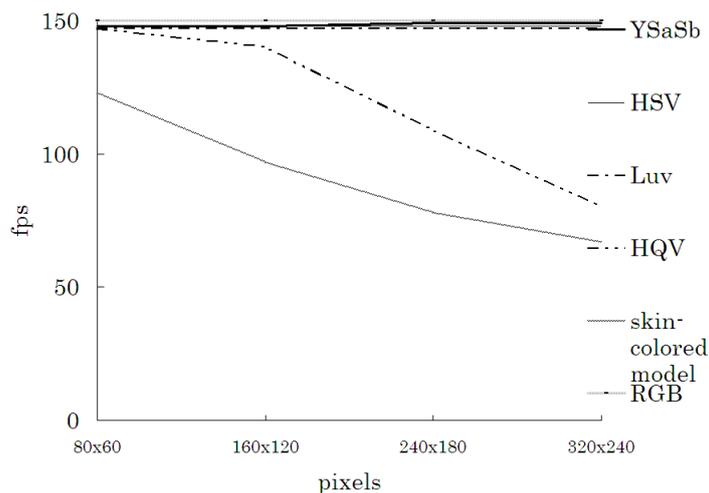


図 4.1: 速度評価実験

インタフェースとして、肌色抽出をする場合には、高速な処理が必要である。特に、指先入力の処理には大きな計算量がかかるため、抽出時は次の処理に殆ど負担がかからない程度、つまり最低でも 100fps 以上の速度が必要だと思われる。実験結果では、RGB、HSV、YSaSb、Luv の順に並び、4 つの表色系が基準を満たす事がわかった。

4.1.3 背景別

次に、背景環境別の精度実験を行った。結果を図 5 に示す。ただし、背景によって肌色が変わることは殆ど無いため、肌色未検出率と背景部分誤認識率は記載しない。背景が黒の時に一度閾値設定し、閾値を変えないまま、他背景で抽出実験を行った。照明環境は蛍光灯で、取得画像の解像度は 320x240 画素とした。縦軸は画面全体からの誤認識率なので、数値が低いほど精度が良いと言える。手の画面上の割合はおよそ 3 割であり、特に指先のみは割合は 1 割に満たない。つまり、インタフェース操作で必要となる指先を認識するためには、誤認識率は悪くとも 1 割以下でなければならない。

今回の実験から、個人別表色系 YSaSb のみが基準をみたすことがわかった。実験に使用したサンプル画像すべてで YSaSb 表色系での抽出精度の分散は殆どなかった [14]。また、どの表色系も赤の背景の時には誤認識率があがっている。これは、肌色自体が赤に近い色であるためと予想できる。つまり、図 3.2-(a) で示した様な非肌色領域に赤が多いと考えられる。

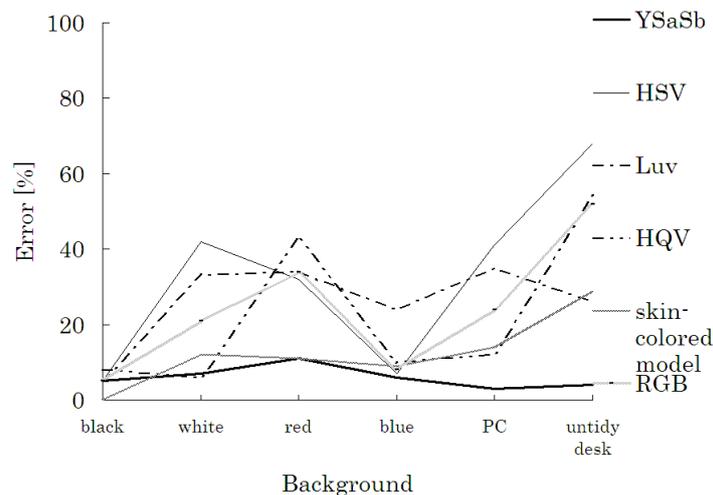


図 4.2: 背景別誤認識率

4.1.4 照明別

照明環境別の抽出精度実験を行った。背景は特定せず、雑多に様々な生活用品が写り込んだ環境とした。図 4.3 に照明環境別の評価実験結果を示す。

図 4.3-(a) は式 (a) で求めた抽出精度結果である。YSaSb 以外の抽出方法は、白熱灯と青のライトの精度が悪い。特に RGB は輝度も含まれているので、屋内太陽光と蛍光灯の照明環境時以外では全て精度が落ちる。図 4.3-(b) は式 (b) で求めた肌色未検出率である。蛍光灯の環境下で閾値調整をしているので、青ライトの環境下以外では、激しい肌色未抽出は起こっていない。ただし、青のライトが照射されたとき肌の色も青みがかってしまうので、赤み成分を抽出する HQV の精度が極端に落ちている。図 4.3-(c) は式 (c) で求めた背景領域誤認識率である。このグラフから、他の抽出方法では白熱灯や青ライトなどの色付きの光環境の時に精度が落ちる事が読み取れる。色付きの光環境の場合、背景を肌色と誤認識率してしまう割合がとても高い。閾値設定を行った場合でも、2.2.3 節での仮定の通り、非肌色領域を多く抽出してしまったからだと思われる。しかしながら、YSaSb 表色系がすべての環境において誤認識率 9%未満の精度を保ちながら肌色抽出できることを確認できた。また、実験に使用したサンプル画像すべてで YSaSb 表色系での抽出精度の分散は殆どなかった。

