

飛行ロボットによる人物検索に関する研究

著者	米澤 亨
学位授与機関	Tohoku University
学位授与番号	11301甲第17087号
URL	http://hdl.handle.net/10097/64047

博士学位論文

飛行ロボットによる人物捜索に関する研究

平成27年度

(平成28年2月16日提出)

東北大学大学院情報科学研究科

応用情報科学専攻

米澤 亨

Person Search by Flying Robots

Toru YONEZAWA

Abstract

This thesis describes person search by flying robots. In this thesis, person search means to find target persons from a crowded area; for example, shopping malls or amusement parks.

In recent years, security cameras have been used in various applications. It has become possible to determine the behaviors and trajectories of criminals, and to find missing persons by analyzing images captured by cameras. As a result, the number of installed cameras is increasing annually. Since the amount of cases solved by footage from security cameras is increasing, additional camera installations are expected.

However, conventional fixed security cameras often fail to capture images of people from required angles or at suitable resolutions. Furthermore, a security camera fixed in one position cannot observe people passing through blind spots. This problem can be alleviated to some degree by increasing the number of cameras, but there are limits in terms of cost and location.

These obstacles can be resolved by performing surveillance using a mobile camera. Flying robots are a suitable method for capturing images. Images of a target person can be obtained at the required angle and resolution by capturing images while moving. Since movement provides ability to cover a wide range, the amount of stationary cameras can be reduced, and the problem of blind spots is solved. If the number of security cameras is decreased, the associated system can be simplified.

There are many possibilities with a camera-equipped flying robot. In this study, the focus is on person searches that can contribute to everyday life, like searching for missing children in amusement parks or other crowded areas, or searching for criminals in cities. When searching for lost children or criminals, manpower is commonly distributed in the locations most likely to succeed. The purpose of this study is creating a system with flying robots that can search for persons autonomously. The system is required to ensure a safe and secure society.

In large amusement parks, the comings and goings of visitors is managed by the entry and exit control gate. In this study, person searches are performed in such facilities. The flow of the person search system is described as follows. First, images of visitors are obtained at the entry and exit control gate. Person searches are performed by comparing these images with those captured by a flying robot. When person searches begin, the system creates the search path of the flying robot in the targeted area. The flying robot captures the images while moving along this path. If the targeted person is found from the images captured, the search is complete.

In order to realize such a system, it is necessary to integrate various technologies, including communications and computer systems. However, the most important point is “How do flying robots perform person searches?” Therefore, this thesis will describe three research themes as follows. 1. Person detection by flying robots. 2. Person identification by flying robots. 3. Path-creation to search for persons using flying robots.

The images captured by flying robots are different from images captured by a fixed camera. Considering safety,

footage cannot be obtained from too close range to the person, so the images are low-resolution. Further, the lower body in images is frequently hidden due to the overlap of multiple persons. Because the flying robot is capturing images while moving, person images contain various angles, and background images change constantly. This thesis' proposed algorithm is to support these flying robot unique image features.

The person detection algorithm is proposed in this study. Considering hidden lower bodies, and increased false positive background results from the perpetually changing backgrounds in person detection, the algorithm was created to detect the upper body of a person with high accuracy. The algorithm's characteristic is a combination of multiple kinds of features with little correlation to each other. By compensation for the weaknesses of each others' features, accuracy of person detection was improved. As a result of evaluation with the public database, it was confirmed that the accuracy was higher than conventional methods. In addition, detection derived from actual footage from flying robots confirmed that the algorithm could be applied to real-life situations.

The person identification algorithm uses appearance-based person identification, because upper body images are taken by various angles from afar. By overlaying identifications from low-resolution images from multiple perspectives, the cumulative appearance-based person identification algorithm can show increased identification accuracy, it can be proposed. The algorithm considers light fluctuation, and does not consider color information. Identification occurs using apparel patterns, and the person's profile between the head and the shoulder line. Experiments with 38 subjects showed that the average identification rate was 85% (false positive rate = 5%) for 15×15 pixel images of the upper body. Additionally, even when all subjects were wearing the same clothes (uniform), the algorithm was confirmed to identify targets by comparing such slight differences as how name tags were placed and wrinkle patterns in the clothes. Further, despite some problems, the experimental results of a person detection (6 subjects) using a flying robot confirmed that it was possible to identify people in the images captured by a flying robot while moving.

By combining person detection and person identification data, a sequential flow system was created. Comparing person identification accuracy to manually doing person detection, it was found that they were equivalent. Additionally, regarding processing time, it was confirmed that the identification results could be output in real time. Therefore, a real environment operable identification system using a flying robot could be created.

Furthermore, to ensure effective and accurate person identification by the flying robot, a path-creation method is proposed by this study. The number of paths is infinite. From these, when using the person identification algorithm proposed by this thesis, the path-creation method identifies subjects more efficiently and accurately. Path-creation was simulated, and effectiveness was confirmed in multiple patterns. Path changes due to targeted person differences (template image) and situations with prior knowledge were confirmed.

In this thesis, to conduct person detection with a flying robot, each of the three proposed methods of person detection, person identification, and flying robot path-creation, as well as simulation and experimental results are described.

目次

Abstract.....	i
第1章 緒論.....	1
1.1 研究の背景.....	1
1.2 本研究の目的.....	3
1.3 本論文の構成.....	6
第2章 飛行ロボットによる人物認証.....	7
2.1 はじめに.....	7
2.2 問題設定.....	8
2.2.1 検出する人物の角度.....	8
2.2.2 検出する人物の解像度.....	8
2.3 関連研究.....	9
2.4 特徴量.....	10
2.4.1 ソベル+ハール特徴.....	10
2.4.2 ガボールフィルタによるエッジ特徴.....	11
2.4.3 ソベルフィルタによるエッジ特徴 (HOG).....	13
2.5 人物検出アルゴリズム.....	15
2.5.1 アルゴリズム全体構成.....	15
2.5.2 各特徴量の識別器の学習と性質.....	16
2.5.2.1 学習データ.....	16
2.5.2.2 学習結果.....	17
2.5.3 拡大縮小処理.....	18
2.5.4 大分類処理 (1 特徴での絞込み).....	18
2.5.5 詳細判定処理 (3 特徴の評価値算出).....	19
2.5.6 最終判定処理 (3 特徴の評価値統合).....	19
2.6 上半身検出アルゴリズムの精度確認実験.....	22

2.6.1	データセット：USC の評価.....	23
2.6.2	データセット：CAVIAR の評価.....	24
2.6.3	データセット：NLPR-HS の評価.....	25
2.6.4	処理時間.....	26
2.6.5	評価結果まとめと，本人物検出アルゴリズムについての考察.....	26
2.7	飛行ロボットが取得した画像での人物検出実験.....	30
2.7.1	人物検出アルゴリズムのための画像変換.....	31
2.7.2	人物検出実験.....	32
2.7.2.1	人物が静止状態での飛行ロボットの撮影による検出実験.....	32
2.7.2.2	人物が移動状態での飛行ロボットの撮影による検出実験.....	36
2.7.3	飛行ロボットが撮影した画像での人物検出実験まとめと考察.....	39
2.8	まとめ.....	39
第3章	飛行ロボットによる人物認証.....	41
3.1	はじめに.....	41
3.2	関連研究.....	42
3.3	飛行ロボットによる人物検索手法.....	44
3.3.1	飛行ロボットによる人物検索システム.....	44
3.3.2	累積アピアランス照合アルゴリズム.....	45
3.3.2.1	1枚画像のアピアランス照合アルゴリズム.....	45
3.3.2.2	累積アピアランス照合アルゴリズムの概要.....	48
3.3.2.3	累積アピアランス照合アルゴリズムの詳細説明.....	50
3.3.2.4	さまざまな入力画像パターンに対する照合値変化.....	52
3.4	累積アピアランス照合アルゴリズムの基本性能評価.....	56
3.4.1	累積アピアランス照合アルゴリズムの評価方法概要.....	56
3.4.2	服装の違いによる認証精度の評価.....	57
3.4.3	テンプレート画像と入力画像の服装が異なる場合の認証精度の評価.....	58
3.4.4	解像度の違いによる評価.....	59
3.4.5	人物の角度バリエーションの変化による評価結果.....	61
3.4.6	誤検出画像の入力の影響.....	63
3.4.7	基本性能評価のまとめと考察.....	64
3.5	飛行ロボットが撮影した画像からの人物検索実験.....	67
3.5.1	飛行ロボットが撮影した画像からの人物検索実験概要.....	67

3.5.2 静止した複数人物を飛行ロボットで撮影した場合の認証.....	68
3.5.2.1 人物の見えの角度の違いによる認証結果の比較.....	69
3.5.2.2 飛行ロボットの高さ（仰角）の違いによる認証結果の比較.....	71
3.5.2.3 検索対象者（テンプレート）の違いによる認証結果の比較.....	72
3.5.2.4 認証精度まとめと，人物捜索用閾値設定.....	72
3.5.3 複数の歩行中の人間を飛行ロボットで撮影した場合の認証実験.....	74
3.5.3.1 通路を歩行中の人物に対して，人物の正面方向を撮影した映像を認証した結果.....	74
3.5.3.2 通路を歩行中の人物に対して，同じ方向に移動しながら，人物の正面方向を撮影した映像を認証した結果.....	76
3.5.3.3 ランダムな方向に移動している人物に対して，移動しながら万遍なく撮影した映像を認証した結果.....	78
3.5.4 飛行ロボットによる人物捜索実験のまとめと考察.....	80
3.6 まとめ.....	82
第4章 人物検出と人物認証の統合による人物捜索.....	83
4.1 はじめに.....	83
4.2 複数フレーム間の人物画像の統合.....	84
4.2.1 複数フレーム間の人物画像の統合手法.....	84
4.2.2 人物シーケンスデータの作成.....	84
4.3 作成したシーケンスデータによる人物認証実験.....	88
4.3.1 静止した複数人物を飛行ロボットで撮影した場合の認証.....	88
4.3.2 複数の歩行中の人間を飛行ロボットで撮影した場合の認証実験.....	89
4.4 リアルタイム処理実験.....	92
4.5 まとめと考察.....	94
第5章 人物認証のための飛行ロボットの軌道生成.....	95
5.1 はじめに.....	95
5.2 関連研究.....	96
5.3 飛行ロボットによる自動人物捜索システム.....	96
5.4 軌道生成手法-概要.....	98
5.4.1 軌道生成手法の考え方.....	98
5.4.2 軌道生成手法概要.....	100

5.5 軌道生成手法-詳細.....	103
5.5.1 画像パターン群 A の作成	104
5.5.2 推定照合値の算出.....	104
5.5.3 軌道候補の作成	105
5.5.4 画像パターン取得確率の算出.....	106
5.5.5 認証対象者の存在確率分布の作成.....	107
5.5.6 照合可能な人物画像取得範囲.....	109
5.5.6.1 カメラ撮影範囲.....	109
5.5.6.2 テンプレート画像と同じ角度の人物画像を取得できる範囲	113
5.5.6.3 照合可能な人物画像が撮影できる範囲.....	114
5.5.6.4 各範囲の統合.....	116
5.6 軌道生成シミュレーション.....	117
5.6.1 照合用テンプレート画像.....	117
5.6.2 軌道生成パラメータ	117
5.6.3 画像取得範囲と人物存在確率作成シミュレーション.....	117
5.6.4 認証対象者 1 人の軌道生成シミュレーション結果.....	119
5.6.5 複数人の認証対象者の軌道生成シミュレーション	121
5.6.6 事前知識がある場合の軌道生成シミュレーション	122
5.6.7 より複雑な軌道作成が可能な場合の演算例.....	124
5.6.8 軌道生成シミュレーションまとめと考察	124
5.7 まとめ.....	124
第 6 章 結論.....	126
6.1 結論	126
6.2 展望.....	127
参考文献.....	128
業績.....	134
謝辞.....	135

