

俯瞰視カラー画像による個人再識別に関する研究

令和3年3月

和歌山大学大学院システム工学研究科

李 鵬

Research on personal Re-identification by Top-View color image

March 2021

Graduate School of Systems Engineering
Wakayama University

PENG LI

論文概要

個人再識別 (Person Re-Identification) はコンピュータビジョンの中でも重要な研究分野である。人物全身の正面画像或いは側面画像は個人差を反映できる特徴量が一番多く含み、ほとんどの研究に利用されている。しかし、現代社会では、人々のプライバシー保護に対する意識が高まり、個人情報の利用は本人の許可を得なければいけなくなった。本論文では、プライバシーを考慮した個人再識別手法について議論する。

人物全身の正面画像或いは側面画像の利用を避けるために、人物の顔を映らない俯瞰視カメラシステムを構築し、人物の頭の上から俯瞰視画像を撮影する。本研究では、1) 短時間内に人物の服と髪の毛の色は変わらない、2) 俯瞰視カメラから地面までの距離はほぼ一定であると仮定する。その上で、個人再識別できる特徴記述法を提案する。

まず、俯瞰視画像上の人物の頭頂部と肩部の領域の RGB 色情報を利用する「色弁別特徴 (CDF)」を提案した。色弁別特徴 (CDF) では、人物の肩部と頭頂部領域の各画素の RGB 色情報によって該当人物の 4 次元特徴量とする色弁別性を計算する。個人再識別実験より、人物の俯瞰視画像を利用する個人再識別の可能性を確認した。

色弁別特徴 (CDF) で同じ色の種類数の場合での識別が不可能という課題を解決するために、水滴が紙を濡らす現象を模倣している特徴量記述法「水滴レンダリングボックス (WDRB)」を提案した。WDRB 法の特徴量の記述は 3 つの段階に分けられる。1) 3 次元の色登録マップを生成し、人物領域の画素色を登録する。2) 色登録マップから登録したすべての色の間の相対的な距離関係を記述する距離マップに変換する。3) 色のヒストグラムにより登録した色を強調する。最後に得られた強調マップを人物の WDRB 特徴量として利用する。複数の実験より、WDRB 法が従来手法よりの有効性と頑健性を確認した。

最後に、個人再識別精度を一層上昇させるために、新たな SHAL 法を提案した。SHAL 法では、輝度の変化が色に影響しないように、RGB 色空間の代わりに、輝度情報と色情報が単独に表示する YUV 色空間を利用する。また、画像上の人物の領域の画素数は人物の体格情報であり、SHAL 法では、RGB-D カメラを利用し、俯瞰視画像から人物の頭頂部領域と肩部領域を検出する。SHAL 特徴量を記述する時に、UV 色平面内に、頭頂部と肩部のそれぞれの画素について、1) 色のヒストグラムと、同じ色の輝度 y の合計値の分布をそれぞれ求める。2) 色の種類数と色空間内の各色の相対的な距離関係を反映するため、色空間におけるユークリッド距離変換を行う。3) 無彩色画素の割合を求める。また、SHAL 特徴量間の類似度の評価に利用する重み付きバタチャリア距離を提案した。複数の個人再識別実験より、SHAL 法はより有効性を持つことを確認した。

Abstract

Personal Re-Identification is an important research field in computer vision. Since the frontal image or the side image of the whole body contains the largest amount of features, it is used in most studies of personal re-identification. However, in modern society, people are becoming more aware of privacy protection, so the use of personal information of people must obtain their permission. In this paper, we discuss with personal re-identification method that takes privacy protection into consideration.

For avoiding the use of front or side images of the whole person, we construct a top-view camera system that can capture from above the person. In order to realize personal re-identification using a top-view image, in this study, we assuming that: 1) the color of person's clothes and hair doesn't change within a short time, 2) the distance from the top-view camera to the ground is constant. Above this, we propose some feature descriptors for personal re-identification.

First, we proposed a descriptor which named "Color Distinctiveness Feature (CDF)". In the CDF descriptor, the color discrimination which calculated by RGB color information of person's head and shoulders area regard as a personal feature .

Second, in order to improve the defects of CDF, we proposed a feature descriptor "Water Drop Rendering Box (WDRB)" that imitates the phenomenon of water drops wet the paper. The description of WDRB feature can be divided into 3 stages. (1) Generate a 3-dimensional color registration map and register the pixel color of the person area into it. (2) Convert the color registration map to the distance map that describes the relative distance relationship between all the registered colors. (3) Emphasize the registered colors using the color histogram.

Finally, we proposed a novel SHAL method to further improve the accuracy of personal re-identification. When describing the SHAL features, 1) calculate the color histogram and the distribution of the total value of the luminance y of the same color in the head area and shoulders area in the UV color plane. 2) Euclidean distance conversion in the color space is performed to reflect the relative distance relationship between the number of color types and each color in the color space. 3) Find the ratio of achromatic pixels as weight. And we proposed a weighted Bhattacharyya distance which used to evaluate the similarity using the weight between SHAL features.

目次

第 1 章	はじめに	1
1.1	研究の背景	1
1.2	本論文の目的	4
1.3	論文の構成	4
第 2 章	関連研究	7
2.1	個人識別	7
2.1.1	顔による個人識別	7
2.1.2	指紋による個人識別	12
2.1.3	掌紋による個人識別	13
2.1.4	静脈による個人識別	14
2.1.5	声紋 (音声) による個人識別	15
2.1.6	虹彩による個人識別	16
2.1.7	網膜による個人識別	17
2.1.8	歩容による個人識別	18
2.2	個人再識別	18
2.2.1	一般画像による個人再識別	19
2.2.2	真上からの画像による個人再識別	21
2.2.3	画像の色情報を用いた個人再識別	22
第 3 章	色弁別マップを用いた人物の再識別	25
3.1	色弁別度 (Color Distinctiveness)	25
3.2	CDF 法による顔情報を利用しない個人再識別	26
3.2.1	CDF:色弁別特徴	26
3.2.2	CDF を用いた個人再識別	27

3.3	実験	30
3.3.1	個人再識別実験	31
3.3.2	比較実験	32
3.4	まとめ	34
第 4 章	RGB 色空間内の色情報の拡張記述による個人識別	35
4.1	CDF 法の問題点	35
4.2	基本アイデアと処理の流れ	35
4.3	WDRB 法:水滴レンダリングボックス (Water-Drop Render Box)	37
4.3.1	色の登録と色のヒストグラム	37
4.3.2	WDRB 特徴の記述	38
4.3.3	WDRB を用いた個人再識別	44
4.4	実験	44
4.4.1	実験環境と実験方法	44
4.4.2	色のヒストグラムを考慮した色の強調の有効性実験	45
4.4.3	WDRB 法を利用した個人再識別の有効性実験	46
4.4.4	被験者が同じ色の服を着ている実験	49
4.4.5	CDF 法との比較実験	51
4.4.6	異なる光源下での比較実験	52
4.5	まとめ	56
第 5 章	UV 色平面内の色情報の拡張記述による個人識別	57
5.1	WDRB 法の問題点	57
5.2	SHAL 特徴記述法	57
5.2.1	頭頂部・肩部別の画素ベクトルと色のヒストグラム	57
5.2.2	頭頂部・肩部の同じ色の輝度 y の合計値分布	58
5.2.3	色の種類数と位置に基づくユークリッド距離変換	60
5.2.4	SHAL 特徴記述法作成プロセス	60
5.3	重み付きバタチャリヤ距離による SHAL 特徴量間の評価	63
5.4	実験	64
5.4.1	予備実験	65
5.4.2	人物領域の自動検出における前処理	65
5.4.3	SHAL 法を用いた個人再識別システム	69

5.4.4	私服を着る実験	70
5.4.5	同色の服を着る実験	74
5.4.6	提案手法の有効性の検証実験	76
5.4.7	比較実験	76
5.4.8	サンプリング間隔変化による安定性実験	78
5.4.9	頭頂部と肩部を一つの領域とする場合の個人再識別	80
5.5	まとめ	82
第 6 章	おわりに	83
6.1	本論文のまとめ	83
6.2	今後の課題	85
	参考文献	87
	研究業績	103
付録 A	CDF 法による物体の再識別	107
付録 B	WDRB 法による物体の再識別	111
付録 C	OiF 法による物体の再識別	115
	C.0.1 段ボールの再識別実験	115
	C.0.2 商品の再識別実験	117
C.1	まとめ	120
	謝辞	121

第 1 章

はじめに

1.1 研究の背景

1990 年代後半以降、情報技術の発展に伴い、あらゆる分野の情報の電子化が進んだ。人々の生活を支える機械（家電製品や ATM など）も知能を持つ機器に進化した。この中には、知能ロボットのように正確に情報の収集・整理及び分析を行うためのセンサーが必要となってきた機器も出てきた。汎用のセンサーとして、コンピュータビジョンの利用が進んできた。これは、CCD カメラを人間の目のように扱う技術で、周囲の環境や物体の認識・識別ができるものである。

コンピュータビジョンとは、静止画もしくは動画のデータをもとに、コンピュータに人間の視覚に近い、もしくはそれ以上の視覚認識機能を付与しようとする研究である。高度情報先進社会では、人にやさしいサービスを提供するために、人間の周囲の環境を検出及び識別する機能が求められている。例えば、ショッピングモール・商店街などのインフォメーションに使われている自動応答ロボットは、目の前に顧客がいるかを判定して挨拶をするや、撮影範囲の人の検出で、店舗への呼び込みなど、人間のように声を掛けることができる。対話をする際には、顔を検出して、顧客に視線を合わせる機能を持っている。これらは、すべてコンピュータビジョンによって実現できる。

近年、社会の発展に伴い、「人がいるかどうか」のように単純な人の検出機能だけでなく、人を見分ける「個人識別」（「個人認証」とも呼ぶ）機能も求められている。個人識別とは、人に関わる情報を利用し、既知の情報との照合により人物を特定することである。利用される人（本人）情報によって、以下の 3 種類がある。

(1) 知識情報：人が知っていること

例：パスワード・暗証番号・生年月日・携帯番号など

(2) 所持情報：人が持っているもの

例：鍵・身分証・判子など

(3) 生体情報：人の自身の特徴

例：指紋・網膜など

センサー利用等により、情報の取得の利便性が進み、また、システム構築コストの低下により、知識情報と所持情報の実応用例はあらゆる領域で使われている。生体情報を利用するには、人の生物学的な身体の特徴を取得するセンサー等の専用設備が不可欠であり、その生体情報を利用した個人識別には、各利用現場に専用設備が必要とされている。その為、生体情報の利用は銀行のATMやビルセキュリティでの普及が中心である。一方、コンピュータビジョンでの生体情報を利用した個人識別は、一般的に人の顔画像を利用している。最近では、PCやスマホの顔認証によるログインは普及している。

個人識別の他の利用用途としては、セキュリティーシステムとして、一部特定の場面で利用する複数カメラ間の人物対応付け（人物照合）を行う個人再識別があり、盛んに研究がなされている。

個人再識別に関する研究では、人物の領域、つまり人物を特定する特徴の抽出領域を決めることが重要である。人物画像の種類により、個人再識別の手法は2種類に大別できる：

- 1) 人体の全身正面・側面画像を利用する手法
- 2) 俯瞰視画像を利用する手法

人体の全身正面（図1.1の(a)）・側面（図1.1の(b)）画像を利用する手法は、人物の顔領域を含んでいる。顔領域には個人差を反映できる特徴を沢山含んでいる。ほとんどの手法では顔領域と身体の局所領域と併用することで個人識別精度を上げている。この種の手法の認識アルゴリズムを評価するために多数の共通データベース（VIPeR, iLIDS, ETHZ など）が公開されている。このデータベースの画像セットではさまざまな角度で撮影した人物の顔や人体画像を含んでいるため、個人に依存した特徴量の識別率を評価し易い。しかし、用途により個人情報情報の漏洩を防ぐ必要があるため、個人情報を含む画像は、許可なく利用することが制限されている。

また、人物の顔領域を撮影しない俯瞰視カメラを利用した人物の俯瞰視画像を利用する個人再識別手法もある。この種類の手法では、人物領域の情報が少ないため、個人の特定精度が落ちるため、利用制限が緩和されることがある。

企業等の自動出勤・退勤管理システムや、福祉施設内の被介護者の見守り等のコン



図 1.1 人物画像の例

コンピュータビジョン応用システムには、本人認証 (Verification of an individual) を目的として、入室・退室の人物画像間で個人再識別が必要とされている (図 1.2)。そのため、個人を特定しやすい顔・指紋・掌紋・網膜の画像が一般的に利用されている。しかし、このような画像には保護しなければならない個人情報を多く含むため、データの収集・取得には本人の同意が不可欠であり、その情報の保管には漏洩防止策が必要である。一方、実際には、厳重に保管・漏洩防止策を施したシステムであっても記録・登録されている個人情報の漏洩が起きる可能性がある。このようなことから、個人情報の含まない人物画像を利用した個人再識別システムの開発が望まれている。

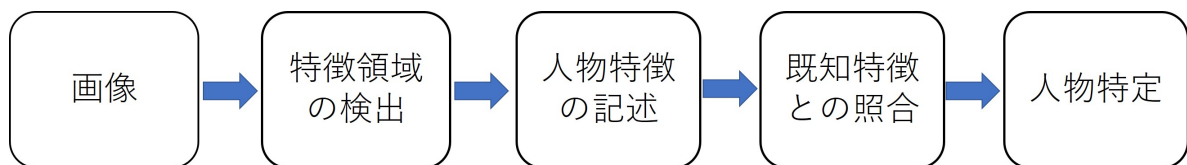


図 1.2 一般個人再識別手法の処理流れ

誰がどこで電車・バスに乗車・下車したかを個人単位で追跡できるシステムや、デパートやスーパーマーケットの平均ショッピング時間の統計データを求めるシステムなどへの応用を想定して、本論文では、カメラに映った人物が再度カメラに映ったときに、同一人物であることを確認 (人物照合) できる個人再識別の方法とそれを利用した俯瞰視システムについて論じる。

1.2 本論文の目的

個人識別手法では、指紋・静脈・音声・虹彩・網膜およびDNAにとって、専用設備が必要なので、これらの特徴を用いたシステムを導入するところでは、設備の設置も必要となる。静脈や掌紋の特徴での識別精度は体調・環境の変化や手のひらの状態に影響される問題がある。虹彩・網膜による識別は読み取り装置が大型・高価で一般の民間応用にはならない。これらの個人識別の特徴はすべて重要な個人情報であり、情報の慎重な使用・管理および保護を行わなければならない。同様に、個人再識別手法に利用される全身画像は撮影角度より、人物の顔が映される場合もよくあるので、データの収集と利用には本人の同意が不可欠であり、そのデータの保管には厳重な漏洩防止策などの個人情報保護対策が必要である。したがって、個人のプライバシーを保護できる特徴記述法が求められている。また、個人識別と比べて、個人再識別手法が撮影角度・撮影環境などの影響で、望ましい特徴の記述は保証できない場合があり、よりチャレンジ性が高い研究分野となる。これで、人物のプライバシーを保護しながら個人の再識別を行うことは本論文の研究目的とした。

本論文では、人物の顔領域の撮影・利用を避けるために、日常生活に人物の顔をよく撮影できない俯瞰視システムを利用する。人物の俯瞰視画像内の「人物の頭と肩領域」の「色情報と画素数」のみを用いて、個人再識別を可能とする特徴記述法を提案する。俯瞰視カメラの下を通った際に一枚の俯瞰視画像を撮影し、その画像から個人の特徴を記述・登録し、再度カメラの下を通過したときの一枚の俯瞰視画像から特徴を記述し、その特徴を持つ登録データと照合することで、再識別を行う。

本研究では、二つの仮定条件を設ける。1) 短時間では人物の服や髪の色は変わらないと仮定する。髪の色や、髪型の変化などには個人性（個人の特徴情報）があり、短時間では不変性があるので識別に有効な情報になる。2) 床から俯瞰視カメラまでの距離が一定であると仮定する。人物の身長や体格が異なると、撮影された俯瞰視画像内の人物領域（頭頂部と肩部）における画素数が異なるので、識別に利用できる。

1.3 論文の構成

本論文は、下記の6つの章から構成される。

第1章：はじめに

本章では、研究背景や目的などを述べる。

第 2 章：

この章では、本研究と似ている個人識別手法を含めて、個人再識別の関連手法について述べる。

第 3 章：

この章では、色弁別特徴 **Color Distinctiveness Feature (CDF)** について説明する。CDF は、俯瞰視画像の人物領域のメインエリアとサポートエリアの 2 つのエリアの色を使用して、ベイズ定理に基づいて計算する。CDF を用いて、俯瞰視システムによる個人再識別を行う。登録された既知人物の CDF と入力 of 未知人物画像から得られた CDF との類似度評価はバッチャリヤ距離またはハミング距離より推定する。

第 4 章：

この章では、CDF 法より高精度の俯瞰視画像用特徴記述法 **Water-Drop Render Box (WDRB)** を提案する。WDRB 法では、人物領域のカラーとそのヒストグラムを利用し、色の登録、距離マップの変換およびヒストグラムを使用した色の強調を行って、4 次元行列を用いて人物の特徴量を記述する。CDF 法との人物の再識別の比較実験より WDRB 法がより有効性を持つことを確認した。さらに、異なる光源下で他人の手法との個人再識別の比較実験より WDRB の安定性も確認した。

第 5 章：

この章では、人物の俯瞰視画像から人物の頭頂部と肩部をそれぞれ検出し、二つの部位の色情報と見えている体格情報を拡張記述する手法 **Shoulders and HeAd Learning (SHAL)** を提案する。この手法では、UV 色平面内に、頭頂部と肩部の画素について、1) 色のヒストグラムと、同じ色の輝度 y の合計値の分布を求める。2) 色の種類数と色空間内の各色の相対的な距離関係を反映するため、色空間におけるユークリッド距離変換を行う。3) 無彩色画素の割合を求める。多数の比較実験で SHAL 法の有効性および安定性を確認した。

第 6 章：おわりに

最後に、本研究の成果を総括し、今後の課題について議論する。

第2章

関連研究

コンピュータビジョンを利用した個人識別の研究について、人間の生体情報を利用するのが多数報告されている。本章では、個人識別手法の各種の方法と、本研究で対象とする個人再識別に関連する手法について述べる。

2.1 個人識別

個人差を反映できる人間の身体的特徴は生体（一般的に人を指す）特徴と呼ばれる。個人識別の研究には、生体特徴を利用するのは一般的であり、特徴の属性によって、個人識別の研究は以下8種類に分けられる：1) 顔識別、2) 指紋識別、3) 掌紋識別、4) 静脈識別、5) 音声識別、6) 虹彩識別、7) 網膜識別、8) 歩容識別。

2.1.1 顔による個人識別

最も人を見分ける情報として人の顔の特徴がよく個人識別の手法に利用されている [1][2][3]。テクスチャー情報は人間の顔全体の特徴の一つであり、よく利用されている [4]。人間の顔には複数の器官（目、鼻、口など）があり、且つそれらの各器官は個人の特徴を反映できる。また、顔局所特徴を利用する手法や顔全体の特徴を利用する手法はそれぞれ多数研究されている。1枚の顔画像に対して、人間の特徴をできる限り細かく記述する場合は、主に局所的な特徴の手法を利用している。例えば、注目画素に対する周辺画素の濃淡パターンを記述する手法LBP(local binary patterns)[5][6]、画像の回転やスケール変化に不変な特徴記述法 SIFT(Scale Invariant Feature Transform)[7]、SIFTに基づく高速化記述法 SURF(Speeded-Up Robust Features)[8]、入力画像をグレースケール化し、各ピクセルの勾配方向と強度を算出する特徴記述法 HOG(Histogram of Oriented Gradients

) [9], 携帯端末に利用する高速な特徴記述法 RIFF [10] が提案されている. また, メモリの消費量や CPU の計算コストを減らすために, パッチの輝度分布から特徴量をバイナリ化する特徴記述法 BRIEF (Binary Robust Independent Elementary Features) [11], BRISK (Binary Robust Invariant Scalable Keypoints) [12], ORB (Oriented FAST and Rotated BRIEF) [13] と FREAK (Fast Retina Keypoint) [14] などが提案されている. 局所特徴を利用する個人識別手法は良好な識別精度を得るために, 入力する人間の顔画像に対して, 顔の各器官・顔の特徴点をよく撮影するには高解像度が必要である. 局所的な特徴を利用する個人識別手法において, 1 枚の人間の顔画像から各器官を検出して利用するのが一般的であり, 識別精度に応じて, 特徴量の生成速度や低計算コストを考慮した研究が中心となっている.

一方, 人間の知的な振る舞いを模倣したコンピュータを構築するため, 手法の効率性の追求ではなくて, 高識別精度を追求する研究者らは, 同一人物の複数枚画像を利用して, 該当人物の顔の特徴量のルールやパターンを発見し, それを利用して個人識別を行う「機械学習」にも注目して研究を進めている. 機械学習では, 学習データの中での重要な要素を手で判断・調整することで識別精度をあげる研究もある. 機械学習は基本的に以下の 3 種類に分けられる.

- (1) 教師あり: 学習データに正解を与えた状態で学習させる手法である. 代表的な手法は「回帰」と「分類」がある. 正解例の学習データがより多量になるほど, 識別精度は上がる性質がある.
- (2) 教師なし: 正解の学習データの代わりにデータに共通する本質的な特徴を分析し, グループ分けやデータの構造・パターンなどを抽出する手法である. 代表的な手法は「クラスタリング」と「次元削減」がある. (例: Ward 法・群平均・K-means 法 (クラスタリング), 主成分分析 PCA・k 近傍法 (次元削減))
- (3) 強化学習: 機械 (コンピュータ) が一連の行動 (パラメータの組合せ) をとった結果ごとに報酬 (識別成功率) を設定し, その報酬 (識別成功率) が最大化するように機械が試行錯誤し, 学習することで精度を上げる手法である.

顔識別に利用される代表的な機械学習アルゴリズムには, 以下のものがある. これらのアルゴリズムは, 学習データとして, 入力画像の画素値のまま利用するか, 他の特徴記述法より記述した特徴量を利用するケースが一般的である. 例えば, 回帰と分類に両方に利用される有名なアルゴリズム SVM (Support Vector) [15][16][17] は, 入力画像をそのまま利用する手法 [18][19], 遺伝的アルゴリズム (GA) によって記述する特徴量を利用する顔識別手法 [20], 顔のボールテクスチャー情報を利用する手法 [21] がある. それ以外

に、PCA に基づく手法 [22][23][24][25]), 顔のボールテクスチャーと Ada-boost 分類器 [26] を組合せた手法 [27], LBP 特徴と Boosted MTL 学習法 [28] の組合せ手法も提案されている。

また、近年コンピュータのハードウェアの発達により、機械学習アルゴリズムの計算時間は段々短くなり、研究者らはより高精度の学習手法「深層学習」を目指して研究している。深層学習（ディープラーニング）は教師ありの機械学習手法を多層ニューラルネットワーク（Neural Network）に適用して発展してきた。ニューラルネットワークでは人間の脳の神経細胞（ニューロン）と神経回路網（シナプス）をそれぞれノードとリンクとして再現して、各層は複数のノードが重み付きのリンクで結ばれる構造をしている（図 2.1 (a) に参照）。人間の脳内のニューロンは電気信号として情報を伝達する。その時にニューロンとニューロンを繋ぐシナプスの繋がり強度によって、情報の伝わりやすさが変わってくる。ニューラルネットワークでは、この情報の伝達ステップをリンクという部分で「重み」と、重みの和を閾値判定する方式で再現する。

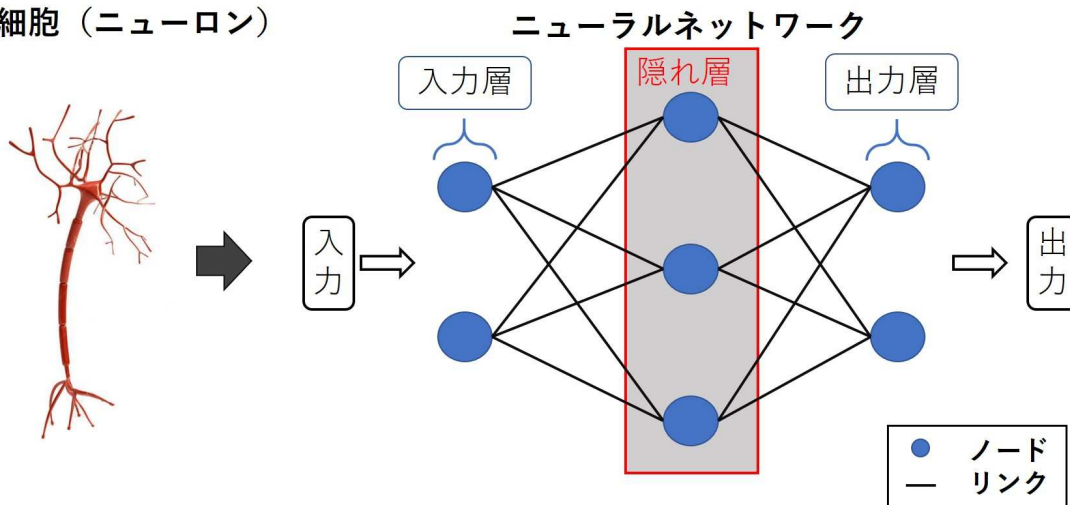
3層ニューラルネットワークは入力層、隠れ層、出力層でネットワークを構成するが、3層で処理、表現できる機能には限界があるため、情報の複雑さに対応するように「層」の数を増やしたものをディープニューラルネットワーク（Deep Neural Network: DNN）と言う（図 2.1 (b) に参照）。深層学習（ディープラーニング）はネットワークの層を増やし、情報の複雑さに対応できるようにしたことと、学習方式の革新で、識別精度が一般の従来手法に比べて、飛躍的に高まっているのが特徴である。ディープラーニングには様々な学習方法があり、活用する分野によってニューラルネットワークの構造も異なる。近年画像処理の分野によく利用されているネットワークの構造は以下の2つがある。

- ① 畳み込みニューラルネットワーク（CNN: Convolutional Neural Network）。CNN の特徴は、隠れ層を「畳み込み」と「プーリング」という2種類の層に分けて定義し、それぞれの層で異なる操作を実行する点である（図 2.2 に参照）。

畳み込み層では、さまざまなフィルタをスライドして計算することにより、入力された画像から「何が特徴か」を検出して抽出する。抽出した特徴を新しい画像データ・特徴マップとして出力する。画像全体をフィルタがスライドするので、特徴がどこにあっても抽出できるので、抽出した特徴が移動不変性または位置不変性を持っている。また、特徴マップのサイズも入力画像より少し小さくなる。

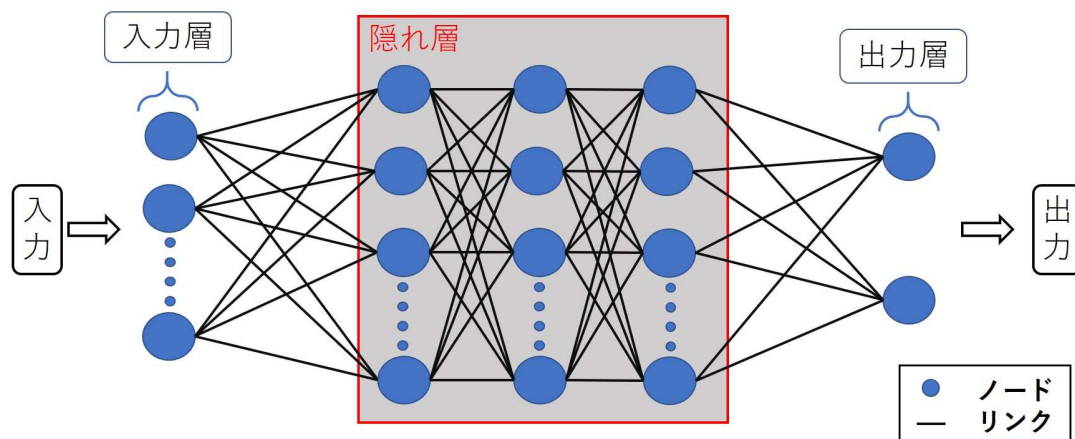
プーリング層では、畳み込み層で抽出された特徴を圧縮し、重要ではない部分を削ぎ落とすことがされている。つまり、特徴として重要な情報を残しながら元の画像を縮小することである。このプーリング処理することによって、画像が数ピクセル

神経細胞（ニューロン）



(a) ニューラルネットワークの構造イメージ

ディープラーニング



(b) 深層学習（ディープラーニング）の構造イメージ

図 2.1 ニューラルネットワークと深層学習の構造イメージ

移動したり、回転したりしても、それらの違いを吸収してほぼ同じ特徴量を抽出してくれるようになる。

- ② 敵対的生成ネットワーク (GAN: Generative Adversarial Networks). GAN を利用することで、入力データから特徴を学習することで、その特徴に類似するが実在しないデータを生成したり、存在するデータの特徴に似せたデータも生成できる。GAN は2つのニューラルネットワーク (生成ネットワークと識別ネットワーク) で構成される (図 2.3 に参照)。元にしたデータ (本物データ) から特徴を抽出し

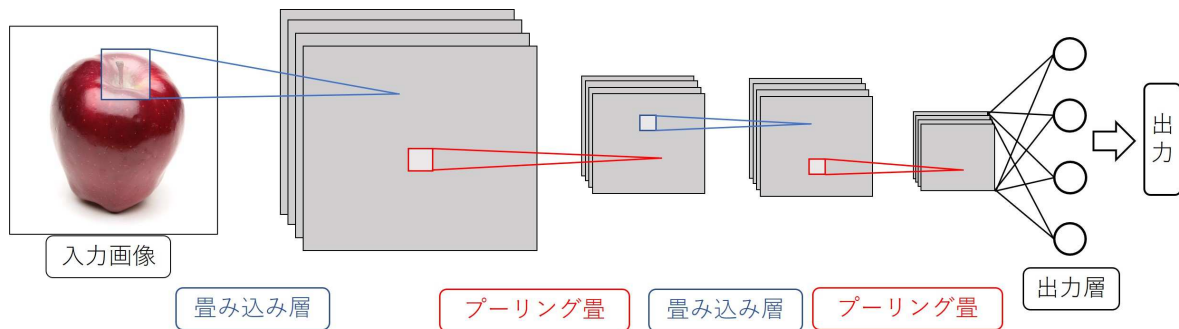


図 2.2 畳み込みニューラルネットワーク (CNN) の処理流れのイメージ

て、似せたデータ（偽物データ）を作成し、識別ネットワークより偽物と本物の真偽を判定する．このように、生成するネットワークと識別するネットワークの相互作用によって、生成したデータは少しずつ本物に近いものになっていく．

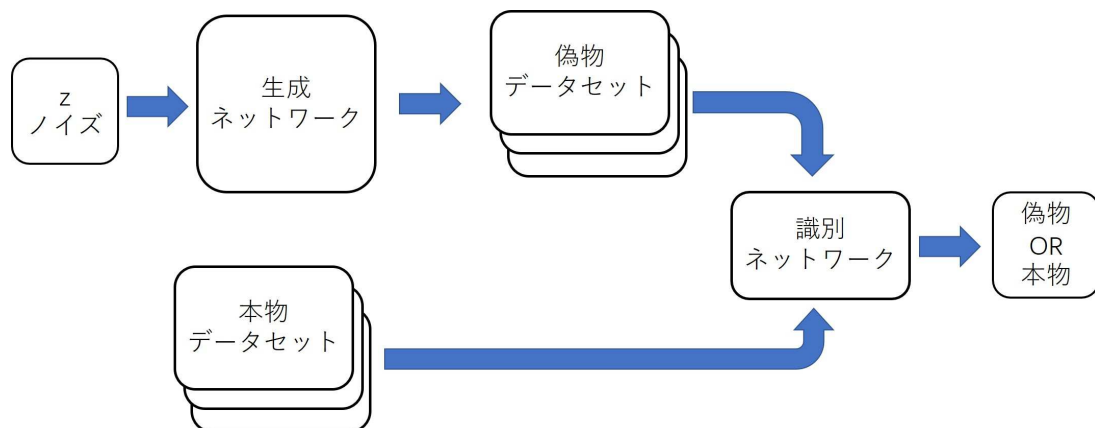


図 2.3 敵対的生成ネットワーク (GAN) の処理流れのイメージ

ディープラーニングの研究の中で最も進められている画像認識、物体検出、領域推定などの分野で、畳み込みネットワーク (CNN: Convolutional Neural Network) がよく活用され、個人識別に関する研究も多数報告されている [29][30]．ほとんどの研究では人の顔画像 (多くの画像要素) をそのままディープラーニングの入力としている．また、ディープラーニングによる人 (物) の特徴量の学習には、同じ人 (物) の複数枚の画像が必要であり、学習精度評価用の汎用の研究専用画像セットがある (LFW[31][32], YTF[33], FERET[34]), WDRRef[35], CelebFaces[36], FaceBook[37], Facenet[38]．

ディープラーニング手法のメリットは学習用の画像の質を保証した上、学習データ数が多いほど識別の精度が高くなることで、他の一般手法で実現できないほど識別精度が高く

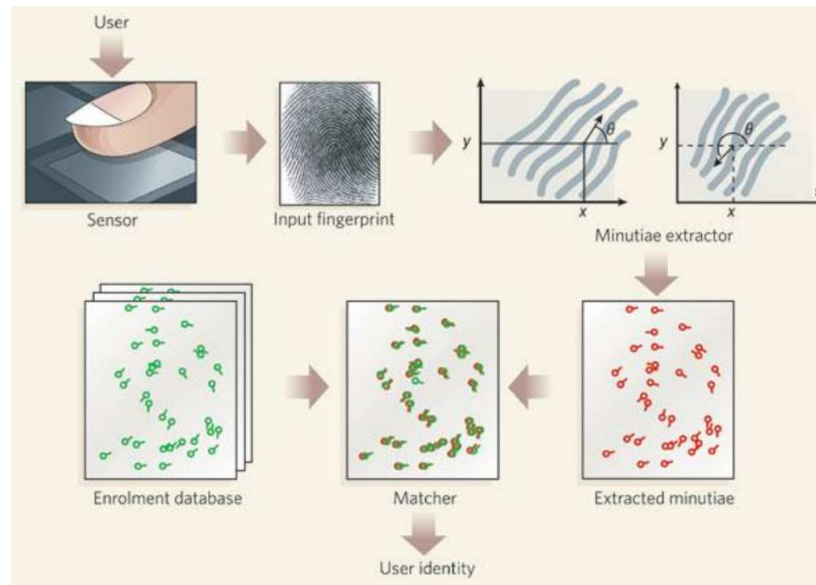


図 2.4 指紋による個人識別の処理流れ [43]

なることである。一方、多数の画像から特徴の抽出は従来の一般手法よりもかなり計算時間がかかるのは当然であり、一つのデメリットとして考えられる。また、人（物）を識別するために、学習用のデータセットを事前に用意しなければならないことで、個人識別システムの稼働中に新人物の特徴を瞬時追加できない欠点がある。

2.1.2 指紋による個人識別

指紋は人間の皮膚紋理の一種として、形状は人によって全て異なり、遺伝子のように、終生不変という特徴を持つと言われている。この不変性のおかげで、指紋による個人識別手法 [39]、および自動化指紋認証システム [40][41] などは多数報告されている [42]。指紋による識別の基本の流れは図 2.4 に示す。

指紋による識別手法では、指紋画像から抽出する微細特徴（Minutiae Feature）の利用は一般的である [44][45][46]。微細特徴の他に、G. Bebis らは指紋画像から抽出した微細特徴に対して、de launay 三角分割を用いて新しいトポロジー構造体を提案した。この構造体によって、重要度の高い微細特徴三角グループを選出することにより、識別精度を犠牲にすることなく識別時間の短縮やメモリの消費量の減少を実現した [47]。識別精度を向上するため、指紋特徴のみならず、他の生体特徴と併用して個人識別を行う手法も提案されている。例えば、指紋特徴と顔特徴との併用手法 [48]、指紋特徴、顔特徴と音声特徴との併用手法 [49]。

他の生体特徴を利用せず、指紋画像のみを利用して識別精度を向上する手法も研究されている。U Park らは、微細特徴に SIFT 特徴量を加えて指紋識別を行っている [50]。特徴点の検出について、SIFT を利用することで、検出した特徴点の数は微細特徴を利用する場合より相当に多くなり（図 2.5 に参照）、微細特徴を利用する手法の成功識別率の向上が実現できた。

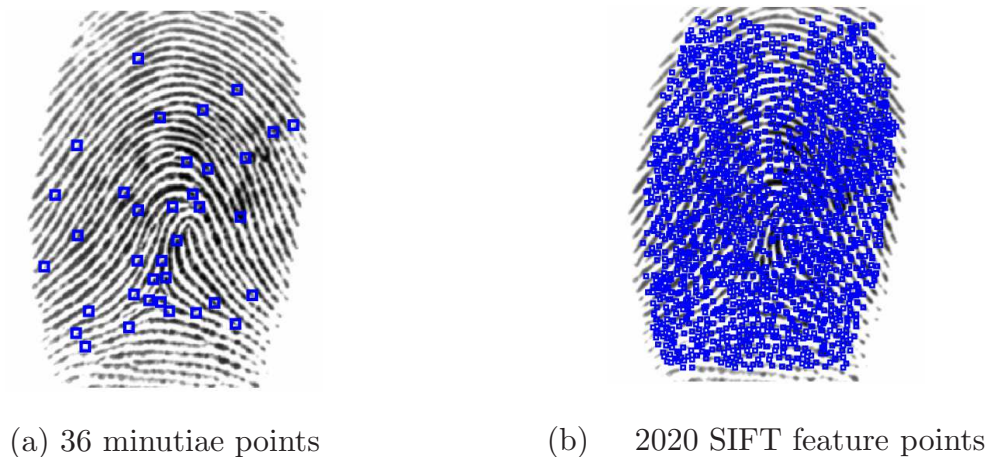


図 2.5 同一画像から検出した Minutiae と SIFT の特徴点 [50]

SIFT 特徴の導入より、識別率は向上できたが、処理コストも高くなったデメリットがある。この問題点を改善するために、Awad らは特徴量のマッチング処理プロセスを減らすことができる Matching Score Matrix (MSM) という評価アルゴリズムを導入した [51]。また、より高精度の識別率を得るために、機械学習手法を利用する手法も発表されている。例えば、[52] で、Shervin M らは平行移動不変性を持つ CNN ベース離散ネットワークを提案し、指紋の識別精度を 98% まで向上できた。

指紋情報の自身が弁別性の高い情報を持っているので、指紋のテクスチャー情報さえしっかり記述できれば、ほとんどの個人識別は失敗しないと考えられる。指紋の特徴の読み取り装置は、現在、小型で導入コストが安価となった反面、指の状態が認識精度に影響している。

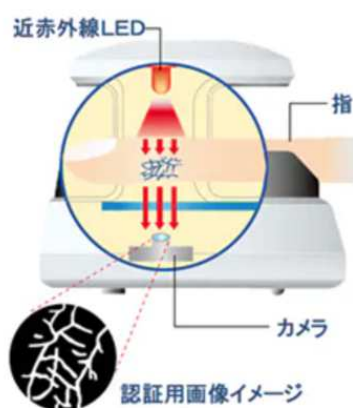
2.1.3 掌紋による個人識別

皮膚紋理特徴量である掌紋も個人識別に利用される。指紋と同様に弁別性の高い情報を持っているので、掌紋を利用した個人識別の研究も少なくない。ほとんどの手法は手掌の画像から掌紋の見え特徴を記述している。例えば、掌紋の幾何情報や手掌の構成を特徴と

する研究 [53][54][55], 掌紋画像に対してフーリエ変換を行い, 掌紋が周波数領域での表現を特徴とする手法 [56], 低解像度の掌紋画像を利用することより, オンラインでの個人識別を実現する手法 [57], Karhunen-Loeve(KL) 変換より, 掌紋特徴をベクトル化した「eigenpalms」特徴記述法 [58] などが提案されている. Manisha P らは, 識別率を高めるために, 掌紋画像を四つ領域に分割し, 各領域に離散コサイン変換 (Discrete Cosine Transform) を適用して掌紋特徴量ベクトルを生成する手法を提案している [59]. 識別精度をさらに向上するために, 単純な掌紋特徴だけでなく, 他の特徴を併用する手法も登場している. Abdallah M らは, 離散コサイン変換 (DCT) や離散フーリエ変換 (DFT) を用いて取得した二種類の特徴量の融合特徴量 [60], 掌紋特徴と指関節の皮膚紋理を併用する手法 [61] を提案している. Shanmukhappa A らは, 掌紋特徴と手の幾何的な特徴も一緒に利用し, SVM 分類器を利用して, 99.19% の識別精度を達成した [62]. 手のひら全体の紋理の特徴は指紋に比べると読み取り装置が大型で, 手のひらの状態が認識精度に影響する問題点がある.