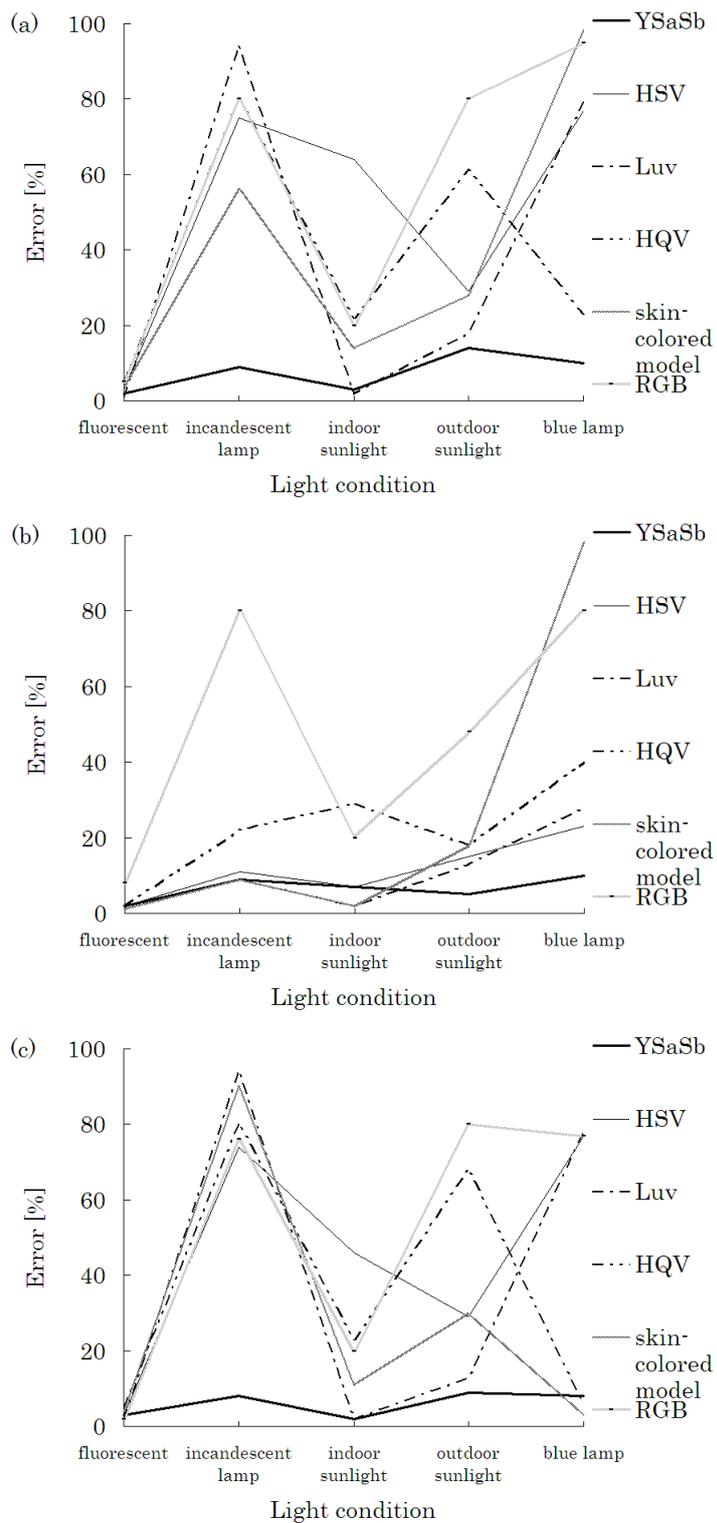


図 4.3: 照明別誤認識率.(a) 誤認識率.(b) 肌色未検出率.(c) 背景部分誤認識率.

4.1.5 肌色分布の切り出し精度についての楕円体モデルとの比較

本節では、個人表色系による抽出手法が、図 3.2-(a) に示す任意の表色系の色差空間 X, Y に比べて楕円体モデルの内外判定による抽出手法に近い精度であることを検証する。図 4.4 に楕円体モデルの内外判定と個人別表色系の色差空間上の切り出しの例を示す。図 4.4(a), (b) とも、太い枠線が切り出しの際に用いる枠である。楕円体モデルの内外判定による抽出手法は、色差空間上 X, Y に肌色分布に近似した傾斜楕円を楕円体モデルとし、色の切り出しを行っている。個人別表色系による抽出手法は、予め傾斜楕円に沿った S_a 軸とそれに直行する S_b 軸を作成し、 S_a と S_b の閾値により矩形枠で、色の切り出しを行っている。

楕円体モデルの内外判定を基準としたときの、任意の表色系の色差空間 X, Y の閾値による切り出しと個人別表色系による切り出しの非肌色領域の割合を算出すると、以下の通りになる。仮に、肌色分布に近似した傾斜楕円の短軸の半径を a 、長軸の半径を $3a$ 、楕円の傾きを $\frac{\pi}{4}$ とすると、楕円体の面積は $3\pi a^2 \doteq 9.425a^2$ 、個人別表色系の $S_a S_b$ 軸による切り出しの面積は $12a^2$ 、色差空間 X, Y の閾値による切り出しの面積は $18a^2$ である。楕円体モデルの内外判定による切り出しに非肌色領域が無いとしたとき、個人別表色系による切り出しの非肌色領域の割合は、27%となる。それに対して、色差空間 X, Y による切り出しでは91%となる。つまり、色差空間による切り出しが圧倒的に抽出精度が低下するのに対して、個人別表色系による抽出手法は、楕円体モデルの内外判定による抽出手法に近い精度が期待できる。

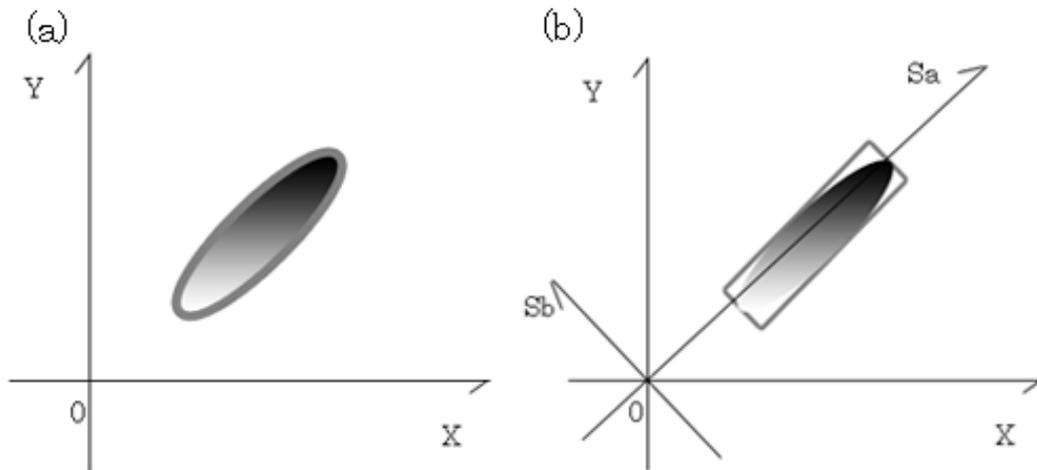


図 4.4: 色差空間上の肌色領域の切り出しの比較。(a) 楕円体モデルの内外判定の例。(b) 個人別表色系の例。

4.1.6 計算量についての楕円体モデルとの比較

楕円体モデルの内外判定による抽出手法よりも個人別表色系による抽出手法の計算量が格段に少ないことを示す．傾斜した楕円体モデルの内外判定による肌色抽出を行った場合の1画素に行う計算式は，以下の通りである．

$$\begin{aligned}x' &= x \cos \theta - y \sin \theta \\y' &= x \sin \theta + y \cos \theta\end{aligned}\quad (4.2)$$

$$\frac{x'^2}{A^2} + \frac{y'^2}{B^2} < 1 \quad (4.3)$$

ただし， x と y は，任意の表色系の色差空間 X, Y 上の点 (x, y) ， A と B と C は楕円体の形状に応じた定数である．楕円体は θ だけ回転しているが， θ の値は定数であるので， $\cos \theta$ と $\sin \theta$ も定数として扱うことが可能である．

次に，個人表色系を予め作成し，閾値による肌色抽出を行った場合の1画素に行う計算式は，以下の通りである．

$$\begin{aligned}A_i &< i < B_i \\A_q &< q < B_q\end{aligned}\quad (4.4)$$

ただし， i と q は，個人別表色系の色差空間 S_a, S_b 上の点 (i, q) ， A_i と B_i と A_q と B_q は肌色分布の大きさに応じた定数である．

楕円体の内外判定に用いる式 (3.2) と式 (3.3) の四則演算の数を数えてみる．このとき，不等式は加減算と処理時間がほぼ等しいので，加減算として換算した．すると，楕円体モデルの内外判定に用いる式 (3.2) と (3.3) の四則演算の数は，加減演算が4，乗除演算6，合計10個となる．個人別表色系での抽出に用いる式 (3.4) の四則演算の数も同じく数えてみると，加減演算が4のみの合計4個となる．つまり，個人表色系による抽出手法は，楕円体モデルの内外判定による抽出手法よりも計算量が半分以上に減少すると推測できる．すなわち，個人表色系による抽出手法は，楕円体の内外判定による抽出手法に近い精度を保ちつつ，高速処理が可能な利点を持っていると言える．

