

## 爪領域周辺に特徴的に生じる色の分布状態を利用した 手画像からの爪検出手法

著者	藤嶋 教彰
発行年	2015
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2014
報告番号	12102甲第7299号
URL	<a href="http://hdl.handle.net/2241/00129421">http://hdl.handle.net/2241/00129421</a>

# 爪領域周辺に特徴的に生じる色の分布状態を 利用した手画像からの爪検出手法

藤嶋 教彰

システム情報工学研究科

筑波大学

2015年3月

## 概要

近年の3次元映像提示装置とAR技術の発展に伴い、3次元空間上で情報を操作できるようになってきた。しかし、現実世界と同じような直観性で情報を操作するためには、現実で物を操作する際に用いる「手」の情報をそのまま使用する、非接触の手指形状推定を利用したジェスチャインタフェースの構築が必要となると考えられる。単眼でRGBカメラを使用した手指形状推定技術の精度向上のために爪情報が利用できると考えられるが、爪領域の色と掌領域の一部の色が類似しているため、掌領域を含む手画像から爪を安定して検出することは困難であり、それを実現したものの存在は確認できなかった。掌領域に対応させるためには新たに、爪らしい色の領域を抽出した後に、それが爪か肌かを判別する過程が必要であった。

著者は、画像空間に着目すると爪領域の周辺には特徴的に生じる色の分布状態があることを発見し、それを利用するための手法を開発した。本論文では、色の分布状態を利用した爪判別手法を提案し、掌領域に対応し、手の裏表に関わらず爪検出可能なシステムの構築方法を提示する。

爪判別アルゴリズムは、爪領域周辺に生じる色の分布状態を大きく2つの性質、色が独立した小領域が存在すること、局所的に見たときに特に爪らしい色を持つ画素が1か所に密集すること、としてまとめ、それぞれを用いた判別法で爪と判別された領域だけを爪として検出した。

評価実験では日本人大学生5名を被験者とし、前腕回旋角度と爪検出率の関係、および提案した爪判別アルゴリズムの除去性能を評価した。評価実験の結果、本手法を用いたシステムは前腕回旋角度 $-90^{\circ}$  から $-50^{\circ}$ 、 $40^{\circ}$  から $90^{\circ}$  の範囲で、爪検出率85%以上を達成した。つまり、掌側に対応できる、手の裏表に関わらず爪検出可能なシステムを実現した。また、除去性能を評価した結果、爪判別アルゴリズムに使用した2つの性質は共に欠点を補完し合う関係にあり、合わせると誤検出する肌領域を95%を超える確率で除去できることが判明した。すなわち、2つを共に使用することが重要であると確認された。

本論文は全6章で構成されている。第1章では研究の背景について扱い、手指形状推定の必要性和爪検出システム発展の重要性を述べる。第2章では本研究に関連のある分野の知識についてまとめる。第3章では爪を用いた手指形状推定について扱い、爪が本当に手指形状推定の精度改善に重要なのか、およびどのようにシステムを組むと良いかについて述べる。第4章では爪検出システムを扱い、先行研究の問題を解決するために、本研究が必要であることを示し、画像空間上の色の分布状態を利用した爪判別アルゴリズムを提案する。そして、掌側に対応した手の裏表に関わらず爪検出可能なシステムの構築手法を示す。ここまでの、主題を閉じ、第5章では爪検出システムを他のシステムに応用する例として仮想ネイルアートシステムを扱う。ここでは、指のモデルを用いた奥行き回転角推定と仮想奥行き座標を用いたネイルチップ変形方法を提案し、前腕回旋を含む手画像にも自然に重畳できる仮想ネイルアートシステムの構築手法を示す。最後に第6章で本論文の総括を行う。

# 目次

第1章 はじめに	1
第2章 手・領域抽出・姿勢推定に関する研究	4
2.1 単眼 RGB カメラを用いた手指形状推定に関する研究	4
2.2 肌・爪の構造や反射モデルに関する研究	6
2.3 色テクスチャ分析に関する研究	7
2.4 数理的手法を利用した色物体認識に関する研究	8
2.5 照明変動や影がある場合の色物体抽出に関する研究	9
2.6 単眼画像からの3次元形状および姿勢推定の研究	10
第3章 爪を用いた手指形状推定の構想	11
3.1 手指形状推定システムの概要	11
3.2 手指形状推定システムの抱える問題と解決案	14
3.2.1 輪郭線のみを用いた特徴抽出における問題	14
3.2.2 マッチングに用いる HLAC アルゴリズムが持つ性質に関わる問題	15
3.2.3 ブロック分割をしてから HLAC を使用する手法に関わる問題	15
3.2.4 各問題に対する解決案	16
3.3 爪を用いた形状推定システムの先行研究	18
3.4 爪の使用方法に関するリスクの検討	20
3.5 爪を用いた手指形状推定システムの構想	21
3.5.1 構想システムの概要	21
3.5.2 評価関数と可変パラメータ設定方法の提案	22
3.6 構想システムの実験結果	25
3.7 第3章のまとめ	29
第4章 爪検出システム	30
4.1 爪領域抽出に言及した先行研究と課題	30
4.2 爪らしさを表す指標軸の生成	34
4.2.1 手領域画素が作る画素分布の傾向	34
4.2.2 爪らしさを表す指標軸の生成	35
4.2.3 指標軸の評価	37
4.3 爪検出システムの動作環境と処理の流れ	41
4.4 爪らしい色を持つ領域の抽出と注目点の設定	43
4.5 局所的な情報を利用した爪と肌の判別	45
4.5.1 爪と肌の判別を実現するための情報	45
4.5.2 領域の独立性に着目した爪と肌の判別手法	45
4.5.3 色のまとまり方に着目した爪と肌の判別手法	46

4.6 評価実験	53
4.7 爪領域候補の抽出過程を改善する試み	57
4.8 第4章のまとめ	63
第5章 爪検出システムを利用した別の応用例	64
5.1 仮想ネイルアートシステム構築の背景	64
5.2 仮想ネイルアートシステムの構成	66
5.3 ネイルチップ画像の標準化	68
5.4 ネイルチップ画像重畳のためのパラメータ設定手法	69
5.5 仮想奥行き座標を利用したネイルチップ画像の変形	72
5.6 ネイルチップ画像の手画像への重畳	76
5.7 仮想ネイルアートシステムの評価	78
5.8 第5章のまとめ	81
第6章 まとめ	82
謝辞	83
参考文献	84
付録A 一般表色系	91
著者論文リスト	96

# 図目次

図 3-1	本論文で対象とする形状推定システムの概要図	11
図 3-2	HLAC 特徴量を使用したデータマッチング用 1600 次元ベクトルの算出	13
図 3-3	輪郭線形状が似ていることが原因で誤推定を起こす例	14
図 3-4	HLAC アルゴリズムが持つ性質によって誤推定を起こす例	15
図 3-5	HLAC を用いたアルゴリズムにより生じる問題の概要図	16
図 3-6	画素の付加処理によるブロックまたぎの防止	17
図 3-7	輪郭線+爪情報による推定	18
図 3-8	爪をふるいわけに使用した推定	18
図 3-9	構想システムの概要図	21
図 3-10	曲率を考慮した HLAC 用重みデータ列の生成	24
図 3-11	構想システムの計測実験時における撮影環境	25
図 3-12	指を間違える推定結果の数の比較	26
図 3-13	爪の情報が推定ミスを防いだ例	27
図 3-14	指の曲げの浅さ・深さに関する推定の改善	27
図 3-15	爪情報が原因と考えられる指間違いの例	28
図 3-16	爪や輪郭線の曲率情報を利用してても正確な推定が困難な形状の例	29
図 4-1	色情報のみで爪検出を行った場合における前腕回旋角度と爪検出率の関係	32
図 4-2	爪領域より大きく抽出された肌領域が存在する場合	32
図 4-3	手領域画素が作る画素分布 (1)	34
図 4-4	手領域画素が作る画素分布 (2)	34
図 4-5	画素分布取得に用いる手画像	35
図 4-6	主成分ベクトルを基底に採用した場合の画素分布	36
図 4-7	主成分ベクトルが表す色	37
図 4-8	線形判別分析によって得られたベクトルが示す色	39
図 4-9	ROC 曲線の概要	39
図 4-10	爪と肌を対象とした ROC 曲線	40
図 4-11	爪検出を行う撮影環境	41
図 4-12	爪検出システムの処理の流れ	42
図 4-13	爪として誤検出しやすい肌領域の分布位置	43
図 4-14	抽出割合と頻度の例	44
図 4-15	第 3 主成分得点を可視化したマップ画像	45
図 4-16	結合領域の生成手法	46
図 4-17	ROI 中心と色が類似する領域の抽出結果	47

図 4-18	中心付近の画素と中心から離れている画素の持つ色の分布の違い.....	48
図 4-19	被験者 1 から 3 における肌色画素の除去率と密度の関係.....	49
図 4-20	被験者 4 から 6 における肌色画素の除去率と密度の関係.....	50
図 4-21	上位 20% の第 3 主成分得点を持つ画素の抽出結果.....	51
図 4-22	評価実験で用いた手の例.....	54
図 4-23	前腕回旋角度と爪検出率の関係.....	54
図 4-24	爪検出システムの概要図と着目点.....	57
図 4-25	爪領域候補の抽出時における問題.....	57
図 4-26	第 3 主成分得点のマップ画像とマップ画像の平滑化画像.....	58
図 4-27	二次元共起ヒストグラム画像.....	59
図 4-28	CH 画像を用いた画素の除去結果.....	59
図 4-29	新しい抽出手法による爪領域候補の抽出結果.....	60
図 4-30	爪領域候補抽出の改善手法を用いた時の爪検出率.....	61
図 4-31	爪領域候補抽出の改善手法の有無による爪検出率の増減.....	61
図 4-32	爪領域と指の側面領域の結合による爪領域候補抽出の失敗.....	61
図 4-33	回旋角度の変化による平均第 3 主成分得点の変遷.....	62
図 5-1	仮想ネイルアートシステムの動作環境.....	66
図 5-2	仮想ネイルアートシステムの処理の流れ.....	67
図 5-3	ネイルチップ重畳に必要な 5 つのパラメータ.....	67
図 5-4	回轉變換後の座標.....	68
図 5-5	標準化ネイルチップ画像.....	68
図 5-6	指領域ごとの ROI 設定と ROI 内の 2 値化画像.....	69
図 5-7	指の回転軸取得のための処理.....	70
図 5-8	指の回転軸方向から見た指のモデル.....	70
図 5-9	奥行き座標を持たない画像を奥行き方向に回転した時の変化.....	72
図 5-10	仮想奥行き座標を利用した偏心の実現.....	73
図 5-11	$z_{\max} = 5$ としたときの回轉變換の様子.....	74
図 5-12	$z_{\max} = 10$ としたときの回轉變換の様子.....	74
図 5-13	$z_{\max} = 15$ としたときの回轉變換の様子.....	75
図 5-14	$z_{\max} = 20$ としたときの回轉變換の様子.....	75
図 5-15	ネイルチップの重畳結果 (右回旋時).....	76
図 5-16	ネイルチップの重畳結果 (左回旋時).....	77
図 5-17	推定値と実測値のユークリッド距離の差.....	78
図 5-18	指先座標の振動性と推定ブレの様子.....	79
図 5-19	$\varphi_i$ と $z_j$ の値が 0 の時.....	79
図 5-20	提案手法 ( $z_{\max} = 10.0$ ).....	80

付録

付録図 1	$xy$ 色度図 .....	92
付録図 2	色相・彩度平面 .....	93

## 表目次

表 4-1	軸評価のための予備実験に参加した被験者と手に関するデータ .....	39
表 4-2	抽出比率の違いによる密集度調査に用いる被験者の性別と手の色.....	48
表 4-3	爪検出精度の評価実験に用いた手の情報とパラメータの数 .....	53
表 4-4	肌を正しく除去できた割合.....	55
表 4-5	爪領域抽出の改善手法評価のために使用した参加者の手とパラメータの情報	60

# 第1章 はじめに

コンピュータを操作する入力インタフェースの発展は、人々の生活を大きく変化させてきた。コンピュータのディスプレイを見ながらプログラムベースで操作する方式として、初めて登場したのはキーボード入力によるコマンド実行で動作命令をする **Character User Interface** 方式であった。この方式はコマンド入力を行うためには多くの専門知識が必要である。そのため、利用できる人が限られた。その後、大きな変革が生じた。1つ目は現実世界の2次元と画面上の2次元を対応させる入力インタフェース、マウスの登場である。2つ目はコンピュータグラフィックスとポインティングデバイスを組み合わせ、アイコンなどにより2次元の情報表示をする **Graphical User Interface** 方式の登場である。これらが Lisa (Apple 社, 1983 年) というコンピュータで統合され、販売された。これにより、コンピュータの直観的操作が可能となり、一般家庭へコンピュータが普及させる要因の1つとなった。また、コンピュータの普及によって、情報の取得・操作・発信が容易になり、情報社会の基盤が形成された。2000年代後半からはさらに、直観的に操作できる **Natural User Interface** として、様々な分化を見せている。これまでは特に、携帯性と遠隔性に長けた技術が進化してきた。携帯性を発展させた要素はタッチパネルの開発と接触位置を正確に取得する技術の開発であった。これらにより、キーボードやマウスなしに、指を用いて様々な操作が行えるようになった。ユーザはコンピュータと自身の手があればよくなり、携帯性が著しく向上した。直観性も高まった。そして、タッチパネル方式を用いたスマートフォンやタブレット端末、そしてキーボードとの分離可能なノート PC の登場により、コンピュータを持ち歩く時代となった。遠隔性を発展させた要素は、掌の重心をトラッキングしてその重心の位置と距離を用いてスクリーン上の対象を操作するような、手の位置情報を用いたハンドトラッキングインタフェースの開発であった。このインタフェースは物に触れなくても操作できる遠隔操作性を持つ。また、不特定多数を対象にできる。島根産業技術センター ICT 技術開発プロジェクトチームはサイネージ・PC・エンターテイメントを目的として **Gesture-cam** [1]を開発した。その技術は東映太秦映画村に設置された次世代デジタル案内板「てぶりカム」[2]や島根県しまね海洋館のシロイルカとのインタラクション提供アプリケーション「**Gesture-Cam** 水族館用」[3]など、公共機関で実用化されている。

近年、コンピュータを用いた情報の利用方法にもう一つ、大きな変革が起き始めている。それを引き起こした大きな理由は情報の3次元化が可能となったことである。3D ディスプレイやヘッドマウントディスプレイなどの3次元情報提示装置の発展、実空間と仮想空間を融合する複合現実技術が向上したことにより、3次元情報を可視化して扱えるようになった。コンピュータ情報の活用範囲が3次元になると、情報の活用幅が大幅に広がり、また人々に恩恵を与えることができる。しかし、上記で述べたインタフェースは現実世界と同じ次元で自由に情報を操作できる自由度を持たない。現実世界と同じような直観性でコンピュータを操作するためには、現実で物を操作する際に用いる「手」の情報をそのまま使用する新しいインタフェースである、ジェスチャインタフェースの構築が必要となる。そして、このインタフェースを構築するための基盤となる、手の姿勢を推定する技術の発展も必要である。特に、ユーザに負担を強くない非接触型の手指形状推定手法が望まれている。

非接触型の手指形状推定技術では、2種類のカメラのどちらかが主に用いられる。1つは深度

カメラ, もう 1 つは RGB カメラである. 深度カメラは 3 次元空間情報を取得し, 平面方向を画像平面で, 奥行き方向を画素値で表現する. RGB カメラは 2 次元平面情報と色情報を取得し, 平面方向を画像平面で, 撮影物体の色を画素値で表現する. 近年では, Oikonomidis et. al による手指形状推定手法[4]のように深度カメラを用いるものが多く研究されている. 深度カメラを用いると現実世界の 3 次元位置情報をすべて取得できるため, 精度を出しやすい. しかし, 文献[4]で Intel i7 950 CPU, 6 GBs RAM, Nvidia GTX 580 GPU with 1581 GFlops processing power, 1.5 GBs メモリで 15 [fps]と示されているように, その情報量の大きさから高性能のコンピュータを用いても速度が遅くなる. 主原因は 2 つある. 1 つは奥行き方向を取得するセンサのノイズが強く, 大きなフィルタを用いて平滑化する必要があることである. もう 1 つは 3 次元情報の有効活用には 3 次元モデルによるフィッティングが必要で, 計算コストが高いシステムとなることである. 手を用いるアプリケーションでは高速で手を動かしても画像にぶれが生じないレベルである 60 [fps]は最低でも確保すべきである. 今後, コンピュータ性能の改善でこの速度は改善されるだろうが, 処理が本質的に重くなるシステムではインタフェースとして問題がある. シルエットの見えなどの情報で高速に手指形状の大まかな形を推定し, 3 次元情報は微調整をするためにだけ適用する方が理想的である. 高速な手指形状推定は RGB カメラを用いた手法で発展している. よって, もう一つの RGB カメラを用いた手形状推定技術に着目した.

RGB カメラを用いる場合も, 複数のカメラを用いる多視点方式と, 1 台の単眼方式がある. 複数のカメラを用いる手法は普及を目的としたインタフェースに不向きである. よって, 著者はジェスチャインタフェース構築のための単眼 RGB カメラを用いた手形状推定技術に対象を絞った.

単眼 RGB カメラを使用した手指形状推定の技術で高速性を持つ手法は手の外側の輪郭線情報を用いて 2 次元の見えに関してマッチングを行っていた. しかし, 外側輪郭線情報のみを利用することには大きな問題がある. 最も大きな問題は掌側に曲げた指の情報が外側の輪郭線に含まれないことが原因で, 形状推定が不安定となることである. これを解決する手段は大きく 2 つある. 内部のエッジ情報を利用すること, または爪の位置情報を利用することである. 著者は爪を用いる方が良いと考え, 爪に着目した. 内部エッジは手のシワと区別をつけにくく, 判別のための計算コストが高い. 一方で, 爪は人が目で見てすぐにわかるように, 色やテクスチャに肌とは違う性質を持つと考えられる. ゆえに, 内部エッジよりも計算コストが低くできる可能性が高い. また, 領域の重心情報という 1 次元で表現できるため, 情報が扱いやすい.

しかし, 爪検出に関する先行研究を調査すると, 利用できる範囲に大きな問題があることが判明した. これまで提案された手法の利用条件は爪と肌の色が異なることであったが, 掌に条件を満たさない領域が存在し, その領域を誤抽出する現象がみられた. つまり, 掌側に対応できるシステムではなかった. これでは, 手指形状推定に爪を利用できない. ゆえに, 爪を用いた手指形状推定を実現するために, 掌領域が写っていても爪検出できるシステムを新たに開発しなければならなかった. ここに重要性を感じ, 著者は手の表裏に関わらず使用できる爪検出システムの開発を主題として研究した. そして, 研究の結果, 爪領域周辺には, 肌に無い特徴的な色の分布状態が存在し, それが爪らしい色を持つ領域が爪か肌か判別する手掛かりになることがわかり, 掌側に対応できない問題を解決する大きな進歩を実現した.

本論文では, 主題として手指形状推定のための, 掌領域に対応した, 手の裏表に関わらず利用

可能な爪検出手法を提案する。本手法では爪らしい色を持つ領域を抽出し、その領域が爪か肌かを爪周辺に存在する特徴的な色の分布状態を利用して判別する 2 段階検出手法をとることで、爪を安定に検出する。

本論文は全 6 章によって構成される。第 2 章では本論文で扱う研究分野に近い領域として事前調査した手、色情報を用いた領域抽出、姿勢推定に関する先行研究について述べる。第 3 章では爪を用いた手指形状推定について扱い、爪情報が手指形状の推定精度向上のために有用であることを示す。まず、単眼 RGB カメラを用い、輪郭線の情報を利用する高速な手指形状推定を一つ取り上げ、そのアルゴリズムの概要を述べる。そして、輪郭線のみを利用すると、掌側に指を曲げた形状を認識することが困難になるという問題点を、例を挙げて述べる。そして、その問題点を解決する手法として爪情報の利用について言及する。爪を利用した形状推定の先行研究と問題点に触れた後、新しい手指形状推定アルゴリズムの構想を述べる。最後に実際の実装結果を基に爪が有用であることを示し、同時に生じうる問題点と課題を提示する。そして、どのようにシステムを組むべきかを言及する。第 4 章では画像から爪を検出する研究を扱い、爪周辺に起こる特徴的な分布状態を利用することで、掌領域が存在しても爪検出できるシステムが構築できることを示す。ここではまず、これまでの先行研究について挙げる。そして、先行研究で前提としていた、爪領域は適切な閾値調整ができれば取得できるという前提が掌側を含む手画像に当てはまらない問題が存在すること述べる。そしてその問題点を解決するために、爪らしい色の領域を取得した後に、それが爪か肌かを判断する過程を入れるという 2 段階検出をする方針を示す。著者が構築した爪検出手法の詳細部を述べるとともに、研究主題である爪か肌か判別するアルゴリズムを提案し、掌領域を含む手画像にも対応できる爪検出システムを実現する。以上の章までで、手指形状推定に関する研究の話題を閉じる。第 5 章では、爪検出システムの別の応用を目的とし、仮想ネイルアートシステムの開発を取り上げる。ここでは、先行研究において、前腕回旋について考慮していないという問題について触れる。そして前腕回旋がある時に生じる、画像の奥行き方向への変化に対応するためのパラメータ設定手法、およびネイルチップの変形手法を提案する。最後に 6 章にて本論文の総括を行う。

## 第2章 手・領域抽出・姿勢推定に関する研究

### 2.1 単眼 RGB カメラを用いた手指形状推定に関する研究

単眼 RGB カメラはカメラ撮影平面に対して被写体が存在する場所を示す 2 次元座標、RGB 各色に対する画素値の 5 次元の情報を取得する。手の形を全く知識なしに正しく知るためには上下・左右・前後の 3 次元情報が必要である。しかし、RGB カメラは前後の奥行き情報を取得できない。つまり、RGB カメラによる手指形状推定は不良設定問題となる。この問題を解決するためには制限をつけるか先見知識を用いて奥行き情報の欠落分を補う必要がある。具体的には、制限は装着物を手に付けること、先見知識は手のモデルを持っておくこと、または見えと形状の対応関係をデータベースとして所持することである。

単眼 RGB カメラで最も精度を出しやすいのは手にマーカをつけた手袋をはめる装着型である。装着型は全体にカラーパターンを配置したグローブを用いて、撮影された手画像の色のパターンから手形状を推定する[5]のような全体情報を利用する手法と、形状を推定する際に重要な指・関節・手首に色の目印をつけ、その目印の見え方のパターンから形状を推定する[6]のように部分的な情報を利用するパターンがある。装着型の長所は手の形状を特定するための情報を確実に取得でき、高精度なシステムを比較的容易に構築できることである。短所は装着をユーザに強いることである。

マーカを使用しない非装着型の形状推定は、主に 2 つの方式に分かれる。1 つは 3D model-based 方式で、もう 1 つが 2D appearance-based 方式である。入力と共にカメラ映像から得られるシルエットや輪郭線であるが、比較するデータの保持方式が異なる。

3D model-based 方式はあらかじめ計算機内で構築した 3 次元形状モデルを保持する。モデルを利用してパーティクルフィルタを利用して手画像と照合する[7]、モデルのオプティカルフローと手指画像のオプティカルフローからトルクを生成し、手指角度の逐次推定をする[8]などの手法を用いてモデルと入力画像の照合を行う。長所は見えのみでマッチングする手法と比較してより精密な形状推定をしやすいことである。短所はモデルベース構築に時間がかかることや、モデルを動かす際に複雑な計算が必要となり、発生する計算コストが大きく速度が遅くなりやすいことである。文献[7]では 2.16GHz デュアルコア CPU で 10.2 [fps]であった。Erol et. al が 2007 年にまとめたサーベイで紹介される 3D model-based 方式は 10 [fps]以下のものがほとんどであった[9]。コンピュータの性能は日々上がっているが、現在でも一般のコンピュータを用いた際にリアルタイムで使用することが困難と考えられる。

2D appearance-based 方式は入力画像から抽出された特徴量ベクトルと、参照データベース内に入っている特徴量ベクトルを比較することで最も類似している画像を見つけ、参照特徴量ベクトルとセットとして付随する関節角度などの姿勢情報を出力する方式である。この方式は、様々な手指形状をデータベースに格納すると探索に計算コストがかかるが、探索するデータは工夫により削減可能である。ゆえに、長所は高速化が期待できることである。例えば、谷本らは自己組織化マップを用いてデータベースをあらかじめ効率の良いデータベース探索ができるようにクラス化しておく手法[10]を提案し、また、過去の推定結果から探索範囲を絞ることで 2.8GHz のコンピュータで平均 150.09 [fps]、誤差の絶対値平均 8.132°を達成した。また、富田らは、画像の 3 つ

の形状比率情報を手掛かりに大まかな探索を行って似ている形状のデータのみを残し、その後  $64 \times 64$  [pixel]のサイズに縮小して高次局所自己相関関数を用いて特徴量化し、マッチングを行う手法[11]を提案した。この処理はノート PC を用いた場合で処理速度が約 80 [fps]で推定誤差平均  $-2.11^\circ$ 標準偏差 $\pm 14.14^\circ$ の推定精度を実現した。これはインタフェース化にも耐えうる速度であると言える。短所はシルエットや輪郭線などの概形情報を基にしたマッチングを行うため、複雑な形状の推定精度を高くすることが困難であるという課題を持つことである。

本論文ではユーザへの負担、および高速性を考え、2D appearance-based に着目する。具体的にはノート PC で操作ができた富田ら[11]の手法をベースとして新たな形状推定を考える。

## 2.2 肌・爪の構造や反射モデルに関する研究

肌は医学、生理学だけでなく、工学でも CG での肌表現の改善や化粧品の効果検証などを目的に研究されている。肌は大まかに表皮層、真皮層、皮下組織に分けられる。皮膚の色はメラニン、ヘモグロビン、カロチン、ビリルビンがあるが主に考慮されるのがビリルビン以外の3つの色素である。吉川によると、メラニンとヘモグロビンは部位によって量が異なる。メラニンが多いほど黒くなり、ヘモグロビンが多いほど赤くなり明度が下がる[12]。カロチンは透過のよい黄色の色素であり、部位ごとに大きな変化はないとされている。

肌色の分光特性や肌色の見え方については古くから研究されている。Buck et. al の研究では、分光反射率のグラフ形状はあまり変わらないが、男子よりも女子の方が、分光反射率が高いことや、身体各部の分光反射率は異なり、普段から日を浴びる顔と比較して、尻は中短波長の反射率が高いことを示した。また、日焼けをした肌や黄色・黒色人種と比較して、日焼けをしていない肌や白色人種の方が、分光反射率が高いことを示した[13]。加齢による色の変化は複雑である。吉川らは日本人女性について、美白が美とされてきた時代を過ごした 1980 年代生まれの人は、日焼けが美とされてきた時代を過ごしたそれ以前の人より有意的に黄みより、かつ高明度低彩度であることを示した。これは紫外線が有害であることが広く認知されて人々が避けるようになり、結果として日焼けにより生成されるメラニンの量が減ったことが原因と考えられている[14]。このように、色は時代背景が関わってくる。一方、肌の状態には性差や年齢差が現れやすく、鳥居らによると、顔画像について空間周波数のパワー値を用いて肌の明暗変化を調べることで判別可能である[15]。

文献[13]などで導かれた分光反射率と画像で得られる情報の結びつけも行われている。中井らは、肌画像に対する主成分分析の第1主成分が肌色、第2主成分が男女差、第3主成分が部位差を表す傾向があり、第3主成分には酸化ヘモグロビンの吸収スペクトルに似た性質が現れることを示した。また、カロチン、メラニン、ヘモグロビンを独立変数とした重回帰分析により肌色の質感表現を可能にした[16]。津村らは、画像の RGB 値から分光反射率を推定した後、独立成分分析 (ICA) を行うことでメラニンとヘモグロビンの色の影響を表す軸を生成し、色素の量による肌色変化の描画を実現した[17]。さらに、モンテカルロ・シミュレーションで導出した理論上の分光反射率に ICA を行い、画像から算出した ICA の結果と比較することで光散乱の影響を求め、その影響を除くことで更に正確な色素を表す軸を生成し、肌色表現を可能とした[18]。

Gupchup et. al によると、爪は爪甲、爪床、爪母の3つの要素でできている。爪母が爪甲を生成し、爪の根元にある白色部は爪半月と呼ばれる新しい爪である。爪床は表皮が無い肌の一部である。爪甲の主成分は無色透明な硬ケラチンである[19]。つまり、爪の色は表皮を除いた肌内部の色に光沢を含んだものとなる。爪はメラノサイト増加で黒色、カロチンの過剰摂取や栄養障害などで黄色、緑膿菌の日和見感染で緑色、糖尿病などで白色になることがある[20]。工学分野では指が何かを押した際に爪床が圧迫され、色が変化する構造的性質を利用することにより、指の押す力を推定する研究が主流として行われている[21]。爪領域を画像から抽出する研究は8件のみ確認できている。これらは本論文の主題と密接に関係しているため、4.1 節で述べる。

## 2.3 色テクスチャ分析に関する研究

画像全体のエッジ情報を見るものや、局所的な照合を用いるような研究について調査を行った。これらは陰影の分析、画像探索、そして物体の照合などで利用される。

画像全体のエッジ情報を見る代表的例の1つが、周波数解析である。Ramamoorthi et. al は砂利のような複雑なテクスチャに生じる陰影解析を目的とし、カラー画像に畳み込みとフーリエ解析を用いた[22]。まず対象の上部にカメラを設置し、光源を、半円を描くように複数回移動させながら複数枚の画像を取得する。そして各画像の輝度とヘビサイドの階段関数を畳み込み、フーリエ解析し、固有のスペクトルと輝度の固有モードを算出した。実験結果では、陰影と畳み込み公式には理論上の関係性があることが示された。Manjunath et. al はカラー画像検索システムの高精度化を目的として、ガボールフィルタにウェーブレット変換を組み合わせるテクスチャが類似する画像探索を行う手法、および事前に高速フーリエ変換で取得した周波数画像とデータベースにある画像の差分をとり、適するフィルタサイズを取得する手法を提案した[23]。これにより、従来のピラミッドやツリー構造を利用したウェーブレット変換より画像検索の精度が向上した。この研究は、後に Zhang らにより検索前に画像が持つテクスチャの主なエッジ方向を回転変換で揃える手法が付加され、回転に対応した[24]。周波数以外の方法として、Fu らは、ガボールフィルタに回転不変な特徴記述子である Zernike モーメントを組み合わせる画像検索システムを考案し、ガボールフィルタのみの検索より精度を向上させた[25]。

局所的な情報の積算からテクスチャ分析する手法は、特徴照合用のデータベースか学習データを持つ。まず Otsu et. al の高次局所自己相関関数[26]がある。この関数は対象が画像内のどこにいても特徴量変化がないという位置不変性と、特徴量の足し合わせが可能な加法性を持っている。この手法はガンの診断や顔の認識など様々な認識に利用されている。また、画像の局所領域で取得したエッジを特徴量とする手法としては、関心領域内の勾配と強度の関係をヒストグラムとして特徴量化する Histograms of Oriented Gradient (HOG)特徴量[27]や、エッジの勾配ごとの画像を生成して2方向のエッジ強度の比を特徴量とする Edge of Orientation Histograms (EOH)特徴量[28]、エッジの繋がりに着目し、連続するエッジ情報を特徴量とする Edgelet 特徴量[29]が挙げられる。さらに、2つのセル領域で作成した色ヒストグラムを特徴量とする Color Self-Similarity (CSS)特徴量[30]と HOG を組み合わせ、エッジ勾配に色の類似度情報を含めるようにした CS-HOG 特徴量[31]のように組み合わせによる特徴量も多数提案されている。

近年の注目は Bag-of-keypoints (BoK)という概念[32]である。これは、自然言語処理分野で重要な Bag-of-Words を画像処理に応用したもので、1つの特徴量算出フィルタで得られた特徴量を単語、複数のフィルタを用いて得られた別々の特徴量を束ねたものを辞書と定義して関連付けを行う。この手法は現在も確率モデルやカーネル関数などを用いてクラスタリングによる認識精度向上を図る研究が盛んに行われている。

## 2.4 数理的手法を利用した色物体認識に関する研究

数理的手法では、画素の情報をそのまま用いるのではなく、ヒストグラムや超空間座標への射影、確率等、数学の理論を用いる。

前景抽出は特に重要な物体認識技術である。前景と背景で明らかに色や輝度が異なる単純な分割問題では、Otsuのように2峰性ヒストグラムの山が最もきれいに分かれる閾値を判別分析法により統計的に取得する方法[33]がよく用いられる。一方、公園など複雑背景で環境変化が小さい場所での移動物体検出は、2次元濃度ヒストグラムという、背景のみの画像と、対象物を含む画像の差分変化を統計的に表す手法が利用できる。喜多はこの手法で画像間の色の変化が大きい領域を抽出し、その領域の大きさと色の勾配変化から重要変化か判別して移動物体を抽出した[34]。この手法は背景のみの画像と検出対象の画像の2枚のみ要求するため、例えばカメラを設置してすぐに利用できるなど、経済性が良い。環境変化が大きい場合は学習を用いる。木の枝葉の揺らぎなど、背景変動にパターンがある場合は、画素分布をカーネルトリック法により超空間を扱うことで、背景領域の画素分布を単純化でき、外れ値を持つ前景領域画素を抽出しやすくなる。川端らは、さらに主成分分析によりノイズの影響低減と情報圧縮を行い、精度向上と高速化を実現した[35]。さらに道路などの不確定要素の大きい場所でも安定した前景検出を可能にするために、統計的に外れ値となる値の情報の重みを小さくするM推定を用いる手法が、島井らによって提案された[36]。

1枚の画像に写る様々な物体の領域を分割する一般物体認識に関する研究では、確率やエネルギーを利用する手法が用いられる。Comaniciuは観測点から任意の半径を設定し、その半径内に存在する画素の分布に関して確率密度を考え、重心点を探索することで、クラスタリングを行うMean-shift法を提案し、領域分割を達成した[37]。同じクラスタリングとして有名なものに、クラスタの重心位置を初期設定した後、画素のクラスタ分けとクラスタ重心の計算を位置移動が収束するまで繰り返すことでクラスタ分割するk-means法がある[38]。この手法は近年、領域分割にソフトコンピューティングの知識を取り入れ柔軟性を増加させたFuzzy c-means法[39]などが開発されている。その他、マルコフ確率場を発展させたMarkov field aspect model [40]や混合ガウスモデルに確率伝搬法（ベータ近似）を用いてパラメータを極力減らしてクラスタリングによる領域分割を実現した手法などがある[41]。

一方、エネルギーを利用する手法として、Arbeláez et. alは向きを考慮したエッジ勾配の取得を行った後、エッジの強弱を山と谷と見立てて山と山の間に分割線を引くWatershed法を適応することにより不特定多数の画像に対して領域分割ができる手法を提案した[42]。また、Koykov et. alはグラフ理論を用いて各画素をノードで結び、ノードを切断したときに損失するエネルギーを評価することで最適な境界を割り振るGraph cut法を用いて複雑背景下での特定の物体の領域取得やCT画像の正確な画像領域分割などを可能とした[43]。

## 2.5 照明変動や影がある場合の色物体抽出に関する研究

画像処理に適する画像を生成し、認識のロバスト性を向上させることを目的とした色補正や影の除去手法を調査した。画像認識分野で、色補正は照明変化による色の変動に強くする目的で使用される。一方、影の除去は防犯カメラに映った人の体の領域を正確に抽出するためや、道路の白線等を正確に取得することが目的とされている。

まず、画像認識で問題となる色飽和をなくす手法として、ハイダイナミックレンジ (HDR) がある。HDR は低露光の色がつぶれるが画像色飽和が無い画像と、高露光で色相や彩度を取得できるが色飽和のある画像を組み合わせることで色が飽和しない画像を生成する手法[44]が有名であり、カメラの機能として実装されている。

局所的な影に対応するには、物体が持つ表面反射率を計算し、物体がもつ固有の色を取得する手法を用いる。表面反射率の求め方はカメラの画素値の決まり方を考慮する方法と、人間の視覚特性を考慮するものがある。カメラが取得する色は画素値がセンサの分光感度、表面反射率、光源の分光強度が乗算されたものとなる。センサの分光感度は既知である。そのため、光源の分光強度を取得し、その後、表面反射率を求める。光源の分光強度取得は分光測色計で直接取得する手法、撮影されたカラーチェッカーの多数の色の変化から取得する手法[45]、大きさが既知の直方体を撮影して色や影の付き方から光源方向と光源色を推定する手法[46]などがある。Finlayson et. al は輝度の色合いの変化具合から影の部分と前景を分け、背景から光源の分光強度を推定し、影を画像から消去することに成功している[47]。人間のものの見方を考慮する方法は Retinex 理論[48]と呼ばれ、明るさは反射率に光源の明度を乗算したものと考える。輝度は通常滑らかに変化するため、低周波数成分を輝度と仮定できる。広範囲の画素を対象とするガウシアンフィルタを複数パターン使い、低周波成分を抜き出す Multi-Scale Retinex (MSR) [49]やより高速な 2 次計画法 (QP 法)[50]などが提案されている。

抽出対象のテクスチャに普遍的特徴がある場合、学習により、照明変化や影に対し、ロバストにする手法も存在する。例えば Viola et. al によって提案された、いくつかの基本分類器を組み合わせるカスケード型分類器[51]がある。文献[51]では目、口など特定部位周辺に生じる大きな明暗差を見る Haar-like 特徴を基本分類器として示した。この特徴は局所的な明るさの大きな差を見るため、照明変動に影響しにくい。道満らは道路標識の検出のために、経年による退色、反射などを含む様々な色条件を基本分類器にした。これにより輝度のみ利用した手法より必要なサンプル数が 96%削減できることが報告されている[52]。また、多数の学習サンプルを主成分分析し、統計的に似ているものを抽出し、それを識別のためのデータとして用いる手法がある。例えば固有顔[53]による顔画像認識が有名である。固有顔は大量のサンプルにより照明成分にもロバストにできる。照明成分を元の画像と低次元固有画像から推定し、それに直交なベクトルを求め、固有顔を直交分解することでよりロバスト性を高める方法も提案されている[54]。これらの学習サンプルは、サンプルごとの位置ズレが少なくなるよう必ず正規化する。

## 2.6 単眼画像からの 3 次元形状および姿勢推定の研究

単眼画像を用いて 3 次元形状・姿勢の推定を行う場合、2 次元情報から 3 次元推定を行う不良設定問題が存在する。この問題を解決するには、何かの情報を付加する必要がある。

2 次元の見えをデータベース化して利用する 2D appearance-based マッチングで良く用いられるものとして、3 次元パラメトリック固有空間法を用いて姿勢を推定する手法[55]が挙げられる。固有空間上では物体は回転角を媒介変数とした一つの関数として表され、入力画像の情報を固有空間に点として写せば、最近傍に存在する関数値を取得することで物体形状と回転角が同時に推定できる。別の手法として、天野らは学習を行う際に、物体の姿勢情報などを表す情報トラックを学習サンプルに付加する EbC 法を用いた 3 次元姿勢推定手法を提案した[56]。この手法は欠損を含む 1 枚のから画像のみから固有空間により画像記述ルールを抽出し、画像補完を行う手法である BPLP[57]を応用することで情報トラックを取得する。画像補完とパラメータ推定の計算を 2 枚の画像 (EbC 画像対) を用いて 2 回の画像の内積演算と 1 回の簡単な三角関数演算で実現できるため、高速性がある。また、固有空間からサポートベクトル回帰を用いた推定式を用いて姿勢パラメータを求める手法も提案されている[58]。手画像の場合、各指の関節角度などを考えると、非常にパラメータ数が多くなるため、上記の手法ですべてを推定するのは困難である。よって実際に使用されている手法は富田ら[11]のように、輪郭線の特徴量を比較用、求めたい関節パラメータを出力用として一括りにしたデータセットを各姿勢で作成し、まとめたデータベースを基にマッチングを行う手法となる。

三次元モデルを活用する 3D model-based マッチングでは、モデルを変形させてシルエット画像を求め、2D appearance-based マッチングのようなフィッティングを行う手法と直接当てはめを行う手法が存在する。ただし、本論文では単眼カメラの 2 次元画像を扱うため、シルエットを用いる方法に限定して述べる。Lowe はエッジの共線性、並行性、近接性に着目し、モデルのワイヤフレームモデルを逐次行いながら対象画像のエッジ画像のマッチングを行うことでマッチングを行う手法を提案し、重なり合ったカミソリのうち、前面に出ている 4 本について姿勢を取得することに成功した[59]。トラックのように特徴のある幾何形状を持つ場合には、その幾何をエッジ画像から抽出し、モデルと照合することで高速化が可能である[60]。

特色のある手法としては、カメラを複数にするのではなく、照明の光源を複数にして光の当て方により変化する輝度の変化を基に 3 次元物体の形状を復元するという手法が挙げられる[61]。文献[61]の研究では復元までで、姿勢の推定は行っていないが、対象物体の 3 次元モデルがあれば、推定したデータとの 3 次元マッチングによる姿勢推定が可能になると考えられる。しかし、照明を意図的に変える手法は特殊環境であるため、ジェスチャインタフェースのように一般環境で使用を想定する場合には利用が難しい。

# 第3章 爪を用いた手指形状推定の構想

## 3.1 手指形状推定システムの概要

本論文で形状推定システムのベースモデルとする富田らが提案した手指形状推定システム[11]の概要を述べる. 図 3-1 はシステムの簡略的な概要である. 入力の手画像, 出力は指 20 自由度 (屈曲伸展と開閉動作), 手首 2 自由度, 前腕回旋 1 自由度の合計 23 自由度である.