2.1.4 静脈による個人識別

人間の手や指の皮膚の下に血管・毛細血管の分布は個人差があるので, 静脈による個人識別の研究も盛んである. 人間の手や指にある静脈 (血管) の特徴を利用するのが一般的である. 図 2.6 (a) と (b) はそれぞれに指と手掌の静脈の採集デバイスの例を示している.



(a) 指の静脈の採集装置 [63]



(b) 手のひらの静脈の採集装置 [64]

図 2.6 指や手の静脈を採集するデバイスの例

この静脈識別の研究では, 採集装置より取得した静脈画像から特徴の記述を中心として

いる。Miura Nらは、指の静脈画像（グレー画像）から画素の輝度値より静脈の検出を行い、検出結果の画像を二値化して特徴量として利用する手法 [65] および静脈検出の頑健性を強化する手法 [66] を提案している。Meng Xらは、画素の勾配方向を利用し、静脈画像から指の静脈特徴を記述するローカル特徴記述法（LDC）を提案している [67]。指の他に、手の平と裏の静脈特徴を利用する手法も多数提案されている。Lin CLらは、赤外線（IR）カメラを入力装置として、手のひらの熱画像（温度分布画像）を取得し、熱画像の性質に基づく流域変換ツール（watershed transformation tool）を修正することによって、静脈の特徴点を抽出して利用する手法を提案している [68]。Zhang YBらはマルチスケールフィルターを提案している [69]。このフィルターにより手のひらの静脈画像から大きい静脈だけでなく、非常に細かい静脈までの検出もでき、識別の成功率を向上させている。Zhou Yらは、入力手掌の静脈画像に対して、1) マルチスケールより検出した静脈、2) 二値化した静脈図、3) 局所ラドン変換をした静脈図、4) 直交線順序法より算出した血管間の関係を融合した特徴記述法 [70] を提案し、識別精度をさらに高めた。また、彼らは指の静脈と指紋特徴との併用手法 [71] や掌紋と手掌の静脈との併用手法 [72] も提案している。静脈特徴を利用する識別精度が高く、特徴量の偽造・模倣・なりすましなどが困難だが、静脈を読み取る装置が大型で体調変化や環境が識別精度に影響する問題点がある。

2.1.5 声紋（音声）による個人識別

人の声が個人特有の情報を含んでいるので、スピーチコンテストや講演会などの場面に利用される声紋による個人識別（話者識別）も相当に人気がある研究分野である [73]。音声識別において、個人識別を正確に評価するために、音声特徴を登録する際に、同じ発声原稿（スピーチ内容）を利用して行うテキスト依存型手法（Text-dependent）と発声原稿によらないテキスト独立型手法（Text-independent）がある [74]。初期の研究では、テキスト依存型手法が主流であり、欧米では盛んに研究されていた。近年の研究では、応用範囲の広さの利点からテキスト独立型の研究にシフトしている。

どんな音声識別手法においても、音声特徴抽出と特徴のマッチングの二つの処理が必要である。音声は波形の形式で記録されるが、波形からの特徴量を記述する方法にはいくつかある。例えば、Linear Prediction Coding(LPC)[75]、Mel-Frequency Cepstrum Coefficients (MFCC)[76]、Perceptual Linear prediction (PLP)[77]、Hidden Markov Modeling (HMM)[78] 法などがある。一方、記述された二つの音声特徴量は同一人物から発声したものかを判定する類似度の計算方法も重要である。例えば、Dynamic Time Warping (DTW)、Vector Quantization (VQ) 法などがある。特徴記述法と類似度計算

法の組合せによる個人識別の手法は多数報告されている。例えば、LPC + Oracle 分類器 [75], MFCC + ユークリッド距離 [79], CC (Cepstrum Coefficients) + DTW [80], MFCC + VQ [81], MFCC + DTW [82].

また、多数の特徴量に応用できる学習モデルとして、混合ガウスモデル (Gaussian mixture models (GMM)) のような特徴の学習モデルも提案されている。GMM は入力データに対して、どのようにデータ点が分布しているかのパラメータを推定する手法である。GMM をベースにして、多数の個人識別手法が提案されている [83][84][85][86][87]. 他の個人識別研究と同様に、ニューラルネットワークのおかげで、より高い識別精度の手法が登場してきた [88][89][90][91].

音声 (声紋) による個人識別は人が講演中の状態でも手軽に行えるが、体調変化や周辺の音が認識精度に影響する問題点がある。

2.1.6 虹彩による個人識別

人間の目の虹彩は指紋のように、人によって異なる特性を持っているので、様々な場面で個人識別に使われている [92][93][94][95].

Zhu Y らは Gabor フィルター + ウェーブレット変換 (wavelet transform) 重み付きユークリッド距離分類器 (Weighted Euclidean Distance classifier) を考案している [96]. この研究では、虹彩の全範囲のテクスチャー特徴を利用するが、虹彩の細かい特徴の反映はできない問題点がある。この問題点を解決するため、Ma L らは虹彩の全範囲のテクスチャー特徴だけでなく、局所テクスチャー特徴も記述する手法を提案している [97]. また、Ma L らは Gabor フィルターベースの circular symmetric filter (CSF) フィルター + the nearest feature line (NFL) 分類器の手法 [98], ウェーブレット変換より虹彩の領域から明暗変化の激しい場所を特徴として利用する手法 [99], エルミートガウス関数 (Hermite Gaussian function) より、虹彩領域のスペクトル強度の局所的変化を特徴とする手法 [100] も提案している。これ以外、ガボールのコンプレックスパスバンドフィルター (Gabor's complex 2D Passband filters) とハミング距離 (Hamming distance) と併用する手法 [101], 局所特徴ベースの特徴マッチングアルゴリズムとカスケード方式の虹彩識別法 [102], より弁別性の高い虹彩特徴の取得および高速なマッチングを実現するため、虹彩のセグメンテーション・画像の増強・特徴記述および評価を含む総合的な手法 [103], SIFT を利用する手法 [104], および近年のニューラルネットワークを利用する手法 [105][106][107] も提案されている。

目の虹彩の特徴は経年変化が少なく、精度も非常に高いが、照明が認識精度に影響する。

2.1.7 網膜による個人識別

網膜画像は眼科医が目の治療のために利用する眼底カメラを使用してスキャンされる画像である。網膜をスキャンするために、人は目をカメラのレンズの非常に近くに付ける必要がある。スキャンプロセス中に、人は目の静止状態を維持しなければならない。また、ノイズ干渉を避けるためにメガネを取り除かなければならない。

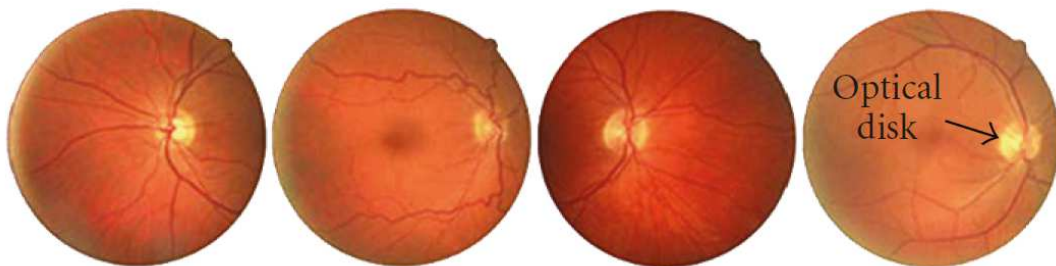


図 2.7 網膜スキャンナーより得られた網膜画像例 [108].

図 2.7 は網膜スキャンナーより得られた網膜画像の例である。網膜の特徴は血管のパターンであり、指紋や虹彩のように、2 人の人間が同じ網膜パターンである確率は極めて低いので、個人識別には非常に信頼性の高い特徴である。網膜識別に関する最も初期の研究の 1 つは、Robert H によるものである [109]。それを基本に多数の網膜識別に関する研究は報告されている。Shahnazi M らは不連続勾配方向分析を利用し、ウェーブレットベースの網膜識別法を提案している [110]。Farzin H らは血管セグメンテーション、特徴生成、特徴マッチングを含む 3 つの主要なモジュールで構成される網膜識別システムを提案している [108]。このシステムでは、ウェーブレット変換を利用し、光ディスク周りの各円形の関心領域 (ROI) から回転不変のテンプレートが作成される。次に、これらのテンプレートを 3 つの異なるスケールで分析し、直径サイズに応じて血管を分離する。最後に、各スケールでの血管の位置と方向を使用して、データベース内の各被験者の特徴ベクトルを生成する。Akram M. U らは 4 つのステップの網膜識別システムを提案している [111]。このシステムでは、より弁別性の高い血管特徴を抽出するために、二つの前処理を加えた。1) 背景差分法より暗い背景からの網膜領域の分離、2) HSI 色空間を利用してノイズの除去。Qamber S らは網膜の血管の終末点と分岐点を特徴とするシステムを提案している [112]。Alexandru P.C らは血管の分岐点特徴を抽出するために、SIFT ベースの特徴記述法とスパース分類器との併用手法を提案している [113]。

同様に、ニューラルネットワークの利用より、多量の網膜画像セットから、より頑健な

網膜識別を行う手法が報告されている [114][115].

網膜画像の取得は一般的に医学界で目の病気の治療の目的で利用されている非常に高価な専用機器が必要である。網膜による個人識別は精度が極めて高いが、専用機器の普及が難しいという問題点がある。

2.1.8 歩容による個人識別

無意識のうちには人は個人に特有な歩き方をしている。つまり、動作は個人性を反映する特徴の一つとも言え、歩容に関する研究は盛んに進められている [116][117]. 歩容の特徴を記述するために、被験者の時系列の全身画像を利用するのが一般的である。Collin R T らの研究 [118] や Wang L らの研究 [119] では人の輪郭より歩容の特徴を記述している。Lee L らは入力画像から検出した人物領域の重心と人体各部位の比例配分で、人物領域を7つのサブエリアに分割している。さらに、各サブエリアにおいて、楕円フィッティングで人体部位を楕円でマークし、全部で7つの楕円を求め、その重心・長短軸の長さ・長軸の傾きなどの情報を纏めて特徴量として利用する手法を提案している [120]. BenAbdelkader C らは人物を追跡し、一定期間内の移動距離によって推定される歩幅とカメラキャリブレーションより推定された歩調を利用する手法を提案している [121]. Kale A らは、人の歩行の時系列画像から生成した人体が画像上の輪郭より算出した各部位の幅情報を含む幅ベクトルをベースにして、ダウサンプリングベクトル、平滑化された幅ベクトル、速度プロファイルなどの歩行特徴を記述して個人識別を行っている [122]. これ以外に、Kale A らは人の時系列画像から生成した人の幅と人の画像上の二値化輪郭情報と併用する手法 [123] や、顔特徴と画像全体のローカル距離尺度との併用する手法 [124] も試していた。Begg RK らは歩行者の歩行中の最小足クリアランス (MFC) 情報を利用し、SVM 分類器より個人識別の手法も提案している [125].

歩容から特徴を記述するには、時系列の歩容画像が必要のため、画像セットの準備や撮影角度条件等の問題がある。

2.2 個人再識別

人物再識別 (Person Re-Identification) とは、複数のカメラビュー、または同じカメラビューのさまざまな角度や場面で特定の人物を照合することを目的とする研究である。この研究は、人物のクロス (複数) カメラ追跡などの多くの重要な応用システムを容易にする。

近年、社会システムへの知的なサービスの充実のため、さまざまな社会情報が収集されている。バス・電車の乗客における移動統計や大型公共施設の混雑度計測など人数カウントにコンピュータビジョンが利用され、特定人物の追跡などにもその利用場面が増えている。これらに対応した個人再識別技術の研究では、次の3種類の画像が利用されている。1) 図 2.8 (a) に示す正面、側面や後ろ姿の全身画像（以下、一般画像と呼ぶ）、2) 図 2.8 (b) に示す俯瞰視カメラで撮影された画像（以下、俯瞰視画像と呼ぶ）、3) 図 2.8 (c) に示す俯瞰視カメラの光軸上の画像（以下、真上からの画像と呼ぶ。俯瞰視画像の特例）である。

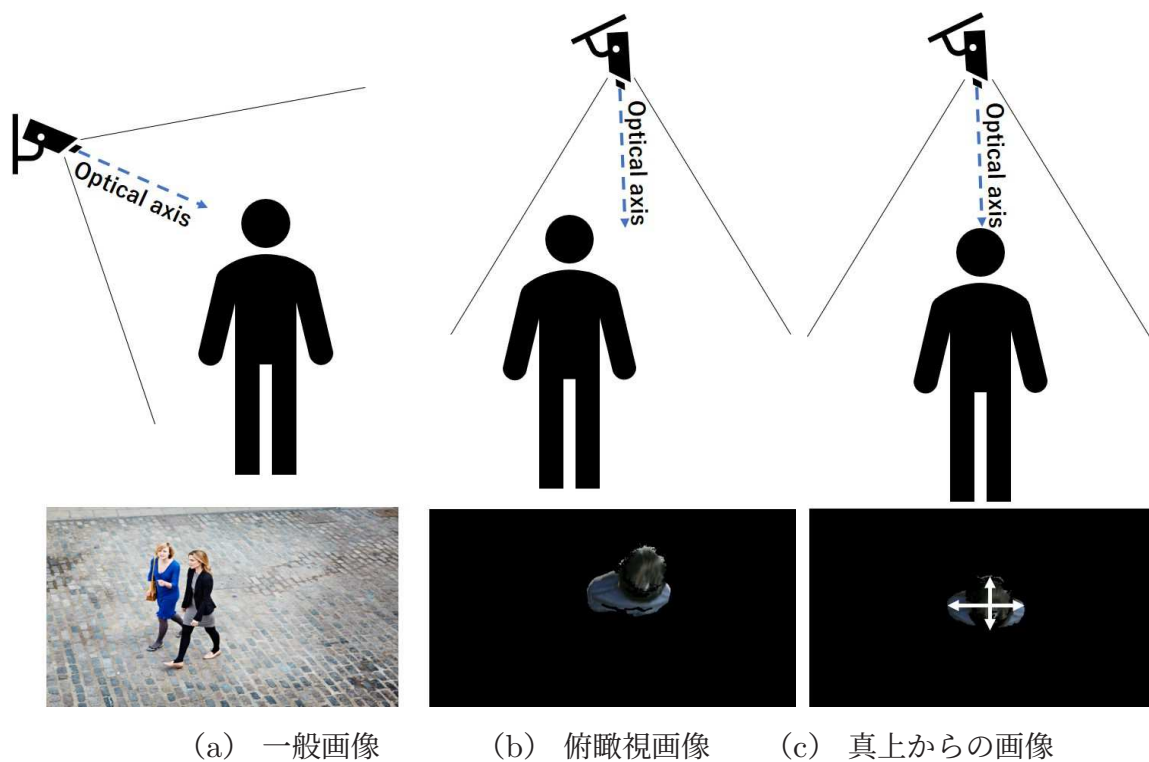


図 2.8 識別用画像の撮影方法

2.2.1 一般画像による個人再識別

一般画像を対象とする個人再識別は研究の主流であり [126]、近年、Deep Learning も画像認識によく利用されている [127]。Zhong ら [128] の提案手法では、既存の CycleGAN 手法を利用して、カメラペアで撮影された多数の人物画像のペアを用いて「姿勢画像転換モデル (style transfer model)」を学習している。学習済みのカメラペアで撮影された元

画像と転換された画像を拡張学習セットとし、Zhong らが提案している個人再識別 CNN モデルを学習で生成し、Pool-5 レイヤーの出力を画像特徴の記述子として、ユークリッド距離で記述された特徴間の類似度評価によって個人識別を行っている。Zhong らの実験 [128] では共通データベースに多数のカメラで撮影された人物の全身が映っているカラー画像を利用している。俯瞰視画像よりも、上下の服、腕、脚を含む人物全身画像は、色情報およびテクスチャー情報を多く含んでいる。

低解像度の一般画像を利用する個人再識別方法も提案されている [129][130]。図 2.9 は低解像度を利用する個人再識別の例である。Bazzani ら [129] は人物画像の解像度の低下、明暗変化、人物の姿勢変化に対応するために、HPE (Histogram Plus Epitome) 特徴記述子を提案し、各個人の連続画像のシーケンスより計算した平均 HSV ヒストグラムをグローバル特徴量とし、人物のテクスチャーと形状および見え (appearance) のエピソード推定結果をローカル特徴量としている。特徴間の類似度をバタチャリア距離で評価している。

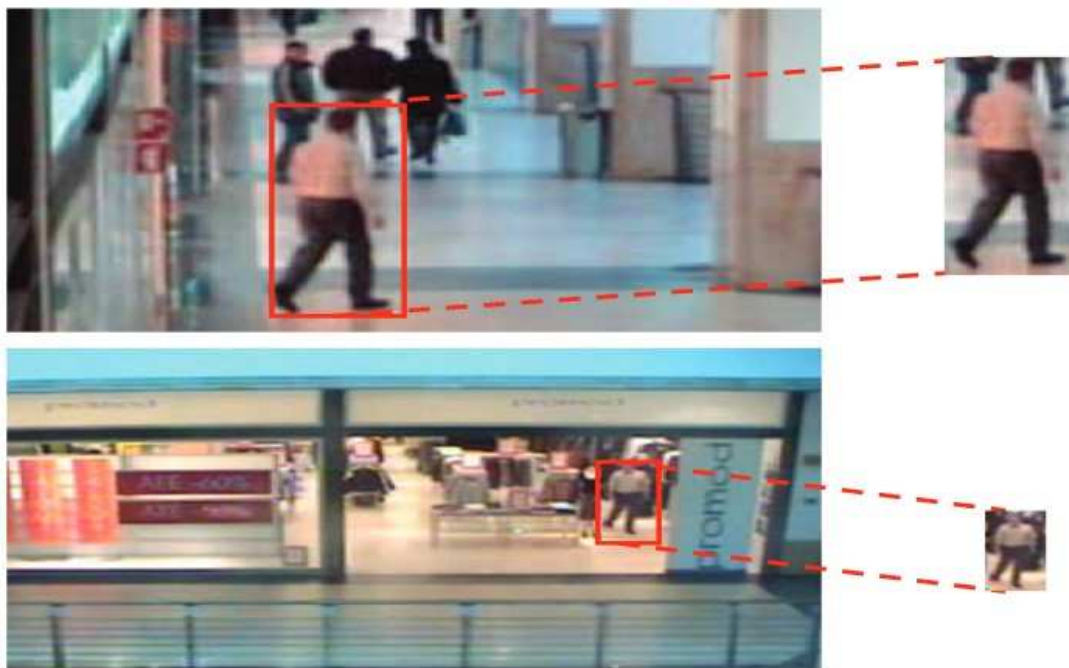


図 2.9 低解像度画像を利用する個人再識別 [130]

Xiang ら [130] は、JUDEA (Joint Multi-scale Discriminant Component Analysis) と呼ばれるマルチスケール判別分析モデルを提案している。異なるスケールの画像をスケールごとに画像セットを生成して、同じ人物の各スケールの画像に対して、8 チャンネルの

色情報 (RGB, YCbCr, HS) と, LBP (Local Binary Pattern) および HOG (Histogram of Oriented Gradients) を特徴量ベクトルとしている. マルチスケール判別距離マトリックス学習モデルを利用して人物間のマッチングを行っている.

全身を映している画像には個人情報も多く含むので, 個人を表現できる特徴量を得やすいが, 許可なく収集・利用することが制限される問題もある. また, モデルの学習, もしくは人物の特徴量を生成するには, 一般的に, 多数の人物画像が必要である.

2.2.2 真上からの画像による個人再識別

俯瞰視画像を対象とする個人再識別はあまり研究されていない. 主に俯瞰 RGB-D カメラで撮影された真上からの画像 (図 2.8 (c) に参照), 特に深度画像を積極的に利用した人数カウント [131] や, 個人再識別のアルゴリズム [132][133] が提案されている.

Liciotti D ら [132] は約 4 メートル高さの天井に俯瞰視 RGB-D カメラを設置し, 真上からのカラー画像から人物の頭部と肩部における HSV 空間内の 2 つのヒストグラムと, 深度画像から人物の頭部から地面までの距離, 肩部から地面までの距離, 頭の面積, 頭囲, 肩周り, 肩幅, 胸の前後の深さ, 合計 9 つのパラメータを個人の特徴として記述している.

Kouno D ら [133] は, 真上からの深度画像 (図 2.10 に参照) を利用して, カメラと頭頂部間の距離, 検出された頭と肩の面積, x 方向の肩幅と y 方向の頭の幅における最大値, 全画面の深度のヒストグラム, 合計 4 つのパラメータで個人を表現する特徴量を記述している. ある特定の人に対して AdaBoost 学習法で識別器を構築し, 得られた識別器を用いて該当人物の個人再識別を行っている. 被験者が増えるたびに全員分の識別器を再構築することが必要である.

Nakatani R ら [134] と中谷ら [135] は本論文と同様に距離を特徴として使用せず, 真上からのカラー画像のみを利用する方法を提案している. Nakatani R らは, 前処理で検出された人物領域に基づいて, x 方向と y 方向の幅をボディサイズとし, X 軸の輝度値が最大になる座標 x と Y 軸の頭部領域の中央となる座標 y から成る座標 (x, y) を仮の人物中心座標とした HOG (Histograms of Oriented Gradients) 法による検出されたエッジ情報から髪型とつむじを抽出し, 頭部領域 (50×50 画素) の輝度値の合計を髪の色とする. ボディサイズ (size), 髪型 (style), つむじ, 髪の色 (color), 合計 4 つのパラメータで個人を表現する特徴量を記述し, AdaBoost 学習法で識別器を構築している. この手法は本研究と最も似ているため, 本論文では, すべての提案手法とこの手法との比較実験を行う.

Liciotti D ら [132], Kouno D ら [133] および Nakatani R ら [134] の手法では, 識別用

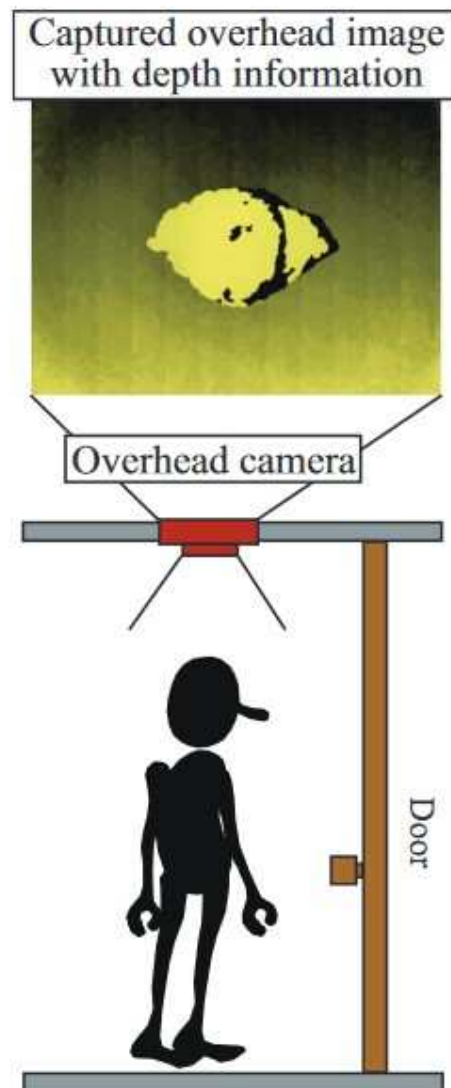


図 2.10 深度画像を撮影する環境 [133]

の特徴量には体格を含む矩形サイズが含まれており、再識別精度を保証するために、撮影時カメラの光軸の真下から被験者が頭と肩の向きを保持しながら通る [132][134] 必要がある。このうち Liciotti D ら [132] と Kouno D ら [133] では身長などの距離情報も特徴量に含めているので、カメラの真下で一時停止して撮影する必要がある。これらの個人再識別手法の再識別の成功率は 91.0%~94.4% であった。

2.2.3 画像の色情報を用いた個人再識別

個人再識別に関する研究では、識別機能向上に有用な色情報が重要な役割を果たしている。

使用するカメラの特性によって、同一人物を再撮影時に色の不一致の問題点があるので、この問題を改善するために、個人再識別の初期では、いくつかの色補正や輝度伝達関数 (BTF) を利用する手法が提案されている [136][137],[138],[139],[140]。特徴量の記述方法の強化による対応する手法も多数提案されている。

Gheissari ら [141] は、安定している前景領域 (人物領域) を検出するために時系列の空間セグメンテーション法を提案している。ローカル領域の場合、HS ヒストグラムと Edgel ヒストグラムが計算される。Edgel ヒストグラムは、主要なローカル境目のオリエンテーションと Edgel 両側の RGB 比を記述する。図 2.11 はこの手法を利用した個人再識別の実験環境である。

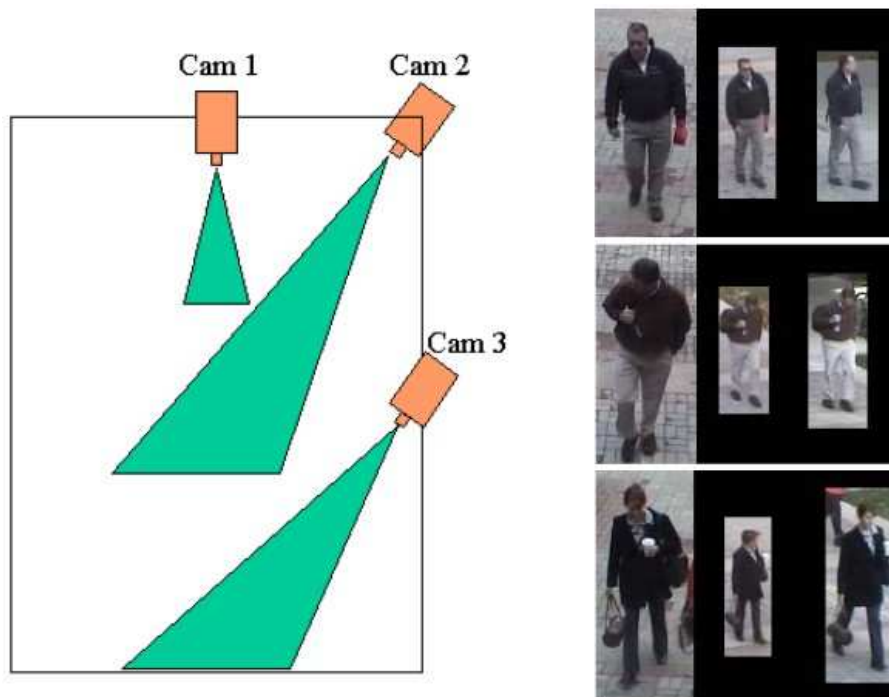


図 2.11 個人再識別の実験環境 [141]

Gray ら [142] は、8 つの色チャンネル (RGB, HS, および YCbCr) と 21 つのテクスチャーフィルターを輝度チャンネルに使用し、人物を複数の水平ストライプに分割している。手法 [143], [144] および [145] は、この [142] と同じ特徴セットを採用している。

Farenzena ら [146] では、人物領域を背景からセグメンテーションし、各ボディパーツに対して対称軸が計算される。画像上の体の構成に基づいて、重み付き色ヒストグラム (WH)、最大安定色領域 (MSCR) および周期的高構造化パッチ (RHSP) も計算される。WH は、対称軸の近くにあるピクセルに大きな重みを割り当て、各パーツの色ヒストグラムを構成する。MSCR 領域では色、面積および重心などの特徴を抽出する。RHSP では、周期的なテクスチャーパッチを取得している。

Mignon ら [147] は RGB, YUV, HSV チャンネルからの特徴ベクトルと、横縞の LBP テクスチャーヒストグラムを特徴量として利用している。

Zhao らの手法 [148][149][150] では、5 ピクセルのステップサイズでサンプリングした各 10×10 のパッチから、32 次元の LAB 色ヒストグラムと 128 次元の SIFT 記述子を抽出する。

Pedagadi ら [151] は PCA による次元数を削減する前に、HSV および YUV 空間から色ヒストグラムとモーメントの抽出を行っている。

Das ら [152] は [129] で検出した人物の輪郭の頭部、体、肢部の HSV ヒストグラムを利用している。

Liu ら [153] は各ローカルパッチの HSV ヒストグラム、勾配ヒストグラム、LBP ヒストグラムを抽出して利用する。

Yang ら [154] は被験者のグローバル特徴を記述するために、色名ベースの色記述子 (SCNCD) を提案している。

Liao ら [155] は、色と SILTP のヒストグラムを含む LOMO 記述子を提案している。

色情報は特徴量として記述される際に、様々な形式で活用できるため、研究の目的に合わせて改善方法も多数研究されている。

第3章

色弁別マップを用いた人物の再識別

本章では，最近傍 (Nearest Neighbor) 分類器 [156] に基づく特徴記述法を提案する．この記述法より，単に人の肩部 (本章では「メイン領域」と呼ぶ) と頭頂部 (本章では「サポート領域」と呼ぶ) だけの色情報を利用して，顔が見えない人物の俯瞰視画像による個人再識別を行う．図 3.1 は本手法による個人再識別の結果の例を示している．画像センターにいる人は，データベースに登録されている特徴量と一致する人物である．この場合，肩の領域が白く表示されているが，これは照合度の高い識別値を示している例である．



(a) 入力画像

(b) 再識別結果 (ハイライト)

図 3.1 提案手法を利用した個人再識別の結果例

3.1 色弁別度 (Color Distinctiveness)

和田はオブジェクト追跡のために色の弁別性を推定する方法を提案した [157]．その論文では，オブジェクトのターゲット色との類似性とノンターゲット色との非類似性はベイ

ズ規則の枠組みに統合できることが示されている。色 c の類似性と非類似性は、それぞれ確率密度関数 (PDF) $p(c|\Omega_t)$ と $p(c|\Omega_n)$ で表すことができると仮定している。事後確率 $P(\Omega_n|c)$ は式 (3.1)

$$P(\Omega_t|c) = \frac{P(\Omega_t)p(c|\Omega_t)}{P(\Omega_t)p(c|\Omega_t) + P(\Omega_n)p(c|\Omega_n)} \quad (3.1)$$

に表される。ここで、 $P(\Omega_t)$ と $P(\Omega_n)$ はそれぞれターゲット色とノンターゲット色の事前確率である。 $P(\Omega_t) = P(\Omega_n) = 0.5$ と仮定すると、式 (3.1) は式 (3.2) に変換される。

$$P(\Omega_t|c) = \frac{p(c|\Omega_t)}{p(c|\Omega_t) + p(c|\Omega_n)} \quad (3.2)$$

$p(c|\Omega_t)$ と $p(c|\Omega_n)$ の定義はそれぞれ式 (3.3) と式 (3.4) の通りとなる。

$$p(c|\Omega_t) = \frac{1}{d(c, NN(c, S_t))} \quad (3.3)$$

$$p(c|\Omega_n) = \frac{1}{d(c, NN(c, S_n))} \quad (3.4)$$

ここで、 S_t と S_n はそれぞれターゲットの色セットとノンターゲットの色セットであり、 $NN(c, S)$ は色セット S の中に最も色 c に近傍する色である。 $d(x, y)$ は x と y の距離である。式 (3.3) と式 (3.4) を式 (3.2) に代入することより、次の式 (3.5) となり、色の弁別度が計算できる。この値も Ω_t に属する色 c の事後確率である。

$$D(c) = P(\Omega_t|c) = \frac{d(c, NN(c, S_n))}{d(c, NN(c, S_t)) + d(c, NN(c, S_n))} \quad (3.5)$$

学習ステージでは、式 (3.5) を用いて、3次元 LUT (Look Up Table) 中のすべての要素は、最も近いターゲット領域色と最も近いノンターゲット領域色から始まる色空間での距離変換を行うことによって計算される。追跡ステージでは、LUT に参照することで、ビデオレートで未知の入力色 c をターゲットの色か、ノンターゲットの色に分類する。

