

2023 年度(令和五年度)

創造工学セミナー II Final Report

人物追従ロボットのための
対象識別機能に関する研究

知能機械研究室

s520049 晝岡雅希

指導教員

金丸隆志 教授

目次

1. 緒言	4
1.1 背景・目的	4
1.2 先行研究	5
1.2.1 二台のカメラを用いた OpenPose の骨格抽出による移動ロボットの 3 次元人物追跡	5
1.2.2 人物追従ロボットのための深度センサによる頭部と胴体情報を用いた追従制御法の提案とその評価	6
1.2.3 ヒトを活動的にする移動支援ロボット -ID を発信する赤外 LED タグと全方位カメラを用いた利用者の追跡-	7
1.3 研究目標	8
2. 人物識別システム	9
2.1 GPU 搭載デスクトップ PC	9
2.2 OpenPose	9
2.3 画像処理用ライブラリ OpenCV	10
2.3.1 calcHist 関数	10
2.3.2 compareHist 関数	11
2.4 使用する色空間	12
2.5 実装想定	13
2.6 識別システムの仕組み	14
4. 実験 1：色分布によるテンプレートでの人物識別	16
4.1 実験概要	16
4.2 実験に使用する動画	16
4.3 実験結果	19
4.3.1 実験 1.a 服の違いが大きく、明るさの変化が小さい状況での識別(動画①) ..	19
4.3.2 実験 1.b 服の違いが大きく、明るさの変化が大きい状況での識別(動画②) ..	21
4.3.3 実験 1.c 服が似ていて、明るさの変化が大きい場所での識別(動画③④⑤) ..	22
4.4 考察	24
5. 実験 2：人物識別中のテンプレートの更新	26
5.1 テンプレートの更新	26
5.2 実験 2	26
5.2.1 実験 2.a 動画②での実験	27
5.2.2 実験 2.b 動画③での実験	28
5.2.3 実験 2.c テンプレート更新間隔の変更	29
5.3 実験 2.d テンプレートの更新手法の変更	31

6. 結言	34
6.1 結論	34
6.2 課題とその解決案.....	34
6.2.1 課題.....	34
6.2.2 解決案1 骨格推定情報の活用	34
6.2.3 解決案2 深度情報や座標情報の併用	35
参考文献	37
謝辞.....	38

1. 緒言

1.1 背景・目的

近年、人々の生活や作業などを手伝えるロボットの技術が発達し、実用化が進んでいる。人々の身近なところで実用化されているロボットとして、図1のようなロボット掃除機[1]やレストランの配膳ロボット[2]などがある。典型的なロボット掃除機は、ToF(Time-of-Flight)を利用した超音波センサにより障害物回避を行う。また、SLAM(Simultaneous Localization and Mapping)による地図作成を行う機種もある。そのような機種は、LiDAR(Light Detection and Ranging)センサや慣性計測装置などによる自己位置推定を行う。また、レストランの配膳ロボットは、赤外線センサによるビジュアルSLAMとLiDARによるLiDAR SLAMの選択式で自己位置を推定する。また、深度カメラにより3次元を知覚して障害物を回避することができる。それにより、人間とは独立した作業を行う自律型のロボットである。

それに対し、本研究では人間が行っている作業を完全にロボットに置き換えるのではなく、人間の作業の負担を減らすことを目的とした、補助的な役割を持つロボットに注目する。その中でも、本研究では荷物を載せ、人間を追従する移動ロボットについて研究を行う。

空港や物流倉庫などでは、カートに荷物を載せ、人間がそれらを自分で押すことで荷物を運んでいる。この作業を追従ロボットが行うことで、人間の負担を軽減できると考えた。現状ではそのような機能を持つロボットは実用化されていない。実用化を妨げる主な要因として、追従の際に安全性を十分に確保できないことが挙げられる。重い荷物を載せているロボットが利用者を正確に追従できない場合、利用者本人や周りの人々に危険が及ぶ恐れがある。また、前述した安全性以外にも、利用者を正確に追従することのできない移動ロボットは利便性が低く、人間の負担を減らすという目的にそぐわない。そこで、移動ロボットが利用者を正確に識別することが出来れば、安全性や利便性が向上し、実用化に近付くと考えている。そのため、本研究では利用者を正確に識別する仕組みを実現する。



図1 ロボット掃除機[1](左)と配膳ロボット[2](右)

1.2 先行研究

本節では追従ロボットに関する先行研究について説明する。

1.2.1 二台のカメラを用いたOpenPoseの骨格抽出による移動ロボットの3次元人物追跡

松原らは2台のステレオカメラとOpenPoseによる骨格推定を使用し、色分布を基に人物に追従するロボットを作成した[3]。この研究は、照明による明るさの変化があり、追跡対象人物以外の人や障害物の無い場所を実験環境としている。本研究では、この研究を参考にした人物識別の手法を用いる。松原らは図2のように、追従人物の左右の肩と足首を結んだ四角形部分の色分布を基に人物追従を行った。その際、腰の部分で四角形を2つの領域に分けているが、色分布の取得領域内に背景が含まれてしまうことが課題であると考えた。



(a) OpenPose による骨格検出



(b) 人物候補領域

図2 松原らの色分布取得領域[3]

1.2.2 人物追従ロボットのための深度センサによる頭部と胴体情報を用いた追従制御法の提案とその評価

下山らは図3のように深度センサXtionと、ライブラリであるOpenNIとNITEから得られる骨格情報から、各人物を識別するための図形を設定し、利用者の識別と追従を行う仕組みを作成した[4]。識別のための図形は図4で示される黄色の部分であり、これをトルソと呼ぶ。トルソは頭と両肩を結ぶ三角形と、両肩の幅とカメラ位置を原点とした肩の高さからなる四角形を合わせた図形である。トルソの面積は身長や肩幅だけでなく首の長さなどを反映した量であるため、人物間に十分な違いが生じ、それによって人物の識別が可能となっている。ただし、対象の身長が高すぎる場合に頭部が見切れてしまい、正確な値が取得できないことや、体格が似ている人物どうしの識別が難しいことなどが課題として挙げられている。

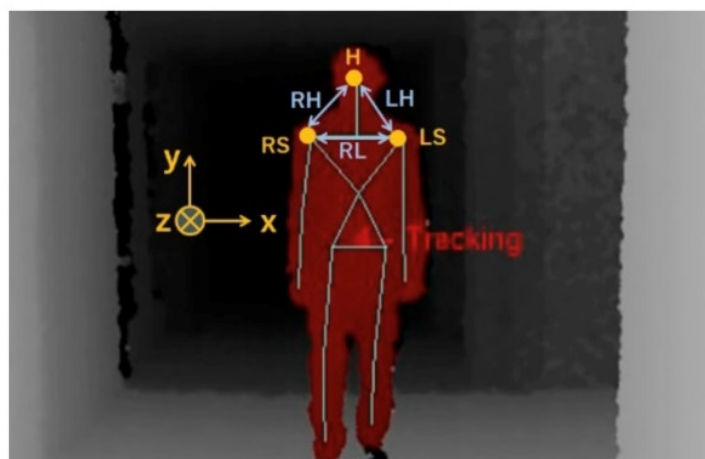


図3 深度センサによる骨格推定情報[4]

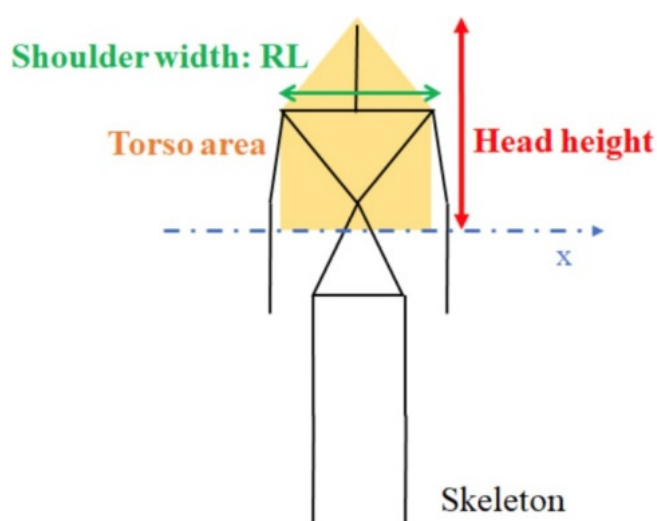


図4 トルソの定義(黄色の領域)[4]

1.2.3 ヒトを活動的にする移動支援ロボット —IDを発信する赤外LEDタグと全方位カメラを用いた利用者の追跡—

小山らは利用者の足首に特定パターンで発光する赤外LEDタグ(図5)を付け、LEDタグの発光パターンによって利用者を識別するロボットを作成した[5]。図6のように輝度によって画像を二値化することで画角内の光を検出し、パーティクルフィルタによってそれぞれの発光パターンを識別する。この研究のように、識別用の道具を利用者に装着させる手法は安定性が高い。しかし、利便性が損なわれる点や、道具を装着することに対する利用者の抵抗感などが課題であると考えられる。

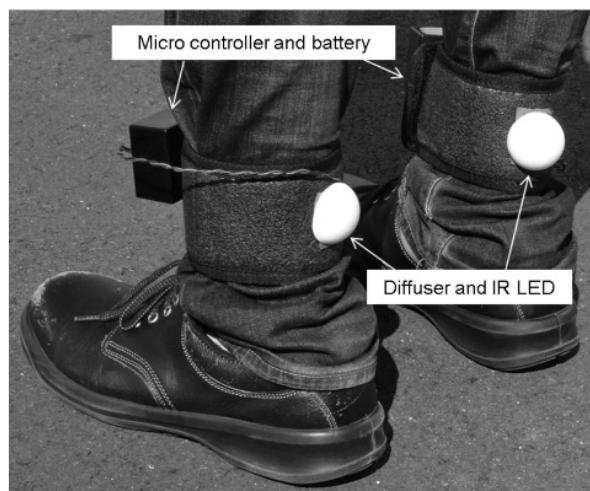


図5 識別用LEDタグを足首に装着[5]