手画像を入力後, 2 値化する. 次に, 手首より下の腕領域に着目し, 輪郭の傾きから手の傾き角度を算出して回転の補正をする. 次にシルエットを距離変換し, 手領域の基準点と, 基準点の輪郭線からの距離を取得する. そして, 基準点からその距離分のピクセルだけ下げた位置を手首位置と定義し, 手領域の下端とする. 上端, 左端, 右端は手領域の最大座標値で定める. 手領域のみを切り取った後, 輪郭線のみを残し,  $64 \times 64$  ピクセルに正規化する. この画像がデータベースマッチング用画像となる. この画像から算出される画像形状比率と画像特徴量をデータベース照合し, 最も近い特徴量を持つデータセットを選択し, 23 自由度のデータを出力する.

データベースはマッチング用の画像形状比率, 画像特徴量と, 出力するデータ列で構成されたデータセットの集まりである. データセットは形状推定を行う前にデータグローブを用いて作成する. 形状比率と画像特徴量は右手をカメラ撮影した画像を基に算出される. 出力用データは

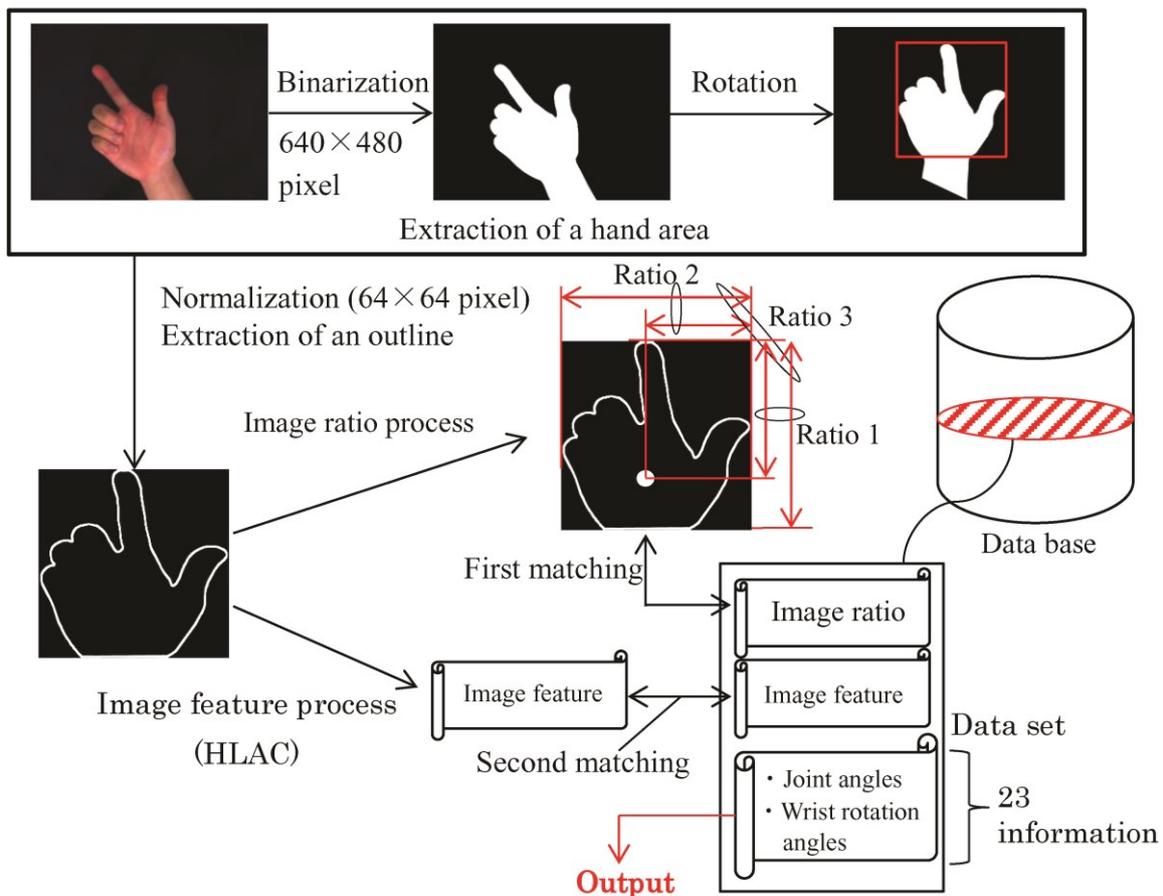


図 3-1 本論文で対象とする形状推定システムの概要図

左手にグローブをはめ、右手と同じ動きを左手で行うことで取得される。

画像形状比率は手領域をちょうど囲む矩形と基準点の関係をを用いることで表現する。マッチングで使用する。上端－基準点間距離と上端－下端間距離の比率  $r_{top-heavy}$  を上長さ、右端－基準点間距離と左端－右端間の距離の比率  $r_{right-biased}$  を右長さ、矩形枠の周の長さの半分と上端－左端間距離の比率  $r_{tall}$  を縦長さとして、定義する。そして、以下の3つの式(3-1)、式(3-2)、式(3-3)が全て満たされたとき、参照データセットは入力画像と類似した性質を持っていると判定する。

$$th_{tall} > |R_{tall}[i] - R_{tall}| \quad (3-1)$$

$$th_{top-heavy} > |R_{top-heavy}[i] - R_{top-heavy}| \quad (3-2)$$

$$th_{right-biased} > |R_{right-biased}[i] - R_{right-biased}| \quad (3-3)$$

ここで、

$R_{tall}[i]$  : データセット番号  $i$  の縦長さ

$R_{top-heavy}[i]$  : データセット番号  $i$  の上長さ

$R_{right-biased}[i]$  : データセット番号  $i$  の右長さ

閾値は通常、0.2 程度の値が用いられる。形状比率に含まれる情報は大まかな形状情報のみである。しかし開いた手と閉じた手など、明らかに形状が異なるものは判別できる。また、この情報は高速で取得できる。ゆえに、この情報は巨大なデータベースから本当にマッチングすべきデータセットをふるい分ける用途で使用され、本手法によって通過したデータセットのみ、次に述べる画像特徴量による類似度判定を行う。

画像特徴量は、Otsu の高次局所特徴量(HLAC) [26]を応用して、特徴量化したものを利用する。まず、図 4-2 のように、 $64 \times 64$  ピクセルの正規化画像を  $8 \times 8$  ピクセルの 64 ブロックに分割する。HLAC は 2 次、25 パターンまでを用いる。ゆえに HLAC 特徴量は合計 1600 次元となる。評価関数  $E[j]$  の式は式(3-4)となる。

$$E[j] = \sum_{l=1}^8 \sum_{c=1}^8 \sum_{h=1}^{25} (x_{current_{lch}} - x_{dataset[j]_{lch}})^2 \quad (3-4)$$

ここで、

$j$  : データセット番号

$l, c$  : 分割画像番号

$h$  : HLAC パターン番号

$x_{current_{lch}}$  : 入力画像の画像特徴量

$x_{dataset[j]_{lch}}$  :  $j$  番目のデータセットの保持する画像特徴量

評価関数は値が小さいほど類似していることを表す。よって、ふるいわけが残ったデータセット

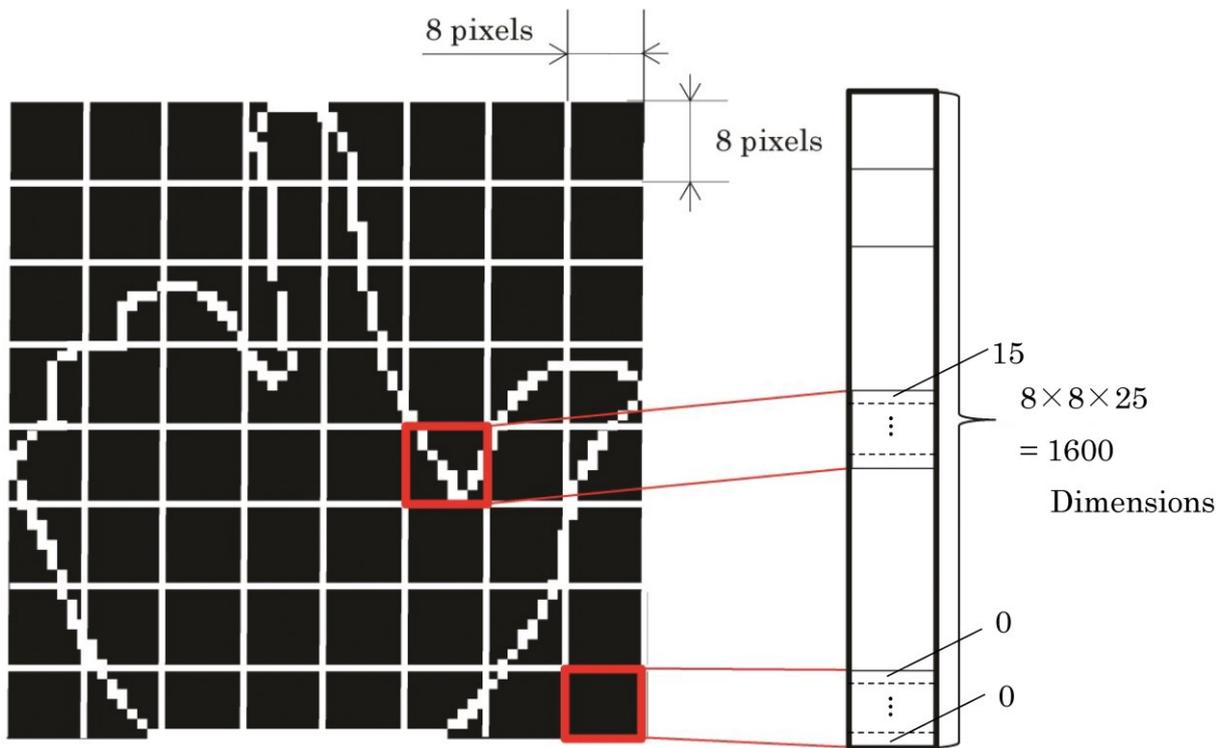


図 3-2 HLAC 特徴量を使用したデータマッチング用 1600 次元ベクトルの算出

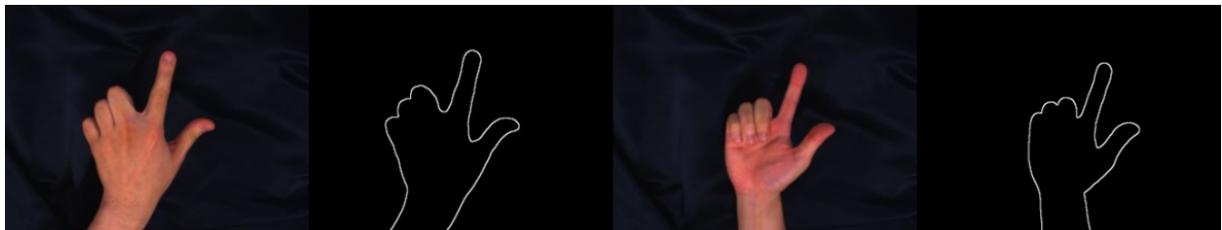
との比較で、式(3-4)の右辺が最小になったデータセットを最類似形状のデータセットと判定し、特徴量に関連づけられた 23 自由度のデータを出力する。

## 3.2 手指形状推定システムの抱える問題と解決案

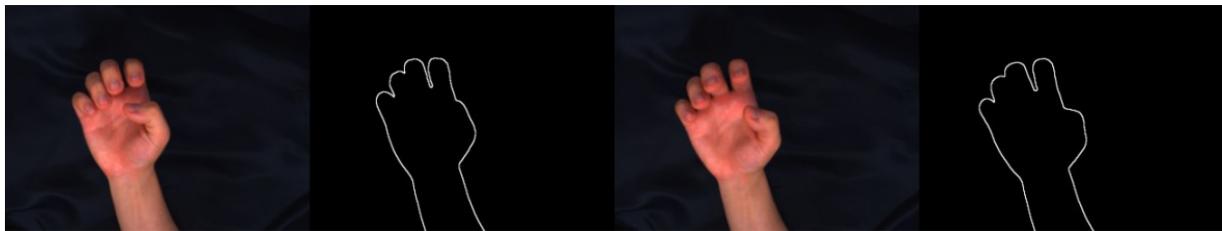
富田らの手法[11]は図 3-3 と図 3-4 に示す 6 種類の形状の違いを区別することが困難である。図 3-3 は輪郭線形状が類似して判別できないことが主原因であり、図 3-4 はマッチングに用いる HLAC アルゴリズムが持つ性質に主原因がある。また、ブロック分割をしてから HLAC を使用する手法では、推定精度のブレを生じさせる別の問題も存在する。以下では、これらに関して 3 つの項に分けて述べる。そして、それらを解決するための案を提示する。

### 3.2.1 輪郭線のみを用いた特徴抽出における問題

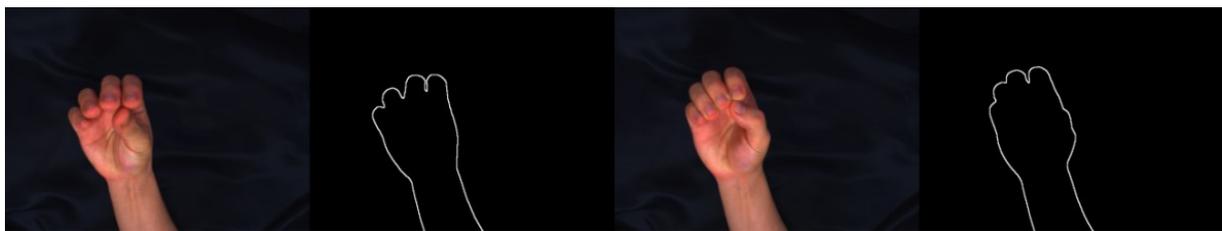
輪郭線のみを利用する手法の問題は手のシルエット内部に存在する指の位置情報が失われてしまうことである。図 3-3 に示すように(a)手の裏表、(b)握りこみの深さの違い、(c)つまみの深さの違いは輪郭線形状には反映されず、類似度が高い。また、図 3(d)の指同士の遮蔽を起こした



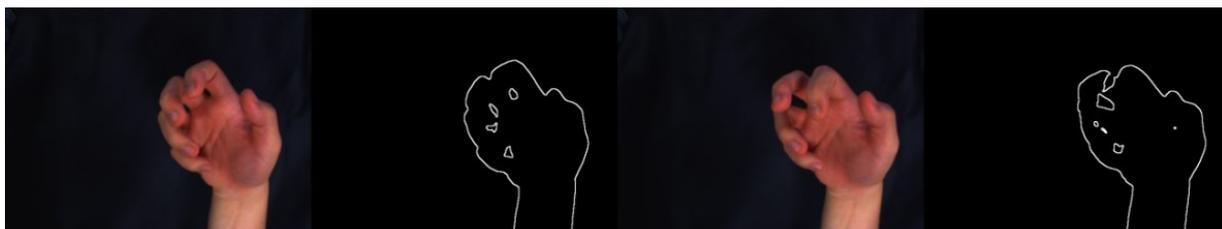
(a) 手の裏表



(b) 握りの強さの違い

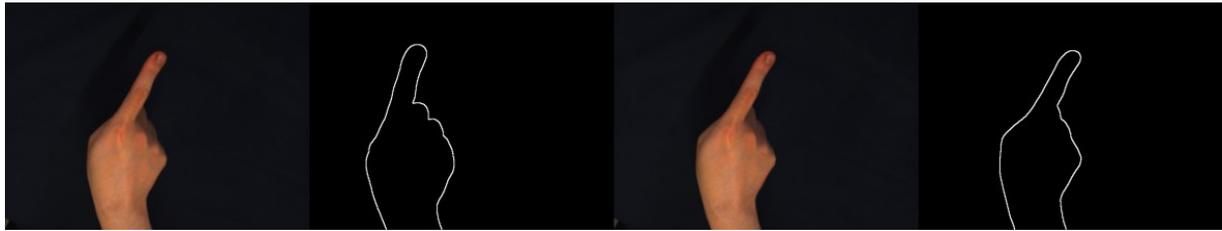


(c) つまみの強さの違い

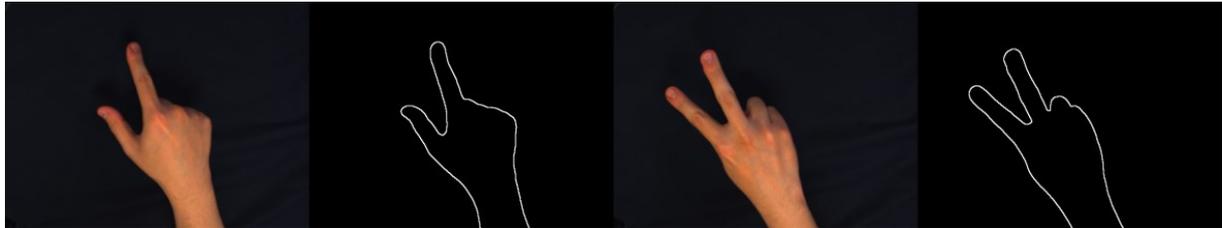


(d) 指同士の遮蔽

図 3-3 輪郭線形状が似ていることが原因で誤推定を起こす例



(a) 指の傾きの違い



(b) 指の構成要素の違い

図 3-4 HLAC アルゴリズムが持つ性質によって誤推定を起こす例

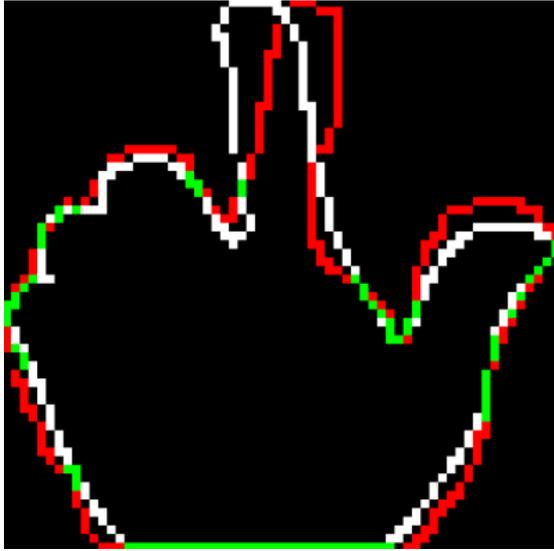
画像では、指と指の間のわずかな隙間や、局所的な影の影響で内部にも輪郭線が現れ、形状差が出ている。しかし、この差は個人差、個人内差でも形状は変化すると考えられる。また、この輪郭線は指の位置を表すものではなく、指の先端がどこにあるかを輪郭線画像だけでは確認できない。精度の良い推定を実現する大前提は、手の形状を示す情報を復元するのに必要な情報を取得できる事と、それをロバスト性のある特徴量化によって数値にすることである。シルエットの輪郭線だけではその大前提を崩してしまう。ゆえに、輪郭線以外の情報を活用する必要がある。

### 3.2.2 マッチングに用いる HLAC アルゴリズムが持つ性質に関わる問題

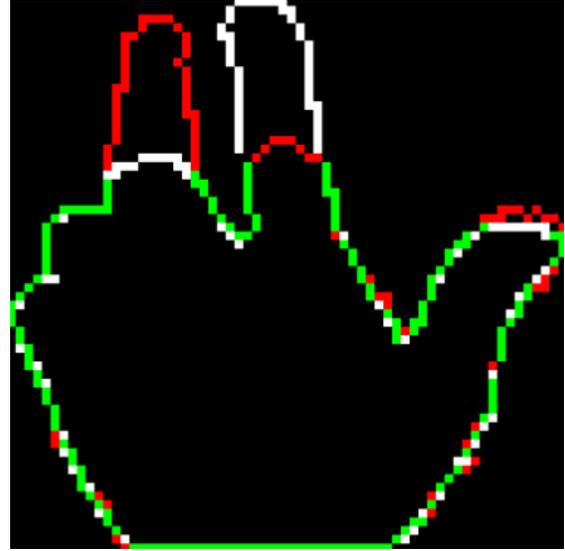
図 3-4 の 2 つの図のように、指の向く角度や指自体を間違える問題は HLAC アルゴリズムが持つ性質に関わる問題として挙げられる。富田らのシステムでは HLAC による値のずれの積算は全て同じ重みで行われおり、これが原因を引き起こす。例を図 3-5 に示す。白が入力画像の輪郭線の画素、赤が参照したデータセットの、特徴量化前の輪郭線で入力画像とずれている画素、緑が入力とデータセットの輪郭線が一致している場合である。図 3-5 (a) のように同じような形状でも撮影距離や個人差などで手の大きさが異なるなどの理由で大きなズレが生じることがある。これが生じると、類似度が低くなる。逆に、図 3-5 (b) のように、違う形状だが多くの輪郭線が一致してしまい、立っている指が異なるにもかかわらず総合的な類似度が高くなる時もある。これが指を間違える原因、そして指の位置ズレを引き起こす原因ではないかと考えられる。

### 3.2.3 ブロック分割をしてから HLAC を使用する手法に関わる問題

この問題は推定が困難な形状を生じさせる主要原因を作るものではなく、推定のブレや、3.2.2 節で述べた推定ミスを生発する問題である。ブロック分割を行うことにより生じる問題は輪郭線またぎである。画像は高性能なカメラで撮影したのものでも、背景と前景の境界に存在するエッジ付近で取得が不安定となる傾向がある。その影響で、輪郭線は画像ごとに絶えず数 pixel 単位で振動的に移動する。それは正規化画像にした際にも影響を及ぼし、一部の輪郭は  $k-1$  番目のブロ



(a) 輪郭線の一致が少ないが似ている画像



(b) 輪郭線の一致が多いが似ていない画像

図 3-5 HLAC を用いたアルゴリズムにより生じる問題の概要図

ックと  $k$  番目のブロックを行き来する現象が生じる。富田らの手法では各ブロックは独立処理される。そのため、輪郭線を含むブロックの特徴量は絶えず変化する。これが推定のブレに繋がっていると考えられる。

### 3.2.4 各問題に対する解決案

3.2.1 項から 3.2.3 項で述べた問題に対してどのような解決ができるかを考え、案を提示する。

1 つ目の輪郭線に情報が含まれていない指の位置が不定となり、推定が不安定になる問題には、爪情報の付加が重要と考えられる。輪郭線から情報が欠落しているが、カメラにその指が写っている状態では、深く握りこんでいない限り、その指の爪が画像に表れる。爪は指の先端を表すため、その指の爪を指の先端位置として利用することができる。手は幾何拘束が大きいので、指の先端位置が判明すれば形状の特定が容易となる。ゆえに、爪の情報を推定時に適切に反映させることでこの問題は解決しうると考えられる。ただし、付け爪をユーザに要求するとユーザへの負担を増加させるため、マーカレスで爪検出できることが重要である。なお、手の裏表に関しては、手の甲側と比較して掌側にはシワや手相などが多数存在し、かつ手の甲側の方が掌側よりメラニンが多いため黒くなることから、手領域全体に対して色やエッジを調べることで対応できる可能性がある。

2 つ目の輪郭線の情報が均一であることが原因で生じる問題は、爪情報だけでは解決できず、輪郭線の重み自体を変化させる処理が必要になる。指の曲げ角度を間違える問題は爪が見えず、指腹側が写る場合にも生じるためである。ゆえに、爪情報の活用方法だけでなく、輪郭線の情報にも重みをつけて対応する必要がある。輪郭線の中で特に重要なものは、手の形状を特定できる急激な変化を起こす、指の先端を表す曲率の大きい部分である。ゆえに、曲率に応じて重みを変化させるような処理を加えるべきであると考えられる。

3つ目の画像ごとに輪郭線がブロック間を行き来する問題は図 3-6 のように、またぎやすい位置にある画素を対象として、またいだ際に移動する位置にも画素を付加してしまうという処理をデータベース、入力画像ともに行うと解消できると考えられる。

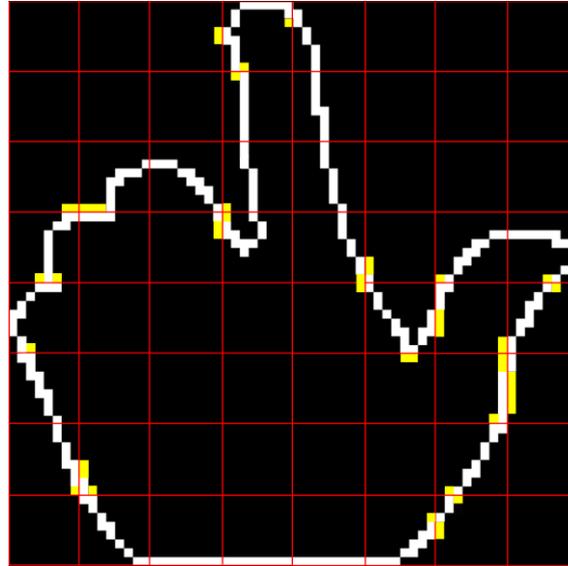


図 3-6 画素の付加処理によるブロックまたぎの防止

### 3.3 爪を用いた形状推定システムの先行研究

爪を利用した手指形状推定は、これまで玉城ら[62]と佐野ら[63]が試みている。それぞれのアルゴリズムの概要を図 3-7、図 3-8 に示す。両研究とも、データベース作成時と評価実験時には右手にデータグローブをはめ、その上に白い手袋、色分離が可能な単色の付け爪をつけていた。玉城らの手法では爪領域抽出法も提案されていた。これについては次章にて述べる。玉城らの研究と、佐野らの研究の大きな違いは爪情報の表し方と爪情報を単独で使用するか輪郭線と同時に使用するかの違いである。玉城らの手法では爪の情報は、爪領域の重心が持つ座標として定義し、輪郭線情報と爪情報を一つの評価関数でまとめ、マッチングを行った。一方、佐野らの手法では、爪の情報は爪領域重心の所属するブロックの番号(画像を 36 分割して番号付したもの)とした。データベースマッチング時には先に、爪のブロック番号が完全に一致するものを探索してデータのふるい分けをし、通過したデータセットのみ輪郭線情報と照合した。玉城らの手法では精度向上が見られ、佐野らの手法では、示指の PIP 関節は推定精度が良いものの、拇指 CM 関節の追従がうまくできなかったことを報告している。また、佐野らの研究では、爪に関するブロック数を増加させると、真値との距離が増加、つまり推定精度が悪くなることを示した。この原因の 1 つとして、爪の位置をブロック番号で表したことが挙げられる。爪の位置をブロック番号とすると、3.2.3 節と同様、ブロック間を重心が行き来して番号の変化を起こす現象が生じる。佐野らの使用法ではブロック番号が完全に一致したデータセットでなければ推定結果にならない。データセットの粒度が大きければ、類似している形状がブロックまたぎ問題によって通過できない可能性が

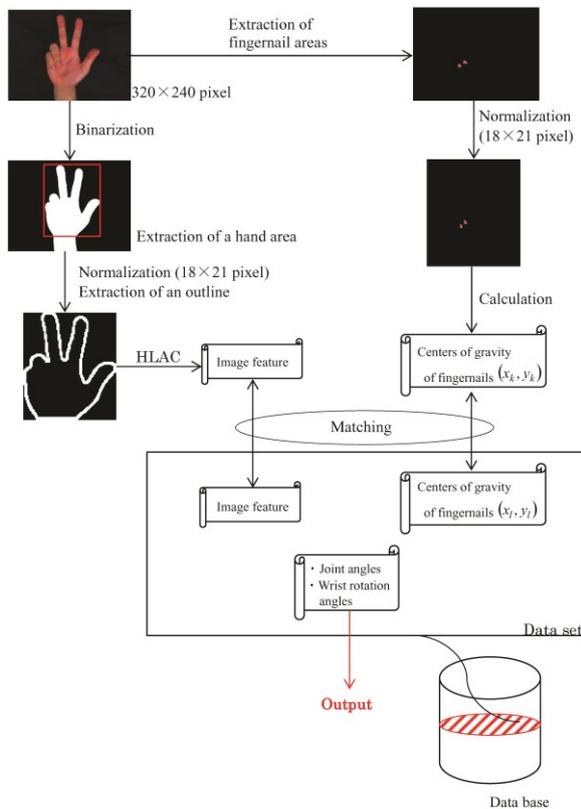


図 3-7 輪郭線+爪情報による推定

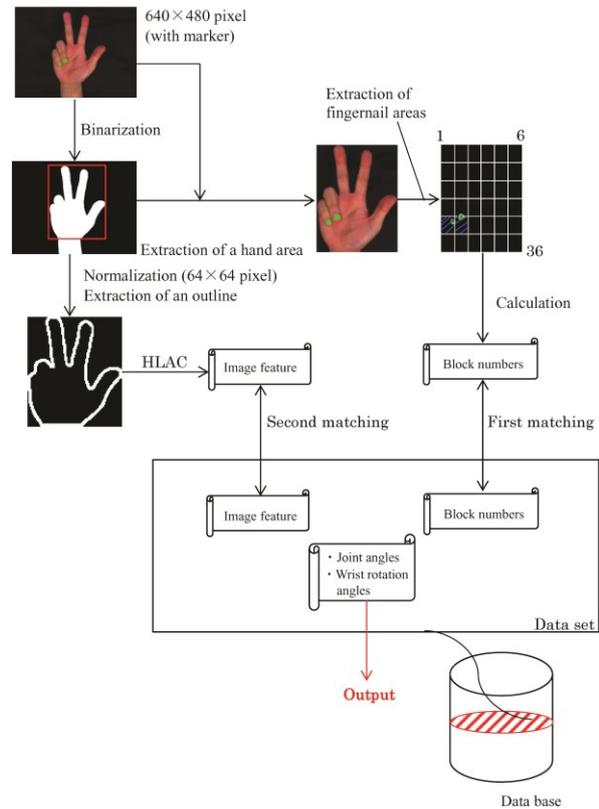


図 3-8 爪をふるいわけに使用した推定

ある。つまり、正しい推定ができなくなる可能性があるという事である。データベースを多く作成しても形状によってはブロック移動をしたものと、していないもの両方のデータが取得できない可能性がある。このことから、爪を利用する際には情報をブロックという単位で表さない方が良いと考えられる。

### 3.4 爪の使用方法に関するリスクの検討

爪情報は指先を定めるための強力な情報と考えられるが、同時にリスクも抱える。特に大きなリスクは前腕回旋がかかり、爪領域が小さくなった結果、抽出できなくなるリスクである。

爪の情報のなまし方は玉城ら[62]のように、輪郭線の情報と爪の情報を評価関数という形で結び付け、一つの特徴量でマッチングする手法を用いる。ここで、玉城らは式(3-5)のように、爪の重心位置のユークリッド距離誤差と、輪郭線のユークリッド距離誤差を同じ重みで加算していた。

$$E(x) = \sum_{k=1}^{64} \sum_{l=1}^{25} (p^{kl}(x) - P_{input}^{kl})^2 + \sum_{m=1}^M \|\mathbf{r}^m(x) - \mathbf{r}_{input}^m\|^2 \quad (3-5)$$

ここで、

$p^{kl}(x)$  : データベース内の  $x$  番目のデータセットの  $k$  番目のブロックに関する HLAC 番号  $l$  の特徴量

$P_{input}^{kl}$  : 取得画像の  $k$  番目のブロックに関する HLAC 番号  $l$  の特徴量

$M$  : 爪の数

$\mathbf{r}_m = (x_m, y_m)$  : データベース内の  $x$  番目のデータセットの  $m$  番目の爪の位置

$\mathbf{r}_{input}^m = (x'_m, y'_m)$  : 取得画像の  $m$  番目の爪の位置

この式(3-5)は改善すべき点が2つある。1つは、爪は1つの情報しか無く輪郭線は多数の情報があるため、そのまま加算すると明らかに輪郭線の影響の方が大きい。つまり、爪の情報の影響が弱い。ゆえに、輪郭線情報の和に対して、輪郭線の数を除することで平均化するべきである。2つ目は爪の枚数が異なる時の対応を明確に定義していないことである。文献[62]には、爪が存在しないとき、仮の座標(-1,-1)を付加することが定義されている。しかし、入力画像に含まれる爪が1枚で、参照するデータベースの爪が2枚の時には言及がない。式(3-5)の爪の枚数  $N$  は入力画像基準か、データベースに含まれる爪情報の基準か明確な記載はない。入力画像基準の場合は、爪の枚数以上のマッチングを行わないため、入力画像以上の爪がデータベースに含まれていても、入力画像の爪の位置付近に爪が存在すれば類似度が高くなる。つまり、爪が多い時の対応が甘くなる。一方で、データベース内の爪の枚数を基準とすると、入力画像に爪の検出ミスがあった場合、1枚以上の爪が明らかに離れた爪とマッチングする現象が生じる。つまり、爪検出の精度がほぼ100%になることを望むシステムとなる。しかし、これを面積が小さい爪で実現するのは困難である。爪をマーカレスで抽出することを考えると、肌と爪の色が似ているため、さらに困難になると言える。つまり、どちらの爪の枚数を基準としても、問題が生じうる。これらに対応するためには、爪に対しての式はもう1つ項が必要となる。

最も重要なことは、爪の重心位置は指先を明確に示す重要な情報だが、取得できないリスクも持っていることである。つまり、爪の情報の重みを強くしながら、取得が難しいものは重みを低くするような対策が必要となる。

### 3.5 爪を用いた手指形状推定システムの構想

ここでは、これまで述べた問題点や検討を踏まえ、新しい、爪を用いた手指形状推定システムを構築するための構想について説明する。はじめに、システム全体の概要について述べる。次にアルゴリズム成立に不可欠な、データベース作成方法およびマッチング方法について述べる。最後に、構想システムが要求する爪検出システムの要求仕様について述べる。

#### 3.5.1 構想システムの概要

構想システムの概要を図 3-9 に示す。このシステムでは輪郭線処理する過程と、爪を検出して重心を求める手法を最後のマッチングまで分離する。ゆえに、デュアルスレッドで処理をすれば処理の効率化が図れる。

初期探索として画像形状比率による大まかな形状の違いによるふるいわけをかけ、第2探索でHLACによる画像特徴量と、各爪領域の重心座標を基にマッチングを行う。構想システムの最も重要な所は第2段階のマッチングの際にパラメータを変更するデータをデータセット内に含める