3.2 CDF 法による顔情報を利用しない個人再識別

3.2.1 CDF:色弁別特徴

ほとんどの人は、日常生活で頻繁に髪の色や服を変えることはない。これを仮定として、人の肩や頭の色は短時間内で変化しないと考える。この場合、人物の肩部を色弁別度

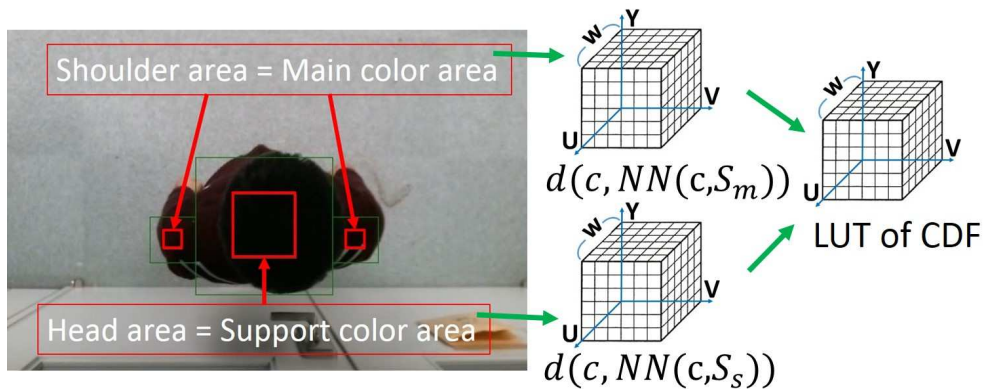


図 3.3 人物のメイン領域とサポート領域および CDF の生成

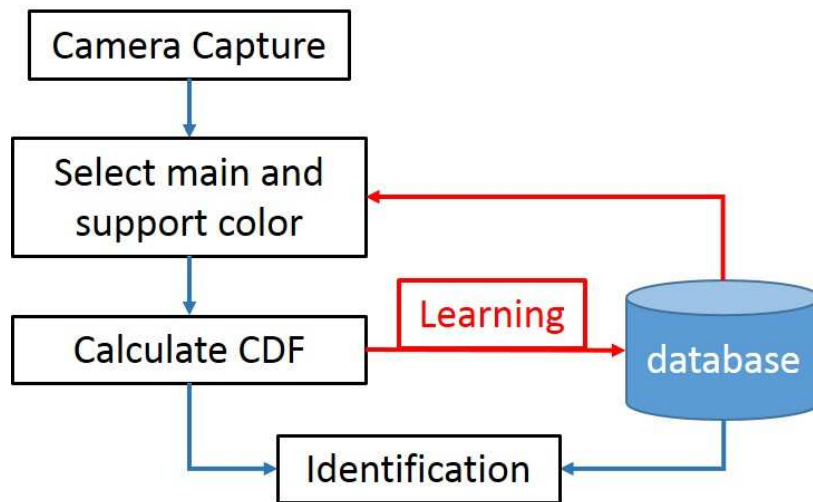


図 3.4 提案手法を用いた個人再識別の処理流れ

$$CDF_{(k)} = P(\Omega_{m_k}|c) = \frac{d(c, NN(c, S_{s_k}))}{d(c, NN(c, S_{m_k})) + d(c, NN(c, S_{s_k}))} \quad (3.6)$$

ここで、 S_{m_k} と S_{s_k} はそれぞれ k 番目の人物のメイン領域の色セットとサポート領域の色セットであり、 $NN(c, S_{s_k})$ と $NN(c, S_{m_k})$ はそれぞれ色セット S_{s_k} と S_{m_k} の中に最も色 c に近傍する色である。この色は色 c との距離をそれぞれ $d(c, NN(c, S_{m_k}))$ と $d(c, NN(c, S_{s_k}))$ で計算する。 $P(\Omega_{m_k}|c)$ は色 c の類似性と非類似性より計算した事後確率である。

この後、該当人物の $CDF_{(k)}$ をデータベースに保存し、 $CDF_{i(k)}$ とする（図 3.3 の右側）。

再識別のステージでは，人物が部屋を出て画像上に現れるとき，式 (3.7) より未知人物の色弁別特徴量 CDF_l を計算する．

$$CDF_l = P(\Omega_{m_l}|c) = \frac{d(c, NN(c, S_{s_l}))}{d(c, NN(c, S_{m_l})) + d(c, NN(c, S_{s_l}))} \quad (3.7)$$

ここで， S_{m_l} と S_{s_l} はそれぞれ未知人物のメイン領域の色セットとサポート領域の色セットであり， $NN(c, S_{s_l})$ と $NN(c, S_{m_l})$ はそれぞれ色セット S_{s_l} と S_{m_l} の中に最も色 c に近傍する色である．この色は色 c との距離をそれぞれ $d(c, NN(c, S_{m_l}))$ と $d(c, NN(c, S_{s_l}))$ で計算する． $P(\Omega_{m_l}|c)$ は色 c の類似性と非類似性より計算した事後確率である．

バタチャリア距離による類似度評価

2つの分布間のオーバーラップの量を計算することにより，2つの分布の類似度を評価するバタチャリア距離を取得できる．バタチャリア距離が長い，つまり，分布のオーバーラップが大きいほど，二つの分布はお互いに似ていることを示している（図 3.5 に参照）．

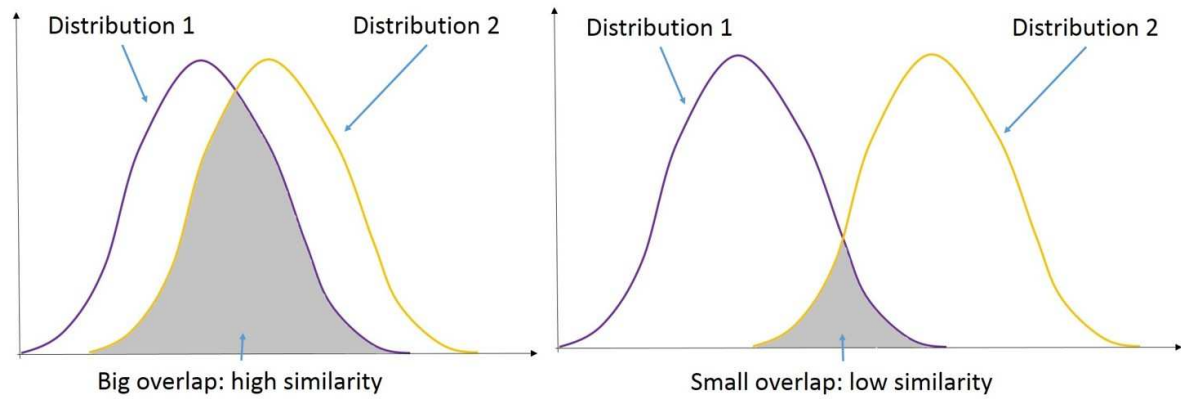


図 3.5 2つの分布間のオーバーラップのイメージ

提案手法を用いた個人再識別は，未知人物の特徴量 CDF_l とデータベース内の各人物の $CDF_{i(k)}$ との比較により行う． CDF_l と $CDF_{i(k)}$ はともに LUT (Look Up Table) で構築した辺長が w である 3次元色空間の離散確率分布であるため，本章では，未知人物 CDF_l とすべての人物の $CDF_{i(k)}$ との類似度はバタチャリア距離を用いて式 (3.8) で評価する．LUT の辺長 w (図 3.3 の右に参照) は 0 から 2^n までに調整できる．ここで， $n = 8, 7, 6, 5$ である．

$$D(CDF_{i(k)}, CDF_l) = \sum_{y,u,v \in \{0, \dots, w\}} \sqrt{CDF_{i(k)}(y, u, v) CDF_l(y, u, v)} \quad (3.8)$$

ここで、 y, u, v はそれぞれ3次元色空間内の座標値である。すべての CDF ペアの類似度を計算した後、式 (3.9) 最も高い類似度を持っているデータベースの CDF の ID 番号 k を識別成功人物の番号と見なす。

$$ID = \arg \max_{k \in \{1, 2, \dots, N\}} D(CDF_{i(k)}, CDF_l) \quad (3.9)$$

ハミング距離による類似度評価

ハミング距離を用いて、2つの分布間にどれだけの位置が異なるかを計算することができる。提案手法では、 $CDF_{i(k)}$ と CDF_l とのハミング距離 d_h は式 (3.10) で計算される。

$$d_h = \sum_{y, u, v \in \{0, \dots, w\}} \delta(CDF_{i(k)}(y, u, v), CDF_l(y, u, v)) \quad (3.10)$$

ここで、 y, u, v はそれぞれ3次元色空間内の座標値であり、値域は $y, u, v \in \{0, \dots, w\}$ である。 $\delta(CDF_{i(k)}(y, u, v), CDF_l(y, u, v))$ は $CDF_{i(k)}(y, u, v) = CDF_l(y, u, v)$ の場合が0、 $CDF_{i(k)}(y, u, v) \neq CDF_l(y, u, v)$ の場合が1を出力する関数である。

2つの CDF のハミング距離は0に近いほど、互いにより類似していることを意味している。

本章では、ハミング距離による2つの CDF の類似度の値域を0から1までに正規化した上、非類似・類似の基準をそれぞれ0と1に逆転させている。これで、類似度 (D_h) は式 (3.11) で定義される。

$$D_h(CDF_{i(k)}, CDF_l) = 1 - \frac{1}{w^3} d_h \quad (3.11)$$

ここで、 w^3 は LUT で構築した3次元空間内の元素の総数である。

未知人物の CDF_l とデータベースに登録したすべての CDF_i とのペアの類似度を計算した後、最も高い類似度を持っているデータベースの $CDF_{i(k)}$ の ID 番号 k を式 (3.9) で計算し、識別成功人物の番号と見なす。

3.3 実験

提案手法の有効性を検証するために、俯瞰視カメラで撮影した人物の俯瞰視画像を利用して個人再識別実験を行った。Intel Core i7 4770 CPU と 8GB メモリを搭載した Windows 7 の PC を使用している。カメラは一般市販 Web カメラ Logicool c615 であ

表 3.1 w の変化による CDF 特徴量生成時間, 特徴量間のマッチング時間, 再識別成功率

w	128		64		32		other	
距離	バタチャリア	ハミング	バタチャリア	ハミング	バタチャリア	ハミング	バタチャリア	ハミング
CDF 生成時間 (ms/人)	23.5		2.9		< 1.0		< 1.0	
照合時間 (ms/人)	421.0	< 1.0	47.0	< 1.0	16.0	< 1.0	< 1.0	< 1.0
再識別の成功率	同一レベル				50% 以下			

る。今回の試作システムで使用するハードウェアの制限により、LUT の w を 256 に設定する場合には動作できなかったため、 w を 128, 64, 32 のそれぞれに設定し、一人分の CDF のバタチャリア距離とハミング距離での生成時間や、同じ条件で 2 つの CDF 間の照合時間も記録した (表 3.1 に参照)。本章では、システムの処理スピードを重視するので、生成時間と照合時間が両方とも速い方、 $w = 64$ の下で実験を行った。

3.3.1 個人再識別実験

個人再識別実験には、被験者が 9 人参加した。入室とき、各人物の CDF を計算してデータベースに登録する。ある人物が退室する際に該当人物の CDF を計算し、データベースに登録された各人物の CDF との類似度を計算する (図 3.6 に参照)。バタチャリア距離とハミング距離を使用した個人再識別実験の結果の混同行列はそれぞれ表 3.2 と表 3.3 に示す。各行において最も高い類似度は赤色でマークしている。

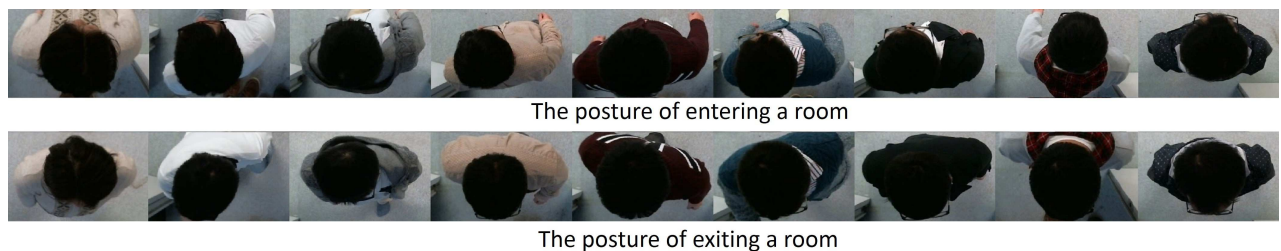


図 3.6 9 人の被験者の入室画像 (上の行) と退室画像 (下の行)

表 3.2 と表 3.3 から、CDF 法は顔情報を利用しなくても個人再識別が可能であることが示された。

表 3.2 バタチャリア距離を用いた個人再識別実験の結果 (%)

PersonNo	1	2	3	4	5	6	7	8	9
1	99.99	99.95	99.64	99.89	99.32	99.81	99.36	99.72	99.76
2	99.96	99.98	99.54	99.80	99.20	99.75	99.25	99.64	99.69
3	99.67	99.43	100.00	99.84	99.85	99.92	99.90	99.93	99.97
4	99.87	99.68	99.89	99.99	99.69	99.92	99.71	99.91	99.93
5	99.34	99.08	99.86	99.60	100.00	99.75	99.98	99.81	99.81
6	99.81	99.71	99.84	99.81	99.69	99.96	99.75	99.89	99.94
7	99.35	99.10	99.88	99.59	99.98	99.77	100.00	99.80	99.83
8	99.70	99.49	99.94	99.86	99.82	99.92	99.83	99.99	99.95
9	99.72	99.51	99.98	99.85	99.82	99.94	99.88	99.94	99.99

表 3.3 ハミング距離を用いた個人再識別実験の結果 (w=64)

PersonNo	1	2	3	4	5	6	7	8	9
1	0.52	0.00	0.16	0.00	0.00	0.27	0.00	0.43	0.18
2	0.00	0.62	0.00	0.00	0.00	0.01	0.00	0.13	0.00
3	0.12	0.00	0.76	0.03	0.00	0.49	0.06	0.35	0.48
4	0.00	0.00	0.07	0.85	0.00	0.08	0.00	0.10	0.02
5	0.01	0.01	0.01	0.01	0.84	0.01	0.01	0.50	0.01
6	0.00	0.01	0.01	0.00	0.00	0.80	0.02	0.01	0.04
7	0.01	0.01	0.16	0.01	0.01	0.37	0.56	0.31	0.37
8	0.17	0.15	0.23	0.03	0.07	0.24	0.01	0.75	0.24
9	0.07	0.05	0.53	0.04	0.04	0.72	0.38	0.47	0.79

3.3.2 比較実験

3.3.1 項の画像セットを用いて、実験用検証ソフト (Weka[158]) で Nakatani R らの手法との比較実験を行った。提案手法の場合、各被験者の入室と退室画像に対して、それぞれに肩部と頭頂部を手動で 10 回選択して、合計 180 回分 (9 人 × (入室 10 回 + 退室 10 回)) の CDF 特徴量セットを獲得した。Nakatani R らの手法の特徴量の記述、Weka での分類器 (C4.5 アルゴリズム) および他の設定はすべて Nakatani R らの論文 [134] と同様

として、10 分割交差検証で実験を行った。また、本実験では、Nakatani R らの手法において、再認識の成功率が一番高い *size + style + color* の特徴量の組み合わせを利用する。

提案手法と Nakatani R らの手法の再認識の成功率はそれぞれ 97.22% と 72.22% であり、両手法の成功回数・失敗回数を表す混同行列は図 3.7 と図 3.8 である。この比較実験より、提案手法は従来手法より再識別率が高いことを確認できた。

```

=== Confusion Matrix ===

  a  b  c  d  e  f  g  h  i  <-- classified as
19  0  0  0  0  0  1  0  0 | a = p1
 0 19  0  0  0  0  1  0  0 | b = p2
 0  0 20  0  0  0  0  0  0 | c = p3
 0  0  1 19  0  0  0  0  0 | d = p4
 0  0  0  0 20  0  0  0  0 | e = p5
 0  1  0  0  0 19  0  0  0 | f = p6
 0  0  0  0  0  0 20  0  0 | g = p7
 0  0  0  0  1  0  0 19  0 | h = p8
 0  0  0  0  0  0  0  0 20 | i = p9

  Correctly Classified Instances   97.2222 %
  Incorrectly Classified Instances  2.7778 %

```

図 3.7 提案手法の結果

```

=== Confusion Matrix ===

  a  b  c  d  e  f  g  h  i  <-- classified as
14  0  2  0  0  3  0  1  0 | a = p1
 2 14  1  2  0  1  0  0  0 | b = p2
 0  1 15  1  0  0  3  0  0 | c = p3
 1  3  0 15  0  1  0  0  0 | d = p4
 0  0  0  0 15  3  2  0  0 | e = p5
 3  2  0  0  0 11  0  1  3 | f = p6
 0  0  5  0  1  2 12  0  0 | g = p7
 0  1  1  0  1  0  0 17  0 | h = p8
 0  1  1  0  0  0  1  0 17 | i = p9

  Correctly Classified Instances   72.2222 %
  Incorrectly Classified Instances 27.7778 %

```

図 3.8 Nakatani R ら手法の結果

3.4 まとめ

本章では、人が短時間内に髪の色と服の色の変更はないことを仮定し、人物の俯瞰視画像から人物の特徴量を記述する CDF 記述子を提案した。本手法を用いることで、顔画像の個人情報を使わず個人再識別を実現できた。スーパーマーケット、企業の入退場ゲート、遊園地など、出入り口のある公共の場所で人の通行量のカウントシステムとして利用できる。個人再識別実験より、提案手法が個人再識別に使用できることを確認した。また、物体の再識別実験より、提案手法は物体再識別にも利用できることを確認した（付録 A に参照）。一方、2 人が同じ色の異なる柄の服を着ている場合、正確に判別できない課題が残っている。

第 4 章

RGB 色空間内の色情報の拡張記述による個人識別

4.1 CDF 法の問題点

3 章では CDF 法について述べた。CDF 特徴量の生成には俯瞰視画像の人物領域の色情報を利用するので、人物領域内の色の種類数が多いほど、より弁別性の高い CDF が得られる。しかし、人物領域内の色の種類数が少ないケースはよくある。この原因で、人物領域の色情報が異なっても、CDF 法の特徴 LUT マトリックスでは、同じ要素がほとんどなので、評価とする類似度も大きな差はなかった。また、CDF 法を利用する際に、メインとサポート二つの領域をそれぞれに選ぶ必要があり、応用システムの稼働には余計な操作時間がかかる。したがって、色の種類数が少ない場合であっても、類似度でその僅かな差をしっかりと反映できると共に、前処理も軽減できる手法が求められる。

4.2 基本アイデアと処理の流れ

本章では、3 章の研究の前提条件と同様に、短時間内に人物の服や髪の色は変わらないと仮定し、人物の俯瞰視画像を利用して個人再識別を行う。人物の俯瞰視画像には、人物の特徴と関係があるのは髪と肩部（以降はターゲット領域と呼ぶ）しかない。ターゲット領域の中に、肉眼では、はっきり認識できない僅かな差の色情報も含んでいる。その部分の情報も活用できれば、人の識別能力を超える個人識別に有用な特徴と考えられることから、色ベースの新たな特徴記述方法を提案する。この手法を水滴レンダリングボックス (**Water-Drop Render Box**) と呼ぶ（以下は WDRB と略称する）。WDRB の処理は水

滴が紙を濡らす現象を模倣している。即ち、水滴（ピクセルの色）が紙を垂らすと（色を登録する）、紙上の水滴の落下のところは濡れる、紙上の同じ場所に複数の水滴が落下すると、濡れた面積（色の基本範囲）がさらに拡大される（ヒストグラムに基づいて色がさらに強調される）という現象を利用して人物の特徴を記述する。WDRB法を簡単に理解するために、特徴の記述を3つの段階に分けて説明する。(1)色登録、(2)色登録マップから距離マップへの変換、(3)登録した色の強調。

WDRB法による個人再識別の処理流れを図4.1に示す。まず、俯瞰視カメラの下を通った人物の単眼俯瞰視画像からターゲット領域を選択する。そして、図4.1の緑枠に示す、4.3.1項に述べる色の登録と各色のヒストグラムを求める処理と、4.3.2項に述べる距離マップの変換や色のヒストグラムによる色の強調処理があり、これらの処理で該当人物の4次元WDRB特徴を記述する。

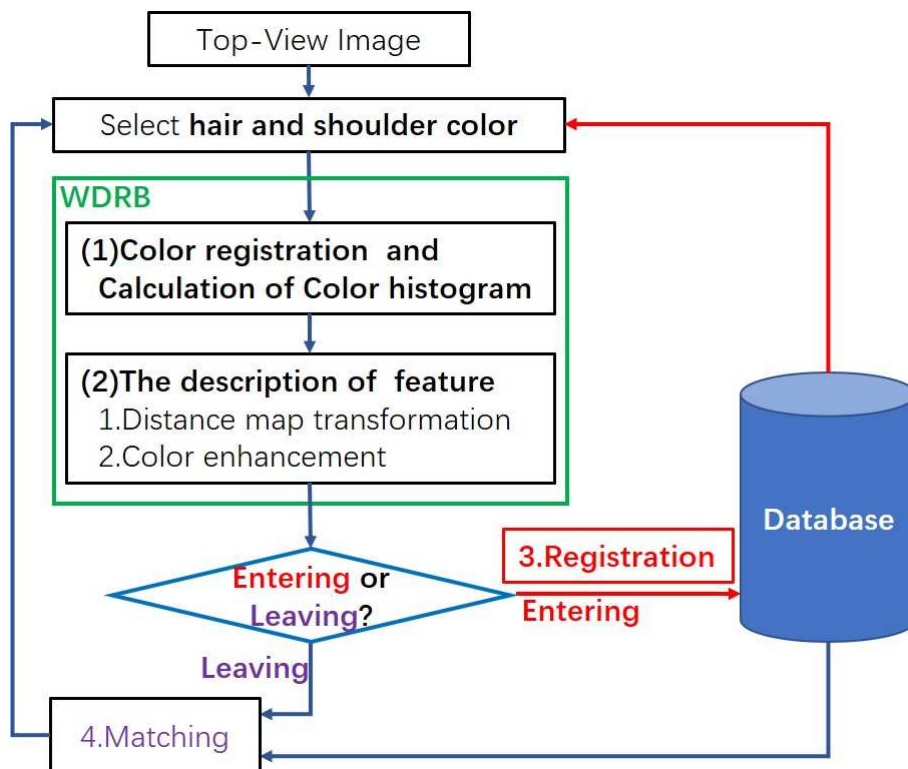


図4.1 WDRB法による個人再識別の処理流れ

4.3 WDRB 法:水滴レンダリングボックス (Water-Drop Render Box)

WDRB 法を利用する個人再識別では，再識別精度と，速度・メモリ容量のトレードオフを両立するために， $255 \times 255 \times 255$ の 3 次元色空間を最大の特徴記述空間として，それを分割して処理単位を小さくする．分割後の単位 ($2^n \times 2^n \times 2^n$; $n(= 0, 1, \dots, 7)$) を s で表記し，「セル」と呼ぶ．本章では，提案手法内のすべての処理はセル単位で行う．分割後の 3 次元色空間内のセルの個数を I で表記する．

4.3.1 色の登録と色のヒストグラム

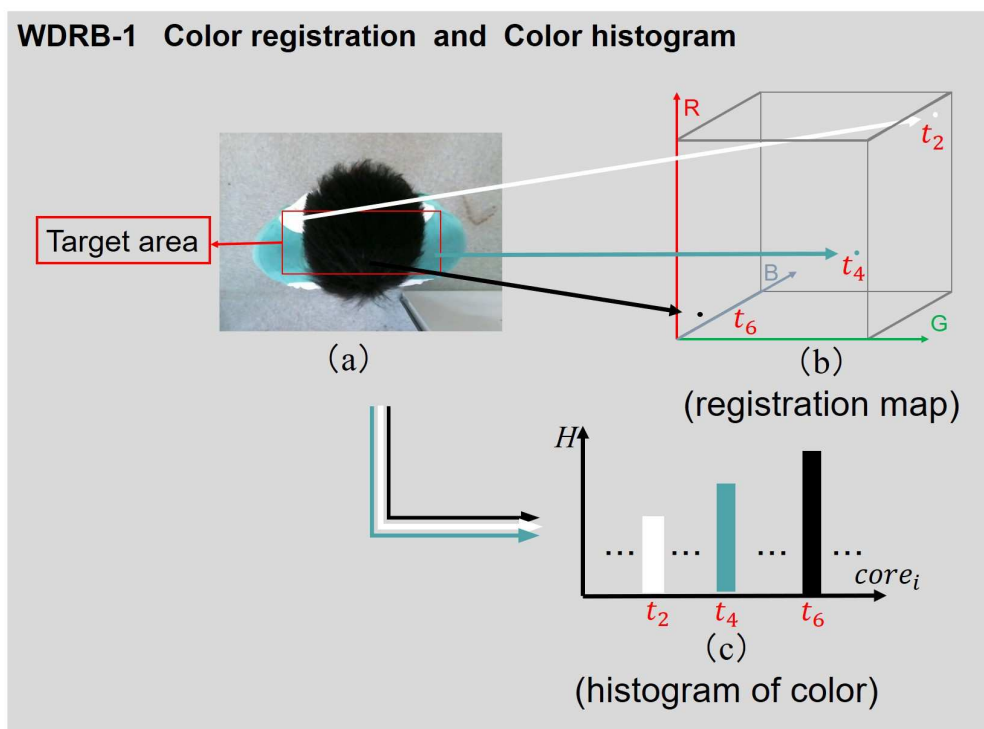


図 4.2 色の登録と色のヒストグラム

色の登録と色のヒストグラムは以下の 4 ステップで求める．

Step 1: n を決定して，3 次元色空間を分割し，セル単位を定め，各セルの座標値を (s) で初期化する．

Step 2: 人物の俯瞰視画像からターゲット領域（髪や肩部位を含む赤色矩形）を選定する

(図 4.2 中の (a) 参照). なお, ターゲット領域の選択処理は 4.4.1 項に述べる.

Step 3: RGB 情報を, RGB 色空間に登録し, 登録された色セルの座標値を (\mathbf{t}) とする.

Step 4: 全画素の色を登録した後, 色登録マップ (図 4.2 の (b)) が得られる. 各画素の色を登録すると同時に, 色 ($\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_s$) に該当するセル毎に色のヒストグラムを計算する. 図 4.2 の (c) は人物の俯瞰視画像から計算した色のヒストグラムのイメージである. 横軸は RGB 空間内の色 (\mathbf{t}_i) を表し, 縦軸は色のヒストグラム (H) を示す.

4.3.2 WDRB 特徴の記述

色登録マップから距離マップへの変換

同じ人物の俯瞰視画像から得られたターゲット領域内の髪や肩部の服の色であっても, 環境照明の変化によって, 色登録マップ内に登録された個々のセルの位置が変化する可能性がある. しかし, 光源による物の演色性が変わっても, 登録されたセル間の相対位置は比較的变化が少なく, その情報は該当人物における色特徴情報を含むと考えられるので, 各個人のセル間の相対位置情報を活用するために, 色登録マップから該当人物の距離マップへのマップ変換を行う.

色登録マップとマップ変換後の距離マップ ($\mathbf{D}_{(r,g,b)}$) はそれぞれ 3 次元のデータであるが, 分かりやすく説明するために, 2 次元への投影をそれぞれ図 4.3 の (a) と (b) に示す. 距離マップへ変換する際に, 色登録マップにある各セル ($\mathbf{s}_i, i \in \{1, \dots, I\}$) を一個ずつ中心とし, m 個の登録色のセル ($\mathbf{t}_j, j \in \{1, \dots, m\}$) までのユークリッド距離を計算し, 一番近い距離を式 (4.1) で求め, マップ変換距離 $D_i, i \in \{1, \dots, I\}$ に保存する.

$$D_i = \arg \min_{i \in \{1, \dots, I\} j \in \{1, \dots, m\}} \|\mathbf{s}_i - \mathbf{t}_j\|_{L_2} \quad (4.1)$$

登録した色の強調

WDRB 法では, 登録した各色の強調は RGB の 3 次元空間内で立方体の形式で表現する. まず, 距離マップと同じサイズの強調マップ ($\mathbf{D}_{\mathbf{e}(r,g,b)}$) を用意した上, 各色登録マップに登録した色における基本範囲及び色のヒストグラムより色の強調範囲を求める. 次に, 各登録した色のセルを中心とし, 辺長が各色の強調範囲 $\Omega_{\mathbf{t}}$ である立方体を作る. j 番目の色における色の強調範囲 $\Omega_{\mathbf{t}_j}$ は式 (4.2) で定義する.

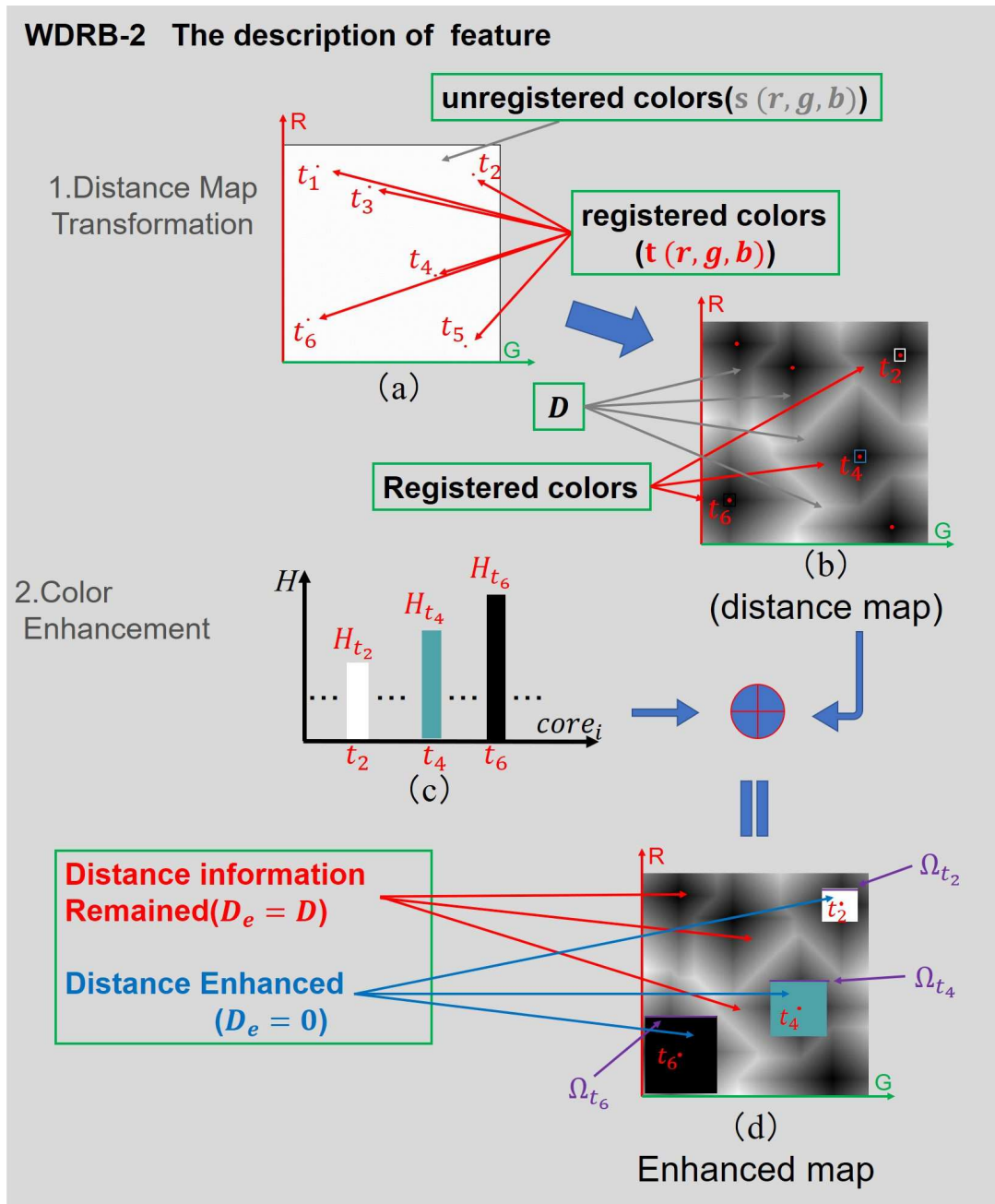


図 4.3 距離マップの変換および登録した色の強調

$$\Omega_{t_j} = \alpha \times \left(1 + \frac{H_{t_j}}{\sum_{k=1}^m H_{t_k}} \right) \quad (4.2)$$

ここで、 Ω_t と α はそれぞれ色の強調範囲と基本範囲である。 H_{t_j} は j 番目のセルにおける色のヒストグラム（図 4.3 の (c) 参照）であり、 m はターゲット領域に含まれる色の総数である。式 (4.2) の右側の第 2 項は j 番目のセルにおける色の割合である。この割合の利用によって、図 4.4 が示すような色が同じ、模様が異なる画像間の識別は可能になる。図 4.4 を利用した第 2 項の有効性実験は 4.4.2 項で示す。

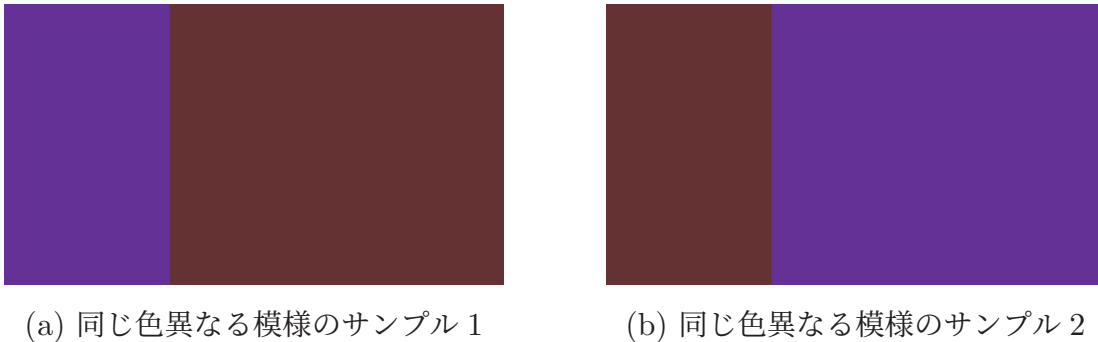


図 4.4 同じ色・違う模様の画像例

また、基本範囲 α は CCD・CMOS などのセンサーの特性や照明変換によって観測誤差の許容範囲と関係する定数である。この観測誤差の発生は人間の視覚の特異性に相当する。人間の視覚の特異性とは、異なる色に対して人間の目による感受性が異なることである。たくさんの CCD・CMOS センサーを収集して、各センサーの観測誤差の計測・統計は困難であるため、本研究では、人間の視覚の特異性を利用して基本範囲 α の値を決定する。

人間が異なる色に対して感受性の差がどれだけあるかを調査するために、27 枚のグラデーションカラー画像を利用して（図 4.5 に参照）予備実験を行った。

各グラデーションカラー画像は 0 番から 8 番までの 9 つのブロックに分割される。0 番ブロックの RGB 値を基準値として設定し、1 番から 8 番までのブロックの RGB 値は、R, G, B 中の任意の一つチャンネル値を等差値で 1 つずつ減らしていく。図 4.6 は G チャンネル値を調整したグラデーションカラー画像例である。この予備実験では、チャンネル値の等差を -5 とした。12 名の被験者で、各被験者が各グラデーションカラー画像に