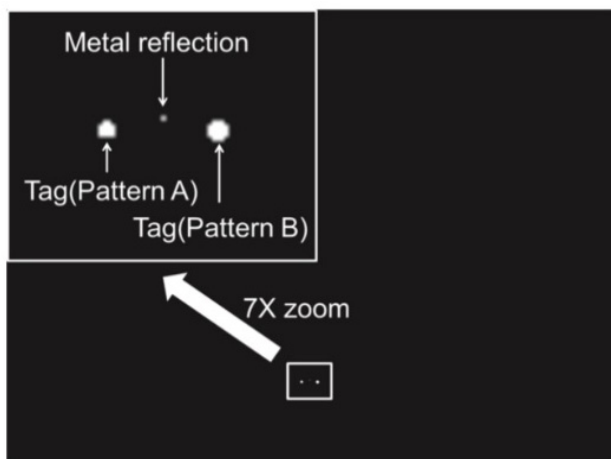


図6 輝度による画像の二値化によって光を検出[5]

1.3 研究目標

1.1節で述べたように、利用者の荷物を運ぶ追従ロボットが実用化されていない原因は安全性や利便性が十分に確保できない事である。そして、それらの問題を解決するためには、追従ロボットが利用者を正確に識別し、追従し続けられることが重要であると考えている。以上のことから、本研究では以下の2つの条件を満たす追従ロボットのための人物識別システムを作成することが目標である。

- ・利用者を周りの人物や障害物から正確に識別することができる。
- ・正確な識別を継続的に行うことができる。

また、識別手法は1.2.1項で説明した松原ら[3]の手法を参考にし、OpenPose[6]による骨格推定と、OpenCVで取得する色分布[8]のみを使用した識別システムの実現を目指す。OpenPoseとOpenCVについては第2章にて詳しく説明する。

2. 人物識別システム

本章では本研究で実現する人物識別システムについて説明する。

2.1 GPU搭載デスクトップPC

今回実験に使用するパソコンのスペックは以下の通りである。

- ・ CPU : intel core i7-12700
- ・ GPU : NVIDIA Geforce RTX-3060
- ・ メモリ : 16GB

2.2 OpenPose

本研究では骨格推定システムであるOpenPose[6]を使用する。OpenPoseはカーネギーメロン大学で開発されたディープラーニングによる骨格推定システムであり、2次元画像から図7で示す18個の骨格点の座標を取得できる。本研究では18個の骨格点のうち、2番(右肩)、5番(左肩)、8番(右腰)、11番(左腰)の4点を人物識別に使用する。また、OpenPoseは図8のように複数人の同時認識が可能である。

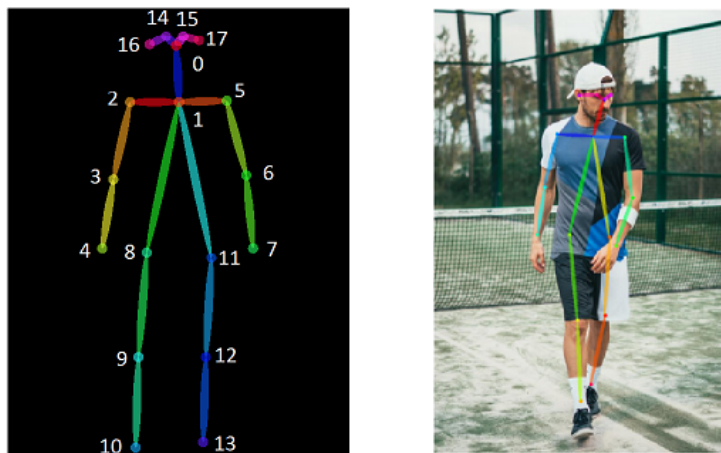


図7 OpenPoseで取得できる骨格点(左)と認識結果を示した画像(右)[7]



図8 OpenPoseによる複数人の同時認識の様子[6]

OpenPoseの骨格推定機能を使用するためのプログラムとして、Hzzone氏のpytorch-openpose[7]を使用する。このプログラムはOpenPoseで検出した骨格点を表示することができる。そして、このプログラムを基に、色分布を取得する機能や取得した色分布によって対象を識別する機能などを追加することで人物識別システムを作成する。

2.3 画像処理用ライブラリOpenCV

本節では、画像処理用ライブラリのOpenCVのうち、今回作成した人物識別システムで使った機能について説明する。OpenCVは様々な言語で利用できるが、本研究ではPython用のOpenCVを用いる。

2.3.1 calcHist関数

calcHist関数はヒストグラムを計算する関数であり[8]、本研究では色分布を取得するために使用する。OpenCVでは画像を、高さ×幅×色の3次元配列として読み込む。そして、calcHist関数は対象領域内の指定された色チャンネルにおける値の要素数を算出し、色の値×1の2次元配列を結果として出力する。なお、色の値とは指定された色チャンネル(本研究ではH(0)、S(1)、V(2))の値であり、0から255の範囲の値である。得られた色の値×1のデータが色分布であるが、本研究では要素数をテンプレート領域の面積(ピクセル数)で割って規格化することで色分布を割合で表し、人物毎の体格の違いに対応している。図7のテンプレート領域のヒストグラムを取得した結果が図9である。

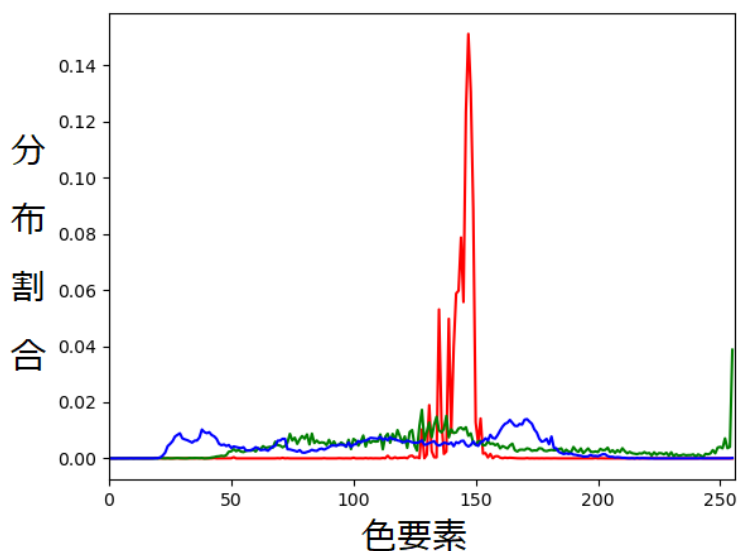


図9 図7のテンプレート領域のヒストグラム(H(赤)、S(緑)、V(青))

2.3.2 compareHist関数

compareHist関数は2つのヒストグラムの類似度を取得する関数であり[8]、比較する2つのヒストグラムと比較手法を引数に持つ。本研究では比較手法として、松原らの研究[3]と同様にピアソンの相関係数(式1)[8]を使用した。ピアソンの相関係数は-1から1の値をとる。本研究では $r = (d + 1)/2$ の値を類似度として定義した。なお、compareHist関数で得られる類似度は、各色要素の位置に関わらず、各色要素が範囲内に分布する割合のみで決まる。そのため、図10のように違うもののように見える場合でも類似度が100%となる場合がある。

$$d(H_1, H_2) = \frac{\sum_l (H_1(l) - \bar{H}_1)(H_2(l) - \bar{H}_2)}{\sqrt{\sum_l (H_1(l) - \bar{H}_1)^2 \sum_l (H_2(l) - \bar{H}_2)^2}} \quad (\text{式1})$$

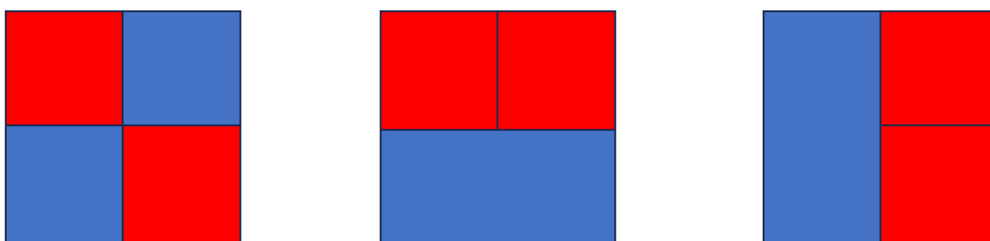


図10 色分布によって類似度が100%となる組み合わせの例

2.4 使用する色空間

OpenCVで画像や動画を読み込んだ場合、基本設定ではRGB色空間で表現されたデータが読み込まれる。RGB色空間は赤、緑、青の3色の組み合わせによって色を表現する。RGB色空間では、明るさの変化によって色が変わった場合などに全く別の色として認識されるという欠点があるため、本研究の目的に適さない。

そこで、本研究ではHSV色空間を使用する。HSV色空間は色相(Hue)、彩度(Saturation)、明度(Value)の3要素で表される。明るさが独立しているため、明るさで見え方が変わった場合でも色相と彩度のみを使えば、同じ系統の色として認識することができるという特長がある。HSV色空間は図11のように、色相を円周方向の角度、彩度を円の中心からの半径方向の距離、明度を高さとする三角錐で表現される。OpenCVでは各値が取り得る値の範囲が色相は0~179、彩度と明度は0~255である。しかし、本研究ではRGB色空間で読み込まれたデータをHSV色空間に変換する際に、COLOR_BGR2HSV_FULLを指定することで、色相も0~255の範囲で表現するように設定している。

