

# 防犯用カメラを用いた不審者検出法の基礎的検討

## Fundamental Study of Detecting Strangers from Surveillance Camera

溝口 慶範†      大谷 淳†  
Yoshinori Mizoguchi      Jun Ohya

### 1. まえがき

画像処理による人物認証では代表的なものとして歩容特徴を用いたものと顔特徴を用いたものがあるが、歩容特徴で個人の特徴をはっきりと証明するにはまだ課題が残っている。顔特徴を用いた認証は個人を特定する上で最も確証性の高い手法であるといえる。

通常の防犯カメラは広域での監視を目的としており、入ってきた人物の全身を検出することが可能だが、その状態では画像の解像度が低く人物の顔から正確に個人を特定できない。

本稿ではパン・チルト・ズーム (以下 PTZ) カメラを使用して全体画像からズームングで顔をとらえるために必要となる、全身画像からの顔の向き・位置推定法の検討を行った。

### 2. 防犯カメラによる顔認識における問題点

本研究はカメラでとらえた人物が承認されている人物か侵入者かを判断することが目的であるが、防犯カメラの特性上、鮮明な顔画像は取得することは難しく、低解像度な顔画像であるほど認識率が低下する。OpenCV で提供されている Haar-like 特徴による cascade 分類器を使うと人間の全身を検出できる最小の解像度が  $14 \times 28$  (画素) である一方、正面顔は  $30 \times 30$  が最小であった。これは一般の web カメラを用いる場合、カメラから水平 3m 以内の距離にいないと顔検出がされない、程度の精度である。(図 1-1 から 1-7 を参照。) 他手法でも同様で、顔の識別には全身像よりも高解像度の画像が必要と考えられる。



図 1-1~7: 左から、Web カメラから水平距離 0.2m, 0.5m, 1.0m, 1.5m, 2.0m, 2.5m, 3.0m 離れた位置での顔認識処理結果。3.0m 以上離れると認識が出来ない。

この結果から、PTZ カメラを使用し、ズームアップされた顔に対する照合を行うようなシステムについて検討する。

視点固定型 PTZ カメラによる物体追跡の有効性は[1]のように様々な方向から検証されてきた。ズームングを用いた画像処理にはズームの対象をどこに置かず広域画像から決定しなくてはならない。顔認識が不可能な距離で人をとらえた場合、全身画像の追跡中に検出対象のどの位置に顔画像があるかを推定する必要がある。

人物の追跡技術では一般的であり最も有効なのがパーティクル・フィルタである。関連研究[2]の中で、人物の全身画像を多数プロットした多様体マップ上の座標点をパーティクル・フィルタの追跡要素として利用している例がある。前フレームの確率分布を更新する際にその精度を高める役割を果たしている。

本稿ではそのようにパーティクル・フィルタでの併用が期待される多様体マップを作成する技術の 1 つである多次元尺度構成法 (以下 MDS) を使用して、全身画像のグループ化を行う。パーティクル・フィルタで使う全身像の MDS グラフから顔の向きや位置を割り出せるか検証する。顔位置を取得することで PTZ カメラによるズームング処理を円滑に行うことが出来ると期待される。

### 3. 多次元尺度構成法 (古典的 MDS)

MDS は多変量解析手法の 1 つである。比較対象である多次元のデータ配列を有した座標がある場合、座標同士の関係性を低次元で示すために、本手法では次元ごとの座標間ユークリッド距離を算出して最終的に座標間の相対的な距離差で表現する。例えば座標  $i(x_{i1}, x_{i2}, \dots, x_{in})$  と  $j(x_{j1}, x_{j2}, \dots, x_{jn})$  のユークリッド距離は (1) 式のように

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (1)$$

で表せる。

このような距離値を全ての座標間に算出し、最終的には 2 次元、もしくは 3 次元空間上でプロットすることができる。この統計グラフでは近い特徴を持った座標はより近く、反対に親和性の薄い座標同士は離れてプロットされることになる。MDS では相対的な距離差によって比較的高速に、また視覚的に関係性を把握することが出来る。

### 4. 実験

今回の実験では統計的背景差分法を用いて、得られた前景画像の面積が基準を満たしたものを移動物体として認識する。前景画像を含むエリアを抽出したバウンディングボックス (リサイズし、今回は  $30 \times 60$  サイズとした) が、これから学習や検証に用いる対象画像である。

各バウンディングボックスは HoG 処理を行う。今回は 1 セルを 5×5 セル, 1 ブロックは 3×3 セル, 勾配方向は 9 方向とした。

画像 1 枚が持つ次元数の総計は (2) 式から

$$\text{dim} = \left(\frac{\text{Image}_w}{\text{cell}_{\text{num}}} - 2\right) \times \left(\frac{\text{Image}_h}{\text{cell}_{\text{num}}} - 2\right) \times \text{block}_{\text{num}}^2 \times \text{Orientation} \quad (2)$$