目次

1.1	多数の監視カメラ設置例（新幹線乗り場）	1
1.2	飛行ロボットを用いた人物検索システム	4
2.1	上半身の人物検出アルゴリズム適用結果	7
2.2	検出対象角度の人物画像	8
2.3	検出対象解像度の人物画像（15×15画素）	8
2.4	ソベルフィルタ適用例	11
2.5	ソベル画像に Haar-like フィルタを適用する例	11
2.6	左図：ガボールフィルタ（フィルタ値を濃淡で表現） 右図：離散化した7画素×7画素のフィルタ（実演算時）	12
2.7	ガボール特徴量	13
2.8	ソベルフィルタからエッジ方向ヒストグラムの作成	14
2.9	正規化ヒストグラムの作成	14
2.10	人物検出アルゴリズムフロー	16
2.11	学習データサンプル（左：人物上半身，右：風景画像等）	16
2.12	各特徴量識別器の学習データに対する ROC グラフ	17
2.13	拡大縮小イメージ図	18
2.14	詳細判定エリア作成	18
2.15	詳細判定処理フロー	19
2.16	最終判定値処理フロー	20
2.17	データセット：USC の評価結果（ROC カーブ）	23
2.18	USC の検出結果（グリーン：検出成功，赤：誤検出）	23
2.19	データセット：CAVIAR の評価結果（ROC カーブ）	24
2.20	データセット：CAVIAR の検出結果（グリーン：検出成功，赤：誤検出）	24
2.21	データセット：NLPR-HS 評価データ	25
2.22	データセット：NLPR-HS の評価結果（DET カーブ）	26

2.23 特徴量の評価値間で差が大きかった例	27
2.24 特徴量の評価値間で差が大きかった例	28
2.25 飛行ロボットが撮影したデータサンプル.....	30
2.26 人物検出のための画像変換.....	31
2.27 飛行ロボットが撮影した入力画像サンプル	32
2.28 俯角 15° の評価結果 (ROC カーブ)	33
2.29 俯角 30° の評価結果 (ROC カーブ)	33
2.30 俯角 50° の評価結果 (ROC カーブ)	33
2.31 集計結果 (FRR-FAR グラフ)	34
2.32 俯角 15° の検出結果画像.....	35
2.33 俯角 30° の検出結果画像.....	35
2.34 俯角 50° の検出結果画像.....	36
2.35 各移動パターンでの評価結果 (ROC カーブ)	37
2.36 パターン 1 の検出結果画像.....	37
2.37 パターン 2 の検出結果画像.....	37
2.38 パターン 3 の検出結果画像.....	37
2.39 人物が移動中の 3 パターンの評価結果 (FRR-FAR グラフ)	38
2.40 パターン 1 の検出失敗画像例.....	38
3.1 飛行ロボットによる人物検索のイメージ図.....	41
3.2 飛行ロボットによる人物認証システムのフローチャート.....	44
3.3 人物認証システムのための入力画像.....	46
3.4 エッジ特徴量.....	47
3.5 特徴量ブロック	47
3.6 入力画像の特徴量ブロックのマージン m	48
3.7 テンプレート画像と入力画像	49
3.8 テンプレート画像.....	52
3.9 (a):本人の入力画像が 1 枚入力されたときの結果 ($V= 0.46$)	52
3.10 (b):変化の少ない本人の入力画像が 2 枚入力されたときの結果 ($V= 0.63$)	53
3.11 (c):変化の少ない本人の入力画像が 3 枚入力されたときの結果 ($V= 0.71$)	53
3.12 (d):異なる角度の本人の入力画像が 2 枚入力されたときの結果 ($V= 0.79$)	53
3.13 (e):異なる角度の本人の入力画像が 3 枚入力されたときの結果 ($V= 1.08$)	54
3.14 (f):すべての本人の入力画像(29 枚)が入力されたときの結果 ($V= 2.18$)	54
3.15 (g):他人の入力画像(48 枚)が入力されたときの結果 ($V= 0.37$)	54

3.16 (h):すべてのテンプレート画像(28 枚)が入力されたときの結果 ($V=21.71$)	55
3.17 (i):誤検出画像が 3 枚含んだときの結果 ($V=0.71$)	55
3.18 3 人の実験対象者のサンプルデータ	56
3.19 38 人の実験対象者の後ろ姿	57
3.20 テンプレート画像と入力画像の服装が同じ場合の照合結果	57
3.21 テンプレート画像と入力画像の服装が異なる場合の照合結果	58
3.22 異なるサイズの入力画像例	59
3.23 夏私服-人物認証結果 (ROC カーブ)	60
3.24 冬私服-人物認証結果 (ROC カーブ)	60
3.25 制服-人物認証結果 (ROC カーブ)	61
3.26 4 パターンの入力画像図	62
3.27 4 パターンの照合結果 (ROC カーブ)	62
3.28 10 枚のランダムな誤検出画像を含んだ入力画像例	63
3.29 10 枚のランダムな誤検出画像を含んだ入力画像の照合結果 (ROC カーブ)	63
3.30 制服テンプレート画像例	65
3.31 3 人の入力画像例と照合結果 (ヒストグラム)	65
3.32 1 枚画像照合の分布分析	66
3.33 飛行ロボットによる撮影実験の様子と撮影に用いた飛行ロボット	67
3.34 ID3 のテンプレート画像と ID1~ID6 のテンプレート画像	68
3.35 飛行ロボットの動きと人物位置模式図	68
3.36 画像サンプルと人物抽出結果	69
3.37 時系列での画像入力と照合値の変化	70
3.38 時系列での照合値の変化 ((a)0°, (b)90°, (c)180°, (d)270°)	70
3.39 各向きでの ID3 の入力画像	71
3.40 時系列での照合値の変化 ((a)15°, (b)50°)	72
3.41 各向きでの ID3 の入力画像 (3 画像ごとに抜粋) ((a)15°, (b)50°)	72
3.42 各テンプレートの時系列での照合値の変化	73
3.43 人物が静止状態での全パターンの FAR, FRR グラフ	74
3.44 飛行パターンの模式図と取得画像抜粋	75
3.45 各テンプレートの時系列での照合値の変化	75
3.46 各実験対象者の入力画像	76
3.47 飛行パターンの模式図と取得画像抜粋	77
3.48 各テンプレートの時系列での照合値の変化	77

3.49	各実験対象者の入力画像	78
3.50	飛行パターンの模式図と取得画像抜粋	79
3.51	各テンプレートの時系列での照合値の変化	79
3.52	各実験対象者の入力画像	80
4.1	3 フレームでのシーケンスデータ作成	83
4.2	シーケンスデータ例	85
4.3	1 人の人物に対して複数個の検出の例	86
4.4	ドア部の誤検出多数の例	86
4.5	静止状態の全シーケンスの FAR,FRR グラフ	88
4.6	歩行パターン 1, ID1 の各シーケンスによる照合値の変化とそのときの入力画像	90
4.7	歩行パターン 2, ID1 の各シーケンスによる照合値の変化とそのときの入力画像	91
4.8	リアルタイム処理アプリケーション画面	92
5.1	人物認証のための飛行ロボットの複数の軌道	95
5.2	飛行ロボットによる自動人物認証	97
5.3	飛行ロボットによる人物検索フロー	97
5.4	テンプレート画像と照合値の上がり易さを示したヒストグラム(実験対象者 A) ...	99
5.5	テンプレート画像と照合値の上がり易さを示したヒストグラム(実験対象者 B) ..	100
5.6	軌道生成フロー	103
5.7	軌道候補	105
5.8	人物画像取得確率表	107
5.9	人物存在確率のイメージ図	109
5.10	カメラ撮影範囲のイメージ図	110
5.11	テンプレート画像の人物画像と同じ角度の人物画像を取得できる範囲	113
5.12	照合可能な人物サイズが撮影できる範囲	114
5.13	画像取得範囲シミュレーション	118
5.14	合成範囲	118
5.15	人物存在確率分布	118
5.16	軌道候補と軌道評価値	119
5.17	飛行シミュレーション図 (軌道番号 2878)	119
5.18	軌道候補と軌道評価値	120
5.19	飛行シミュレーション図 (軌道番号 198)	120
5.20	軌道候補と軌道評価値	121
5.21	飛行シミュレーション図 (軌道番号 1166)	121

5.22 事前知識ありの軌道候補と軌道評価値	122
5.23 飛行シミュレーション図（軌道番号 1215）	122
5.24 事前知識なしの場合の飛行シミュレーション結果図（軌道番号 189）	123
5.25 手動で設定した軌道	123
5.26 飛行シミュレーション図（手動で設定した軌道）	123

表 目次

2.1 人物検出パラメータ	22
2.2 各特徴量の評価値間の相関係数.....	27
3.1 累積アピランス照合アルゴリズムフロー	51
4.1 人物検出パラメータと，人物追跡パラメータ	85
4.2 各シーンのシーケンスデータ数.....	85
4.3 手動入力シーケンスに対する画像数の割合	87
4.4 手動シーケンスによる照合結果.....	89
4.5 人物検出アルゴリズムのシーケンスによる照合結果	89
4.6 使用 PC，アルゴリズム設定パラメータ等	93
3.1 累積アピランス照合アルゴリズムフロー	51

第1章 緒論

本研究の研究背景と研究目的，本論文の構成について述べる。

1.1 研究の背景

近年，社会の安全・安心を脅かす危険や脅威が顕在化してきている．例えば，大災害や事故の多発化，米国同時多発テロをはじめとした国際的なテロの頻発，治安の悪化などである．科学技術がこれらの危険や脅威の防止，排除に貢献し，安全・安心な社会の構築へ寄与することは重要なことである．人々の科学技術への期待も高まっている．また，日本においても文部科学省の科学技術・学術審議会で，安全・安心な社会の構築に資する科学技術政策に関する懇談会[1]も開かれており，国を挙げて安心・安全の社会構築について議論されている．

日常生活における安全・安心な社会にフォーカスすると，近年では防犯カメラの設置による犯罪の未然防止，犯罪捜査への利用など，防犯カメラによる見守り社会が挙げられる．世界的に最も防犯カメラの整備が進んでいると言われているイギリスでは，420万台以上のカメラが存在すると言われており，観光客は1日に約300回も防犯カメラに写されると言われている[2]．イギリスで防犯カメラが設置され始めたのは1990年代の初頭で，防犯カメラによる凶悪犯罪の事件解決に重要な役割を果たしてから，イギリス国民に広く認知されるようになった．日本でも昨今の凶悪犯罪の増加に伴い設置台数は年々増加している（図1.1）．設置場所も，街頭，商店街，ショッピングモール，学校，地下鉄，その他公共施設等，街中のあらゆる場所に設置されている．犯罪捜査にカメラ映像が用いられ，防犯カメラに映った映像を元に事件が解決した例も増えてきている．またそのカメラ映像も一般公開され，警察が情報提供を呼びかけるといったように，利用頻度や利用方法も増えている．今後もますます設置台数は増加すると考えられる．