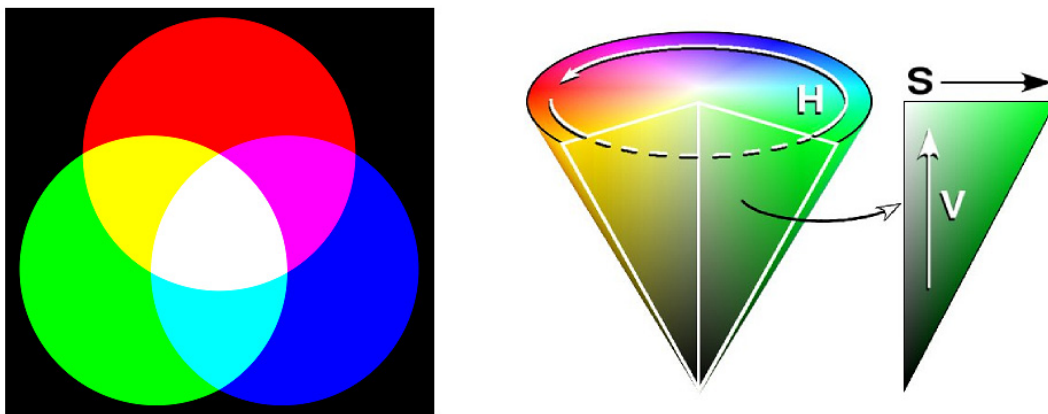


図11 RGB色空間(左)[9]とHSV色空間(右)[10]のイメージ図

2.5 実装想定

本節では識別システムをロボットに搭載し、実用化する場合の実装方法の想定について説明する。

図12は堀越らが作成した人物認識によりジェスチャーで操作する移動ロボットである[11]。人物認識を行うためのカメラとパソコンを移動ロボットに搭載するという構成が本研究と似ているため、本研究で作成する識別システムを実際にロボットに搭載する場合は、同様の構成になると考えている。堀越らは人物認識用センサとしてMicrosoft社製のAzure Kinect、移動ロボットはTurtleBot3 Waffle Piを使用している。また、Azure Kinectが常に対象人物の全身を映すようにするために、金属製のフレームによってAzure Kinectを床から920mmの高さに設置している。

なお、本研究では識別システムのみを作成し、ロボットの制御やロボットのハードウェアの実装は行わない。



図12 堀越らが作成した移動ロボット[11]

2.6 識別システムの仕組み

本節では作成した識別システムについて説明する。本システムではOpenPoseで得られるテンプレート領域の色分布をテンプレートとし、各人物のテンプレート識別領域の色分布をテンプレートと比較した時の類似度で人物識別を行う。テンプレート領域は、2.2節で述べた両肩と両腰の4点から成る四角形の領域である。システムの処理の流れを図13に示す。

はじめに、識別開始時を1フレーム目とし、1フレーム目から図14のように識別対象のみが写るテンプレート作成用画像を作成する。そして、テンプレート領域の色分布をOpenCVのcalcHist関数で取得し、規格化した上で識別対象のテンプレートとする。なお、一般的にはテンプレートを作成する際、テンプレート作成用画像に映っている人物は1名だけであるべきである。しかし、本研究では1つの動画でテンプレート作成と認識の両方を行うため、「画面中央の人物からテンプレートを作成する」という前提を設けている。

続いて、テンプレートを作成した後の識別処理について説明する。識別処理はカメラに写っている人物全員に対して行う(図15)。まず、OpenPoseによってテンプレート比較領域を検出し、その範囲の色分布を取得する。そして、取得した色分布をOpenCVのcompareHist関数でテンプレートと比較し、各人物の類似度を求める。フレーム内の全員に対して処理を行った後は次のフレームを取得し、同様の処理を繰り返す。

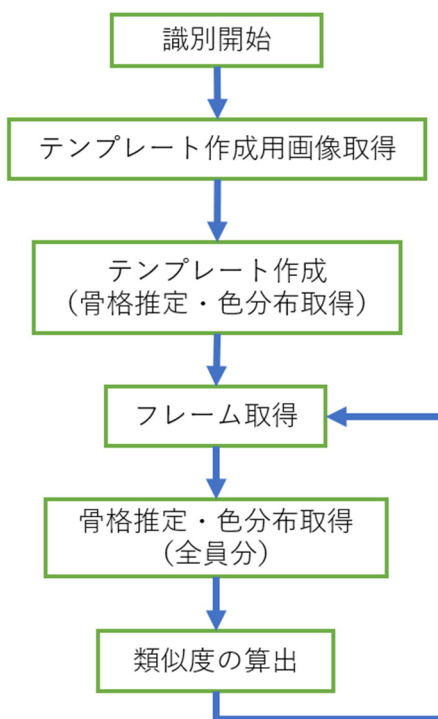


図13 システムの流れ

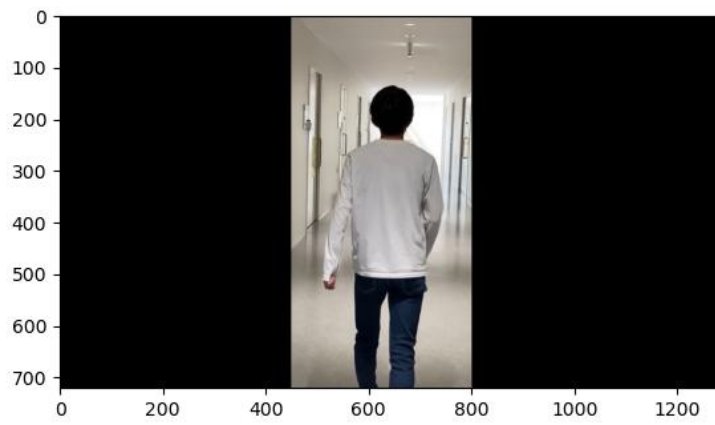


図14 識別対象のテンプレートを取得するためのテンプレート作成用画像

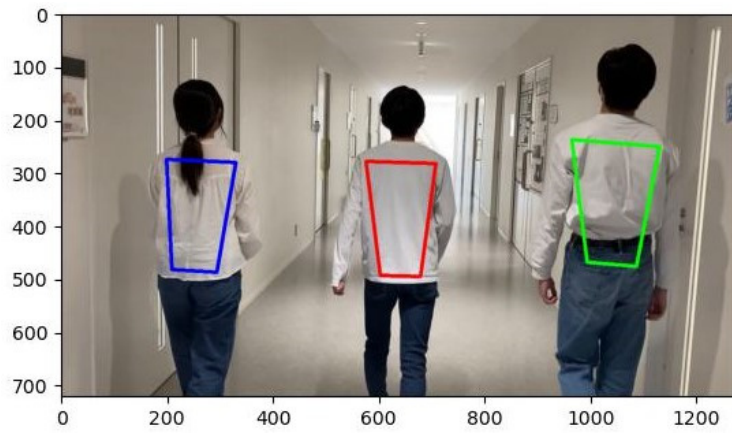


図15 テンプレート比較領域の四角形を表示

4. 実験1：色分布によるテンプレートでの人物識別

実験1として、識別開始時に作成する1つのテンプレートのみを使用して識別を行う仕組みの性能について調べた。

4.1 実験概要

識別開始時に作成する1つのテンプレートを基に動画に対して人物識別を行う。識別開始時の1フレーム目から、図16のように識別対象者のみが写る初期テンプレート作成用画像より、テンプレート領域の色分布のテンプレートを作成する。そして、10秒間（300フレーム）のすべてのフレームに対して、テンプレートと画角内に映る人物との色分布の類似度を取得する。



図16 1フレーム目から得られるテンプレート領域(赤い四角形)

4.2 実験に使用する動画

本システムを実用化する場合は人物を追従するためにリアルタイムで人物を識別することを想定しているが、本研究では事前に用意した動画を使用して実験を行う。実験に動画を使用する理由は2つある。1つめは、処理を行う際の条件を揃えやすいからである。2つめは、OpenPoseの処理速度が遅く、全フレームに対しての処理とリアルタイムでの処理を両立できないからである。本システムでは1フレームの処理に約0.8秒かかるため、実用化する場合に想定しているリアルタイムでの実装が行えない。この問題は、より処理が軽い lightweight-OpenPose[12]を骨格推定に使用することで解決できると考えている。

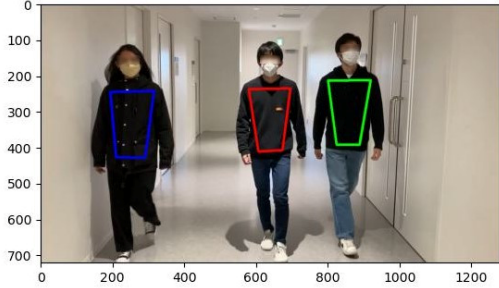
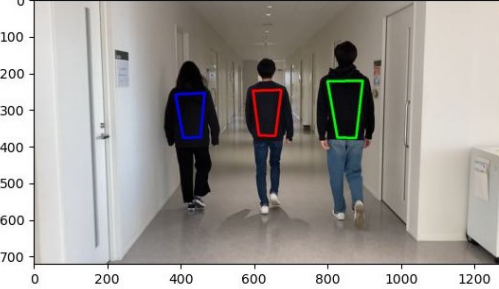
本研究では5つの動画を使用する。すべての動画のフレームレートは30FPSであり、サイズは1260×720ピクセルである。また、実験では10秒間（300フレーム）分のデータを取得する。

続いて、各動画の内容について説明する。各動画の1フレーム目にテンプレート領域とテンプレート比較領域を描画した画像、照明などによる明るさの変化の大きさ、識別対象人物

と周りの人物の服装の違いの大きさ、撮影方向をまとめたものが表1である。動画①は著作権フリー動画[13]である。動画②～⑤は工学院大学八王子キャンパス4号館8階の廊下にて、スマートフォンのカメラを使用して撮影したものである。また、対象の正面を後退しながら撮影した動画と、カメラが対象を追うように背面から撮影した動画があるが、いずれも色分布を取得する部分がすべて写る距離を一定に保ちながら撮影したものである。肩と腰の骨格点是对象の向きに関わらず検出することができ、人物の前後方向の向きは実験に影響が無いと判断した。そのため、実際にロボットで追従する場合は、正面と背面のどちらの映像を用いても影響は無いと考えている。