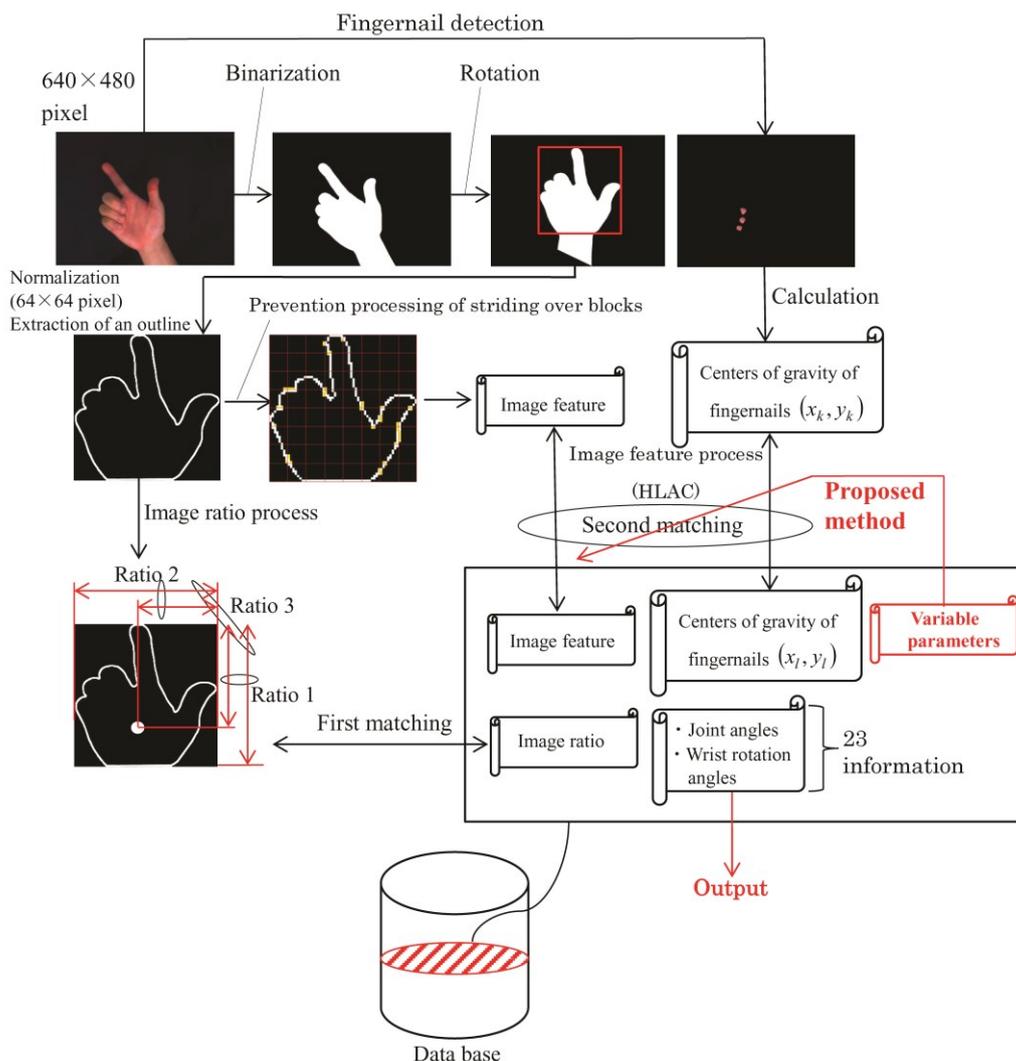


図 3-9 構想システムの概要図

ことを提案することである。データセットごとに変更するパラメータは HLAC の加算時の重みと爪情報の重みの係数である。

各データセットの作成時には、手に白い手袋と緑色の付け爪をつけたものをはめた。これにより、データセット作成時に手領域と爪領域を確実に取得する。基本的な処理は図 3-9 の処理からマッチングの過程を抜いたものと等しい。基本的な処理として、輪郭線画像のブロック分けと爪領域の取得をした後、可変パラメータの設定を行う。設定方法は次項で説明する。

### 3.5.2 評価関数と可変パラメータ設定方法の提案

3.2 節から 3.4 節までに述べた内容を精査し、新たな評価関数を提案する。また、評価関数に含むパラメータの設定方法について述べる。これまでに述べた重要な内容を箇条書きで示す。

- (a) 輪郭線で得られない指先情報を爪情報で補足することで、形状を特定できる
- (b) HLAC では、すべてを平等の重みで加算しているため、指の位置が全く異なるが、多くの輪郭線がマッチングしている画像 A が、指の位置が近いが、手の外形とマッチングしない部分が多い画像 B より類似度が高いとして推定結果になることがある。
- (c) 爪領域が大きく、取得しやすい爪の重みは強くすべきであるが、爪領域が小さく、取得が困難なものは重みを弱くすべきである。

(a)と(c)は爪情報の重みを変える問題であり、(b)は HLAC が均等重みであることを問題視している。ゆえに、HLAC と爪情報について重みづけができれば、この問題を解決に導くことができる。

まず、HLAC の問題について考える。これがうまく機能しない最大の原因は、手の個人差や個人内差が出やすく、手指姿勢に必要な情報をあまり含まない指以外の領域も、姿勢を知る上で重要な指先の輪郭線情報と同じ重みを持つことである。本来ならば、指先情報の重みを重くすべきである。著者らは重要か否かを判別する条件として曲率を用いるべきと考えている。手姿勢を知る手掛かりは急峻な変化を起こす部分であるからである。データベース作成時に曲率に応じて重要度を定める事で優位性のある情報の重みを高くする。一方、爪情報の重みは、データベース作成時に抽出された各爪の領域の大きさをパラメータとしてシグモイド関数で算出する。以上のことを踏まえた新しい評価関数は以下の式(3-6)となり、パラメータは式(3-7)と式(3-8)となる。

$$E(x) = \frac{A}{64} \sum_{k=1}^{64} \alpha_k \left( \sum_{l=1}^{25} |p^{kl}(x) - P_{input}^{kl}| \right) + \frac{1}{M} \sum_{m=1}^M \beta_m \| \mathbf{r}^m(x) - \mathbf{r}_{input}^m \| + \gamma(N-M)\delta_{MN} \quad (3-6)$$

$$\delta_{MN} = \begin{cases} 1 & \text{if } M < N \\ 0 & \text{else} \end{cases} \quad (3-7)$$

$$\alpha_k = \alpha \left( \theta_k^{mean} \right) = \frac{\alpha l}{\theta_{max}^{mean}} \theta_k^{mean} + (1 - \alpha l) \quad (3-8)$$

$$\beta_m = \beta(S_m) = \frac{1}{1 + \exp(-\beta_1(S_m - \beta_2))} \quad (3-9)$$

ここで、

$A$  : 輪郭線と爪の重みを調整するための定数パラメータ

$\alpha_k$  :  $k$  番目のブロックに関する曲率によって設定された輪郭線情報の重みパラメータ

$\beta_k$  : 爪領域の大きさに基づいて設定された爪に関する重みパラメータ

$\theta_k^{mean}$  :  $k$  番目のブロックに属する輪郭線が持つ曲率の平均

$\theta_{max}^{mean}$  :  $\alpha_k$  の値を最大にする曲率

$\alpha_1$  : 曲率の有無によって重みを変えるためのパラメータ

$S_m$  :  $m$  番目の爪の面積

$\beta_1$  : シグモイド曲線の勾配を決定するパラメータ

$\beta_2$  : 関数値が 0.5 になる時の面積値

式(3-6)の第 1 項は HLAC 特徴量の差分に関する項、第 2 項は爪の重心座標の距離に関する項、第 3 項は入力画像の爪の個数に対してデータセットの爪の個数が多い時に与えられるペナルティの項である。本論文では、ペナルティの大きさを決める定数を  $\gamma=10000$  と大きくとった。輪郭線番号  $i$  の画素がもつ曲率値  $\theta_i$  は丸谷らが定義した曲率[64]に絶対値を付加した式(3-10)を使用した。

$$\theta_i = \left| \tan^{-1} \left( \frac{y_{i+j} - y_i}{x_{i+j} - x_i} \right) - \tan^{-1} \left( \frac{y_i - y_{i-j}}{x_i - x_{i-j}} \right) \right| \quad (3-10)$$

ここで、

$j$  : 輪郭線上の注目画素と参照画素の間隔

また、

$$\| \mathbf{r}^m(x) - \mathbf{r}_{input}^m \| = r_{lim} \quad \text{if} \left( \| \mathbf{r}^m(x) - \mathbf{r}_{input}^m \| > r_{lim} \right) \quad (3-11)$$

とし、比較する爪のペアが正常ではない組み合わせになった際に、距離が極めて大きい値になる問題を、一定値のペナルティを与える状況に置き換えた。爪の座標は手領域切り出し後の左端を原点と定義して求めた。この定義により、手が画像に対して並行移動しても爪の座標が変化しなくなる。

輪郭線の曲率に関する重みは図 3-10 に示すように、HLAC 特徴量算出で分けるブロックごとに設定する。輪郭線が存在しないブロックは、本来輪郭線が存在してはいけないため、重みは最大の 1.0 とした。 $\theta_{max}$  を超える平均曲率となった場合、 $\alpha_k$  の値を 1.0 とした。提案した評価関数などで、手動で設定が必要なパラメータは  $A$  ,  $\alpha_1$  ,  $\beta_1$  ,  $\beta_2$  ,  $\theta_{max}$  ,  $r_{lim}$  である。

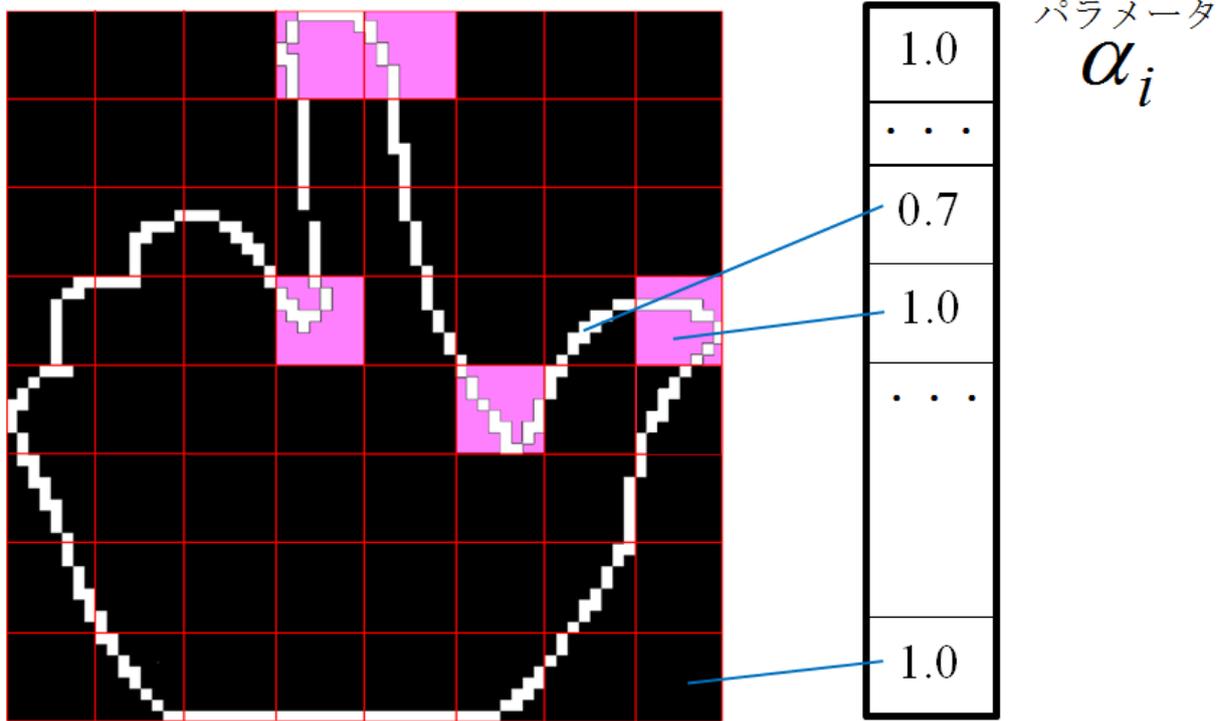


図 3-10 曲率を考慮した HLAC 用重みデータ列の生成

### 3.6 構想システムの実験結果

構想システムを実装し、実験により動作を確認した。動作を確認した時のシステム環境を図 3-11 に示す。使用したカメラは Point Gray Research 社製高速カメラ Flea3 である。画像サイズは  $648 \times 488$  [pixel]、撮影速度は 60 [fps] とした。照明は天井の蛍光灯で色温度 5000 [K] であった。推定結果の出力は 23 個の関節角度データではなく、データベース作成時に実際に撮影された手画像を出力した。この画像を目視で確認し、指を間違えるような大きなミスをした回数を計測した。データベースが持つデータセット数は 12500 枚、入力画像は掌側が主に写る手画像 2100 枚、手の甲側が主に写る手画像 350 枚とした。入力画像には図 3-3、図 3-4 の形状を持つ手画像を含めた。確認した出力結果は(1)HLAC のみの時、(2)HLAC と曲率による重み付けを行った時、(3)HLAC と爪による重み付けを行った時、(4)HLAC と曲率と爪による重み付けを行った時の 4 種類、およびそれらに対してまたぎ防止処理を付加した 4 種類の合計 8 種類で確認した。パラメータは  $A=40.0$ 、 $\alpha_1=0.7$ 、 $\beta_1=0.1$ 、 $\beta_2=50$ 、 $\theta_{\max}=2.0$ 、とした。このパラメータ値は調整した中で最も良いふるまいを見せたものである。

実験結果を図 3-12 に示す。この図から、爪情報を利用した 3 番、4 番、7 番、8 番が元の HLAC のみの手法、そして他の手法と比較して有意差を持って失敗数が低いことが確認できる。つまり、爪は手指形状推定の改善のために有効性が高いと言える。爪情報を用いないものは図 3-13 に示す



図 3-11 構想システムの計測実験時における撮影環境

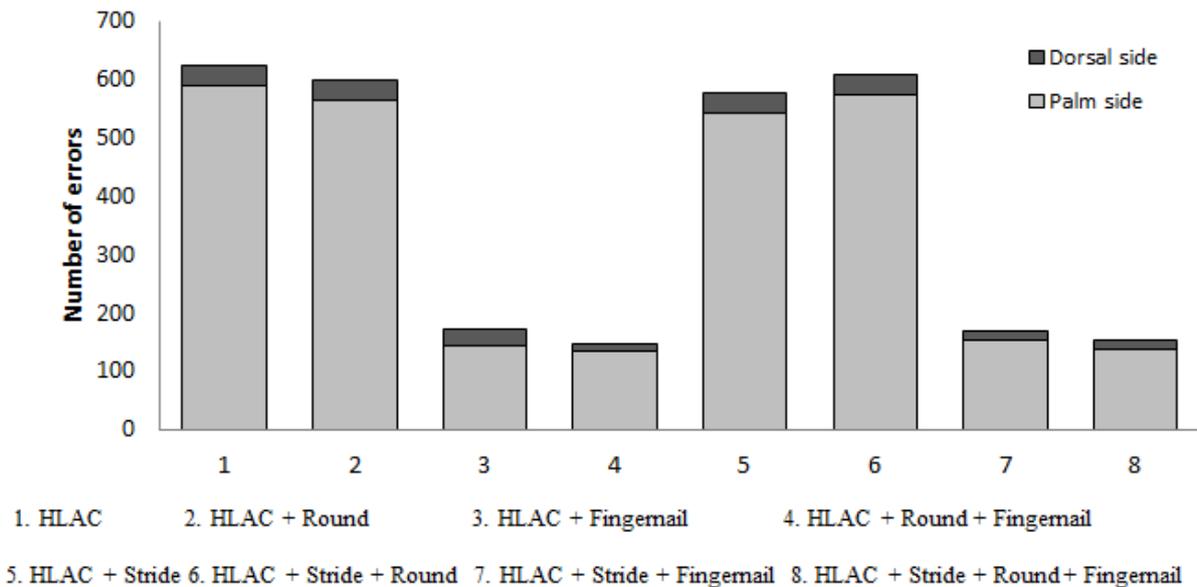


図 3-12 指を間違える推定結果の数の比較

ような指が曲がっているか否かを間違える、曲げている、または伸ばしている指の本数を間違える、手の裏表を間違えるという問題が頻出していた。爪の重心情報はこれらを防ぐ役割を果たしていた。なお、図 3-12 には現れない他の利点として、図 3-14 に示すように本構想手法を用いると指の曲げの浅さ、深さの表現が細かく表現できるようになった。爪情報が原因と思われる問題として、まれに図 3-15 に示すような、爪の位置が同じで曲がっている指が異なる間違いを起こすことが確認できた。また、どの手法においてもつまみと握りを間違える問題が頻出することを確認した。これは、爪の面積が小さいために爪の重みが小さくなり、指先の特定が困難になったためと考えられる。これに関しては、実際の爪検出自体も困難であることが予想されるため、シグモイド曲線を調整するだけでは対応が困難であると考えられる。この問題は、今後検討すべき課題とする。

著者らは上記の実験とは別に、輪郭線情報と爪の情報の重みを変化させることで、どのような変化が生じるか確認を行った。それにより、輪郭線情報と爪情報の利用方法の検討を行った。その結果、式(3-6)の  $A$  の値を変化させて輪郭線と爪の情報の重み付けを変化させる方法のみでは、輪郭線の曲率情報と爪情報を生かすことができないことが判明した。 $A$  の値が小さいと、爪情報の影響が強く、輪郭線情報が軽くなる。すると、曲げた指を間違えなかったならば、その指の先端位置がブレずに推定できるようになった一方、データセットの爪と入力画像の爪の位置があれば、輪郭線の形状が大きく異なっても類似したものと判定されるようになった。特に生じる問題は図 3-15 で述べた指を間違いの問題であった。一方で、 $A$  の値が大きいと、曲げた指の指先情報が小さくなり、曲げた指の推定が正しく行えなくなった。形状推定精度を上げるためには輪郭線情報から手の概形と指先位置、爪情報から指先位置の情報を取得し、指先位置情報を、概形情報より少々強めの情報として扱うことであると考えられる。しかし、現在の方法では輪郭線の指先位置情報と爪の指先位置情報のいずれかが弱められてしまう。これを防ぐためには、今回のような 1 つの評価関数に収めるのではなく、情報を別々に利用することが必要である。ただし、佐野ら[63]のようなブロック番号のような記号的数値を用いないことが必要である。よって爪の重心座標を利用し、データベースの爪の重心と、入力画像の爪の重心の距離が閾値以下のものを通すという方式とする事がより精度を改善するために必要であると結論づけた。

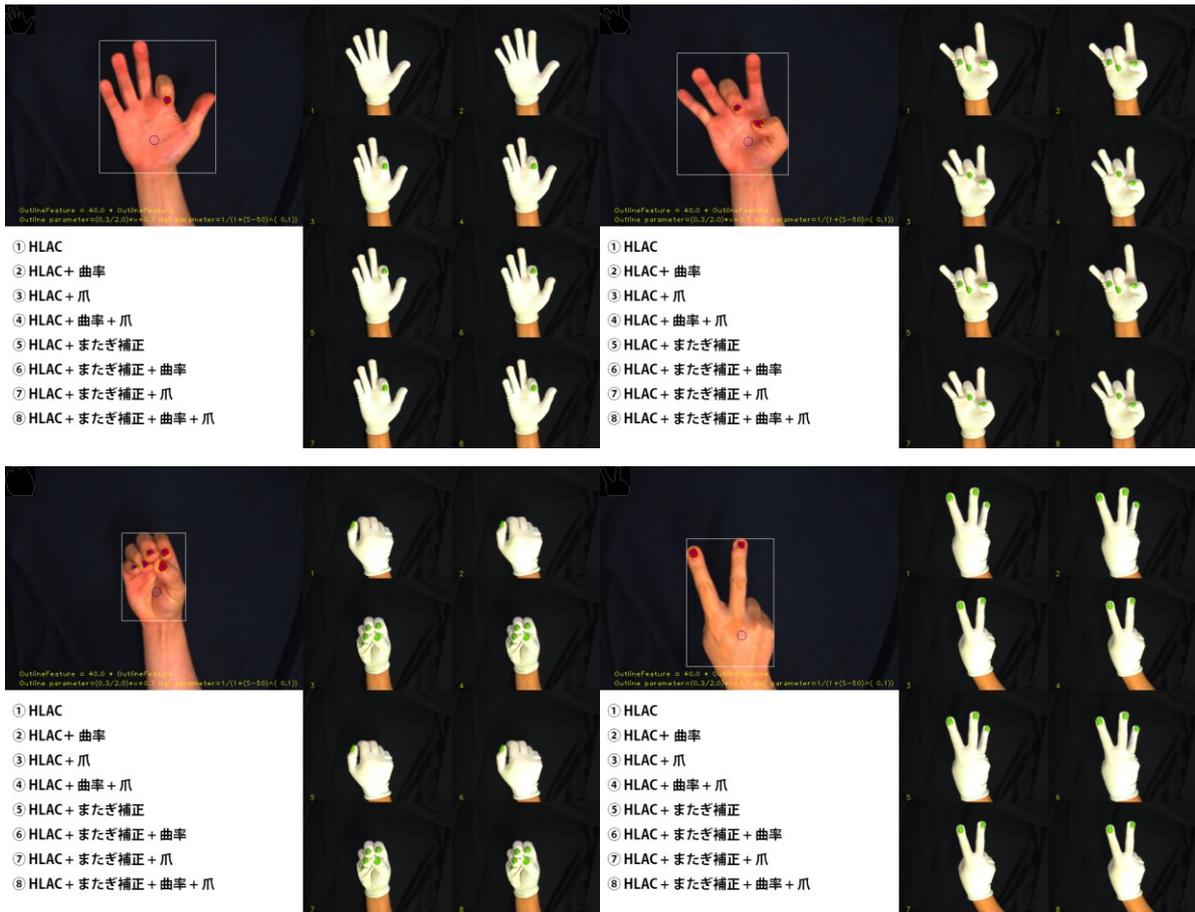


図 3-13 爪の情報が推定ミスを防いだ例



図 3-14 指の曲げの浅さ・深さに関する推定の改善

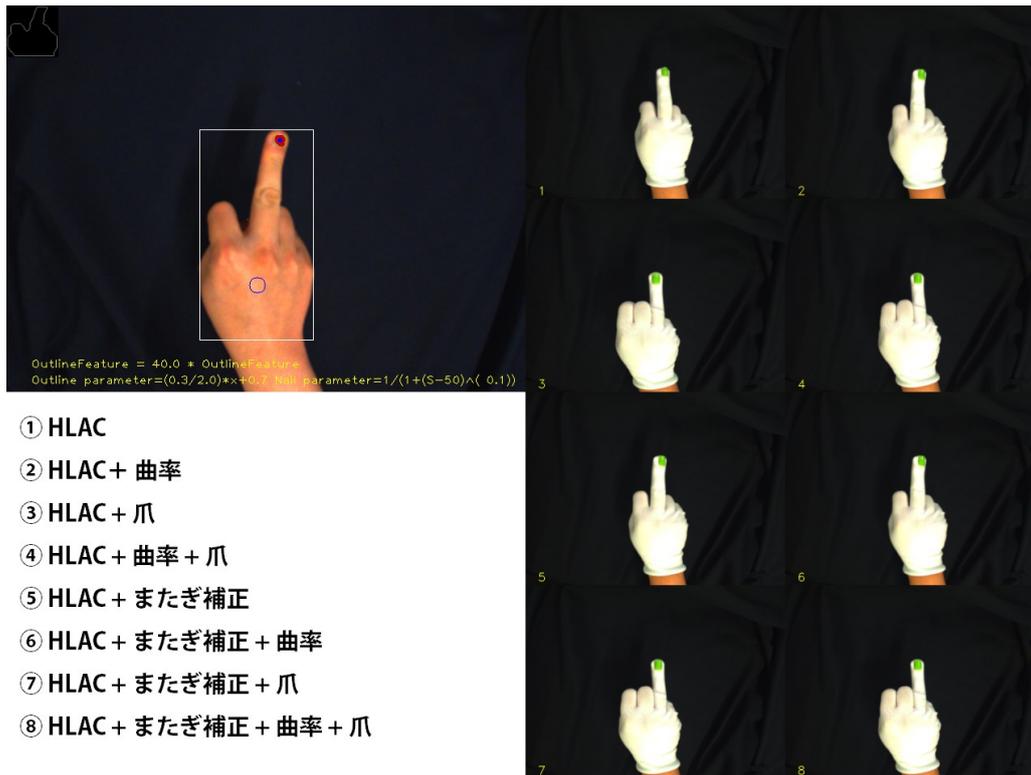


図 3-15 爪情報が原因と考えられる指間違いの例

### 3.7 第3章のまとめ

本章では、RGB カメラ 1 台を用いた手指形状推定の精度改善に爪情報が有効性を持つことを確認した。まず、従来の手指形状推定の概要に触れ、その手法に関わる大きな3つの問題と、それが主原因と考えられる推定間違いの傾向を述べた。次に、各問題に対して爪の情報を利用することで解決できること、できないことについて考え、爪情報と輪郭線の曲率による重み付けの変化、輪郭線のブロックまたぎ防止処理の3つで対応する方針を立てた。そして、その方針と爪を用いた形状推定手法の先行研究、爪を使うことにより生じるリスクの検討を基にして新しい爪を用いた手指形状推定手法のアルゴリズムを構想した。

構想したシステムを実際に実装し、指の曲げ伸ばしを間違える大きなミスがどの程度減少したかを計測する評価実験を行った。その結果、爪の情報を利用した手法が利用していない手法と比較して、優位的に精度改善が見られ、爪情報の有効性が確認された。一方で、輪郭線情報と爪情報を1つの評価関数で表現することにより、それぞれの情報を有効に活用できないことも確認された。このことから、爪情報は輪郭線情報でのマッチングとは別の処理で用いるべきであるという結論を得た。

なお、今回は図 3-16 に示すような、爪が見えず、複雑な形状をしているものや、爪が隠れてしまっているものは想定していない。これらは爪が取得できず、輪郭線情報もうまく活用できない形状のため、応用システム構築時に利用が想定されるならば、別の処理を考える必要がある。



図 3-16 爪や輪郭線の曲率情報を利用して正確な推定が困難な形状の例

## 第4章 爪検出システム

### 4.1 爪領域抽出に言及した先行研究と課題

爪領域抽出は、爪検出によって得られる爪領域の重心を知る際に必ず行う作業である。爪領域抽出に言及した研究は前例が少ないものの、いくつか確認されている。本節では確認されている爪領域抽出手法について紹介をし、各手法の問題点について述べる。その後、全体として何が不足しているかについて述べる。

爪領域抽出に関して最も提案数が多いのが津田らで、爪に楕円フィッティングをして仮想ネイルアートをするという目的のもと、3手法を提案している。1つの手法は色相を利用する。通常の状態では撮影した画像1枚と、フラッシュを用いて撮影した画像1枚を用いて、HSV表色系の色相変化量  $\Delta H$  の大きさを閾値として、爪領域を抽出する手法を提案した[65]。この手法は爪が肌より光を反射しやすい性質を利用したもので、撮影面に平行で照明に近い爪は正確に抽出できるが、光量が十分に届かない爪や、奥行き方向に傾いた爪は正確に抽出できないという報告がされている。つまり、前腕回旋に対応することは困難である。残りの2つは輝度を利用している。これらの手法は共に、(1) 指が1本写っている画像を使用する、(2) 指と爪が、指の軸に対して対称になっている、ことを前提としている。また、爪が左右対称に丸みを帯びるために、輝度の分布が対象線に対して環状に分布することを利用している。初めに提案された手法では、まず高輝度成分を一定の bin 幅でスライスした各領域重心が爪部分で重複するため、複数回重複した重心を爪候補点とした。爪候補を決定した後は、指画像に対して Canny エッジをかけ、爪候補点からエッジに到達するまで領域成長法を適応することで爪領域抽出を行った[66]。もう一つの手法では、文献[66]はライトが指の正面に無く、輝度分布に偏りが出る問題に弱い、ということ指摘した後、指の軸と傾きを利用して輝度分布を左右対称にする手法を提案した。また、領域成長法から Watershed 法に変更し、抽出精度を文献[66]より 10.6%改善させた[67]。これらの手法が持つ最大の問題点は、指の対象線を探す際に、指の軸を推定することである。軸推定が正しくできなくなるアブダクションを閉じた形状で指が2本以上重なる時や、掌側に曲げた指に対応することが困難となる。

津田らのように、一般的な表色系を用いた例は他に2件ある。古味らは、HLS表色系における色相  $H$  と彩度  $S$  に対し、適切に閾値を4つ定めることで爪領域が抽出できることを示唆した[68]。ただし、閾値設定は画像ごとに異なり、また、掌側においては、閾値設定が適切でない場合、掌領域が大量に抽出されることを、その後の報告で述べている[69]。Wang et. al は指先検出を目的として肌領域と爪領域の分布を UV 空間内で調査した。その結果、V 軸に対して  $86.41^\circ$  傾いた軸が分離するのに最適な軸という結果が現れ、それが  $90^\circ$  に近似できることから、爪は U 軸の負方向を正にした軸で表せることを述べた。この研究では手の甲側が写る画像を対象とし、16人のアジア人を対象として VGA サイズの画像で 48 枚実験したところ、98.7%の確率で爪による指先検出に成功している[70]。

一般的な表色系以外を利用する手法も提案されている。Chou et. al はハイパースペクトルカメラを用いて撮影したモノクロの手画像に対して k-means 法による教師なし学習を用いた分類、k-最近傍法による教師あり学習を用いた分類、第1主成分ベクトルを用いた輝度マップ構築による

肌と爪の分離を試みた[71]. その結果, k-means 法, k-最近傍法では爪領域の抽出が全くできず, 第 1 主成分ベクトルを用いると, 手領域の平面がカメラの撮影面に対してほぼ垂直の時のみ, 爪領域と肌領域で差が生じることを示した. これらの結果は, モノクロ画像での爪検出が困難であることを示唆している. 玉城らは手領域を抽出しやすくする個人別肌色表色系を表す軸に対して垂直な軸は肌と爪を分離しやすい軸であるとし, 爪領域抽出に利用した[62]. 評価実験では背景が青の環境以外では爪検出率が 90%を超えることを述べている. しかし, 爪ごとの抽出は 20%できればよいと定め, かつ各環境下での実験が 5 枚の爪を含んだ VGA サイズの画像を 3 枚でしか実験していない. また, 手の表裏, 前腕回旋の有無の記載がなく, どこまで対応できるシステムか不透明である.

従来研究全体で 2 つの大きな問題がある. 1 つは爪領域を抽出する際に用いる軸で最適なものは何かを具体的に検証した例がないことである. もう一つは全研究において, 爪の色は肌の色と異なっており, 閾値設定が適切であれば様々な手画像で爪領域が抽出できる, という前提を持っていることである.

著者は 2 つ目に挙げた問題の前提がいつ成り立つのか確認する予備実験を行った. 実験では爪領域画素が持つ RGB 値を既知とした. その既知の RGB 値に対して,  $\pm 2$  の誤差を許可して画素を抽出し, メディアンフィルタをかけて, ノイズ除去と領域生成を行った場合, 爪検出精度がどのようになるか調査した. メディアンフィルタのサイズは  $3 \times 3$  から  $4[\text{pixel}]$  ずつ大きくした. 残った領域が爪のみなら成功とした. 爪領域が消えてしまった場合と肌領域が抽出された場合は失敗とした. これは, 学習により, 爪領域画素の RGB 値を超高精度に推定し, 抽出を行った状態を想定している.  $\pm 2$  の許容誤差は, センサノイズにより各画素に対して常時起こると考えられるレベルの誤差である. RGB の各値に対して誤差を持たせるため,  $5 \times 5 \times 5 = 125$  色を 1 色と同じ色として許容したことになる. 画像は爪が一枚写る手画像を各回旋角度 100 枚, 合計 1900 枚を使用した. 前腕回旋は掌領域がカメラ平面と平行になる角度を  $-90^\circ$ , 手の甲領域がカメラ平面と平行になる角度を  $90^\circ$  と定義した.  $0^\circ$  は掌側が写る手画像と定義し, 曲げた指の爪領域を抽出した. 図 4-1 に  $7 \times 7$  から  $27 \times 27$  までのメディアンフィルタをかけた場合の予備実験の結果を示す.  $3 \times 3$  のメディアンフィルタは精度がどの回旋角度でも 0%に近く,  $31 \times 31$  以降のメディアンフィルタはサイズが大きすぎるために多数の爪領域が抽出できず, 精度が急降下するため除外した. 図 4-1 から, 掌側が主に写る手画像に対しての爪検出率は 70%未満となることが確認できる. また, この時, 手の甲側が主に写る手画像に対しての爪検出率は急降下していることがわかる. また, たとえば回旋角度が  $-90^\circ$  の時のみを利用するという大きな制限を課しても, 図 4-2 のように, 爪領域よりも大きな領域で抽出される肌領域も存在するためフィルタ除去による対応には限界がある. ここで行った実験は爪領域の RGB 値がほぼ特定されている理想状態で行っている. 値が未知の場合は, 爪領域を全て抽出できるとは限らないうえ, 肌領域をより抽出する可能性が高い. つまり, 爪検出精度はさらに低下する. 以上のことから, これまでの色情報を用いて画素単位で判定を行うことで画素抽出を行い, 平滑化処理をして領域を抽出する手法だけでは掌側が写る手画像に対応できないといえる. 第 3 章で述べたように, 単眼 RGB カメラを用いた手指形状推定において, 爪の情報は掌側が写る手画像の曲げた指の位置を特定するのに重要である. ゆ

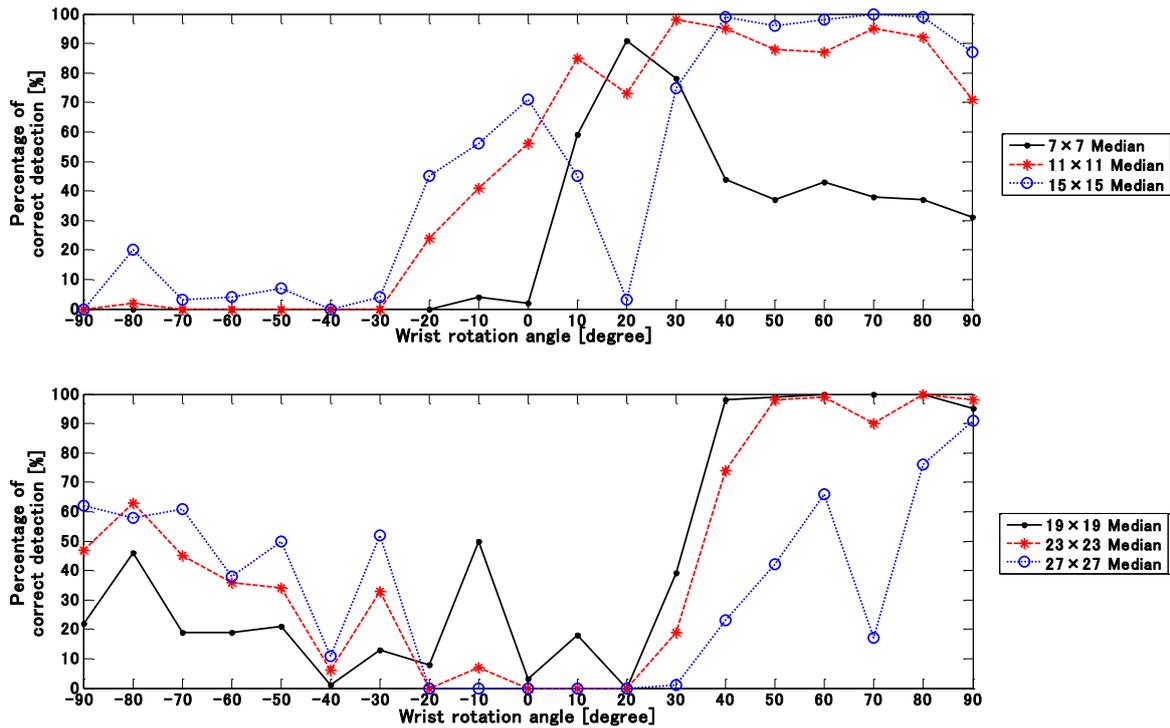


図 4-1 色情報のみで爪検出を行った場合における前腕回旋角度と爪検出率の関係

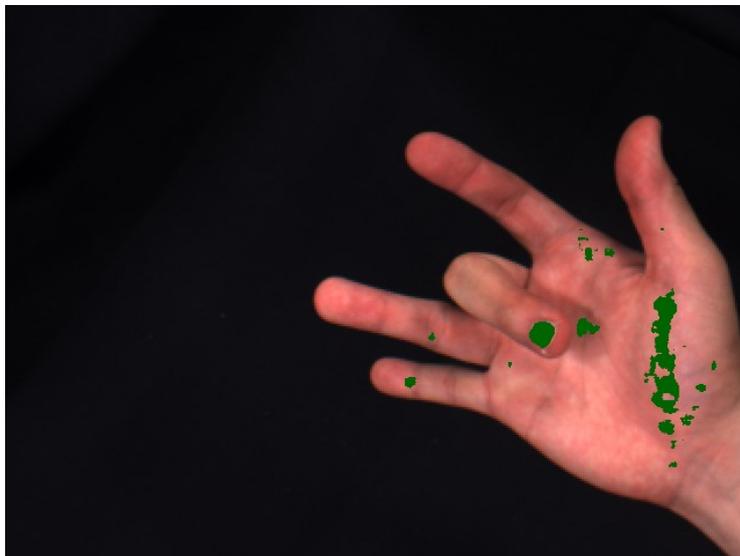


図 4-2 爪領域より大きく抽出された肌領域が存在する場合

えに、掌側が写る手画像に対応しなければならない。よって、2 つ目の前提は手指形状推定への使用を想定した爪検出システムを構築する場合に適用できない。この前提が成り立たない場合にも爪検出できるようになるためには、爪に似た色の領域が爪か肌か判別するアルゴリズムが必要である。以上より、本論文では爪か肌かを判別するアルゴリズムを新たに提案することを研究の主題とした。

本章において、4.2 節では、本論文で使用する爪領域抽出のための軸の提案する。そしてその軸と他の軸との比較を行うことで優位性を確認する。4.3 節では、従来研究で提案されていない、爪領域と、爪に似た色の肌領域を判別する手法を加えた新たなアルゴリズムに関する発想と具体的な実装方法について述べる。4.4 節では、爪らしい色を持つ領域の抽出法について述べる。4.5 節では爪らしい色の領域が爪か肌か判別する手法について述べる。そして4.6 節では、爪検出の精度と、爪判別アルゴリズムの働きについて評価した結果を示す。最後に4.7 節では、提案したシステムが抱える課題に対処するための1例を示す。