図1.1 多数の監視カメラ設置例（新幹線乗り場）[3]

一方、防犯カメラにはプライバシーの問題がついて回る。例えば、ゴミ置き場に設置された防犯カメラの映像から、ゴミ出しのちょっとしたルール違反者を周囲の住民に公開され、その人物は引越しまで余儀なくされた。こういった例が増え、ドイツなどでは監視カメラ規制が強められ、設置できる場所や設置できるカメラの種類なども類型化されるようになってきている。日本では、法規制的には、個人情報保護法で、一定人数以上の個人情報を扱う民間人が監視カメラで集めた個人情報を目的外利用することが禁じられているぐらいであり、その適用もケースバイケースである。民間人設置の場合は、同意原則が当てはまるので、「監視カメラ作動中」という警告なしに撮影することはできない。しかし、警察など公共機関については、その縛りはない[4]。

このように現在の防犯カメラは、犯罪防止、犯罪捜査に役に立つというメリットと、プライバシーを侵害する可能性があるデメリットが混在している状態であるが、その設置台数は増加している。設置台数の増加は、防犯、監視できるエリアを増やし、安全・安心なエリアを増やすためである。

防犯カメラの設置台数が増えるにつれて、撮影された映像のデータ量が膨大になってきている。ホテル、企業の建物内、ショッピングモールなど、リアルタイムに映像を監視し、異常状態を見つけ対応する必要がある施設では、映像を監視するために人手が必要である。また犯罪捜査においては、事件が起こった場所の周辺の防犯カメラ映像を集めて、手がかりとなる部分を膨大な量の映像内から探し出している。これも捜査員の手が必要である。そこで、それらの人的に行っている行為を自動化する研究が従来から行われてきた。防犯カメラから異常状態の映像だけを抽出する異常判定技術[5] [6]、動いている部分のみを抽出する動体検知技術[7]、捜したい人物を検索する人物認証技術[8]などである。

一方、防犯カメラの設置方法としては、コスト面からも、基本的には1台で広い範囲を撮影できるように設置される。そのため、犯罪捜査等、写った映像を解析し利用する目的に用いられる場合には、撮影したい角度の人物映像や、必要な解像度の映像を撮影出来ていないことが多い。前段で述べた自動化技術も、適用できる対象の角度や映像の解像度に制限があり、適用できない場合が多い。この問題を解決するための1つの案は、カメラの設置台数を増やすことが考えられる。しかし、設置場所、コスト面、インフラの整備などで限界があると考えられる。また、固定されたカメラである以上、どうしても死角になる部分が残る。

本研究では、この問題に対して、移動できるカメラを用いることが解決策になると考えた。視点を移動させながら撮影することで、必要な角度と解像度の映像を取得することが出来る。また、移動することで、広範囲をカバーすることが出来るため、防犯カメラの設置台数を削減することができ、死角の問題も解決できる。カメラ台数が削減できれば、それに伴うシステムも簡素化できる。しかし、その移動できるカメラの操作を人手で行っていないと解決にならない。よって、異常検知や人物捜索といった目的に対して、必要な映像を能動的に移動しながら撮影し認識する自律移動ロボットを作成することが解決案である。また、自動的に認識することが出来れば、最終的なアウトプットに用いる映像以外は人が直接見る必要がない

ので、プライバシーの問題も解決できる可能性がある。システム内部に一時的に保存する映像についても、認証に用いる不可逆な特徴量に変換しておくことで視認できないようにすることが出来る。

能動的に移動しながら撮影できるカメラとしては、地上走行ロボットと飛行ロボットが考えられるが、飛行ロボットが適していると考えた。その理由は、第一に足場の影響を受けない点である。人物はあらゆる障害物を回避し、路面状況をあまり気にせず移動できるが、地上走行ロボットは段差や障害物があると移動できない場合がある。飛行ロボットは、障害物、路面状況は関係なく移動可能である。第二に、ルートに制限がない点である。地上を移動すると、障害物以外の人物等も避ける必要があるが、上空ではその必要がない。第三に、上空から広視野で撮影できる点である。ターゲットを見つけやすく、追跡しやすい。さらに上空から素早く回り込んで撮影することで、地上走行ロボットでは移動できずに死角になる場所も撮影可能である。

カメラを搭載した飛行ロボットによって、従来の防犯カメラの映像を用いて行われている特定エリアの監視や人物捜索が高度化するだけでなく、人に関すること以外にも用途は広がる。例えば、高所など人が行きにくい場所の検査や、テロ対策のための不審物検知なども考えられる。

自律移動できることのメリットは大きく、さまざまな用途が考えられるが、本研究では、生活の中で身近に貢献できる人物捜索に注目した。遊園地や大型ショッピングセンターなどにおける迷子捜しや、街中での犯人捜索などである。現在迷子や犯人を捜索しようとした場合、捜索対象が存在しそうな場所に人数を振り分けて人的に捜索するのが一般的であろう。この人的に行っている捜索行為を、飛行ロボットが自律的に行えるようにすることが本研究の目的である。

飛行ロボットにはデメリットも考えられる。例えば、ペイロードの問題が考えられる。現在のドローンやヘリ型飛行体では、搭載できる重さに制限があり、多数のセンサを搭載することは難しい。また、燃料の問題もあり、連続稼働時間も地上走行ロボットと比較すると短い。さらに安全性の問題として、何らかのアクシデントで飛行不可能状態に陥った場合、最悪の場合、墜落事故が発生する可能性がある。しかしながら、安全性を考慮した飛行ロボットの研究[9]は進んでおり、小型化の研究[10]も行われている。近い将来、安全で長時間飛行可能な飛行ロボットが十分実現可能であると考えられる。

1.2 本研究の目的

本研究では、日常生活での利用、例えば、遊園地やショッピングモールでの迷子探し、犯人捜しを飛行ロボットが行えるようにすることを目的とする。こうすることで安全・安心な社会の構築に飛行ロボットを役立せることが出来る。

我々が迷子や犯人などの人物検索を行うときは、その検索対象の人物の視覚的な特徴を手がかりとして、その特徴を持った人物を目視で探索する方法が一般的であろう。人間は、その検索対象の人物に詳しくれば詳しいほど、細かな特徴、例えば、ほくろの位置や歩き方などの情報を使って検索するが、ケースバイケースであり、それを飛行ロボットに行わせることは難しい。よって本研究では、検索対象の画像を手がかりに検索することを考えた。図1.2に本研究で想定する飛行ロボットを用いた人物検索システムを示す。

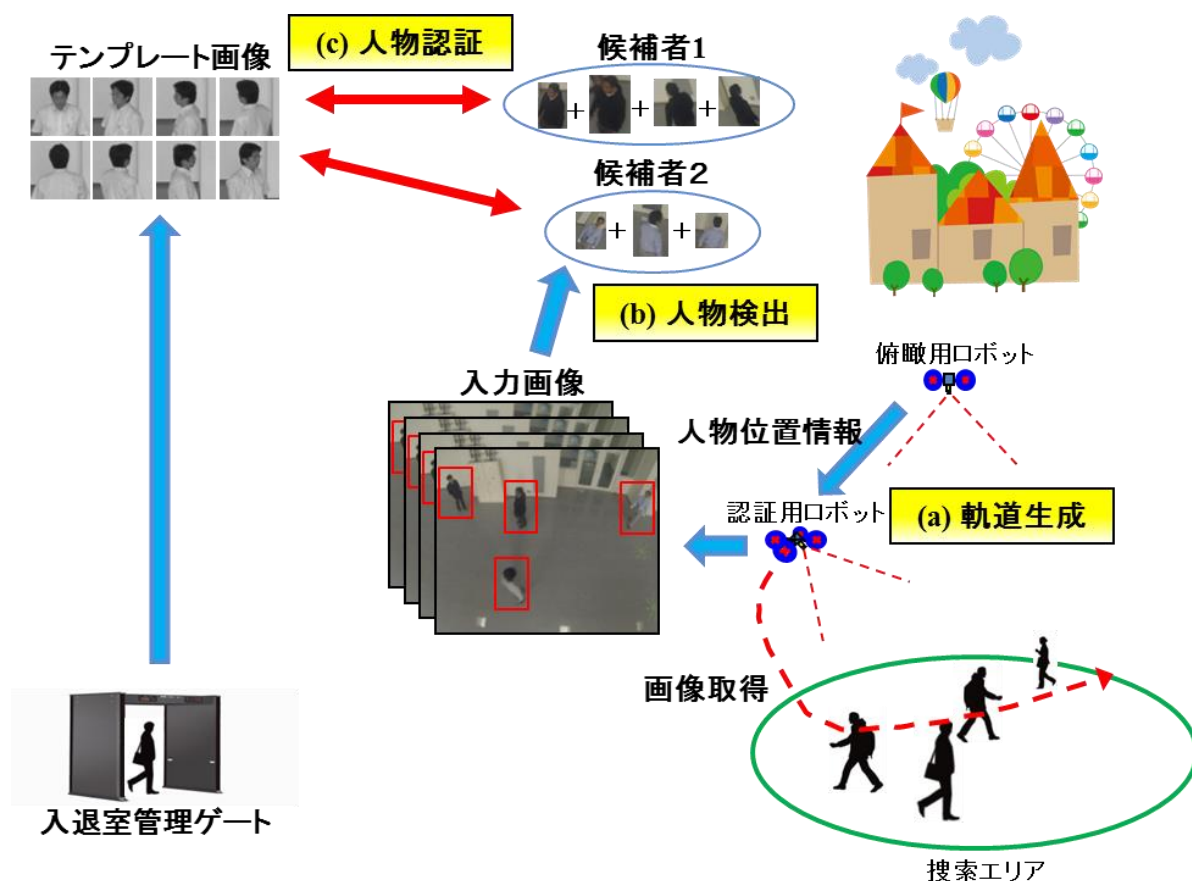


図1.2 飛行ロボットを用いた人物検索システム

近年、セキュリティに対する意識の向上から、大型の遊園地では入出管理ゲートにより人の出入りを管理している[11]。今後も遊園地やショッピングモール等でこういった入出管理ゲートは増えると考えられる。本研究ではこのような施設において、人物検索を行うことを考えた。図1.2に示した人物検索システムの流れについて述べる。最初に、遊園地やショッピングモールの入出管理ゲートにおいて、探索に用いる入場者全員の人物の画像を取得しておく。これらの画像が探索時に用いるテンプレート画像となる。人物検索開始時に、取得しておいたテンプレート画像群の中から、検索したい人物のテンプレートを指定する。次に飛行