表1 実験で使用する動画一覧

	1フレーム目	明るさの変化	服装の違い	向き
動画① フリー動画		小	大	背面
動画② 白色正面		大	大	正面
動画③ 白色背面		大	小	背面

<p>動画④ 黒色正面</p>		<p>大</p>	<p>小</p>	<p>正面</p>
<p>動画⑤ 黒色背面</p>		<p>大</p>	<p>小</p>	<p>背面</p>

4.3 実験結果

本節では動画①～⑤に対して実験を行った結果を示す。

4.3.1 実験1.a 服の違いが大きく、明るさの変化が小さい状況での識別(動画①)

まず、動画①に対して識別を行う。動画①は青色の服を着た人物とオレンジ色の服を着た人物を背後から撮影した動画である。動画に写っている2人の服は人間の目で見ても明らかに色の違いが大きいことが分かる。また、明るさの変化が小さいため、識別が行いやすい状況の動画といえる。

図17は1フレーム目の青色の服の人物を識別対象者とし、10秒間の識別を行った時の類似度の変化を示したグラフである。横軸はフレーム数、縦軸は類似度である。また、グラフの色が各人物の服の色に対応している。そして、識別対象をオレンジ色の服の人物に変えて識別を行った結果が図18である。図19は代表的なフレームにおける色相の分布を示したヒストグラムである。

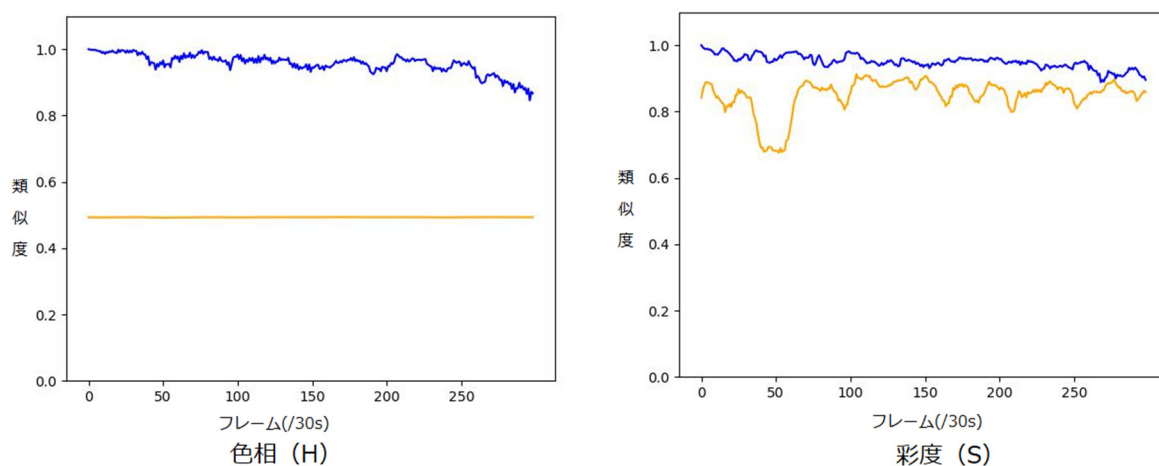


図17 左の人物を識別対象とした時の類似度の変化

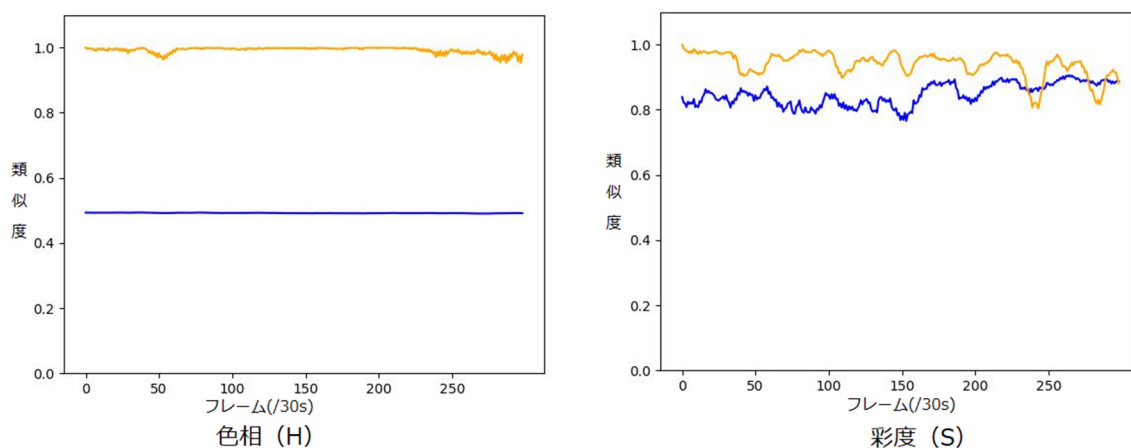


図18 右の人物を識別対象とした時の類似度の変化

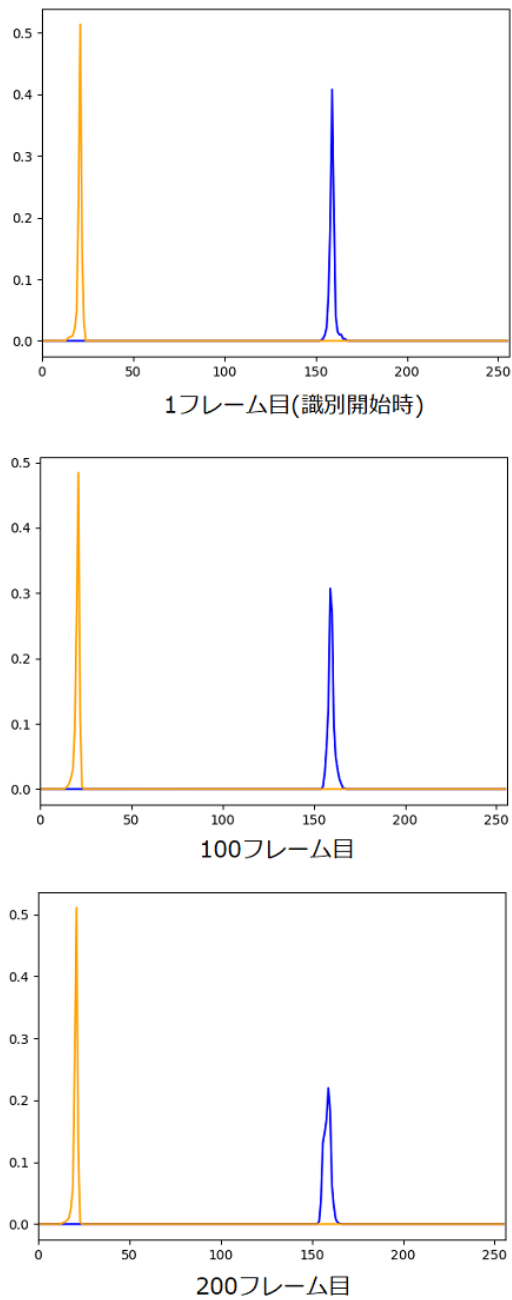


図19 代表的なフレームにおける色相の分布

10秒間を通して、識別対象者の色相の類似度が他の人物の識別対象よりも常に大きくなっている。そして、識別対象者以外の色相の類似度は殆ど変化していない。一方で、色相の類似度と比較すると彩度の類似度は変化が大きく、識別対象者の類似度が他の人物の類似度を下回ることがあった。また、図19より、それぞれの色相が一定の値に集中し、互いの分布が重なっていないことが分かる。それにより、識別対象者の類似度は時間経過で変化しているものの、識別対象者ではない人物の類似度は殆ど変化していない。以上のことから、色相の類似度のみで識別が可能であると考えられる。

4.3.2 実験1.b 服の違いが大きく、明るさの変化が大きい状況での識別(動画②)

続いて、動画②に対して識別を行った。識別対象者である中央の人物の服は左右の人物の服との模様の違いが大きいものの、天井の照明によって、短時間で明るさが大きく変化する動画である。図20の中央の人物のテンプレートを用いて10秒間の識別を行った結果が図21である。1フレーム目の画像に示したテンプレート領域およびテンプレート比較領域の色(図20)に、グラフの色が対応している。

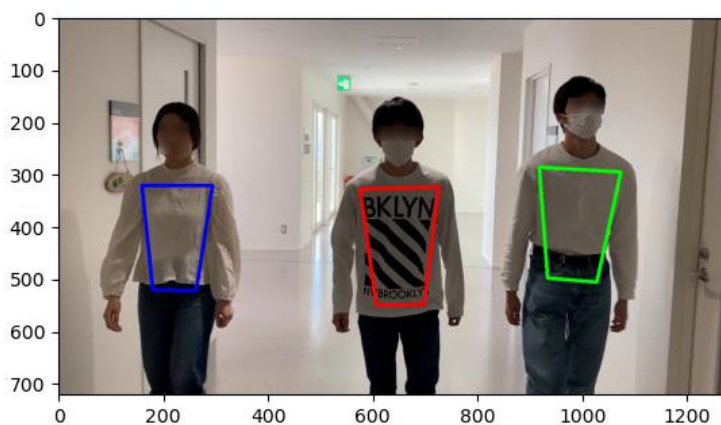


図20 1フレーム目のテンプレート領域(赤)およびテンプレート比較領域(青・緑)