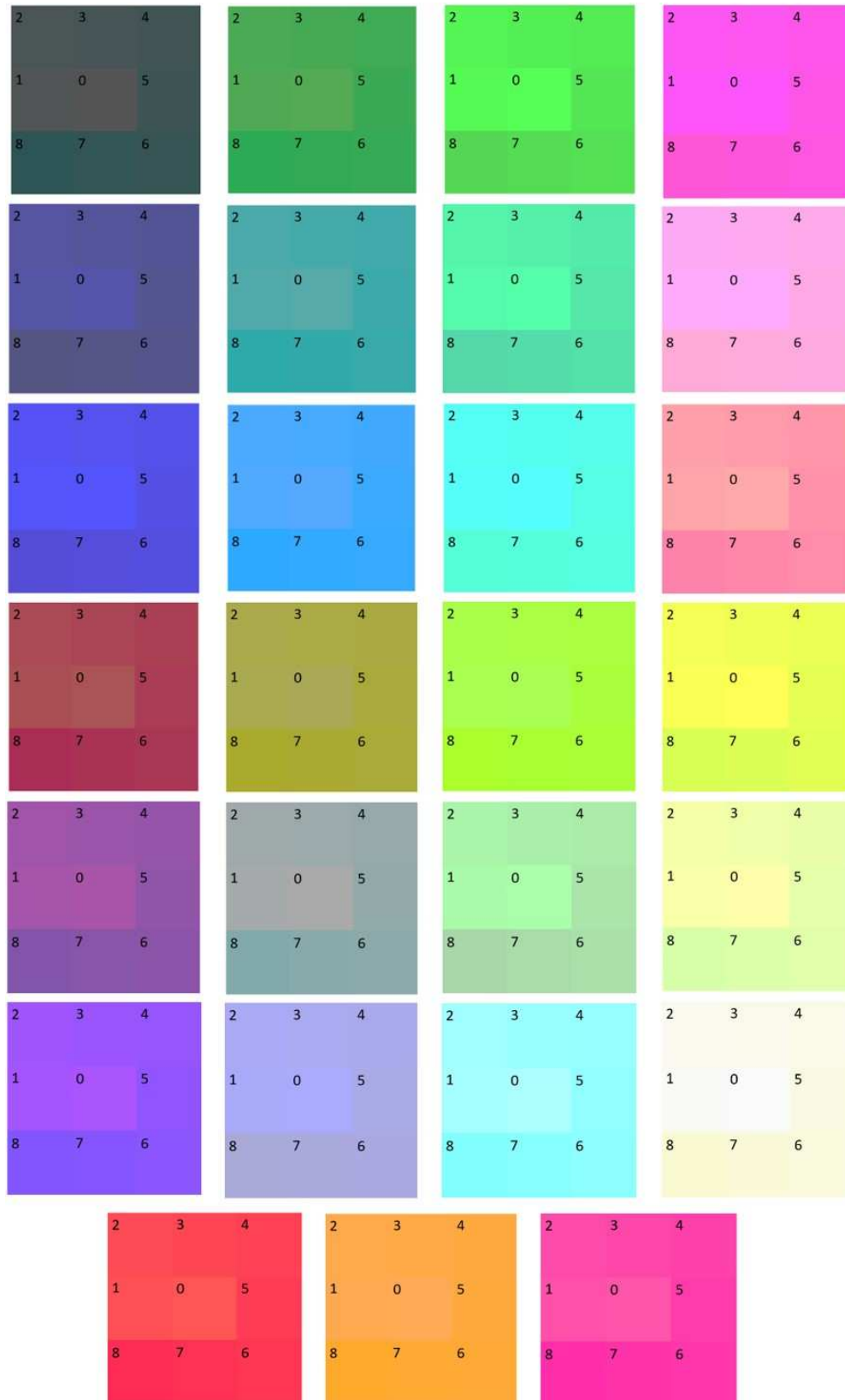


図 4.5 27 枚のグラデーションカラー画像

2	R:255 G:75 B:85	3	R:255 G:70 B:85	4	R:255 G:65 B:85
1	R:255 G:80 B:85	0	R:255 G:85 B:85	5	R:255 G:60 B:85
8	R:255 G:45 B:85	7	R:255 G:50 B:85	6	R:255 G:55 B:85

図 4.6 グラデーションカラー画像のチャンネル調整例

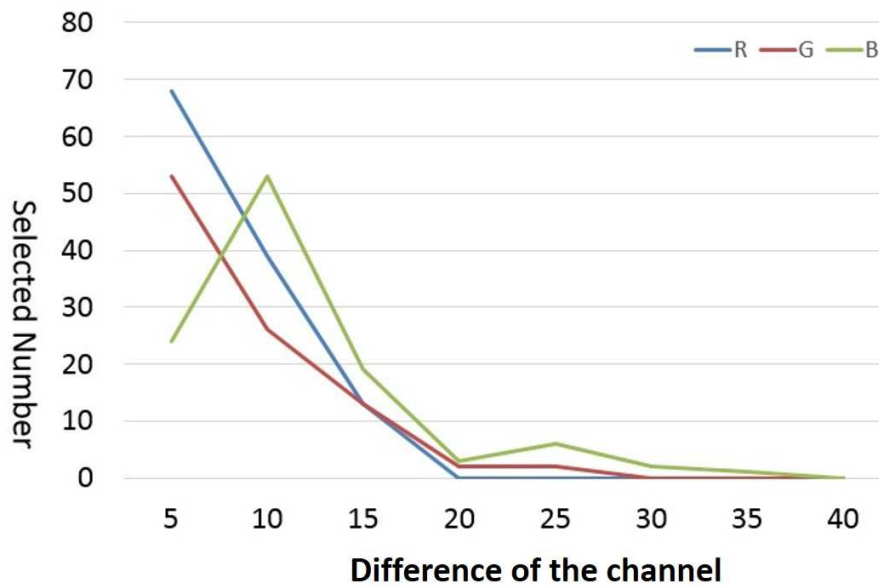


図 4.7 色に対する人の感受性のアンケート結果

対して、1番から8番までのブロックを順番に0番のブロックと比較し、色の違いを意識できる最初のブロック番号を記録する。すべての被験者の記録データを収集・統計して、チャンネルごとに結果を纏めった。図4.7はこの予備実験の結果グラフである。横軸はグラデーションカラー画像の1番から8番までブロックと0番のチャンネル値の差である、縦軸は各ブロックが選択された回数である。この結果より、チャンネルの差値を5か10に設定する場合、ほとんどの被験者が色の違いに気づき始めたことは分かった。この差値を利用し、基本範囲 α の値を10に設定する。つまり、強調マップに対して、式(4.2)で決めた Ω_{t_j} に基づいて j 番目のセルの強調範囲を辺長 Ω_{t_j} の立方体へ拡大する(図4.3の(d)の白色・緑色・黒色の立方体)。この後、アルゴリズム1の処理によって、全ての立方体内の各セルの距離値 D_e を0と設定し、立方体外の各セル距離値 D_e に距離マップの対応するセルの距離値 D を付与する。

Algorithm 1 各登録色の強調範囲(立方体)内の各セルの距離値の設定

```

1: for  $i = 1$  to  $I$  do
2:   for  $j = 1$  to  $m$  do
3:      $d(s_i, t_j) = \sqrt{(r_{s_i} - r_{t_j})^2 + (g_{s_i} - g_{t_j})^2 + (b_{s_i} - b_{t_j})^2}$ 
4:     if  $d(s_i, t_j) < \Omega_{t_j}/2$  then
5:        $D_{e(r_{s_i}, g_{s_i}, b_{s_i})} = 0$ 
6:     else
7:        $D_{e(r_{s_i}, g_{s_i}, b_{s_i})} = D_{(r_{s_i}, g_{s_i}, b_{s_i})}$ 
8:     end if
9:   end for
10: end for

```

ここで、 $r_{s_i}, g_{s_i}, b_{s_i}$ と、 $r_{t_i}, g_{t_i}, b_{t_i}$ はそれぞれ i 番目のセルと j 番目登録色のセルの RGB 色空間内の座標値である。 $d(s_i, t_j)$ は i 番目のセルと j 番目登録色のセルの RGB 色空間内のユークリッド距離である。 $D_{(r_{s_i}, g_{s_i}, b_{s_i})}$ は距離マップ内の i 番目のセルに記録する距離値である。 $D_{e(r_{s_i}, g_{s_i}, b_{s_i})}$ は強調マップ内の i 番目のセルに記録する距離値である。

この処理で計算した強調マップ ($\mathbf{D}_{e(r,g,b)}$) をある個人を表現する特徴量 $\mathbf{WDRB} = \mathbf{D}_{e(r,g,b)}$ とする。

4.3.3 WDRB を用いた個人再識別

提案手法では、人物のプライバシーを利用しないので、3章と同じ俯瞰視システムを利用して歩行者の俯瞰視画像を撮っている（図 3.2 に参照）。個人再識別を行う時、時系列上一度カメラで撮影された人物の俯瞰視画像のターゲット領域から人物の特徴行列 $\mathbf{WDRB}_d(d : \text{database})$ を求め、 ID 番号と共にデータベースに登録する。再びカメラに撮影された際、該当人物の俯瞰視画像のターゲット領域からの特徴行列 $\mathbf{WDRB}_u(u : \text{unknown})$ を求める。 \mathbf{WDRB} 個人同士の \mathbf{WDRB} 特徴量間の類似性は、両分布間のバタチャリア距離を利用して評価する。再び出現した人物の \mathbf{WDRB}_u と、データベースに登録しているすべての人物の \mathbf{WDRB}_d をそれぞれ用いて、式 (4.3) でバタチャリア距離 D_B を計算する。

$$D_B(\mathbf{WDRB}_d, \mathbf{WDRB}_u) = \sum_{r,g,b \in \{0, \dots, 2^{(7-n)} - 1\}} \sqrt{\mathbf{D}_{ed(r,g,b)} \mathbf{D}_{eu(r,g,b)}} \quad (4.3)$$

最終的に、式 (4.4) より一番大きい類似度を持つデータベースの ID 番号を選び出し、再び出現した人物の ID とする。

$$ID = \arg \max_{i \in \{1, 2, \dots, n\}} D_B(\mathbf{WDRB}_{di}, \mathbf{WDRB}_u) \quad (4.4)$$

ここで、 i はデータベースに登録した特徴量の番号である。 n はデータベースに登録した人物の \mathbf{WDRB} の総数である。

4.4 実験

4.4.1 実験環境と実験方法

提案手法を用いた個人再識別の有効性と安定性を確認するために、Logicool c615 ウェブカメラ（図 3.2 に参照）とパソコンをそれぞれ一台用いて実験用俯瞰視システムを構築した。使用したパソコンの CPU はインテル Core-i7 4770、メモリは DDR3 PC1600 16GB、OS は Windows 7 Professional 64bit、プログラミングソフトは Visual C++ 2013(32bit) である。この俯瞰視システムの稼働環境は室内でした。

予備実験によって、4.3 節で述べたセルの大きさ $2^n \times 2^n \times 2^n$ の n が 0 から 7 までのバタチャリア距離での照合時間と再識別の成功率の関係を調べた（表 4.1 に参照）。 $n = 0$

表 4.1 n の変化による、特徴量間のマッチング時間、再識別成功率

n	1	2	3	4	5	6	7
照合時間 (ms/人)	421.0	47.0	16.0	0.31	0.047	0.0063	0.00094
再識別の成功率	同一レベル		50% 以下				

の場合は、今回の試作システムで使用するハードウェアの制限により動作できなかった。 $n = 1$ と $n = 2$ の場合、照合時間の差があるが、再識別の成功率は両方とも同一レベルである。 $n = 3$ と $n = 7$ の場合、照合時間がかなり短縮できたが、再識別の成功率はすべて 50% 以下であり、利用できない。提案手法が計算速度を重視するため、今回 n を 2 とした。

提案手法の有効性を確認するために、17 人分の俯瞰視画像セットを利用し、同一照明条件下でターゲット領域の異なる実験を行った。

提案手法の安定性を確認するために、1) 同じ色の服を着ている 6 人分の俯瞰視画像を用いてそれぞれ 20 回以上ターゲット領域を変化させた実験、2) 異なる光源下で Nakatani R らの手法 [134] との比較実験、3) 同一照明条件下で CDF 法との人物再識別実験を行った。

4.4.2 色のヒストグラムを考慮した色の強調の有効性実験

この実験では、登録色のヒストグラムによる色の強調の有効性を検証した。色が同じ、模様が異なる図 4.4 のサンプル 1 とサンプル 2 に対して、登録色のヒストグラムを利用した場合 (式 (4.2) の通りに利用する) と利用しない場合 (式 (4.2) の右側第 2 項はなく、単純に $\Omega_i = \alpha$ を利用する)、強調マップ中のそれぞれの色の立方体の様子を可視化して記録した。

図 4.4 のサンプル 1 とサンプル 2 はそれぞれ同じ 2 種類の色がある。図 4.8 の (a) と (b) はサンプル 1 とサンプル 2 それぞれの色のヒストグラムを利用しない場合の色の立方体である。2 枚のサンプル画像の特徴量が同一で、お互いに識別できなかった。

一方、提案手法の場合、各画像の 2 つ色のヒストグラムの差で、強調した立方体の大きさも異なるので、2 つの画像をそれぞれに識別できるようになった (図 4.9 の (a) と (b) に参照)。

この実験より、提案する式 (4.2) を用いた色のヒストグラムの利用は、色が同じ、模様

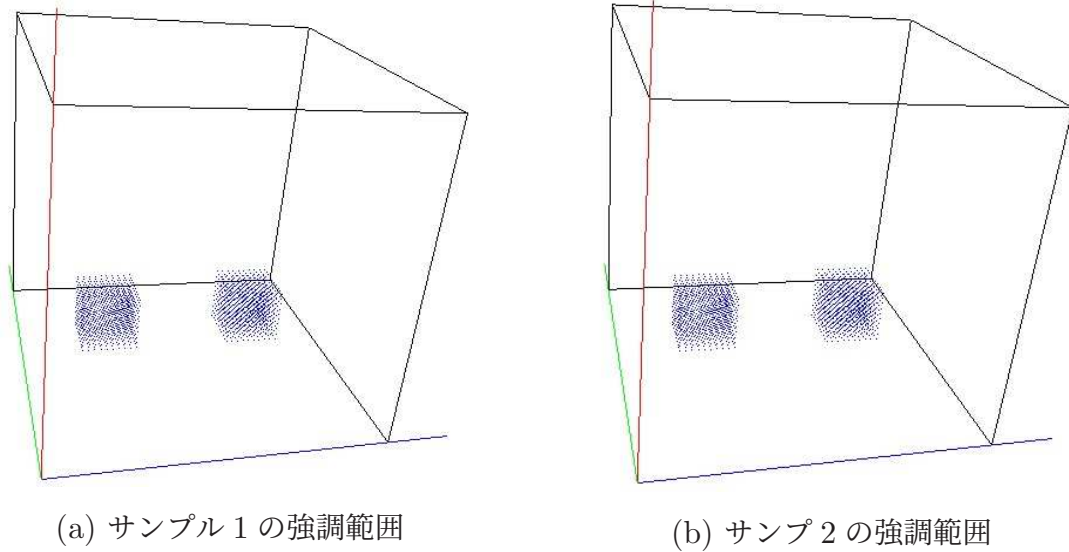


図 4.8 色のヒストグラムの利用しない場合の強調範囲

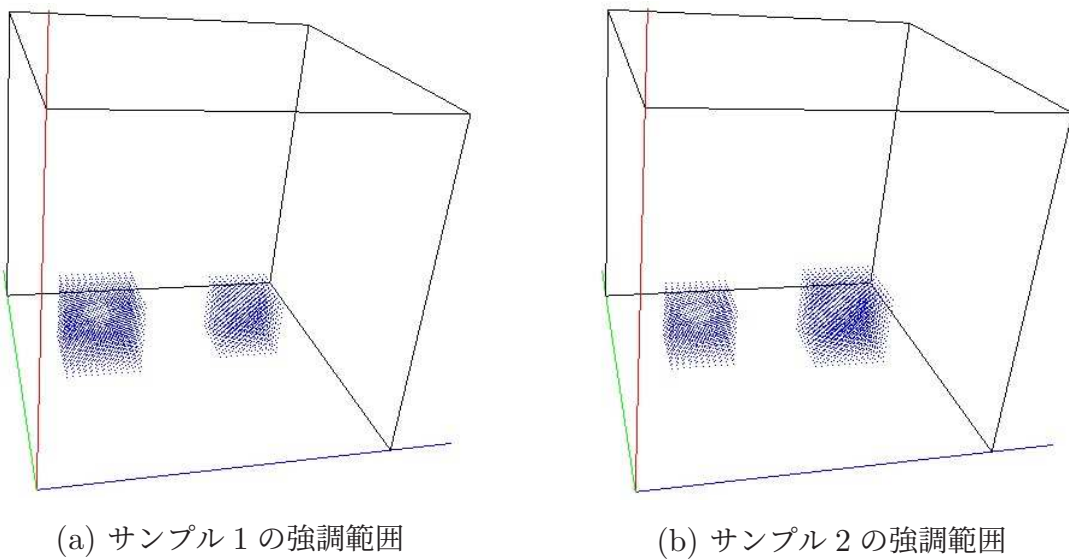


図 4.9 色のヒストグラムの利用した場合の強調範囲

が異なる画像間の識別に有効であることを確認できた。

4.4.3 WDRB 法を利用した個人再識別の有効性実験

この実験では、被験者 17 人（男性 13 人分、女性 4 人分）、一人の被験者に対して同じ光源下の入室（図 4.10 の (a)）と退室（図 4.10 の (b)）時に各一枚ずつ撮影し、俯瞰視

画像セットとした。



(a) 入室俯瞰視画像セット



(b) 退室俯瞰視画像セット

図 4.10 人物の入室俯瞰視画像セットと退室俯瞰視画像セット

各被験者の入室画像から、背景を含まず人物の髪の毛や肩部のみ含む最大矩形を手動で選択してターゲット領域とした (図 4.11 の黄色枠). このターゲット領域より該当被験者の **WDRB** を生成し、撮影順に ID 番号を付けてデータベースに登録する. 各人物が退室画像から、図 4.12 の中に示す黄色枠の中から手動でランダムに人物の髪の毛や肩部を含む異なるターゲット領域 (赤色点線の矩形) を 20 回指定し、各ターゲット領域で **WDRB** を生成し、17 人のデータベースとのマッチングを繰り返し行った. 図 4.13 に 17 人分の再認識の成功率を示す.

図 4.13 の横軸は被験者の番号である, 縦軸は 20 回のマッチングの内の成功回数より計算した再識別成功率である. この結果によると, 9 番と 16 番の被験者を除いた被験者の再識別の成功率は 85% 以上であった. 図 4.10 に示す俯瞰視画像セットによると, 8 番,

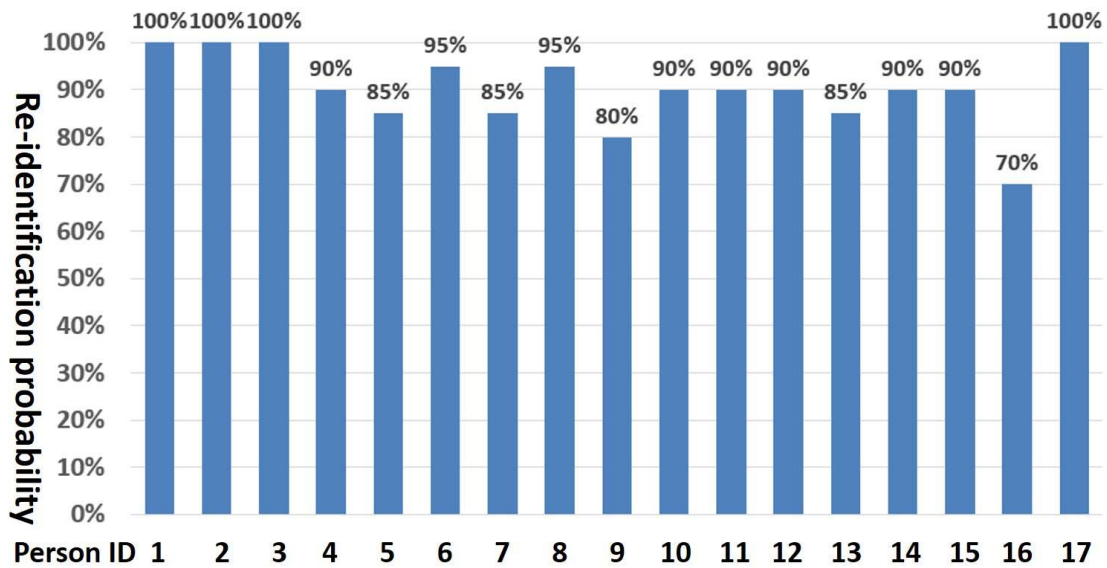


図 4.13 17 人 × 20 回の個人再識別の識別率

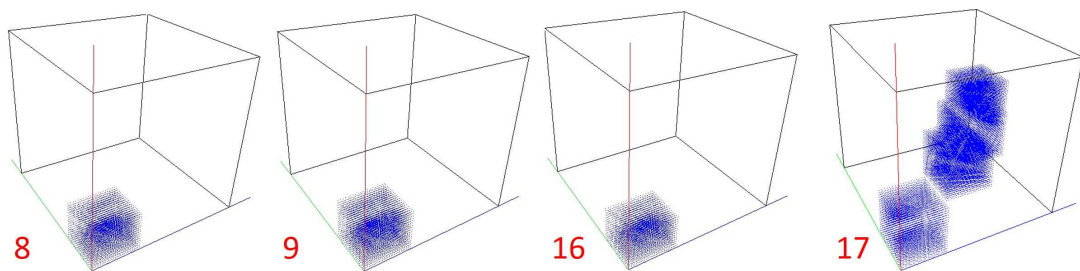


図 4.14 8 番, 9 番, 16 番と 17 番の色のエネルギー範囲

4.4.4 被験者が同じ色の服を着ている実験

提案手法を用いて、「対象となる人物が同じ色の服を着ている場合」という厳しい条件の下での再識別の実験を行った。

この実験では、男性被験者 6 人を対象とした。図 4.15 に示すように、同じ光源下で被験者の入室と退室の俯瞰視画像セットを撮影し、被験者毎に 41 回の個人再識別実験を行った。また、退室画像内の各被験者の **WDRB** 特徴生成は 4.4.3 項に述べた実験と同じく、黄色のターゲット領域から毎回異なる範囲を選択した。再識別の成功率は図 4.16 に示す。

図 4.16 の横軸は被験者の番号である、縦軸は 41 回のマッチングの内の成功回数より計算した再識別成功率である。この結果によると、被験者の個人再識別の成功率は最高が

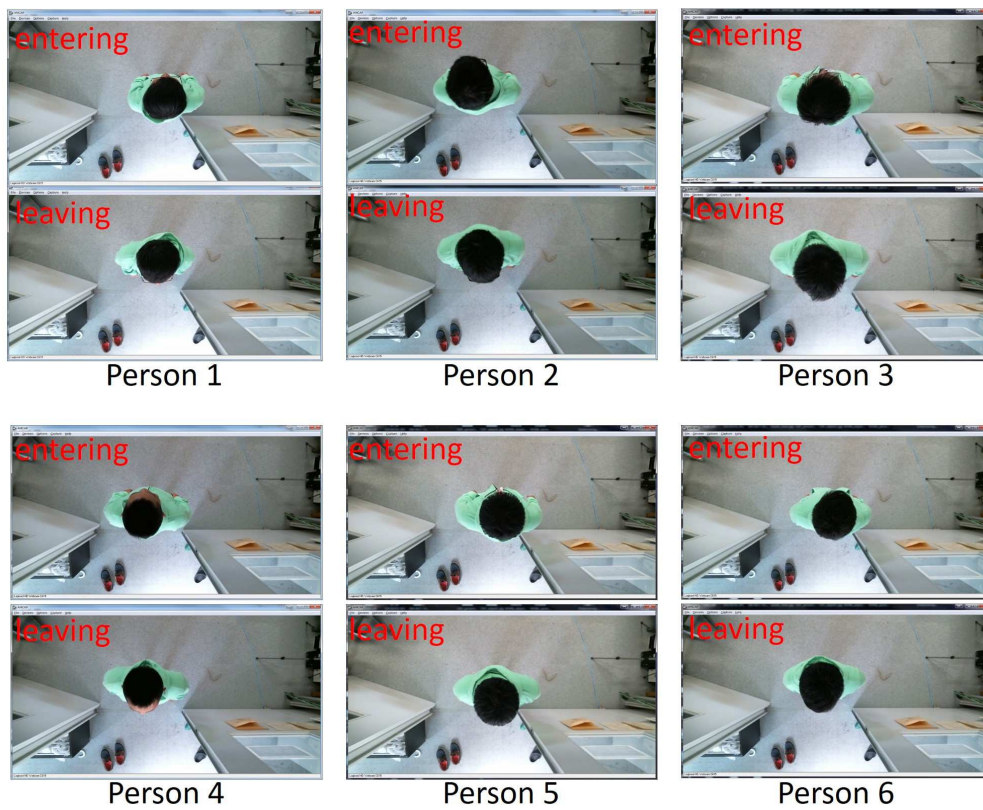


図 4.15 同じ色の服を着る被験者の入室や退室の俯瞰視画像

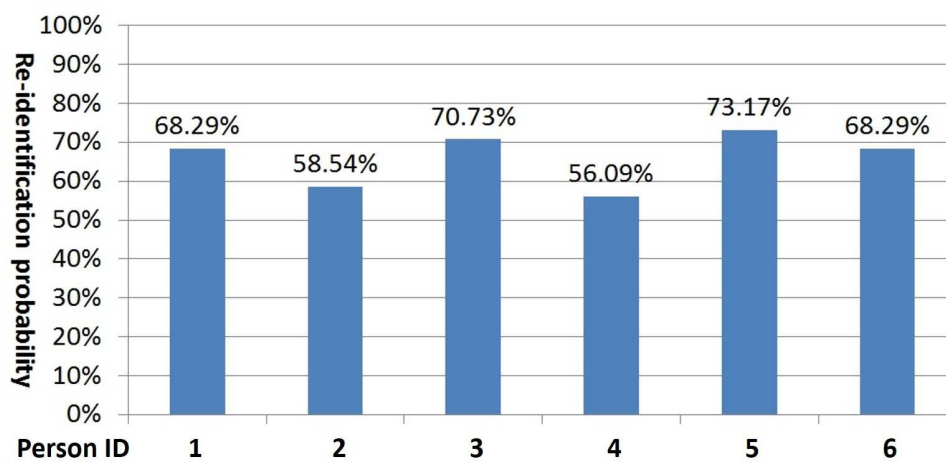


図 4.16 同じ色の服を着る 6 人 × 41 回の再識別の識別率

73.17%で、最低が56.09%であった。6人とも黒い髪かつ同じ色の服を着てもこの識別率が得られる理由としては、図4.17で示す通りに、登録された被験者の髪の色セルの強調範囲（RGB色空間の原点付近の小さいブロック）と肩の服色のセルの強調範囲（RGB色空間の真ん中付近の大きいブロック）がお互いに分かれていることである。

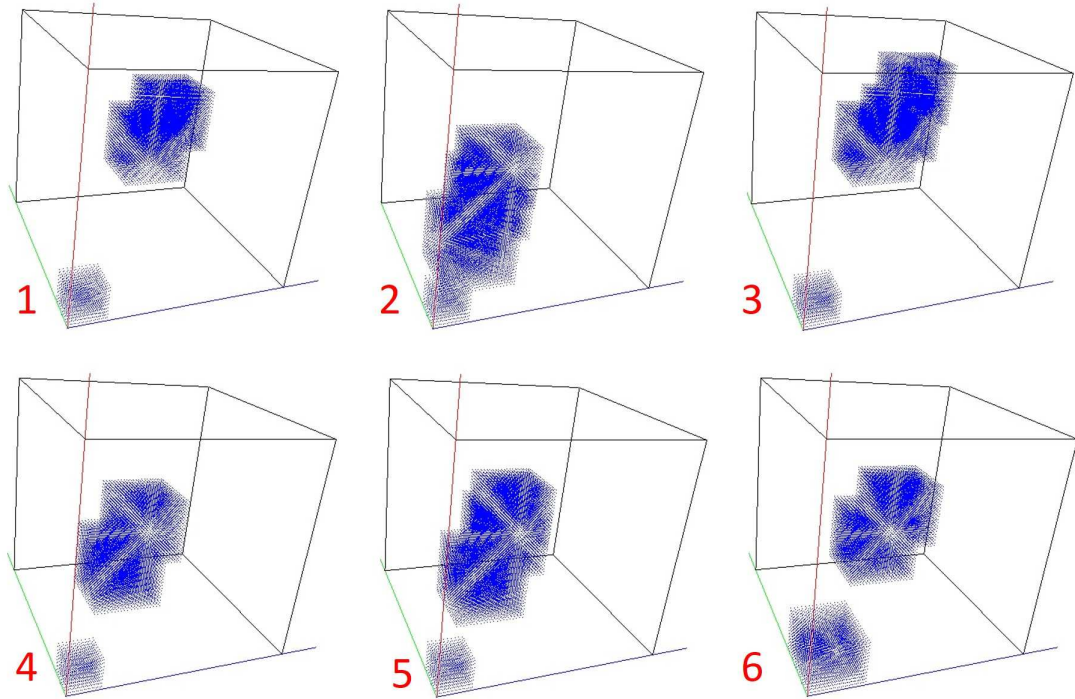


図 4.17 退室の6人の色の強調範囲

一方、髪の色微妙な個人差は、詳細な色情報が俯瞰視画像に映されており、提案手法を用いることでその差異がうまく強調できているため、厳しい条件の下でも個人再識別ができていると考えられる。

4.4.5 CDF 法との比較実験

この実験では、有効性実験（4.4.3項）と同じデータベースを利用した。同様に、各被験者のWDRBとCDF特徴量は、被験者が入室画像（図4.10(a))により生成した特徴量を既知人物の特徴量としてデータベースに登録する。退室画像（図4.10(b))より生成した特徴量を未知人物の特徴量としてデータベースの各被験者の特徴量とのマッチングを行う。WDRBのターゲット領域は背景を含まず被験者の髪の毛と肩部のみを含む最大矩形を利用する（図4.18(a)に参照）。CDF法の場合、人物の髪の毛のみを含む矩形をサ

ポート領域とし、肩部のみを含む矩形をメイン領域とする（図 4.18 (b) に参照）。類似度の計算はバタチャリア距離より行った。

WDRB 法と CDF 法を利用した個人再識別実験の結果はそれぞれ図 4.19 (a) と (b) に示す。図 4.19 (a) と (b) は混同行列であり、各サブ実験の結果を行ごとに示している。各サブ実験の中に、最も高い類似度の結果は赤枠で示されている。ID.1~17 は、対応する人物の入室画像を用いて計算したデータベースの WDRB と CDF である。Exp 番号に対応する人物の退室画像を使用して、データベースの中の各人物の特徴量とのバタチャリア距離の類似度は計算した。類似度を見やすくするために、すべての類似度はグレー値に変換され、結果を表示する箇所は白いほど類似性が高いことを意味している。自分自身の特徴量間の類似度が最も高いはずなので、左上から右下への対角線上の結果が最も白い線になるのは正確だ。図 4.19 に示すように、WDRB 法の対角線上の結果は最も白くて高精度である。一方、ほとんどの CDF 法の結果が各人物を明確に再識別できず、誤識別が多数出られた。

この実験より、WDRB 法は CDF 法より再識別率が高いことが分かった。

4.4.6 異なる光源下での比較実験

提案手法で記述する特徴量の照明変化に対する頑健性を確認するために、図 4.20 に示すような屋内の照明環境およびレイアウトの下で異なる光源下の人物の入室・退室の画像セットを用意し、Nakatani R らの手法 [134] との比較実験を行った。部屋の中の明るさは二本一セットの蛍光灯光源 ($L1$) と、一本の蛍光灯光源 ($L2$) と、窓からの自然光 ($L3$)、四種類の光源の組み合わせ ($L1 + L2 + L3$ 、 $L1 + L3$ 、 $L2 + L3$ 、 $L3$) とし、照

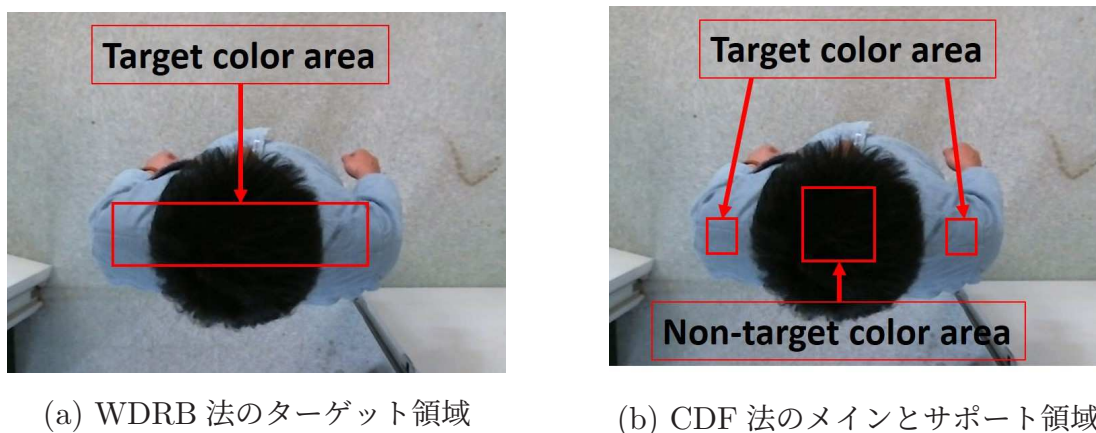
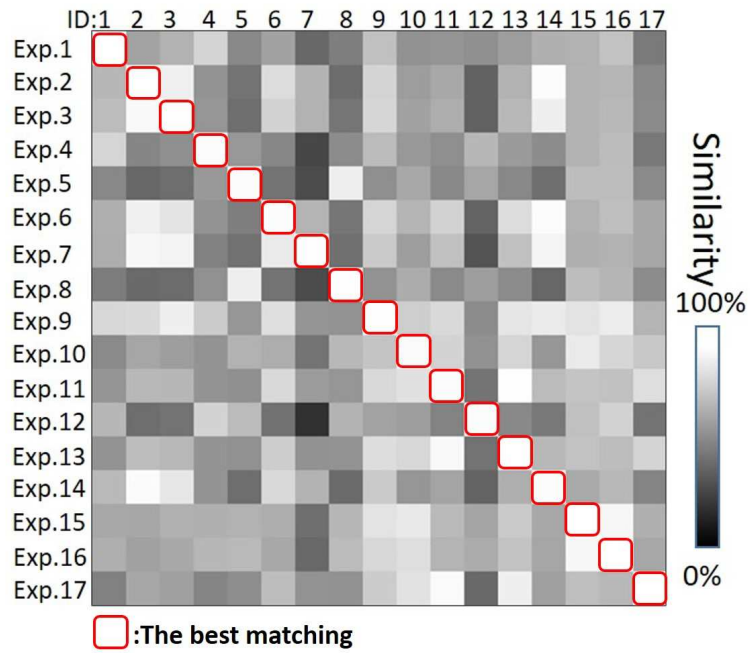
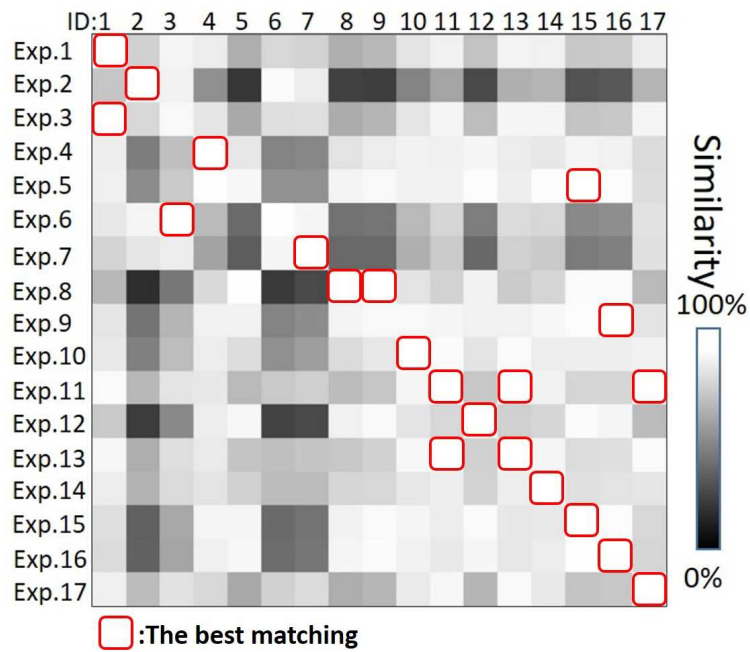


図 4.18 WDRB 法と CDF 法の人物領域の選択



(a) WDRB 法の結果



(b) CDF 法の結果

図 4.19 WDRB 法と CDF 法による個人再識別実験の結果

明条件を変化させた。

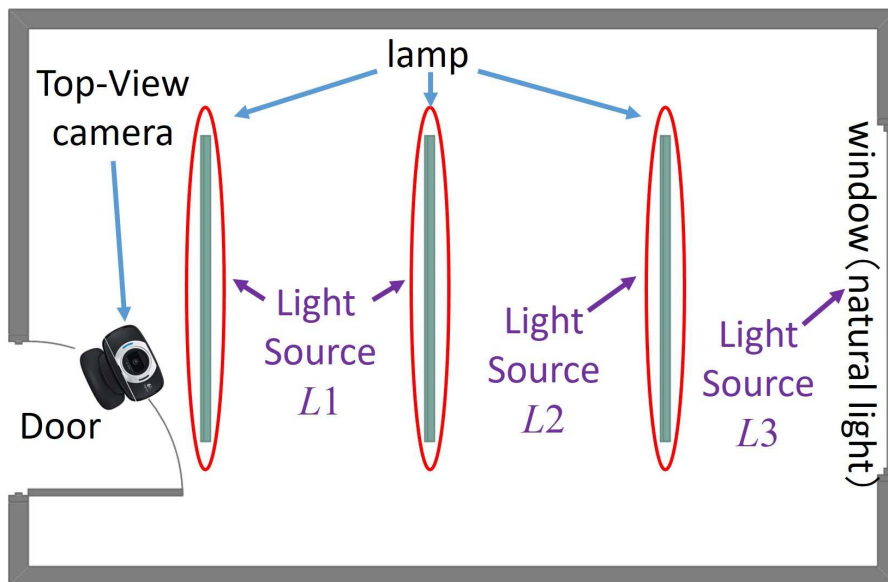


図 4.20 実験現場の光源配置

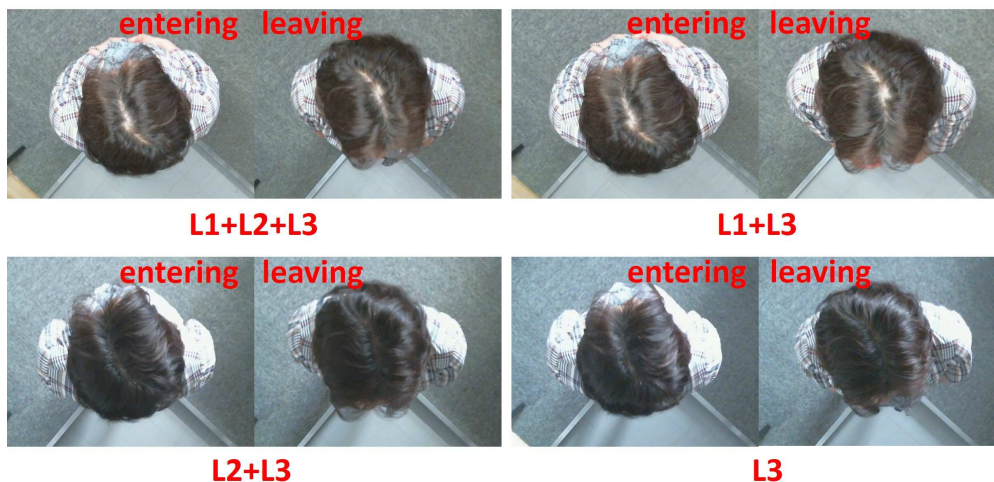


図 4.21 異なる光源下の人物の入室・退室の俯瞰視画像サンプル

被験者は全部 12 人 (男性 6 人・女性 6 人), 1 人あたりに各光源下において入・退室 10 枚 (入室時 5 枚と退室時 5 枚) の俯瞰視画像を撮影し, 計 480 枚の画像セットを利用した (図 4.21 のサンプルに参照). 3 章の比較実験 (3.3.2 項) と同様に, 実験用検証ソフト (Weka[158]), 分類器 (C4.5 アルゴリズム) 及び他の設定はすべて Nakatani R らの論文 [134] と同様であり, 10 分割交差検証で実験を行った. Nakatani R らの手法において,

再認識の成功率が一番高い *size + style + color* の特徴量の組み合わせを利用する.