## 4.2 爪らしさを表す指標軸の生成

### 4.2.1 手領域画素が作る画素分布の傾向

図 4-3 と図 4-4 に掌領域が主に写り、爪が 4 枚写る手画像中の手領域が作る画素分布を例として示す。なお、爪領域の画素が少ないため、爪領域画素のプロットを強調のために 1 段階大きく設定している。これらの図から、爪領域画素は肌領域画素が構成する画素分布の分布が最も薄い方向に乘るような形で分布していることがわかる。また、肌領域の分布の飽和した画素を除去し、過疎でない場所のみに着目すると、薄い楕円体形状に近い図形となる。ゆえに、楕円体の厚みが薄い方向に平行な軸を引くことができれば、爪らしさを表す指標軸，すなわち爪尤度を示す軸を生成することができる。

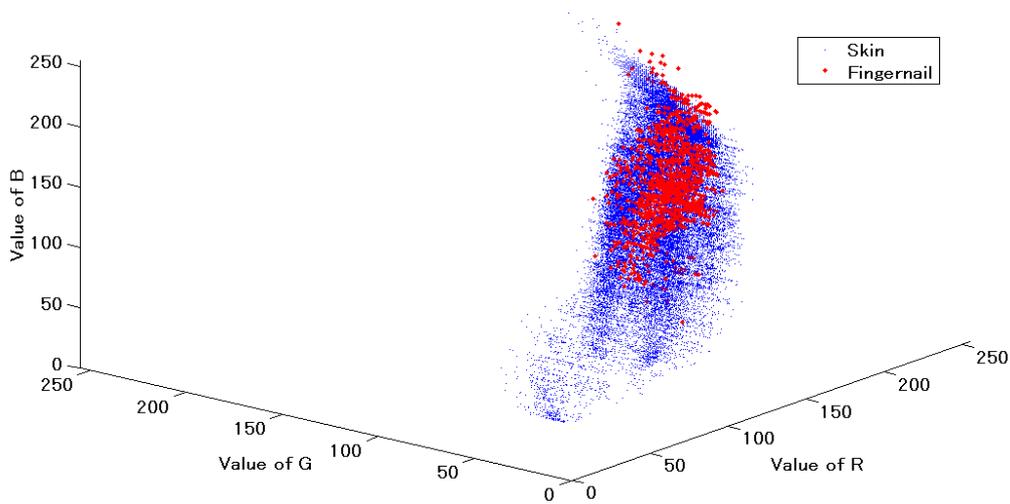


図 4-3 手領域画素が作る画素分布 (1)

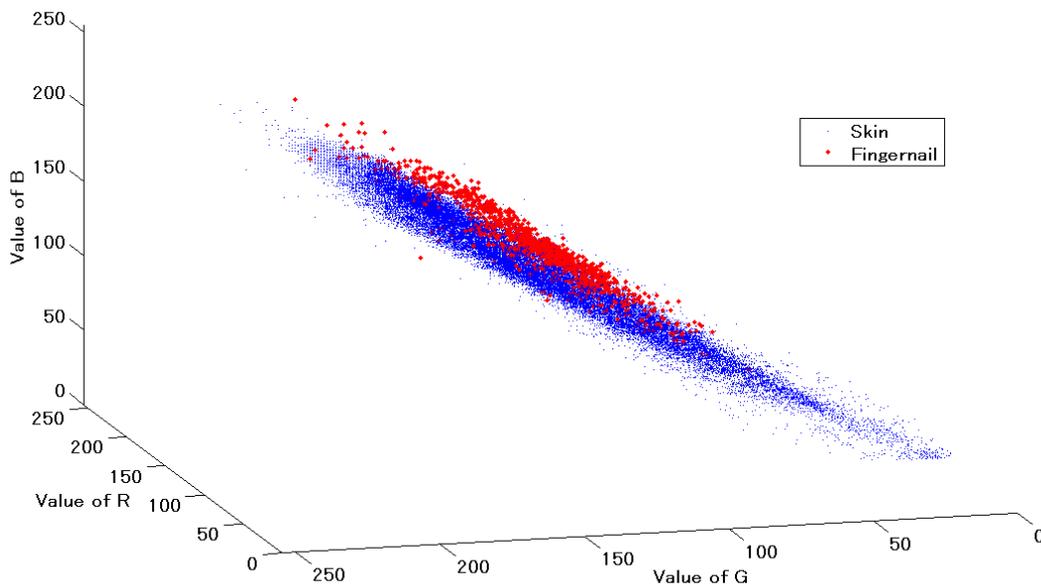


図 4-4 手領域画素が作る画素分布 (2)

#### 4.2.2 爪らしさを表す指標軸の生成

楕円体の厚みが最も薄い方向に平行なベクトルは、その楕円体を構成する要素を主成分分析して得られる第3主成分ベクトルと等しい。よって、本節では肌領域に対して主成分分析を行い、第3主成分ベクトルに対してどのように肌領域と爪領域が分布するかを示す。ここで、画素分布の形状は画像ごとの急激な変化は考えなくても問題がないため、爪検出システムではキャリブレーションであらかじめ軸を生成する。楕円体の形状を正しく取得するために、図4-5のような撮影ゲインが低く色飽和がほとんどない掌のみを撮影した手画像を用いる。まず、画像空間上の座標 $(x,y)$ を持つ画素 $xy$ の画素値 $\mathbf{x}_{xy}$ を縦に並べたデータ行列 $\mathbf{X}_D$ を式(4-1)のように定義して生成し、その分散共分散行列 $\Sigma$ を式(4-2)から求める。

$$\mathbf{X}_D = \left\{ \mathbf{x}_{xy}^T \mid \mathbf{x}_{xy} = (R_{xy}, G_{xy}, B_{xy})^T, R_{xy} \in (0, 255), G_{xy} \in (0, 255), B_{xy} \in (0, 255), \text{ if } xy \in C_{hand} \right\} \quad (4-1)$$

$$\Sigma = \frac{1}{N} (\mathbf{X}_D - \mathbf{x}_{mean})(\mathbf{X}_D - \mathbf{x}_{mean})^T \quad (4-2)$$

ここで、

$C_{hand}$ : 手領域画素の集合

$\mathbf{x}_{mean}$ : 手領域画素全体の平均値ベクトル

$N$ : 手領域の画素数

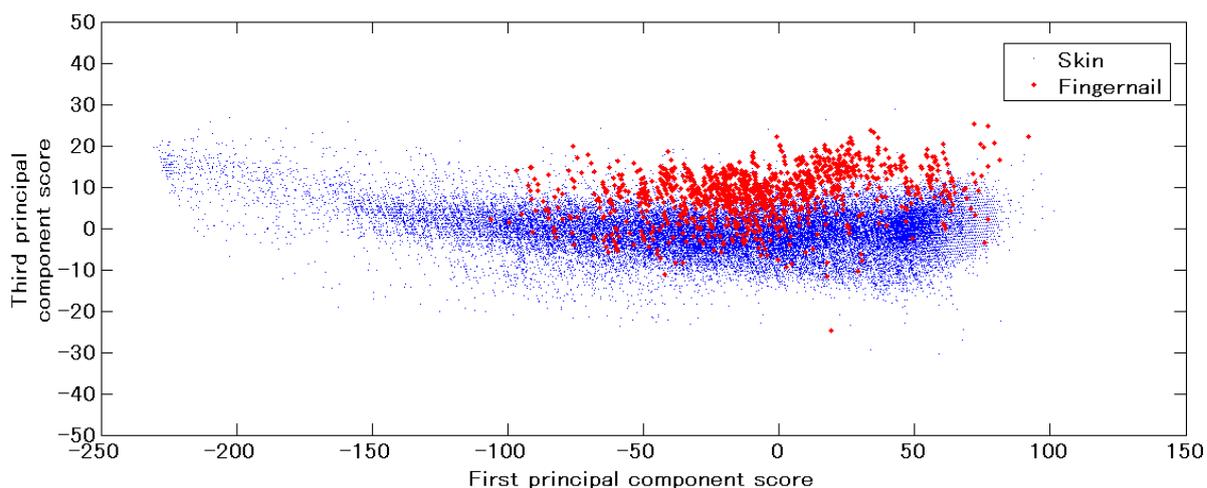
$\Sigma$ の固有ベクトルが主成分軸となる。最も大きな固有値に対応するベクトルは第1主成分ベクトルで、最も小さな固有値に対応するベクトルは第3主成分ベクトルである。これらのベクトルを基底とする新しい座標 $\mathbf{x}_{xy}^{PCA}$ は式(4-3)で定義される行列 $\mathbf{P}_{PCA}$ を用いて式(4-4)により算出する。

$$\mathbf{P}_{PCA} = (\mathbf{e}_{first} \mathbf{e}_{second} \mathbf{e}_{third}) \quad (4-3)$$

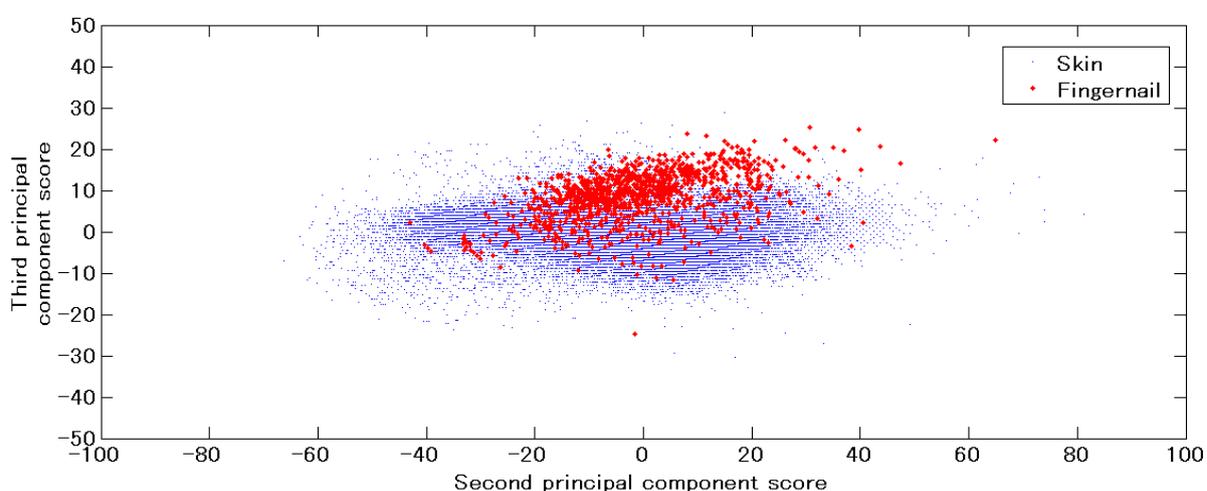
$$\mathbf{x}_{xy}^{PCA} = \mathbf{P}_{PCA}^{-1} (\mathbf{x}_{xy} - \mathbf{x}_{mean}) \quad (4-4)$$



図 4-5 画素分布取得に用いる手画像



(a) 第1-第3主成分ベクトル方向の写像結果



(b) 第2-第3主成分ベクトル方向の写像結果

図 4-6 主成分ベクトルを基底に採用した場合の画素分布

式(4-4)を用いた結果，図 4-4 の画素分布は図 4-6 のように写像され，爪領域の画素が抽出しやすくなることを確認できる．なお，図 4-5 を処理して得られた主成分ベクトルは式(4-5)の通りであり，第3主成分ベクトルは色の飽和が起こりやすい赤色に影響されにくい軸であることがわかる．

$$\mathbf{P}_{PCA} = \begin{pmatrix} 0.806 & -0.599 & -0.012 \\ 0.433 & 0.592 & -0.679 \\ 0.414 & -0.679 & 0.734 \end{pmatrix} \quad (4-5)$$

また，原点を画素分布の重心としてベクトルがどのような色を表すか解析すると図(4-7)のように第1主成分ベクトルが肌色の輝度，第2主成分ベクトルが赤からシアン方向に輝度を増加させながら変化する軸，第3主成分が黄色から青紫方向へ変化する色相軸であることがわかる．黒い線は原点を表す．長さが異なるのは変域が異なるためである．

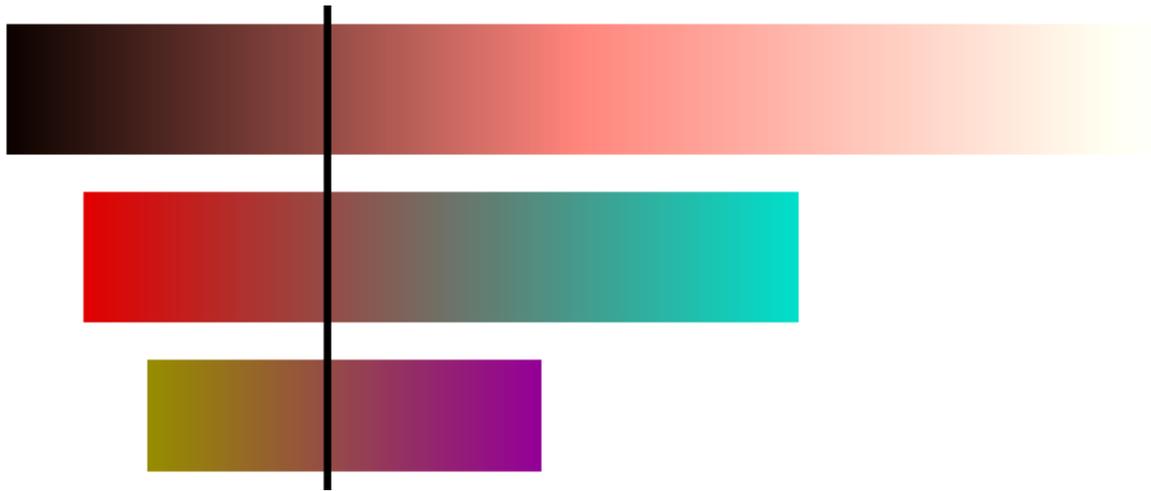


図 4-7 主成分ベクトルが表す色

#### 4.2.3 指標軸の評価

第3主成分ベクトルを利用した軸が、他の軸と比較した際に最適になるか評価を行った。評価で比較した表色系は RGB 表色系, XYZ 表色系, YIQ 表色系, YUV 表色系, HSV 表色系, HLS 表色系, L\*a\*b\*表色系, L\*u\*v\*表色系, YSaSb 表色系[62], 第1・第2主成分軸, 線形判別分析軸である。YSaSb 表色系, 主成分軸と判別分析軸以外は CIE で定められた表色系である。

YSaSb 表色系は 4.1 節で触れた, 玉城らの個人別肌色軸に関する表色系である。Y は輝度軸, Sa は肌色軸, Sb は輝度と肌色の傾向に左右されない軸として定義されている。ここでは, 参考文献[62]に従い, 各軸を定義する。

Sa, Sb 軸の交点は I, Q 軸の交点と一致する。原理としては, R, G, B の各値を 255 で除し, 0 から 1 の範囲とする。次に式(4-6)で表される  $\gamma$  補正をかけ, 補正後の値を R', G', B' とする。

$$\begin{aligned}
 R' &= \begin{cases} 4.5 \times R & \text{if } R < 0.018 \\ 1.099 \times (R^{0.45}) - 0.099 & \text{if } R \geq 0.018 \end{cases} \\
 G' &= \begin{cases} 4.5 \times G & \text{if } G < 0.018 \\ 1.099 \times (G^{0.45}) - 0.099 & \text{if } G \geq 0.018 \end{cases} \\
 B' &= \begin{cases} 4.5 \times B & \text{if } B < 0.018 \\ 1.099 \times (B^{0.45}) - 0.099 & \text{if } B \geq 0.018 \end{cases}
 \end{aligned} \tag{4-6}$$

そして, R'G'B' を YIQ に変換する。この時に得た Y 軸が YSaSb 軸の Y 軸となる。最後にキャリブレーションによって得た変換パラメータを基に回転変換を行う。キャリブレーションでは, 十分多量の掌領域の肌画素を取得する。そして肌画素の座標を縦軸 I, 横軸 Q としたときの画素分布の重心 (Q', I') を求める。

原点からこの重心までを結ぶベクトルの単位ベクトルが Sa 軸の基底となる。(Q, I) 座標から (S<sub>a</sub>, S<sub>b</sub>) 座標へ変換する, 回転変換のパラメータ  $\alpha$  は式(4-7)となる。

$$\alpha = \tan^{-1} \left( \frac{I'}{Q'} \right) \tag{4-7}$$

したがって、座標値の計算式は式(4-8)となる。

$$\begin{pmatrix} S_a \\ S_b \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} Q \\ I \end{pmatrix} \quad (4-8)$$

線形判別分析軸は、線形判別分析法により求める軸である。線形判別分析法を用いると、判別したいクラスがわかっている際に、各クラス内での分散が小さく、クラス間の分散が大きくなる最も良い線形軸を求めることができる。あらかじめキャリブレーションで掌領域のみを写した手画像から肌領域に属する画素（以下、肌画素）を、爪のみが写る手画像から爪領域に属する画素（以下、爪画素）をあらかじめ 5000 [pixel]以上取得する。そして、クラス内分散  $S_w$  を式(4-9)、クラス間分散  $S_b$  を式(4-10)で求め、 $S_w^{-1} S_b$  の最大固有値に対応する固有ベクトルを導出する。この固有ベクトルが線形判別分析軸となる。

$$S_w = \frac{1}{N} \sum_{i=1,2} \sum_{l \in C_i} (\mathbf{x}_l - \mathbf{m}_i)(\mathbf{x}_l - \mathbf{m}_i)^T \quad (4-9)$$

$$S_b = \frac{1}{N} \sum_{i=1,2} n_i (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T \quad (4-10)$$

ここで、

$C_i$  :  $i$  番目のクラス

$\mathbf{x}_l$  :  $C_i$  に属する  $l$  番目の画素の RGB 座標

$\mathbf{m}_i$  :  $C_i$  に属する全画素の平均 RGB 座標

$n_i$  :  $i$  番目のクラスの画素数

$N$  : 全画素数

線形判別分析軸を色解析した結果を図 4-8 に示す。図 4-7 と図 4-8 を比較すると、線形判別分析軸のベクトルは第 3 主成分ベクトルに輝度成分が含まれたような軸であることがわかる。

軸の評価には Receiver operating characteristic (ROC) 曲線を用いた。図 4-9 に ROC 曲線の概要を示す。各クラスのヒストグラム面積が 1 になるように正規化しておく。ある閾値を設定し、抽出することを考える。抽出したいクラス(本論文では爪)がその閾値によって抽出できる割合を真陽性率(敏感度)と定義する。また、抽出しないクラス(本論文では肌)がその閾値によって抽出してしまう割合を偽陽性率(1-特異度)という。すると、1つの閾値で真陽性率と偽陽性率の1組のデータを取得することができる。このデータを、閾値を可変させることでデータ列として取得できる。これを横軸に偽陽性率、縦軸に真陽性率としてグラフ化させると曲線を描くことができる。これが ROC 曲線である。偽陽性率が低く、真陽性率が高いほど、ROC 曲線は左上に急峻に立ち上がる。よって、係数の大きい逆指数関数上の形になるほど、その軸が抽出に優れていることを表す。被験者として学生 7 人分の手を用意した。被験者 1 から 5 が男性、6 と 7 が女性である。表 4-1 は爪、掌領域、手の甲領域の肌の平均 RGB 値をまとめたものである。画像は各被験者 100



図 4-8 線形判別分析によって得られたベクトルが示す色

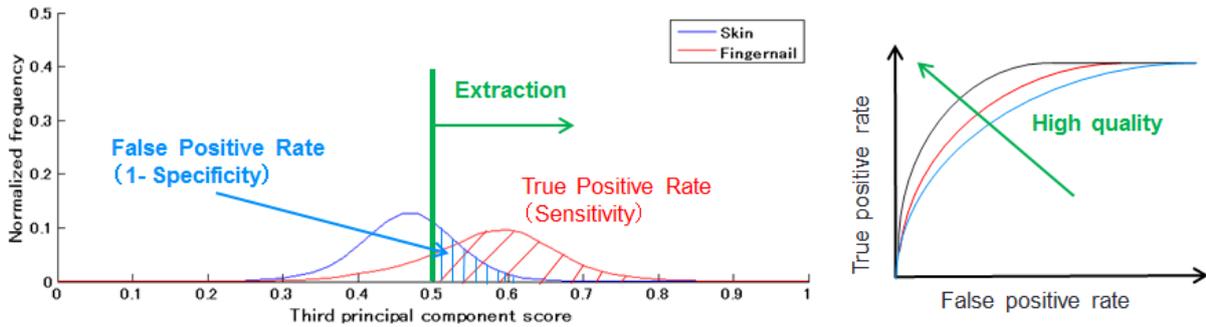


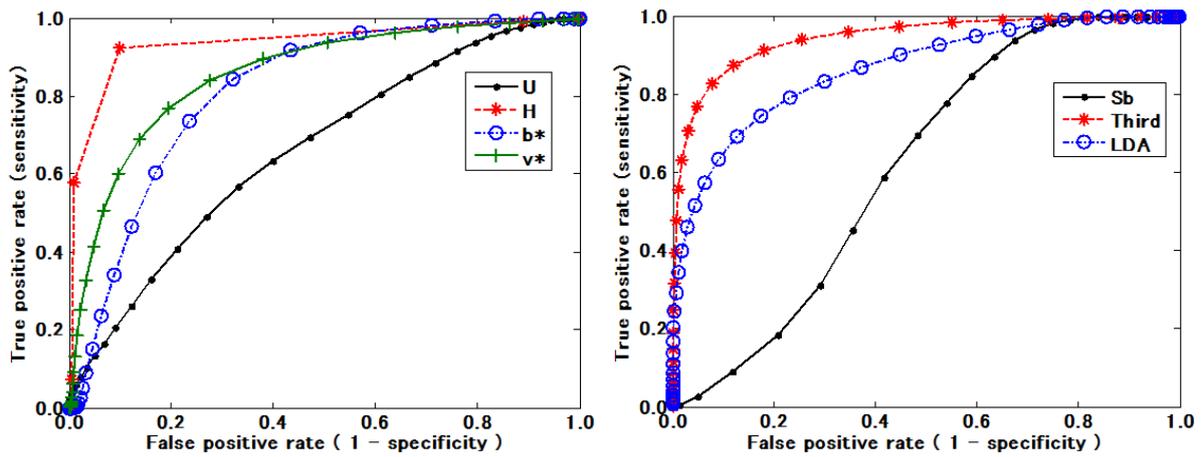
図 4-9 ROC 曲線の概要

表 4-1 軸評価のための予備実験に参加した被験者と手に関するデータ

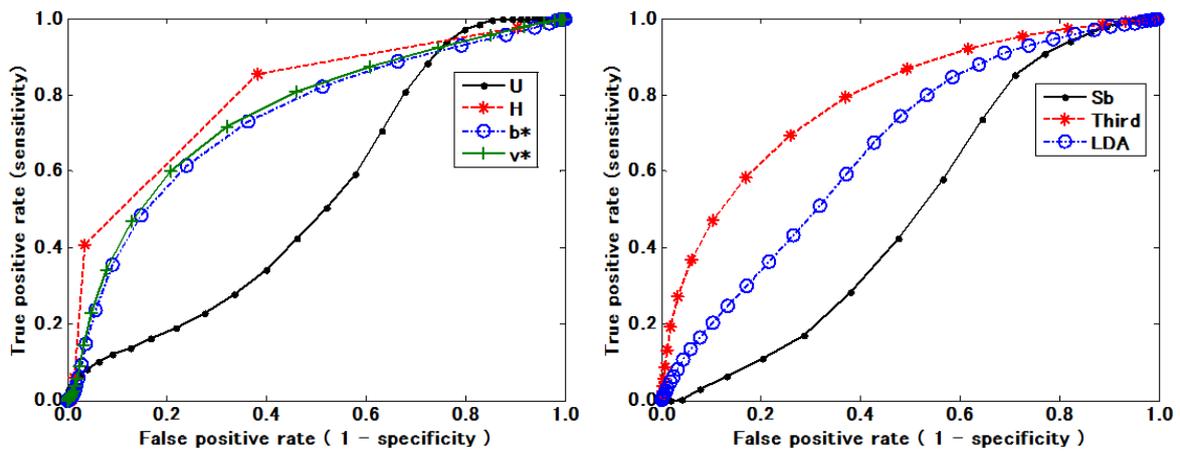
Subject number	1	2	3	4	5
Gender	Man	Man	Man	Man	Man
Fingernail	(229,126,126)	(197,123,126)	(248, 130, 129)	(224,125,128)	(238,170,167)
Palm side	(227,129,120)	(182,125,118)	(231,133,122)	(197,121,115)	(203, 118, 106)
Dorsal side	(199,126,105)	(168,121,108)	(193,113,98)	(184, 126, 113)	(168, 93, 74)
Subject number	6	7			
Gender	Woman	Woman			
Fingernail	(244,166,157)	(241,137,134)			
Palm side	(234,142,129)	(214,122,116)			
Dorsal side	(225,154,132)	(193,121,101)			

枚である。各画像には爪が1枚含まれるものを使用した。この予備実験前、被験者1のヒストグラムデータを確認し、明らかに爪が正しく抽出できないものは除外した。ここでは、正しく抽出できる可能性があるとした。U, H,  $b^*$ ,  $v^*$ , Sb, 第3主成分軸 (Third)および判別分析軸(LDA)のみを対象とする。それぞれの軸の変域の大きさの違いが評価に影響を及ぼさないように、ヒストグラムは、手領域画素が持つ最小値が0, 最大値が1となるように正規化をかけた。閾値の分解能は0.02とした。(0,0)から(1,1)へ変化する際のプロット数が一か所に固まっているほど、閾値設定をピンポイントな値で行わなければならない、パラメータ調整が難しいことを示す。

図4-10に予備実験の結果を示す。図4-10から、手の裏表に関わらず色相Hが一番立ち上がりのよい軸であることがわかる。しかし、立ち上がり過程でプロットされている点が2点しかないことから、閾値設定が困難であることがわかる。第3主成分軸は2番目に立ち上がりがよい軸である。そして、プロット数が多いことから、閾値設定に余裕があることがわかる。ゆえに、軸の使いやすさも考慮に入れると、第3主成分軸が最も優れた指標軸になるといえる。なお、線形



(a) 手の甲側が主に写る手画像



(b) 掌側が主に写る手画像

図 4-10 爪と肌を対象とした ROC 曲線

判別分析軸の精度が高くならなかった理由は、軸の情報に輝度が含まれたためと考えられる。輝度情報は津田ら[66][67]の研究でも使われたように、大きな情報が含まれる可能性がある。しかし、輝度は照明条件や手の自己遮蔽による影に影響を受けやすい。それが問題を引き起こしたと考えられる。

### 4.3 爪検出システムの動作環境と処理の流れ

本論文で述べるシステムが動作する環境を図 4-11 に示す。カメラは Point Gray Research 社製の Dragonfly Express を使用した。画像サイズは  $640 \times 480$  [pixel]、撮影速度は最大 120 [fps] である。黒布は背景分離を行った後に本システムを使用することを想定しているため、背景分離がしやすいように使用した。照明は天井に取り付けてある色温度 5000 [K] の蛍光灯である。

図 4-12 に爪検出システムの基本動作のフローチャートを示す。著者が構築したシステムは起動時、個人ごとにキャリブレーションを行い、第 3 主成分軸を生成する。メインシステムでは、初めに画像を入力し、色ノイズを除去する。次に、爪尤度の高い領域を抽出し、注目点を設定する。そして、注目点周りの情報を用いてその領域が爪か肌かを判別し、爪のみを検出する。最後に検出した爪の注目点を着色した画像を出力する。なお、本システムをアプリケーションに応用する場合には、注目点を爪領域の重心として出力する。画像の入力後から画像の出力前までの処理速度は Intel Core i5-2500 CPU 3.30GHz, 8GB メモリのコンピュータで最大速度 250 [fps]、最低速度 150 [fps] となった。



図 4-11 爪検出を行う撮影環境

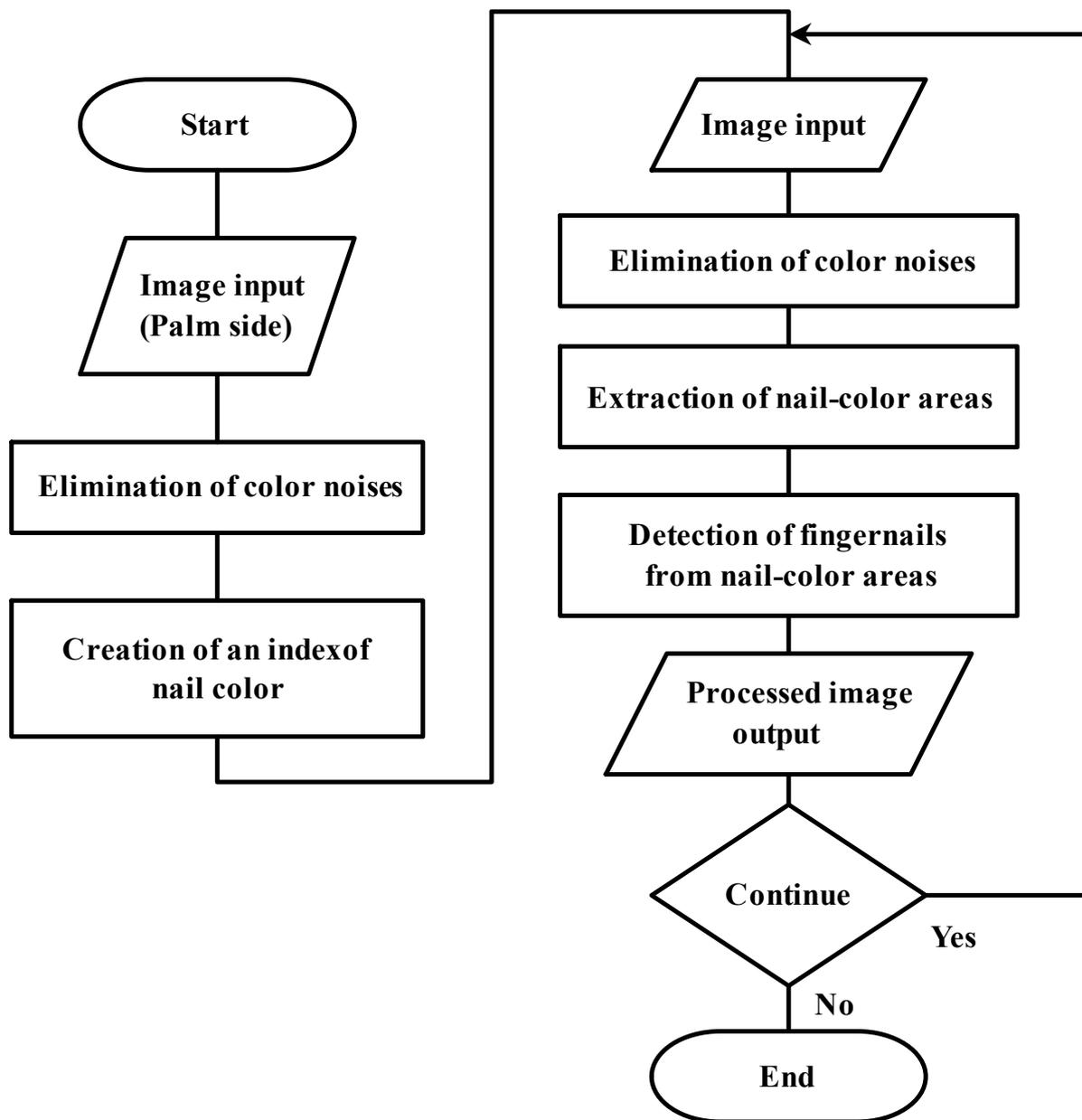


図 4-12 爪検出システムの処理の流れ

## 4.4 爪らしい色を持つ領域の抽出と注目点の設定

爪尤度の高い領域，つまり爪の候補となる領域を画像上から抽出する．そして，爪か肌かを判別する際に使用する注目点を設定する．まず，各画素の第3主成分得点を算出する．画素  $pixel_i$  の第3主成分得点は，式(4-11)で定義した．

$$Third_i = \begin{cases} (\mathbf{x}_i - \mathbf{x}_{mean})^T \mathbf{e}_{third}, & \text{if } pixel_i \in C_{hand} \\ -100, & \text{if } pixel_i \notin C_{hand} \end{cases} \quad (4-11)$$

画素が手領域でないときに使用する  $-100$  は，その範囲が手領域でないことを示すための象徴として利用している．一般的に第3主成分得点の値は  $-50$  を下回ることがないため，その値よりも低ければどのような値を使用してもよい．この数値を持つ画素はその後の処理を行う際に無視される．画素の抽出は事前に設定した固定閾値  $Th_{third}$  以上のものを抽出し，2値画像の白点とする．本論文では，著者らの経験をもとにし基準閾値は  $Th_{third}=6.0$  としている．抽出後は平滑化，およびラベリングを行うことで爪尤度の高い領域を取得する．最後に，ラベリングにより取得した各領域の注目点を設定する．注目点の座標は領域に属する画素中の，特に第3主成分得点の高い上位100画素の平均座標とする．この処理は，爪と指の側面領域の画素が連結する問題に対処するための処理である．取得領域が100画素以下の場合には領域の重心点を注目点とする．この100画素という基準は予備実験を基に決めた値である．この注目点は通常，真の爪領域の重心点に近くなるため，最終出力の爪領域重心の座標はこの注目点の座標とする．

図4-13は著者の経験と実験における，様々な人の手で特に抽出されやすい肌の位置を表している．参考データとして，表4-1の被験者1に対し，最適と考えられる閾値で設定した時に実際に抽出された肌の抽出割合と画像1枚当たりの抽出数を調査した結果を図4-14に示す．個人差や照明の当たり方によってこの割合は変化する．本論文で使用する環境では，被験者によらず図4-14のように拇指球，指腹，指の側面の順で高くなる傾向が確認された．なお，光の当たり方によっては小指側にあるふくらみの小指球や手の甲側のMP関節が出ることもある．小指球は拇指球と，MP関節は指の側面部と色の分散の傾向が類似している．

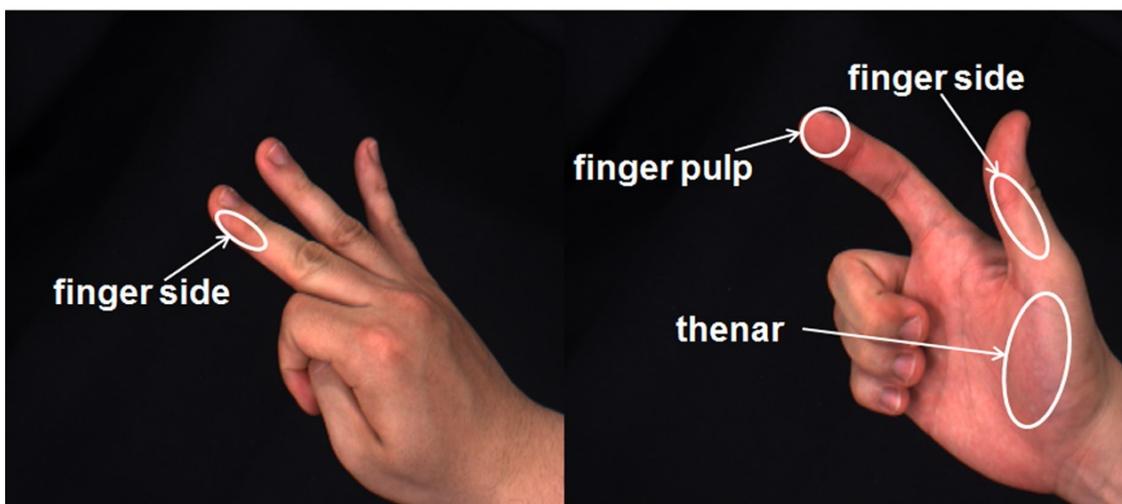


図 4-13 爪として誤検出しやすい肌領域の分布位置

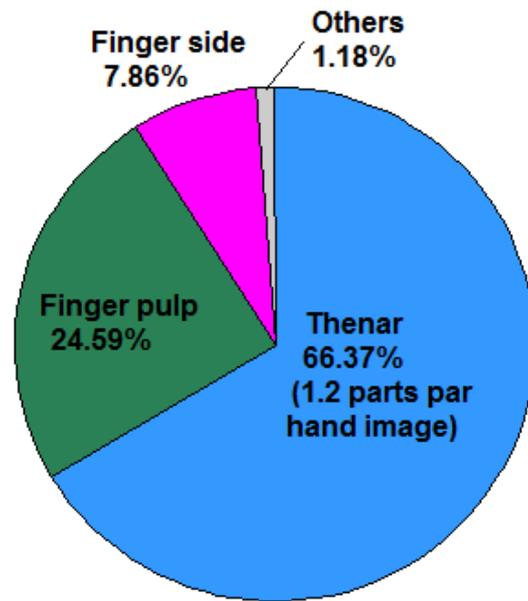


図 4-14 抽出割合と頻度の例

## 4.5 局所的な情報を利用した爪と肌の判別

### 4.5.1 爪と肌の判別を実現するための情報

色情報を画素単位で利用して用いるだけでは爪尤度が高い肌色領域を抽出する問題が生じる。ここでは、爪と爪尤度が高い肌を区別するために使用する情報について述べる。

図 4-15 に第 3 主成分得点を可視化した画像を示す。この図では第 3 主成分得点が 8.0 以上の時 255, -10 以下の時 0 とし, その間の範囲は 0 から 255 の間の整数値に等間隔で振り分けた。上限 8 と下限 -10 という値は, 爪の可能性が高い画素値と爪の可能性がほとんどない画素値として, 著者が経験をもとに判定している値で設定した。図 4-15 から, 爪領域は第 3 主成分の値が低い領域に囲まれた, 小さい独立領域として存在していることがわかる。また, 爪と誤検出されやすいと述べていた肌領域を含め, 肌全体の性質として, 色は大域的に値の平均が変化する傾向が見られる。また, 肌領域を局所的にみると色のまばらさが目立つ傾向がみられる。これらの性質は, (1) 色が独立している小領域が存在するか否か, (2) 局所的に見た時, 色は分散しているか, 密集しているか, の 2 つの性質に集約できる。本節では, この 2 つの性質を各項に分け, 考えのコンセプトと性質を数値化し, 判定に用いるためのアルゴリズム, そしてどのような場合に強く, どのような場合に弱いと想定されるかについて述べる。なお, これらの性質を用いた判定を行う際には, 4.4 節で取得した領域を爪候補領域と定義し, 注目点を中心として矩形の関心領域 (ROI) 画像を生成する。4.5.2 項と 4.5.3 項で述べる処理は ROI 画像ごとに行う。ROI は正方形で爪領域全体を含む大きさで設定する。本論文では爪の小さな被験者は 1 辺 41[pixel], 大きな被験者は 1 辺 51[pixel]と設定した。