ロボットが搜索エリアに飛び立ち、搜索を開始する。このとき、2種類の飛行ロボットを用いて人物搜索を行うことを考えている。俯瞰用ロボットと認証用ロボットである。俯瞰用ロボットの役割は、搜索エリアに存在する人物全員の位置や動きを追跡することである。認証用ロボットの役割は、認証するための人物画像を取得することである。俯瞰用ロボットは人物の位置、動きの情報を認証用ロボットに渡す。認証用ロボットは受け取った情報を元に認証に適した画像を取得するための軌道生成を行う(図1.2(a))。認証用ロボットはその軌道に沿って飛行しながら映像を撮影する。その映像が人物搜索を行うための入力映像となる。人物認証を行うために、人物検出アルゴリズムを用いて入力映像から人物部の画像のみを抽出する(図1.2(b))。そして、抽出された候補者ごとの人物画像とテンプレート画像を照合する(図1.2(c))。設定した閾値以上の照合値が出力されれば、その候補者が搜索している人物であるとして、搜索依頼元に画像や位置を送信し、搜索依頼元に確認を行ってもらう。その結果、その人物が搜索対象者であれば搜索完了であるし、間違っていれば、再搜索等の次のアクションを行う。また、全ての候補者の照合値が閾値以下であれば、再度そのエリアを搜索するか、別のエリアに移動するか判断し、搜索を続ける。以上が飛行ロボットを用いた人物搜索システムの流れである。尚、ロボットは撮影やデータ転送のみを行い、軌道生成や人物検出、人物認証等の演算処理を別途用意したPCで行うことは勿論可能である。

このようなシステムを実現するためには、通信、システム面も含めて、さまざまな技術を融合する必要がある。しかし、人物搜索が目的であるので、最も重要なポイントは、“飛行ロボットがどのように人物搜索を行うか”，であると考えられる。よって、本研究では次に示す3つの研究課題に取り組む。1.「飛行ロボットによる人物検出」(図1.2(b))、2.「飛行ロボットによる人物認証」(図1.2(c))、3.「人物認証のための飛行ロボットの軌道生成」(図1.2(a))である。

1. 「飛行ロボットによる人物検出」

飛行ロボットが撮影した映像内から人物部分を抽出するアルゴリズムである。その映像は上空から撮影する映像なので、人物の足部はオクルージョンとなる可能性が高い。そのため、本研究では、画像内から人物の上半身を検出する人物検出アルゴリズムを提案する。また、飛行ロボットは移動しながらあらゆる場所を写すため、誤検出が増えると考えられる。固定カメラのように背景がある程度分かれば、パラメータ等の調整で誤検出を減らせる可能性もあるが、背景は完全にランダムであると考えられる必要がある。認証アルゴリズム側でも誤検出画像に対するロバスト性は考慮するが、人物検出アルゴリズムにおいても出来るだけ誤検出は減らしたい。よって、上半身を高精度に検出するアルゴリズムを提案する。

2. 「飛行ロボットによる人物認証」

本研究での認証は、安全に、且つ、その認証対象者の行動を出来る限り阻害することなく行うことを目指している。犯人搜索を想定すると気付かれないほうがよい。そのため、飛行ロボットは認証対象者に接近せず、意識させないように撮影したい。遠方から人物を撮影す

るため、認証に用いることが出来る人物画像は低解像度が予想される。且つ、あらゆる視点から撮影した上半身の人物画像である。誤検出画像に対するロバスト性も必要である。本研究では、このような飛行ロボット特有の画像に対応する人物認証アルゴリズムを提案する。

3. 「人物認証のための飛行ロボットの軌道生成」

本研究では、飛行ロボットが撮影した映像に人物認証アルゴリズムを適用することで、人物検索を行う。撮影される映像は、その飛行ロボットの軌道によってさまざまに変化する。飛行ロボットはホロノミックな移動体であるので、全方向に移動することができ、その軌道は無限大に存在することになる。よって、提案する人物認証アルゴリズムを用いたときに、効率的かつ、高精度に人物認証を行うことができる軌道生成手法について提案する。

また、その人物検索の過程では、映像取得から、人物検出、人物認証と演算負荷が高いと考えられる処理を一連の流れで行う必要がある。本研究では、その実現性の検証や、人物検出アルゴリズムと人物認証アルゴリズムを組み合わせたときの人物認証精度の検証についても行う（「人物検出と人物認証の統合による人物検索」）。

1.3 本論文の構成

本論文の構成は以下の通りである。

1章では、本研究の背景，研究目的について述べた。

2章では、飛行ロボットによる人物検出について述べる。

3章では、飛行ロボットによる人物認証について述べる。

4章では、人物検出と人物認証の統合による人物検索について述べる。

5章では、人物認証のための飛行ロボットの軌道生成について述べる。

6章では、結論と今後の展望について述べる。

第2章 飛行ロボットによる人物検出

2.1 はじめに

人物検索を行うために飛び立った飛行ロボットが撮影する画像は、上空から地上を撮影した映像となる。人物はさまざまな角度で撮影されるが、その中から検出したい人物画像は、人物認証を行うためのものである。しかし、人物を真上から撮影した画像では人物認証は難しいので、やや斜め上から撮影した画像内から人物部を検出したい。このような画像は、複数人物の重なりや、その撮影方向によって、人物の足元はオクルージョンになりやすい。また、想定している人物認証アルゴリズム（第3章）はあらゆる方向を向いた人物画像を用いるため、後ろ姿のような顔が写っていない人物画像も抽出したい。よって、検出対象は、やや斜め上から撮影したあらゆる方向の人物の上半身である。

また安全面から人物に接近することはできないので、撮影する画像は低解像度である。

飛行ロボットは移動しながら撮影するため、背景は常に変化する。そのため、従来からある背景差分を基本とした動体検知アルゴリズム[1]は適用できない。画像1枚から人物を検出する必要がある。

加えて、背景が常に変化するということは、背景が完全にランダムであると考えられることから、人物検出において、背景を誤検出する確率が増えると考えられる。従来の防犯カメラのような固定されたカメラでは、背景がある程度想定できるため、パラメータ等の調整で誤検出を減らせる可能性もあるが、飛行ロボットが撮影する画像では行うことは出来ない。よって従来法よりも高精度に検出したい。

まとめると、低解像度のあらゆる角度の上半身画像を高精度に検出する必要がある。本章では、それらの条件を満たす人物検出手法を提案する。尚、第3章で述べる人物認証アルゴリズムも上半身のみで行う。図2.1に上半身の人物検出のイメージ図を示す。



図2.1: 上半身の人物検出アルゴリズム適用結果（データは[26]のデータベースから抜粋）

2.2 問題設定

2.2.1 検出する人物の角度範囲

飛行ロボットは人物に対して、真横方向から真上方向（人物の仰角 $0^{\circ} \sim 90^{\circ}$ ）までさまざまな角度で撮影することが出来る。しかし、目的は人物認証であり、真上方向では目視でも人物を見分けることが難しい。そのため、本研究の人物認証アルゴリズムは、人物を認証できる範囲として、仰角： $0^{\circ} \sim 45^{\circ}$ ，方位角 $0^{\circ} \sim 360^{\circ}$ を想定した。人物検出アルゴリズムも同範囲の人物を検出できるようにする。図2.2に検出対象の人物画像を示す。

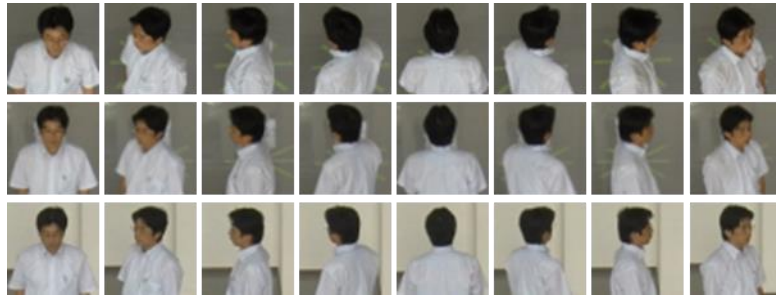


図2.2: 検出対象角度の人物画像（自ら撮影）

2.2.2 検出する人物の解像度

飛行ロボットは遠方から人物を撮影するため、本研究に用いる人物認証アルゴリズムは、人物の頭から腰までの長さが約15画素の画像でも照合に用いることを想定している。よって、人物検出アルゴリズムも、頭から腰までの長さが約15画素の画像でも検出できるようにする。図2.3に検出対象解像度の人物画像を示す。

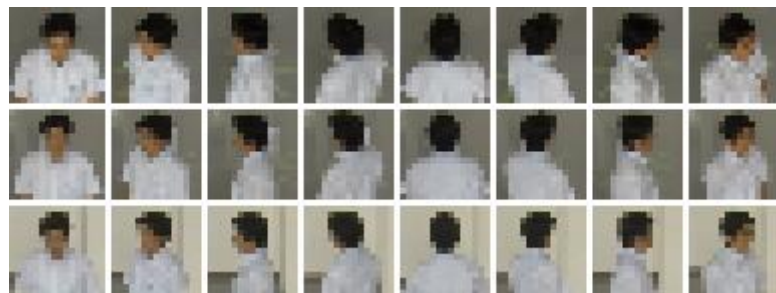


図2.3: 検出対象解像度の人物画像（ 15×15 画素）

2.3 関連研究

動画像から人物を検出するアルゴリズムとして古くから研究されているのは、背景差分法を代表とする動体検知アルゴリズム[1]である。しかしながら飛行ロボットは移動しながら撮影を行うため、背景は常に変化する。よって動体検知アルゴリズムは適用できない。

そのため画像1枚から人物を検出するアルゴリズムが必要である。カラー情報（肌色）を用いた検出手法[2]もあるが、カラーは日照条件により変化し、さらに服装や撮影角度によっては肌色が見えない場合もある。そのため、カラー情報による検出手法も適用は難しい。

画像1枚から顔を検出するアルゴリズムとして、ViolaとJonesによる手法[3]が有名であり、特徴量はHaar-like特徴で、識別器はAdaboost[4]である。Haar-like特徴量とは、矩形領域の平均明度の差分値として求められるスカラ量である。Adaboostとは統計的学習手法Boostingの中の1つであり、識別に用いるには弱い特徴（弱識別器）を多数組み合わせることで高精度な識別を行うものである。目と頬の明度差や、口と頬の明度差(Haar-like特徴)など、それだけで顔と判断するには弱いと考えられる特徴を、数千単位の数で組み合わせ顔を検出する。彼らの手法はさまざまなところで応用されている。本研究でも顔が見える場合は適用可能であるが、背面方向や低解像度で顔が写らない場合は検出できないため、適用範囲が限られる。

よって、人物の姿を特徴として検出する人物検出アルゴリズムが挙げられる。従来研究されている人物の姿による人物検出アルゴリズムは、上で述べたViolaとJonesによる手法を応用したものが多い。人物検出に有効な特徴量としては、局所領域における勾配方向をヒストグラム化した Histograms of Oriented Gradients (HOG)[5]、設定した領域の累積エッジ強度の比である Edge Orientation Histograms (EOH)[6]、エッジ同士を連結させ作成した線である Edgelet[7]などが挙げられる。このように、人物の輪郭情報を取得するために、エッジ情報に着目した特徴量が多く提案されている。また、それらの特徴量間の共起に基づく手法である Joint-Haar-like特徴[8]や、Shaplet特徴[9]、HOG特徴の共起表現による手法[10]が提案されている。さらに、リーマン多様体上のクラスタリングとLogit-Boostを組み合わせた手法[11]も提案されている。これらは高精度な人物検出を実現している。また、背景画像や複数枚の連続画像を用いて高精度な人物検出を実現している手法[12][13][14]もある。