Correctly Classified Rate: 80.625%												
a	b	c	d	e	f	g	h	i	j	k	l	
33	0	0	0	0	0	3	0	4	0	0	0	a = p1
0	33	0	0	7	0	0	0	0	0	0	0	b = p2
0	0	35	0	0	0	0	5	0	0	0	0	c = p3
1	0	0	32	0	5	0	0	0	0	1	1	d = p4
0	6	0	0	34	0	0	0	0	0	0	0	e = p5
0	0	0	6	0	31	0	0	2	0	0	1	f = p6
7	0	0	0	0	0	33	0	0	0	0	0	g = p7
0	0	5	0	0	0	0	30	0	2	0	3	h = p8
3	0	0	1	0	2	1	0	27	0	6	0	i = p9
0	0	1	0	1	0	0	3	0	34	0	1	j = p10
1	0	0	0	0	0	0	0	6	0	33	0	k = p11
0	0	0	0	0	3	0	3	1	1	0	32	l = p12

(a) Nakatani R ら手法の結果

Correctly Classified Rate: 89.375%												
a	b	c	d	e	f	g	h	i	j	k	l	
37	0	0	0	0	0	0	0	0	2	0	1	a = p1
0	39	1	0	0	0	0	0	0	0	0	0	b = p2
0	0	29	2	1	0	0	0	5	0	3	0	c = p3
0	0	2	34	1	0	0	0	0	0	3	0	d = p4
0	0	2	3	31	0	0	1	0	0	3	0	e = p5
0	0	0	0	0	40	0	0	0	0	0	0	f = p6
0	0	2	0	0	0	38	0	0	0	0	0	g = p7
0	0	0	0	0	0	2	38	0	0	0	0	h = p8
0	0	4	0	0	0	0	0	34	0	1	1	i = p9
1	0	0	0	0	0	0	0	0	39	0	0	j = p10
0	0	1	1	3	0	0	1	1	0	33	0	k = p11
0	0	1	0	0	0	0	1	1	0	0	37	l = p12

(b) 提案手法の結果

図 4.22 異なる光源下での比較実験結果

Nakatani R らの手法と提案手法の再認識の成功率はそれぞれ 80.625% と 89.375% であり、両手法の成功回数・失敗回数を表す混同行列は図 4.22 の (a) と (b) である。構

内・室内などの一般的な明暗変化に対する頑健性について、提案手法は Nakatani R らの手法より高い再識別能力を示した。

4.5 まとめ

本章では、人物のプライバシーなどの利用を回避するために、髪の毛や服の色を利用した俯瞰視画像による再識別方法として WDRB 特徴記述法を提案した。WDRB 法の有効性を検証するために、WDRB 法および CDF 法の個人識別実験がを実施し、バタチャリヤ距離を使用して特徴量間の類似度評価を行った。多数の実験結果から、WDRB 法は CDF 法よりも再識別精度が高いことを示した。また、Weka による比較実験より、WDRB 特徴量は CDF 特徴量より高い頑健性を持つことを示した。CDF 法と物体再識別の比較実験より、提案手法は物体の再識別にも利用できることを確認した（付録 B に参照）。

第 5 章

UV 色平面内の色情報の拡張記述による個人識別

5.1 WDRB 法の問題点

俯瞰視システムを利用すると、カメラの設置場所の変化がなければ、俯瞰視画像上の人物領域の大きさは人物の体格を反映する重要な情報である。WDRB 手法は、人物の領域を手動で長方形を選択しているため、人物の体格情報が正確に得られない。また、WDRB 手法では、画像ピクセルの RGB 情報を利用している。RGB 情報は画像の撮影環境の明るさ変化の影響を受けやすいため、再識別の精度への影響も出る。この二つの問題点を克服できれば、人物の再識別の精度を向上できると考えられる。

5.2 SHAL 特徴記述法

本章では、一枚のカラー画像から得られる人物の頭頂部と肩部の画素情報を用いて、個人再識別可能な特徴記述法（SHAL:Shoulder and HeAd Learning）を提案する。

5.2.1 頭頂部・肩部別の画素ベクトルと色のヒストグラム

SHAL 法では、照明変化に頑健な特徴を記述するために、俯瞰視画像から得られた各画素の色情報 (r, g, b) を式 (5.1) で RGB 空間から輝度情報 (y) と色情報 (u, v) が独立に表

現できる YUV 色空間へ変換して利用する.

$$\begin{cases} y = 0.299 \times r + 0.587 \times g + 0.114 \times b \\ u = -0.169 \times r - 0.3316 \times g + 0.500 \times b \\ v = 0.500 \times r - 0.4186 \times g - 0.0813 \times b \end{cases} \quad (5.1)$$

ここで, r , g と b の値域はそれぞれ $0 \sim 255$ であり, 求めた y の値域が $0 \sim 255$, u と v の値域がそれぞれ $-128 \sim 127$ である. 以下, Y , U , V は YUV 色空間の三つの座標軸を表す.

再識別精度と, 速度・メモリ容量のトレードオフを両立するために UV 色平面を $2^n \times 2^n$ サイズの小さいセルに分割した. 分割後のセル数は $2^{7-n} \times 2^{7-n}$ 個である. ここで, $n = 0, 1, \dots, 7$ である. 以下の処理はすべて同様に量子化された UV 色平面で行う.

前処理 [159] (5.4.2 「人物領域の自動検出における前処理」に参照) によって俯瞰視画像から人物の頭頂部と肩部をそれぞれ検出して, 図 5.1 に示すように, 頭頂部の画素をベクトル $\mathbf{P}_{\mathbf{h}_q} = \{ (y_{\mathbf{h}0}, u_{\mathbf{h}0}, v_{\mathbf{h}0}), \dots, (y_{\mathbf{h}q}, u_{\mathbf{h}q}, v_{\mathbf{h}q}) \}$ に, 肩部の画素をベクトル $\mathbf{P}_{\mathbf{s}_p} = \{ (y_{\mathbf{s}0}, u_{\mathbf{s}0}, v_{\mathbf{s}0}), \dots, (y_{\mathbf{s}p}, u_{\mathbf{s}p}, v_{\mathbf{s}p}) \}$ に記録する. ここから, \mathbf{h} を頭部 (head), \mathbf{s} を肩部 (shoulders) と表し, q と p は頭頂部と肩部の画素総数とする.

また, 頭頂部と肩部の画素における UV 色平面内の色のヒストグラム ($\mathbf{C}'_{\mathbf{h}(u,v)}, u, v$) と ($\mathbf{C}'_{\mathbf{s}(u,v)}, u, v$) (以下, 色登録マップと呼ぶ) を生成する. uv 色のヒストグラムのビン (bin) サイズはセルのサイズと同じである. 図 5.4 (a) は生成された色登録マップ例を示す. $\mathbf{C}'_{\mathbf{h}(u,v)}$, $\mathbf{C}'_{\mathbf{s}(u,v)}$ 内のあるビンの度数値域は $0 \sim +\infty$ であり, 度数が非零である色 (u, v) を登録色と呼び, 度数が零である色 (u, v) を非登録色と呼ぶ. 色登録マップ例の UV 色平面への俯瞰図を図 5.3 (a) に示す, 中の点は登録色の位置である.

5.2.2 頭頂部・肩部の同じ色の輝度 y の合計値分布

頭頂部に対応する画素は無彩色である. 紺や白, ベージュなどの服を着ると, 肩部に対応する画素も無彩色になる. 色によらず, 重要な輝度情報を活用するために, 画素ベクトル $\mathbf{P}_{\mathbf{h}_q}$ と $\mathbf{P}_{\mathbf{s}_p}$ から同じ uv 値の画素における輝度成分の合計値 $\mathbf{B}_{\mathbf{h}(u,v)}$ と $\mathbf{B}_{\mathbf{s}(u,v)}$ を式 (5.2) でそれぞれ計算する.

$$\mathbf{B}_{\mathbf{e}(u,v)} = \sum_{k=1}^f Y_{(u,v)_k} \quad (5.2)$$

ここで, \mathbf{e} は \mathbf{s} あるいは \mathbf{h} である. f は同じ uv 値の画素数である. 求めた輝度成分の合計値 $\mathbf{B}_{\mathbf{e}(u,v)}$ を UV 色平面の該当座標にそれぞれ格納する. 色の輝度 y の合計値を格納す

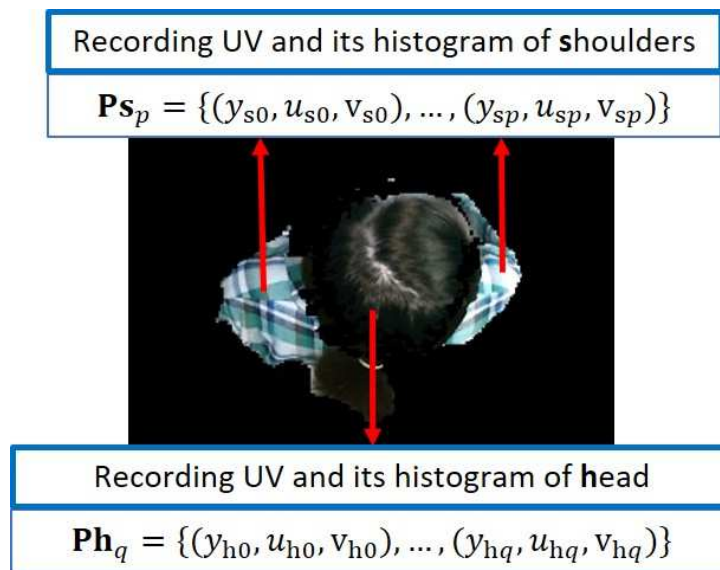
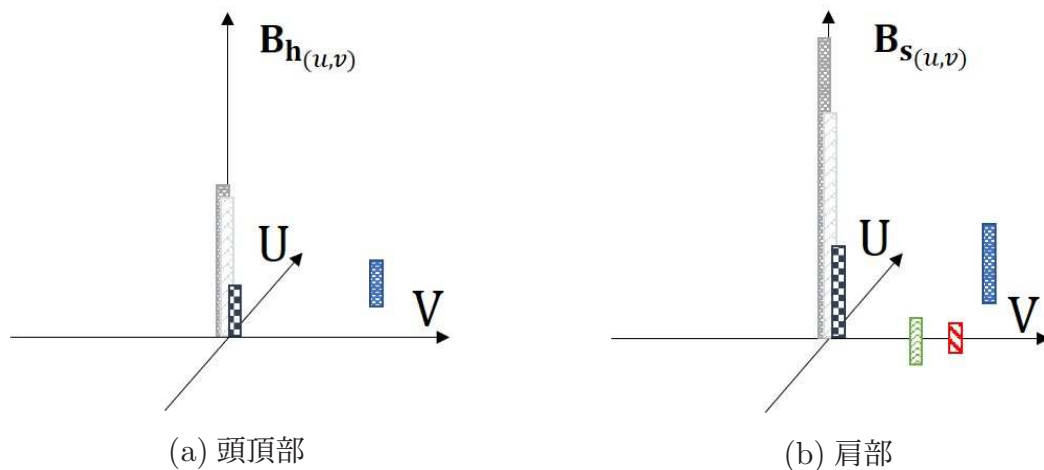


図 5.1 頭頂部・肩部別の画素ベクトルと色のヒストグラム

る座標 (u, v) 位置と色登録マップ内の登録色の座標位置に対応している. 図 5.2 (a) と (b) はそれぞれ得られた頭頂部・肩部の同じ色の輝度 y の合計値分布例 $(\mathbf{B}_{h(u,v)}, u, v)$, $(\mathbf{B}_{s(u,v)}, u, v)$ である.

図 5.2 頭頂部・肩部の同じ色の輝度 y の合計値分布例

5.2.3 色の種類数と位置に基づくユークリッド距離変換

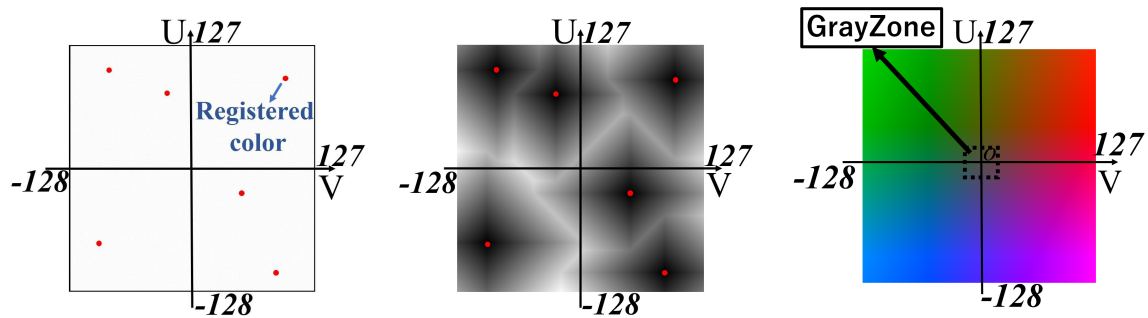
UV 色平面内に、各個人の頭頂部と肩部の色情報に基づいて生成する頭頂部と肩部に対応した2つの色登録マップには色の種類数、相対的な距離関係を登録している。これらはすべて個人の特徴情報を示す重要な情報である。また、色登録マップにおいて非登録色は、登録色に比べて、圧倒的に多いため、個人再識別率の向上に利用することが考えられる。そこで、色登録マップを色距離マップに変換する。

UV 色平面に登録色の座標値を \mathbf{t} 、非登録色の座標値を \mathbf{r} とすると、UV 色平面内の各非登録色 $\mathbf{r}_i (i = 1, \dots, I)$ におけるすべての \mathbf{t} までのユークリッド距離を計算し、さらに、各非登録色 \mathbf{r}_i について、 \mathbf{t} までの最短距離 D_i を式 (5.3) で計算する。

$$D_i = \arg \min_{j \in \{1, \dots, m\}} \|\mathbf{r}_i - \mathbf{t}_j\|_{L_2} \quad (5.3)$$

ここで、 i と I は要素 \mathbf{r} の番号と総数で、 j と m は要素 \mathbf{t} の番号と総数である。非登録色 \mathbf{r}_i から最も近い登録色までの距離を D_i として、色距離マップ ($\mathbf{D}_{(u,v)}, u, v$) に登録する。

図 5.3 (b) には色登録マップ (図 5.3 (a)) に対してユークリッド距離変換を行って、得られた色距離マップ例を示す。色距離マップでは、 \mathbf{r}_i についての距離の値が大きいほど白くなるようにしたグレースケールで距離が示されている。



(a) 色登録マップ例 (b) 色距離マップ例 (c) Grayzone の定義

図 5.3 色登録マップと色距離マップ例及び Grayzone の定義の平面図

5.2.4 SHAL 特徴記述法作成プロセス

色登録マップの拡張および色距離マップとの融合

5.2.1 項から 5.2.3 項までの処理によって、頭頂部/肩部の色の種類数・相対距離は、色距離マップで、体の見えに関連する色ヒストグラムを含む色登録マップで、それぞれ独立

して表現している。SHAL 法では、人物の頭頂部と肩部の各色を特徴とする頑健性を強調するために、色登録マップ内の色ヒストグラムを利用して、色登録マップ内で各登録色の影響範囲を拡大する「拡張処理」を行い、拡張された色登録マップと色距離マップを融合して色特徴量として利用する。

色登録マップを拡張するために、色登録マップの中に l 番目の登録色を中心とする正方形の辺長 $\Omega_{e(u,v)_l}$ (単位：セル) を式 (5.4) で求める。

$$\Omega_{e(u,v)_l} = \alpha \times \left(1 + \frac{C'_{e(u,v)_l}}{\sum_{k=1}^m C'_{e(u,v)_k}} \right) \quad (5.4)$$

ここで、 e は s あるいは h である。 α はセルと関連する定数である。 $C'_{e(u,v)_l}$ は l 番目の登録色に対応する色累積度数であり、 $\sum_{k=1}^m C'_{e(u,v)_k}$ は全 m 個登録色に対応する色の累積度数の和である。度数 $C'_{e(u,v)_l}$ が大きい登録色ほど、大きく拡張された影響範囲 (正方形範囲) を有する。さらに、式 (5.4) の括弧内の第 2 項の導入によって、2 種類以上の色を含む 2 枚の画像同士、色成分が同じでも、色同士の相対的な画素数が異なるならば、求められた正方形範囲が異なるので、区別することが可能になる。

色登録マップにおける登録色 $(u, v)_l$ を中心とする UV 色平面上の $\Omega_{e(u,v)_l}$ の正方形範囲に含まれる非登録色に、登録色 $(u, v)_l$ の度数 $C'_{e(u,v)_l}$ と同じ値の度数 $C_{e(u,v)_l}$ を付与する。図 5.4 (b) は人物の頭頂部 ($C_{h(u,v), u, v}$) と、肩部 ($C_{s(u,v), u, v}$) が式 (5.4) で求めた $\Omega_{e(u,v)_l}$ によって、色の影響範囲が拡張された色登録マップ例である。影響範囲が拡張された色登録マップでは、持つ色の特徴を強調して記述する。

色登録マップ内の近隣登録色を中心とした正方形範囲と重なる部分がある場合、それらの累積度数の合計値を重なる部分に設定する (図 5.5 に参照)。

融合処理として、色の影響範囲が拡張された色登録マップ (図 5.6 (b)) 及び色距離マップ (図 5.6 (a)) を重畳させる。融合したイメージを図 5.6 (c) に示す、以下、これを色融合マップ ($CF_{e(u,v), u, v}$) と呼ぶ。

無彩色画素の割合による重み設定

輝度 y の合計値分布には、頭頂部/肩部の登録色ごとの輝度 y の合計値を記録している。UV 色平面の原点付近の登録色は無彩色なので、輝度の合計値 (輝度特徴量) が識別に重要な識別情報になる。一方、原点付近から離れた登録色は有彩色なので、色融合マップの結果 (色特徴量) が識別に有効である。

SHAL 法には、特徴量を評価する際に、UV 色平面における位置に応じて、色と輝度の重

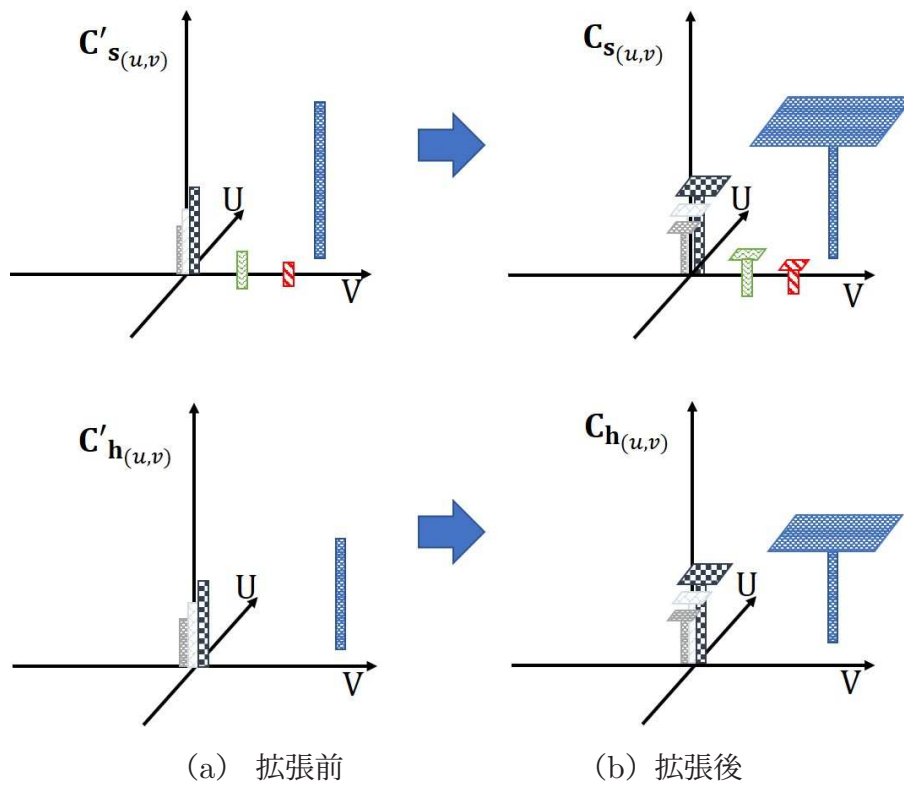


図 5.4 3次元で表示する頭頂部と肩部の色登録マップと拡張された色登録マップ

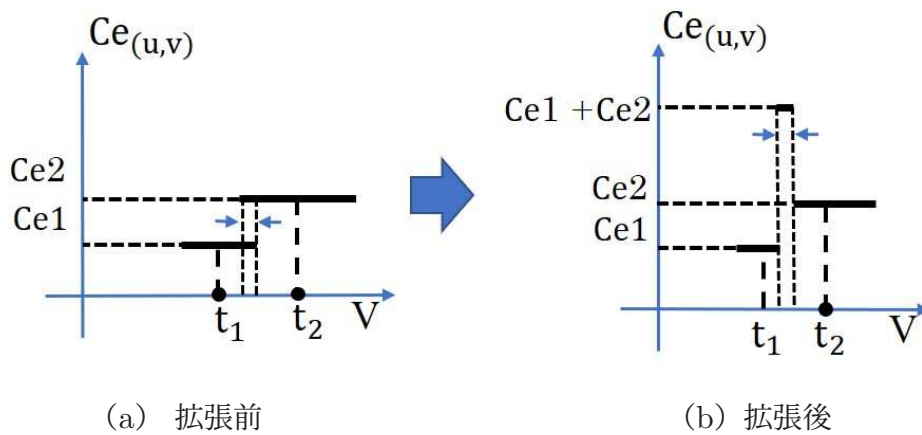


図 5.5 色登録マップに近隣登録色に強調範囲が重なる場合の設定方法の側面図

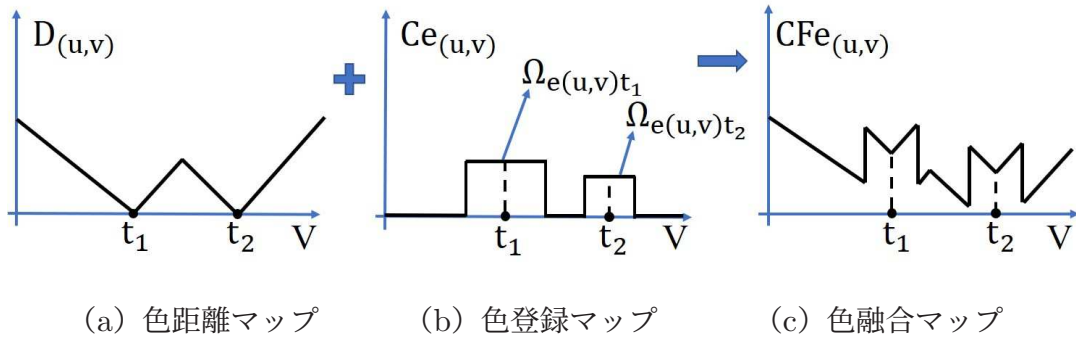


図 5.6 色距離マップと拡張された色登録マップと色融合マップの側面図

要度が異なることを利用する．そこで，図 5.3 (c) に示すように，無彩色を囲む Grayzone を設定し，色登録マップを二つの領域に分ける．Grayzone の内側に登録している無彩色の全登録色の累積度数の和 N_{e_y} と，Grayzone 外側に登録している有彩色の全登録色の累積度数の和 $N_{e_{uv}}$ をそれぞれ求めた後，式 (5.5) より，Grayzone 内側の全登録色の累積度数の和の割合 W_{e_y} と，Grayzone 外側の全登録色の累積度数の和の割合 $W_{e_{uv}}$ をそれぞれ求め，色と輝度の重要度を考慮した重みとする．

$$\begin{cases} W_{e_y} = N_{e_y} / (N_{e_y} + N_{e_{uv}}) \\ W_{e_{uv}} = 1 - W_{e_y} \end{cases} \quad (5.5)$$

ここで， e は s あるいは h である． $(N_{e_y} + N_{e_{uv}})$ は色登録マップにおける Grayzone 内外の全登録色の累積度数の合計値である．

個人特徴量 SHAL の決定

本論文では，輝度 y の合計値分布 ($\mathbf{B}_{e(u,v)}$)，色融合マップ ($\mathbf{CF}_{e(u,v)}$)，及び色と輝度の重要度を考慮する重み (W_{e_y} , $W_{e_{uv}}$) は人物の個人性を記述する $\mathbf{SHAL} = \{\mathbf{B}_{e(u,v)}, \mathbf{CF}_{e(u,v)}, W_{e_y}, W_{e_{uv}}\}$ 特徴量とする．ただし， e は s あるいは h である．

5.3 重み付きバタチャリヤ距離による SHAL 特徴量間の評価

俯瞰視カメラの下を 1 回目 (例えば：入室時) 通る時，撮影された人物俯瞰視画像から該当人物の \mathbf{SHAL} 特徴量を記述し，登録順番の id 番号と共にデータベースに記録する ($\mathbf{SHAL}(id)$)．データベースに登録している総人数は ID_{all} で表す．

俯瞰視カメラの下を 2 回目以降 (例えば：退室時，室内移動時) 通る時，撮影された人物

俯瞰視画像から未知人物の特徴量 $\mathbf{SHAL}(x)$ を記述する．未知人物の特徴量 $\mathbf{SHAL}(x)$ とデータベースに登録している ID_{all} 人分の $\mathbf{SHAL}(id)$ の間，色と輝度の重要度をバタチャリヤ距離と結合して，重み付きバタチャリヤ距離による類似度評価を行う．

まず，バタチャリヤ距離で，式 (5.6) で輝度 y の合計値分布 ($\mathbf{B}_{\mathbf{e}(u,v)}(x)$ と $\mathbf{B}_{\mathbf{e}(u,v)}(id)$) の類似度 $D_{Y\mathbf{e}}(x, id)$

$$D_{Y\mathbf{e}}(x, id) = \sum_{u,v \in \{-128, \dots, 127\}} \sqrt{\mathbf{B}_{\mathbf{e}(u,v)}(x) \mathbf{B}_{\mathbf{e}(u,v)}(id)} \quad (5.6)$$

と式 (5.7) で色融合マップ ($\mathbf{CF}_{\mathbf{e}(u,v)}(x)$ と $\mathbf{CF}_{\mathbf{e}(u,v)}(id)$) の類似度 $D_{UV\mathbf{e}}(x, id)$

$$D_{UV\mathbf{e}}(x, id) = \sum_{u,v \in \{-128, \dots, 127\}} \sqrt{\mathbf{CF}_{\mathbf{e}(u,v)}(x) \mathbf{CF}_{\mathbf{e}(u,v)}(id)} \quad (5.7)$$

をそれぞれ計算する．ここで， \mathbf{e} は \mathbf{s} あるいは \mathbf{h} である． $id \in \{1, \dots, ID_{all}\}$ である．次に， \mathbf{SHAL} に記述されている重みを用いて，式 (5.8) で未知人物とデータベースに登録している id 番目の人物間の総合類似度 $D_B(\mathbf{SHAL}(x), \mathbf{SHAL}(id))$ を求める．

$$\begin{aligned} D_B(\mathbf{SHAL}(x), \mathbf{SHAL}(id)) = & \\ & (W_{\mathbf{h}_y}(x) + W_{\mathbf{h}_y}(id)) \times D_{Y\mathbf{h}}(x, id) \\ & + (W_{\mathbf{h}_{uv}}(x) + W_{\mathbf{h}_{uv}}(id)) \times D_{UV\mathbf{h}}(x, id) \\ & + (W_{\mathbf{s}_y}(x) + W_{\mathbf{s}_y}(id)) \times D_{Y\mathbf{s}}(x, id) \\ & + (W_{\mathbf{s}_{uv}}(x) + W_{\mathbf{s}_{uv}}(id)) \times D_{UV\mathbf{s}}(x, id) \end{aligned} \quad (5.8)$$

最後に，式 (5.9) で一番大きい類似度 $D_{B_{maxid}}$ を見つけ，対応する id 番号を未知人物の ID とする．

$$D_{B_{maxid}} = \arg \max_{id \in \{1, \dots, ID_{all}\}} D_B(\mathbf{SHAL}(x), \mathbf{SHAL}(id)) \quad (5.9)$$

5.4 実験

提案手法の有効性を検証するために，1台の俯瞰視 RGB-D カメラ (Kinect V2) と1台のパソコンを利用して俯瞰視システムを構築した (5.4.3 項に参照)．実験環境と処理の流れをそれぞれに図 5.7 (a) と図 5.7 (b) に示す．使用したパソコンの CPU はインテル Core-i7 4770，メモリは DDR3 PC1600 16GB，OS は Windows 10 Professional 64bit，

プログラミングソフトは Visual C++ 2013 (32bit) である。室内照明環境と窓からの自然光のみで、異なる日の異なる時間帯で撮った俯瞰視画像を用いて実験を行った。

5.4.1 予備実験

5.2 節で述べたセルの大きさ $2^n \times 2^n$ の n が 1 から 7 まで前処理時間、一人当たりの特徴量生成時間、SHAL 間のマッチング時間と再識別の成功率の関係を（「私服を着る実験」のデータベースの 1 セットを用いて）調べる予備実験を行った。実験結果を表 5.1 に示す。前処理時間は、 $n(= 1, \dots, 7)$ とは関係ないため、いずれも同じ時間となっている。SHAL の生成時間及びマッチング時間は短縮したが、再識別の成功率は下がることが分かった。今回、再識別の成功率を重視し、 n を 1 とした。

表 5.1 n の変化による前処理時間、SHAL 特徴量生成時間、特徴量間のマッチング時間、再識別成功率

n	1	2	3	4	5	6	7
前処理時間 (ms/フレーム)	106.80						
生成時間 (ms/人)	5.41	4.41	4.24	3.85	3.79	3.59	3.70
マッチング 時間 (ms)	196.02	46.95	11.75	2.66	0.72	0.18	0.06
再識別の成功率	97.94%	90.88%	89.71%	87.35%	81.18%	75.00%	73.53%

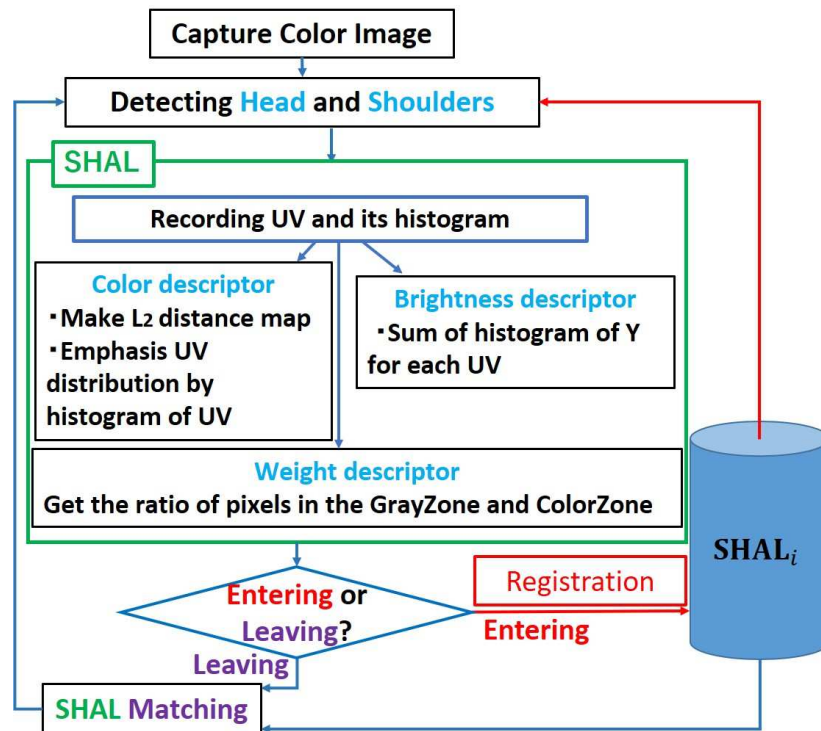
また、 $n = 1$ の下で、「私服を着る実験」のデータベースのうちの 3 セットをそれぞれ用いて、式 (5.4) で述べた定数 α の大きさと再識別の成功率との関係を調べる予備実験を行った。実験結果を図 5.8 に示す。ここで、横軸は α の値であり、縦軸は個人再識別の成功率である。図 5.8 において、 α を 10 から 60 まで変化させると、3 セットの成功率ともに少しずつ下がっていくが、少なくとも 96.5% 以上の成功率があることを確認した。今回、3 セットいずれも最高の成功率を得たときの α 値を利用し、 α を 20 とした。

5.4.2 人物領域の自動検出における前処理

再識別精度を保証できる SHAL の特徴量を生成するために、俯瞰視画像から背景を含まず、人物の頭頂部と肩部だけの画素ベクトル (\mathbf{Ph}_q と \mathbf{Ps}_p) を自動的に取得する前処理