### 4.5.2 領域の独立性に着目した爪と肌の判別手法

爪領域は小さい独立領域として存在し, 肌は大域的に色に変化している集合領域である。よって, 爪候補領域が小さい独立領域なのか, 集合領域の一部に属するのか判定できれば爪か肌かを判別することができる。もし, 爪候補領域が爪領域ならば, 注目点, すなわち ROI の中心は爪領域内に存在する。つまり, つまり色が周辺と異なり独立した領域内に存在する。一方, 爪候補領域が肌ならば, 周辺とそれほど色に差がない領域に ROI の中心が来る確率が高い。ゆえに, ROI



図 4-15 第 3 主成分得点を可視化したマップ画像

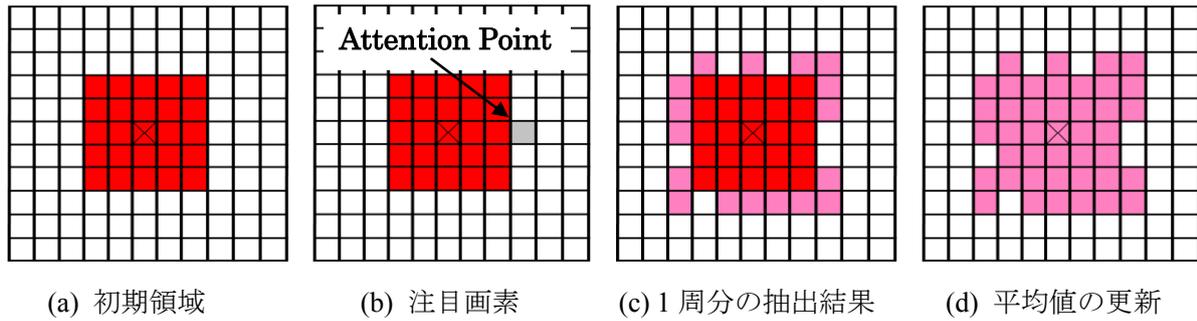


図 4-16 結合領域の生成手法

の中心から色の類似する画素を領域成長法により結合させた時、爪ならば小さい範囲のみ抽出されて処理が終了するが、肌は大域にわたり領域抽出が起きる。よって、領域抽出が行われた領域の広さが爪か否かを判別する情報として利用可能である。この手法での判別が難しい部位の代表は指腹といえる。理由は関節のしわと領域の小ささにある。指腹上に注目点を置くと、小指など、大きさが小さい指の先端は ROI 内に含まれるため、領域抽出が止まる。そして、手の中心に向かう方向には関節部のしわが抽出を止める現象が確認されている。

実際の処理は、図 4-16 のように行う。まず ROI 画像の中央を中心とした 5×5 の領域を初期の「結合領域」と定義する。その際、結合領域の平均の第 3 主成分得点の平均値を算出する。次に、注目点から  $j(\geq 3)$  番目の外周上を内側から順に探索する。 $j$  番目の外周上の画素  $i$  を注目画素とする。この画素が結合領域と隣合っており、次の式(4-12) を満たすとき、この画素を結合させる。

$$Third_i^j - Third_{connectionAreaMean}^{j-1} > Th_{connection} \quad (4-12)$$

ここで、

$Third_i^j$  :  $j$  番目の外周上の画素  $i$  の第 3 主成分得点

$Third_{connectionAreaMean}^{j-1}$  :  $j-1$  番目の外周上まで探索して得られた結合領域の平均第 3 主成分得点

$Th_{connection}$  : 結合するか否かを判定する閾値

外周を 1 周したら結合領域の平均値を更新する。この作業を結合画素がなくなるか、爪でないと判定される閾値を超えるまで行う。本稿では、閾値は ROI のサイズを使用している。図 4-17 は  $Th_{connection}$  の値を  $-3$  とし、結合領域の増加がなくなるまで抽出した結果の例である。爪領域を対象としたものは領域内で収まり小さくなった。一方で、肌領域を対象としたものは ROI に収まりきらない巨大な領域を形成した。このように幅が大きく異なるため、判定を安定して行いやすいのが特徴である。

#### 4.5.3 色のまとまり方に着目した爪と肌の判別手法

爪領域の周辺は第 3 主成分得点の値が低いこと、肌領域は色がまばらになりやすいことを利用して爪と肌の判別を行う。ここで使用するのは高爪尤度画素、つまり高い第 3 主成分得点を持つ

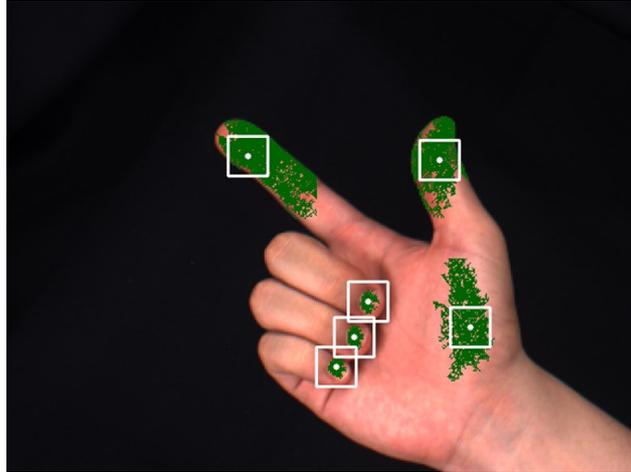


図 4-17 ROI 中心と色が類似する領域の抽出結果

画素の分布である。爪領域重心を中心を持つ ROI 画像の場合、爪領域と、第 3 主成分得点が低い肌領域を持つため、高爪尤度の画素のみを抽出すると、抽出される画素は爪領域から選ばれる。ゆえに抽出画素の密集度が高くなる傾向がある。一方、肌の場合は色がまばらのため、抽出画素の密集度が低くなる傾向がある。図 4-18 はそれを実際に調査したものである。ROI サイズは  $41 \times 41$  [pixel] とした。矩形枠は中心より 8 近傍距離で 10 進んだ画素に引いてある。実線が矩形枠内の領域を対象として正規化ヒストグラムをとったものであり、中心付近の画素の色の傾向を表している。点線が矩形枠外（境界含む）の領域を対象として正規化ヒストグラムをとったものであり、中心から離れた画素の色の傾向を表している。この 2 つのヒストグラムが離れるほど中心と外側で色が明確に異なることを意味し、中心領域にある画素を集中的に抽出できる。図 4-18 では、爪を含む ROI 画像では、中心付近の値が外側の値より有意的に第 3 主成分得点が高いことがわかる。つまり、ROI 内で爪尤度のランキングをとり、上位のみを抽出すると、あまり大きくないリスクで大量の爪領域を抽出できる。一方で、肌は、中心付近とそれ以外であまり値が変わらないため、上位のみを抽出しても様々な位置の画素が抽出され、拡散して抽出する傾向がある。特に拡散するのは拇指球で、拇指球の除去には特に効果を発揮する可能性があることが示唆されている。

上位何%程度をとるべきか調査した。抽出数は少なすぎると爪内部で抽出される所とされない所で分散し、正しい判別ができないため、抽出比率は出来るだけ多い方が良い。しかし比率を多くしすぎると、肌領域も徐々に密集度を上げるため限度が存在する。よって、どの程度の抽出比率なら爪の方が肌よりも密集度が高くなるのかを調査した。調査に使用した被験者の性別と手の甲、掌側の平均 RGB 値の情報を表 4-2 に示す。ここでは、密集度を調べるために式(4-13)と式(4-14)を用いた。ROI のサイズは全て  $41 \times 41$  [pixel] とした。各被験者は画像数  $M=100$  枚で爪の大きさで結果が変わらないように、1 つの指を曲げた画像とし、曲げている指も 25 枚ずつで変更した。爪は示指から小指を対象とした。対象とする部分は手の甲が主に写る手画像中の爪、掌が主に写る手画像中の爪、拇指球、指腹、指の側面とした。爪を 2 パターンにするのは、手の甲側が主に写る場合は爪と手の甲側領域が ROI 画像内に存在するが、掌側が主に写る場合は、爪・手の甲側領域・掌側領域の 3 つが ROI 画像内に存在し、抽出画素が拡散しやすくなることを考慮した。実験結果を図 4-19 と図 4-20 に示す。

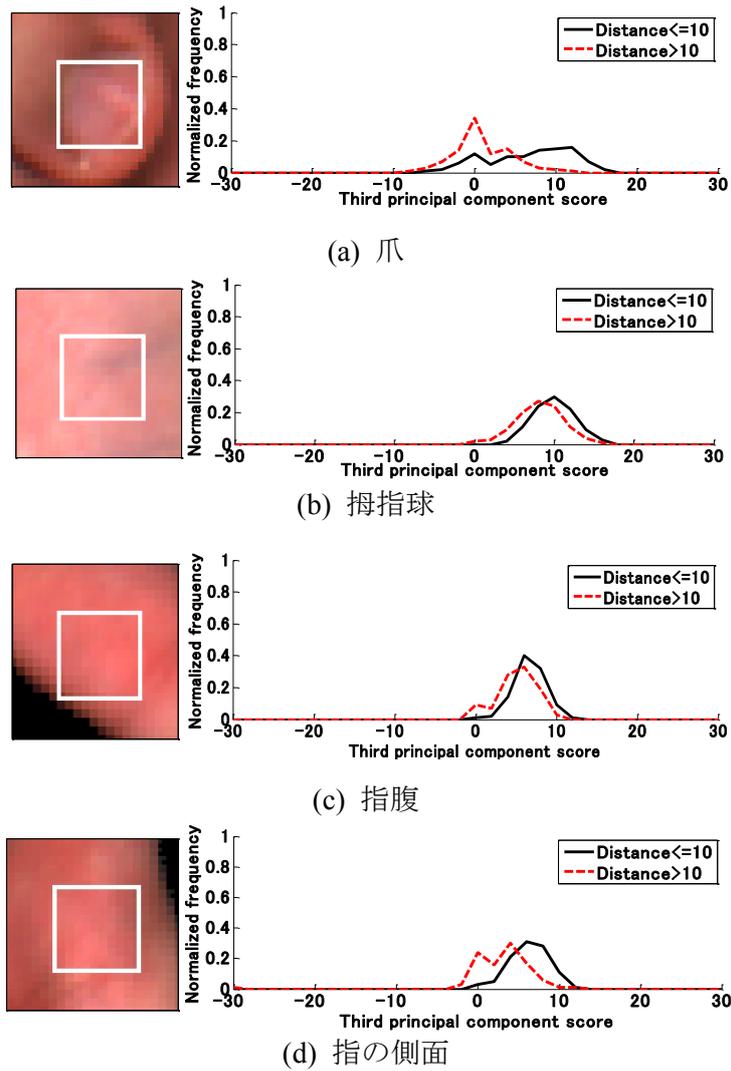


図 4-18 中心付近の画素と中心から離れている画素の持つ色の分布の違い

$$Density = 1 - \left( \frac{1}{20\sqrt{2MN}} \sum_{j=1}^M \sum_{i=1}^N \sqrt{(x_{ij} - x_{mj})^2 + (y_{ij} - y_{mj})^2} \right) \quad (4-13)$$

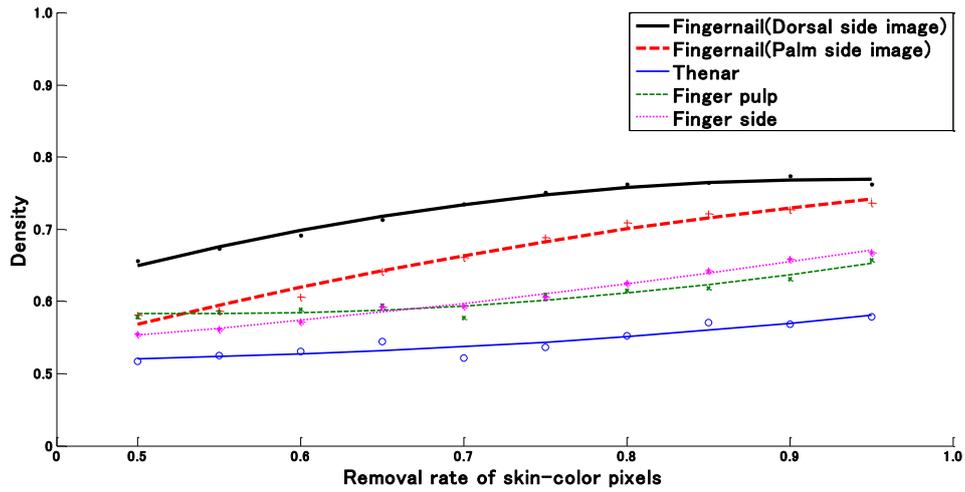
$$x_{mj} = \frac{1}{N} \sum_{i=1}^N x_{ij}, \quad y_{mj} = \frac{1}{N} \sum_{i=1}^N y_{ij} \quad (4-14)$$

ここで,

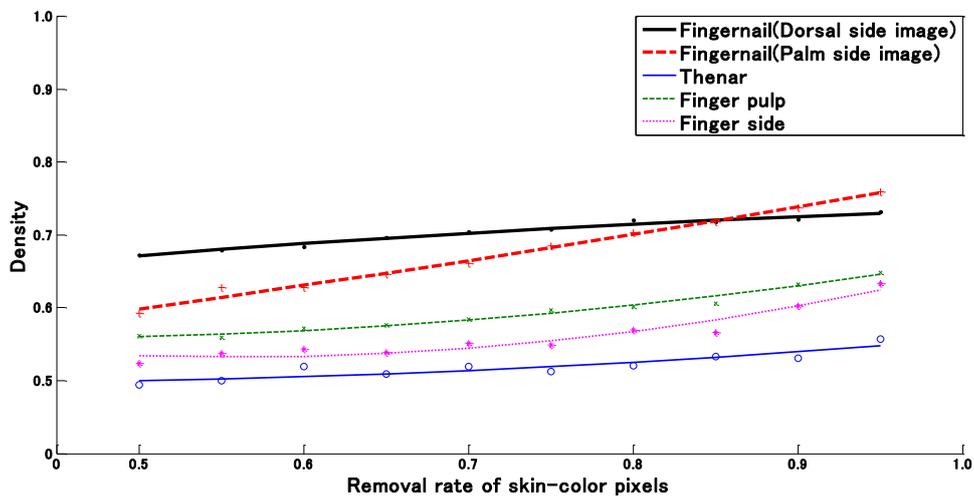
$N$  :  $j$  番目の ROI 画像中で抽出される手領域画素の数

表 4-2 抽出比率の違いによる密集度調査に用いる被験者の性別と手の色

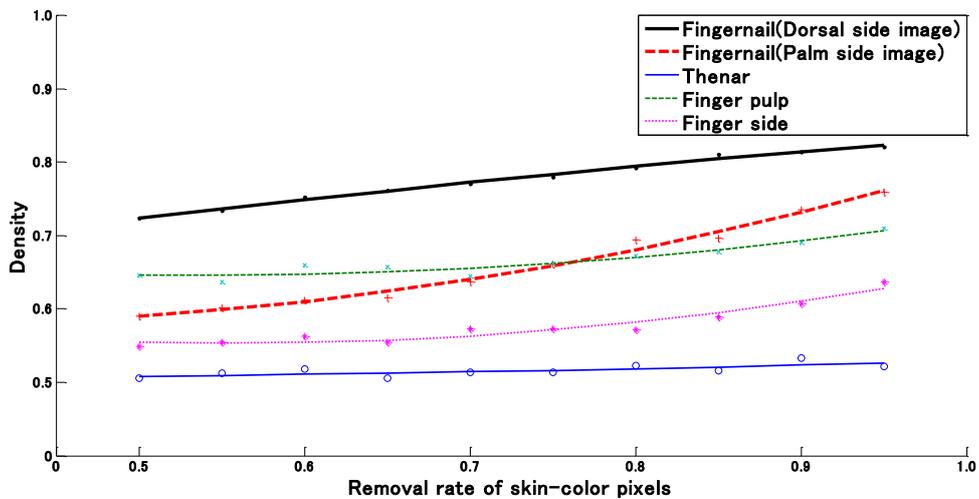
Subject number	1	2	3	4	5	6
Gender	Man	Man	Woman	Man	Man	Man
Fingernail	(229,126,126)	(248,130,129)	(234,132,126)	(227,121,120)	(204, 148, 131)	(215, 128, 122)
Palm side	(227,129,120)	(230,132,122)	(228,129,110)	(228,129,110)	(203, 118, 106)	(233, 131, 119)
Dorsal side	(199,126,105)	(200,118,103)	(201, 130, 106)	(201, 130, 106)	(168, 93, 74)	(175, 101, 79)



(a) 被験者 1

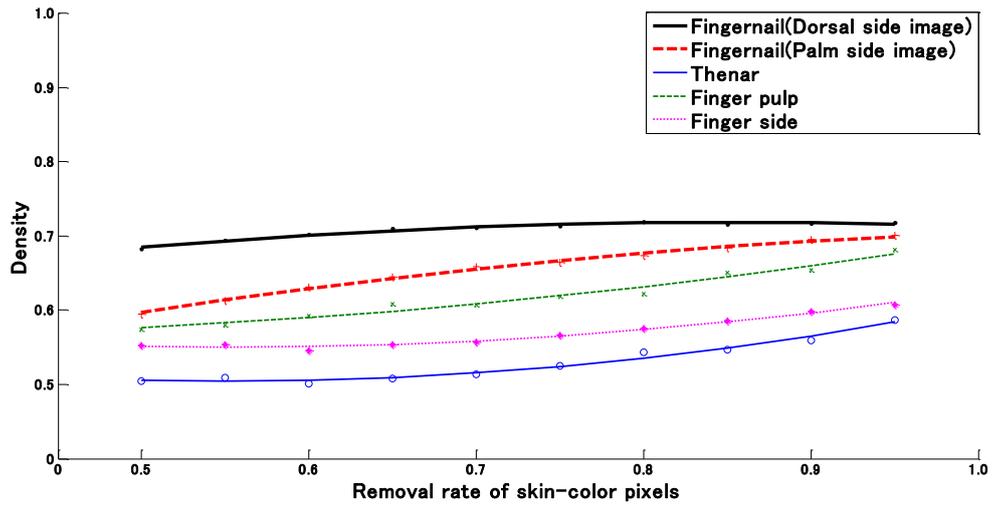


(b) 被験者 2

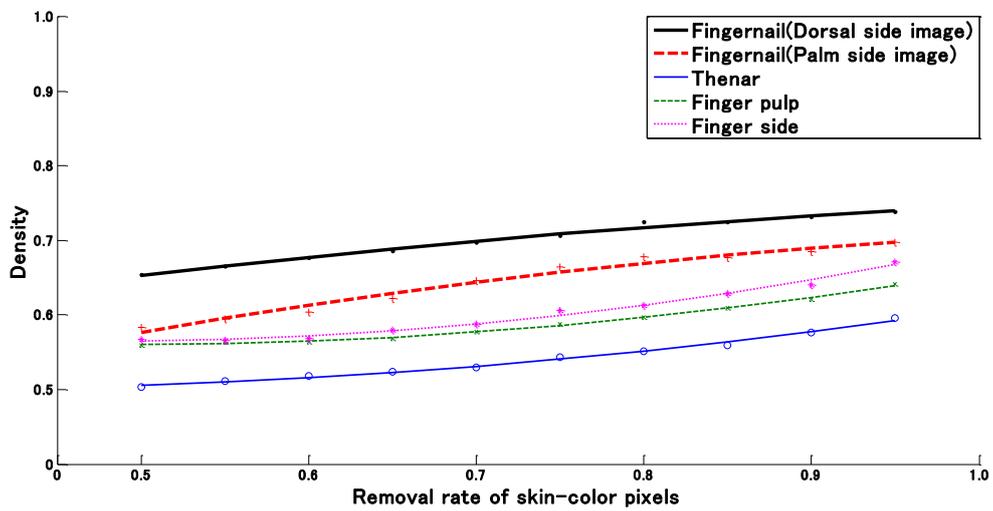


(c) 被験者 3

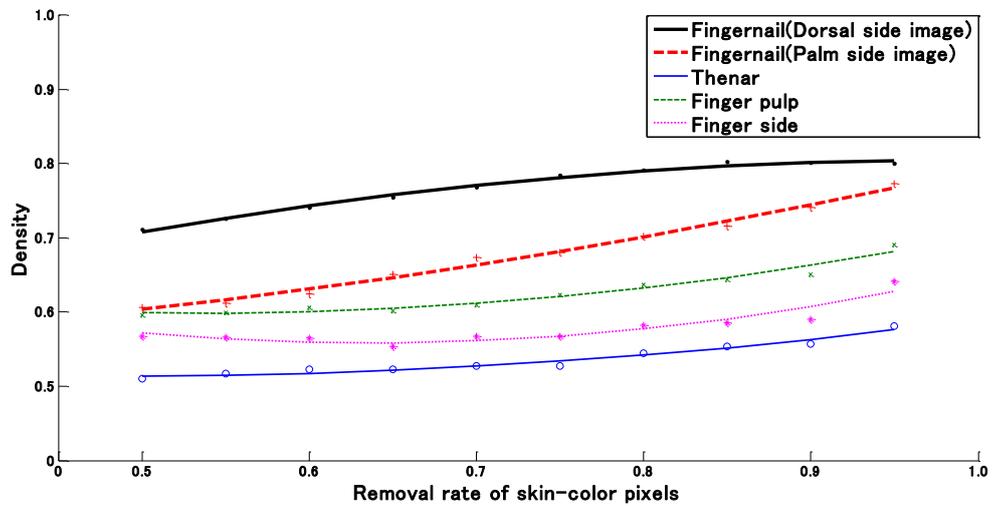
図 4-19 被験者 1 から 3 における肌色画素の除去率と密度の関係



(a) 被験者 4



(b) 被験者 5



(a) 被験者 6

図 4-20 被験者 4 から 6 における肌色画素の除去率と密度の関係

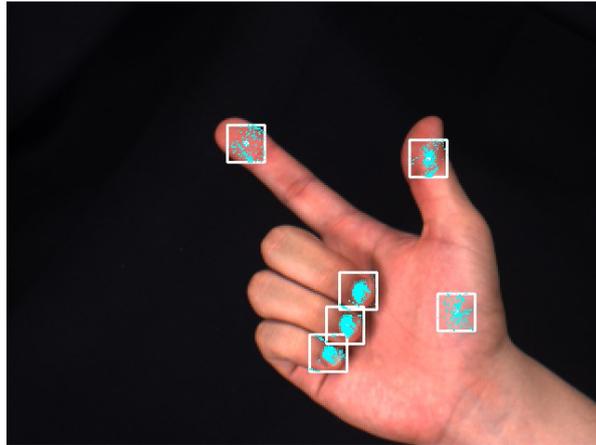


図 4-21 上位 20%の第 3 主成分得点を持つ画素の抽出結果

図 4-19 と図 4-20 から、爪と他の肌の密度が同じになる限界の最も厳しい所が肌の除去率約 75%，すなわち上位約 25%残した状態であることがわかる。また、指腹と指の側面のどちらが密集しやすいかは被験者ごとに異なっているが、拇指球は共通して最も密集度が低くなっていることがわかる。これは図 4-18 で述べたことと一致しており、密集度を用いると拇指球の除去精度が高くなる確率が大きい。上位 25%は限界点のため、それより少ない上位 20%に属する第 3 主成分得点を閾値として画素を抽出する。図 4-21 が爪尤度の上位 20%の画素を対象として抽出した実際の結果である。この図は予想通りの結果が現れたことを示している。本手法はランキングにより相対的に評価するため、輝度低下により爪と肌の色が近くなる問題の影響を受けにくい性質を持つ。一方で、前腕回旋角度によっては肌に色飽和が大きい照りのある部分が生じる時に問題が生じる。照りのある部分は局所的に爪尤度が高くなる傾向を持つため、肌であっても高爪尤度画素が密集分布することがある。そのため、照りの生じやすい指の側面部を爪として誤検出することがある。また、図 4-19 と図 4-20 で述べたように、個人差により指腹の方が密集する人も存在する。

密集度を数式化して爪か肌か判別する。式(4-13)と式(4-14)のような画素すべてを対象にした計算法だと、ROI の範囲の小ささから元々密集しやすく、画素数が少ないことから数値変化も大きい。ゆえに、図 4-19 と図 4-20 のような平均では差が出て、実際には誤検出することが多かった。よって、著者は別の手法として、密集画素と拡散画素の比を特徴量とする手法について提案する。この特徴量は画素の位置関係の情報を含む。まず、 $k$  番目の ROI 画像内で爪尤度が高い上位  $Rate_{upper}\%$  の画素を抽出して、2 値画像に白色点を打つ。この 2 値画像を  $O_k$  と定義する。そして、この画像を  $7 \times 7$  のメディアンフィルタにて平滑化をかけた画像を  $C_k$  と定義する。フィルタサイズは経験を基にした最適な値である。この画像は密集していた画素だけでなく、密集部の中にあつた黒色ノイズ画素も白色点に補完され含まれる。そして、この 2 つの画像の画素ごとの排他的論理和を取った画像を  $S_k$  と定義する。排他的論理和をとるのは、密集していない画素だけでなく、白色点に補完された画素も  $S_k$  に含ませるためである。これにより画素の補完が生じたことによる密度変化の影響を減少させる。この時、次の式(4-15)が真ならば、その ROI 画像は爪を含んでいると判定する。

$$\frac{N_S^k}{N_C^k} < Th_{CS} \quad (4-15)$$

ここで,

$N_C^k$  :  $C_k$  内の白色点数

$N_S^k$  :  $S_k$  内の白色点数

$Th_{CS}$  : 爪か肌かを区別するための閾値

## 4.6 評価実験

評価実験では2つの事を評価した。1つ目は前腕回旋角度と爪検出の成功率の関係である。これを求めることで提案システムがどの前腕回旋の範囲で使用できるのか、および精度低下が緩やかなのか急なのかを知ることができる。2つ目は、提案した爪か否かを判定する手法の性能である。これを求めることで提案手法がどのように働き、どのような有効点を持つのかを知ることができる。評価は実験1で求めた提案手法の有効範囲内で行い、肌を爪ではないとして除去した割合を求めた。また、2つの判別方法をそれぞれ使用した場合と共に使用した場合の違いについても調査した。さらに、ROIサイズの違いでどのような影響が出るかについても調査を行った。

2つの評価実験では人差し指から小指までの4指を対象とした。拇指の爪は他の指と前腕回旋による影響が異なり、回旋の影響を推測しづらくするため、除外した。視認できる4指の爪を重心ずれ10pixel以内で全て検出でき、肌を爪として検出しなかったとき、検出に成功したと定義した。被験者数は日本人学生5名とした。被験者ごとの性別、爪、掌領域の肌、および手の甲領域の肌の平均RGB値、そして検出に使用したパラメータは表4-3にまとめた。ROIのサイズは爪が小さい被験者は41×41 [pixel]、爪が大きいため足りなかった被験者は51×51 [pixel]とした。その他の値は著者の経験を基にした最適値を設定した。撮影時刻は外部光の入らない夜間とした。撮影距離は60cmとした。掌がカメラ正面を向くときを-90°とした。前腕角度は-90°から90°までを10°刻み、19パターンで変化させた。-90°から0°の間では曲げた指の爪を、10°から90°の間では伸ばした指の爪を検出した。指を曲げる際は完全に曲げきり、指を伸ばす際は完全に伸ばしきった。1人あたり、各回旋角度で合計500枚、総合計9500枚の画像から評価した。爪が写らない爪0枚の画像から4指すべてが写るまでの爪4枚の画像まで、100枚単位で爪の枚数を増加させた。図4-22は実際の各被験者の手である。

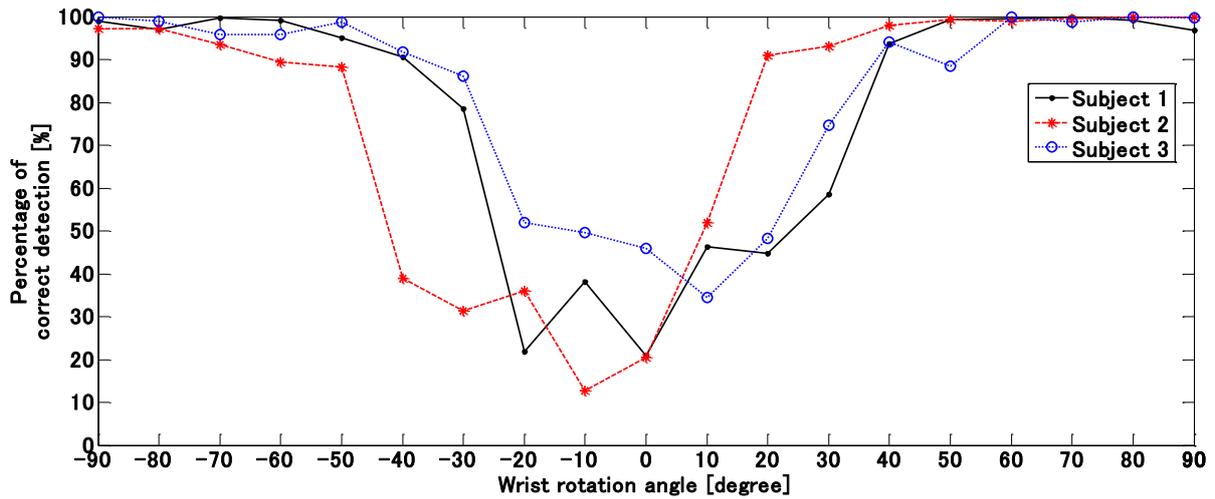
図4-23は提案手法を利用した際の実験結果である。提案システムは個人差で精度にばらつきがあるものの、掌側が主に写る-90°から-50°の間、手の甲側が主に写る40°から90°の間で少なくとも85%の確率で爪を検出できた。この値は画像1枚で爪検出をした場合の確率であり、数枚の画像の時系列処理を加えると、より精度向上することを確認している。これは掌の有無に

表 4-3 爪検出精度の評価実験に用いた手の情報とパラメータの数

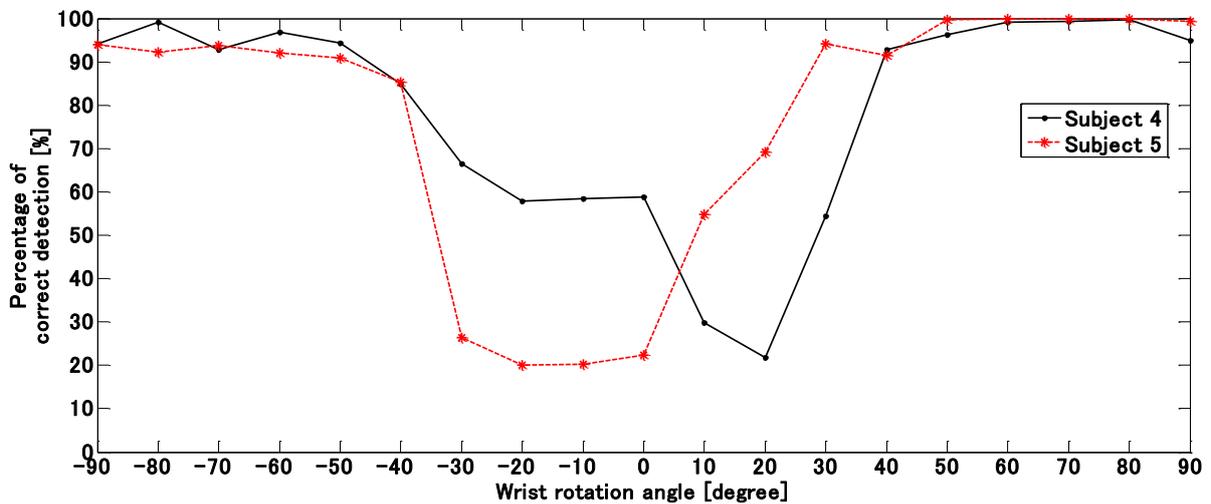
Subject number	1	2	3	4	5
Gender	Man	Man	Woman	Man	Man
Fingernail	(229,126,126)	(197,123,126)	(246,134,125)	(248,130,129)	(224,125,128)
Palm side	(227,129,120)	(182,125,118)	(244,156,123)	(231,133,122)	(197,121,115)
Dorsal side	(199,126,105)	(168,121,108)	(229,151,119)	(193,113,98)	(184,126,113)
$Th_{layer}$	6.0	6.0	6.0	6.0	6.0
ROI size	41 × 41	41 × 41	41 × 41	51 × 51	51 × 51
$Th_{connection}$	-3.0	-3.0	-3.0	-3.0	-3.0
$Rate_{upper}$	20.0	20.0	20.0	15.0	15.0
$Th_{CS}$	1.2	1.2	1.2	1.2	1.2



図 4-22 評価実験で用いた手の例



(a) ROI サイズが  $41 \times 41$  となった被験者の爪検出結果



(b) ROI サイズが  $51 \times 51$  となった被験者の爪検出結果

図 4-23 前腕回旋角度と爪検出率の関係

関わらず爪検出できることを示している。また、指が大きく重ならないような、手指形状推定に利用される実用範囲の多くを含んでいる。よって、本手法は手指形状推定に利用するために有効な範囲を確保できていると考えられる。 $-40^\circ$ から $30^\circ$ の区間では、手の回旋により画像に写る爪のサイズが小さくなる。また、光の反射が弱くなる。それにより、爪候補を決める段階で一部の

表 4-4 肌を正しく除去できた割合

Percentage of skin elimination [%]			
Independency of areas			
ROI size	Thenar	Finger pulp	Finger side
41 × 41	97.00	88.83	93.32
51 × 51	93.28	61.19	99.68
Distribution density			
ROI size	Thenar	Finger pulp	Finger side
41 × 41	99.52	89.04	66.48
51 × 51	100.0	95.29	89.95
Proposed method			
ROI size	Thenar	Finger pulp	Finger side
41 × 41	99.93	95.56	97.88
51 × 51	100.0	97.41	99.87

爪が抜き出せないという現象が生じていた。これが精度急降下の主原因であった。これはここまで述べた手法では解決が困難な原因であり、精度の限界を作り出すものである。よって、提案手法の有効範囲は $-90^{\circ}$ から $-50^{\circ}$ 、 $40^{\circ}$ から $90^{\circ}$ の間であると結論付けた。なお、爪の面が、カメラの撮影平面に対して並行でないことが理由で光の反射が弱くなり、爪領域が爪の候補にならなかったケースは有効範囲内にもあり、これが検出精度を低下させる一因となっていた。第3主成分は色相軸のため、輝度の変化による影響は受けにくい。しかし、反射が弱いと色がつぶれて色相自体が変化することがあるため、完全に影響を受けないわけではなかった。よって、取得できない爪領域をなくすために、爪領域候補の抽出過程を改善することが、別の課題として重要である、と言える。これについて取り組んだ成果は次節で述べる。

表 4-4 は 2 つ目の評価実験の結果である。この評価実験から、提案した手法の弱点、有効点、および補完関係がわかる。まず爪領域の色の独立性を用いると、4.5.2 項で述べたように、指腹の除去性能が悪くなっていた。画像を確認したところ、しわが影響していることが確認された。また、ROI のサイズが大きくなると、性能が大幅に低下する可能性があることもわかった。これは、ROI のサイズが大きくなるほど、しわなどの色が急激に変化する部位を含める危険性が高まるためであると考えられる。一方で、拇指球や指の側面部の除去は良好であった。高爪尤度画素の分布密度を用いる手法に関しては、前者の優劣点と状況が逆転する現象が見られた。この手法では 4.5.3 項で述べたように、照りによる指の側面部の除去失敗が多く見られた。一方で、拇指球や指腹の除去性能は良好であった。この手法では ROI サイズが小さくなるほど、性能が低下する現象が見られた。ROI が小さいと、たとえ中心から離れていても密集する危険性が高まるということが影響していると考えられる。以上よりわかることは、2 つの爪の判別方法はそれぞれの弱点を補い合う性質を持っていることである。すなわち、指腹は分布密度の判定で、指の側面部は領域の独立性の判定で爪でない判定されるため、共に用いると補完現象がおきる。これが、誤検出しやすいすべての肌領域を、95%を超える精度で除去できるようにした大きな理由と考えられ

る。以上のことから、爪検出精度を高めるためには、本手法で提案した2つの判別手法を共に用いることが重要であると言える。

以上で述べたシステムは現在、インタフェース構築のための手指形状推定[72]と手話認識のための手の姿勢推定[73]を目的として外部グループによって利用されている。今後、さらに爪検出精度と使用可能範囲を改善し、より応用が可能なシステムを目指したい。

## 4.7 爪領域候補の抽出過程を改善する試み

本節では図 4-24 に示すように、これまでの爪か肌かを判別して爪検出できるようにする過程より前の、爪領域候補の抽出を行う過程に新しく着目点を写し、改善への試みを行った。ゆえに、本節は本論文の主題である 4.6 節までとは離れた話題となる。4.6 節までは従来の「爪と肌の色合いは異なるので画素ごとの色情報だけで爪が検出できる」を否定し、「爪と肌の色には違いがあるが、爪と似た色を持つ肌領域があるため、それを除去するアルゴリズムを付加することで爪が検出できる」とした。本節の内容はさらに「爪と肌の色には違いがあり」という前提を否定し、「爪と肌の色が同じような色でも」とすることを目指すための処理である。

これまでの爪領域候補抽出段階は 4.4 節で述べたように、第 3 主成分得点を固定閾値で抽出していた。ゆえに、図 4-25 のように画像によっては爪の一部が抽出できない場合や、大量の領域