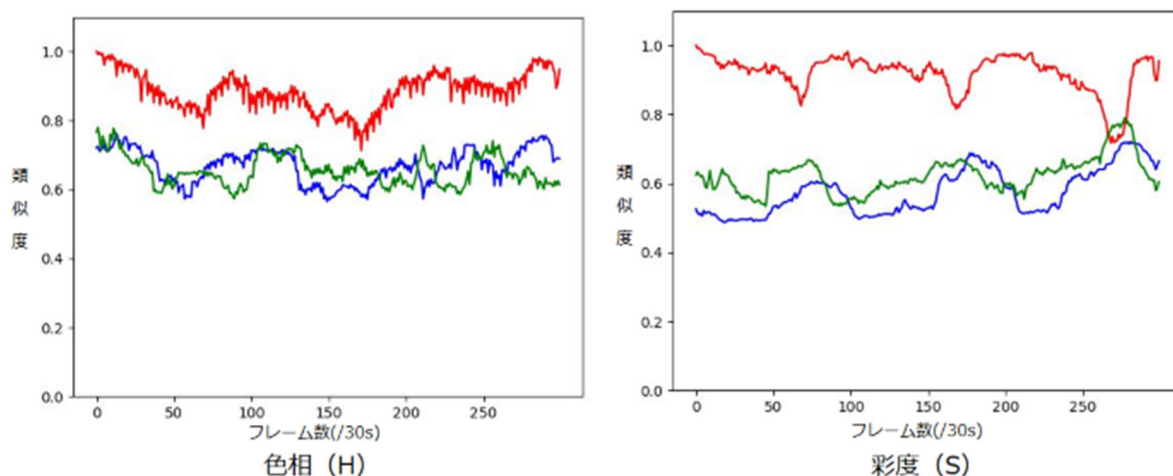


図21 中央の人物を識別対象者とした時の類似度の移り変わり

動画①に対して行った実験1.aと同様に、色相の類似度は常に識別対象者のものが最大となっているため、色相のみで識別が可能であると考えられる。ただし、実験1.aと比較すると全員の類似度の変化が大きく、識別対象者とそれ以外の人物の類似度の差も小さくなっている。しかし、服装の違いが大きいことで他の人物の類似度の最大値は小さくなり、識別対象者の類似度が一時的に低下する場合でも、色相では識別対象者の類似度が常に最大となっている。

4.3.3 実験1.c 服が似ていて、明るさの変化が大きい場所での識別(動画③④⑤)

最後に、動画③④⑤に対して識別を行った。動画②に対して行った実験1.bと同様に明るさの変化が大きいことに加え、動画③④⑤は色分布を取得する部分の色や模様の違いが小さい。いずれも中央の人物を識別対象者とし、図22は各動画の1フレーム目の画像にテンプレート領域とテンプレート比較領域を描画した画像である。そして、図23は各動画での類似度の変化を示したものである。図23のグラフの色は図22の四角形の色に対応している。

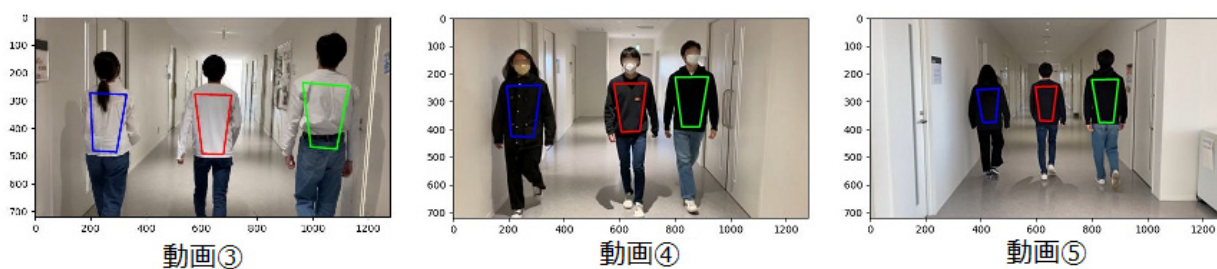


図22 各動画の1フレーム目にテンプレート領域とテンプレート比較領域を描画したもの

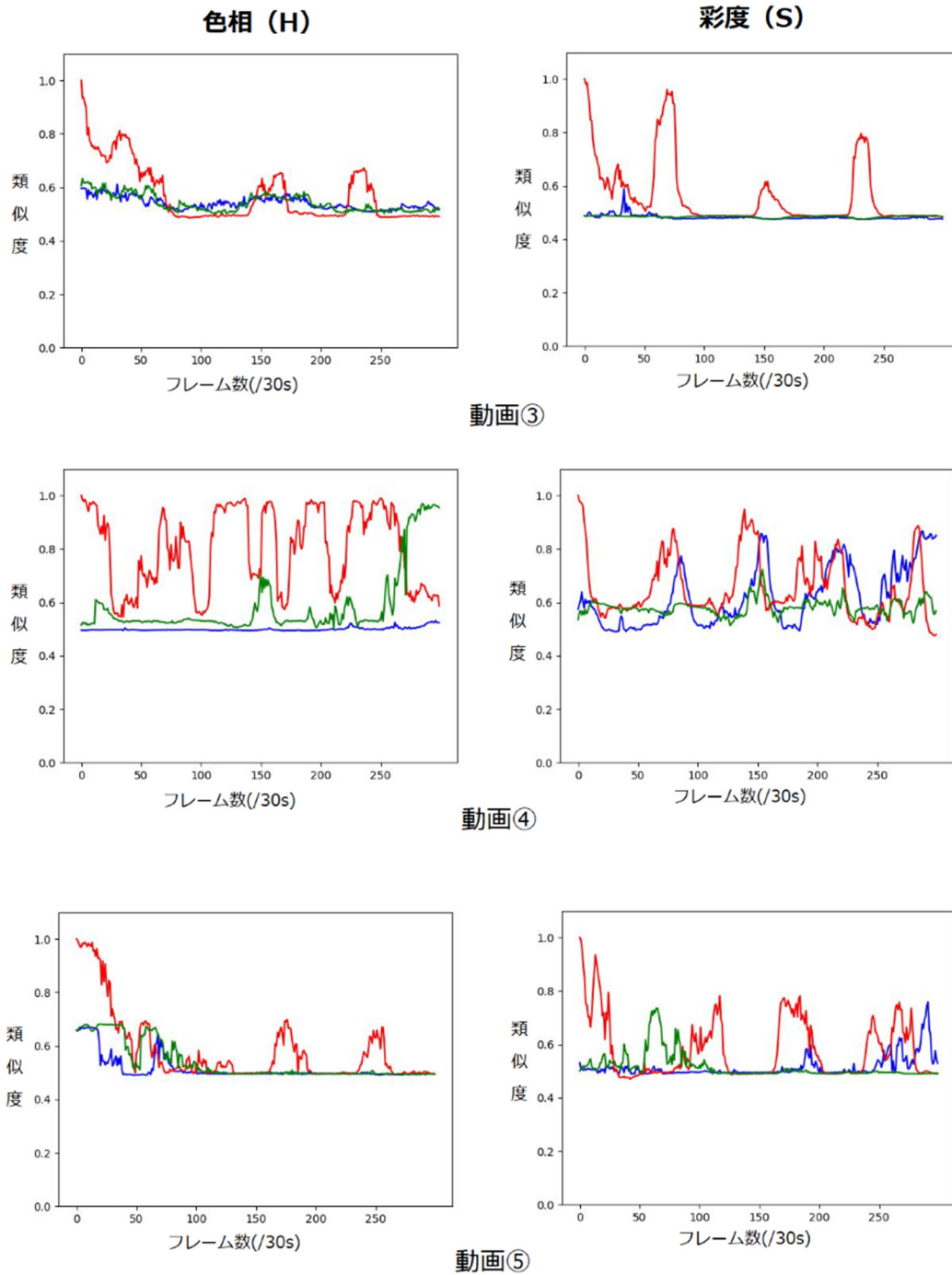


図23 中央の人物を識別対象者とした時の各動画の類似度の移り変わり

本実験では色相と彩度ともに時間経過による類似度の変化が大きく、識別対象者の類似度がその他の人物の類似度が下回ることもあった。そのため、前述した動画①②での実験とは異なり、類似度を使用した識別は困難であることが分かった。

4.4 考察

識別対象者の服装が周囲の人物のものと大きく異なる場合は、照明による明るさの変化が大きい状況でも色相の類似度によって継続的に識別を行うことが出来た。一方で、識別対象者の服装が周囲の人物のものと似ている場合は、識別開始直後は識別が可能であるが、継続的な識別は難しいことが分かった。本節では、そのような状況で類似度が低下している原因を考察する。図24は動画③で識別を行った際の色相の類似度の変化を示すグラフである。そして、図25は類似度が低下していく際の3枚のフレームを抽出した画像である。図25で画像を抽出したフレームの位置を図24に丸印で示している。さらに、図26は図25の画像の彩度と明度を最大にすることで、色相の違いを分かりやすくした画像である。

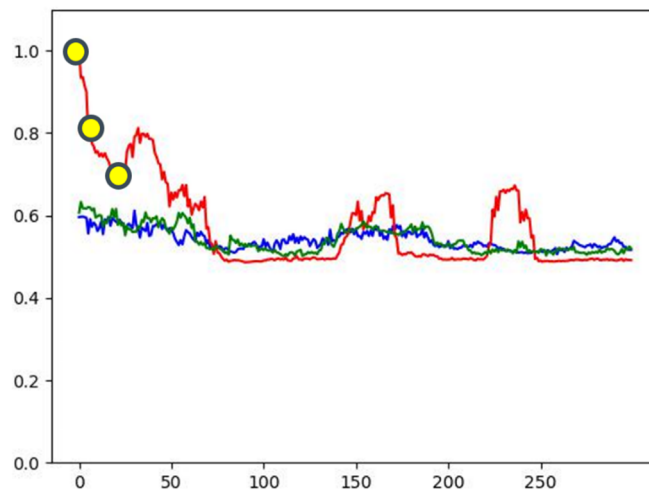


図24 動画③の色相の類似度の変化(図25の画像抽出位置に○を表示)