(a) 実験環境



(b) 処理の流れ

図 5.7 俯瞰視システムの全体像および個人再識別実験の処理の流れ

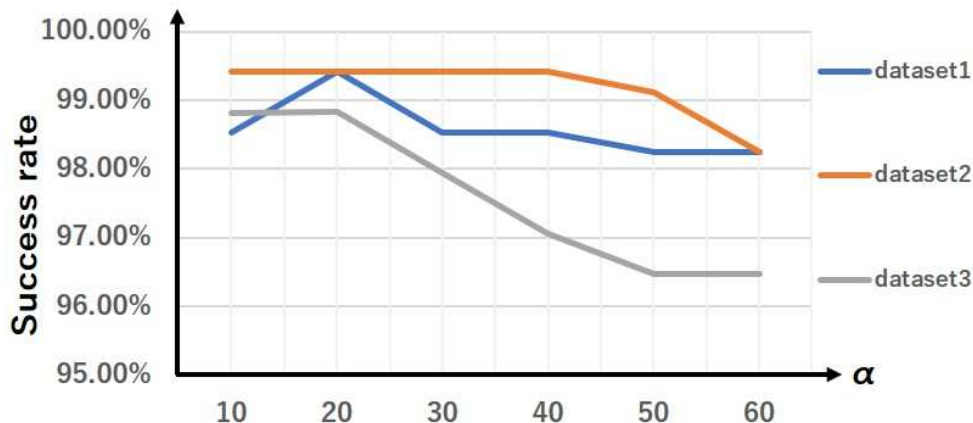


図 5.8 α の大きさと再識別の成功率との関係

が必要である。今回の実験では、この目的に合致する奥 剛一の手法 [159] を利用する。

カメラの深度情報よりカメラから人物の身長の高点まで距離を計測できる (図 5.9 (a) 参照)。奥 剛一の手法では、俯瞰視カメラの真下を通過する人物に対して、計測された人物の高点の軌跡より、人物の現在位置と移動方向を計測できる。図 5.9 (b) に示すように、時系列で人物の高点の移動軌跡順は線 L1 から線 L3 への場合は「入室」とし、逆に、L3 から L1 への場合は「退室」とする。それぞれの場合、人物の高点が線 L2 を通る瞬間の深度画像 (512 × 424 ピクセル) とカラー画像 (1920 × 1080 ピクセル) をキャプチャする。

子供から成人まで個人差があるが、頭頂部から肩部までの距離ばらつきは一定の範囲内である。奥剛一の手法では、頭頂部から耳あたりまでの距離を 100mm とし、図 5.9 (a) に示すよう、キャプチャされた画像を用いて、

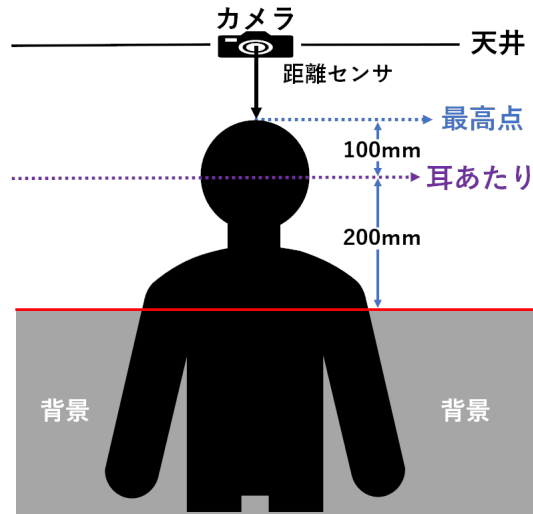
1) 最高点 +100mm という自動閾値 TH_e を用いて、深度画像内の距離は TH_e 以下の領域を頭頂部領域とする。

2) 最高点 +300mm という自動閾値 TH_b を用いて、深度画像内の距離は TH_b 以上の領域を背景領域とする。

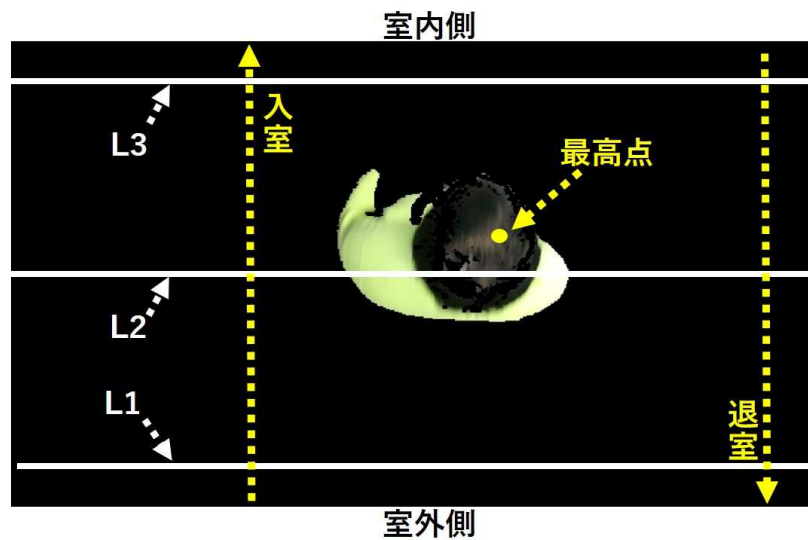
3) 深度画像内の距離は TH_e から TH_b までの範囲内の領域を肩部領域とする。

そして、ラスタースキャンの順で、深度画像をカラー画像と対応しながら、カラー画像上の頭頂部領域の画素をベクトル $\mathbf{P}_{h_q} = \{(y_{h0}, u_{h0}, v_{h0}), \dots, (y_{hq}, u_{hq}, v_{hq})\}$, 肩部領域の画素をベクトル $\mathbf{P}_{s_p} = \{(y_{s0}, u_{s0}, v_{s0}), \dots, (y_{sp}, u_{sp}, v_{sp})\}$ にそれぞれ記録する。以上の処理は前処理とし、所要時間は約 107ms である (表 5.1 の一行目に参照)。

尚、この処理以外、SHAL 特徴記述法と類似度評価法には距離情報を利用していない。



(a) 頭頂部と肩部の切り出す深度範囲



(b) 入退室判定と画像撮影の方法

図 5.9 頭頂部と肩部の切り出す深度範囲と入外室判定と画像撮影の方法

また、可視化するため、背景領域の画素値は0とし、その他の画素値はカラー画像の対応値とした画像を生成できる（例えば：図 5.1，図 5.9 (b)）。

この手法で収集した俯瞰視画像データベースは「関連ページ」を参照されたい。

5.4.3 SHAL 法を用いた個人再識別システム

SHAL 法, 重み付きバタチャリア距離と奥剛一の手法を結合し, 実験用個人再識別システムを構築した (図 5.10 に参照). このシステムは, 俯瞰視カメラの画像・再識別結果を表示するウィンドウ (図 5.10 の上部), システムコントロールパネル (図 5.10 の中部) および可視化特徴量を表示するウィンドウ (図 5.10 の下部) で構成する.

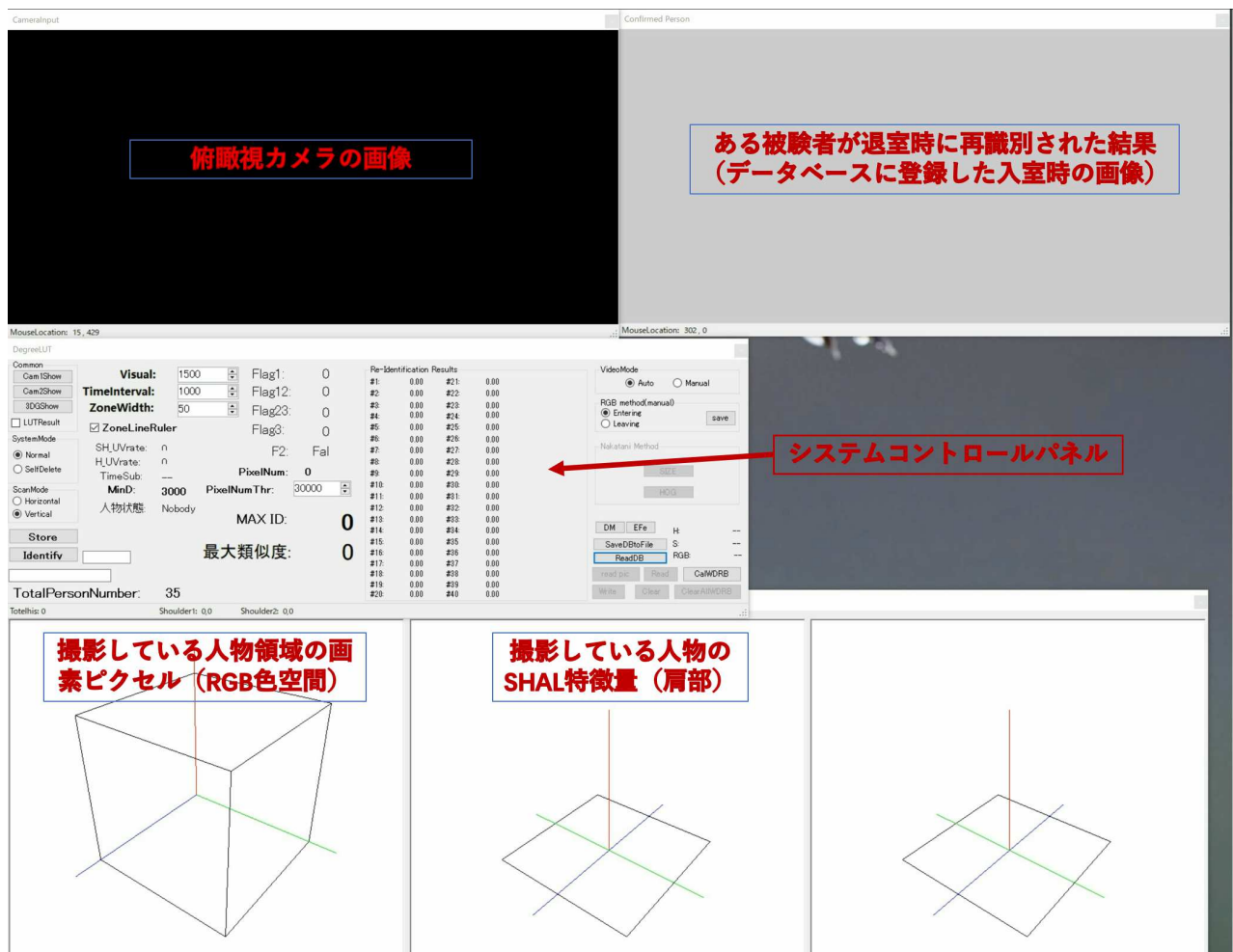


図 5.10 提案手法を用いた個人再識別システム

システムのコントロールパネルでは, データベースに登録した人数・再識別結果とする最大類似度と対応する ID 番号・退室人物とデータベースに登録したすべての人物との類似度の確認ができる (図 5.11 に参照).

人物が俯瞰視カメラの下を通過する際に計算した SHAL 特徴量はこのシステムで可視

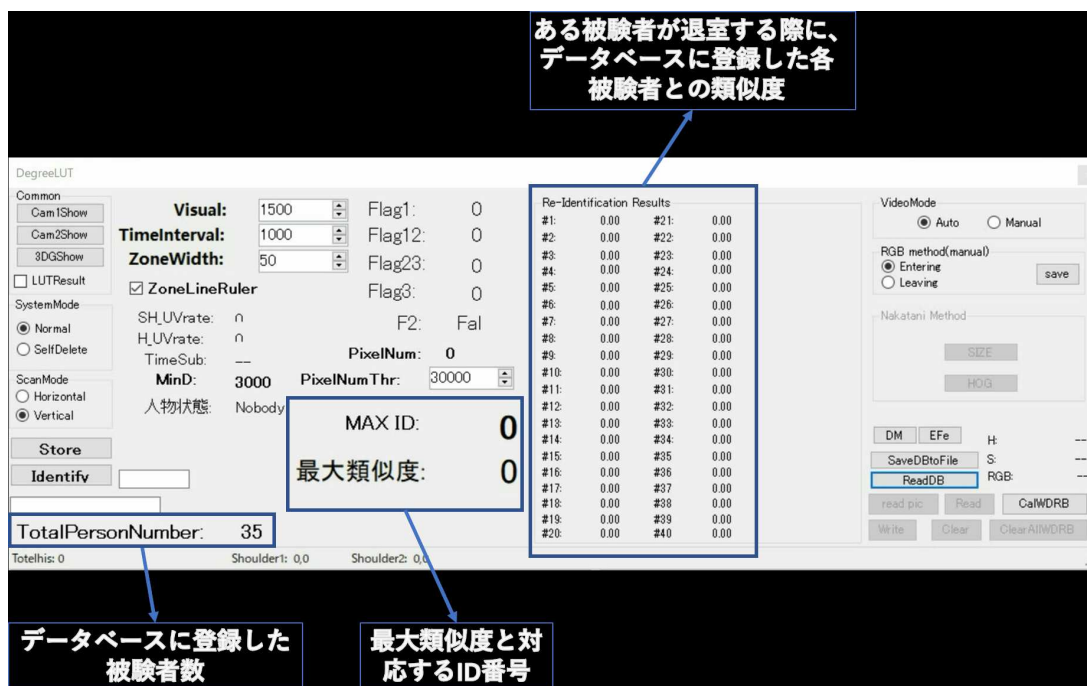


図 5.11 個人再識別システムのコントロールパネル

化することができる。図 5.12 は可視化した人物の **SHAL** 特徴量のサンプルである。

退室人物に対して、**SHAL** 特徴量を記述し、データベースのすべて既知人物の特徴量との類似度を計算する後、最大類似度を持つ *ID* 番号を見つけて再識別の結果とする。最大類似度と相応する *ID* 番号をコントロールパネルで表示する同時に、再識別結果の表示ウィンドウに再識別結果に相応する人物の入室時の画像も表示する（図 5.13 に参照）。

これからのすべての実験（Weka による比較実験を除き）はこの個人再識別システムを利用して行う。

5.4.4 私服を着る実験

被験者が 34 人（男性 14 人，女性 20 人），1 人当たり 10 回入室し，撮影順に 34 人 × 1 回入室 × 10 セットの **SHAL**(*id*) を作成し，データベースに登録した（図 5.14 (a) はデータベースの 1 セットである）。また，同じ順に被験者がそれぞれ 10 回退室し，340 サンプル（34 人 × 10 回退室）の **SHAL**(*x*) データセットを用意した（図 5.14 (b) は各被験者の退室時のサンプル画像である）。個人再識別実験は全部で 10 回行い，毎回異なる **SHAL**(*id*) のセットを利用し，340 サンプルの **SHAL**(*x*) と 3400 回照合した。被験者毎の再識別の平均成功率及び全員の再識別の平均成功率（ALL）は表 5.2 に纏めた。

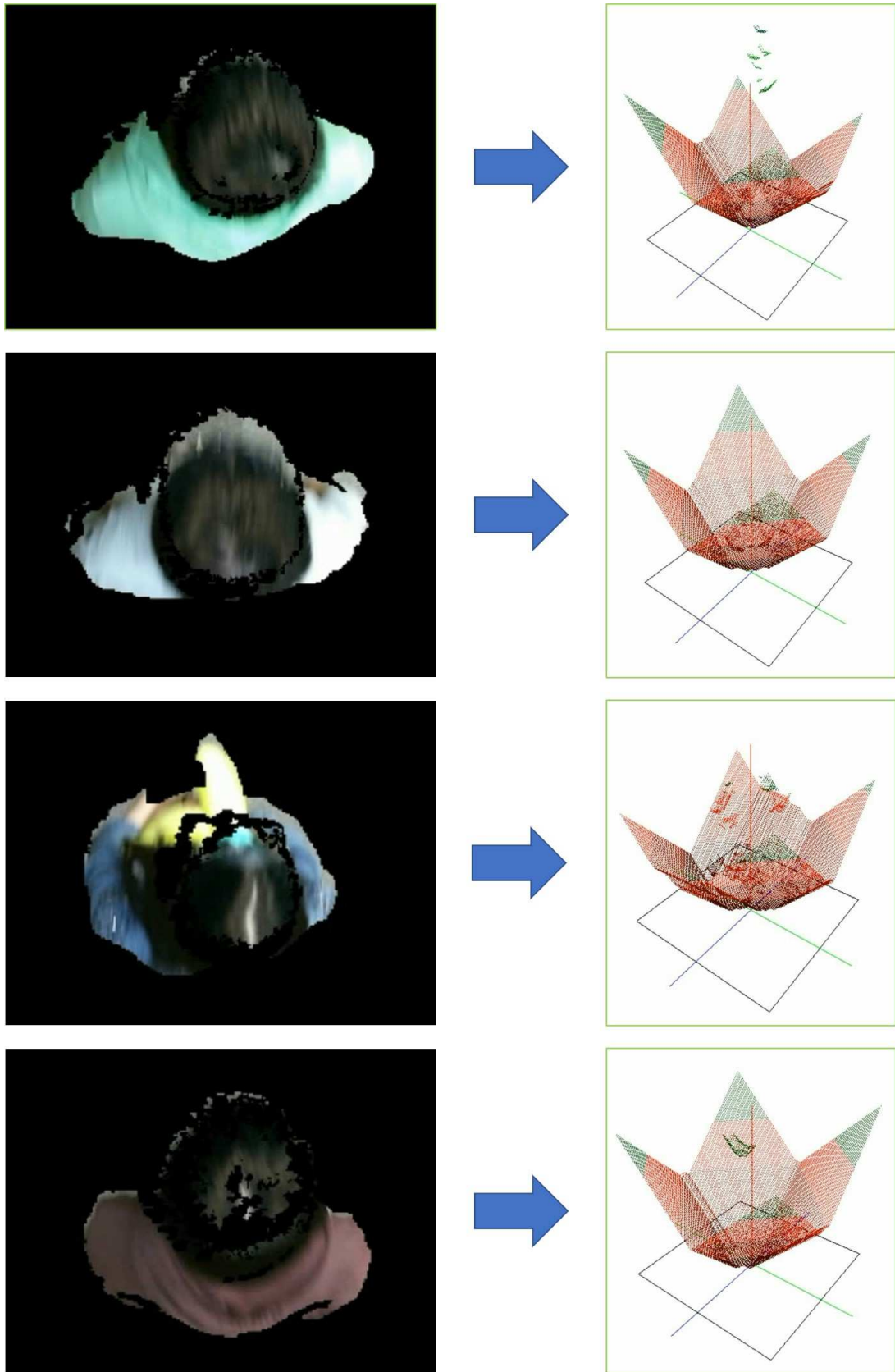


図 5.12 可視化した人物の SHAL 特徴量

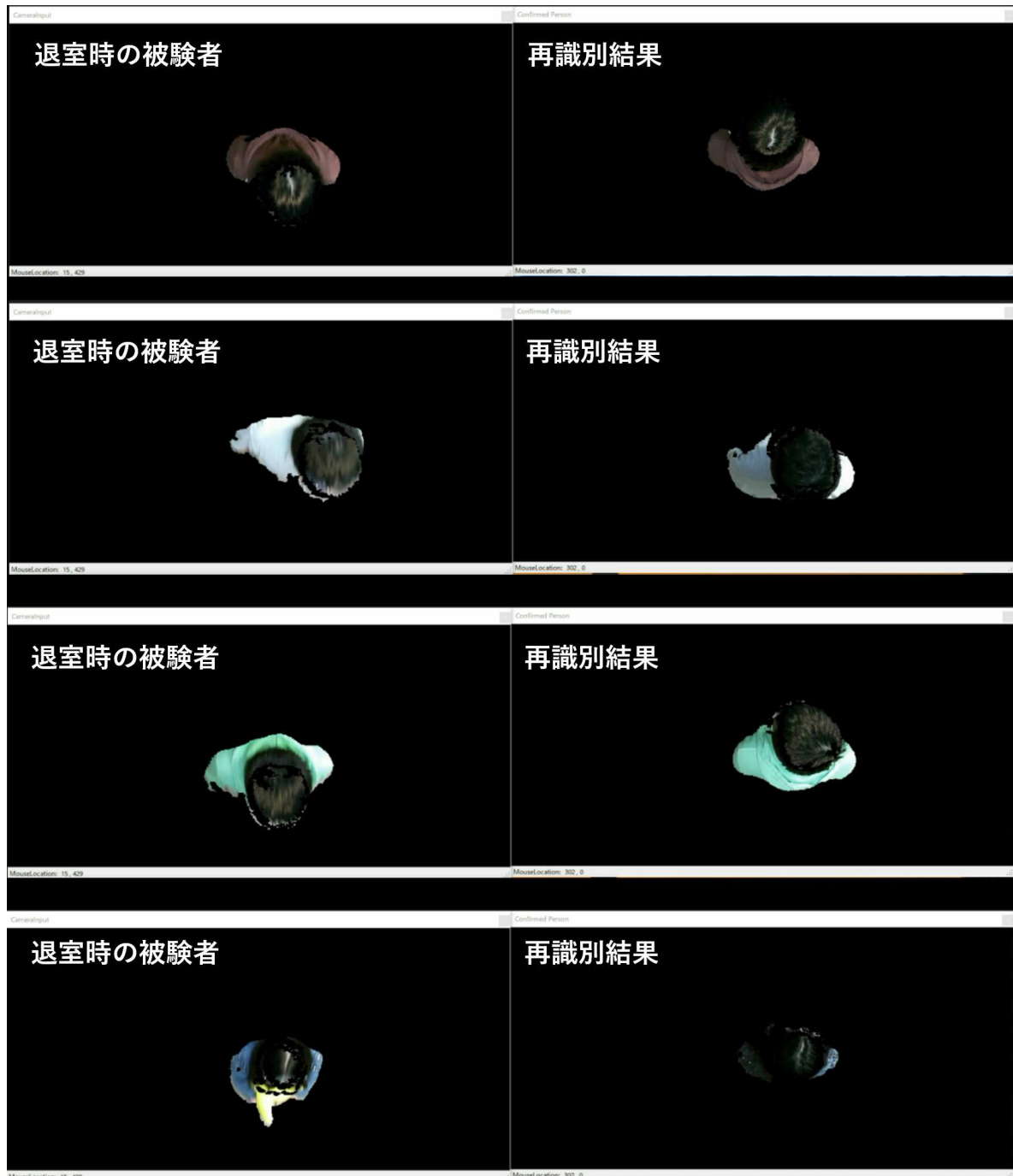
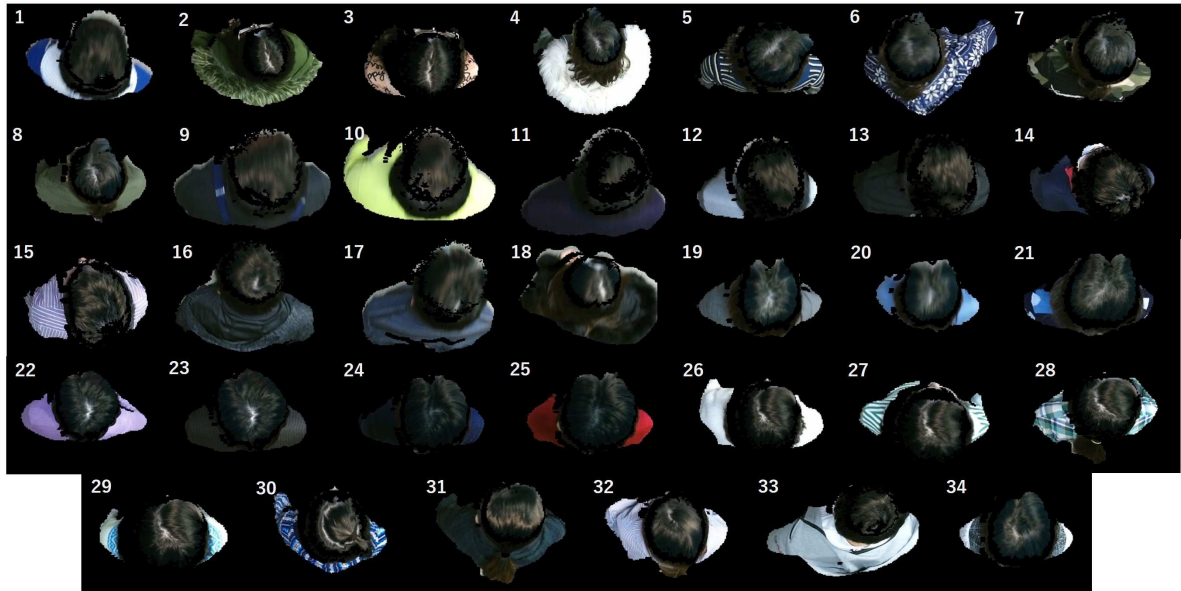
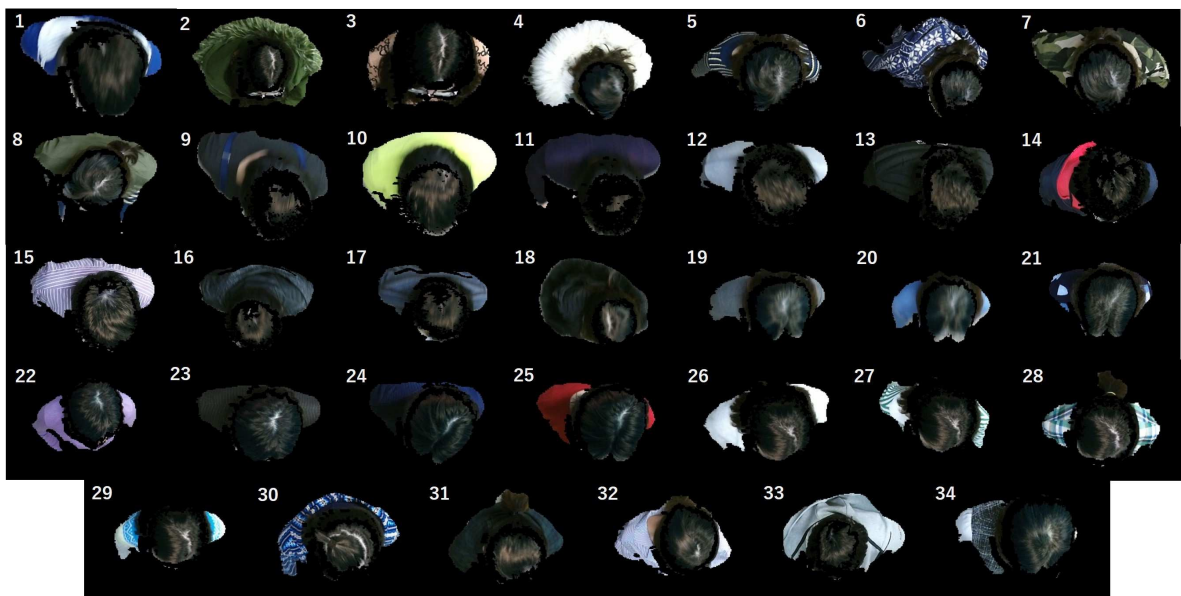


図 5.13 再識別結果に相応する人物の入室画像の表示



(a) 私服の入室画像セット



(b) 私服の退室画像セット

図 5.14 被験者の入室・退室のサンプル画像

表 5.2 私服を着る 10 セット分の平均再識別成功率 ($\alpha = 20$)

ID	1	2	3	4	5	6	7
成功率	90.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
ID	8	9	10	11	12	13	14
成功率	100.00%	90.00%	100.00%	72.00%	100.00%	100.00%	100.00%
ID	15	16	17	18	19	20	21
成功率	100.00%	75.00%	98.00%	100.00%	100.00%	100.00%	100.00%
ID	22	23	24	25	26	27	28
成功率	100.00%	100.00%	97.00%	100.00%	100.00%	77.00%	100.00%
ID	29	30	31	32	33	34	ALL
成功率	100.00%	100.00%	100.00%	100.00%	93.00%	93.00%	96.62%

表 5.2 の奇数行と偶数行はそれぞれに被験者の入室時の登録 ID 番号と再識別の平均成功率である。表 5.2 から、11 番、16 番と 27 番の被験者の平均成功率は 70% 台で、それ以外の被験者の平均成功率は 90% 以上となった。全体の平均成功率は 96.62% であった。11 番の被験者は退室時の移動スピードが速く、画像内の人物領域はぶれたので、色情報を正確に取れなかったことが、平均成功率が低下の原因と考えられる。16 番と 27 番の場合、頭頂部（髪の毛の色）と肩部（服の色）の色情報は共に類似であり、さらに体の見えも似ていることが誤認識の原因と考えられる。

5.4.5 同色の服を着る実験

被験者は 7 人である（男性:4 人、女性:3 人）。あらかじめ用意した 3 着の上着（黄緑色服、灰色服、黒色服）にそれぞれ着替えて、1 人当たり各色の服で 10 回入室し、撮影順に、(7 人 × 3 着の服 × 1 回入室 × 10 セット) の $\mathbf{SHAL}(id)$ を作成して、データベースに登録した。また、被験者が同じ順に 10 回退室し、服毎に 70 サンプル (7 人 × 10 回退室 × 同色の服) の $\mathbf{SHAL}(x)$ データセットを作成した (図 5.15 の (a) (b) (c) はそれぞれ黄緑色、灰色、黒色服の 1 セットの入室画像セットである)。

3 種類の服毎に同色の服を着るときの個人再識別実験は 10 回行い、毎回異なる $\mathbf{SHAL}(id)$ のセットを利用し、対応する服の 70 サンプルの $\mathbf{SHAL}(x)$ と 700 回照合した。3 種類の服別で被験者毎の平均成功率及び 2100 回の平均成功率 (ALL) は表 5.3 に纏めた。

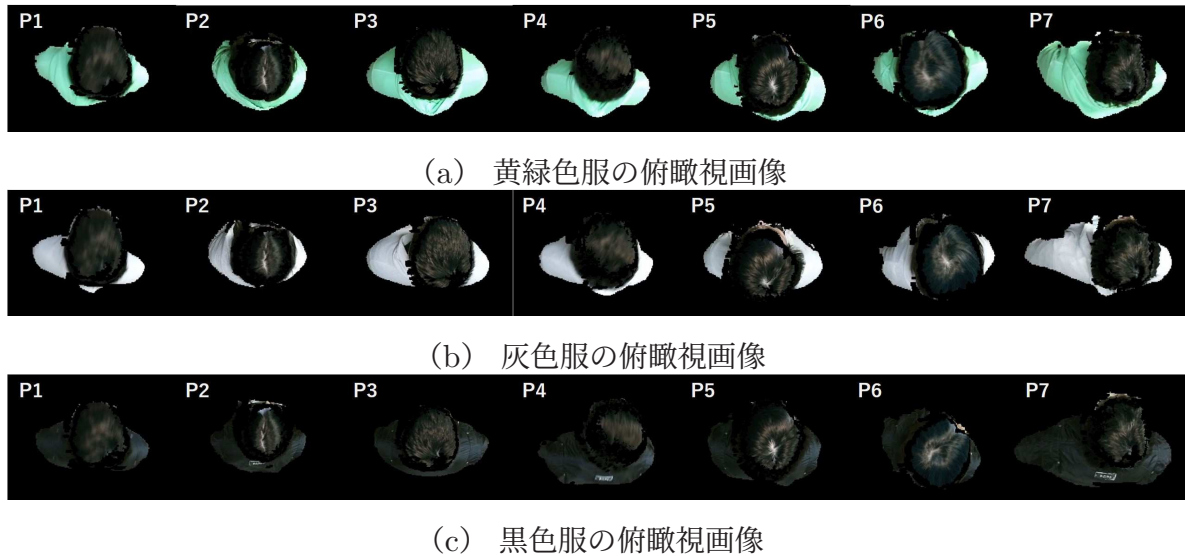


図 5.15 7 人の入室俯瞰視画像のサンプル

表 5.3 同色の服を着る 10 セット分の平均再識別成功率 ($\alpha = 20$)

ID	1	2	3	4	5	6	7
黄緑色服 成功率	63.00%	100.00%	97.00%	82.00%	95.00%	100.00%	73.00%
灰色服 成功率	91.00%	100.00%	100.00%	84.00%	100.00%	100.00%	100.00%
黒色服 成功率	82.00%	100.00%	100.00%	100.00%	100.00%	100.00%	97.00%
ALL	93.52%						

表 5.3 の一行目は被験者の ID 番号である。2 行目から 4 行目までは各被験者の平均成功率である。5 行目は全員の平均成功率である。

表 5.3 より、黄緑色服の 1 番と 7 番の成功率は最も低かった。撮影された 1 番と 7 番の被験者の俯瞰視画像内の見えがそれぞれ 3 番と 4 番に似ていることが成功率の低下原因と考えられる。それ以外の平均成功率は 80% 以上だったので、同じ色の服を着ても提案手法で識別できることを確認した。

5.4.6 提案手法の有効性の検証実験

「同色の服を着る実験」と同じ画像セットを用いて、色のヒストグラムによる色登録マップの拡張 (式 (5.4)), 及び色と輝度の重要度を考慮する重み (式 (5.5)) は、個人再識別の成功率向上に効果があることを確認した。結果は表 5.4 に纏めた。一列目 (None) は色登録マップを拡張しない場合、重要度を考慮しない場合 (重みはすべて 1) の結果である。二列目は色登録マップを拡張する (With Eq. (5.4)) が、重要度を考慮しない場合の結果である。三列目 (With Eq. (5.4) + (5.5)) は、すべてを考慮している提案手法の結果である。表 5.4 の結果より、認識率は、何も考慮しない場合、色登録マップを拡張する場合、色登録マップを拡張して重要度を考慮した場合で徐々に再識別成功率が上昇している。色登録マップの拡張と、色と輝度の重要度を考慮する提案手法の有効性を確認できた。

表 5.4 提案手法の有効性の検証実験結果 ($\alpha = 20$)

	None	With Eq. (5.4)	With Eq. (5.4) + (5.5)
成功率	84.43%	88.57%	93.52%

5.4.7 比較実験

Weka[158] を利用する比較実験

本研究と一番近い Nakatani R ら [134] の手法との比較実験を行った。Nakatani R らが図 2.8 (c) のような真上からの画像を利用した実験結果を示したが、本方式の実験には図 2.8 (b) に示す環境で収集した 5.4.4 項で利用した画像データセットを用いた (被験者: 34 人, 各被験者の入退室画像合計: 20 枚)。

比較用の特徴量の記述について、Nakatani R らの論文 [134] で一番高い認識率が得られた *size + style + color* の組み合わせのものとし、提案手法は SHAL 法で記述されたものとした。

特徴量間のマッチング手法について、Nakatani R はデータ解析・予測モデリングソフト Weka を利用しているので、まず、論文 [134] に記載しているパラメータで再インプリメントし、C4.5 アルゴリズムの分類器と、10 分割交差検証で、記述された 2 種類の特徴量を用いて比較実験を行った。平均成功率は表 5.5 の 2 行目と 3 行目に示し、両手法の成

功回数・失敗回数を表す混同行列をそれぞれ図 5.16 (a) と (b) に示す。混同行列の横軸と縦軸は被験者を同じ順番に表し、各行は該当被験者の 20 枚の画像セットを利用し、1 枚ずつすべて被験者とのマッチングを行った累積成功回数を示している。混同行列の対角線上にある数字は認識の成功回数であり、対角線外の数字は誤認識回数である。Nakatani R 手法の結果は対角線外の誤認識が多数となり、全体の再認識成功率は 40% 台であった。提案手法の場合は、対角線外の誤認識回数がほとんど 0 で、対角線上の認識成功回数が多く、全体の再認識成功率が 93.82% に達した。この比較実験結果から、提案手法は従来手法よりも有効であることを確認できた。

提案手法を利用する比較実験

表 5.5 の 4 行目には提案手法を用いた場合の平均成功率を示す。表 5.5 より、図 2.8 (b) に示す撮影状況の場合、SHAL 法で記述された特徴量は Nakatani R らの特徴量 (*size + style + color*) より再識別成功率が高く、重み付きバタチャリア距離での評価方法も Weka より再識別成功率が高いことを確認できた。したがって、被験者がカメラの真下を通るやカメラの下で一時停止して撮影が必要などの制限をかけていない提案システムはより幅広い個人再識別に利用可能と思われる。

表 5.5 比較実験結果

特徴量と (評価方法)	成功率
<i>size + style + color</i> と (Weka)	40.00%
SHAL と (Weka)	93.82%
SHAL と (重み付きバタチャリア距離)	96.62%

5.4.8 サンプリング間隔変化による安定性実験

この実験では、私服を着る実験に使われる 34 人の入室時に作成したデータベースの 1 セット、10 回退室時に作成したデータセットを利用し、計 340 回 (34 人 × 10 回退室) の個人再識別を行った。毎回の実験では、元画像から検出した人物領域に基づいて記録されている頭頂部の画素のベクトル $\mathbf{P}_{h_q} = \{(y_{h0}, u_{h0}, v_{h0}), \dots, (y_{hq}, u_{hq}, v_{hq})\}$ と、肩部の画素のベクトル $\mathbf{P}_{s_p} = \{(y_{s0}, u_{s0}, v_{s0}), \dots, (y_{sp}, u_{sp}, v_{sp})\}$ に対して、異なるサンプリング間隔で \mathbf{P}_{h_q} と \mathbf{P}_{s_p} から画素を選んで、新しい $\mathbf{P}_{h_{q'}}$ と $\mathbf{P}_{s_{p'}}$ ベクトルを生成した後、SHAL 特徴量を生成し、重み付きバタチャリア距離で再識別の実験を行った。



図 5.17 サンプル間隔変化による安定性実験結果 ($\alpha = 20$)

サンプリング間隔 0 ~ 500 ピクセルの再識別実験を行い、実験結果は図 5.17 に纏めて示す。図 5.17 の横軸はカッコ前が間隔 (interval) であり、カッコ内が $\mathbf{P}h_{q'}$ と $\mathbf{P}s_{p'}$ の概ねの合計画素数であり、単位が両方ともピクセルである。縦軸は再識別の成功率 (単位: パーセント) である。画像の解像度は Bazzani ら [129] の手法 ($128 \times 48 = 6144$ 画素) と、Xiang ら [130] の手法 ($64 \times 24 = 1536$ 画素) と同程度の低い解像度の場合でも再識別率がそれぞれ 97.00% と 96.18% となっていることを確認できた。

この俯瞰視画像データベースを撮影した俯瞰視カメラは床から 2.4 メートルの所に設定され、得られた一人分の頭頂部と肩部の合計画素数は概ね 60000 ピクセルであった (ピクセル数は個人差がある)。図 5.17 に示すように、70 ピクセル間隔でサンプリングして得られた概ね 880 ピクセルを用いた場合でも 95 % 程度の再識別率が得られた。透視投影の関係より、この場合の仮想カメラから床までの距離 $D_{CG} = \sqrt{\frac{60000}{880}} \times 2.4 = 19.8(m)$ となる。従って、実俯瞰視カメラは、床から 19.8 メートル程度設置されても 95 % 程度の再識別率が得られると考えられる。

また、エッジ情報を利用する手法 (例えば Nakatani R[134]) にとって、人物領域の解像度がエッジ情報の検出が困難な程度 (約 $300 \times 200 = 60000$ 画素) までに下がると、HOG 特徴量の記述はできなくなる。一方、提案手法では、人物領域の画素の色情報 (UV 色平面内の位置と登録色の数) と画素数を利用して個人の特徴情報を記述するので、解像度が下がると、色のヒストグラムにより求めた各登録色の影響範囲も計測対象全員について同程度縮小されるが、登録色の数はほとんど変化がない。この為、より低い解像度の場合でも提案手法は利用可能と考えられる。

5.4.9 頭頂部と肩部を一つの領域とする場合の個人再識別

SHAL 法では, depth 情報と RGB 画像を同時に獲得できる RGB-D カメラを利用して奥 剛一の手法により人物の俯瞰視画像を撮影している. しかし, 3 章の CDF 法と 4 章の WDRB 法で利用した一般の市販 Web カメラを利用する場合, depth 情報を利用できず, 人物の RGB 俯瞰視画像しか利用できなくなる. この時, 人物の領域は頭頂部と肩部をそれぞれ分けて記述できなく, 一つの領域として利用することになる. SHAL 法の応用範囲を拡大するために, RGB-D カメラの代わりに一般市販 Web カメラも利用できる人物の頭頂部と肩部を一つの領域とした特徴量を記述する方法を試みる. この方法を **Object identification Feature (OiF)** と名付ける. OiF 特徴量の記述手順が人物の頭頂部と肩部を分ける場合と同様であるが, 頭頂部と肩部を一つの領域とし, 3次元特徴空間の数は4つから2つまで減る. 変更点を纏めて以下の5点がある.