4.2 形状推定

4.2.1 実験手法

データグローブを装着した状態で手指形状推定を行い，データグローブから出力された関節角度と，本システムから出力された関節角度を比較して評価実験を行う．ただし，複雑なテクスチャを含んだデータグローブが手指全体を覆ってしまうため，白い手袋と赤い付け爪を装着した．手袋と付け爪の色については，人工色のため，どの色においても単色であることに変わらず，手袋と付け爪色によって抽出精度が変化することは無いので，任意の色とした．

被験者は先述の高速カメラの前方 1m 程度の位置に手をかざし、腕の位置を固定した状態で前腕と手指を自由に動かした。CPU が Pentium IV、クロック周波数 2.8GHz の PC で動画を 30fps で撮像し、撮像とは別スレッドで手指の形状推定を行った。また、データベース内のデータ数は 8934 個であり、その時の手指の形状推定の処理速度の平均は 100fps であった。

4.2.2 結果と考察

前腕回旋動作時に手指関節角度の推定精度実験を行った。初めに、中指第 2 関節角度推定の実験結果を図 4.5 に示す。ただし、図 4.5-(a) の縦軸は関節角度、横軸はフレーム数を示しており、図中の太い実線はデータグループによる手指関節角度の測定値、細い実線は本システムによる手指関節角度の推定値、破線はそのときの前腕回旋角度の推定値を示している。図 4.5-(b) の縦軸は実験時の中指第 2 関節角度の推定誤差、横軸はフレーム数を示している。実験時には、推定値は 10 フレーム前からの移動平均を求めることにより、平滑化を行っている。前腕回旋動作は 30 度ごとの 6 段階で計測されており、右手の掌がカメラに向いているときの前腕回旋角度を 0 度、手の甲がカメラに向いているときの前腕回旋角度を 180 度とした。

図 4.5-(a) より、前腕回旋時でも中指第 2 関節角度が高精度に推定できていることが確認できる。しかし、図 4.5-(b) より、前腕回旋角度が 0 度で固定の時に比べ、動作時は推定誤差の標準偏差が 8.40 度増加し、精度が低下する。また、前腕回旋角度が 180 度で中指第 2 関節の角度の値が高いとき、つまり、爪が見えないときは、実際よりも値が低く推定され、推定誤差が大きくなっている。しかしながら、前腕回旋時の中指第 2 関節角度推定の誤差平均は 8.22 度、標準偏差 6.09 で、手指形状推定としてジェスチャ認識に必要な推定精度に達している。

次に、前腕回旋時の親指第 1 関節角度の推定精度実験を行った。実験結果を図 4.6 に示す。ただし、図 4.6-(a) の縦軸は関節角度、横軸はフレーム数を示しており、各図の太い実線はデータグループによる手指関節角度の測定値、細い実線は本システムによる手指関節角度の推定値、破線はそのときの前腕回旋角度の推定値を示している。図 4.6-(b) の縦軸は実験時の中指第 2 関節角度の推定誤差、横軸はフレーム数を示している。実験時の推定値は、10 フレーム前からの移動平均を求めることにより、平滑化を行っている。

図 4.6 より、前腕回旋時の親指第 1 関節角度が、推定できていることが確認できる。実験時の、前腕回旋時の親指第 1 関節角度推定実験での推定誤差平均は 12.80 度、標準偏差は 11.12 であった。中指第 2 関節角度推定のとくと同様に、前腕回旋角度が 0 度で固定の時は、比較的推定誤差が小さく、前腕回旋時には誤差が大きくなっている。親指第 1 関節角度の推定が、中指第 2 関節角度の推定より精度が劣る原因は、単にデータセットの数が足りないのではなく、前腕回旋角度が 30 度から 160 度の間、親指動作時の輪郭特徴量の変化が小さいからだと考えられる。輪郭特徴量の変化が極端に小さい場合、データベース作成時に同じデータセットと認識して削除されていることや、爪の位置情報に比べ、輪郭線特徴量の値がデータとして大きいことが原因だと考えられる。よって、データベース作成時に爪の位置情報の重みを調整し、式 (2.7) での類似度算出の際には、手指特徴量と爪の位置についてそれぞれ重みを置くなどして対策する必要がある。

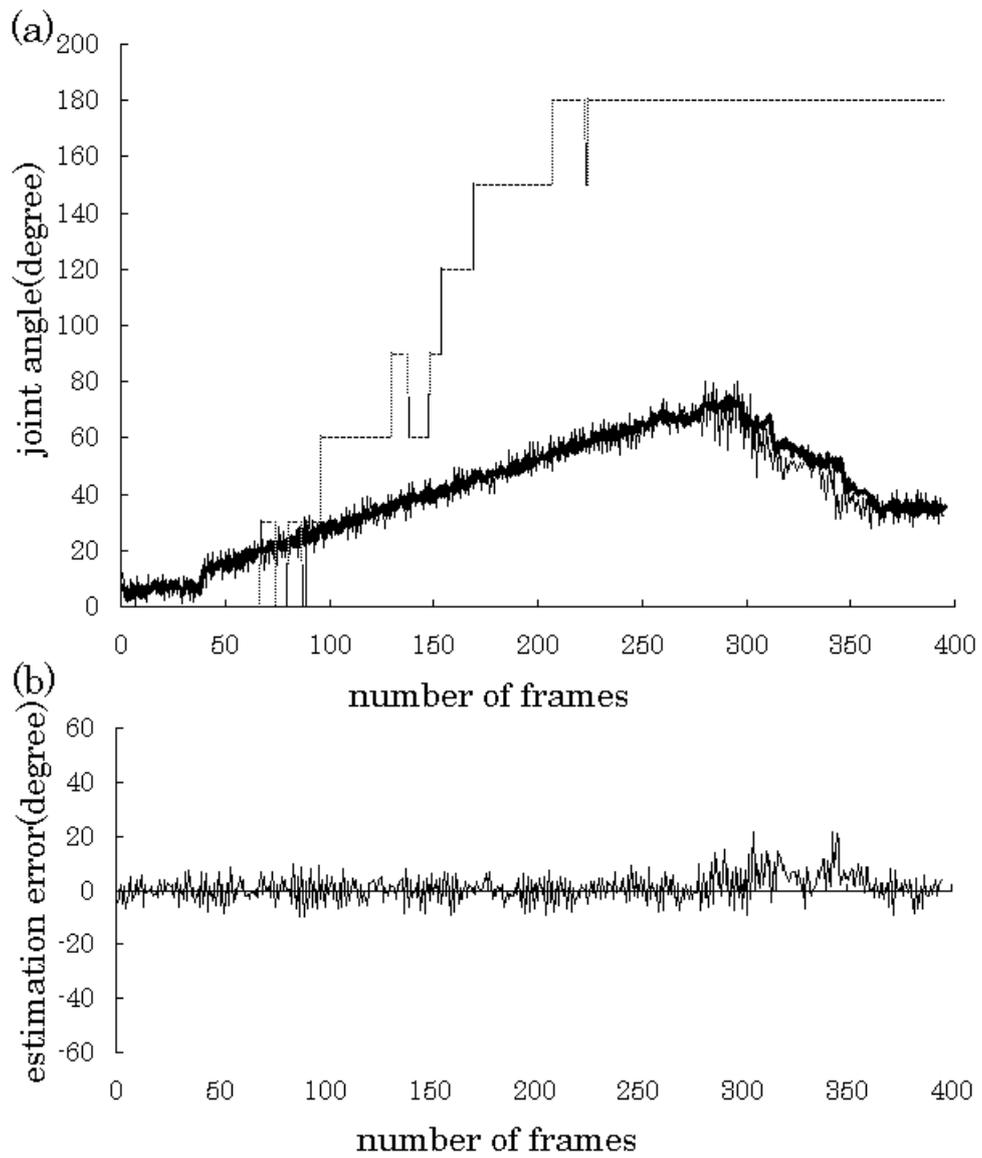


図 4.5: 前腕回旋時の中指第 2 関節角度.(a) 推定精度.(b) 推定誤差.