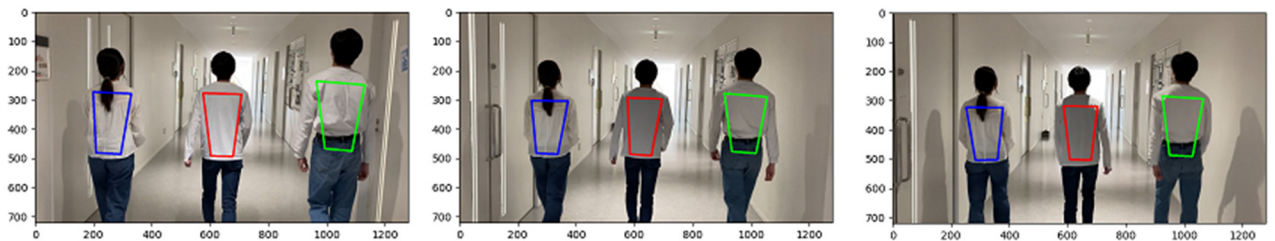


図25 図24の丸印の位置のフレーム画像

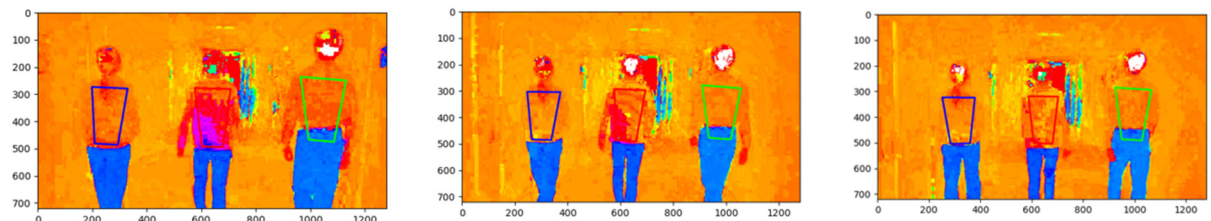


図26 彩度と明度を最大にし、色相の違いを可視化

まず、図26の3人の色相を1フレーム内で比較すると、同じような白色の服に見えても実際は色相に違いがあり、この違いによって似ている服装でも識別が可能となっている事が分かる。一方で、今度は識別対象者に注目し、3枚の画像を比較すると、時間経過によって色相が変化している事が分かる。これにより、識別開始から時間が経つと識別対象者の類似度が低下し、類似度による継続的な識別が困難となる事が分かる。動画①や②のように識別対象者とその他の人物の服装が大きく異なれば、識別対象者以外の人物の類似度が大きくなることは無い。そのため、時間経過によって識別対象者の類似度が小さくなっても、その他の人物の類似度が識別対象者の類似度を超えることが無く、それによって継続的な識別が可能となる。

5. 実験2：人物識別中のテンプレートの更新

本章ではテンプレートの更新による識別機能の改良の実験について説明する。

5.1 テンプレートの更新

実験1では識別対象者の服装に関わらず、識別開始直後は識別が正しく行えることが分かった。しかし、時間経過によって識別対象者の類似度が低下するため、識別対象者の服装が他の人物のものと似ている状況では継続的な識別が困難だった。そこで、一定時間毎にテンプレートを更新することで、時間経過による類似度の低下に対応できるようになると考えた。

5.2 実験2

実験2では、一定時間毎のテンプレートの更新の効果を調査する。一定時間毎に、そのフレーム内で色相の類似度が最大である人物の色分布を新たなテンプレートとして更新する。

5.2.1 実験2.a 動画②での実験

最初に、実験2.aとして、実験1で色相による継続的な識別が成功した動画②に対して90フレーム(3秒)毎にテンプレートの更新を行った場合の色相の類似度の変化を調べた。30fpsの動画で10秒間の識別を行ったため、3・6・9秒の3回、テンプレートの更新を行っている。図27がテンプレートの更新を行わなかった場合(実験1.b)の色相の類似度の変化を示すグラフ、図28が90フレーム毎にテンプレートの更新を行った場合の色相の類似度の変化を示したグラフである。

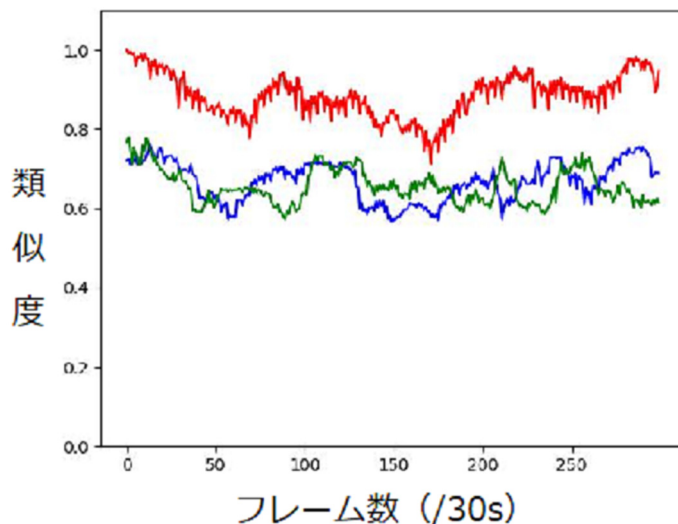


図27 テンプレートの更新を行わなかった場合(実験1)の色相の類似度の変化

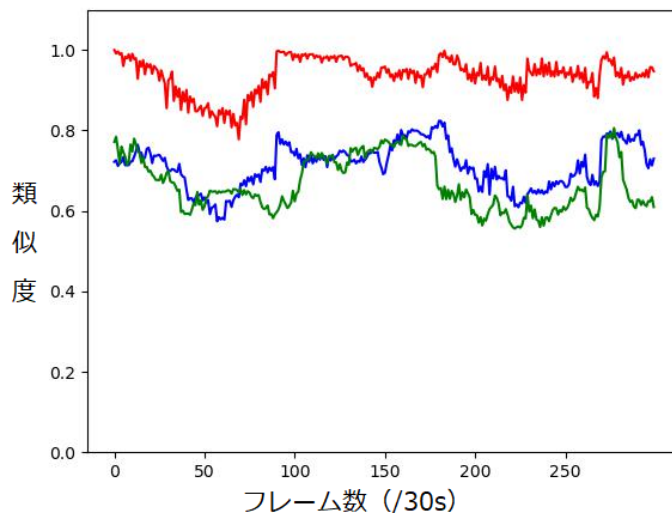


図28 90フレーム毎にテンプレートの更新を行った場合(動画②)の色相の類似度の変化

図27と図28を比較すると、低下した類似度がテンプレートの更新によって回復し、識別対象者の類似度がその他の人物の類似度に接近することが減少している。そのため、より安定した識別が可能となり、テンプレートの更新が有効であることが分かった。

5.2.2 実験2.b 動画③での実験

続いて、実験1で継続的な識別が難しいと判断した動画③に対して、90フレーム(3秒)毎にテンプレートの更新を行った場合の色相の類似度の変化を調べた。図29がテンプレートの更新を行わなかった場合(実験1.c)の色相の類似度の変化を示すグラフ、図30が90フレーム毎にテンプレートの更新を行った場合の色相の類似度の変化を示したグラフである。

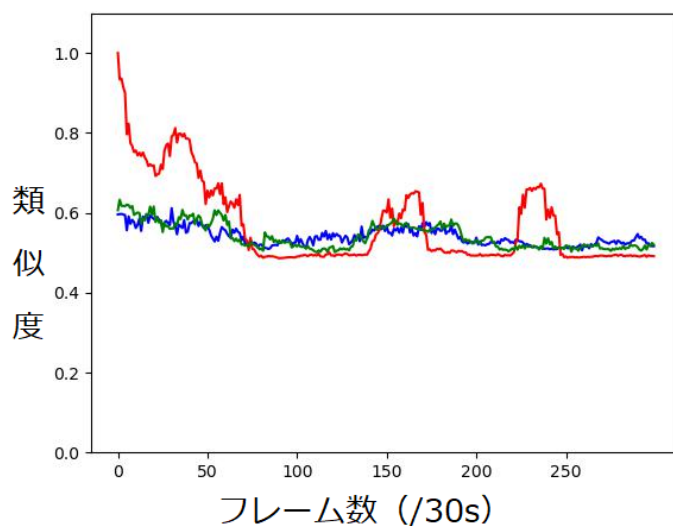


図29 テンプレートの更新を行わなかった場合(実験1.c)の色相の類似度の変化

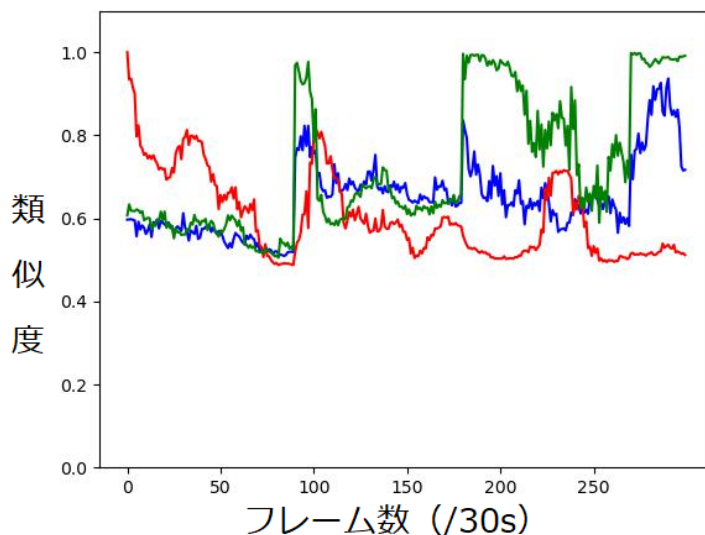


図30 90フレーム毎にテンプレートの更新を行った場合(動画③)の色相の類似度の変化

図30では識別開始直後は識別対象者の類似度を示す赤色のグラフが最大であったが、一回目の更新を行った90フレーム目では緑色のグラフが最大となったため、新たなテンプレートが緑色のグラフの人物の色分布へ入れ替わってしまった。

5.2.3 実験2.c テンプレート更新間隔の変更

実験2.bでは識別対象者の類似度がテンプレートの更新前に他の人物の類似度を下回り、その結果、識別対象者以外の色分布が新たなテンプレートとして選ばれてしまうという問題が生じた。そこで、更新間隔をより短くすれば、識別対象者の色分布を新たなテンプレートとして更新し続けられるようになると考えた。