- 1: 頭頂部の画素をベクトル \mathbf{Ph}_q と肩部の画素をベクトル \mathbf{Ps}_p の合計より, 一つの UV 色平面内の色登録マップ $\mathbf{C}'(U, V)$ を生成する.
- 2: \mathbf{Ph}_q と \mathbf{Ps}_p の合計より, 一つの人物領域の同じ色の輝度 y の合計値分布 $\mathbf{B}(U, V)$ を生成する.
- 3: 色登録マップ $\mathbf{C}'(U, V)$ から変換する色距離マップ $\mathbf{D}(U, V)$ を生成する.
- 4: 色の影響範囲が拡張された色登録マップ $\mathbf{C}(U, V)$ を生成する. $\mathbf{C}(U, V)$ と色距離マップ $\mathbf{D}(U, V)$ との融合処理による色融合マップ $\mathbf{Cf}(U, V)$ を生成する.
- 5: Grayzone 内側の全登録色の累積度数の和の割合を W_y とし, 式 (5.10) より計算する. 式 (5.10) では, N_y と N_{uv} はそれぞれ Grayzone の内側に登録している無彩色の全登録色の累積度数の和と, Grayzone 外側に登録している有彩色の全登録色の累積度数の和である.
- 6: 人物の OiF 特徴量は $\mathbf{OiF} = \{\mathbf{B}(U, V), \mathbf{Cf}(U, V), W_y\}$ となる.
- 7: 重み付きバタチャリヤ距離による輝度 y の合計値分布の類似度 $D_y(x, id)$, 色融合マップの類似度 $D_{uv}(x, id)$ および OiF 特徴量間の類似度 $D_b(\mathbf{OiF}(x), \mathbf{OiF}(id))$ の計算式は, それぞれ式 (5.11), 式 (5.12) と式 (5.13) となる.

$$W_y = N_y / (N_y + N_{uv}) \quad (5.10)$$

$$D_y(x, id) = \sum_{u, v \in \{-128, \dots, 127\}} \sqrt{\mathbf{B}(U, V)(x) \mathbf{B}(U, V)(id)} \quad (5.11)$$

$$D_{uv}(x, id) = \sum_{u, v \in \{-128, \dots, 127\}} \sqrt{\mathbf{Cf}(U, V)(x) \mathbf{Cf}(U, V)(id)} \quad (5.12)$$

$$\begin{aligned} D_b(\mathbf{OiF}(x), \mathbf{OiF}(id)) = & \\ & (W_{\mathbf{y}}(x) + W_{\mathbf{y}}(id)) \times D_y(x, id) \\ & + (2 - W_{\mathbf{y}}(x) - W_{\mathbf{y}}(id)) \times D_{uv}(x, id) \end{aligned} \quad (5.13)$$

私服を着る個人再識別実験

OiF 法が SHAL 法より、特徴量の記述分布の個数は半分に減ったので、個人再識別への影響はどれだけ出るかを確認するために、SHAL 法を用いた実験（5.4.4 項に参照）と同様な条件下で私服を着る個人再識別実験を行った。実験に利用する人物の俯瞰視画像セットも同様である（図 5.14 に参照）。被験者毎の再識別の平均成功率及び全員の再識別の平均成功率（ALL）は表 5.6 に纏めた。

表 5.6 OiF を用いた、私服を着る 10 セット分の平均再識別成功率（ $\alpha = 20$ ）

ID	1	2	3	4	5	6	7
成功率	99.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
ID	8	9	10	11	12	13	14
成功率	100.00%	65.00%	91.00%	72.00%	98.00%	39.00%	95.00%
ID	15	16	17	18	19	20	21
成功率	96.00%	53.00%	100.00%	84.00%	88.00%	86.00%	88.00%
ID	22	23	24	25	26	27	28
成功率	100.00%	100.00%	100.00%	100.00%	98.00%	74.00%	100.00%
ID	29	30	31	32	33	34	ALL
成功率	100.00%	100.00%	93.00%	100.00%	99.00%	84.00%	91.24%

表 5.6 の奇数行と偶数行はそれぞれに被験者の入室時の登録 ID 番号と再識別の平均成功率である。表 5.6 の結果と表 5.2 と比較すると、OiF 法の再識別の平均成功率が SHAL 法の 96.62% から 91.24% に低下したことは分かった。被験者の頭頂部と肩部を一つの領域として利用することで、人物の体格情報の記述が二つの領域を分ける場合より精度が悪いため、再識別精度が低下した原因と考えられる。OiF は特徴量の記述分布の個数が SHAL の半分であり、システムメモリの使用量・特徴量の生成時間・マッチング時間がそれぞれ半分に短縮できたメリットがある。また、再識別率は 9 割以上であり、SHAL より

若干低くなるが，一般市販 Web カメラで撮影した俯瞰視画像を利用する場合の個人再識別に応用できることが分かった。

5.5 まとめ

本論文では，人物の俯瞰視画像を対象とし，頭頂部と肩部の画素情報を利用して個人の特徴を表現できる特徴記述法と，色と輝度の重要度を考慮した重み付きバタチャリヤ距離の評価法を提案した．本論文の個人再識別システムを用いて，多数の検証実験及び比較実験を行い，提案手法の有効性と安定性を確認した．また，応用範囲を拡大するために，RGB-D カメラの代わりに一般市販 Web カメラも利用できる人物の頭頂部と肩部を一つの領域とした特徴量を記述する方法も試み，個人再識別の実験を行った．人物の頭頂部と肩部を分ける場合より，個人再識別の成功率が若干低下したので，頭頂部と肩部を分ける SHAL 法は個人再識別により有効的であることを確認した．また，物体の再識別実験（付録 C に参照）より，OiF 法は物体再識別にも利用できることを確認した．

第 6 章

おわりに

6.1 本論文のまとめ

本論文では、プライバシーを考慮した個人再識別を実現するために、1) 短時間内に人物の服と髪の毛の色は変わらないこと、2) 俯瞰視カメラから地面までの距離はほぼ一定であると仮定して、個人再識別できる特徴記述法を複数提案した。

人物全身の正面画像或いは側面画像の利用を避けるために、人物の顔が映らない俯瞰視カメラシステムを構築し、人物の頭の上から俯瞰視画像を撮影する。

まず、第 3 章では、人物の頭頂部と肩部、この二つの領域の RGB 色情報を利用する「色弁別特徴 Color Distinctiveness Feature (CDF)」を提案した。人物の俯瞰カラー画像を撮影するために、市販の Web カメラを利用して俯瞰視システムを構築した。色弁別特徴 (CDF) では、まず、人物の肩部をメイン領域、頭頂部をサポート領域とし、人物の俯瞰視画像から手動で選択する。そして、メイン領域とサポート領域の各画素の RGB 色情報を利用して、ベイズ定理に基づき、3次元 RGB 色空間内で色弁別性を計算する。この色弁別性を持つ 3 次元色空間を該当人物の 4 次元色弁別特徴量 (CDF) とし、俯瞰視システムより撮影した人物の俯瞰視画像を利用して個人再識別を行う。再識別の際に、未知人物の俯瞰視画像から計算した CDF とデータベースに登録したすべての既知人物の CDF との類似度評価を行う。色弁別特徴量は 3 次元色空間内の分布であり、類似度の計算には二つの分布間の類似度評価に対応できるバタチャリア距離を利用する。また、バタチャリア距離での類似度の正確性を確認するために、バタチャリア距離の他に、ハミング距離も類似度の評価に利用する。事前に撮影した人物の俯瞰視画像からメイン領域とサポート領域を複数回選択する個人再識別実験や、Weka を利用する比較実験により、「色弁別特徴 (CDF)」は俯瞰視画像を利用する個人再識別に利用できることを確認した。色弁

別特徴 (CDF) によって色の種類数が異なる特徴量間の識別は可能であるが、色の種類数が同じ場合は、識別は不可能という課題が残った。

第4章では、色弁別特徴 (CDF) の課題を解決するために、色のヒストグラムも利用する新たな手法「水滴レンダリングボックス (WDRB)」を提案した。WDRB の処理は水滴が紙を濡らす現象を模倣している。WDRB 法の特徴記述は、(1) 色登録、(2) 色登録マップから距離マップへの変換、(3) 登録した色の強調、という3つの段階に分けられる。色登録では、人物領域の画素の RGB 情報を利用し、3次元色空間内で該当色をマップし、色登録マップを得る。色登録マップから距離マップへの変換により、色登録マップに登録したすべての色の間の相対的な距離関係を記述する距離マップが得られる。登録した色の強調では、登録した各色のヒストグラムを利用して、3次元色空間内に各色が占める範囲を拡大し、4次元の強調マップを得る。この強調マップは WDRB の特徴量として利用する。人物の俯瞰視画像から手動で人物領域を選択することで、CDF 法との比較実験を行い、WDRB 法はより有効性を持つことを証明した。また、物体の再識別実験より、WDRB 法は物体再識別にも利用できることを確認した。さらに、Weka を利用し、異なる光源下で撮影した人物の俯瞰視画像セットを用いて、他人の手法との個人再識別の比較実験より、WDRB 法が明暗変化に対する頑健性は従来手法より高いことも証明した。WDRB 法が画素の RGB 色情報を利用しているため、明るさが大幅に変化される場合、再識別はよく失敗する問題点が残っている。また、画像上の人物の領域の画素数は人物の体格情報であり、WDRB 法では手動で人物領域を選択するのが人物の体格情報の反映が不十分である課題は残った。

第5章では、WDRB 法の課題を解決するために、SHAL (Shoulders and Head Learning) 法を提案した。SHAL 法では、人物の体格情報を十分に記述するために、RGB-D カメラを利用し、俯瞰視画像から人物の頭頂部領域と肩部領域のセグメンテーションを行った。輝度の変化が色に影響しないように、RGB 色空間の代わりに、輝度情報と色情報が単独に表示する YUV 色空間を利用する。特徴量を記述する時に、UV 色平面内に、頭頂部と肩部のそれぞれの画素について、1) 色のヒストグラムと、同じ色の輝度 y の合計値の分布を求める。2) 色の種類数と色空間内の各色の相対的な距離関係を反映するため、色空間におけるユークリッド距離変換を行う。3) 無彩色画素の割合を求める。SHAL 特徴量間の類似度を評価するために、無彩色画素の割合を重みとして、重み付きバタチャリア距離を提案した。私服を着る個人再識別実験、同色の服を着る個人再識別実験および Weka を利用する比較実験より、個人再識別における SHAL 法の有効性を証明した。さらに、応用範囲を拡大するために、RGB-D カメラの代わりに一般市販 Web カメラも利用できる人物の頭頂部と肩部を一つの領域とした特徴量を記述する方法 OiF も試み、個

人の再識別の実験を行った。特徴量の個数が半分減ったことにより、人物の頭頂部と肩部を分ける場合と比べて、個人再識別の成功率は若干低下した。しかし、システムメモリの使用量・特徴量の生成時間・マッチング時間がそれぞれ半分に短縮できた上、一般市販 Web カメラで撮影した俯瞰視画像を利用する場合の個人再識別に利用できることも検証できた。

付録 A, 付録 B と付録 C では、CDF 法を用いた物体の再識別実験、WDRB 法と CDF 法との物体再識別の比較実験、頭頂部と肩部を一つの領域とした特徴量を記述する方法 OiF を用いて商品の再識別実験をそれぞれ行った。これらの実験より、CDF 法、WDRB 法および OiF 法は物体の再識別への利用可能性を検証できた。

6.2 今後の課題

俯瞰視画像のブレについて：俯瞰視システムの下を通った人物の走行スピードおよび利用されるカメラの仕様によって、撮影した俯瞰視画像はブレる場合がある。ブレ画像では、人物領域の色情報だけでなく、人物の体格情報も正確に反映できなくなる。提案手法の現状では、画像が少しぶれている場合、再識別精度への影響は出なかったが、大きくぶれている場合は、提案手法では対応できないと考えられる。このため、ぶれ・ノイズの効率的な解消手法について考察が必要である。

再識別成功率を低下させない頭頂部と肩部を一つの領域とする方法について：5.4.9 項「頭頂部と肩部を一つの領域とする方法」での個人再識別と付録 C 「OiF 法による物体の再識別」で物体再識別実験を説明した。頭頂部と肩部を一つの領域としたことで、再識別成功率が少し低下するが、メモリの消費量・特徴量の処理（生成・マッチング）時間は半減できるメリットがある。個人再識別システムにとって、リアルタイムで処理するためには、メモリの消費量・特徴量の処理時間の半減は非常に魅力的なポイントである。このため、頭頂部と肩部を一つの領域としたものも、人物の体格情報を頭頂部と肩部に分ける場合と同様、或いはそれ以上の識別能力を持つ記述方法についても考察が必要である。

参考文献

- [1] Wenyi Zhao, Rama Chellappa, P Jonathon Phillips, and Azriel Rosenfeld. Face recognition: A literature survey. *ACM computing surveys (CSUR)*, Vol. 35, No. 4, pp. 399–458, 2003.
- [2] AS Tolba, AH El-Baz, and AA El-Harby. Face recognition: A literature review. *International Journal of Signal Processing*, Vol. 2, No. 2, pp. 88–103, 2006.
- [3] Yoshio IWAI, Shihong LAO, Osamu YAMAGUCHI, and Takatsugu HIRAYAMA. A survey on face detection and face recognition. *CVIM2005*, Vol. 149, pp. 343–368, May 2005.
- [4] Yingjie Wang and Chin-Seng Chua. Face recognition from 2d and 3d images using 3d gabor filters. *Image and Vision Computing*, Vol. 23, No. 11, pp. 1018 – 1028, 2005.
- [5] Timo Ojala, Matti Pietikainen, and David Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In *Proceedings of 12th International Conference on Pattern Recognition*, Vol. 1, pp. 582–585. IEEE, 1994.
- [6] Timo Ojala, Matti Pietikainen, David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, Vol. 29, No. 1, pp. 51 – 59, 1996.
- [7] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, Vol. 2, pp. 1150–1157. IEEE, 1999.
- [8] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Computer vision and image understanding*, Vol. 110, No. 3, pp. 346–359, 2008.
- [9] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human

- detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, Vol. 1, pp. 886–893. IEEE, 2005.
- [10] Gabriel Takacs, Vijay Chandrasekhar, Sam Tsai, David Chen, Radek Grzeszczuk, and Bernd Girod. Unified real-time tracking and recognition with rotation-invariant fast features. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 934–941. IEEE, 2010.
- [11] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features. In *European conference on computer vision*, pp. 778–792. Springer, 2010.
- [12] Stefan Leutenegger, Margarita Chli, and Roland Y Siegwart. Brisk: Binary robust invariant scalable keypoints. In *2011 International conference on computer vision*, pp. 2548–2555. IEEE, 2011.
- [13] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *2011 International conference on computer vision*, pp. 2564–2571. IEEE, 2011.
- [14] Alexandre Alahi, Raphael Ortiz, and Pierre Vandergheynst. Freak: Fast retina keypoint. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 510–517. IEEE, 2012.
- [15] Vladimir Vapnik. Pattern recognition using generalized portrait method. *Automation and remote control*, Vol. 24, pp. 774–780, 1963.
- [16] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144–152, 1992.
- [17] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, Vol. 20, No. 3, pp. 273–297, 1995.
- [18] Bernd Heisele, Thomas Serre, and Tomaso Poggio. A component-based framework for face detection and identification. *International Journal of Computer Vision*, Vol. 74, No. 2, pp. 167–181, 2007.
- [19] Hongjun Jia and Aleix M Martinez. Support vector machines in face recognition with occlusions. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 136–141. IEEE, 2009.
- [20] Kyunghee Lee, Yongwha Chung, and Hyeran Byun. Svm-based face verification with feature set of small size. *Electronics Letters*, Vol. 38, No. 15, pp. 787–789,

- 2002.
- [21] Yap Wooi Hen, Marzuki Khalid, and Rubiyah Yusof. Face verification with gabor representation and support vector machines. In *First Asia International Conference on Modelling & Simulation (AMS'07)*, pp. 451–459. IEEE, 2007.
 - [22] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, Vol. 3, No. 1, pp. 71–86, 1991.
 - [23] Peter N. Belhumeur, João P Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 19, No. 7, pp. 711–720, 1997.
 - [24] Zhimin Cao, Qi Yin, Xiaoou Tang, and Jian Sun. Face recognition with learning-based descriptor. In *2010 IEEE Computer society conference on computer vision and pattern recognition*, pp. 2707–2714. IEEE, 2010.
 - [25] Priyanka Wagh, Roshani Thakare, Jagruti Chaudhari, and Shweta Patil. Attendance system based on face recognition using eigen face and pca algorithms. In *2015 International Conference on Green Computing and Internet of Things (ICGGIoT)*, pp. 303–308. IEEE, 2015.
 - [26] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pp. 23–37. Springer, 1995.
 - [27] Mian Zhou and Hong Wei. Face verification using gaborwavelets and adaboost. In *18th International Conference on Pattern Recognition (ICPR'06)*, Vol. 1, pp. 404–407. IEEE, 2006.
 - [28] Xiaogang Wang, Cha Zhang, and Zhengyou Zhang. Boosted multi-task learning for face verification with applications to web image and video search. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 142–149. IEEE, 2009.
 - [29] Guosheng Hu, Yongxin Yang, Dong Yi, Josef Kittler, William Christmas, Stan Z Li, and Timothy Hospedales. When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition. In *Proceedings of the IEEE international conference on computer vision workshops*, pp. 142–150, 2015.
 - [30] Asifullah Khan, Anabia Sohail, Umme Zahoora, and Aqsa Saeed Qureshi. A

- survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, pp. 1–62, 2019.
- [31] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. 2008.
- [32] Tianyue Zheng, Weihong Deng, and Jiani Hu. Cross-age lfw: A database for studying cross-age face recognition in unconstrained environments. *arXiv preprint arXiv:1708.08197*, 2017.
- [33] Lior Wolf, Tal Hassner, and Itay Maoz. Face recognition in unconstrained videos with matched background similarity. In *CVPR 2011*, pp. 529–534. IEEE, 2011.
- [34] P Jonathon Phillips, Hyeonjoon Moon, Syed A Rizvi, and Patrick J Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, No. 10, pp. 1090–1104, 2000.
- [35] Dong Chen, Xudong Cao, Liwei Wang, Fang Wen, and Jian Sun. Bayesian face revisited: A joint formulation. In *European conference on computer vision*, pp. 566–579. Springer, 2012.
- [36] Neeraj Kumar, Alexander C Berg, Peter N Belhumeur, and Shree K Nayar. Attribute and simile classifiers for face verification. In *2009 IEEE 12th international conference on computer vision*, pp. 365–372. IEEE, 2009.
- [37] Yi Sun, Xiaogang Wang, and Xiaoou Tang. Deep learning face representation from predicting 10,000 classes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1891–1898, 2014.
- [38] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815–823, 2015.
- [39] Masahiro Kawagoe and Akio Tojo. Fingerprint pattern classification. *Pattern recognition*, Vol. 17, No. 3, pp. 295–303, 1984.
- [40] Babu M Mehtre. Fingerprint image analysis for automatic identification. *Machine Vision and Applications*, Vol. 6, No. 2-3, pp. 124–139, 1993.
- [41] Lin Hong. Automatic personal identification using fingerprints. 1999.
- [42] Neil Yager and Adnan Amin. Fingerprint verification based on minutiae features: a review. *Pattern Analysis and Applications*, Vol. 7, No. 1, pp. 94–113,

-
- 2004.
- [43] http://research.nii.ac.jp/~iechizen/official/BiometricJammerE_rev.pdf.
- [44] FA Afsar, M Arif, and M Hussain. Fingerprint identification and verification system using minutiae matching. In *National Conference on Emerging Technologies*, Vol. 1, pp. 141–146, 2004.
- [45] Neil Yager and Adnan Amin. Fingerprint verification based on minutiae features: a review. *Pattern Analysis and Applications*, Vol. 7, No. 1, pp. 94–113, 2004.
- [46] Daniel Peralta, Mikel Galar, Isaac Triguero, Daniel Paternain, Salvador García, Edurne Barrenechea, José M Benítez, Humberto Bustince, and Francisco Herrera. A survey on fingerprint minutiae-based local matching for verification and identification: Taxonomy and experimental evaluation. *Information Sciences*, Vol. 315, pp. 67–87, 2015.
- [47] George Bebis, Taisa Deaconu, and Michael Georgiopoulos. Fingerprint identification using delaunay triangulation. In *Proceedings 1999 International Conference on Information Intelligence and Systems (Cat. No. PR00446)*, pp. 452–459. IEEE, 1999.
- [48] Lin Hong and Anil Jain. Integrating faces and fingerprints for personal identification. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 20, No. 12, pp. 1295–1307, 1998.
- [49] Anil K Jain, Lin Hong, and Yatin Kulkarni. A multimodal biometric system using fingerprint, face and speech. In *2nd Int’l Conf. AVBPA*, Vol. 10, 1999.
- [50] Unsang Park, Sharath Pankanti, and Anil K Jain. Fingerprint verification using sift features. In *Biometric Technology for Human Identification V*, Vol. 6944, p. 69440K. International Society for Optics and Photonics, 2008.
- [51] Ali Ismail Awad and Kensuke Baba. Evaluation of a fingerprint identification algorithm with sift features. In *2012 IIAI International Conference on Advanced Applied Informatics*, pp. 129–132. IEEE, 2012.
- [52] Shervin Minaee and Yao Wang. Fingerprint recognition using translation invariant scattering network. In *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–6. IEEE, 2015.
- [53] Wei Shu and David Zhang. Automated personal identification by palmprint.

- OptEn*, Vol. 37, pp. 2359–2362, 1998.
- [54] Dapeng Zhang and Wei Shu. Two novel characteristics in palmprint verification: datum point invariance and line feature matching. *Pattern recognition*, Vol. 32, No. 4, pp. 691–702, 1999.
- [55] Wei Shu, Gang Rong, Zhaoqi Bian, and David Zhang. Automatic palmprint verification. *International Journal of Image and Graphics*, Vol. 1, No. 01, pp. 135–151, 2001.
- [56] Wenxin Li, David Zhang, and Zhuoqun Xu. Palmprint identification by fourier transform. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 16, No. 04, pp. 417–432, 2002.
- [57] David Zhang, Wai-Kin Kong, Jane You, and Michael Wong. Online palmprint identification. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 25, No. 9, pp. 1041–1050, 2003.
- [58] Guangming Lu, David Zhang, and Kuanquan Wang. Palmprint recognition using eigenpalms features. *Pattern Recognition Letters*, Vol. 24, No. 9-10, pp. 1463–1467, 2003.
- [59] Manisha P Dale, Madhuri A Joshi, and Neena Gilda. Texture based palmprint identification using dct features. In *2009 Seventh International Conference on Advances in Pattern Recognition*, pp. 221–224. IEEE, 2009.
- [60] Abdallah Meraoumia, Salim Chitroub, and Ahmed Bouridane. Efficient person identification by fusion of multiple palmprint representations. In *International Conference on Image and Signal Processing*, pp. 182–191. Springer, 2010.
- [61] Abdallah Meraoumia, Salim Chitroub, and Ahmed Bouridane. Fusion of finger-knuckle-print and palmprint for an efficient multi-biometric system of person recognition. In *2011 IEEE International Conference on Communications (ICC)*, pp. 1–5. IEEE, 2011.
- [62] Shanmukhappa A Angadi and Sanjeevakumar M Hatture. Biometric person identification system: a multimodal approach employing spectral graph characteristics of hand geometry and palmprint. *International Journal of Intelligent Systems and Applications*, Vol. 8, No. 3, p. 48, 2016.
- [63] https://www.hitachi-hightech.com/hmt/product_detail/?pn=hmt024.
- [64] <https://www.itmedia.co.jp/enterprise/articles/1802/14/news082.html>.

-
- [65] Naoto Miura, Akio Nagasaka, and Takafumi Miyatake. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine vision and applications*, Vol. 15, No. 4, pp. 194–203, 2004.
- [66] Naoto Miura, Akio Nagasaka, and Takafumi Miyatake. Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEICE TRANSACTIONS on Information and Systems*, Vol. 90, No. 8, pp. 1185–1194, 2007.
- [67] Xianjing Meng, Gongping Yang, Yilong Yin, and Rongyang Xiao. Finger vein recognition based on local directional code. *Sensors*, Vol. 12, No. 11, pp. 14937–14952, 2012.
- [68] Chih-Lung Lin and Kuo-Chin Fan. Biometric verification using thermal images of palm-dorsa vein patterns. *IEEE Transactions on Circuits and systems for Video Technology*, Vol. 14, No. 2, pp. 199–213, 2004.
- [69] Yi-Bo Zhang, Qin Li, Jane You, and Prabir Bhattacharya. Palm vein extraction and matching for personal authentication. In *International Conference on Advances in Visual Information Systems*, pp. 154–164. Springer, 2007.
- [70] Yingbo Zhou and Ajay Kumar. Contactless palm vein identification using multiple representations. In *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pp. 1–6. IEEE, 2010.
- [71] Ajay Kumar and Yingbo Zhou. Human identification using finger images. *IEEE Transactions on image processing*, Vol. 21, No. 4, pp. 2228–2244, 2011.
- [72] Yingbo Zhou and Ajay Kumar. Human identification using palm-vein images. *IEEE transactions on information forensics and security*, Vol. 6, No. 4, pp. 1259–1274, 2011.
- [73] George R Doddington. Speaker recognition-identifying people by their voices. *Proceedings of the IEEE*, Vol. 73, No. 11, pp. 1651–1664, 1985.
- [74] Sadaoki Furui. An overview of speaker recognition technology. In *Automatic speech and speaker recognition*, pp. 31–56. Springer, 1996.
- [75] John Makhoul. Linear prediction: A tutorial review. *Proceedings of the IEEE*, Vol. 63, No. 4, pp. 561–580, 1975.
- [76] Steven Davis and Paul Mermelstein. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE transactions on acoustics, speech, and signal processing*, Vol. 28, No. 4, pp. 357–366,

- 1980.
- [77] Hynek Hermansky. Perceptual linear predictive (plp) analysis of speech. *the Journal of the Acoustical Society of America*, Vol. 87, No. 4, pp. 1738–1752, 1990.
- [78] M. Jack Xuedong Huang and Y. Ariki. Hidden markov models for speech recognition. *ISBN 978-0-7486-0162-2.2*, 1990.
- [79] Vibha Tiwari. Mfcc and its applications in speaker recognition. *International journal on emerging technologies*, Vol. 1, No. 1, pp. 19–22, 2010.
- [80] Medha Pandit and Josef Kittler. Feature selection for a dtw-based speaker verification system. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, Vol. 2, pp. 769–772. IEEE, 1998.
- [81] Md Rashidul Hasan, Mustafa Jamil, MGRMS Rahman, et al. Speaker identification using mel frequency cepstral coefficients. *variations*, Vol. 1, No. 4, 2004.
- [82] Lindasalwa Muda, Mumtaj Begam, and Irraivan Elamvazuthi. Voice recognition algorithms using mel frequency cepstral coefficient (mfcc) and dynamic time warping (dtw) techniques. *arXiv preprint arXiv:1003.4083*, 2010.
- [83] Douglas A Reynolds and Richard C Rose. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE transactions on speech and audio processing*, Vol. 3, No. 1, pp. 72–83, 1995.
- [84] Olivier Thyges, Roland Kuhn, Patrick Nguyen, and Jean-Claude Junqua. Speaker identification and verification using eigenvoices. In *Sixth International Conference on Spoken Language Processing*, 2000.
- [85] Tomi Kinnunen, Evgeny Karpov, and Pasi Franti. Real-time speaker identification and verification. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 14, No. 1, pp. 277–288, 2005.
- [86] Imran Naseem, Roberto Togneri, and Mohammed Bennamoun. Sparse representation for speaker identification. In *2010 20th International Conference on Pattern Recognition*, pp. 4460–4463. IEEE, 2010.
- [87] Manas A Pathak and Bhiksha Raj. Privacy-preserving speaker verification and identification using gaussian mixture models. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 2, pp. 397–406, 2012.

-
- [88] Noor Almaadeed, Amar Aggoun, and Abbes Amira. Speaker identification using multimodal neural networks and wavelet analysis. *IET Biometrics*, Vol. 4, No. 1, pp. 18–28, 2015.
- [89] Chao Li, Xiaokong Ma, Bing Jiang, Xiangang Li, Xuewei Zhang, Xiao Liu, Ying Cao, Ajay Kannan, and Zhenyao Zhu. Deep speaker: an end-to-end neural speaker embedding system. *arXiv preprint arXiv:1705.02304*, Vol. 650, , 2017.
- [90] Joon Son Chung, Arsha Nagrani, and Andrew Senior. Voxceleb2: Deep speaker recognition. *arXiv preprint arXiv:1806.05622*, 2018.
- [91] Sarthak Yadav and Atul Rai. Learning discriminative features for speaker identification and verification. In *Interspeech*, pp. 2237–2241, 2018.
- [92] Geert Molenberghs and Yingzi Eliza Du. Review of iris recognition: cameras, systems, and their applications. *Sensor review*, 2006.
- [93] Richard Yew Fatt Ng, Yong Haur Tay, and Kai Ming Mok. A review of iris recognition algorithms. In *2008 International Symposium on Information Technology*, Vol. 2, pp. 1–7. IEEE, 2008.
- [94] John Daugman. How iris recognition works. In *The essential guide to image processing*, pp. 715–739. Elsevier, 2009.
- [95] Nivedita S Sarode and AM Patil. Review of iris recognition: an evolving biometrics identification technology. *International Journal of Innovative Science and Modern Engineering*, Vol. 2, No. 10, pp. 34–40, 2014.
- [96] Yong Zhu, Tieniu Tan, and Yunhong Wang. Biometric personal identification based on iris patterns. In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, Vol. 2, pp. 801–804. IEEE, 2000.
- [97] Lia Ma, Yunhong Wang, Tieniu Tan, et al. Iris recognition based on multichannel gabor filtering. In *Proc. Fifth Asian Conf. Computer Vision*, Vol. 1, pp. 279–283, 2002.
- [98] Li Ma, Yunhong Wang, and Tieniu Tan. Iris recognition using circular symmetric filters. In *Object recognition supported by user interaction for service robots*, Vol. 2, pp. 414–417. IEEE, 2002.
- [99] Li Ma, Tieniu Tan, Yunhong Wang, and Dexin Zhang. Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image processing*, Vol. 13, No. 6, pp. 739–750, 2004.
- [100] Li Ma, Tieniu Tan, Yunhong Wang, and Dexin Zhang. Local intensity variation

- analysis for iris recognition. *Pattern recognition*, Vol. 37, No. 6, pp. 1287–1298, 2004.
- [101] Christel-loic Tisse, Lionel Martin, Lionel Torres, Michel Robert, et al. Person identification technique using human iris recognition. In *Proc. Vision Interface*, Vol. 294, pp. 294–299, 2002.
- [102] Zhenan Sun, Yunhong Wang, Tieniu Tan, and Jiali Cui. Improving iris recognition accuracy via cascaded classifiers. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 35, No. 3, pp. 435–441, 2005.
- [103] Mayank Vatsa, Richa Singh, and Afzel Noore. Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 38, No. 4, pp. 1021–1035, 2008.
- [104] Fernando Alonso-Fernandez, Pedro Tome-Gonzalez, Virginia Ruiz-Albacete, and Javier Ortega-Garcia. Iris recognition based on sift features. In *2009 First IEEE International Conference on Biometrics, Identity and Security (BIDS)*, pp. 1–8. IEEE, 2009.
- [105] Abhishek Gangwar and Akanksha Joshi. Deepirisnet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition. In *2016 IEEE international conference on image processing (ICIP)*, pp. 2301–2305. IEEE, 2016.
- [106] Kien Nguyen, Clinton Fookes, Arun Ross, and Sridha Sridharan. Iris recognition with off-the-shelf cnn features: A deep learning perspective. *IEEE Access*, Vol. 6, pp. 18848–18855, 2017.
- [107] Alaa S Al-Waisy, Rami Qahwaji, Stanley Ipson, Shumoos Al-Fahdawi, and Tarek AM Nagem. A multi-biometric iris recognition system based on a deep learning approach. *Pattern Analysis and Applications*, Vol. 21, No. 3, pp. 783–802, 2018.
- [108] Farzin Hadi, Abrishami-Moghaddam Hamid, and Mohammad-Shahram Moin. A novel retinal identification system. *EURASIP Journal on Advances in Signal Processing*, Vol. 2008, , 05 2008.
- [109] Robert “Buzz” Hill. *Retina Identification*, pp. 123–141. Springer US, Boston, MA, 1996.

