

SIFTを用いた人物対応付けの向き変化に対する耐性の検証
 Investigate robustness of SIFT feature descriptor
 to variation in Human pose for re-identification

岡本 大和† Yamato Okamoto 川西 康友‡ Yasutomo Kawanishi 椋木 雅之‡ Masayuki Mukunoki 美濃 導彦‡ Minoh Michihiko

1. まえがき

防犯やマーケティングの観点から、複数のカメラによる広範囲の人物追跡を実現することは有用である。そのために、異なるカメラ映像間で同一人物を対応付ける研究が盛んに行われている。

本田ら[1]は、映像中の各人物領域から抽出した色ヒストグラムを比較することで人物を対応付けた。しかし、カメラによって明るさが異なる場合や、服装の配色が似たような人物に対して、色のみでは対応付けが困難である。

一方、画像の局所的な勾配情報に基づく SIFT 記述子は識別力が高く、明るさやスケールの変化に頑健であることから近年さかんに利用されている。人物対応付けでは人物の服の模様や形が手掛かりとなるが、勾配情報に基づく SIFT 記述子はこれらをキーポイントとして検出・比較して類似性を捉えることができる。また、防犯カメラ映像では人物領域の明るさや大きさが異なることがあるので、これらに影響を受けにくい SIFT 記述子は人物対応付けに有利である。しかし、観測された人物の向きが同じでなければ特徴的な部分の見え方が変化するため対応付けできない。そのため映像間の人物対応付けに SIFT 記述子はあまり用いられてこなかった。

本研究では通路や廊下などに設置された防犯カメラ映像といった、映像間で人物の向き変化が小さい場合の人物対応付けを想定する。人物対応付け結果に影響を及ぼさない程度の向き変化しかない場合には、明るさやスケールの変化に頑健な SIFT 記述子を優先的に用いるべきである。そこで、向き変化がどの程度までであれば SIFT 記述子によるカメラ映像間の人物対応付けができるのか、人物を各角度で撮影したデータを用いて検証する。

2. 異なる映像間の人物対応付け

人物対応付け手順について説明する。映像 A 中の対応付けしたい人物領域 x^a に対して、映像 B 中の同一人物の可能性のある人々を対応付け候補者と呼び、それぞれの人物領域を $x_1^b, x_2^b, \dots, x_n^b$ と表記する。 $x_1^b, x_2^b, \dots, x_n^b$ それぞれから抽出した特徴を x^a の特徴と比較して、どの程度類似しているのかを示す類似度 $f(x^a, x_i^b)$ を計算する。同じ人物からは同じ特徴が抽出できるはずなので、最も類似度が大きい人物 \hat{x} を

$$\hat{x} = \arg \max_{x_i^b} f(x^a, x_i^b) \quad (1)$$

で算出し、 x^a と対応付ける。 SIFT 記述子同士の類似度 $f(x^a, x_i^b)$ は、各人物領域から検出されたキーポイント同士を比較して、類似したキーポイントの個数とする(図 1)。