実験2.cでは、更新間隔を90フレーム、60フレーム、30フレーム、15フレーム、5フレームに変更した場合の類似度の変化を調べる。使用する動画は実験2.bと同様に動画③である。また、動画③のフレームレートは30fpsであるため、秒に換算すると3秒、2秒、1秒、0.5秒、0.17秒である。それぞれの更新間隔で識別を行った際の色相の類似度を示したグラフが図31である。

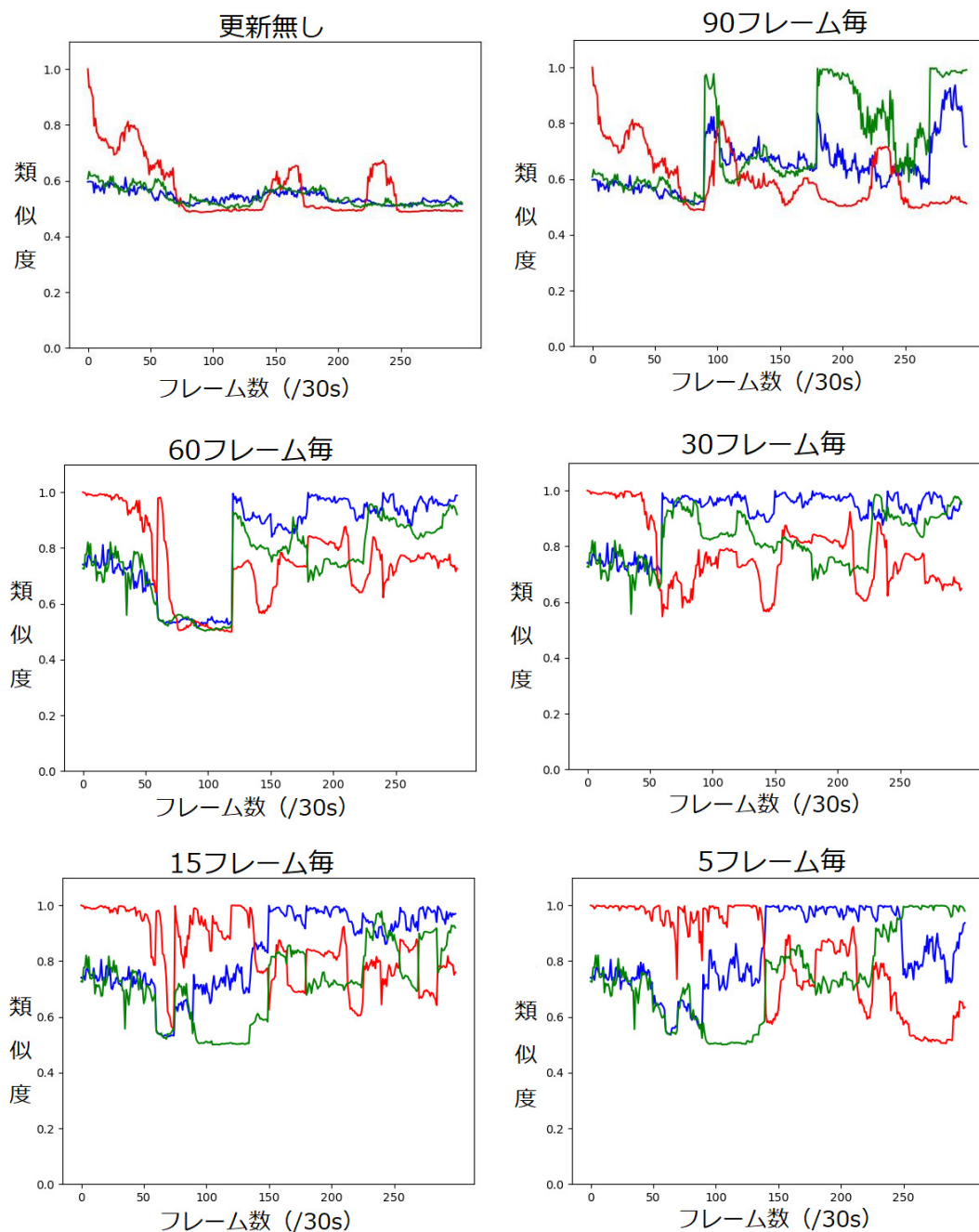


図31 動画③での5種類のテンプレート更新間隔による識別結果

最も更新頻度が高い5フレーム毎の識別でも継続的な識別は行うことが出来なかった。識別対象者の類似度が最大でなくなり、テンプレートの取り違いが起こる場合の類似度の変化として、2パターンの特徴が見られる。

1つめは、識別対象者の類似度が時間経過によって減少し、最終的には識別対象者以外の人物の類似度と同様の変化をするようになることで取り違いが起こるパターンである。これは、90フレーム、60フレーム毎の更新の時に見られた。これは、照明などによって色相が

変化することで、識別対象者の類似度が識別対象者以外の人物たちの類似度と同じ値に収束していくためであると考えられる。

2つめは、識別対象者の類似度が低下していくと同時に、識別対象者以外の類似度が向上した場合に最大のもので入れ替わってしまうパターンである。これは、短い時間でテンプレートの更新が行われる際に発生していた。テンプレートの更新の直後に急な環境の変化が生じた場合、テンプレートの更新は識別対象者の色分布で行われても、その直後の環境の急な変化によってテンプレートの更新の効果が小さくなってしまおうと考えている。

また、実際に人物に追従する場合はリアルタイムで処理を行う必要がある。それに従い、テンプレートの更新もリアルタイムで処理を行える間隔にしなければならない。そのため、30fpsの動画の全フレームに対して処理を行うことは難しく、また、本実験では5フレーム毎の更新でも継続的な識別を行うことはできなかった。したがって、識別対象者とそれ以外の人物の服装が似ていて、照明による明るさの変化が大きい動画③のような状況では、一定時間毎にテンプレートを更新する手法は有効でないと言える。

5.3 実験2.d テンプレートの更新手法の変更

実験2.a～2.cでは単一のテンプレートを一定時間毎に更新する手法が、照明による明るさの変化の対策として有効でないことが分かった。そこで、複数のテンプレートを用意し、それぞれのテンプレートと比較した中で最も高い類似度を、そのフレームでの類似度とする手法について実験を行う。この手法は今井ら[14]が使用していた手法である。今井らは直近の5フレームの色分布を5つのテンプレートとしていた。本実験では処理速度を考慮し、10フレーム間隔の5枚をテンプレートとし、10フレームごとに最も古いテンプレートを新たなテンプレートへ更新するものとした。動画③④⑤に対して処理を行った時の色相の類似度の変化を示したグラフが図32である。

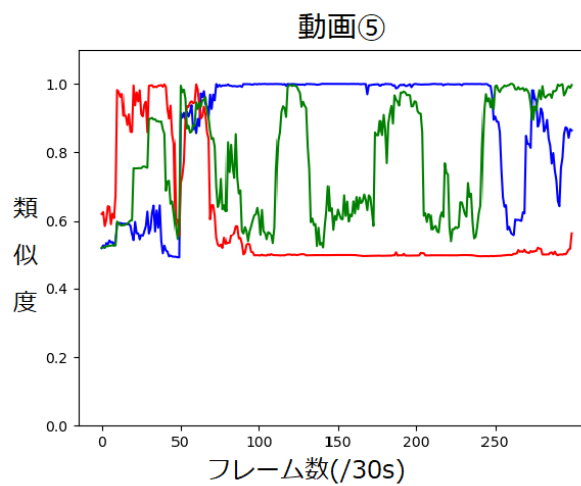
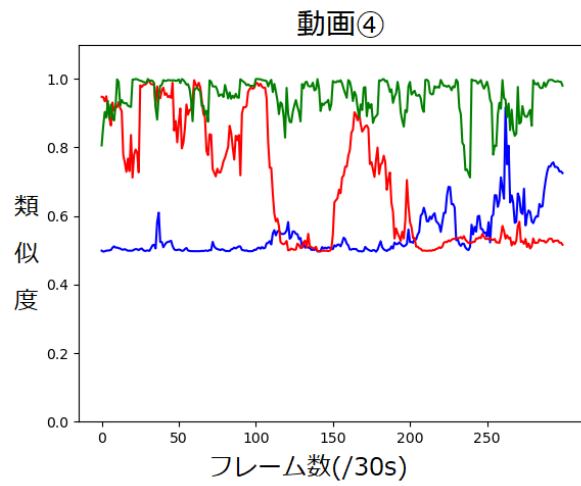
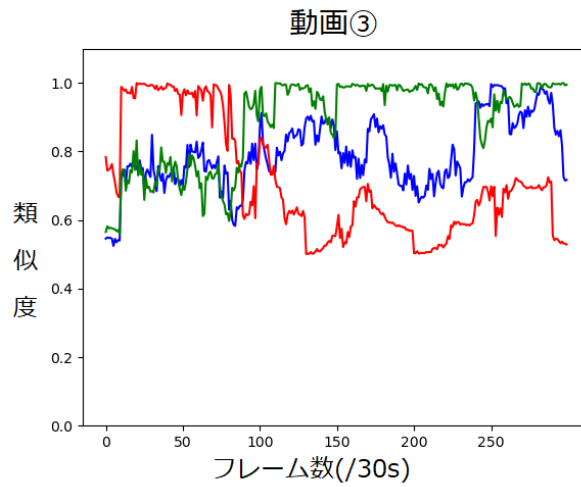


図32 複数のテンプレートを基にした識別

赤いグラフが示す人物が識別対象者だが、いずれの動画でも早い段階でテンプレートの取り違いが起きているため、本来の目的である識別は失敗している。しかし、別の人物に識別対象が移ってしまった後のグラフを見ると、新たなテンプレートとなった人物の類似度