本研究では上半身の画像を用いて人物検出を行いたい。上半身に特化した人物検出手法としては、HOGを用いた手法[15][16]、Local Binary Pattern (LBP) を用いた手法[17]、Region Covariance Matrix (RCM)を用いた手法[18]などがある。筆者らは、複数の周波数の Ω 型エッジ特徴を捉え、独自の学習手法であるLDAArray法[19]（特許文献）による識別を行う手法によって、高精度化を目指した。

また、カメラ画像以外のセンサを用いた人物検出に関連する研究としては、距離画像に基づく人物検出手法[20]、レーザレンジスキャナによる人物検出と追跡[21]、レーザスキャナと画像センサの組み合わせによる人物検出と追跡[22][23][24]等が挙げられる。しかし、これらは撮影距離や人物への危険性の問題から適用することは難しい。

2.4 特徴量

本研究では従来研究と同様に、低解像度の上半身の人物画像の特徴として Ω 型のラインを捉えることで人物検出を行うことを考えた。飛行ロボットが撮影する人物画像は、照明環境や、カメラのピントの合い方、人物との距離、人物の動き、カメラの動き、カメラの種類、特性などにより、撮影される人物画像の Ω 型のラインはブレやボケの影響を大きく受けると想定される。特に飛行ロボットは、人物までの距離が大きく変化することが想定されるので、ピントはずれる傾向にあると考えられる。そのため、1つの周波数のエッジ特徴だけでは対応しきれないと考え、異なる複数の周波数、及び異なる種類のエッジ特徴を組み合わせ、補うことで精度向上を図った。本研究では3つの特徴量を用いた。ガボールエッジ特徴[25](7×7)、ソベルエッジ特徴 (HOG)、ソベルエッジ特徴とハール特徴を組み合わせたソベル+ハール特徴である。ガボール特徴 (7×7) で低周波のエッジ特徴を、ソベルエッジ特徴 (HOG) で高周波のエッジ特徴を、ソベル+ハール特徴で、エッジの集まりや繋がりを取得できると考えた。入力画像はグレイ画像変換され、それぞれの特徴量に変換される。

2.4.1 ソベル+ハール特徴

ソベル+ハール特徴は、ソベルフィルタによりエッジ強度画像を作成し、その画像からHaar-like特徴量を作成する。Haar-like特徴量の候補数は、学習画像のサイズと、それに適用するHaar-like特徴フィルタの種類、サイズ、位置パターンにより決まる。検出時に適用するフィルタは、ViolaとJonesによる顔検出手法[3]と同様に、学習時に学習アルゴリズムによって選択する。

まず、ソベルフィルタにより、4方向のエッジ強度画像を作成する。ソベルフィルタの4方向 ($0^\circ, 45^\circ, 90^\circ, 135^\circ$) を F_{s_i} エッジ方向の添え字を i 、入力画像を I_m 、その画素番号を k とすると、各画素におけるエッジ強度 $s_{i,k}$ は以下の式で表される。

$$s_{i,k} = |F_{s_i} * I_{m,k}| \quad (2.1)$$

$$F_{s_0} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \quad F_{s_1} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}, \quad F_{s_2} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \quad F_{s_3} = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} \quad (2.2)$$

図2.4に入力画像から4種類のソベルエッジ画像を作成した例を示す。

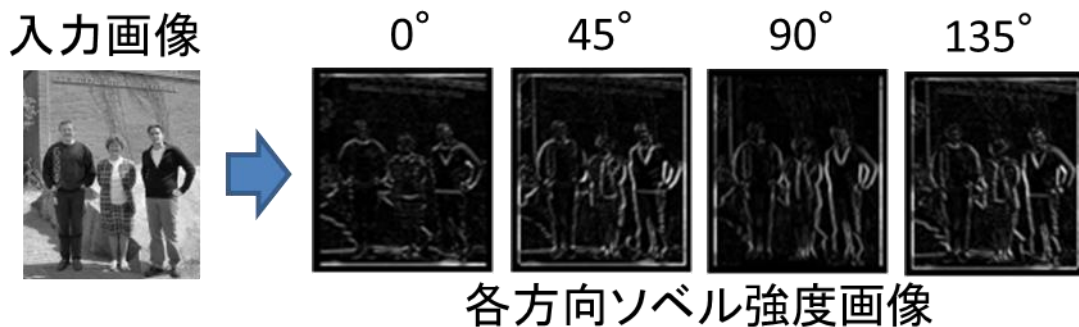


図2.4: ソベルフィルタ適用例 (左図は[7]のデータベースから抜粋)

作成されたソベル強度画像から, Haar-like特徴を作成する. そのHaar-like特徴作成の例を図2.5に示す.



図2.5: ソベル画像にHaar-likeフィルタを適用する例

Haar-like特徴量は, 図2.5に示した白の領域 $r2$ と黒の領域 $r1$ の2つの領域の輝度差を算出したものである. Haar-like特徴量 H は, 以下の式で表される.

$$H(r1, r2) = S(r1) - S(r2) \quad (2.3)$$

$S(r)$ は, 領域 r の輝度平均を算出する関数である. 2つの領域の組み合わせパターンは多数ある. 縦方向, 横方向, 斜め方向の明暗差を取ることができる. 本研究において, 識別に用いるこの2つの領域のパターンとその数は, 学習アルゴリズムLDAArray法[19]を用いて学習することで求めた.

2.4.2 ガボールフィルタによるエッジ特徴

ガボールフィルタは, ヒトをはじめとする哺乳類の脳の視覚野において, モノや空間を認識するために, その処理が行われていることが知られている[25]. ガボールフィルタは, 任意の向き, 任意の周波数のエッジ特徴を取得することが出来る. 以下に2次元ガボールフィル

タの式を示す.

$$F_g(x, y, \lambda, \varphi) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cdot \exp(2\pi\lambda i(x \cos \varphi + y \sin \varphi)) \quad (2.4)$$

λ は周波数で, φ は回転角を表す. この関数は, 実部と虚部の成分を持つので, 式(2.5)に実部を式(2.6)に虚部の式を示す.

$$F_g^{real}(x, y, \lambda, \varphi) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cdot \cos(2\pi\lambda(x \cos \varphi + y \sin \varphi)) \quad (2.5)$$

$$F_g^{imag}(x, y, \lambda, \varphi) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cdot \sin(2\pi\lambda(x \cos \varphi + y \sin \varphi)) \quad (2.6)$$

本研究では4方向 (0° , 45° , 90° , 135°) の向きのエッジ特徴を求めた. 図2.6に各方向, 実部, 虚部のガボールフィルタの値を濃淡で表した画像を示す. また, 図2.6で示したように, 実演算では, 7画素 \times 7画素の大きさでフィルタを作成する.

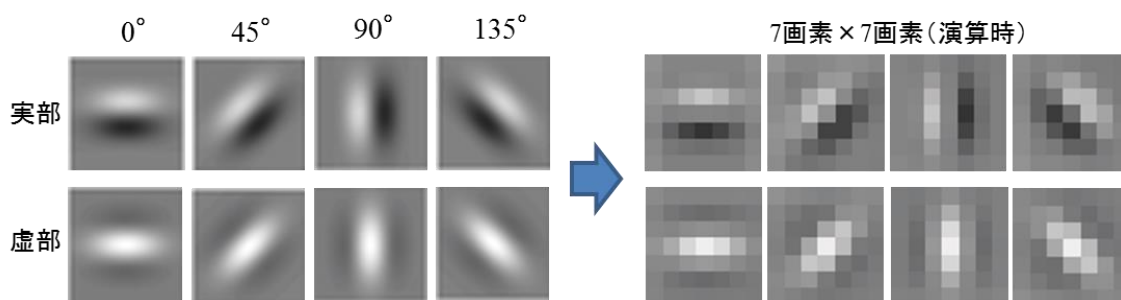


図2.6: 左図: ガボールフィルタ (フィルタ値を濃淡で表現)
右図: 離散化した7画素 \times 7画素のフィルタ (実演算時)

周波数 λ は, 7画素 \times 7画素の大きさのフィルタに1周期分収まるように設定した. ソベル特徴と同様に, 入力画像 I_{in} の画素番号を k , エッジ方向の添え字を i , 作成したフィルタを実部 $F_{g,i,k}^{real}$, 虚部 $F_{g,i,k}^{imag}$ と定義すると, 各画素におけるエッジ強度 $g_{i,k}$ は以下の式で表される.

$$g_{i,k} = \sqrt{(g_{i,k}^{real})^2 + (g_{i,k}^{imag})^2} \quad (2.7)$$

$$g_{i,k}^{real} = F_{g,i,k}^{real} * I_{in,k} \quad (2.8)$$

$$g_{i,k}^{imag} = F_{g,i,k}^{imag} * I_{in,k} \quad (2.9)$$

人物検出アルゴリズムで用いる特徴量は、4方向のエッジ強度の合計値と、そのエッジ強度の合計値で正規化した4方向のエッジ特徴を組み合わせた、1画素あたり5次元の特徴量である。各特徴量 $G_{0,k} \sim G_{4,k}$ は、以下の式で表される。

$$G_{0,k} = g_{0,k} + g_{1,k} + g_{2,k} + g_{3,k} \quad (2.10)$$

$$G_{1,k} = \frac{g_{0,k}}{G_{0,k}} \quad G_{2,k} = \frac{g_{1,k}}{G_{0,k}} \quad G_{3,k} = \frac{g_{2,k}}{G_{0,k}} \quad G_{4,k} = \frac{g_{3,k}}{G_{0,k}} \quad (2.11)$$

図2.7に各特徴量のサンプルイメージを示す。

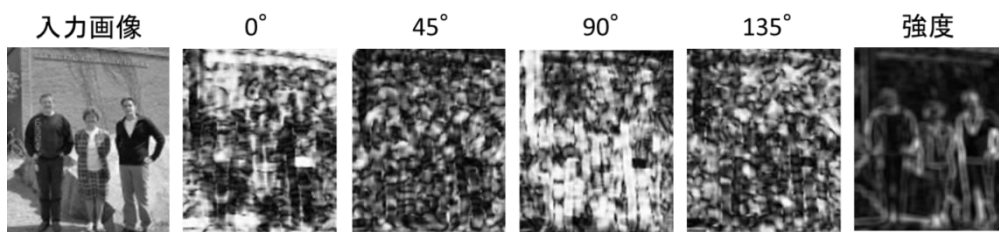


図2.7: ガボール特徴量

2.4.3 ソベルフィルタによるエッジ特徴 (HOG)