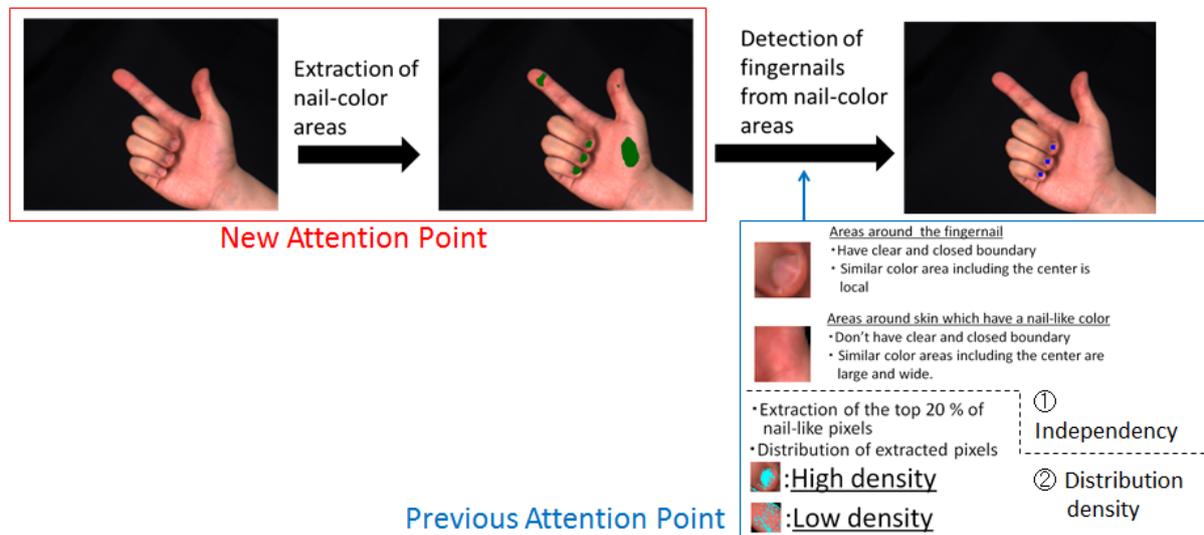


図 4-24 爪検出システムの概要図と着目点

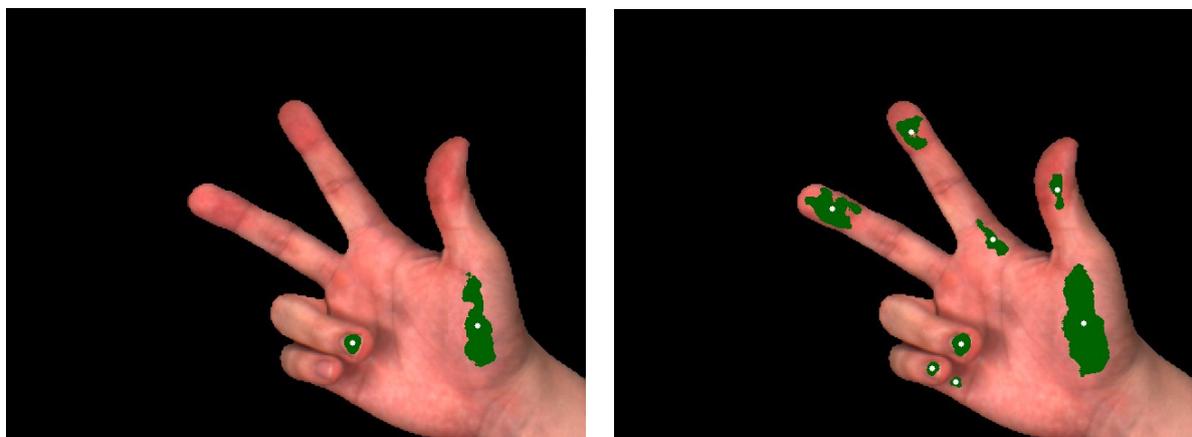
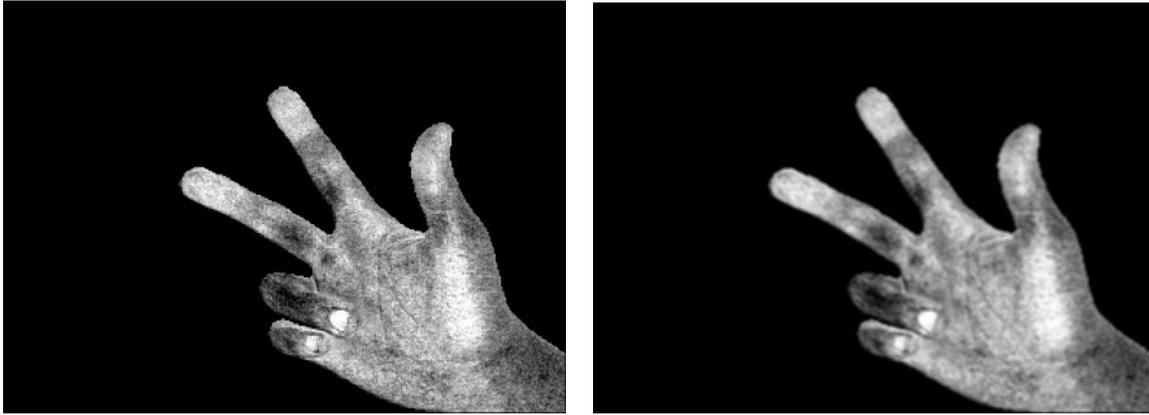


図 4-25 爪領域候補の抽出時における問題



(a) マップ画像

(b) マップ画像を平滑化した画像

図 4-26 第 3 主成分得点のマップ画像とマップ画像の平滑化画像

が抽出される場合が生じる。よって、適応的な閾値処理が必要となる。しかし、一般的な、Otsu の判別分析法[33]を応用した適応的閾値処理は用いることができない。適応的閾値処理のアルゴリズムはヒストグラムが二峰性の山にならないとクラスの分割を自動的に行うことができないために判別分析が適切に行われず閾値が正しく設定できない。しかし、手領域中に存在する爪領域の割合が微小のため、ヒストグラムを作ると爪領域の情報が反映されず単峰性となる。ゆえに使用できない。図 4-15 のマップ画像で爪領域がはっきりわかる大きな理由は爪と肌で周辺画素との色合いの様子が大きく異なることである。つまり 4.5.3 項で用いた、爪領域は同じような色合いでまとまっている一方で肌領域は色合いがまばらである、という座標空間を考慮した局所的な色合いの違いを上手く利用することが、爪領域を取得するための大きな鍵となる。ただし、4.5.3 項と同じことを画像全体で行おうとすると、ROI の設定と密集度の計算を最大で画素分の 307,200 回、背景や明らかに異なると判定できるところを除いても膨大な回数が必要となる。ゆえに、空間的な情報を含む新たな手法で実現する必要が生じた。

著者は喜多[34]のように、2 次元濃度共起ヒストグラム(Co-occurrence Histogram : CH)画像[74]を応用することによって、実現を試みた。まず、式(4-16)を用いて画素値を整数にする。そして、-10.0 から 8.0 が 0 から 255 になるように整数値を割り当て、図 4-26(a)のマップ画像を生成する。

$$Third'_{xy} = \begin{cases} -10.0 & (Third_{xy} \leq -10.0) \\ Third_{xy} & (-10 < Third_{xy} < 8.0) \\ 8.0 & (Third_{xy} \geq 8.0) \end{cases} \quad (4-16)$$

次に、この画像に対して 5×5 の単純平滑化フィルタをかけ、図 4-26(b)の画像を生成する。平滑化画像は周りの画素との差が影響として現れる。爪の場合、周辺領域と色が近いので、領域の淵を除き画素値は変化しにくい。一方で肌の場合、周辺と色が異なる傾向があることから、平滑化すると画素値の変化が大きくなりやすい。CH 画像は式(4-17)で定義した。画像は一辺 256 ピクセルの正方形である。



図 4-27 二次元共起ヒストグラム画像

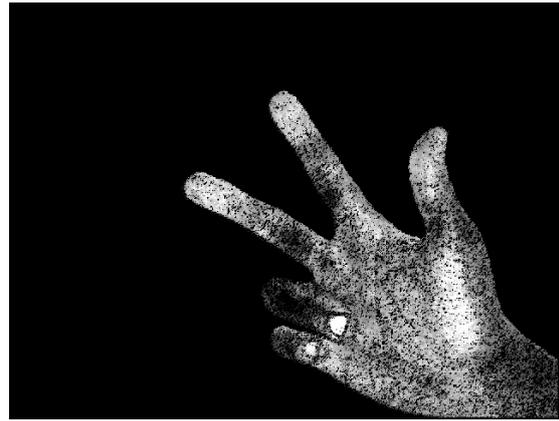


図 4-28 CH 画像を用いた画素の除去結果

$$CH:(P_{mn})=hh(Value_{xy}^{original}, Value_{xy}^{smoothing}) \quad (4-17)$$

$$(0 \leq m \leq 255, 0 \leq n \leq 255)$$

ここで,

$Value_{xy}^{original}$  : 正規化画像の座標  $(x, y)$  が持つ値

$Value_{xy}^{smoothing}$  : 平滑化画像の座標  $(x, y)$  が持つ値

$Value_{xy}^{original}$  と  $Value_{xy}^{smoothing}$  の値の差が大きいほど、つまり周辺画素の影響で値が変わってしまった画素程 CH 画像において画像の対角線から離れた位置に分布する。ゆえに、対角線からどれだけ離れたかが周辺とどれだけ色が異なるかを知るパラメータとなる。爪領域は基本的に高い第 3 主成分得点を持つ確率が高いことを考慮し、著者は CH 画像の左上端を起点として対角線に対して角度  $\pm\theta$  だけ傾けた直線を引き、その直線よりも外側に位置する画素をマップ画像から除去、すなわち画素値を 0 とした。図 4-28 は閾値  $Th_{\theta} = \pm 5^{\circ}$  で除去を実行した結果である。この処理によってまだら模様が鮮明になったことがわかる。よって、この画像に再度平滑化をかけると、爪領域の値はほとんど変化せず、肌領域の多くの領域で画素値が低下する。以上の処理により、肌寄りの爪領域も全体と比較すると相対的に高い値を持つようになる。最後に画素値に対して固定閾値を用いて領域抽出を行い、平滑化、ラベリングで領域抽出を行う。本論文では閾値を 200 とした。領域の抽出結果は図 4-29 に示すように、これまでの爪らしい色を持つ部分に肌寄りの色を持つ爪領域が追加されて抽出されることがわかる。

評価実験では、爪と肌を判別する処理を加えた完全な爪検出システムで行い、4.6 節までに述べたシステムと爪検出率を比較した。実験環境及び条件は 4.6 節と同じである。被験者は日本人学生 2 名とした。表 4-5 に被験者の性別、領域別の RGB 平均色、及び各パラメータの値を示す。なお、被験者 1 は 4.6 節の被験者 1、被験者 2 は 4.6 節の被験者 4 である。

図 4-30 に実験結果を示す。実験結果から本節提案手法の全体性能は  $-90^{\circ}$  から  $-50^{\circ}$  の間で 90%以上、 $50^{\circ}$  から  $90^{\circ}$  の間で 85%以上の爪検出率となった。図 4-31 が 4.6 節までに述べたシ

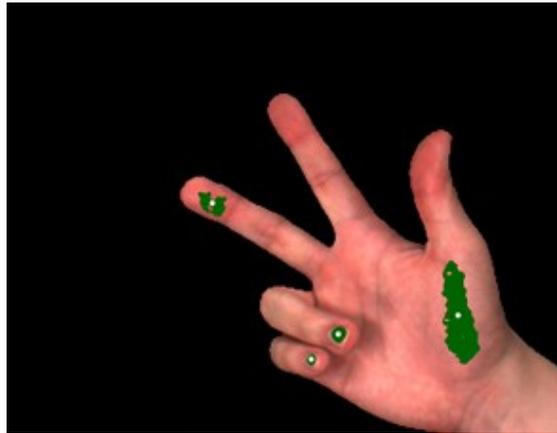


図 4-29 新しい抽出手法による爪領域候補の抽出結果

表 4-5 爪領域抽出の改善手法評価のために使用した参加者の手とパラメータの情報

Subject number	1	2
Gender	Man	Man
Fingernail	(229,126,126)	(248,130,129)
Palm side	(227,129,120)	(231,133,122)
Dorsal side	(199,126,105)	(193,113,98)
ROI size	41 × 41	51 × 51
$Th_{connection}$	-3.0	-3.0
$Rate_{upper}$	20.0	15.0
$Th_{CS}$	1.2	1.2
$Th_{\theta}$	±5.0	±3.0

ステムとの検出率の差である。−40° から−20° までは検出率の向上が見られた。−40° から−20° で精度向上することは、これまで第1段階で抽出できなかった爪が抽出され、爪検出できたことを表しており、本節で提案した手法の貢献を表す。一方、本節の提案手法は10° から90° の手画像では逆に精度が低下するという大きな問題が存在することが判明した。この問題が生じる主原因は爪と指の側面領域が結合しやすいことである。本手法での爪領域抽出法は第3主成分の値が低い爪領域の画素を抽出し、同じ値を持つ肌画素を抽出しないことが可能であるが、元々抽出されやすい拇指球、指腹、指の側面の抽出範囲は逆に広く、大きくなる傾向が確認された。そのため、図4-32のように爪領域と指の側面領域が結合してしまい、重心がずれ、肌領域に重心がくる現象が従来よりも生じた。これにより、第2段階の爪か肌かの判別で肌として判別され、除去される爪が多くなった。また、前腕回旋角度が0度に近すぎる時や、指の曲げが浅すぎて爪の色が著しく変化している場合では、爪領域画素の一部が除去されるため、抽出できないことも判明した。手の甲側が主に写る手画像において精度を改善するためには事前に掌側が写っているか、手の甲側が写っているかを判別する手法を組み込むことが有効であると考えられる。それにより、掌側が写っている際は本手法を、手の甲側が写っている際は従来的手法を用いるように分岐処理が可能となる。判別は、各画像の画素分布中心の第3主成分得点を調査することで可能である。式(4-11)のRGB平均値を第3主成分軸の生成に用いた輝度を落とした掌側だけの画像の

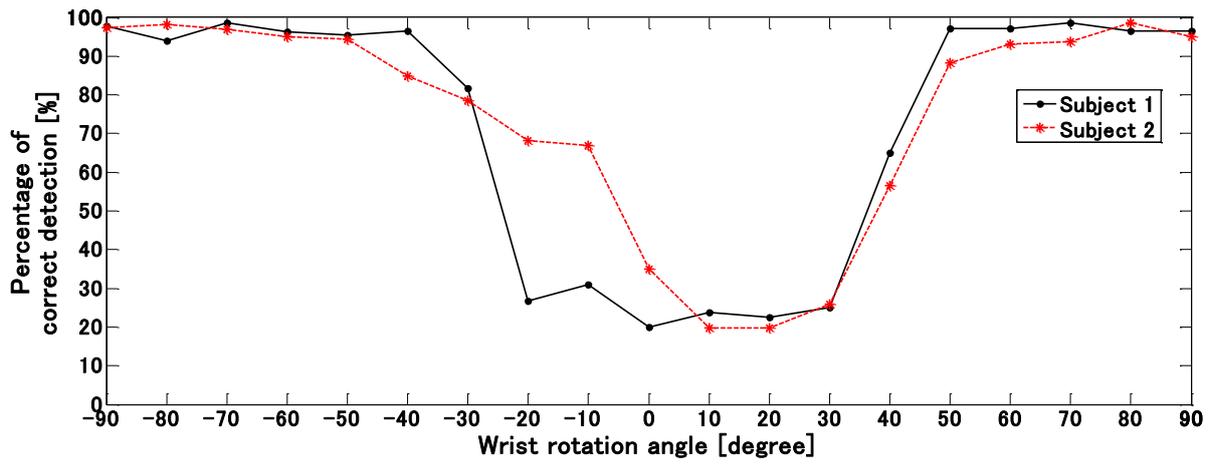


図 4-30 爪領域候補抽出の改善手法を用いた時の爪検出率

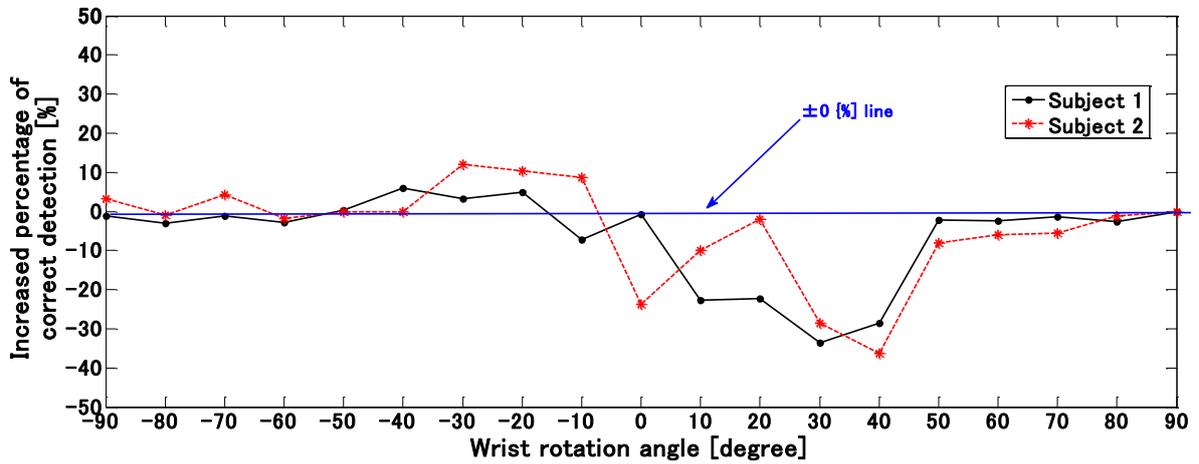


図 4-31 爪領域候補抽出の改善手法の有無による爪検出率の増減

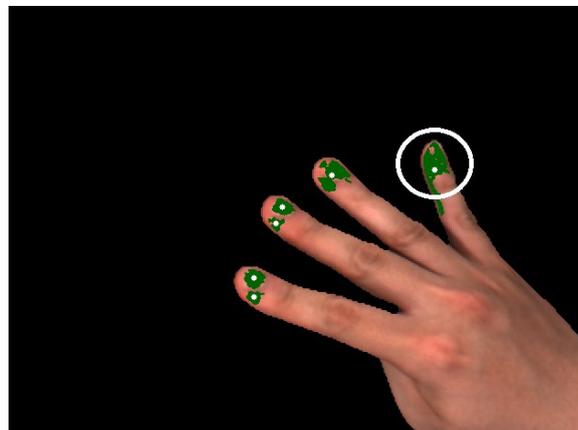


図 4-32 爪領域と指の側面領域の結合による爪領域候補抽出の失敗

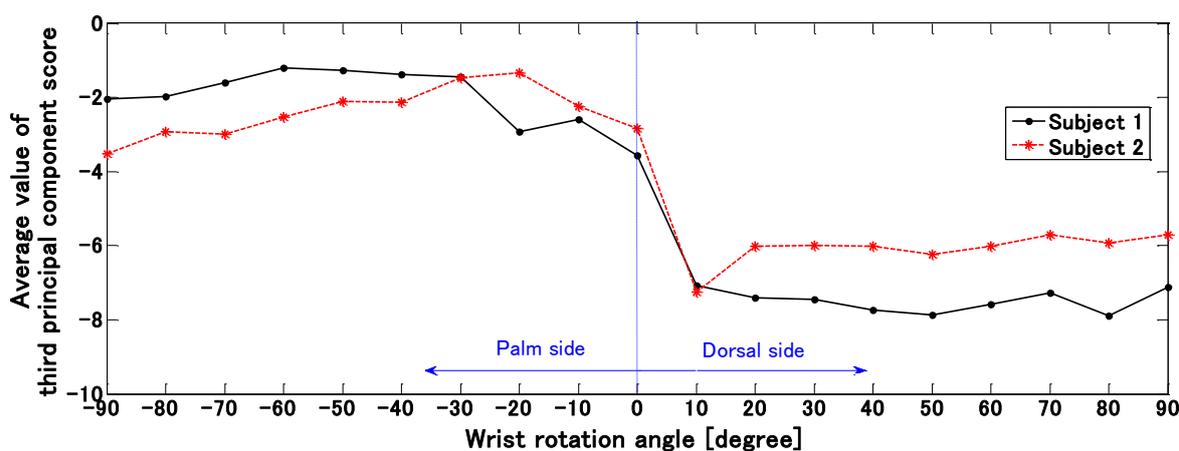


図 4-33 回旋角度の変化による平均第 3 主成分得点の変遷

RGB 平均値である  $\mathbf{x}_{mean}^{cal}$  を基準として用い、平均第 3 主成分得点  $Third_{ave}$  を式(4-18)で定義する.

$$Third_{ave} = (\mathbf{x}_{mean} - \mathbf{x}_{mean}^{cal})^T \mathbf{e}_{third} \quad (4-18)$$

すると、前腕回旋角度との関係は図 4-33 のように掌側と甲側で値が急激に変化することがわかる。つまり、キャリブレーションで掌側と甲側の画像を 1 枚ずつ撮影し、差を調査することで、判別閾値を設定できる。そして、それを用いて手の向きを求められ、爪領域抽出法の切り替えが可能となる。

次に、回旋が 0 度付近の時と、指の曲げが浅すぎるときの問題であるが、この問題が生じる主原因は、CH 画像に輝度の情報が含まれないことがある。この状態の爪は輝度が小さくなるだけでなく、色がかなり肌色寄りとなり、第 3 主成分座標値が小さくなっている。本手法は第 3 主成分座標が小さいほど、抽出される確率が一定比率低くなる方法を用いているため、輝度情報なしではこの問題に対応できない。一方で、輝度情報を大きく活用すると照明に大きく影響され検出率は下がると考えられる。よって、この問題を解決するためには輝度の情報を適度に利用する手法を新たに提案する必要があると考えられる。

以上の通り、爪領域抽出の改善のための技術はまだ大きな発展が必要である。しかし、本節の方法によって色のまとまり具合を利用した方法の利点と問題点を取得できた。この知見により、現在著者はより抽出精度の高い分布密度の利用方法を見つけつつある。また、輝度情報と第 3 主成分得点の相対的關係性を用いることにより、第 3 主成分得点の小さい爪領域を確実に残しながら、肌領域の大部分を除去できる方法を発見している。今後、これらを精査してより高性能な手法を構築し、よりロバスト性の高い爪検出システムを作る手法、および手掛かりを社会へ提示したいと考えている。

## 4.8 第4章のまとめ

本章では主題として、爪領域周辺に特徴的に生じる色の分布状態を利用して爪か肌かを判別する手法を提案し、掌領域を含む手画像でも爪検出可能な爪検出システムを実現する方法について述べた。

最初に、爪検出を行う際に必ず行う行程である爪領域抽出に関する先行研究についてまとめ、これらの研究ではどの軸が最も爪らしさを表す軸であるかを評価していないこと、爪領域と肌領域では色が異なる、という前提を持っていることを述べた。そして、その前提条件がいつ成り立つかを検証実験により確かめた結果、掌側が主に写る手画像では成り立たないということを示した。これにより、掌側が主に写る手画像で爪検出を行うためには、本研究で提案する爪判別アルゴリズムが必要であることを示した。次に、肌領域と爪領域の画素分布から、肌領域の第3主成分軸が爪らしさを表す軸として利用できる可能性があることを示し、予備実験で他の軸と比較することにより、第3主成分軸の優位性を示した。

爪判別アルゴリズムの提案では、画像空間上の爪らしい色を持つ画素の分布に着目し、爪領域周辺では(1)色が独立している小領域が存在する、(2)局所的に見た時、特に爪らしい色を持つ画素は密集している、という2つの性質について述べ、それぞれを特徴量化し、判別アルゴリズムを提案した。

評価実験では日本人被験者5名により、前腕回旋角度と爪検出率の関係、および爪判別アルゴリズムによる、誤検出の多い肌領域の除去率について調査した。その結果、本提案手法を用いた新しい爪検出システムは掌側が主に写る $-90^{\circ}$ から $-50^{\circ}$ の間、手の甲側が主に写る $40^{\circ}$ から $90^{\circ}$ の間で爪検出率85%以上の確率で爪検出できること、つまり掌領域を含む手画像にも対応でき、手の裏表に関わらず爪検出が可能になったことが確認された。また、肌の除去率を確認したところ、提案した2つの判別アルゴリズムには欠点を補う補完関係があり、共に用いることにより誤検出しやすいすべての肌領域に関して除去率95%以上となることが確認された。

最後に、爪検出システムに残る課題である、肌寄りの色を持つ爪の領域抽出法の改善に触れ、色の分布密度を利用すると掌側が主に写る手画像への爪検出は改善が示唆される一方、手の甲側が主に写る手画像への爪検出には課題が残ることを示した。

## 第5章 爪検出システムを利用した別の応用例

### 5.1 仮想ネイルアートシステム構築の背景

爪はファッションを行う1つのパーツとして古くから親しまれているパーツの1つである。尹は美容サービス産業の現状と課題について研究を行い、ネイル分野についても調査した[75]。それによると、世界では、爪にネイル・ポリッシュを塗る歴史が紀元前 3000 年以上前の古代エジプト時代から存在し、ギリシャ・ローマ時代には上流階級の男女に流行し、中世・ルネサンス時代にはマニキュア術が発達し、手の化粧が最重要視されていた。日本では平安時代に記録が残っている。爪を簡単に装飾できる「ネイルチップ」は 1970 年代後半、ハリウッドで開発された。日本では、1985 年に日本ネイリスト協会が設立され、本格的な普及が進んだ。手の爪を装飾することをマニキュア、足の指を装飾することをペディキュアという。本論文ではマニキュアを扱う。また、爪の装飾はネイル・ポリッシュや爪に直接装飾するものなど、様々な方法が存在するが、本論文では脱着可能で、気軽にファッションが楽しめる、ネイルチップを用いたネイルアートに着目して話を進める。

ネイルチップを選び、購入する方法はネイルチップが置いてある店に行くか、インターネット上のカタログを見ることが主流である。つまり、実際に装着した印象がわからず購入することが多い。ネイルチップは通常、ネイルグルーという接着剤や両面テープを使用してネイルチップと爪を接着する。ゆえに、試着をすると商品であるネイルチップを傷つけるおそれがある。近年ではネイルチップに低接着性の粘着剤がついており、取り付け・外しが簡単に出来るワンタッチネイルが登場したが、不特定多数の人に使用することを考えると衛生面上の問題が生じる。ゆえに、店舗での試着は一部に限られ、有料となる場合が多い。

著者は、手の画像の爪にネイルチップを貼ることができる仮想ネイルアートシステムが構築できれば、試着が難しいという問題を解決できると考えた。また、手の画像とネイルチップがあれば利用できるシステムは、インターネット上で手画像を入力し、付けたいネイルチップを選択すれば合成表示してくれるサービスや、試着したいネイルチップを選択した後、カメラで動画撮影すれば、リアルタイムでネイルチップの様子を確認できるアプリが提供できるようになる。この利便性は、ネイルチップの選び方に大きな良い変化をもたらすと考えている。よって、著者らは第4章で述べた爪検出システムの手指形状推定以外の別の応用例として、仮想ネイルアートシステムの構築に関わろうと考えた。

仮想ネイルアートシステムを構築する上で重要な要素技術は、(1) 手領域を画像から抽出する技術、(2) 手領域から爪の情報を取得する技術、(3) ネイルチップを自然に見えるように重畳する技術、(4) 角度によって変化する照明条件に対応させて、ネイルチップの色を変化させる技術、の4つである。(1) はジェスチャ関連の技術で多数の研究がある。(2) は第4章で述べた研究に該当する、(4) は CG 関連で研究されている。そして、(3) に関する研究が未発達である。よって、著者はネイルチップを重畳する技術についての研究を行った。

ネイルアートにおける重畳を考えた研究は第4章でも取り上げた津田らの研究[65-67]がある。津田らの手法は、爪領域が指の軸に対して爪領域も指領域も常に対称であるという条件を想定している。また、爪が指先から飛び出した指のみを対象とし、爪領域を楕円フィッティングし、そ

の領域を塗りつぶすことのみ行っている。つまり、爪のサイズより大きいネイルチップには適用できない。また、前腕回旋がかかると指の軸に対する爪や指領域の対称性が崩れる。ゆえに前腕回旋も想定できない。著者が知る限り、前腕回旋に対応したネイルアートシステムが存在しないため、著者是对応するためのアルゴリズム構築を行うこととした。

## 5.2 仮想ネイルアートシステムの構成

仮想ネイルアートシステムの動作環境を図 5-1 に、フローチャートを図 5-2 に示す。システムはキャリブレーションとして、1 枚のネイルチップ画像を入力し、ネイルチップの変形がしやすいように標準化したネイルチップ画像を生成する。主要な処理では、まず手画像を入力し、爪領域抽出を行う。重心検出をした後に指の情報を基にネイルチップ重畳のためのパラメータ設定を行う。そして、標準化ネイルチップ画像を変形して手画像上にネイルチップを重畳する。指  $i$  に関するパラメータは図 5-3 に示す 5 つである。  $\phi_i$  以外は画像から取得するパラメータであり、  $\phi_i$  は指のモデルを考えて推定する。本章では、ネイルチップ重畳手法の本質的な性能評価を行うため、爪は緑色のマニキュアをマーカとして利用した。また、ネイルチップは一般的な縦長の形状のものとして想定した。カメラは Point Gray Research 社製の高速カメラ Firefly MV を使用した。画像サイズは VGA、撮像速度は 60 [fps] を使用した。システム動作速度は画像入力後から出力前までの処理において、最大で 30 [fps] 程度である。ただし、システムの最適化を行っていない。処理の実質的な計算コストは少ないため、最適化を行うことで撮像速度以上となると考えられる。

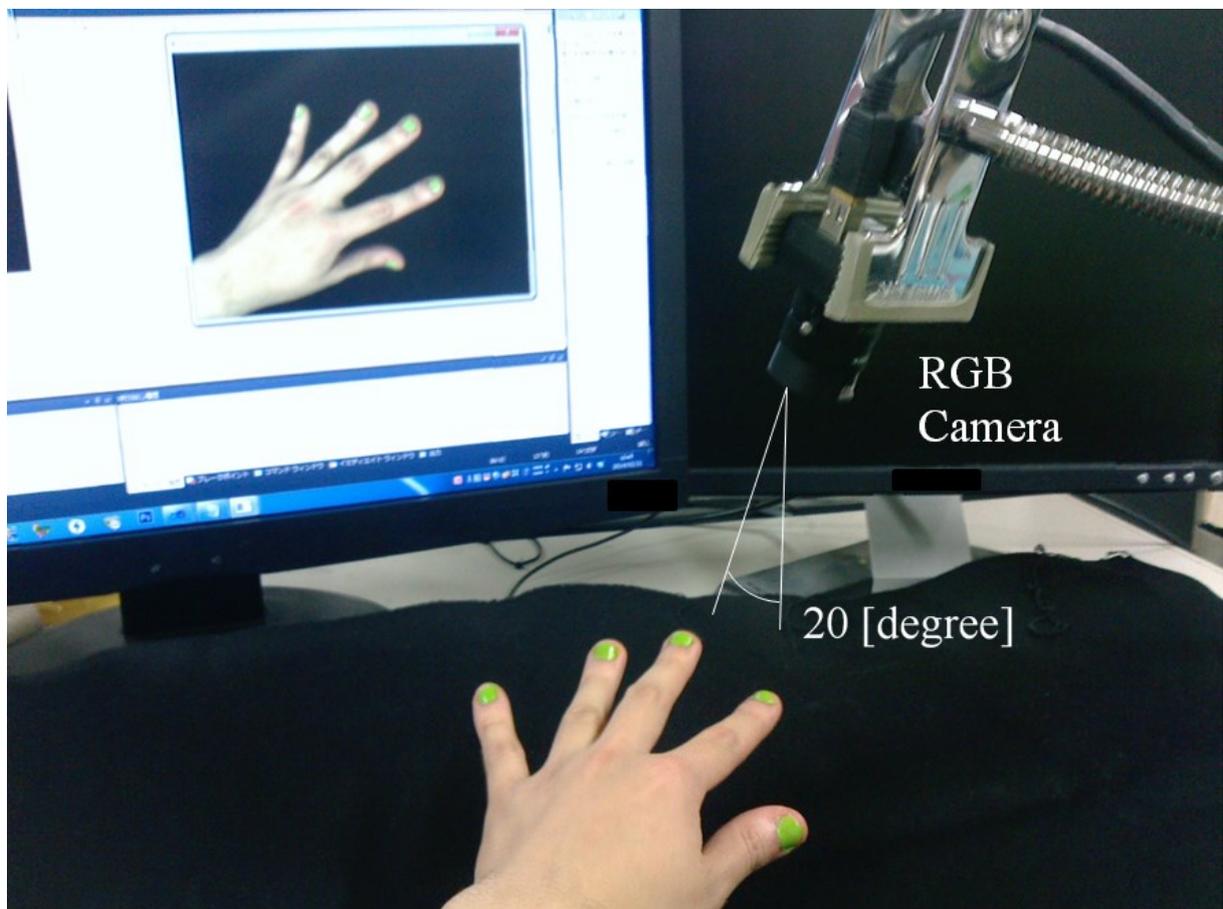


図 5-1 仮想ネイルアートシステムの動作環境

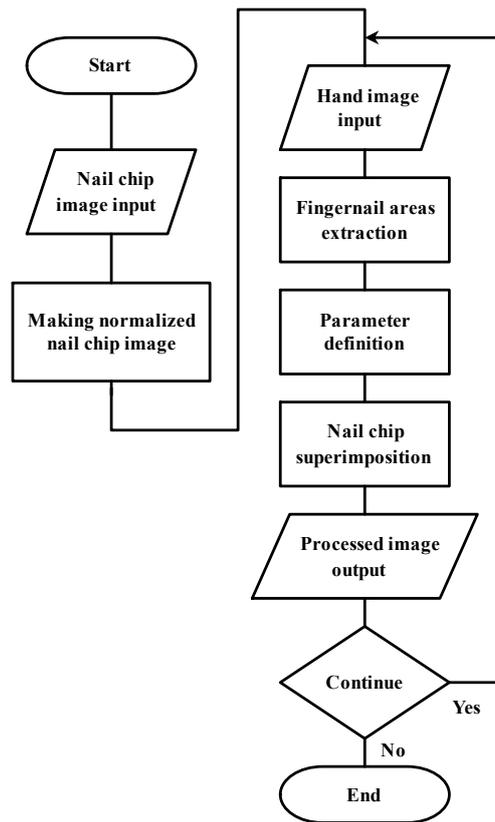


図 5-2 仮想ネイルアートシステムの処理の流れ

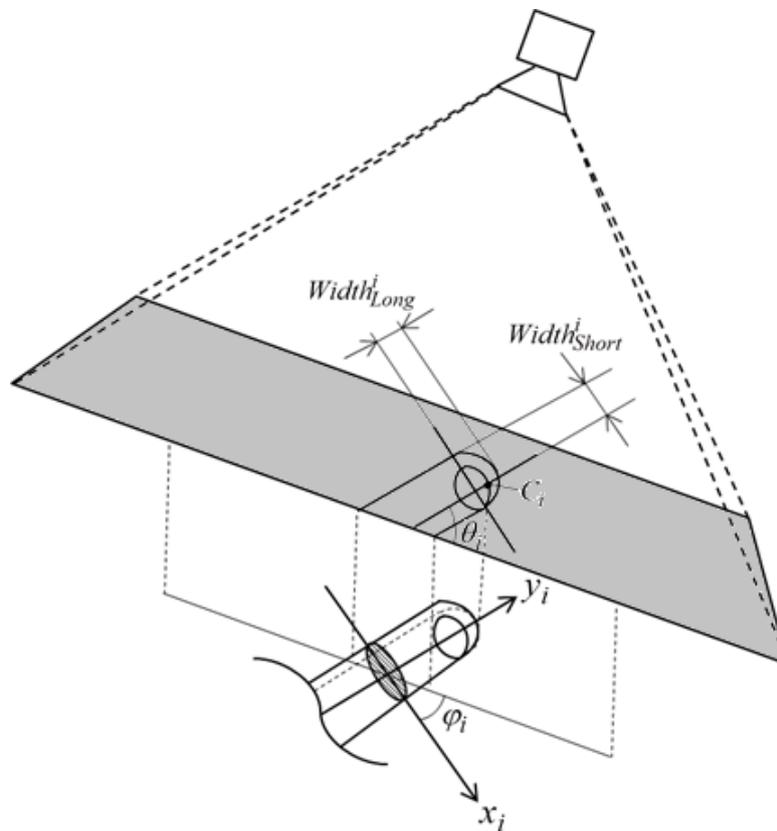


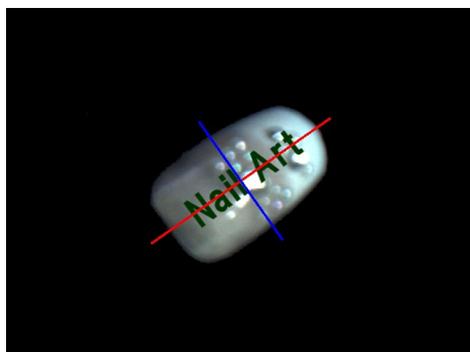
図 5-3 ネイルチップ重畳に必要な5つのパラメータ

### 5.3 ネイルチップ画像の標準化

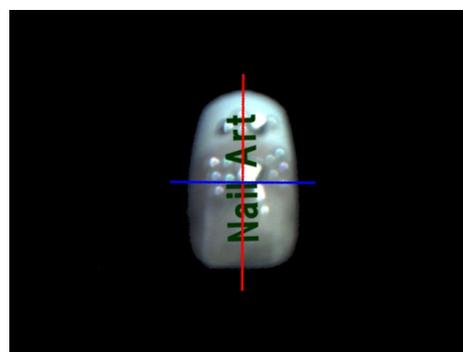
本システムを動作させる前にネイルチップが中央に写り、画像サイズが一定なネイルチップの標準化画像を生成する。まず、 $640 \times 480$  [pixel]の画像を入力し、ネイルチップ領域が存在する座標 $(x_i, y_i)$ を取得する。そして、その座標をデータとして主成分分析し、2本の主成分ベクトルを得る。主成分ベクトルを導出するためのデータ行列 $\mathbf{X}_D^{nailChip}$ は式(5-1)で表される。

$$\mathbf{X}_D^{nailChip} = \left\{ \mathbf{x}_i^T \mid \mathbf{x}_i = (x_i, y_i)^T \text{ if } i \in C_{nailChip} \right\} \quad (5-1)$$