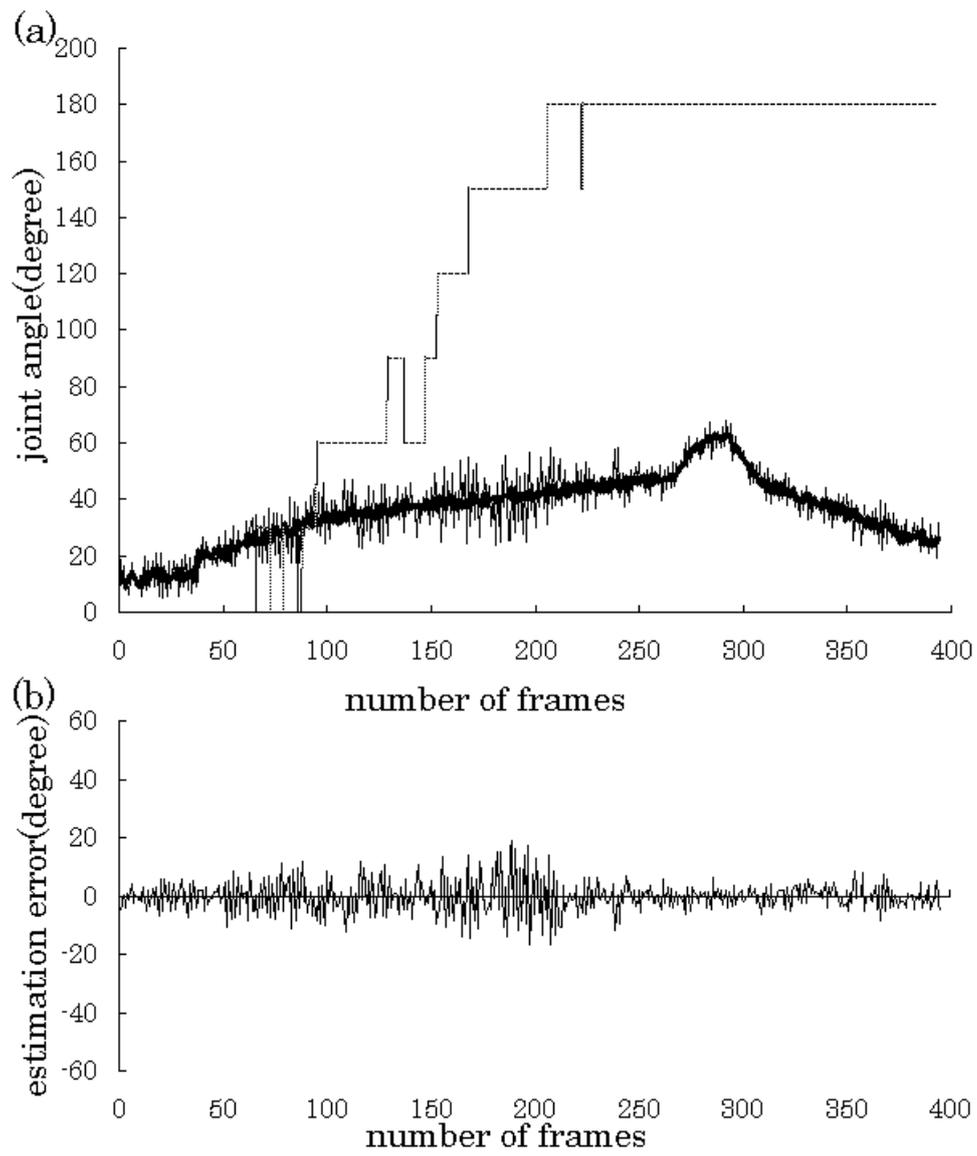


図 4.6: 前腕回旋時の親指第 1 関節角度.(a) 推定結果.(b) 推定誤差

4.2.3 爪の位置情報を持たないシステムとの比較

本システムと爪の位置情報を持たないシステムとの比較のために，輪郭線特徴量と手指関節角度のみをデータセットとし形状推定を行うシステムを構築し，前腕回旋時の推定精度実験を行った．データベース構築手法は，本システムと同じく，形状推定時の類似度算出には次式を用いた．

$$E_r = \sum_{i=1}^N (x_i - x'_i)^2 \quad (4.5)$$

ただし， E_r は類似度， P は高次局所パターン数， x_i はデータベース内のデータセットの手指輪郭の i 番目の特徴量， x'_i は取得画像の手指輪郭の位置 i 番目の特徴量，25 は図 3.6 で示す局所パターンの数を示している． E_r を最小にするデータセットを検索し，そのデータセットが持つ手指の関節角度を出力することによって，3次元手指形状推定を行う．

カメラの前で自由に手を動かした時の人差し指第 2 関節角度推定の実験結果を図 4.7 に示す．ただし，図 4.7-(a) の縦軸は関節角度，横軸はフレーム数を示しており，図中の太い実線はデータグロブによる手指関節角度の測定値，細い実線は本システムによる手指関節角度の推定値，破線はそのときの前腕回旋角度の推定値を示している．図 4.5-(b) の縦軸は実験時の中指第 2 関節角度の推定誤差，横軸はフレーム数を示している．実験時には，推定値は 10 フレーム前からの移動平均を求めることにより，平滑化を行っている．前腕回旋動作は 30 度ごとの 6 段階で計測されており，右手の掌がカメラに向いているときの前腕回旋角度を 0 度，手の甲がカメラに向いているときの前腕回旋角度を 180 度とした．

推定の誤差の標準偏差 22.0 であった．爪の位置情報を含まないシステムは，ジェスチャ認識に必要な推定精度に達していない．前腕回旋角度が 80 度から 180 度の間は，0 度から 80 度のときの推定に比べ，推定誤差平均が 39.7 度増加していた．つまり，前腕回旋角度が 80 度以降の時，推定精度が落ちる．爪の位置情報を含んだ提案システムは，前腕回旋角度 80 度以降も高い推定精度を保っており，前腕回旋動作時の手指形状推定には，爪の位置情報が有用であると確認できる．