-
- [110] Majid Shahnazi, Maryam Pahlevanzadeh, and Mansour Vafadoost. Wavelet based retinal recognition. In *2007 9th International Symposium on Signal Processing and Its Applications*, pp. 1–4. IEEE, 2007.
- [111] M. U. Akram, A. Tariq, and S. A. Khan. Retinal recognition: Personal identification using blood vessels. In *2011 International Conference for Internet Technology and Secured Transactions*, pp. 180–184, Dec 2011.
- [112] S. Qamber, Z. Waheed, and M. U. Akram. Personal identification system based on vascular pattern of human retina. In *2012 Cairo International Biomedical Engineering Conference (CIBEC)*, pp. 64–67, 2012.
- [113] Alexandru Paul Condurache, Johannes Kotzerke, and Alfred Mertins. Robust retina-based person authentication using the sparse classifier. In *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, pp. 1514–1518. IEEE, 2012.
- [114] Tripti Rani Borah, Kandarpa Kumar Sarma, and Pran Hari Talukdar. Retina recognition system using adaptive neuro fuzzy inference system. In *2015 International Conference on Computer, Communication and Control (IC4)*, pp. 1–6. IEEE, 2015.
- [115] Fahreddin Sadikoglu and Selin Uzelaltinbulat. Biometric retina identification based on neural network. *Procedia Computer Science*, Vol. 102, pp. 26–33, 2016.
- [116] Nikolaos V Boulgouris, Dimitrios Hatzinakos, and Konstantinos N Plataniotis. Gait recognition: a challenging signal processing technology for biometric identification. *IEEE signal processing magazine*, Vol. 22, No. 6, pp. 78–90, 2005.
- [117] Jin Wang, Mary She, Saeid Nahavandi, and Abbas Kouzani. A review of vision-based gait recognition methods for human identification. In *2010 international conference on digital image computing: techniques and applications*, pp. 320–327. IEEE, 2010.
- [118] Robert T Collins, Ralph Gross, and Jianbo Shi. Silhouette-based human identification from body shape and gait. In *Proceedings of fifth IEEE international conference on automatic face gesture recognition*, pp. 366–371. IEEE, 2002.
- [119] Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu. Silhouette analysis-based gait recognition for human identification. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 25, No. 12, pp. 1505–1518, 2003.
- [120] Lily Lee and W Eric L Grimson. Gait analysis for recognition and classifica-

- tion. In *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, pp. 155–162. IEEE, 2002.
- [121] Chiraz BenAbdelkader, Ross Cutler, and Larry Davis. Stride and cadence as a biometric in automatic person identification and verification. In *Proceedings of Fifth IEEE international conference on automatic face gesture recognition*, pp. 372–377. IEEE, 2002.
- [122] Amir Kale, Naresh Cuntoor, B Yegnanarayana, AN Rajagopalan, and Rama Chellappa. Gait analysis for human identification. In *International Conference on Audio-and Video-Based Biometric Person Authentication*, pp. 706–714. Springer, 2003.
- [123] Amir Kale, Naresh Cuntoor, B Yegnanarayana, AN Rajagopalan, and Rama Chellappa. Gait analysis for human identification. In *International Conference on Audio-and Video-Based Biometric Person Authentication*, pp. 706–714. Springer, 2003.
- [124] Amit Kale, Amit K RoyChowdhury, and Rama Chellappa. Fusion of gait and face for human identification. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, pp. V–901. IEEE, 2004.
- [125] Rezaul K Begg, Marimuthu Palaniswami, and Brendan Owen. Support vector machines for automated gait classification. *IEEE transactions on Biomedical Engineering*, Vol. 52, No. 5, pp. 828–838, 2005.
- [126] Chunxiao Liu, Shaogang Gong, Chen Change Loy, and Xinggang Lin. Person re-identification: What features are important? In *European Conference on Computer Vision*, pp. 391–401. Springer, 2012.
- [127] Zhun Zhong, Liang Zheng, Zhiming Luo, Shaozi Li, and Yi Yang. Invariance matters: Exemplar memory for domain adaptive person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 598–607, 2019.
- [128] Zhun Zhong, Liang Zheng, Zhedong Zheng, Shaozi Li, and Yi Yang. Camera style adaptation for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5157–5166, 2018.
- [129] Loris Bazzani, Marco Cristani, Alessandro Perina, Michela Farenzena, and Vittorio Murino. Multiple-shot person re-identification by hpe signature. In *2010 20th International Conference on Pattern Recognition*, pp. 1413–1416. IEEE,

- 2010.
- [130] Xiang Li, Wei-Shi Zheng, Xiaojuan Wang, Tao Xiang, and Shaogang Gong. Multi-scale learning for low-resolution person re-identification. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3765–3773, 2015.
 - [131] Filip Malawski. Top-view people counting in public transportation using kinect. *Challenges of Modern Technology*, Vol. 5, No. 4, 2014.
 - [132] Daniele Liciotti, Marina Paolanti, Emanuele Frontoni, Adriano Mancini, and Primo Zingaretti. Person re-identification dataset with rgb-d camera in a top-view configuration. In *Video Analytics. Face and Facial Expression Recognition and Audience Measurement*, pp. 1–11. Springer, 2016.
 - [133] Daichi Kouno, Kazutaka Shimada, and Tsutomu Endo. Person identification using top-view image with depth information. In *2012 13th ACIS international conference on software engineering, artificial intelligence, networking and parallel/distributed computing*, pp. 140–145. IEEE, 2012.
 - [134] Ryota Nakatani, Daichi Kouno, Kazutaka Shimada, and Tsutomu Endo. A person identification method using a top-view head image from an overhead camera. *JACIII*, Vol. 16, No. 6, pp. 696–703, 2012.
 - [135] 中谷良太, 香野大地, 嶋田和孝, 遠藤勉ほか. 頭上方向から撮影されたカメラ画像を利用した人物識別. 画像の認識・理解シンポジウム (MIRU2011) 論文集, Vol. 2011, pp. 593–598, 2011.
 - [136] Fatih Porikli. Inter-camera color calibration by correlation model function. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, Vol. 2, pp. II–133. IEEE, 2003.
 - [137] Omar Javed, Khurram Shafique, and Mubarak Shah. Appearance modeling for tracking in multiple non-overlapping cameras. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 2, pp. 26–33. IEEE, 2005.
 - [138] Andrew Gilbert and Richard Bowden. Tracking objects across cameras by incrementally learning inter-camera colour calibration and patterns of activity. In Bischof Horst Leonardis, Aleš and Axel Pinz, editors, *Computer Vision – ECCV 2006*, pp. 125–136, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
 - [139] Yi-Ping Hung Kuan-Wen Chen, Chih-Chuan Lai and Chu-Song Chen. An adaptive learning method for target tracking across multiple cameras. In *2008 IEEE*

- Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.
- [140] Bryan James Prosser, Shaogang Gong, and Tao Xiang. Multi-camera matching using bi-directional cumulative brightness transfer functions. In *BMVC*, Vol. 8, p. 74. Citeseer, 2008.
- [141] Niloofar Gheissari, Thomas B Sebastian, and Richard Hartley. Person reidentification using spatiotemporal appearance. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, Vol. 2, pp. 1528–1535. IEEE, 2006.
- [142] Douglas Gray and Hai Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In *European conference on computer vision*, pp. 262–275. Springer, 2008.
- [143] Bryan James Prosser, Wei-Shi Zheng, Shaogang Gong, Tao Xiang, and Q Mary. Person re-identification by support vector ranking. In *BMVC*, Vol. 2, p. 6, 2010.
- [144] Wei-Shi Zheng, Shaogang Gong, and Tao Xiang. Reidentification by relative distance comparison. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 35, No. 3, pp. 653–668, 2012.
- [145] Andy J Ma, Pong C Yuen, and Jiawei Li. Domain transfer support vector ranking for person re-identification without target camera label information. In *Proceedings of the IEEE international conference on computer vision*, pp. 3567–3574, 2013.
- [146] Michela Farenzena, Loris Bazzani, Alessandro Perina, Vittorio Murino, and Marco Cristani. Person re-identification by symmetry-driven accumulation of local features. In *2010 IEEE computer society conference on computer vision and pattern recognition*, pp. 2360–2367. IEEE, 2010.
- [147] Alexis Mignon and Frédéric Jurie. Pcca: A new approach for distance learning from sparse pairwise constraints. In *2012 IEEE conference on computer vision and pattern recognition*, pp. 2666–2672. IEEE, 2012.
- [148] Rui Zhao, Wanli Ouyang, and Xiaogang Wang. Unsupervised salience learning for person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3586–3593, 2013.
- [149] Rui Zhao, Wanli Ouyang, and Xiaogang Wang. Person re-identification by salience matching. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2528–2535, 2013.

-
- [150] Rui Zhao, Wanli Ouyang, and Xiaogang Wang. Learning mid-level filters for person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 144–151, 2014.
- [151] Sateesh Pedagadi, James Orwell, Sergio Velastin, and Boghos Boghossian. Local fisher discriminant analysis for pedestrian re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3318–3325, 2013.
- [152] Abir Das, Anirban Chakraborty, and Amit K Roy-Chowdhury. Consistent re-identification in a camera network. In *European conference on computer vision*, pp. 330–345. Springer, 2014.
- [153] Xiao Liu, Mingli Song, Dacheng Tao, Xingchen Zhou, Chun Chen, and Jiajun Bu. Semi-supervised coupled dictionary learning for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3550–3557, 2014.
- [154] Yang Yang, Jimei Yang, Junjie Yan, Shengcai Liao, Dong Yi, and Stan Z Li. Salient color names for person re-identification. In *European conference on computer vision*, pp. 536–551. Springer, 2014.
- [155] Shengcai Liao, Yang Hu, Xiangyu Zhu, and Stan Z Li. Person re-identification by local maximal occurrence representation and metric learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2197–2206, 2015.
- [156] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, Vol. 13, No. 1, pp. 21–27, 1967.
- [157] Toshikazu Wada. Visual object tracking using positive and negative examples. In Makoto Kaneko and Yoshihiko Nakamura, editors, *Robotics Research*, pp. 189–199, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
- [158] <https://sourceforge.net/projects/weka/>.
- [159] 奥剛一. 俯瞰視 rgb-d カメラによる個人再識別用領域の自動取得. 和歌山大学 2019 年度博士前期課程論文
<http://web.wakayama-u.ac.jp/wuhy/syuronOku.pdf>, 2019.

研究業績

本論文の一部は、下記論文の内容を含む。

学術論文

- 1 李 鵬, 呉海元, ”SHAL: 俯瞰視画像から個人再識別ための特徴記述法”, 映像情報メディア学会誌, 74.4: 719-728, 2020.

国際会議（査読あり）

- 2 Peng Li, Haiyuan Wu, Qian Chen, Chongke Bi, ”Person re-identification using color enhancing feature”, 3rd IAPR Asian Conference on Pattern Recognition (ACPR), 086-090, 2015.
- 3 Peng Li, Haiyuan Wu, Qian Chen, ”Color distinctiveness feature for person identification without face information”, Proc.of 19th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2015), Procedia Computer Science (Open Access Journal), Vol.60, pp.1809-1816, 2015.
- 4 Haiyuan Wu, Peng Li, Jiwei Zhang, Qian Chen, ”Inventory Management using Color Images”, Proc. of the 2020 IEEE International Conference on Mechatronics and Automation (IEEE ICMA), pp.610 – 614, 2020.

特許

- 5 Haiyuan Wu, Peng Li, ”識別システム, 識別方法, 及びコンピュータプログラム”, 出願番号: 特願 2019-134318.

その他 (国内発表)

- 6 李 鵬, 奥剛一, 呉海元, ”プライバシーの保護を考慮した個人再識別法と俯瞰視システム, 信学技報, vol. 118, no. 513, PRMU2018-137, pp. 25-30, 2019.
- 7 奥剛一, 李 鵬, 呉海元, ”俯瞰視 RGB-D カメラによる個人再識別用の画像の自動取得, 信学技報”, vol. 118, no. 513, PRMU2018-138, pp.31-35, 2019.
- 8 李 鵬, 鈴木一正, 呉海元, 陳謙, ”色弁別度を特徴量とした個人識別”, 情報処理学会, 第 77 回全国大会, IS-07, pp.2-365-2-366, 2015 (3 月 17 日京都大学吉田キャンパス) .
- 9 李 鵬, 郎彦昆, 陳謙, 呉海元, ”FREAK を用いた顔認証”, 第 20 回画像センシングシンポジウム, pp.IS3-08-1 – IS3-08-8, 2014.
- 10 李 鵬, 郎彦昆, 陳謙, 呉海元, ”FREAK を用いた人物の識別”, 信学技報, Vol.113, No.402, pp.263-268, 2014.

解説記事

- 11 呉海元, 李 鵬, ”SHAL: 俯瞰視画像から個人再識別ための特徴記述法 - プライバシーを考慮した個人再識別法 -”, 画像ラボ, Vol.32, No.2, pp.40-45, 2021.

関連ページ

奥 剛一手法のデモ動画：

http://web.wakayama-u.ac.jp/~wuhhy/Rec_20190129_105607.avi

奥 剛一手法で収集した俯瞰視画像セット：

<http://web.wakayama-u.ac.jp/~wuhhy/ImageSet.zip>

SHAL 手法の「私服を着る実験結果 (For windows)」

(Access Password:SHAL-19614Li)：

<http://web.wakayama-u.ac.jp/~wuhhy/AboutSHAL.zip>

付録 A

CDF 法による物体の再識別

本付録では，CDF が物体の再識別にも利用できるかを確認するために，CDF 法を用いて物体の再識別実験を行った。

実験用物体は 5 種類を準備した．各物体に対して 2 つ異なる撮影角度から画像を一枚ずつ撮影する (図 A.1 に参照)．物体のデータベースに登録用既知 CDF を生成する際に，該当物体の 2 枚画像のうちに任意の 1 枚の画像からメイン色とサポート色を選択する．CDF 法のメイン領域とサポート領域はすべて物体上の画像領域から選ぶ (図 A.2 に参照)．

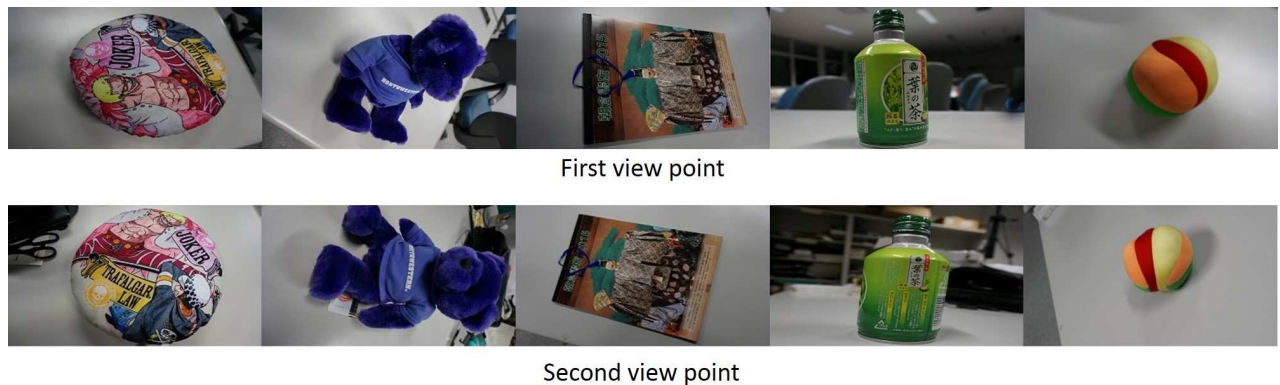


図 A.1 5つの被験物体の異なる角度で撮影した画像

再識別の際に，物体の既知 CDF 生成用画像を除き，もう一枚の画像より CDF を生成し，未知物体の特徴 CDF とする．これから，個人再識別実験のように，未知物体の CDF_l とデータベースの各 CDF_{ik} との照合を行い，バタチャリア距離とハミング距離で類似度評価を行った．表 A.1 と表 A.2 はそれぞれバタチャリア距離，ハミング距離を用

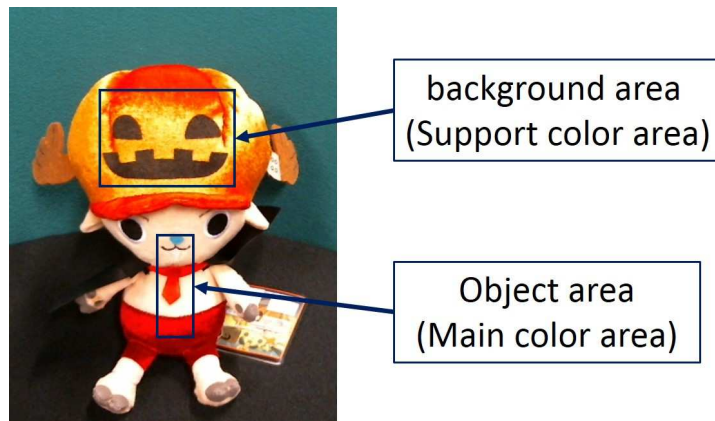


図 A.2 メイン領域とサポート領域の選び方

いた再識別類似度を示す混同行列である。各行において最も高い類似度を赤色でマークしている。

表 A.1 バタチャリア距離を用いた物体の再識別結果 (%)

PersonNo	1	2	3	4	5
1	98.94	98.60	98.84	97.99	98.08
2	98.44	99.51	99.07	98.01	97.89
3	99.34	99.37	99.75	99.10	98.96
4	99.06	99.41	99.22	99.47	99.40
5	98.86	99.18	98.95	99.27	99.83

表 A.2 ハミング距離を用いた物体の再識別結果 ($w = 64$)

PersonNo	1	2	3	4	5
1	0.48	0.00	0.19	0.02	0.01
2	0.01	0.89	0.67	0.00	0.00
3	0.14	0.02	0.76	0.04	0.01
4	0.04	0.00	0.02	0.15	0.00
5	0.31	0.00	0.00	0.00	0.45

表 A.1 と表 A.2 に示されている通り、バタチャリヤの類似度は全体的に高かったが、自己比較（表の対角線上の類似度）の値は最も高かった。一方、ハミング距離による評価し

た物体のすべての再識別結果を見ると、異なる物体間に明らかに大きな差があることは分かった。バタチャリア距離の実験結果と同様、最も高いのは自己比較の結果（表の対角線上の類似度）である。したがって、CDF 法を用いて特徴を記述し、メインとサポート 2 つの領域の色のみ利用するだけで物体再識別にも利用可能のことを示した。

付録 B

WDRB 法による物体の再識別

WDRB 法は、色と色ヒストグラムを利用して特徴量の記述ができるので、人の再識別の他に、物体の再識別にも使用できると考えられる。本付録では、WDRB および CDF 法を利用して物体再識別の比較実験を行った。この実験では、15 個の物体を使用し、各物体に対して、2 つ異なる視点から撮影し、2 つの画像セットを用意した (図 B.1 に参照)。物体の WDRB および CDF 特徴量は、各物体の 2 つの画像を利用し、それぞれ画像の物体領域を選択して計算した (図 B.2 に参照)。

物体の再識別には、物体の 2 つの画像セットを利用し、WDRB 法と CDF 法の両方の実験を行った。2 つ画像セットの内に 1 セットを利用して WDRB と CDF 法の特徴量を生成し、既知物体としてデータベースに登録する。もう 1 つの画像セットを利用して未知物体の WDRB と CDF 法の特徴量を生成し、順番に各データベースの WDRB および CDF との間のマッチングを行った。個人再識別実験と同様に、バタチャリア距離を使用して特徴量間の類似度を計算した。両手法の実験結果はそれぞれ図 B.3 (a) および (b) に示している。実験結果は、個人再識別実験と同じ、混同行列で表示する。No.1 から 15 は、データベースの既知特徴量の番号である。各行は、同じ番号順に該当物体の未知特徴量を利用した物体の再識別実験である。図 B.3 によると、WDRB 法を利用したすべての再識別は成功であった。CDF 法を利用した物体再識別はほとんど正しく識別できたが、特徴量間の類似度を代表する色も全体的に白くなり、特徴量間には明確な差異がなかったことを意味している。したがって、WDRB 法は CDF 法より高い安定性を持つことが分かった。



The 1st view point picture

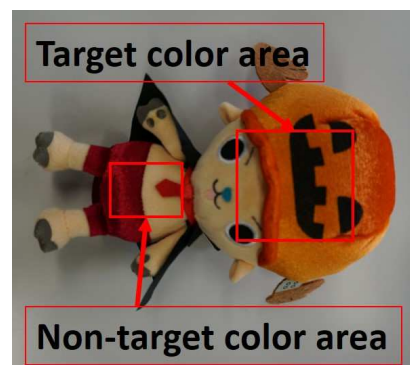


The 2st view point picture

図 B.1 2つの視点から撮影した物体の画像セット

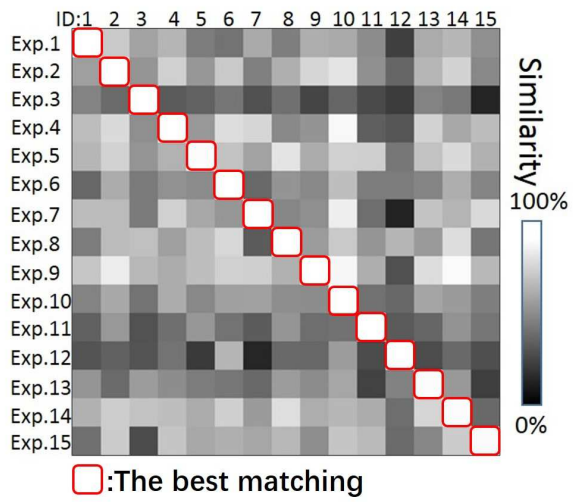


(a) WDRB 法の物体領域の選択

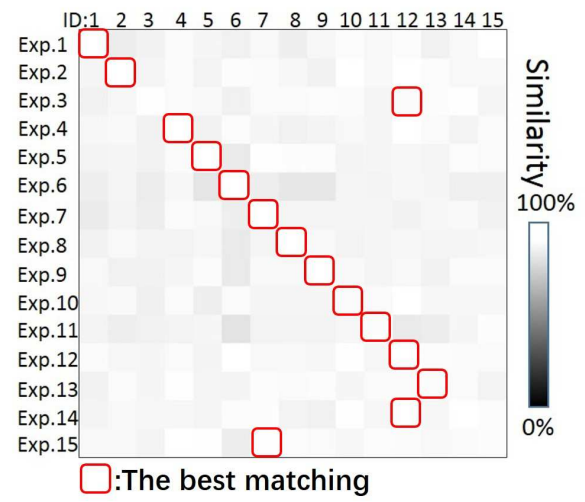


(b) CDF 法の物体領域の選択

図 B.2 WDRB 法と CDF 法それぞれの物体領域の選択



(a) WDRB 法の結果



(b) CDF 法の結果

図 B.3 WDRB 法と CDF 法による物体再識別実験の結果

付録 C

OiF 法による物体の再識別

5.4.9 項では OiF 法を個人再識別に利用する場合，成功率は SHAL 法より低いことが分かった．本付録では，俯瞰視システムを用いて，OiF 法が物体の再識別に利用できるかを確認する．物体の再識別実験は実応用形式によって，

- ① 物流システム用の段ボール再識別実験
- ② 無人商店などの自動会計システム用の商品の再識別実験

の 2 種類の再識別実験に分けて行った．

C.0.1 段ボールの再識別実験

物流分野の最小単位である段ボール箱を用いた実験を行った．提案された手法の有効性と安定性を確認するために，サイズと色が似ている段ボール箱をいくつか利用した．各段ボールに対して，異なる角度の 10 枚の俯瞰視画像を俯瞰視カメラの下で撮影した．画像セット ID_i を図 C.1 に示す．各列は異なる段ボール $ID (= a, b, c, d, e, f)$ であり，各行は該当段ボールの異なる角度で撮った画像セット $i (= 0, 1, \dots, 9)$ である．この実験では，学習画像セットとして一行の画像セットを使用して $\mathbf{OiF}(id)$ を記述してデータベースに登録する．再識別のために，他の 9 セットの画像を使用して未知物体として $\mathbf{OiF}(x)$ を記述し，重み付きバタチャリヤ距離による類似度評価を行った．図 C.1 中の (a) と (b) はそれぞれデータベース用画像セットと未知物体セットを示している．0 番から 9 番までに順番にデータベース用画像セットとして，実験が全部 10 回行った．また，個人再識別実験と同様に式 (5.4) の α を 20 に設定した (図 5.8 に参照)．各段ボール箱の平均成功率とすべての段ボール箱の平均成功率を表 C.1 に纏めた．

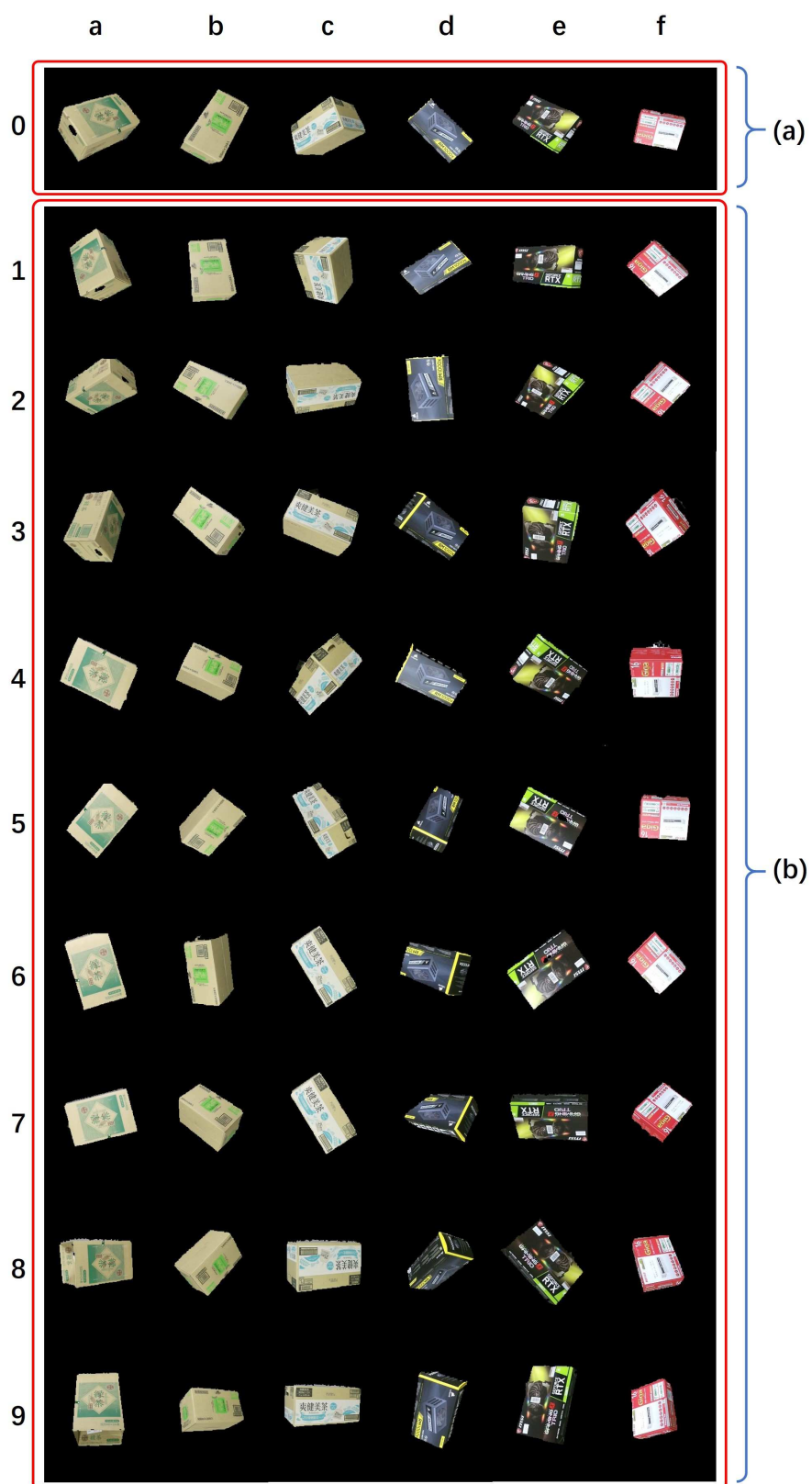


図 C.1 段ボールの画像セット.

表 C.1 段ボールの再識別実験の結果 ($\alpha = 20$)

	a(%)	b(%)	c(%)	d(%)	e(%)	f(%)	Total average(%)
0	88.9	100	100	88.9	100	100	96.3
1	88.9	55.6	100	88.9	100	100	88.9
2	66.7	33.3	100	88.9	88.9	100	79.6
3	88.9	22.2	100	100	100	100	85.2
4	66.7	77.8	100	88.9	100	100	88.9
5	100	100	100	100	100	66.7	94.4
6	55.6	100	100	100	100	100	92.6
7	44.4	100	100	100	100	77.8	87.0
8	44.4	77.8	100	100	100	100	87.0
9	88.9	88.9	100	100	100	100	96.3
ALL							89.62%

表 C.1 の 2 行目から 11 行目までは各回の再識別実験の中で各段ボールの平均成功率である。2 列目から 7 列目までは各段ボールが各回実験の中の再識別成功率である。8 列目は各回再識別実験の平均成功率である。12 行目はすべての 10 回の再識別実験の平均成功率である。すべての実験の最終平均成功率 (ALL) は 89.62% である。この実験結果により、段ボール a と b の再識別は何度も失敗したことが分かった。実際の画像から確認すると、二つの段ボールの色合いがよく似ていることが、再識別成功率が低下した原因と考えられる。

C.0.2 商品の再識別実験

無人店舗・百貨店などの自動会計システムには、商品の再識別はかなり重要な機能である。ここでは、実際の商品を使った再識別実験を行った。図 C.2 に示すように、5 つの異なる商品 $ID (= A, B, C, D, E)$ を手持ちで、俯瞰視カメラの下を渡した。俯瞰視画像上に、手と商品の領域を奥 剛一の手法 [159] によって検出して追跡した。この領域の重心の移動軌跡によって、商品が棚から取り出された状態と棚に戻された状態を判定した (図 C.3 に参照)。

俯瞰視カメラで事前に撮影した各商品の 1 つの画像を利用し、該当商品の $OiF(id)$ を記述してデータベースに保存した。ここで、 $id (= A, B, C, D, E)$ (図 C.4 に参照)。



図 C.2 実験の様子

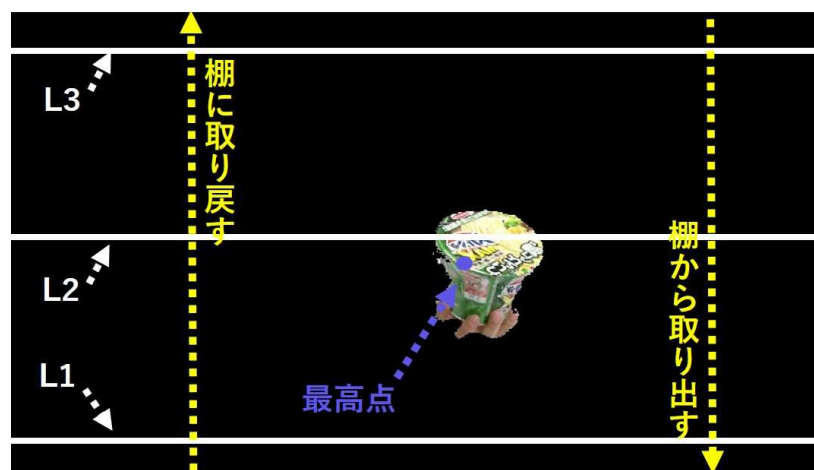


図 C.3 実験様子と商品の取り出す・取り戻すの判定

図 C.5 は 5 つの商品の $\mathbf{Cf}(U, V)$ を示す. 図 C.5 から, 各商品の $\mathbf{O_iF}$ 間の異なることが確認できる. したがって, 提案手法は商品の再識別にも利用可能と考えられる.



図 C.4 実験用の五つの商品

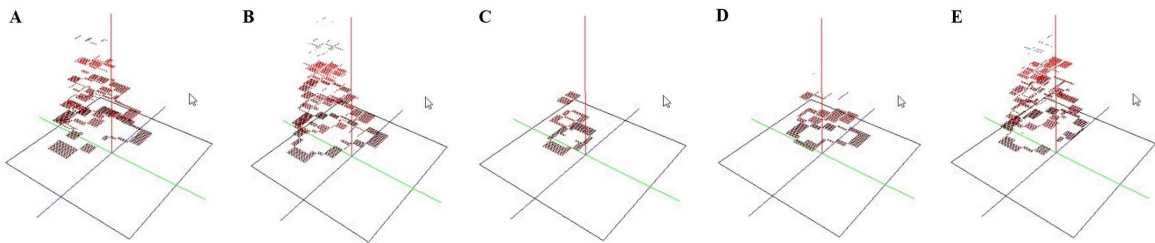


図 C.5 各商品の拡張された色融合マップ

この実験では, 各商品に対して, 俯瞰視カメラの下を手持ちでランダム角度で 20 回通過させた. 各商品の 20 枚の画像の中に, 1 つの画像はデータベース学習用で, 他の 19 枚の画像は未知商品として利用した. 図 C.6 は, 商品 A の画像セットの例である. 実験時に, 商品が 1 回カメラを通過したときに, 該当商品の一枚の画像をキャプチャーして $\mathbf{O_iF}(x)$ を生成し, 重み付きバタチャリヤ距離によって再識別を行った. 各商品の平均成功率とすべての商品の平均成功率は表 C.2 に纏めている. 商品の色が異なるため, 精度の高い再識別結果が得られた.

表 C.2 商品の再識別結果 ($\alpha = 20$)

A	B	C	D	E	Total average
100%	100%	100%	99.47%	99.47%	99.79%

これらの実験結果より, $\mathbf{O_iF}$ 法は物体再識別に応用できることを確認した.

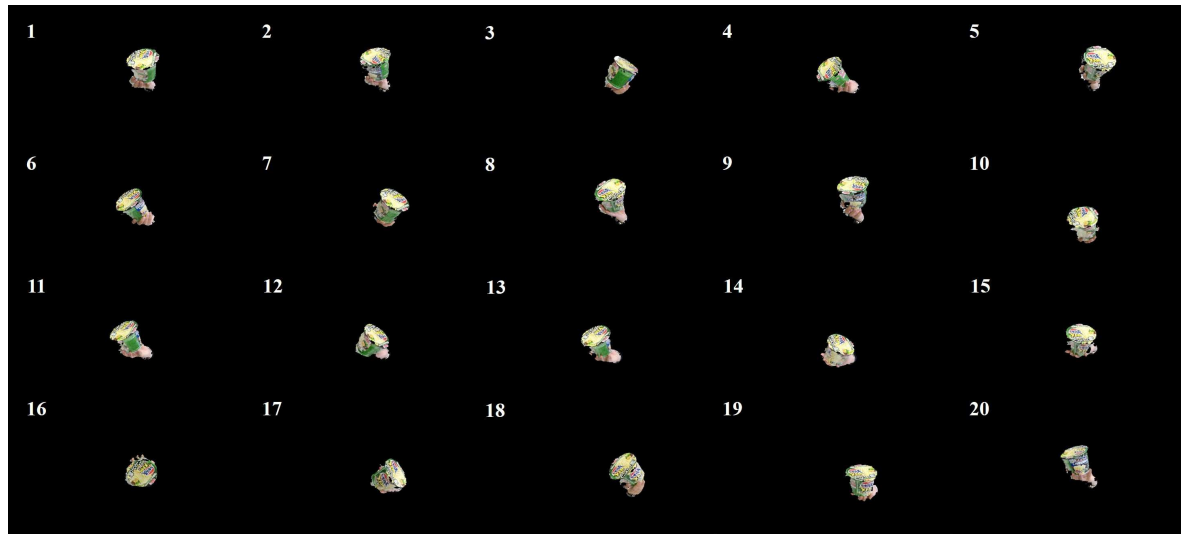


図 C.6 俯瞰視カメラの下にランダム角度で通過する手持ち商品 A の画像例

C.1 まとめ

ここでは、段ボールの再識別実験と商品の再識別実験より、OiF 法は商品の再識別が可能になり、無人店舗などに利用できる可能性を示した。

謝辞

本研究の過程において、指導教員として長きに渡る懇切なる御指導、御鞭撻を賜りました和歌山大学システム工学研究科呉 海元教授に心より感謝致します。また、国内、国際会議での学会発表などの貴重な経験をさせて頂いたこと、ならびに私生活の面での様々なご配慮を頂いたことを重ねてお礼申し上げます。

本研究を進めるにあたり、研究に関わる内容のみならず、プログラミング等その他技術的な面でも多くの御指導、御助言を賜りました陳 謙准教授に心から深く感謝致します。

本論文の作成にあたり、学位論文審査委員として懇切なる御指導を賜りました和歌山大学システム工学部吉野 孝教授と塚田 晃司教授および日本語の修正において丁寧な指導して下さいました和歌山大学学術情報センターの瀧 寛和教授に感謝の意を表します。

最後に、学部から、博士前期課程を経て、博士後期課程修了に至るまで、学生生活を精神的、経済的に支えてくれた著者の家族に感謝の意を表します。