この後は 4.2.2 項の式(4-2)を、データ行列と平均値を変更して用いることで分散共分散行列を生成し、分散共分散行列を固有値分解することで第1、第2主成分ベクトルの2つを取得することができる。図 5-4 は取得した主成分ベクトルを、領域の重心を原点にとって描画した画像である。第1主成分ベクトルは図 5-4(a)に示す長軸、第2主成分ベクトルは同図に示す短軸に平行なベクトルである。画像に対して回転中心をネイルチップ領域の重心、回転角を第1主成分ベクトルの傾きの負方向としてアフィン変換をかけることで、爪の長軸が画像のy軸と一致する図 5-4(b)のような画像に変形できる。最後に、ネイルチップ領域の重心が中心となるような1辺201 [pixel]の正方形領域を設定し、新しい画像として切り取る。図 5-5 が処理後の画像である。切り取った画像を標準化した画像として、仮想ネイルアートシステムの本処理で用いる。



(a) 主成分分析により得た軸



(b) 回転補正後のネイルチップ画像

図 5-4 回転変換後の座標

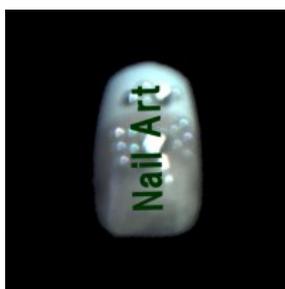


図 5-5 標準化ネイルチップ画像

## 5.4 ネイルチップ画像重畳のためのパラメータ設定手法

本節では、画像平面内を対象として取得するパラメータである、重畳の中心点  $C_i$ 、指の基準長軸長さ  $Width_{Long}^i$ 、指の基準短軸長さ  $Width_{Short}^i$ 、指の傾き角  $\theta_i$ 、および奥行き方向の回転角  $\varphi_i$  を設定する手法について述べる。

初めに、各指の軸を取得する手法を述べる。まず、図 5-6 のように  $i$  番目の指の領域が中心座標上にある 1 辺 81 [pixel] の正方形 ROI を設定する。そして ROI 内に存在する指領域をシレット化させた画像を生成する。次に、図 5-7(a) のように画像中心から半径が 40 [pixel] の円形状に領域を切断し、図 5-7(b) のように、画像の中心点と同じ領域に属さない画素を除去する。これにより、 $i$  番目の指のみが残った画像が生成できる。この処理の後、輪郭線を複数回除去して図 5-7(c) のような指領域が細くなった画像を生成する。この処理は、指領域の座標に対する主成分ベクトルを安定的に取得するために行う。第 1 主成分ベクトルは指の回転軸、第 2 主成分ベクトルはその法線ベクトルと平行となる。よって、これらの軸を指の長軸、短軸として使用する。主成分の導出方法は 5.3 節と同様の方法であり、データ行列は指領域の画像空間における座標である。指の画像空間における傾き  $\theta_i$  は第 1 主成分ベクトルを  $(x_{PCA-1}^i, y_{PCA-1}^i)$  とすると、式(5-2)で算出できる。

$$\theta_i = \tan^{-1} \left( \frac{y_{PCA-1}^i}{x_{PCA-1}^i} \right) - \frac{\pi}{2} \quad (5-2)$$

本論文では、重畳中心  $C_i$  を爪領域重心と指先の中点と定義した。爪領域重心を通り、第 1 主成分ベクトルに平行な直線を引き、指先を取得し、中点を求める。長軸長さ  $Width_{Long}^i$  は爪領域重心と指先の距離と定義した。短軸長さ  $Width_{Short}^i$  は爪領域の重心に対して第 2 主成分ベクトルに平行な直線を引き、その直線に沿って左右の爪領域の長さを求め、その長い方と定義した。

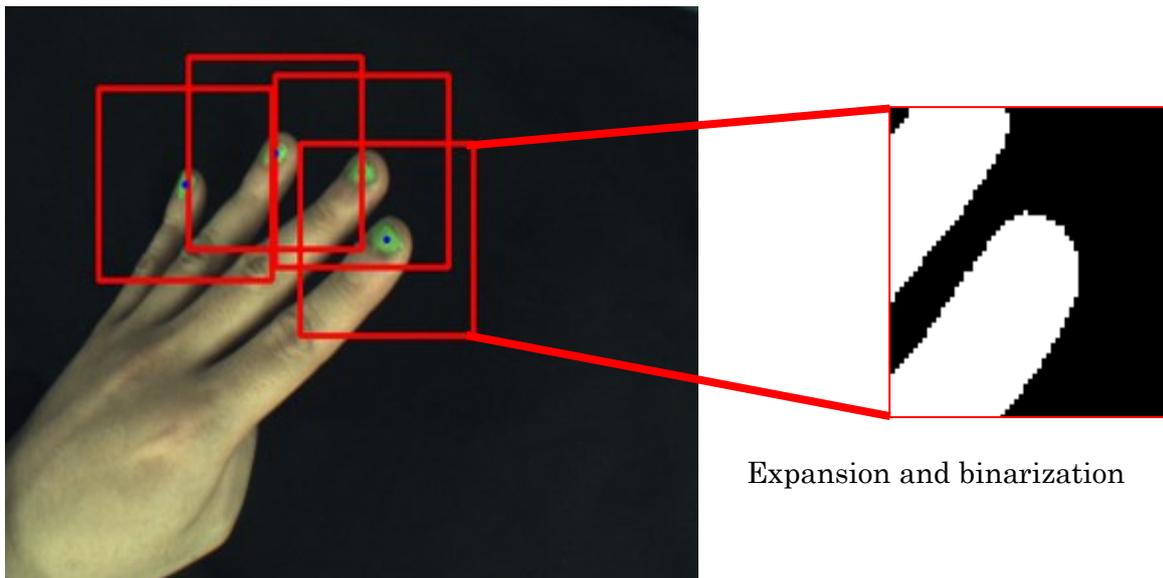


図 5-6 指領域ごとの ROI 設定と ROI 内の 2 値化画像



図 5-7 指の回転軸取得のための処理

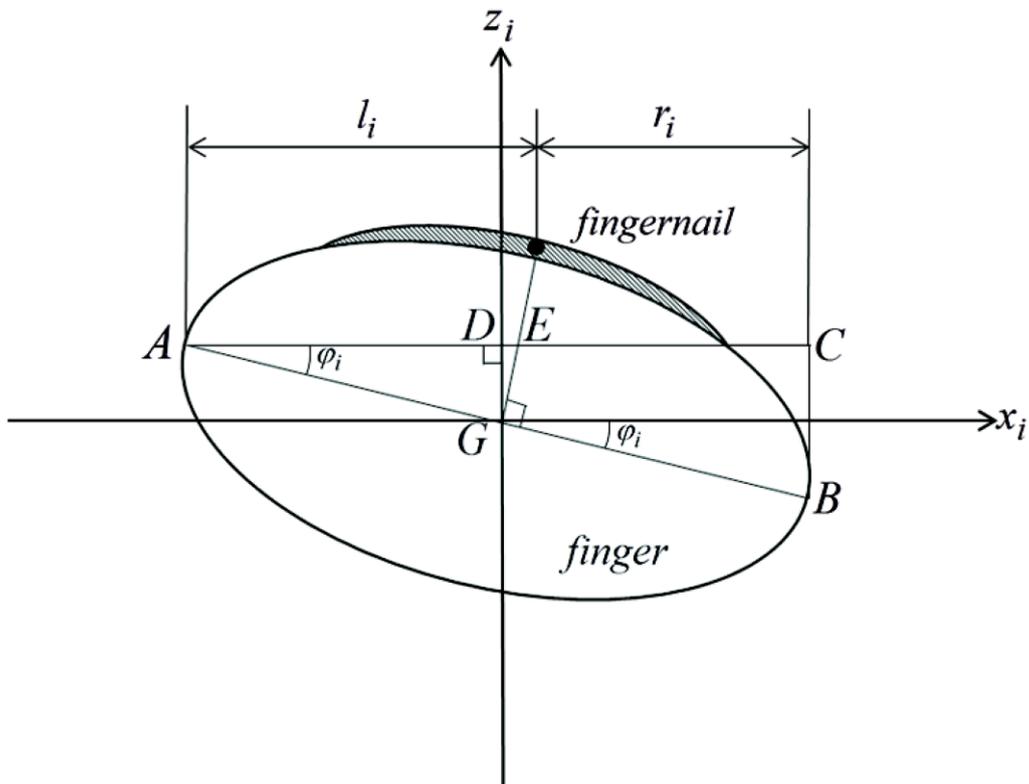


図 5-8 指の回転軸方向から見た指のモデル

奥行き方向の回転角  $\varphi_i$  を図 5-8 に示す指のモデルを基として近似する．画像から得られる値は  $l_i$  と  $r_i$  である．しかし，指の正確な幅が既知でないため， $\varphi_i$  とこれらを結ぶ解析的な式を得ることはできない． $\varphi_i$  は右回旋の場合，式(5-3)のように，未知の長さ  $AB$  と  $l_i$  と  $r_i$  の和である  $AC$  を用いて表せる．よって，できる限り  $AB$  に近い値を  $l_i$  と  $r_i$  を用いて表現する．

$$\varphi = \cos^{-1} \left( \frac{AC}{AB} \right) \quad (5-3)$$

$ADG$  は  $AG$  を斜辺とする直角三角形である．故に  $AG \geq AD$  が成立する．また， $AEG$  は  $AE$  を斜辺とする直角三角形である．ゆえに， $AG \leq AE$  が成立する．ここで， $AE \leq l_i$  より， $AG \leq l_i$  が常に成立する．よって，

$$AD \leq AG \leq l_i \quad (5-4)$$

となる。本論文では AD と  $l_i$  の中間長さを AG の近似長さとして定義した。よって、AB の近似式は式(5-5)となる。

$$AB \approx \frac{3l_i + r_i}{2} \quad (5-5)$$

このとき、指が右回旋ならば式(5-3)と式(5-5)から、式(5-6)の近似式が得られる。

$$\varphi_i \approx \cos^{-1} \left( \frac{2(l_i + r_i)}{3l_i + r_i} \right) \quad (5-6)$$

左回旋についても同様のモデルを立てて計算することで、式(5-7)を導出することができる。

$$\varphi_i \approx -\cos^{-1} \left( \frac{2(l_i + r_i)}{l_i + 3r_i} \right) \quad (5-7)$$

## 5.5 仮想奥行き座標を利用したネイルチップ画像の変形

平面のままのネイルチップに対し、奥行き方向に $\phi_i$ 回転の変換を行うと大きな問題が生じる。図 5-9 はネイルチップの右側面が正面を向いた状態から始め、左回転しながら左側面が正面を向くまで回転した様子を示したものである。図 5-9 に見られるように、奥行き座標を持たない画像を奥行き方向に回転させると、左右均等に圧縮されたような不自然な変換画像が生成される。その原因は、奥行き情報無しでは回転による重心の偏心が生じないことである。本稿では仮想的に奥行き座標 $z$ を与えて、その問題を解決する。概要を図 5-10 に示す。実在するネイルチップの形状は中心が最も $z$ 座標が高く、ネイルチップの端が最も低くなる。そこで、画像の $x$ 軸のみに着目して、中央の $x_{center}=100$ で最大値 $z_{max}$ を、ネイルチップ領域の左端、および右端で最小値 $z_{min}=0$ となるように $z$ 座標を生成する。最も滑らかになるのは円弧を描く三角関数となるため、式(5-8)にて $\phi_i=0$ における画素 $j$ の $z$ 座標 $z_j$ を定義する。

$$z_j = z_{max} \sin\left(\frac{\pi(x_j - 100)}{Width_{chip}} + \frac{\pi}{2}\right) \quad (5-8)$$

次に $y$ 軸回転による奥行き方向への変形を行う。変換式は式(5-9)で表される。

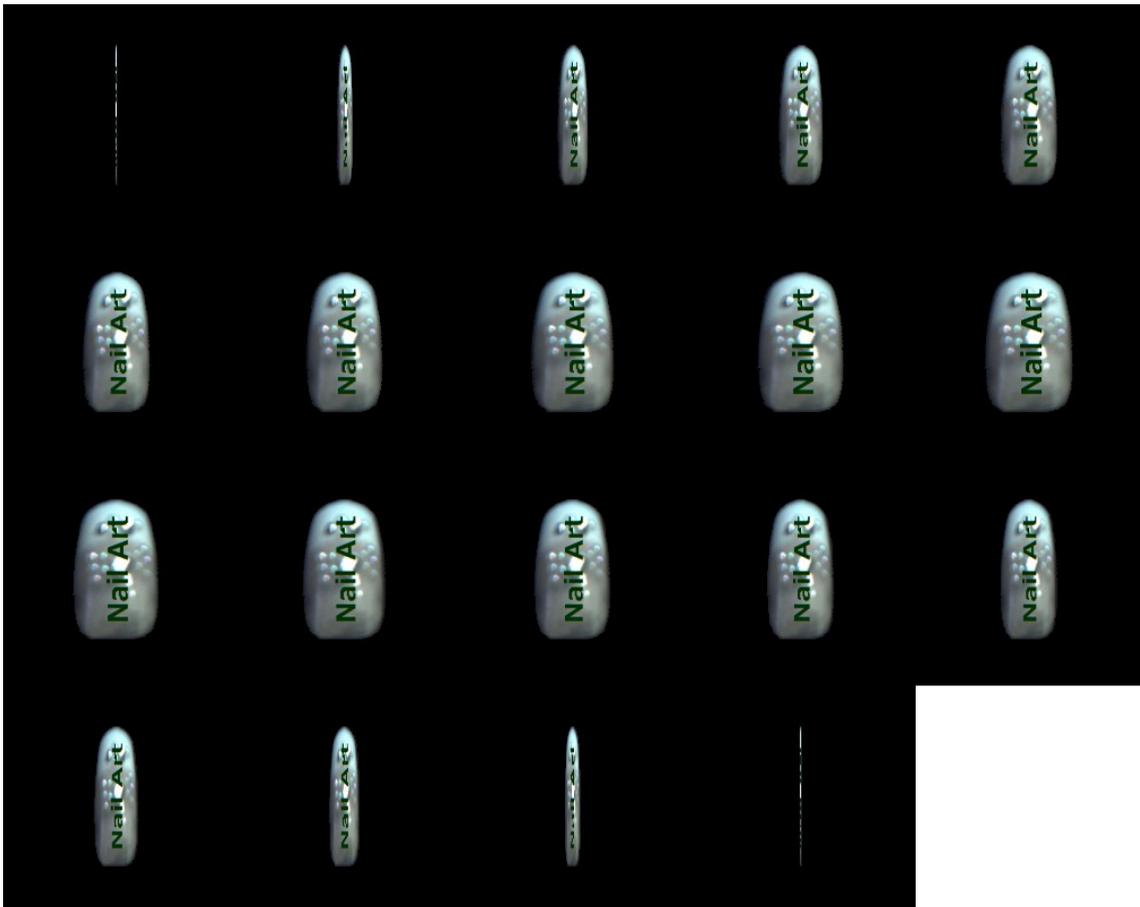


図 5-9 奥行き座標を持たない画像を奥行き方向に回転した時の変化

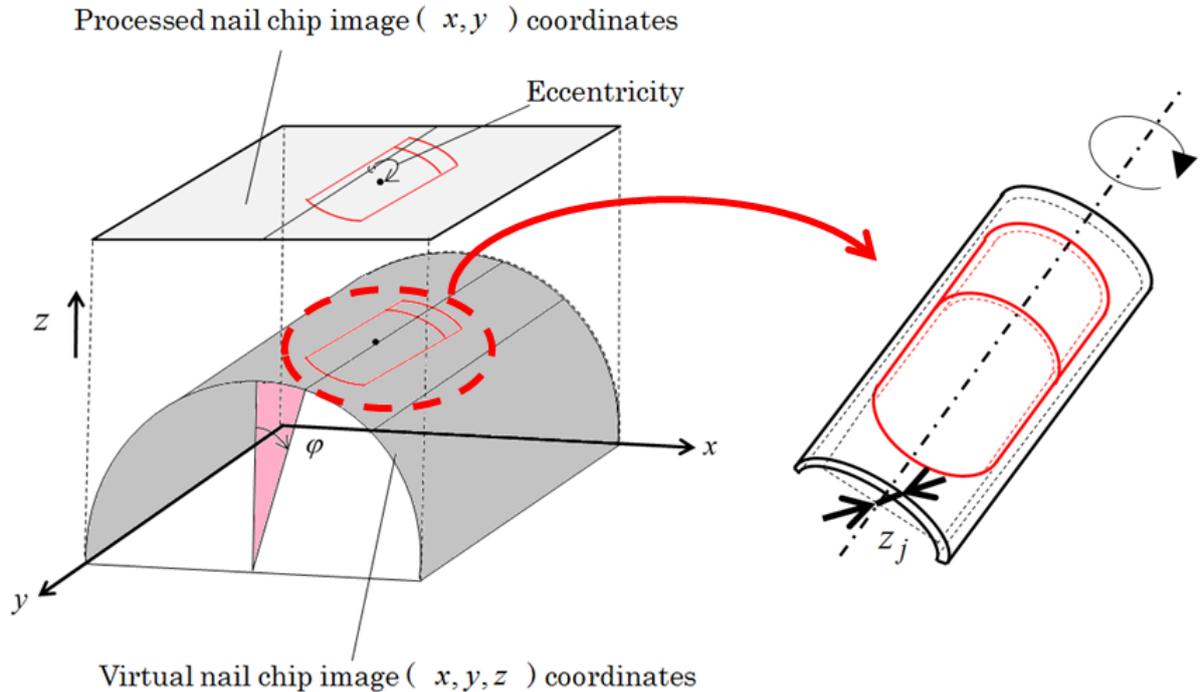


図 5-10 仮想奥行き座標を利用した偏心の実現

$$\begin{pmatrix} x_{1j}^i \\ y_{1j}^i \\ z_{1j}^i \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \varphi_i & 0 & -\sin \varphi_i & 100 \\ 0 & 1 & 0 & 100 \\ \sin \varphi_i & 0 & \cos \varphi_i & z_{\max} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_j - 100 \\ y_j - 100 \\ z_j - z_{\max} \\ 1 \end{pmatrix} \quad (5-10)$$

$x_j$  の値が離散値であるため、そのまま  $\varphi_i$  回転変換を行うと、回転変換後に列座標  $x_{1j}^i$  をとる画素が存在せず、ネイルチップが  $y$  軸に対して並行に割れるという現象が生じる。ゆえに、本稿では式 (5-9) による変換前に  $x_j$  に対してバイリニア法を適用することにより整数変数から実数変数に変換を行い、対応している。

奥行き変換を行った後は  $(x, y)$  座標のみの画像空間に戻す必要がある。この時、複数の画素が同じ  $(x, y)$  座標に写像されることがある。この時は  $z_j$  の値が最も高い画素を選出し、画像空間に写像する。図 5-11 から図 5-14 は最大奥行きを変化させたときの回転変形の様子の違いを示している。これらの図から、本手法を用いるとネイルチップの偏心が実現でき、奥行座標の最大値  $z_{\max}$  の値が大きいほど強く生じることがわかる。ゆえに、本来のネイルチップの曲率に合わせて事前にパラメータ設定を行うことが、自然な重畳を実現するための作業となる。本論文では  $z_{\max}$  の値を 10 とした。

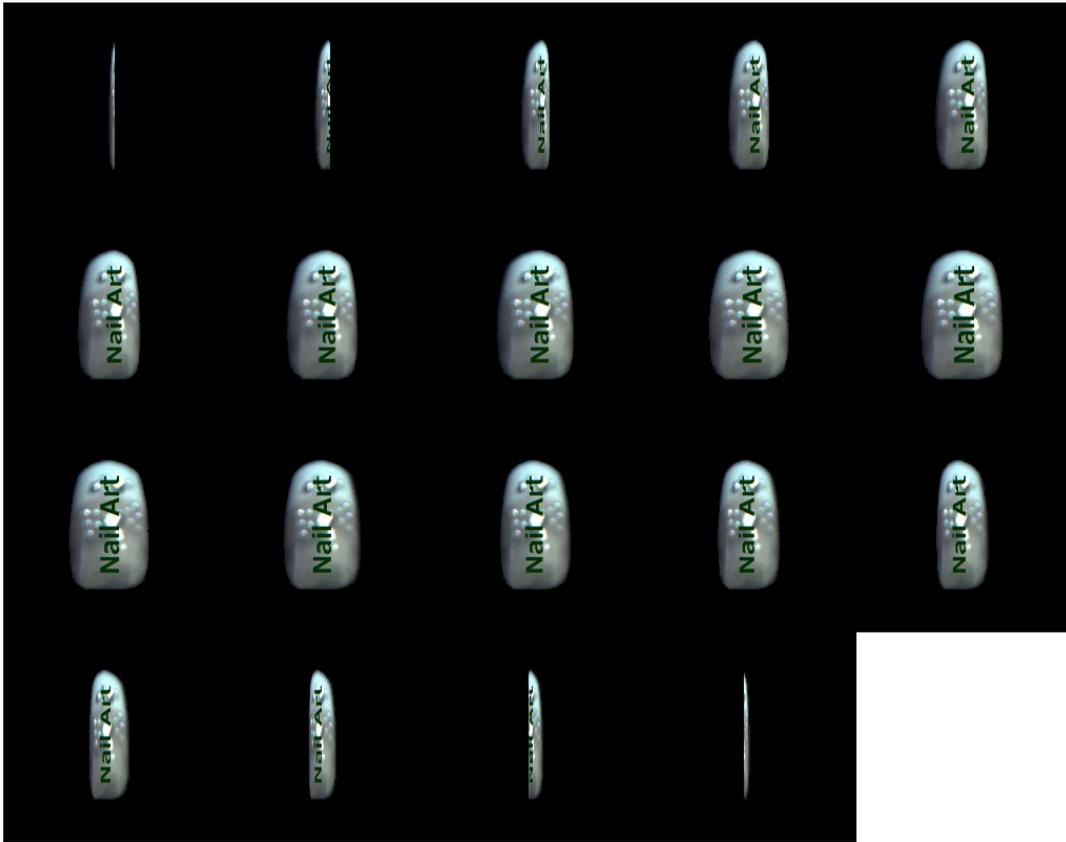


図 5-11  $z_{\max} = 5$  としたときの回転変換の様子



図 5-12  $z_{\max} = 10$  としたときの回転変換の様子



図 5-13  $z_{\max} = 15$  としたときの回転変換の様子

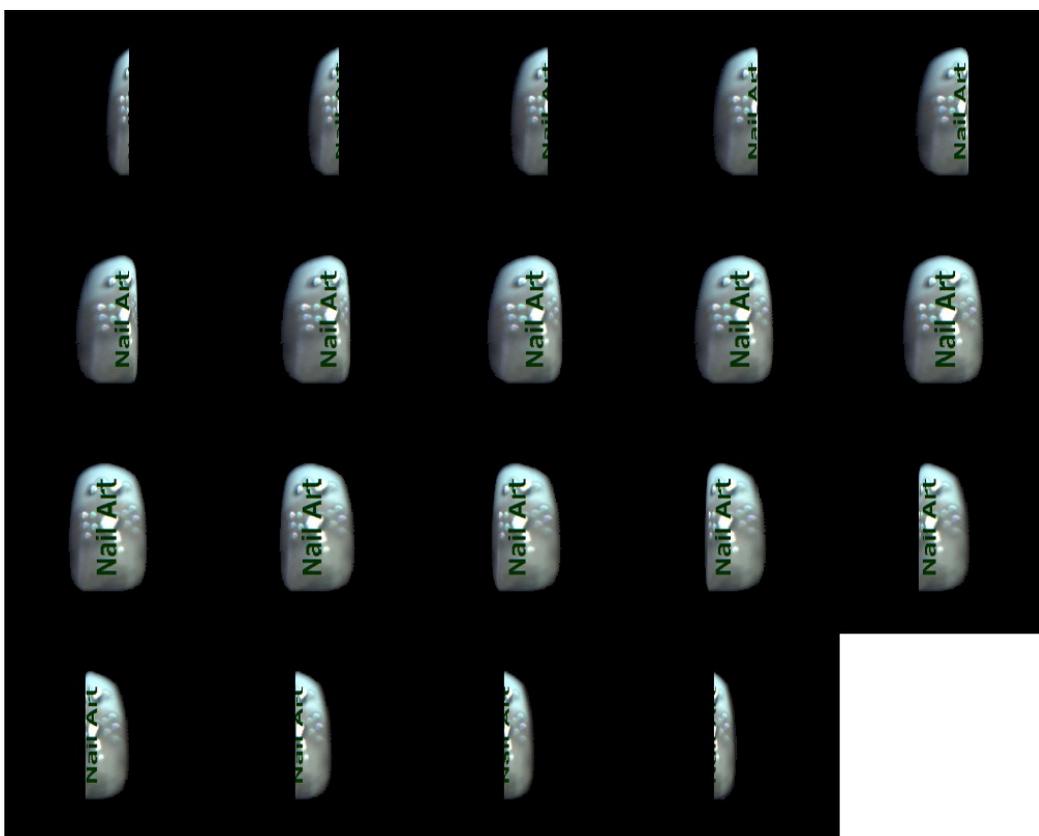


図 5-14  $z_{\max} = 20$  としたときの回転変換の様子

## 5.6 ネイルチップ画像の手画像への重畳

ネイルチップの重畳をするためにパラメータに合わせたネイルチップ画像を生成する。まず、ネイルチップを幅  $Width_{Short}^i$ 、長さ  $Width_{Long}^i$  となるように処理を行う。処理ではネイルチップの比率を壊さずに、長軸の値が  $Width_{Long}^i + \alpha$  になるように縮小処理を行う。この縮小処理で短軸長さが  $Width_{Short}^i$  未満になる場合は、 $Width_{Short}^i$  以上を保てる割合まで縮小を行う。ここで、 $\alpha$  はユーザが主観的に設定する指先からネイルチップを飛び出させたい長さである。その後、ネイルチップが指領域から短軸方向に不自然に飛び出すのを防ぐため、短軸の左右比に合わせてネイルチップの領域を削り、短軸長さを  $Width_{Short}^i$  にする。この作業により、画素は  $x_{1j}^i$  から  $x_{2j}^i$  へ写像される。そして、 $x$  軸正方向から  $\theta_i$  だけ回転させるアフィン変換を、式(5-11)を用いて行うことで、重畳用のネイルチップ画像を生成する。

$$\begin{pmatrix} x_{3j}^i \\ y_{3j}^i \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \theta_i & \sin \theta_i & 100 \\ -\sin \theta_i & \cos \theta_i & 100 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{2j}^i - 100 \\ y_{2j}^i - 100 \\ 1 \end{pmatrix} \quad (5-11)$$

最後に、ネイルチップの重心と  $C_i$  を一致させるように重畳を行う。重畳結果の例を図 5-15 と図 5-16 に示す。

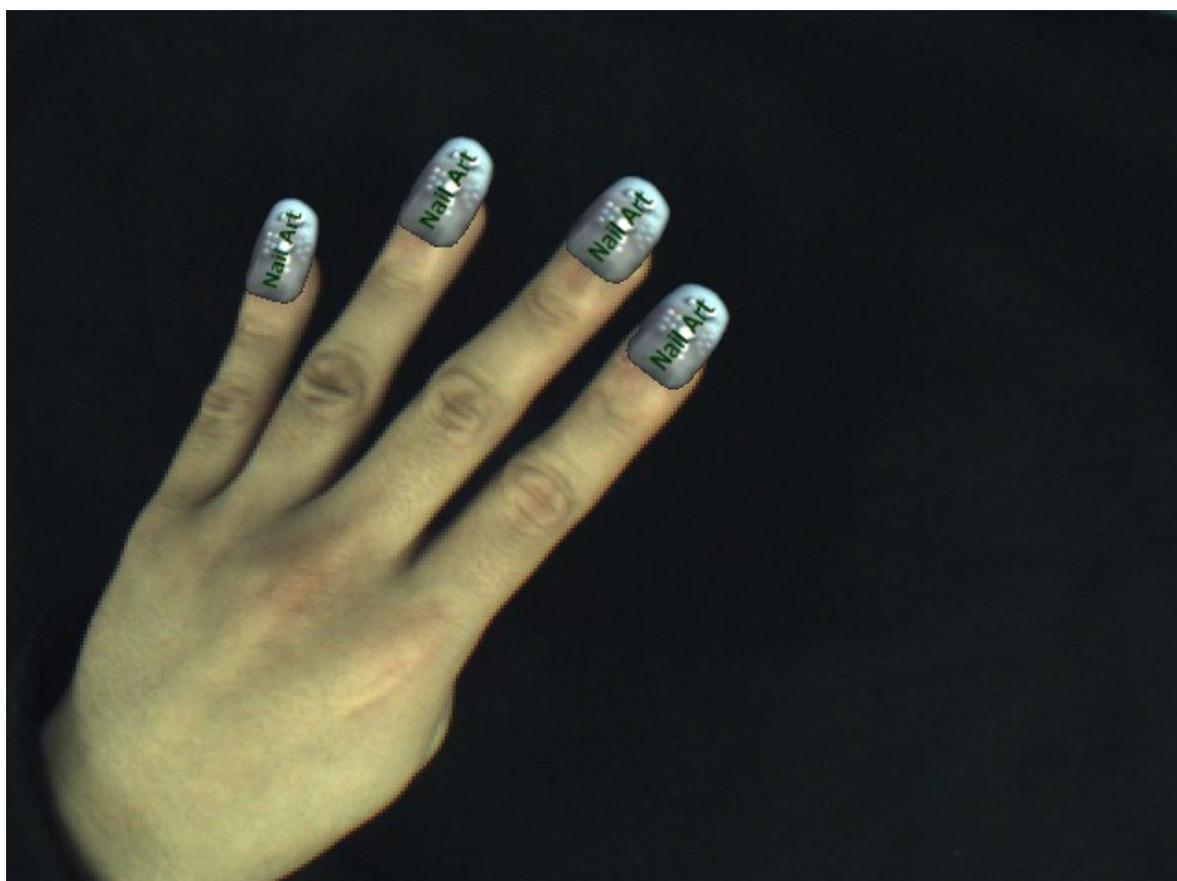


図 5-15 ネイルチップの重畳結果 (右回旋時)



図 5-16 ネイルチップの重畳結果 (左回旋時)

## 5.7 仮想ネイルアートシステムの評価

評価実験では2つの事について調査を行った。1つ目は軸の正確な推定ができたかを表す指先検出の正確さの評価である。これによりどの程度正確に重畳できるかを評価した。2つ目は提案手法を用いる場合と用いない場合の差の評価である。これにより、奥行き情報を用いたときの効果を確認した。評価実験1では、指先端の座標の推定値と実測値の差を算出した。推定値は本稿で構築したネイルアートシステムが算出する指先座標で、実測値は著者が目視により取得した指先座標である。5枚の爪が写る100枚の画像を用いた。前腕は左右に数回往復する回旋を行った。

実験結果は図5-17に示すように、平均に標準偏差を加えても3 [pixel] 程度の検出精度を達成した。これは十分に小さい値であると考えている。5.6節で示した図5-15と図5-16の実際の装着の様子を確認しても、回旋の有無に関わらずネイルチップが理想の方向に重畳できていることが確認できる。つまり、指の軸の推定はシステム内で良く動作していることを表している。

一方で、実験結果の画像を動画として確認すると、ネイルチップが振動的に動いていることも確認された。図5-18は拇指のみを時系列でみたときの指先座標のズレである。この図からも振動は随時こまめに起きていることがわかる。また、振動の偏り方向は前腕回旋角度に依存する傾向が見られている。振動に関しては大きなズレではない。実用で利用する際には、指の先端の座標の移動量が少ない時には移動させないなどの時系列を考慮したプログラミングを利用すべきである。振動の偏りに関しては、爪領域の回旋による欠けなども影響しているため、補正するための手法を今後検討する必要がある。

次に、 $\phi_i$ と $z_j$ が存在するときと、しないときの画像を生成した。結果の一例を図5-19と図5-20に示す。実験結果を比較すると、図5-19では若干ながら“Nail Art”の文字が、指の軸に比べて左または右に存在することがわかる。これは、ネイルチップが指領域から飛び出さないように5.6節の最後に行った短軸比を利用した幅の除去が影響して生じる症状であった。一方、本稿の提案手法を用いると、その補正を行っても、図5-20のように、ネイルチップの文字は指の軸上に乗っていた。つまり、自然な画像を生成することができたといえる。

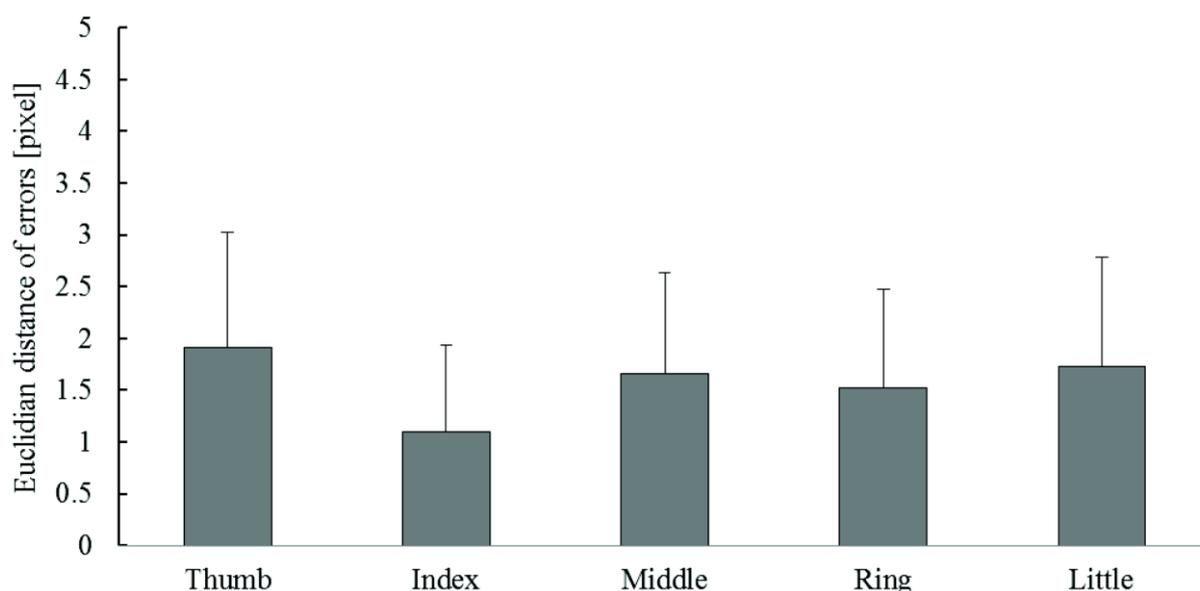


図 5-17 推定値と実測値のユークリッド距離の差

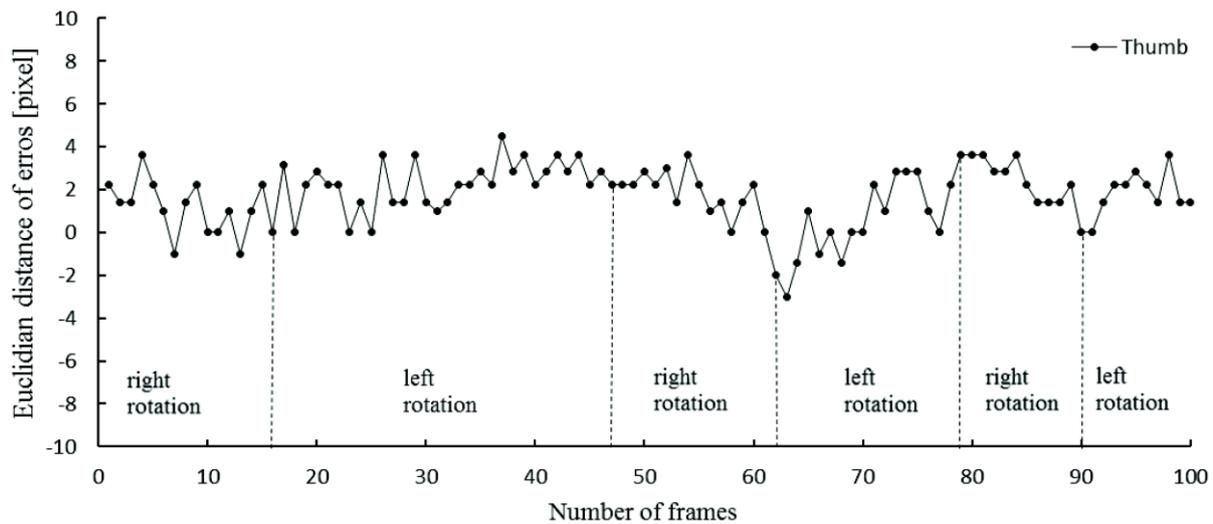


図 5-18 指先座標の振動性と推定ブレの様子



図 5-19  $\phi_i$  と  $z_j$  の値が 0 の時



图 5-20 提案手法 ( $z_{\max} = 10.0$ )

## 5.8 第 5 章のまとめ

本章では爪検出システムの別の応用例として仮想ネイルアートシステムを取り上げ、前腕回旋に対応したシステムにするためのネイルチップ重畳パラメータの設定手法とネイルチップ変形手法を提案した。

ネイルチップ重畳パラメータの設定では、指のモデルを利用して指の奥行き回転角度の推定を実現した。ネイルチップの変形では、仮想的な奥行き座標を付加することにより、1枚の画像でも奥行き回転が自然に表せるようにした。

評価実験では指先検出の正確さと重畳の見えの評価を1名の被験者により評価した。評価実験の結果、指先の推定値と実測値のずれは平均+標準偏差で3 [pixel]程度のずれになることが確認された。また、見えの評価では、提案した奥行き座標の推定と仮想奥行き座標付加によるネイルチップ変形を行わなかった画像と比較し、より自然な見えが実現できることが確認された。

## 第6章 まとめ

RGB 単眼カメラを用いた手指形状推定システム発展のために、爪を指先代わりに利用する手法が考えられるが、従来の爪検出に関わる技術は掌領域が写る手画像に対応できないという大きな問題が存在した。その原因は、爪領域と肌領域は色が異なり、画素ごとの色抽出のみで爪領域が取得できるという前提が掌領域に対して適用できないことであった。つまり、掌領域が写る手画像では爪検出が困難であった。掌領域に対応させるためには新たに、爪らしい色を持つ領域が爪か肌かを判別するアルゴリズムの開発が必要であった。

本論文では、画像空間における色の分布状態が爪周辺とその他で異なる性質に着目した。そして、爪らしい色を持つ領域の重心周りに着目し、「領域の独立性」と「高爪尤度画素の分布密度」の2つ性質を用いて、爪か肌かを判別するアルゴリズムを提案した。

評価実験では日本人大学生5名を被験者とし、前腕回旋角度と爪検出率の関係、および提案した爪判別アルゴリズムの除去性能を評価した。評価実験の結果、本手法を用いたシステムは前腕回旋角度 $-90^{\circ}$  から $-50^{\circ}$  ,  $40^{\circ}$  から $90^{\circ}$  の範囲で、爪検出率85%以上を達成した。つまり、掌側に対応できる、手の裏表に関わらず爪検出可能なシステムを実現した。また、除去性能を評価した結果、爪判別アルゴリズムに使用した2つの性質は共に欠点を補完し合う関係にあることが判明した。すなわち、2つを共に使用することが重要であると確認された。

# 謝辞

本論文を作成するにあたり、日頃より有意義なご指導を与えてくださった、筑波大学システム情報系 星野聖教授に深く感謝の意を表します。また、本論文を査読してくださるとともに、大変貴重なご意見を頂きました筑波大学システム情報系 堀憲之教授、安信誠二教授、星野准一准教授、白川友紀教授に心からお礼申し上げます。星野准一准教授には博士論文執筆時に関しましても、多大なご指導とお時間を頂き、また精神的な支えとなっていただきましたことを、重ねて深く御礼申し上げます。

また、日頃より研究生活を共にし、時に研究の有意義な話を、時に実験の協力を、そして時に日々の生活を楽しくしてくれました医用生体工学研究室の皆様にも深く感謝いたします。

最後にこれまで5年間の長きにわたる大学院生活を精神的にも、環境的にも支えてくださった両親に深く感謝いたします。

## 参考文献

- [1] 泉 賢二, “画像認識装置および操作判定方法並びにプログラム”, WO 2009148064 A1.
- [2] 島根産業技術センター, “島根産業技術センターの開発技術『Gesture-Cam』株式会社インディソフトウェア事業化第一弾、東映太秦映画村が活用開始”, [http:// www3.pref.shimane.jp/houdou/files/DC2458AF-2BA8-4902-9DB1-08C6E6B359EE.pdf](http://www3.pref.shimane.jp/houdou/files/DC2458AF-2BA8-4902-9DB1-08C6E6B359EE.pdf) (2014 年 12 月閲覧).
- [3] 島根県立しまね海洋館, “島根県立産業技術センター、島根県立しまね海洋館 (アクアス) と連携し、プロジェクト開発品「Gesture-Cam」水族館用システムの試験運用を開始”, <http://www.aquas.or.jp/uploads/news/gesture-cam.pdf> (2014 年 12 月閲覧).
- [4] Iason Oikonomidis, Nikolaos Kyriazis and Antonis Argyros, “Efficient model-based 3D tracking of hand articulations using Kinect”, Proc. of the British Machine Vision Conference, pp. “101.1”-“101.11”, September 2011.
- [5] Robert Y. Wang and Jovan Popović, “Real-Time Hand-Tracking with a Color Glove”, Journal of ACM Transactions on Graphics - Proceedings of ACM SIGGRAPH 2009, Vol.28, No.3, Article No.63, August 2009.
- [6] 吉野 和芳, 川島 稔夫, 青木 由直, “色の組み合わせによるジェスチャの直接的推定”, 電子情報通信学会誌 A, Vol.J79-A, No.2, pp.424-431, 1996 年 2 月.
- [7] Thomas Gump, Pedram Azad, Kai Welke, Erhan Oztop, Rüdiger Dillmann and Gordon Cheng, “Unconstrained Real-time Markerless Hand Tracking for Humanoid Interaction”, Proc. of 2006 6th IEEE-RAS International Conference on Humanoid Robots, pp.88-93, December 2006.
- [8] Shan Lu, Dimitaxas Metaxas, and John Oliensis, “Using Multiple Cues for Hand Tracking and Model Refinement”, Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.2, pp.443-450, June 2003.
- [9] Ali Erol, George Bebis, Mircea Nicolescu, Richard D. Boyle and Xander Twombly, “Vision-based hand pose estimation: A review”, Computer Vision and Image Understanding, Vol.108, No.1, pp.52-73, October 2006.
- [10] 谷本 貴頌, 富田 元将, 森 大祐, 以後 直樹, 星野 聖, “多階層データベースを用いた高精度かつ高速の手指形状推定”, 電気学会 C 部門誌, Vol.129-C, No.1 pp.800-805, 2009 年 1 月.