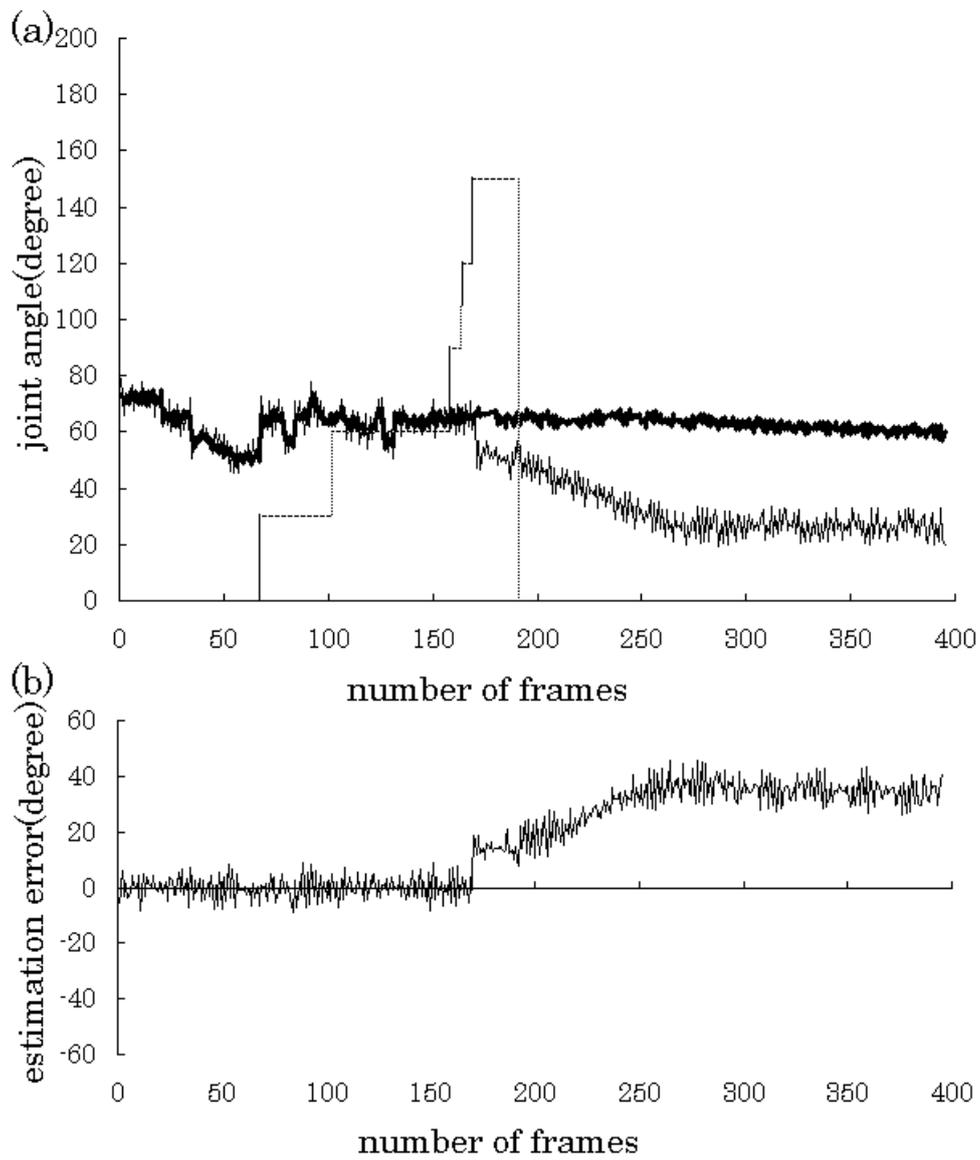


図 4.7: 爪の位置情報を持たないシステムでの人差し第 2 関節角度の推定精度.(a) 推定精度.(b) 推定誤差.

第5章 結論

カメラ動画入力を入力を主体とした，手指に何も装着しない状態での，3次元手指動作による情報入力ユーザインタフェースを実現するためには，前腕回旋動作を含んだ3次元の手指形状推定を高速かつ高精度に行わなければならない．3次元の情報入力ユーザインタフェースを実現するためには，形状推定の誤差の標準偏差 10 度以内の推定精度と 70fps 以上の処理速度が，3次元手指形状推定システムに求められている．

3次元手指形状推定のひとつに，データベースから類似データを高速検索する手法があるが，従来手法では前腕回旋動作を含んだ3次元手指の推定精度に限界があった．また，前腕回旋動作を含んだ手指形状推定には，データベース量増加による処理速度の低下や，指の自己遮蔽の問題や，表裏の認識の問題により，推定精度が著しく低下する問題があった．これらの問題を解決するためには，手指の表裏情報と指先の位置情報が必要と考えられる．

そこで本研究では，第一にヒトが手の表裏情報を含んだ手指形状を理解する際の視覚的特徴を明らかにすることにより，カメラで手指形状推定を行う際に最適な要素について検討した．心理知覚実験の結果，視覚的手掛かりにおける関与度の比率は，爪が 25 %，皺が 20 %，皮膚の色が 15 %であった．

第二に，手指の輪郭の特徴量に爪の位置情報を追加し，手の表裏情報も考慮した手指形状推定システムを構築した．まず，データグローブを装着した状態で手画像を取得し，手指の輪郭線特徴量，爪の位置情報及び関節角度データを持ったデータセットを作成した．次に，データセットを前腕回旋角度と手指関節角度によってクラスタリングし，データベースを構築した．最後に，手指画像がシステムへ入力されると，その手指画像の爪の位置情報と輪郭線特徴量を元にデータベース内の類似データセットを検索し，そのデータセットが持つ手指関節角度を出力する手指形状推定システムを構築した．また，このとき，肌領域と爪領域を高速に抽出できる個人別表色系を提案した．さまざまな照明環境で個人別表色系を用いた肌色抽出精度実験を行ったところ，平均処理速度 149fps で肌色抽出率 94% の高精度処理を実現した．さらに，楕円体の内外判定による抽出手法と個人別表色系による抽出手法の速度と精度についての検証を行い，肌色抽出において必要な精度と速度の有用性を確認した．

データグローブを装着した状態で手指形状推定を行い，データグローブから出力された関節角度と，本システムから出力された関節角度を比較して行った評価実験の結果，8934 個のデータ数に対して，CPU が PentiumIV ，2.8GHz の通常仕様の PC で，前腕回旋動作に対応しつつ手指関節角度の誤差の標準偏差 7.23 度，処理速度 100fps が実現できた．これは，前腕回旋動作を許容しない従来システムと同程度の精度であった [5],[6]．一方，爪の位置情報を持たない推定システムの標準偏差は 22.0 度であり，要求される仕様を満たせない．一連の結果

から，前腕回旋動作時の手指形状推定には，爪の位置情報を追加した本システムが有用であることを確認できた．

謝辞

本論文は、著者が筑波大学在学中に、星野 聖 准教授の御指導のもとに行った研究をまとめたものである。本研究をまとめるにあたり、多大なる御指導、御鞭撻を賜った筑波大学大学院システム情報工学研究科 星野 聖 准教授に心より感謝の意を表します。本論文を完成させるにあたり有意義な御指導と御鞭撻を賜った同研究科 鬼沢 武久 教授、亀田 能成 准教授に深く感謝いたします。

また、本論文を執筆するにあたり、形状認識実験における適切な御指導を頂き、動画像の高速処理のためのシステム設計に関し貴重なご意見、御討論を頂いた、国立大学法人琉球大学 工学部情報工学科 知能情報工学 當間愛晃 助教と同学科 複雑系研究室、並びに並列信頼研究室の皆様、川淵機械技術研究所 川淵一郎 博士に深く感謝致します。最後に、常日頃よりミーティングなどで貴重な意見と御支援頂きました、医用生体工学研究室の皆さんに深く感謝いたします。

