

# パーツ分割を利用したオクルージョンに頑健な人物同定

中島 健太郎<sup>†</sup> 梅澤猛<sup>‡</sup> 大澤範高<sup>‡</sup>

千葉大学工学部<sup>†</sup> 千葉大学大学院融合科学研究科<sup>‡</sup>

## 1. はじめに

アピアランスベースによる人物同定では、色情報を用いる手法が一般的である。Farenzena らは検出対象の領域分割による高精度な同定手法 [1] を提案している。分割した領域それぞれから色情報を取得し、その模様やヒストグラムの空間的組み合わせを考慮することにより同定精度を上げるものである。しかし、この手法は人物の全体像が常に取得できることを前提としているため、オクルージョンに対応できないという問題がある。そこで本研究では、検出対象の領域を分割するという手法を用い、分割された領域の一部からでも同定できるようにすることで、パーツ分割同定でのオクルージョンを考慮した手法を提案する。

## 2. 概要

### 2.1. 人物領域取得

人物領域の取得には動的背景差分法を用いる。動画像の初期の数フレームを基に背景のモデルを作成し、前景を検出する。背景のモデルは、逐次更新を行い、ある程度の輝度変化には追従可能とする。動的背景差分法の輝度値モデルは森田ら [2] の手法を用いた。

### 2.2. 人物同定

取得した人物領域にパーツ分割をし、人物同定を行う。本研究では、簡単化のため上半身と下半身の 2 つのパーツでの分割とした (図 1)。なお図 1~3 および実験においては文献 [3] のものデータセットを利用した。



図 1-a: 上半身



図 1-b: 下半身

図 1: 分割模様

人物領域を HSV 色空間に変換し、分割した人物領域のマスク画像を基に、背景領域以外の HSV ヒストグラムを計算する。そして分割した人物領域それぞれを HSV ヒストグラムによるテンプレートマッチングを行う。テンプレートマッチングによる評価指標には Bhattacharyya 距離を

用いる。Bhattacharyya 距離とは、統計学の分野において、2 つの離散的確率分布間の距離を求める尺度として用いられている指標で、画像間の類似性評価として代表的な指標でもある。そして各チャンネルにおいて計算された Bhattacharyya 距離のユークリッド距離を求めて最終的な類似度とする。類似度がものを対象人物として同定する。ただし、値が 0 に近いほど類似度が高く、画像中で最も類似度が高かったものを対象人物とする。

## 3. 実験

静止画像と動画像での実験を行った。Processing [3] の画像サイズ 360×240 のデータセットを用いた。動画像実験では、このうち 44 フレームを動画入力として用い、静止画像実験で用いた画像は動画像実験で用いたデータの 1 フレームである。ただし、このデータセットではオクルージョンは発生しない。静止画像実験でオクルージョンへの対応を確かめ、動画像実験では、オクルージョンの設定無しに、今後の改善のために必要だと思われる類似度の値や、誤検出率、最終的な結果として用いられたパーツについて調べた。ヒストグラムや類似度や計算方法は変えず、パーツ分割をしないで検出を行う既存手法と比較した。

### 3.1. 静止画像実験

オクルージョンがデータセットでは発生しないため、画像加工により静的に設定した。図 2、3 が結果画像である。



図 2: 提案手法



図 3: 既存手法

図 2 では上半身と下半身の検出枠を正方形で、類似度の高い方に人物領域の大きさの長方形の矩形を表示するようにしている。図 3 で検出できていないものを図 2 では長方形の矩形で正しく検出できているのが分かる。

Robust Human Identification against Occlusion using Part Division

<sup>†</sup>Kentaro Nakajima, School of Engineering, Chiba University

<sup>‡</sup>Takeshi Umezawa, Noritaka Osawa, Graduate School of Advanced Integration Science, Chiba University

### 3.2. 動画実験

評価目的を確かめるためにオクルージョンを設定していない動画像を用いて実験した結果を図4~6に示す。

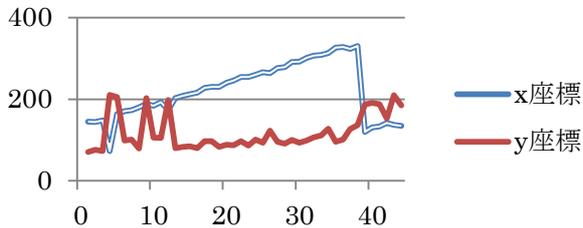


図4: 座標(提案手法)