- [11] 富田 元將, 星野 聖, “前腕を含む手指画像からの3次元手指形状推定”, 映像情報メディア学会誌, Vol.63, No.6, pp.822-828, 2009年6月.
- [12] 吉川拓伸, “肌色を科学するー第1回ー”, 日本色彩学会誌, Vol.29, No.1 pp.31-34, 2005年3月.
- [13] G. B. Buck and H. C. Froelich, “Color Characteristics of Human Complexions”, *Illumination Engineering*, Vol.43, pp.27-49, January 1948.
- [14] 吉川拓伸, 棟方 明博, 高田 定樹, 矢口 博久, “1990年代における日本人女性の肌色変化”, 日本色彩学会誌, Vol.34, No.2, pp.120-130, 2010年6月.
- [15] 鳥居 さくら, 多田 明弘, “顔画像の空間周波数特性による年齢の推定と男女の判別”, 電子情報通信学会論文誌 A, Vol.J97-A, No.10, pp.657-664, 2014年10月.
- [16] 中井 久史, 眞鍋 佳嗣, 井口 征士, “肌のスペクトル分布解析と肌らしさの生成”, 電子情報通信学会論文誌 D-II, Vol. J81-D-II, Vol.9, pp.2139-2145, 1998年9月.
- [17] Norimichi Tsumura, Hideaki Haneishi and Yoichi Miyake, “Independent Component Analysis of Skin Color Image”, *Journal of the Optical Society of America A*, Vol.16, No.9, pp.177-180, September 1999.
- [18] 津村 徳道, 羽石 秀昭, 三宅 洋一, “光散乱を考慮した肌の分光反射率の非線形独立成分分析”, *光学*, Vol.30, No.6, pp.400-406, 2001年6月.
- [19] Gouri V. Gupchup and Joel L. Zatz, “Structural characteristics and permeability properties of the human nail: A review”, *Journal of the Society of Cosmetic Chemists*, Vol.50, No.6, pp.363-385, November 1999.
- [20] 清水 宏, “新しい皮膚科学(第2版)”, 中山書店, pp.350-351, 2011年4月.
- [21] Stephen A. Mascaró and H. Harry Asada, “Measurement of Finger Posture and Three-Axis Fingertip Touch Force Using Fingernail Sensors”, *IEEE Transactions on Robotics and Automation*, Vol.20, No.1, pp.26-35, February 2004.
- [22] Ravi Ramamoorthi, Melissa Koudelka and Peter Belhumeur, “A Fourier Theory for Cast Shadows”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.27, No.2, pp.288-295, February 2005.

- [23] B.S. Manjunath and W. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.18, No.8, pp.837-842, August 1996.
- [24] Dengsheng Zhang, Aylwin Wong, Maria Indrawan and Guojun Lu, "Content-based Image Retrieval Using Gabor Texture Features", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.22, No.8, pp.13-15, August 2000.
- [25] X. Fu, Y. Li, R. Harrison and S. Belkasim, "Content-based Image Retrieval Using Gabor-Zernike Features", Proc. of 18th International Conference on Pattern Recognition, Vol.2, pp.417-420, August 2006.
- [26] Nobuyuki Otsu and Takio Kurita, "A new scheme for practical flexible and intelligent vision systems", Proc. of IAPR Workshop on Computer Vision, pp.431-435, October 1988.
- [27] Navneet Dalal and Bill Triggs, "Histograms of Oriented Gradients for Human Detection", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Vol.1, pp.886-893, June 2005.
- [28] Kobi Levi and Yair Wess, "Learning Object Detection from a Small Number of Examples: the Importance of Good Features", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Vol.2, pp.53-60, June 2004.
- [29] Bo Wu and Ram Nevatia, "Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors", Proc. of the Tenth IEEE International Conference on Computer Vision, Vol.1, pp.90-97, October 2005.
- [30] Stefan Walk, Nikodem Mayer, Konrad Schindler and Bernt Schile, "New Features and Insights for Pedestrian Detection", Proc. of 2000 IEEE Conference on Computer Vision and Pattern Recognition, pp.1030-1037, June 2010.
- [31] 後藤 雄飛, 山内 悠嗣, 藤吉 弘亘, "CS-HOG : 色の類似性に基づいた形状特徴量", 電子情報通信学会論文誌 D, Vol.J96-D, No.7, pp.1618-1626, 2013 年 7 月.
- [32] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski and Cédric Bray, "Visual Categorization with Bags of Keypoints", Proc. of Workshop on Statistical Learning in Computer Vision, 8th European Conference on Computer Vision, pp.1-22, May 2004.
- [33] Nobuyuki Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE Transactions on Systems, Man, and Cybernetics, Vol.SMC-0, No.1, pp.12-66, January 1979.

- [34] 喜多秦代, “二次元濃度ヒストグラムを用いた画像間変化抽出”, 電子情報通信学会論文誌 D, Vol.J90-D, No.8, pp.1957-1965, 2007 年 8 月.
- [35] 川端 聡, 日浦 慎作, 佐藤 宏介, “カーネル PCA による未知物体領域の高速抽出法”, 電子情報通信学会論文誌 D, Vol.J91-D, No.11, pp.2673-2683, 2008 年 11 月.
- [36] 島井 博行, 栗田 多喜夫, 梅山 伸二, 田中 勝, “ロバスト統計に基づいた適応的な背景推定法”, 電子情報通信学会論文誌 D-II, Vol.J86-D-II, No.6, pp.796-806, 2003 年 6 月.
- [37] Dorin Comaniciu, ”Mean Shift: A Robust Approach Toward Feature Space Analysis”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No.5, pp.603-619, May 2002.
- [38] Tapas Kanungo and Nathan S. Netanyahu, “An Efficient k-Means Clustering Algorithm: Analysis and Implementation”, IEEE Transactions on Pattern analysis and Machine Intelligence, Vol.24, No.7, pp.881-891, July 2002.
- [39] Maoguo Gong, Yan Liang, Jiao Shi, Wenping Ma and Jingjing Ma, “ Fuzzy C-Means Clustering With Local Information and Kernel Metric for Image Segmentation”, IEEE Transactions on Image Processing, Vol.22, No.2, pp.573-584, February 2013.
- [40] Jakob Verbeek and Bill Triggs, “Region Classification with Markov Field Aspect Models”, Proc. of Computer Vision and Pattern Recognition, pp.1-8, June 2007.
- [41] Fan Chen, Kazuyuki Tanaka and Tsuyoshi Horiguchi, “Image Segmentation Based on Bathe Approximation for Gaussian Mixture Model, Interdisciplinary Information Sciences, Vol.11, No.1, pp.17-29, April 2005.
- [42] Pablo Arbeláez and Charless Fowlkes, “Contour Detection and Hierarchical Image Segmentation”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.33, No.5, August 2010.
- [43] Yuri Y. Boykov and Marie-Pierre Jolly, “ Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images”, Proc. of Internation Conference on Computer Vision, vol. I , pp.105-112, July 2001.
- [44] Paul E. Debevec and Jitendra Malik, “Recovering High Dynamic Range Radiance Maps from Photographs”, Proc. of ACM SIGGRAPH 2008 classes, Article No.31 (10 pages), August 2008.

- [45] 竹部 啓輔, 伊藤 直己, 中内 茂樹, 臼井 支朗, “色恒常性を有するデジタルカラーイメージング”, 電子情報通信学会誌 D-II, Vol.J83-D-II, No.8, pp.1753-1762, 2000年8月.
- [46] 森 克己, 渡邊 栄治, 渡辺 浩司, 片桐 重和, “直方体を用いた物体色と光源色・光源方向の推定”, 電子情報通信学会誌, D-II, Vol.J87-D-II, No.10, pp.1973-1982, 2004年10月.
- [47] G. D. Finlayson, S. D. Hordley, C. Lu and M. S. Drew, “On the Removal of Shadows from Images”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.28, No.1, pp.59-68, January 2005.
- [48] Edwin H. Land and John J. McCann, “Lightness and Retinex Theory”, Journal of the Optical Society of America, Vol.61, No.1, pp.1-11, January 1971.
- [49] Zia-ur Rahman, Daniel J. Jobson and Glenn A. Woodell, “Multiscale retinex for color rendition and dynamic range compression”, Proc. of SPIE, Vol.2847, pp.183-191, November 1996.
- [50] Ron Kimmel, Michael Elad, Doron Shaked, Renato Keshet and Irwin Sobel, “A Variational Framework for Retinex”, International Journal of Computer Vision, Vol.52, No.1, pp.7-23, April 2003.
- [51] P. Viola and M. Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Vol.1, pp.511-518, December 2001.
- [52] 道満 恵介, 出口 大輔, 高橋 友和, 目加田 慶人, 井出 一郎, 村瀬 洋, “色変動を考慮した生成型学習法による道路標識検出器の構築”, 電子情報通信学会論文誌 D, Vol.J93-D, No.8, pp.1375-1385, 2010年8月.
- [53] L. Sirovich and M. Kirby, “Low-dimensional procedure for the characterization of human faces”, Journal of the Optical Society of America A, Vol.4, No.3, pp.519-524, March 1987.
- [54] 重成 一真, 坂上 文彦, 尺長 健, “固有顔の直交分解と仮想化による照明変動に影響されない顔画像認識”, 電子情報通信学会論文誌 D-II, Vol.J86-D-II, No.7, pp.996-1004, 2003年7月.
- [55] Hiroshi Murase and Shree K. Nayar, “Visual Learning and Recognition of 3-D Objects from Appearance”, International Journal of Computer Vision, Vol.14, No.1, pp.5-24, January 1995.

- [56] 天野 敏之, 玉木 徹, “EbC 画像対を用いた三次元物体の高速線形姿勢推定”, 電子情報通信学会論文誌 D, Vol.J90-D, No.8, pp.2060-2069, 2007 年 8 月.
- [57] 天野 敏之, 佐藤 幸男, “固有空間法を用いた BPLP による画像補完”, 電子情報通信学会論文誌 D-II, Vol.J85-D-II, No.3, pp.457-465. 2002 年 3 月.
- [58] 安藤 慎吾, 草地 良規, 鈴木 章, 荒川 賢一, “サポートベクトル回帰を用いた三次元物体の姿勢推定法”, 電子情報通信学会論文誌 D, Vol.J89-D, No.8, pp.1840-1847, 2006 年 8 月.
- [59] David G. Lowe, “Three-Dimensional Object Recognition from Single Two-Dimensional Images”, *Artificial Intelligence*, Vol.31, No.3, pp.355-395, March 1987.
- [60] 土肥 浩, 石塚 満, “ソリッドモデラと幾何学的推論を組み込んだ 3D ビジョンシステム”, 電子情報通信学会論文誌 D-II, Vol.J73-D-II, No.10, pp.1677-1686, 1990 年 10 月.
- [61] 岡部 孝弘, 佐藤いまり, 于 瓊, 佐藤 洋一, “照明変化ともなう輝度変化の類似度に基づく物体形状復元”, 情報処理学会論文誌 コンピュータビジョンとイメージメディア, Vol.1, No.3, pp.1-11, 2008 年 11 月.
- [62] 玉城 絵美, 味八木 崇, 歴本 純一, “インタラクションシステムのための高精度な 3 次元ハンドジェスチャ認識手法”, 情報処理学会論文誌, Vol.51, No.2, pp.229-239, 2010 年 2 月.
- [63] 佐野 亮介, 富田 元将, 星野 聖, “爪の相対的な位置情報をふるい分けに用いた 3 次元手指形状推定”, 映像情報メディア学会技術報告, Vol.33, No.17, pp.21-24, 2009 年 3 月.
- [64] 丸谷 誠慶, 西川 敦, 宮崎 文夫, “2 次元画像からの指先位置検出に基づく実時間手形状認識と移動ロボット遠隔操縦タスクへの応用”, 情報処理学会研究報告(CVIM), Vol. 2002, No.34, pp.145-152, 2002 年 5 月.
- [65] 津田 藍花, 加藤 晴久, 米山 暁夫, 半谷 精一郎, “ライトによる色相変化量を利用したネールアートシミュレーション”, 電子情報通信学会総合大会 基礎・境界講演論文集(2009), p.290, 2009 年 3 月.
- [66] 津田 藍花, 上野 智史, 米山 暁夫, 半谷 精一郎, “ライトによる輝度分布特性を用いたネールアートシミュレーション”, 映像情報メディア学会年次大会講演予稿集(2009), pp. “14-4-1” - “14-4-2”, 2009 年 8 月.

- [67] 津田 藍花, 上野 智史, 酒澤 茂之, 半谷 精一郎, “照明による輝度変化の影響を考慮した携帯端末用ネイルアートシミュレーションに関する検討”, 電子情報通信学会技術研究報告, パターン認識・メディア理解(PRMU), Vol.110, No.381, pp.59-64, 2011 年 1 月.
- [68] 古味 重幸, 藤嶋 教彰, 星野 聖, “手指画像中における肌と爪の分離”, 映像情報メディア学会技術報告, Vol.24, No.41, pp.21-23, 2010 年 10 月.
- [69] 星野 聖, 古味 重幸, 藤嶋 教彰, “手画像からの爪領域の高精度抽出”, 電子情報通信学会技術研究報告, ヒューマン情報処理(HIP), Vol.110, No.422, pp.87-90, 2011 年 2 月.
- [70] Yaming Wang, Jiansheng Chen and Guangda Su, “Efficient Hand Segmentation and Fingertip Detection Using Color Features of Skin and Fingernail”, IEICE Transactions on Information and System, Vol.E96-D, No.8, pp.1894-1897, August 2013.
- [71] Yi-Ting Chou and Peter Bajcsy, “Toward Face Detection, Pose Estimation and Human Recognition from Hyperspectral Imagery”, National Center for Supercomputing Applications, Technical report TR.20041108-1, November 2004.
- [72] 平沢 拓己, 戸村 豊明, “手指画像における輪郭と爪を用いた手の形状推定”, 2014 年度精密工学会北海道支部学術講演会講演論文集, B-P24, pp.47-48, 2014 年 8 月.
- [73] 三浦 航平, 張 英夏, 向井 信彦, “爪と手首の位置検出に基づく日本語手話の指文字認識”, 映像情報メディア学会技術報告, Vol.37, No.17, pp.199-202, 2013 年 3 月.
- [74] Takayuki Fujiwara, Kazuhiko Yamaashi and Hiroyasu Koshimizu, “A Method of New Filter Design Based on the Co-occurrence Histogram”, Journal of Electrical Engineering in Japan, Vol.166, No.1, pp.36-42, January 2009.
- [75] 尹 五仙, “美容サービス産業の現状と課題”, 立命館経営学, Vol.47, No.6, pp.155-173, 2009 年 3 月.

## 付録 A 一般表色系

ここでは、国際照明委員会(CIE)にて定義された表色系の中で、本論文で扱ったものについて述べる。また、RGB 表色系から表色系変換を行う時の変換式を示す。まず初めに、RGB 表色系について述べる。次に、RGB 表色系を線形変換して得られる XYZ 表色系、YIQ 表色系、YUV 表色系について述べる。そして、RGB 表色系を非線形変換して得られる HSV 表色系、HLS 表色系、L\*a\*b\*表色系、L\*u\*v\*表色系について述べる。なお、本付録中で使用される図は色度図作成用のフリーソフト、Color AC を用いて作成されている。

RGB 表色系は色を赤(700 [nm])、緑(546.1 [nm])、青(435.8 [nm])を原色として色を表現する表色系である。この3原色を均等な強さで重ねると白色になる。それぞれの色の定義は人が2度視野角にて試料の色と、3色の光を一か所にあてて加法した時の色が同じか判定する等色実験により定義した。この表色系は人間の視覚の錐体と感受する波長が近い。また、カメラの標準的な表色系となっている。各原色を256階調(8 [bit])、合計1677万7216色で表現すれば人間が自然な色合いと感ずることから、24 [bit]表現による色表現をフルカラーという。OpenCVなどの画像処理のオープンAPIでもフルカラーでの扱いが標準である。ディスプレイで表現できる範囲では、RGB表色系で大きな問題は生じない。ただし、現実世界ではRGBの混ぜ合わせで表現できない、つまりRGB表色系で表現できない色も存在する。これらの色は試料側にも光を当てて色を合わせ、その後あてた分の光を引き算するため、RGBのいずれかの値が負値となる。

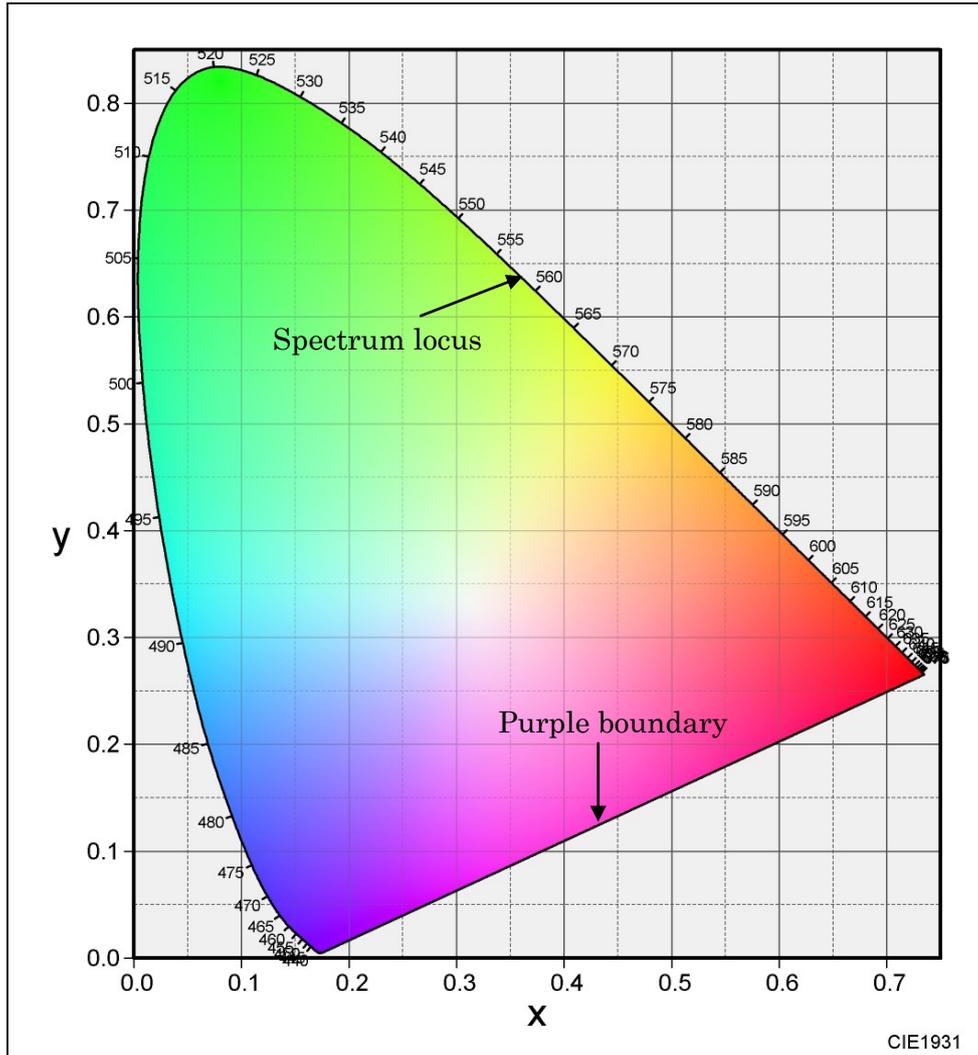
XYZ表色系は3つの刺激値XYZで表現する表色系であり、実在するすべての色が正の値を取る。ただし、色として認知できない範囲も含まれており、現実の色と対応しない仮想色を含んでいる。Xは明度のない赤み、Yは明度のある緑み、Zは明度のない青みである。式(A-1)が変換式となる。

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (\text{A-1})$$

ここで、X、YをXYZの各値の和で除したものを小文字で表す。それをグラフで表現すると付録図1のような図ができる。この図を色度図といい、色彩工学で頻繁に用いられる。色度図の紫から青、緑を経由し、赤へ向かう概形輪郭線は波長のスペクトル軌跡を表している。青と青を結ぶ線は仮想的なスペクトル軌跡であり、純紫軌跡と呼ばれる。

YIQ、YUV表色系は輝度と色差で表現する表色系であり、放送関係にゆかりがある。モノクロテレビからカラーテレビに移り変わる際、放送信号がそのどちらにも対応できる必要があったため、信号のうち一つは色情報を持たない輝度信号となっている。人間の視覚系は輝度と比較して色の近くが鈍い。色差信号は情報を間引き、狭い帯域で送信することで送信帯域を有効に使用できる。

YIQ表色系におけるIはオレンジーライトブルー間の色差、Qは青紫ー黄緑間の色差である。目の感度はQ軸に関して弱いため、Q軸の情報圧縮を行っても質の良い映像が実現できる。RGB表色系からの変換式は式(A-2)となる。

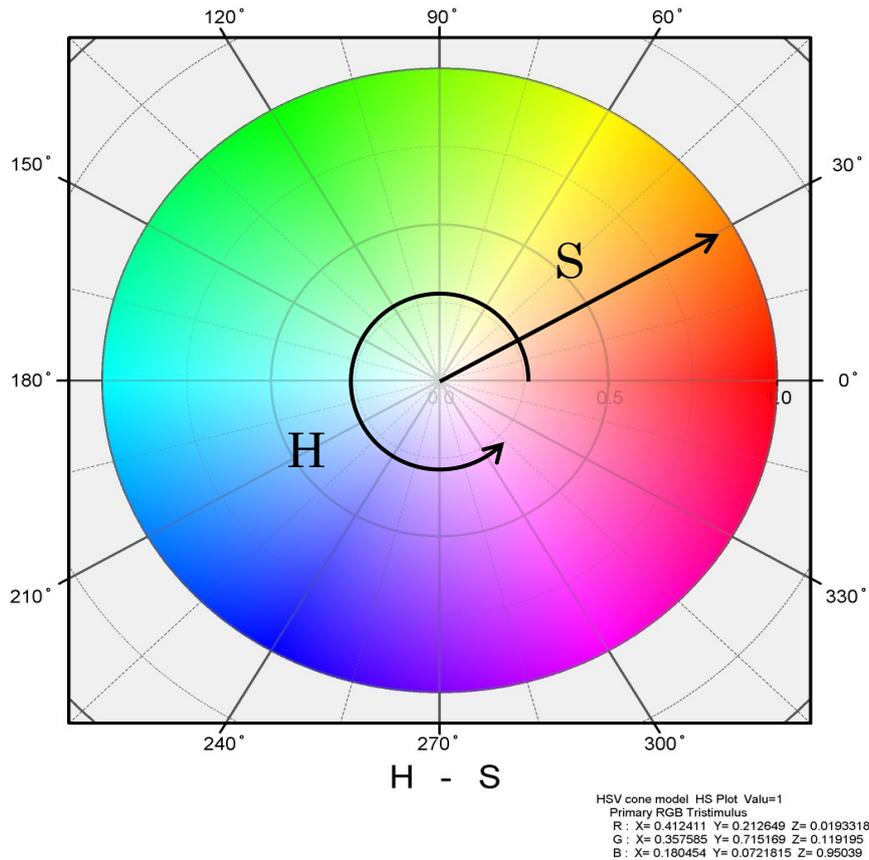


付録図 1  $xy$  色度図

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.522 & -0.311 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (\text{A-2})$$

YUV 表色系における U は寒色系の色差，V は暖色系の色差である．IQ の軸を  $33^\circ$  回転させると UV 軸になる．U 軸の情報は情報圧縮を大きくできる．この表色系は JPEG や MPEG の内部データで使用されている．また，一部のカメラではデータ送信の高速化を目的として YUV フォーマット指定をできるものが存在する．RGB 表色系からの変換式は式(A-3)となる．

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (\text{A-3})$$



付録図 2 色相・彩度平面

HSV, HLS 表色系は色相, 彩度, 明度という 3 つの尺度を軸にして色を表現する. 色相は色の本質的な違いを表現するものである. 彩度は色の鮮やかさの違いを表現するものである. この表色系は人間が色を認識する感覚に近く, 直観的に扱うことができる. 画像処理の分野では同じ色合いの領域を抽出するためにこの表色系を利用することが多い. これらの表色系は柱モデルと錐モデルがあるが, 本論文では推モデルを用いた. 色相・彩度は付録図 2 のように色相は回転方向, 彩度は半径方向を考える極座標方式となっている.

HSV 表色系は色相  $H$ , 彩度  $S$ , 明度  $V$  で構成される. 彩度の基準は明度の値が最大値の所で設定されている. 変換式は式(A-4)から式(A-7)となる. ここで, 式(A-7)は肌と爪の色相が CIE の定義軸では変わり目をまたぐ, という問題を考慮し, 画素値の連続性を確保するために色相  $H$  の変域を定義し直す式である. それぞれの変域は  $-180 \leq H < 180$ ,  $0 \leq S \leq 1$ ,  $0 \leq V \leq 1$  となる.

$$V = \max(R, G, B) \quad (\text{A-4})$$

$$S = \begin{cases} \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} & \text{if } \max(R, G, B) \neq 0 \\ 0 & \text{else} \end{cases} \quad (\text{A-5})$$

$$H' = \begin{cases} \frac{60(G-B)}{S} & \text{if } V=R \\ 120 + \frac{60(B-R)}{S} & \text{if } V=G \\ 240 + \frac{60(R-G)}{S} & \text{if } V=B \end{cases} \quad (\text{A-6})$$

$$H = -180 + H' \quad (\text{A-7})$$

HLS 表色系は色相 H , 明度 L , 彩度 S で構成される。彩度は明度の値が 0 と最大値の間である  $L=0.5$  の所で設定されている。つまり, この表色系は双錐型で定義されている。明度は式(A-8), 彩度は式(A-9)で変換できる。色相は彩度の値を変更して式(A-6), 式(A-7)を用いることで変換できる。

$$L = \frac{\max(R, G, B) + \min(R, G, B)}{2} \quad (\text{A-8})$$

$$S = \begin{cases} \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B) + \min(R, G, B)} & \text{if } L < 0.5 \\ \frac{\max(R, G, B) - \min(R, G, B)}{2 - (\max(R, G, B) + \min(R, G, B))} & \text{else} \end{cases} \quad (\text{A-9})$$

L\*a\*b\*表色系と L\*u\*v\*表色系は色の物理的な差異より人間の知覚の際に着目し, 人間の知覚対して均等になるように軸を設定した均等色空間を持つ表色系である。L\*は明度, そのほかを色度, またはクロマティックネス指数という。

L\*a\*b\*表色系は赤み, 緑み, 青み, 黄みの心理的な 4 原色に着目して作成された。光のスペクトルは考慮されていない。a\*は負側が緑, 正側が赤になる軸である。b\*は負側が青, 正側が黄になる軸である。軸同士の交点は無彩色になる。L\*a\*b\*表色系には均等色空間であることの他に大きな特徴がある。それは色が絶対座標で表せることである。これまでの RGB 表色系などはプリンタやディスプレイ等の出せる色合いの最大値が異なるために, 同じ色でも, その色を指し示す値がデバイス間で異なる。ゆえに, RGB 値などで色を指定して印刷すると, 実際に印刷した色がディスプレイで見ていた色と異なるという問題が生じる。しかし a\*, b\*はデバイスに依存せず, 同じ色ならば同じ値で定義できる。つまり, ディスプレイ上の色と同じ色で印刷できる。ゆえに, コンピュータ内部では, デバイス間で色情報をやり取りする際に, 一度 L\*a\*b\*表色系に変換して情報がやり取りされている。これをカラーマネジメントという。L\*a\*b\*表色系の各値は XYZ 表色系をもとに作成する。まず, RGB 表色系から式(A-1)を用いて XYZ 表色系に変換する。その後式(A-10)から, 式(A-15)を用いて L\*a\*b\*表色系へ変換する。

$$X' = X/0.950456 \quad (\text{A-10})$$

$$Z' = Z/1.088754 \quad (\text{A-11})$$

$$L = \begin{cases} 116Y^{1/3} - 16 & \text{if } Y > 0.008856 \\ 903.3Y & \text{if } Y \leq 0.008856 \end{cases} \quad (\text{A-12})$$

$$a^* = 500(f(X') - f(Y)) \quad (\text{A-13})$$

$$b^* = 200(f(Y) - f(Y - Z')) \quad (\text{A-14})$$

$$f(t) = \begin{cases} t^{1/3} & \text{if } t > 0.008856 \\ 7.787t + 16/116 & \text{else} \end{cases} \quad (\text{A-15})$$

変域は  $0 \leq L^* \leq 100$ ,  $-127 \leq a^* \leq 127$ ,  $-127 \leq b^* \leq 127$  である.

$L^*u^*v^*$ 表色系は光のスペクトルが均一になるように考慮されて作成された. 現在は  $xy$ 色度図の代わりに  $u^*v^*$ 図を利用することがある. 現在, カラーテレビや写真業界で使用されている.  $XYZ$ 表色系から  $L^*$ の値は式(A-12),  $u^*$ ,  $v^*$ の値は式(A-16)から式(A-19)を用いて変換できる.

$$u' = \frac{4X}{X + 15Y + 3Z} \quad (\text{A-16})$$

$$v' = \frac{9Y}{X + 15Y + 3Z} \quad (\text{A-17})$$

$$u^* = 13L(u' - 0.19793843) \quad (\text{A-18})$$

$$v^* = 13L(v' - 0.4831096) \quad (\text{A-19})$$

変域は  $0 \leq L^* \leq 100$ ,  $-134 \leq u^* \leq 220$ ,  $-140 \leq v^* \leq 122$  である.

# 著者論文リスト

## 【査読付き学術論文】

- [1] Noriaki Fujishima and Kiyoshi Hoshino, "Fingernail Detection System Using Differences of the Distribution of the Nail-Color Pixels", Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII), Vol.17, No.5, pp.739-745, September 2013.

## 【査読付き国際会議論文】

- [1] Noriaki Fujishima and Kiyoshi Hoshino, "Fingernail Area Extraction Using Distribution of Colors", Proc. of the 3rd IEEE Global Conference of Consumer Electronics (GCCE2014), Makuhari Messe (Chiba, Japan), pp.762-765, October 2014.
- [2] Noriaki Fujishima and Kiyoshi Hoshino, "Virtual Nail Art System", Proc. of the 3rd IEEE Global Conference of Consumer Electronics (GCCE2014), Makuhari Messe (Chiba, Japan), pp.766-770, Oct. 2014.
- [3] Noriaki Fujishima and Kiyoshi Hoshino, "Fingernail Detection Method from Hand Images including Palm", Proc. of IAPR International Conference on Machine Vision Applications (MVA2013), Ritsumeikan University (Kyoto, Japan), pp.117-120, May 2013.
- [4] Noriaki Fujishima and Kiyoshi Hoshino, "Examination of Regions containing Color similar to that of Nails", Proc. of the 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV2013), Inha University (Incheon, Korea), HybridB-5 (6 pages), January 2013.
- [5] Noriaki Fujishima and Kiyoshi Hoshino, "Fingernail Detection in Hand Images using Difference of Nail-Color Pixels' Density between Vicinity Areas of Fingernails and Skin", Proc. of the 1st IEEE Global Conference of Consumer Electronics (GCCE2012), Makuhari Messe (Chiba, Japan), pp.238-242, October 2012.

## 【報告】

- [1] 藤嶋 教彰, 星野 聖, "仮想ネイルアートシステムのためのネイルチップ重畳手法", 映像情報メディア学会技術報告, 早稲田国際会議場, Vol.38, No.16, pp.47-50, 2014年3月.
- [2] 藤嶋 教彰, 星野 聖, "色の分布特徴を利用した爪領域抽出法", 映像情報メディア学会技術報告, 東京農工大学小金井キャンパス, Vol.38, No.10, pp.33-36, 2014年3月.
- [3] 藤嶋 教彰, 小林 透, 星野 聖, "デジタル写真内に潜む思い出情報の物理的保存法", 映像情報メディア学会技術報告, 機械振興会館, Vol.37, No.44, pp.5-9, 2013年10月.
- [4] 藤嶋 教彰, 星野 聖, "肌領域における色の連続性を利用した爪検出の高精度化", 映像情報メディア学会技術報告, 東京農工大学小金井キャンパス, pp.5-8, Vol.37, No.12, 2013年3月.
- [5] 藤嶋 教彰, 星野 聖, "指先と爪領域重心を用いた仮想ネイルアート", 第18回電子情報通信学会東京支部 学生会研究発表会講演集, 東海大学高輪キャンパス, p.44, 2013年3月.
- [6] 藤嶋 教彰, 星野 聖, "掌領域を含む手画像からの高精度爪認識", FAN2012 予稿集, 沖縄県浦添市でだこホール, 2A-2-8(6 ページ), 2012年8月.
- [7] 藤嶋 教彰, 星野 聖, "手指画像中における爪類似色の密集度に着目した爪認識システム", 第17回電子情報通信学会 東京支部学生会研究発表会講演集, 東海大学高輪キャンパス, p.95, 2012年3月.
- [8] 藤嶋 教彰, 星野 聖, "爪輪郭内外と肌における爪色画素の密度差に着目した手画像からの爪領域検出", 映像情報メディア学会技術報告, 機械振興会館, Vol.35, No.40, pp.9-12, 2011年10月.
- [9] 藤嶋 教彰, 星野 聖, "肌と爪の分離精度向上のための照明光の検討", 映像情報メディア学会技術報告, 東京工科大学蒲田キャンパス, Vol.35, No.15, pp.47-50, 2011年3月.