参考文献

- [1] 島田伸敬. 白井良明. 久野義徳. 三浦純. 緩やかな制約知識を利用した単眼視動画像からの関節物体の形状と姿勢の同時推定. 信学論, Vol. J81-D-2, No. 1, pp. 45–53, January 1998.
- [2] J.Cui. Z Sun. Visual hand motion capture for guiding adexterous hand. *Proc. 6th Int. Conf.on Automatic Face and Gesture Redcognition*, pp. 729–734, 2004.
- [3] 今井章博. 島田伸敬. 白井良明. 輪郭の変化学習による 3-D 指先姿勢の認識. 信学論, Vol. J88-D, No. 8, pp. 1634–1651, 2005.
- [4] K.Hoshino. T.Tanimoto. Real time search for similar hand images from database for robotic hand control. *IEICE Transactions on Fundamentalsof Electronics, Communications and Computer Sciences*, Vol. E88-A, No. 10, pp. 2514–2520, October 2005.
- [5] 谷本貴頌. 星野聖. 人間-ロボット間コミュニケーションのための実時間・高精度ヒト手指形状推定. ヒューマンインタフェ-ス学会論文誌 ISSN, Vol. 7, No. 4, pp. 535–540, November 2005.
- [6] 谷本貴頌. 星野聖. 自己増加型 SOM を用いた画像データベースからの手指形状の実時間推定. 情報処理学会 CVIM 論文誌, Vol. 47, SIG5(CVIM13), pp. 100–107, March 2006.
- [7] Vssilis Athitsos. Jonathan Alon. Stan Sclaroff. Geroge Kollios. BoostMap:A Method for Efficient Approximate Similarity Rankings. *IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, Vol. 2, pp. 268–275, 2004.
- [8] R.Rosales. V.Athitsos. L.Sigal. S.Sclaroff. 3D Hand Pose Estimation using Specialized Mappings. *Proc.of 8th International Conference On Computer Vision*, pp. pp378–387, 2001.
- [9] 玉城絵美. 星野聖. 手首関節動作を含んだ手指形状の 3 次元推定. 電子情報通信学会技術報告書 WIT 福祉情報工学, Vol. 107, No. 179, pp. 59–62, July 2007.
- [10] Emi Tamaki. Kiyoshi Hoshino. 3D Copycat hand with wrist motions- Robot hand imitating human motions at high speed and with high accuracy -. *Proc. 8th Intl. Conf. on Intelligent Technologies*, Vol. 8, December 2007.
- [11] Emi Tamaki. Naruaki Toma. Color system for skin color extraction. *Journal of Artificial Life and Robotics (Springer)*, Vol. 12,December 2007.

- [12] 玉城絵美. 星野聖. 手指形状推定装置及び手指形状推定方法, ならびにプログラム. 特願 2007-165689, 特許出願日 2007 年 6 月 22 日.
- [13] 玉城絵美. 星野聖. カラーリストバンドを用いた前腕回旋角度推定装置及び前腕回旋角度推定方法, ならびにプログラム. 特許出願日 2007 年 10 月 18 日.
- [14] 玉城絵美. 星野聖. 安定した肌色抽出のための個人表色系. 日本バーチャルリアリティ学会論文誌, Vol. 12, No. 4, pp. 471–478, December 2007.
- [15] N.Ostu. T.kurita. A new scheme for practical, flexible and intelligent vision systems. *Proc.IAPR. Worksshop on Computer Vision*, pp. 431–435, 1998.
- [16] A.A.Argyros. M.I.A.Lourakis. Tracking Skin-colored Objects in Real-time, Cutting Edge Robotics Book. ISBN 3-86611-038-3, Advanced Robotic Systems International, 2005.
- [17] 松井望. 山本喜一. バーチャルキーボード:ビデオ画像からの頑健な実時間指先検出の実現. SPA2000, 2000.
- [18] Satoshi Matsushashi. Osamu Nakamura. Toshi Minami. Human-face extraction using modified HSV color system and personal identification through facial image based on isodensity maps. *IEEE CCGEI '95*, pp. 909–912, 1995.
- [19] 川戸慎二郎. 鉄谷信二. 顔領域抽出を目的とした肌色モデルと肌色領域抽出. 電子情報通信学会信学技報, PRMU2001-59, pp. 143–148, 2001.
- [20] 青木茂樹. 大西正輝. 小島 篤博. 山田 憲一. 三宅 克実. 福永 邦雄. 運転者の目領域抽出による視線推定. 電子情報通信学会技術研究報告 PRMU, Vol. 99, No. 450, pp. 15–22, November 1999.
- [21] 鈴木 恒男. 小谷津 孝明. 表情から認知される感情に及ぼす顔色の効果に関する研究. 日本色彩学会誌, Vol. 22, No. 2, pp. 45–52, July 1998.

付録A カラーリストバンドによる前腕回旋角度の検出方法

スペクトルカラーを描画したカラーリストバンドを手首に装着し、カメラで撮影したリストバンドの色から前腕回旋角度の検出する(図 5.1)。前腕回旋角度によってカメラから見える色が変わる。予め作成しておいた色の平均値と前腕回旋角度の対応表から、前腕回旋角度を出力する。このとき、カラーリストバンドの手首への巻き開始地点は、対応表作成時と前腕回旋角度検出時とも同じである。また、カラーリストバンドの枠線の色から、色の切り出し範囲を決定する。切り出し範囲内の色と枠の色を混同しないために、枠線の色は、スペクトルカラー上には描画しない。

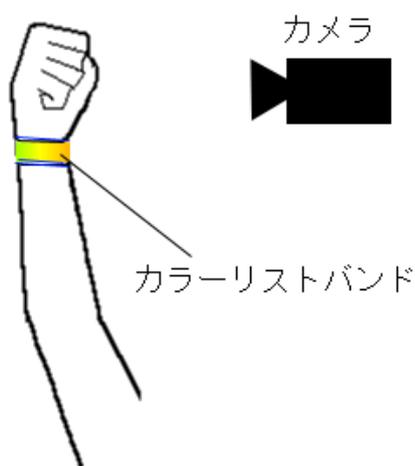


図 5.1: カラーリストバンドによる前腕回旋角度の検出の様子.

図 5.2 に色と前腕回旋角度のテーブル作成処理のフローチャートを示す。はじめに、画像を取得し、平滑化処理によってゴマ塩ノイズを除去する。次に、カラーリストバンドの枠の色によって 2 値化して、その枠内部の色を切り出し、色の平均値を算出する。最後に、色の平均値と前腕回旋角度の対応表を照会し、前腕回旋角度を出力する。図 5.3 にカラーリストバンドのペーパークラフト例を示す。

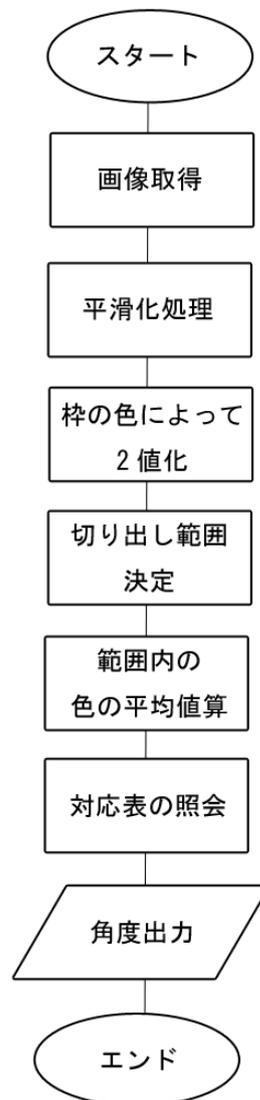


図 5.2: カラーリストバンドによる前腕回旋角度の検出のフローチャート.