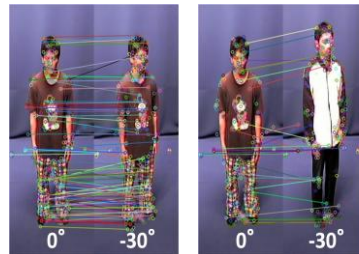


図 1: キーポイントの比較

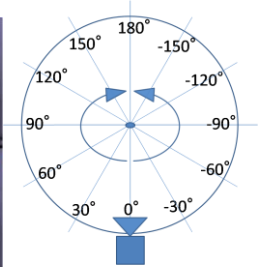


図 2: 撮影環境

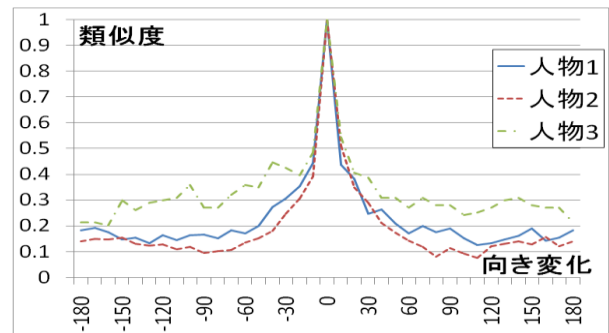


図 3: 人物の向き変化と類似度の関係

3. 向き変化に対する耐性の検証

同一人物において、向き変化によって類似度がどう変化するのか、また、向き変化の大きさがどの程度ならば SIFT 記述子により同一人物を対応付けできるのかを調査する。実験のために人物を様々な向きで撮影した画像を用意する。

3.1 実験データ撮影

まず、実験用に向きの異なる人物画像を撮影した。人物の立ち位置を固定して、カメラに真っすぐ向かっている人物の向きを 0° と定義する。そこから時計回りを正、反時計回りを負とする(図 2)。今回は 10° 刻みで人物の向きを変えて計 36 方向を撮影した。カメラの高さは商業施設に設置された防犯カメラを想定して、上方から見下ろして撮影するように約 3m に設定して固定した。姿勢は施設内で実際に観測される直立、歩行(右手左足前)、歩行(左手右足前)の 3 種類に定めた。9 人についてこれら 36 方向 3 姿勢を撮影した。

3.2 向き変化による類似度の増減

撮影した画像を用いて、同一人物において向き変化によって類似度がどう変化するかを調査した。ある人物画像に対し、同じ姿勢で向きの異なる同一人物の画像それぞれの類似度を計算する。

実験の結果、類似度は同じ向きの場合が最も類似度が高く、向きが異なれば異なるほど類似度が低くなった。図 3 は直立姿勢で向きが 0° の 3 人の画像それぞれについて、向き変化と類似度の関係を示したものである(ただし、人物毎に類似度の最大値が 1 となるよう正規化している)。どの人物でも向き変化と類似度は同じ関係を示し、向き変化

†京都大学大学院情報学研究科

‡京都大学学術情報メディアセンター

が30°を超えたあたりから類似度が急減している。また、人物や向きや姿勢によって類似度の最大値や平均値は異なったものの、向き変化と類似度の関係は同じ結果となった。

3.3 対応付けにおける向き変化の許容範囲

次に、人物の向き変化の大きさがどの程度ならば SIFT 記述子によって同一人物の対応付けができるのか調査した。今回の実験ではある人物画像1枚に対して、対応付け候補者として向きの異なる同一人物の画像1枚と、別人8名を36方向から撮影した計289枚の画像を用意した。姿勢は直立で統一した。対応付け候補者の中で SIFT 記述子の類似度が最大の人物が同一人物であれば対応付けが正しくできたと判定して、この実験を9名36方向の計324件に対して実施した。

実験結果を図4に示す。これは SIFT 記述子による人物対応付けの正解率を向き変化の大きさごとに示している。ただし、今回の実験では向き変化の大きさに注目したため、正方向と負方向を区別していない。グラフの正解率は9名36方向の平均値である。人物の向き変化が20°の場合の正解率は96%となり、30°の場合に正解率は82%となった。向き変化が40°を超えると正解率は60%を下回った。以上から、映像間で人物の向き変化が30°以内の場合には、SIFT 記述子は人物対応付けに有効であるといえる。これは防犯カメラの配置関係によっては十分に想定される。

4. SIFT 記述子による人物対応付け実験

商業施設に設置された防犯カメラ映像に対して、SIFT 記述子を用いた対応付けを試みる。

4.1 実験環境

用いる映像はどちらも商業施設の通路を撮影したもので、人物が通過するとおおよそ同じ向きで観測される。また、撮影領域に重なりを持たず、カメラ間に非撮影領域が存在する。さらに、通路には店舗や分岐点が存在するため、全ての人物が両カメラで観測されるとは限らない。このような関係にあるカメラペアの映像を用いた(図5)。