本研究で用いたソベルフィルタによるエッジ特徴は、HOG特徴[5]の微分特徴作成部分をソベルフィルタで行ったものである。最初に2方向(0°, 90°)のソベルフィルタによりエッジ強度とエッジ方向(正接)を求める。入力画像 I_m の画素番号を k とすると、エッジ強度 h_m とエッジ方向 h_{θ} は、式(2.2)にて定義したソベルフィルタを用いて、以下の式で表される。

$$h_{m,k} = \sqrt{h_{x,k}^2 + h_{y,k}^2} \quad (2.12)$$

$$h_{\theta,k} = \tan^{-1} \left(\frac{h_{y,k}}{h_{x,k}} \right) \quad (2.13)$$

$$h_{x,k} = F_{s2} * I_{m,k} \quad (2.14)$$

$$h_{y,k} = F_{s0} * I_{m,k} \quad (2.15)$$

$h_{x,k}$, $h_{y,k}$ は入力画像にソベルフィルタを適用した結果の値である。

次に、作成されたエッジ強度画像 h_m を、設定したサイズの矩形エリア(セル)に分割する。エッジ方向ごとのエッジ強度のヒストグラムを作成するため、エッジ方向の分割範囲もあらかじめ設定しておく。そして、その矩形エリア内で各画素 k でのエッジ方向 $h_{\theta,k}$ を参照しながら、分割された各方向のビンにエッジ強度の値を累積する。その結果がヒストグラム v_n である。 n

は全セルのヒストグラムの角度番号である．ソベルフィルタからエッジ方向ヒストグラムの作成過程を図2. 8に示す．

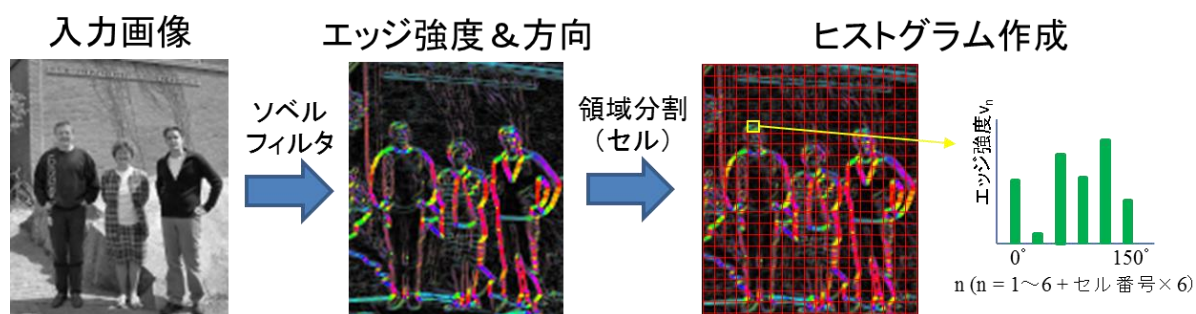


図2. 8: ソベルフィルタからエッジ方向ヒストグラムの作成
(エッジ強度画像の色は、方向を表している)

求めたヒストグラム v_n が特徴量となる．ヒストグラムの各ビンは、複数のセルをまとめて1つのブロックとした1ブロック内のヒストグラムの全体強度で正規化される．その正規化は以下の式で表される．

$$v_n = \frac{v_n}{\sqrt{\left(\sum_{k=1}^{q \times q \times N} v_n^2\right) + 1}} \quad (2.16)$$

分母は1ブロック($q \times q$ セル)に含まれるヒストグラムの総和である． N は分割されたエッジ方向数である．その正規化の模式図を図2. 9に示す．図2. 9は 3×3 セルを1ブロックとしたときの例である．この正規化は照明変動に対応するために行う．

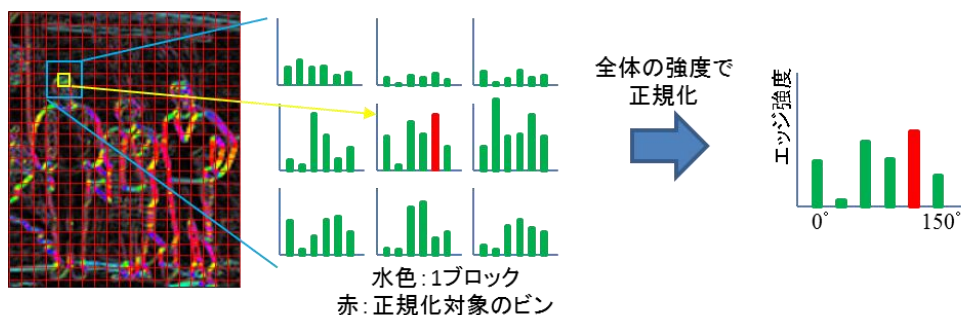


図2. 9: 正規化ヒストグラムの作成

2.5 人物検出アルゴリズム

2.4で述べた特徴量を使って識別器を作成する。人物検出アルゴリズムのアウトプットとして評価されるのは、検出精度と処理時間である。処理時間は短いほどよいが、アルゴリズムによる高速化だけでなく、近年のCPUによる分散処理、GPUを用いたものやFPGAによるハードウェア化による高速化なども考えられる。ハードウェアを用いた高速化は、実運用時に適用場所やシステム構成、コスト面が明らかになった上で検討されることが多い。そのため、本研究では、高速化に関してはアルゴリズムの構成による高速化を図り、ハードウェアによる高速化は今後の検討課題とした。その上で、人物検出の高精度化を目指した。

2.5.1 アルゴリズム全体構成

人物検出アルゴリズムは、人物認証の前処理であるため、出来るだけ高速処理を目指したい。人物検出を行うためには、基本的には入力画像の全画素に対し識別器を適用し判定を行う必要がある。そのため、識別器を高速な処理にしつつ、精度を維持するための手法として、大分類処理と詳細判定処理の2段階の識別を行うことを考えた。図2.10に、人物検出アルゴリズムのフローを示す。処理の流れは次のとおりである。(a): 入力画像が入力されると、カラー情報は用いないため、グレイ画像変換を行う。(b): 次にそのグレイ画像に対し、拡大縮小処理を施し、複数枚の拡大縮小された判定用画像を作成する。検出したい人物のサイズは、テンプレート画像と同じとは限らないため、入力画像を拡大縮小した判定用画像を作成し判定処理を行う。(c): 大分類処理を行い、軽い処理で明らかに人物領域ではない部分をリジェクトし、以後判定しないようにする。大分類処理ではソベル+ハール特徴による判定を行う。識別器はLDAArray法[19]（特許文献）で作成する。LDAArray法はAdaboostを応用したものである。(d): 大分類処理が行われた後に残った画素について、詳細判定処理を行う。詳細判定では、3つの特徴量を求める。ガボール特徴、ソベル特徴(HOG)、そして大分類で求めたソベル+ハール特徴である。各特徴量の識別器を適用し評価値を得る。識別器は大分類処理と同様に、LDAArray法によるものである。(e): 最後に、最終判定では詳細判定で求めた各特徴量の識別器の評価値を統合する処理を行う。

次に大分類処理、詳細判定処理で用いられる、各特徴量の識別器の学習について述べる。

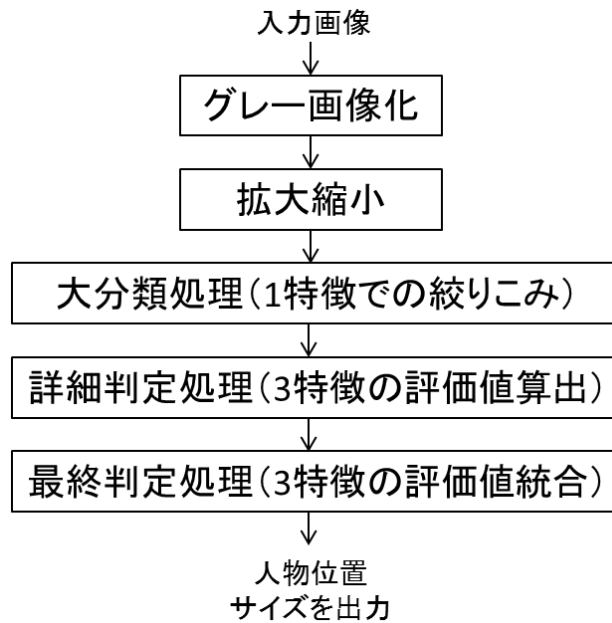


図2. 10: 人物検出アルゴリズムフロー

2.5.2 各特徴量の識別器の学習と性質

2.5.2.1 学習データ

各特徴量の識別器に用いた学習データは共通である。

- ・ ポジティブデータ（人物姿画像）：8627枚（56画素×48画素）
- ・ ネガティブデータ（風景画像）：7395枚

人物画像については、著者らが人物照合用に撮影したデータと、インターネット上からランダムに収集した画像を併せたものである。風景画像については、インターネット上からランダムに収集した画像である。そのサンプルを図2. 11に示す。人物姿画像は56画素×48画素で、人物の上半身が写るように正規化した。

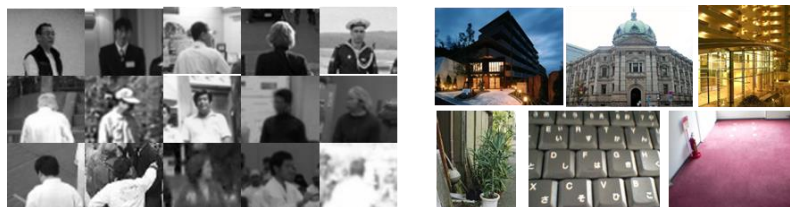


図2. 11: 学習データサンプル（左：人物上半身，右：風景画像等）
（インターネットより取得）

2.5.2.2 学習結果

各特徴量を学習し識別器を作成した。学習に用いた特徴量数はそれぞれ、ガボール特徴：10500、ソベル+ハール特徴：10340、ソベル（HOG）特徴：13500であった。ソベル+ハール特徴では、ハール特徴の、サイズと位置のバリエーションについては、識別精度が落ちない程度に適度に間引いた。ソベル（HOG）特徴については、1セルを5画素×5画素とし、1ブロックを5セル×5セルとした。

学習データに対する評価結果を図2.12に示す。

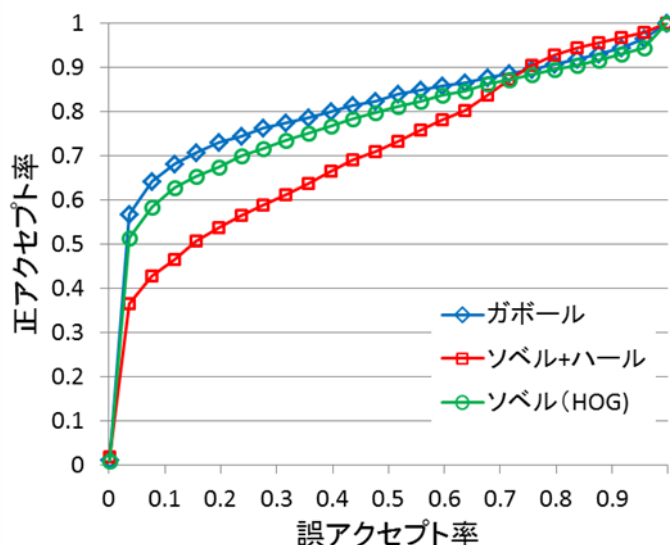


図2.12: 各特徴量識別器の学習データに対するROCグラフ

横軸は誤アクセプト率で、ある閾値を設定したときに、風景画像を誤って人物と判定した割合である。ただし、風景画像の評価については、風景画像全体を走査し、その最大値をその画像の出力値とした。よって誤アクセプト率は、風景画像数に対する各風景画像内で1画素でも人物と判定した画像の割合である。縦軸は正アクセプト率で、ある閾値を設定したときに、人物画像を正しく人物と判定した割合である。図2.12から、誤アクセプト率が低い(0~0.3)ときは、ガボール、ソベル（HOG）、ソベル+ハールの順に正アクセプト率が高く、誤アクセプト率が高い(0.7以上)のときは、ソベル+ハールの正アクセプト率が高いことが分かる。大分類処理では、入力画像内で、人物を人物以外であると判定する誤リジェクトを出来るだけ発生させないようにし、明らかに人物以外であると判定できる部分をリジェクトしたい。どちらか判定しにくい部分については、詳細判定処理で判定するようにする。そのため、大分類処理には正アクセプト率が90%以上のときに、最も誤アクセプト率が低くなるソベル+ハール特徴量が適しているといえる。また、ソベル+ハール特徴は、3つの特徴量の中で最も処理時間が短いため大分類に適している。