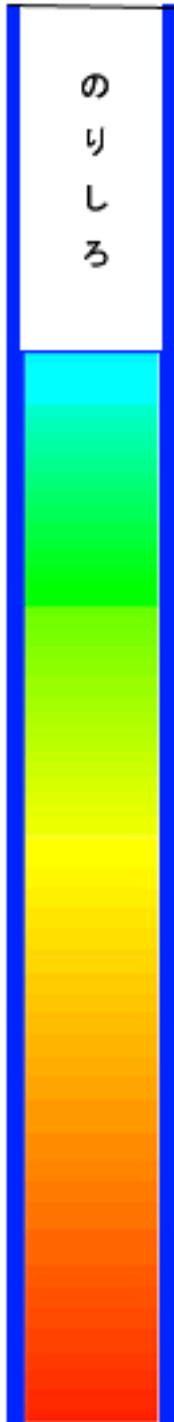


図 5.3: カラーリストバンドのペーパーラフト例.

付録B 形状推定の結果

B.1 手指の各関節を動かさない状態

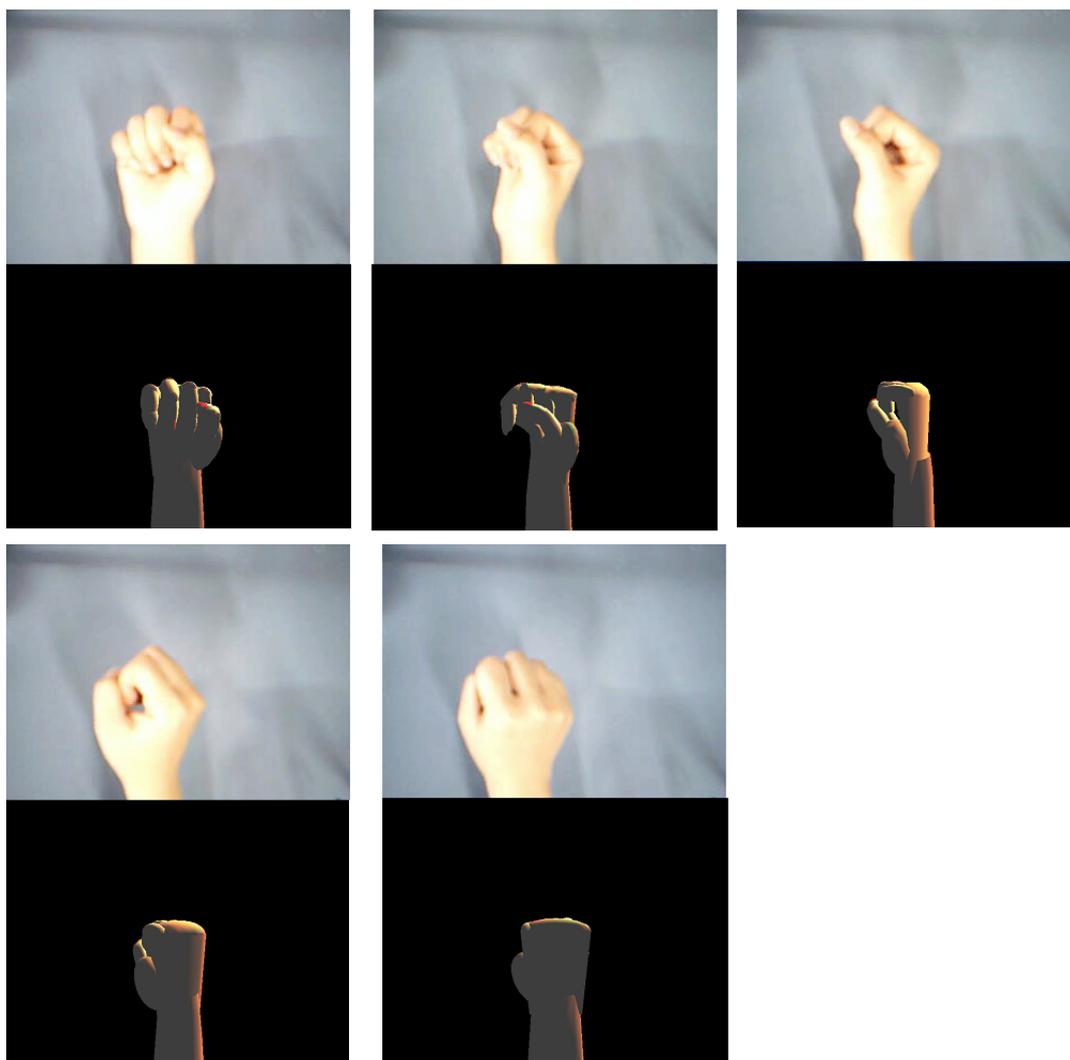
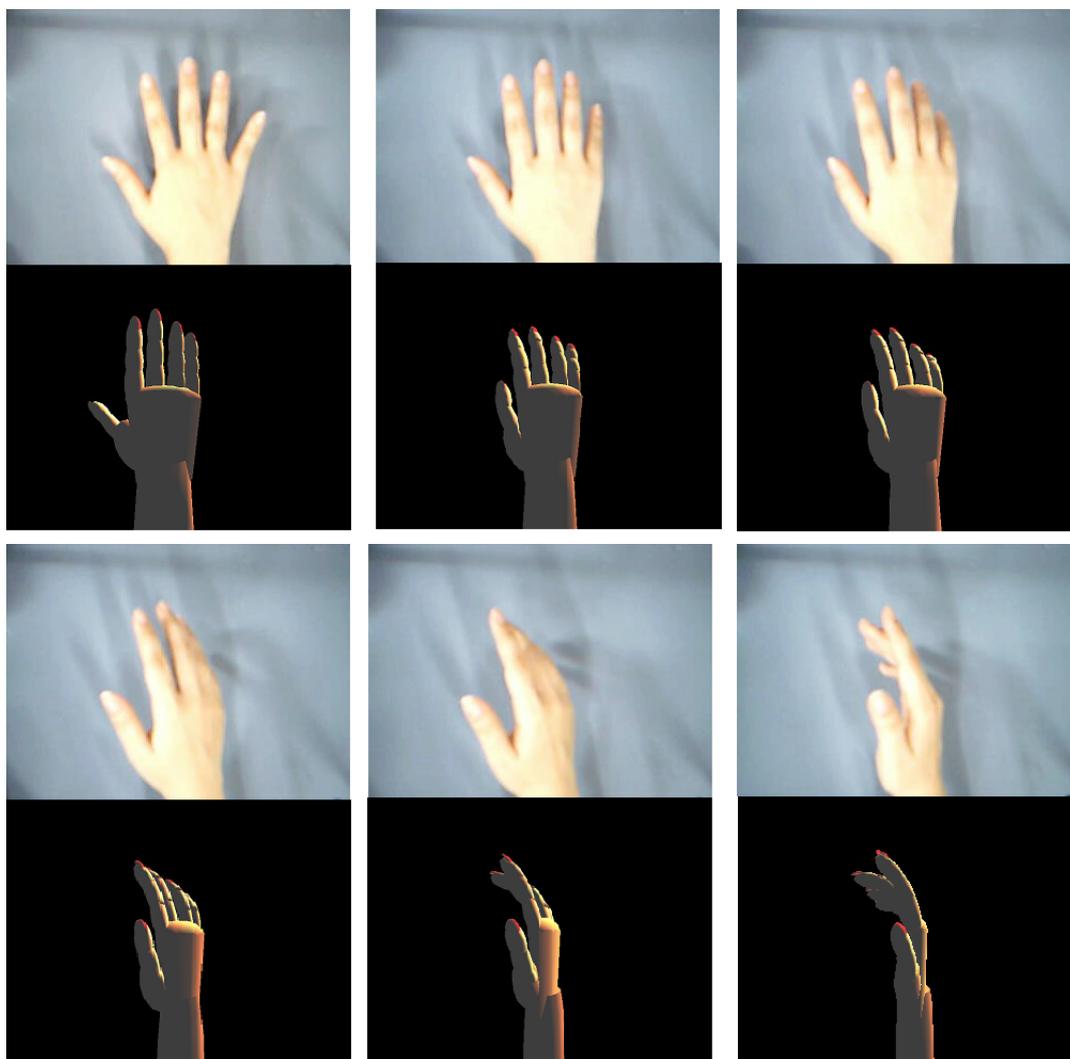
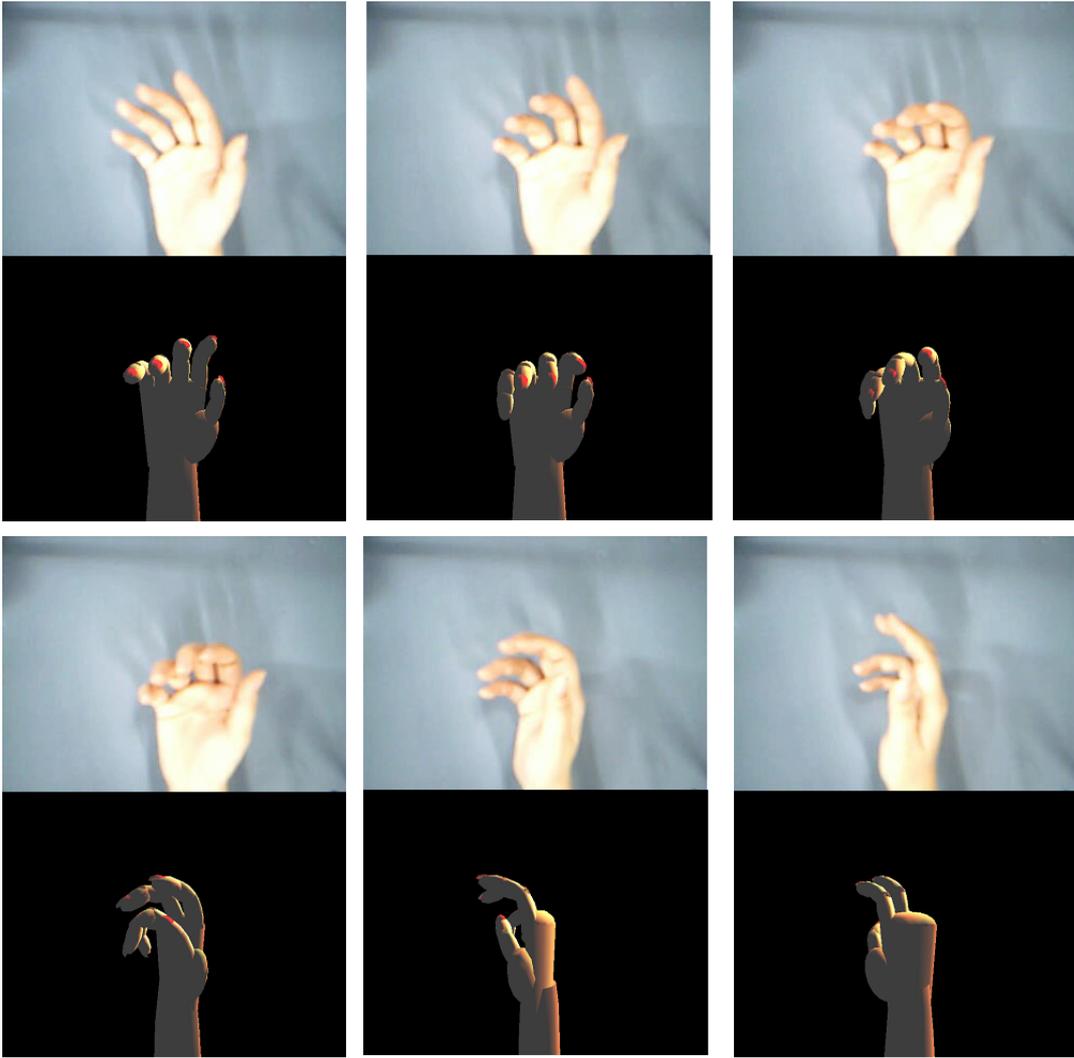