全 3240 次元となるのが分かる。

この特徴量を, MDS を用いて低次元空間に圧縮する。MDS 解析には R 言語を活用する。R 言語とは統計解析に特化したフリーの開発環境である。

#### 4.1 4 方向画像の比較

カメラに対して正面を向いて歩いている人物像を 0 度, 右方向に歩いている人物を 90 度としたとき, 頭頂部から見て反時計回りに 0 度 (Front=F), 90 度 (Right=R), 180 度 (Back=B), 270 度 (Left=L) の 4 種の画像群をそれぞれ 10 枚ずつ用意した。MDS で得られた尺度グラフは図 3 のようにはっきりとグループ間でまとまった結果が得られた。



左から, 図 2-1:正面 (F), 2-2:右向き (R), 2-3:左向き (L), 2-4:背面 (B)

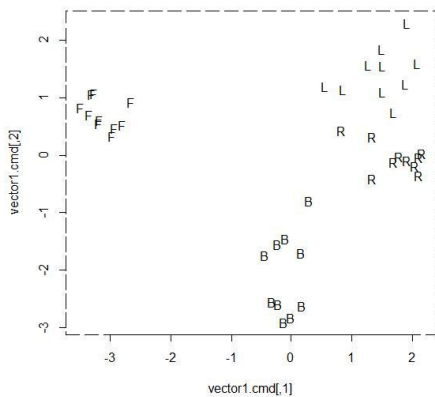


図 3:実験結果 1

#### 4.2 8 方向画像グループと対象画像の比較

4.1 に加えて, 45 度程度 (fr), 135 度程度 (br), 225 度程度 (bl), 315 度程度 (fl) の 8 方向を考える。その学習データに対して X=L, X=fr の 2 つの検証を行う。

X との距離値は比較対象を Y とすると,

$$\text{distance} = (Y_x - X_x)^2 + (Y_y - X_y)^2 \quad (3)$$

の式で算出したものである。

・ X=L の場合 :

表 1:X=L のときの検証結果

	x-dimension	y-dimension		Xとの距離
F	1.37292456	3.72274581	→	39.28652
R	1.16447566	-2.5100096	→	33.07398
L	-2.6305901	-0.95232	→	3.040593
B	1.9404755	-0.998269	→	35.33211
fr	-0.750934	1.72501947	→	11.784
fl	0.47692372	0.78210108	→	19.10133
br	2.25300934	-0.3045802	→	37.76658
bl	0.03832261	-1.7443556	→	19.32953
X=L	-3.8646073	0.27966812		



図 4:

X 画像

・ X=fr の場合 :

表 2:X=fr のときの検証結果

	x dimension	y dimension		xとの距離
F	2.5015052	-2.143094	→	8.932224
R	-2.4307618	0.4650046	→	28.61019
L	-0.7378845	2.918418	→	17.69261
B	-1.4749159	-1.535796	→	24.73665
fr	2.28584246	1.0759469	→	0.451643
fl	0.07987557	0.1202035	→	8.476346
br	-1.2668345	-2.7991	→	30.49974
bl	-1.8632516	1.0803851	→	22.81864
X	2.90642502	0.8180317		



図 5:

X 画像

結果はどちらも, 該当する方向を持ったグループに対する距離値が最も小さく算出され, 同グループに属していると判断できるプロットされていることが分かる。

### 5. 結論

#### 5.1 実験結果, 本提案に関して

実験結果から身体の向きによるグループ化がなされていることが分かる。但し学習データが多くなるとグループ同士の境界線が曖昧になってしまう点や, 単純でない動きが含まれる実環境を考慮すると今回の手法は厳密な判定には遠いと思われる。特に顔の向き・位置推定を行うにあたって所望の情報を全身画像から得られなかった部分がある。同一の身体の向きをしたグループ内であっても顔の位置が違う場合に対処できないケースや, 参照データと実対象の人物が異なる場合にも大きな誤差が生じた。身体の向きから顔の向きを推定するには一層の工夫が必要であることが分かった。

#### 5.2 今後の展望について

今後は人物に対してやはりパーティクル・フィルタで追跡処理を行いながら, 扱うカメラの機械的仕様を考慮して, 知覚 (画像処理) ・行動 (パン・チルト・ズーム) の同期制御を行いたいと思う。人物による複雑な動きやズームの焦点が対象物体の一部分である点などの解決案を考える。

肩から頭部を中心に追跡を行う [3] のような先行研究などを手掛かりに高解像度の顔画像検出を行いたい。

### 6. 参考文献

- [1] 松山隆司, 和田俊和, 物部祐亮, “視点固定型パン・チルト・ズームカメラを用いた実時間対象検出・追跡”, 情報処理学会論文誌, vol40 No.8, Aug1999
- [2] Nir Galili, Fabian Nater, “HoG-Model based Human Tracking and Irregular Behavior Detection”, ETH, 4Jan2010
- [3] 橋本潔, 加賀屋智之, 片岡裕雄, 藤田光子, 田藤雅基, 野上建士, 青木義満, “姿勢変動に伴う身体的特徴変化の統計的モデリングによる遮断に頑健な人物追跡”, ViEW2011, pp.60-67, Dec2011