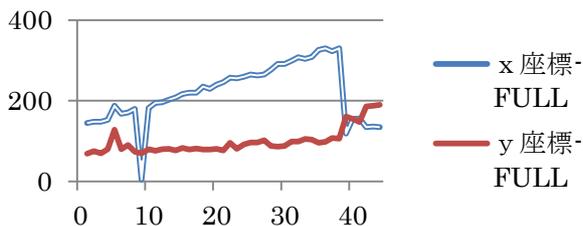


図5: 座標(既存手法)

図4, 5は既存手法と提案手法で検出された矩形の座標をそれぞれグラフ化したものである。この動画像では対象人物は1方向に静止することなく動き続けているため、座標は連続的であるはずである。よって図4, 5にあるように座標が急激に変化しているフレームが、誤検出が起きたフレームだと推定できる。

全44フレーム中でこの誤検出の回数を数え、誤検出率を計算した。その結果が表1である。

表1: 誤検出率

	提案手法	既存手法
誤検出回数	16	7
確率(%)	37.2	16.3

次に提案手法の長方形の矩形の決定にどちらのパーツが用いられたかをフレームごとにグラフ化したのが図6である。

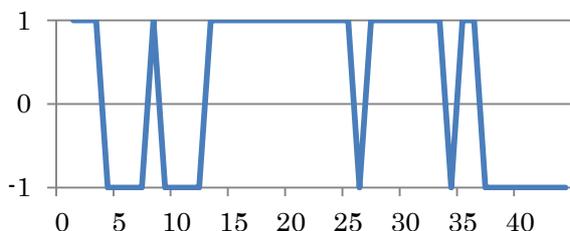


図6: 使用したパーツ(-1:下半身\_1:上半身)

上半身のパーツが多く用いられているのが分かる。だが、図4, 5と図6を比較すると、誤検出が起きているフレームでは下半身が用いられていることが分かる。

### 4. 考察

動画実験に用いたデータセットはオクルージョンが発生しないが、既存手法のほうが精度がよいという結果になった。これはおそらくパーツが個人の特徴を濃く反映しているかに問題があると考えられる。比較したパーツが多量の人物に共通する特徴をもっていた場合、違う人物でも高い類似度を示してしまうことがある。図6と比較してみるとわかるように図5で座標が急激に変化しているフレーム時には下半身の特徴量が用いられている。下半身は上半身に比べて一般的に色の変化に乏しいため、違う人物の下半身でも高い類似度になってしまったと考えられる。

### 5. まとめ

4で述べたように、下半身(色の変化に乏しいパーツ)の扱いを考えていく必要がある。具体的には、色の変化に乏しいパーツには重みづけをして類似度を計算するのがよいと考える。登場人物のパーツのヒストグラムを参考にしきい値を決めることで解決を図っていこうと考えている。重みづけをするためには、画像中に登場する人物のパーツを全て比較する必要がある。

オクルージョンの発生する動画像での実験も行い、オクルージョンに対しての提案手法の有効性を示していく。

発展として、動画像にすることで時間的な特徴も用いることができるようになるので、時間的特徴の利用や、パーツの分割方法についての検討している。

### 参考文献

- [1]M. Farenzena, L. Bazzani, A. Perina, V. Murino, M. Cristani"Person Re-Identification by Symmetry-Driven Accumulation of Local Features", Proc. ECCV, pp. 383-396, 2012
- [2]森田 真司, 山澤 一誠, 寺沢 征彦, 横矢 直和: "全方位画像センサを用いたネットワーク対応型遠隔監視システム", 電子情報通信学会論文誌 (D-II), Vol. J88-D-II, No. 5, pp. 864-875, (2005.5).
- [3]Computer Vision With Processing <http://www.vision.cs.chubu.ac.jp/VU/html/data.html>

### 謝辞

本研究は JSPS 科研費 21300318 の助成を受けたものです