このような空間的に近くに配置されたカメラでは、時間帯が異なるほど同一人物が両方の映像に映っている可能性は低くなる。そのため、今回は同じ時間帯の映像についてのみ対応付けを行った。長さ30~40分程度の映像ペア5セット計180分間に対して、映像中より検出した各人物領域から SIFT 記述子を抽出・比較して人物対応付けを試みた。なお、人物領域の検出には HOG によるテンプレートマッチングと背景差分を用いた。

4.2 結果

実験結果を表1に示す。180分間に72人が観測され、両方の映像に映っていた52人中45人が SIFT 記述子により対応付けられた。正解率は87%となった。映像間の人物の向き変化は目測で20°~30°程度であるため、このような結果になったと考えられる。

SIFT 記述子は人物領域の色を用いていない。無地の服といった勾配特徴の少ない人物に対しては色など他の特徴も用いるべきである。そこで、色特徴として色ヒストグラムを組み合わせて用いる手法も試みた。ここでは各人物領域から抽出した色ヒストグラムの Bhattacharyya 距離の逆数を類似度として、これと SIFT 記述子の類似度の合計値が最大の人物を対応付けた。その結果、対応付け正解率が92%

に向上した。なお、類似度として色ヒストグラムのみを用いた対応付け正解率は44%だった。

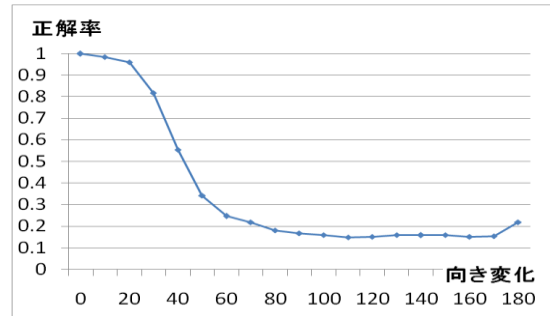


図4:向き変化に応じた人物対応付け正解率

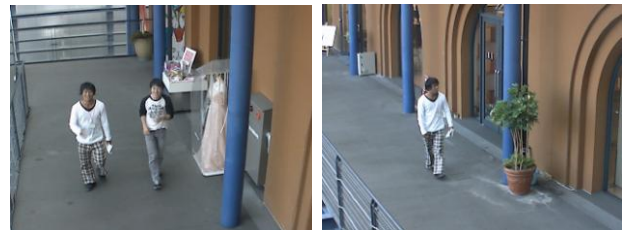


図5:カメラ映像

	対応付けを試みた人数	正しく対応付けられた人数	正解率
SIFT 記述子	52人	45人	87%
色特徴	52人	23人	44%
SIFT 記述子 + 色特徴	52人	48人	92%

表1:異なるカメラ映像間の人物対応付け結果

5. まとめ

SIFT 記述子の向き変化への耐性を検証した結果、映像間で人物の向き変化が30°程度ならば人物対応付けできることがわかった。実際の防犯カメラ映像に SIFT 記述子による人物対応付けを試みると正解率87%となった。さらに色ヒストグラム特徴を組み合わせることで正解率が92%に向上した。複数の特徴を併用する際の、最適な組み合わせ方の検討が今後の課題となる。

謝辞

本研究は文科省、安全・安心な社会のための犯罪・テロ対策技術等を実用化するプログラム「環境適応型で実用的な人物照合システム」の一環として実施したものである。

参考文献

- [1] 本田拓也, 出口大輔, 高橋友和, 井手一郎, 村瀬洋, "入室解析のための色特徴を用いた人物対応付けに関する検討", 電子情報通信学会総合大会講演論文集.D-12-116, Vol. 2009, No.2, pp.225, (2009-03).
- [2] 藤吉弘亘, "Gradient ベースの特徴抽出 -SIFT と HOG-", 情報処理学会研究報告 CVIM 160, pp.211-224, (2007-9).