が常に最大となっている。つまり、取り違えてしまった人物が本来の識別対象者だった場合、このテンプレート更新手法によって継続的な識別が可能であると考えられる。また、服の色や素材によって照明による明るさの変化の影響の受けやすさに違いがあり、影響を受けにくい服ほど類似度が低下しにくいと考えている。実験2.dより、複数のテンプレートを使用することが照明による明るさの変化への対策として一定の効果があることが分かった。

6. 結言

本章では、実験から得られた結論と課題について述べ、その後、それらの課題の解決案を先行研究の例を挙げて説明する。

6.1 結論

本研究では、利用者の荷物を運ぶ人物追従ロボットのための対象識別機能を作成し、その性能について調べた。実験1では、識別対象者の服装が周囲の人物と大きく異なる場合は、照明による明るさの変化が大きい状況でも対象を継続的に識別できることが分かった。しかし、識別対象者の服装が周囲の人物の服装と似ている場合、識別開始直後は正しい識別が可能だが、時間の経過とともに類似度が低下するため、継続的な識別は難しいことが分かった。そこで、テンプレートを定期的に更新することで、時間経過による類似度の低下に対応できると考え、2種類のテンプレートの更新方法を実装し、実験2を行った。識別対象者の服装が周囲の人物と大きく異なる場合は、類似度の変化が小さくなり、より安定した識別が可能となった。一方で、識別対象者の服装が周りの人物のものと似ている場合ではテンプレートの更新の効果は不十分であり、継続的な識別は困難であることが分かった。

6.2 課題とその解決案

本節では実験1と実験2から得られた課題と、先行研究を参考にした解決案について説明する。

6.2.1 課題

実験1と実験2より、識別対象者の服装が周囲の人物の服装と似ている場合、時間経過によって識別対象者の類似度が低下するため、継続的な識別ができないことが課題として残った。また、テンプレートの更新は色分布の類似度による識別の性能向上に一定の効果はあったものの、識別対象者の服装が周囲の人物のものと似ている状況には対応することが出来なかった。そのため、色情報以外の方法によってテンプレートの更新を補助することが必要であると考えた。

6.2.2 解決案1 骨格推定情報の活用

本研究ではOpenPoseで得られる骨格推定情報を、色分布を取得するためのテンプレート領域とテンプレート比較領域を定めるためだけに使用していた。1.2.2項で述べたように、下山らは骨格推定情報から求められるトルソ(図4)によって人物識別を行っていた[4]。その研究結果より、骨格推定情報による人物識別手法は有効であるものの、骨格が似ている人物どうしの識別が課題であるとしていた。そのため、我々が用いた色分布と骨格推定情報を組み合わせることは効果的であると考えられる。

6.2.3 解決案2 深度情報や座標情報の併用

本研究ではロボットの制御は行わなかったため、二次元の動画に対してのみ人物識別を行った。しかし、ロボットで人物を追従するためにはロボットと人物の距離を一定に保つ必要があり、そのためには深度情報を取得することが出来るRGB-Dカメラが有効であると考えている。そして、深度情報をロボットと人物の距離を一定に保つためだけでなく、人物識別にも活用できると考えている。今井らは色情報と深度情報を併用することで、以下の3つの仮定のもとに、物体の見え方の変化に対応できる物体追跡手法を実現した[14]。

- (1)画像上の追跡対象物体の領域における色分布は、連続する2フレーム間で大きく変化しない。
- (2)画像上の追跡対象物体の領域の大きさは、連続する2フレーム間で大きく変化しない。
- (3)同一フレームにおける追跡対象物体の領域内の深度変化は連続的である。

これらの仮定に加え、文献[14]ではパーティクルフィルタを用い、パーティクルの座標の変化が等速変化と揺らぎで構成されることを仮定している。これは、追跡対象物体の2フレーム間の変化が大きくないことも仮定している事を意味する。このように、深度情報と座標情報を併用することで、安定した追跡が可能になると考えられる。

文献[14]における物体追跡の様子が図33である。カメラの映り方によって見え方が変化していても、継続的に対象の物体を抽出できている事が分かる。この手法を用いることで、色分布が大きく変化した場合でもテンプレートを正しく更新することができ、継続的に人物を識別できると考えている。ただし、OpenPoseを使用する場合、仮定(1)(2)を満たすための処理フレームの確保に注意が必要である。

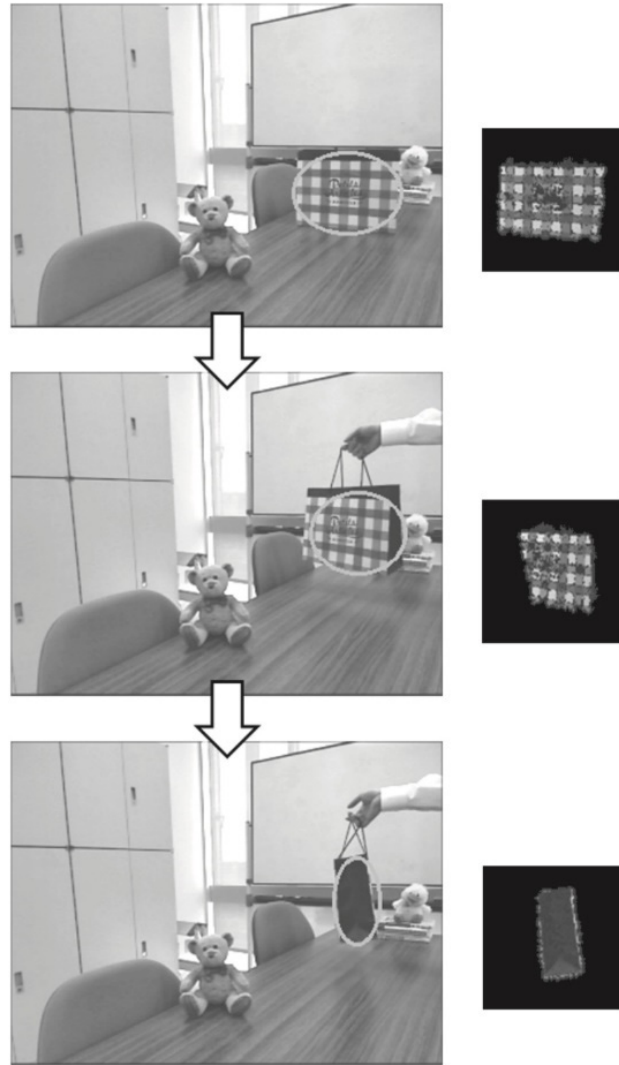


図33 色分布と深度情報の併用による物体追跡の様子[14]

参考文献

- [1] TDK Product Center, “ロボット掃除機に搭載されているセンサの仕組み”, TDK株式会社,
<https://product.tdk.com/ja/techlibrary/applicationnote/how-service-robots-make-sense.html>
- [2] NECプラットフォームズ, “ネコ型配膳ロボット「BellaBot（ベラボット）」”,
<https://www.necplatforms.co.jp/solution/food/store-solution/delivery-robot/bellabot/index.html>
- [3] 松原 慎也, 池 勇勳, 梅田 和昇, “2台のカメラを用いた骨格抽出による移動ロボットの3次元人物追跡”, 精密工学会誌, vol.88, no.3 (2022) pp.276-281.
https://www.jstage.jst.go.jp/article/jjspe/88/3/88_276/_article/-char/ja/
- [4] 下山 未来, 浅田 郁弥, 加藤 宏一朗, 松日楽 信人, “人物追従ロボットのための深度センサによる頭部と胴体情報を用いた追従制御法の提案とその評価”, 日本機械学会論文集 Vol.87, No.901, 2021
https://www.jstage.jst.go.jp/article/transjsme/87/901/87_20-00133/_article/-char/ja
- [5] 小山 渚, 但馬 竜介, 田中 稔, 廣瀬 徳晃, 鋤柄 和俊, 藤井 亮暢, “ヒトを活動的にする移動支援ロボット —IDを発信する赤外LEDタグと全方位カメラを用いた利用者の追跡—”, 計測自動制御学会論文集 Vol.51, No.1,49/56, 2021
https://www.jstage.jst.go.jp/article/sicetr/51/1/51_49/_article/-char/ja/
- [6] Zhe Cao and Tomas Simon and Shih-En Wei and Yaser Sheikh. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. CVPR. 2017
<https://arxiv.org/abs/1611.08050>
- [7] Hzzone, “pytorch-openpose”, github, <https://github.com/Hzzone/pytorch-openpose>
- [8] opencv 2.2 documentation, “imgproc. 画像処理”,
http://opencv.jp/opencv-2svn/cpp/imgproc._image_processing.html
- [9] Wikimedia commons, “File:Synthese+.svg”,
<https://commons.wikimedia.org/wiki/File:Synthese%2B.svg>

[10] Wikimedia commons, “File:HSV cone.jpg”,

https://commons.wikimedia.org/wiki/File:HSV_cone.jpg

[11] 堀越優来, 丸山拓己, 宮崎生望, 目黒新大, 吉武敏, “人物認識によりジェスチャーで操作する移動ロボットの開発”, 工学院大学 知能機械研究室 2021年度卒業論文

https://brain.cc.kogakuin.ac.jp/research/202203_GestureControl.pdf

[12] Daniil-Osokin, “Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose”, arxiv, <https://arxiv.org/abs/1811.12004>

[13] 動画AC, “自転車を押して歩く夫婦の後ろ姿”, <https://video-ac.com/video/2597>

[14] 今井 順一, 柏木 雄平, 木辻 亮, “色情報と深度情報を併用した自己遮蔽に頑健な物体追跡手法”, 計測自動制御学会論文集 Vol.55, No.5, 342/352 2019

https://www.jstage.jst.go.jp/article/sicetr/55/5/55_342/_article/-char/ja/

謝辞

本研究を行うにあたり、2年間ご指導くださった金丸隆志教授、そして、実験に協力してくださった知能機械研究室の皆様深く感謝申し上げます。