図 5.4: 手指の各関節を動かさない状態.

B.2 手指の各関節を自由に動かした状態





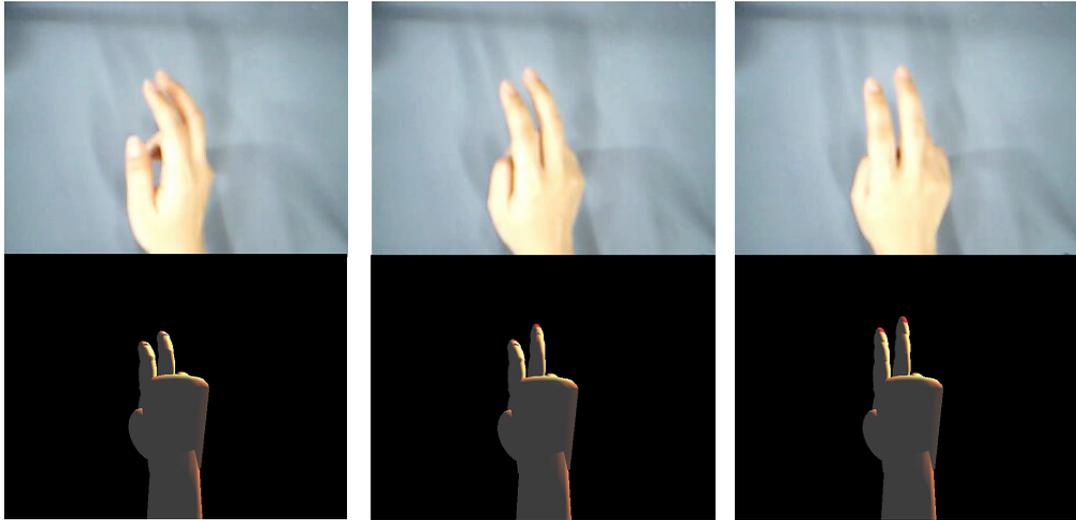


図 5.5: 手指の各関節を自由に動かした状態.