2.5.3 拡大縮小処理

テンプレートとして学習するサイズは56画素×48画素である。しかし、入力画像にそのサイズの上半身の人物が写っているとは限らない。よって、検出したい人物のサイズの範囲を入力パラメータとして指定することにより、あらゆるサイズの人物画像を検出できるようにした。図2.13に拡大縮小のイメージ図を示す。検出したい人物サイズの範囲内で、1.2倍ずつ倍率を変えて画像を作成する。

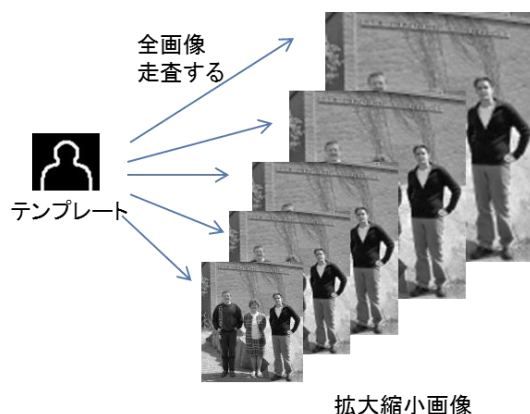


図2.13: 拡大縮小イメージ図

2.5.4 大分類処理 (1特徴での絞込み)

大分類処理では、ソベル+ハール特徴の識別器を用いて、明らかに人物領域ではない部分をリジェクトする。ソベル+ハール特徴の識別器を大分類処理に用いた理由は、2.5.2.2で述べたように、正アクセプト率が90%以上のときに、最も誤アクセプト率が低くなることと、ViolaとJones[3]の顔検出処理のように、積分画像を利用した高速な処理が行えるためである。ソベル+ハール特徴の識別器で拡大縮小画像を走査し、評価値画像を得る。評価値画像をあらかじめ設定した閾値で2値化し、詳細判定を行う画素を決定する。拡大縮小画像から詳細判定を行うエリアを求めた様子を図2.14に示す。

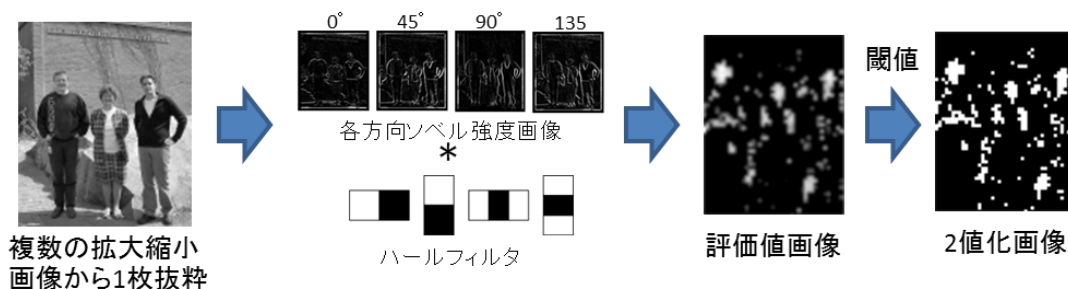


図2.14: 詳細判定エリア作成

2.5.5 詳細判定処理 (3特徴の評価値算出)

詳細判定領域部分のみ3つの特徴量の評価値を求める。ガボール識別器と、ソベル識別器を適用し、評価値を求める。ソベル+ハール識別器の結果は大分類結果と同じである。図2. 15に詳細判定処理の流れを示す。

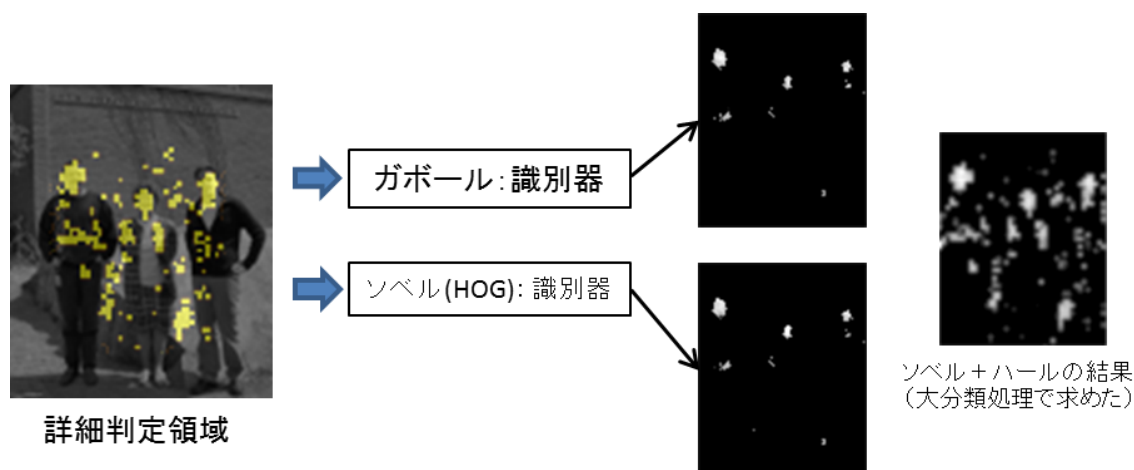


図2. 15: 詳細判定処理フロー

2.5.2で学習した各特徴量の識別関数について、それぞれ、ガボールによる識別関数を L_g 、ソベル (HOG) による識別関数を L_h 、ソベル+ハールによる識別関数を L_s と定義する。入力画像 I_{in} に対し、それら3つの特徴量の識別関数による評価を行った結果の評価値画像をそれぞれ、 V_g 、 V_h 、 V_s とすると、それらは以下の式で表すことができる。

$$V_{gk} = L_g(I_{ink}) \quad (2.17)$$

$$V_{hk} = L_h(I_{ink}) \quad (2.18)$$

$$V_{sk} = L_s(I_{ink}) \quad (2.19)$$

k は入力画像 I_{in} の画素番号である。

2.5.6 最終判定処理 (3特徴の評価値統合)

最終判定処理では、各識別器で出力された評価値 V_g 、 V_h 、 V_s を統合する。識別器が1つの場合には、一般的には、出力された評価値画像内を走査し、設定した閾値以上の値である位置を人物位置として出力する。しかし、本アルゴリズムでは、3つの特徴量の評価値画像が得ら

れる。最も単純な統合処理として3つの評価値画像を加算することが考えられるが、本アルゴリズムでは単に評価値画像を加算するだけでなく、より精度が高くなるような工夫を施した。

画像内の人物が写っている位置を中心として、その周囲数画素の範囲の評価値を分析すると、3つの特徴量の評価値画像でその評価値の表れ方が異なる。1つめの違いは、その数画素の範囲内での最大値を得る位置が異なることである。もう1つは、その最大値を得る画素からの評価値の広がり方が異なることである。特徴量によっては、その最大値だけを評価値として判断したほうが精度が高くなる場合もあれば、その周辺を平均的に評価値として判断したほうが精度が高くなる場合もある。これは、高い評価値が出力される風景画像の場合にも、その評価値の表れ方が人物画像とは異なる場合がある。よって、これらを総合的に判断すれば、より高い精度が得られると考えた。本アルゴリズムでは3つの各特徴量の評価値の最大値とその周辺画素の平均値を特徴量として、線形判別分析で学習する手法を提案する。

具体的な処理の流れは次のとおりである。(a): 大分類処理の結果の評価値画像の中から最大値を得る画素を求める。(b): その最大値を得る画素の周囲9画素×9画素を第一の判定範囲として記憶する。(c): 第一の判定範囲内にある評価値を0で初期化する。(d): (a)-(c)の処理を繰り返し、判定範囲のリストを作成する。得られる最大値が設定した閾値以下になれば繰り返し処理を完了させる。(e): (a)-(d)の処理で作成した判定範囲のリストを参照し、順に最終判定値を求める。最終判定値は、各特徴量の評価値エリア内から最大値と平均値を求め、重み付き積和演算をしたものである。図2.16に、最終判定処理のアルゴリズムのフローを示す。

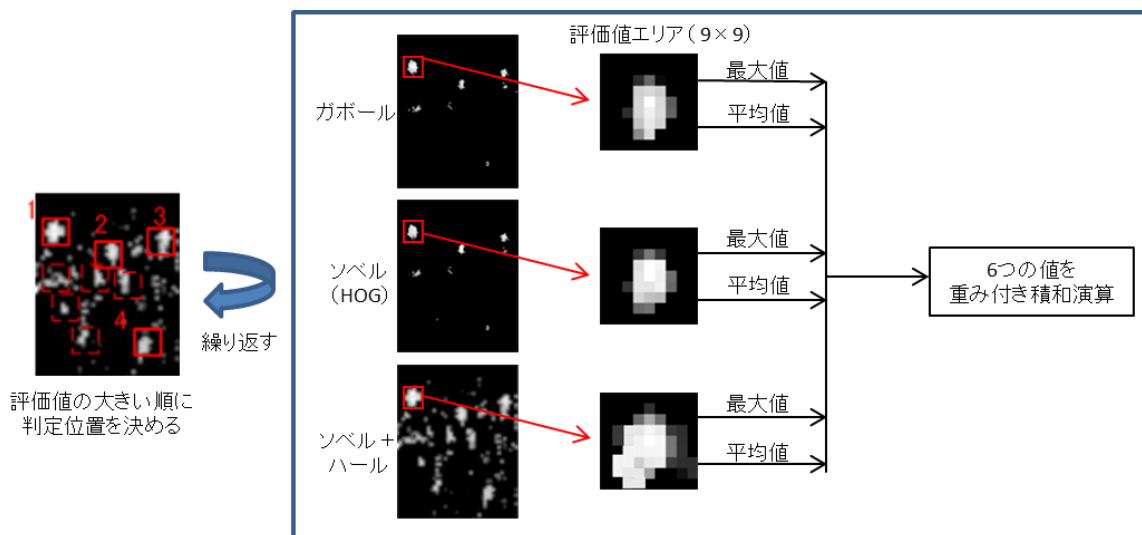


図2.16: 最終判定値処理フロー

画素番号 k において、その周辺画素から最大値を得る関数を F_{max} 、平均値を得る関数を F_{ave} 、ガボール評価値画像から得られる最大値、平均値を v_{g_max} 、 v_{g_ave} 、ソベル（HOG）評価値画像から得られる最大値、平均値を v_{h_max} 、 v_{h_ave} 、ソベル+ハール評価値画像から得られる最大値、平均値を v_{s_max} 、 v_{s_ave} 、とすると、それぞれは以下の式で表すことができる。

$$v_{g_max\ k} = F_{max}(V_{g\ k}) \quad (2.20)$$

$$v_{g_ave\ k} = F_{ave}(V_{g\ k}) \quad (2.21)$$

$$v_{h_max\ k} = F_{max}(V_{h\ k}) \quad (2.22)$$

$$v_{h_ave\ k} = F_{ave}(V_{h\ k}) \quad (2.23)$$

$$v_{s_max\ k} = F_{max}(V_{s\ k}) \quad (2.24)$$

$$v_{s_ave\ k} = F_{ave}(V_{s\ k}) \quad (2.25)$$

最終判定値 V は、これらの値による重み付き積和演算である。式(2.26)に最終判定式を示す。 α は重み係数である。

$$V = \alpha_1 \cdot v_{g_max} + \alpha_2 \cdot v_{g_ave} + \alpha_3 \cdot v_{h_max} + \alpha_4 \cdot v_{h_ave} + \alpha_5 \cdot v_{s_max} + \alpha_6 \cdot v_{s_ave} \quad (2.26)$$

この最終判定値 V に対し、閾値設定を行い、閾値以上となった位置を人物位置として出力する。尚、各拡大縮小画像に対し最終判定値は求められる。そのため、入力画像に写った同じ人物のサイズ違いで重なって出力されることがある。このサイズ違いは、出力されたエリアの重なり度合いを判定し、重なりが大きい場合は、最終判定値が大きいほうが優先される。

以上の処理で、人物検出は終了する。

2.6 上半身検出アルゴリズムの精度確認実験

本章では公開されている異なる3つのデータセットでの評価結果について述べる。これらのデータは従来研究において、頭と肩を用いたディテクターに対して評価が行われているものである。2.6.1ではデータセット：USC [7]の評価について述べる。データセット：USCの評価結果では本アルゴリズムの各特徴量での評価結果についても示す。本アルゴリズムでは、学習データは2.5.2.1で述べたようにインターネット上から集めた画像であるが、本アルゴリズム検討中はUSCのデータを参考にしながら開発した。その他の2つのデータセットについては、アルゴリズムが完全に完成してから評価を行ったものである。尚、2.5.2.1で述べた学習データには本章で述べる評価データは含まれていない。

また、本アルゴリズムを適用するにあたって、パラメータを設定する必要がある。以下にそのパラメータの説明を示し、表2.1に本実験での設定値を示す。

設定パラメータ

- ・サーチ個数：
画像に写っている人物の検出個数。本実験では評価に十分な個数として20を設定した。
- ・サーチサイズ（画素）：
頭の中心から腰中心の位置までのサイズ。
- ・大分類閾値：
本実験では人物検出精度を優先し、人物を誤リジェクトしない値を設定した。
- ・重なり範囲：
本アルゴリズムにより人物検出を行うと、人物位置とサイズが出力される。その位置とサイズから矩形で囲った範囲（図2.18）が人物の上半身（頭，肩）の範囲である。その矩形が重なって出力された場合、1人の人物による結果が複数出力されたものか、2人の人物による出力結果なのかを判断する必要がある。その矩形の重なり率（重なった面積／矩形面積の大きい方の面積）がこの値以下の場合には2つ出力し、この値以上の場合には、最終判定値が高いほうを出力し、低いのは出力しない。
- ・最終判定値閾値：
本実験では精度検証なので閾値設定せず値をすべて出力した。
※図2.18、図2.20の結果例の閾値は61（デフォルト値：学習データでのEERを得る値）

表2.1: 人物検出パラメータ

人物検出パラメータ	サーチ個数	20
	サーチサイズ	15～150
	重なり範囲	0.5
	大分類閾値	20
	最終照合値閾値	※

2.6.1 データセット：USCの評価

データセット：USCは、205枚の画像に対し313人の正面と背面の人物画像が含まれている。図2.17に検出結果のROCカーブを示す。

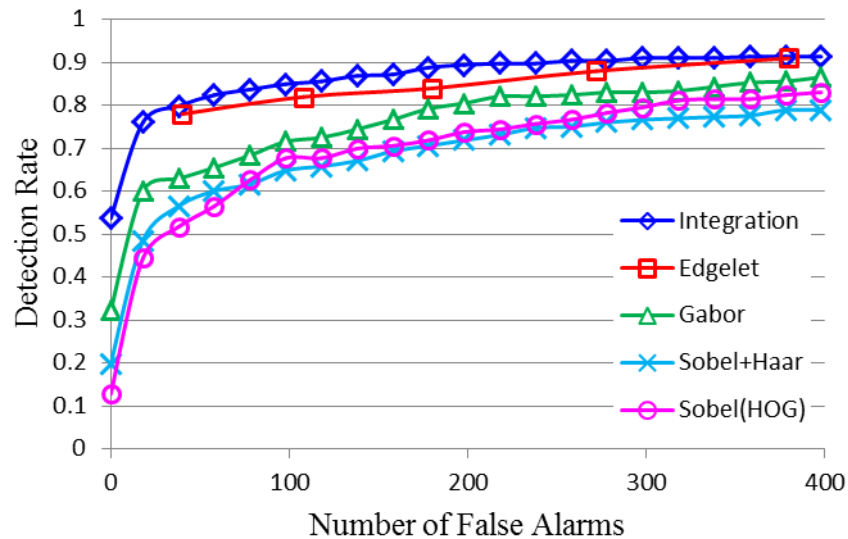


図2.17: データセット：USCの評価結果 (ROCカーブ)

図2.17のEdgeletの結果は、参考文献[7]から結果を抜粋したものである。Gabor, Sobel+Haar, Sobel(HOG)は本アルゴリズムの各特徴量単体での結果であり、Integrationは統合した結果である。図2.17より、各特徴量のみでの評価では、精度はEdgeletよりも低い、統合することで精度が向上していることが分かる。図2.18に検出結果のサンプル画像を示す。



図2.18: データセット：USC[7]の検出結果 (グリーン：検出成功, 赤：誤検出)

2.6.2 データセット：CAVIARの評価

データセット：CAVIAR[26]は、1つのシーンの監視カメラ映像で54フレーム（画像）ある。その中に271人の人物が写っている。75人の人物が部分的に他人と重なりを持ち、18人の人物が部分的にシーンから外れる。図2.19に検出結果のROCカーブを示す。

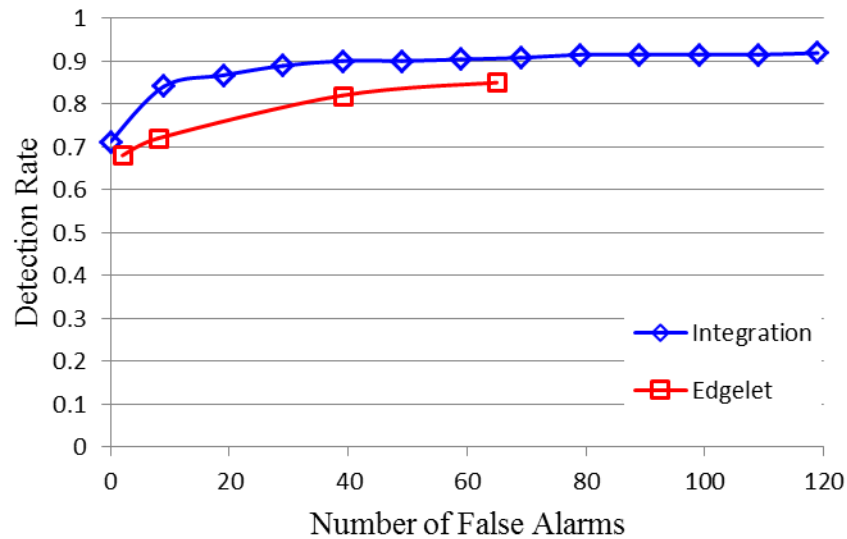


図2.19: データセット：CAVIARの評価結果（ROCカーブ）

図2.19のEdgeletは、図2.17と同様に参考文献[7]から結果を抜粋したものである。図2.19より、20個のエラーを出力する閾値設定で、88%の検出率が得られることが分かる。図2.20に検出結果のサンプル画像を示す。

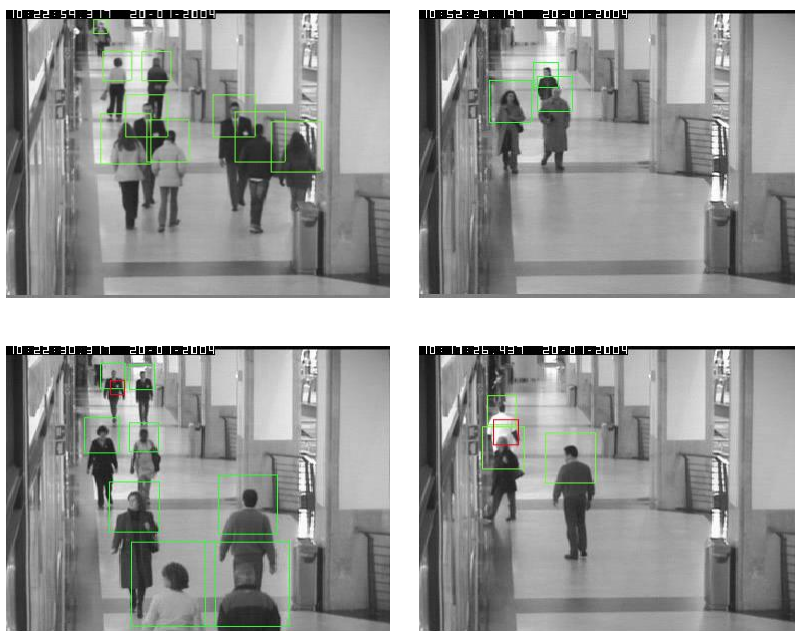


図2.20: データセット：CAVIAR[26]の検出結果（グリーン：検出成功，赤：誤検出）

2.6.3 データセット：NLPR-HSの評価

データセット：NLPR-HS[18]は，1812枚の人物抽出画像と331枚の人物が写っていない風景画像から成る．図2.21にそのサンプル画像を示す．



図2.21: データセット：NLPR-HS[18]評価データ
(上段：ポジティブサンプル，下段：ネガティブサンプル)

人物画像については，抽出済みですべて32画素×32画素である．人物は2枚1セットで左右対称画像が含まれている．人物が写っていない画像のサイズはそれぞれ異なるがおよそ320画素×240画素である．

本実験の評価方法は，Detection Error Tradeoff (DET) カーブ[27]によって行う．DETカーブとは，横軸に誤検出率 (False positives per window) ，縦軸に未検出率 (Miss rate) を両対数グラフで表したものである．誤検出率 (FPPW) は式 (2.27) ，未検出率は式 (2.28) によって表される．図2.17，図2.19のROCカーブは，閾値を変化させることによって，誤検出率 (数) に対する正検出率を求めることができるが，DETカーブでは，誤検出率に対する未検出率を求めることができる．DETカーブでは，左下に向かうほど検出性能が高い．

$$FPPW = \frac{FalsePositive}{TrueNegative + FalsePositive} \quad (2.27)$$

$$Miss\ rate = \frac{FalseNegative}{FalseNegative + TruePositive} \quad (2.28)$$

図2.22に検出結果のDETカーブを示す．

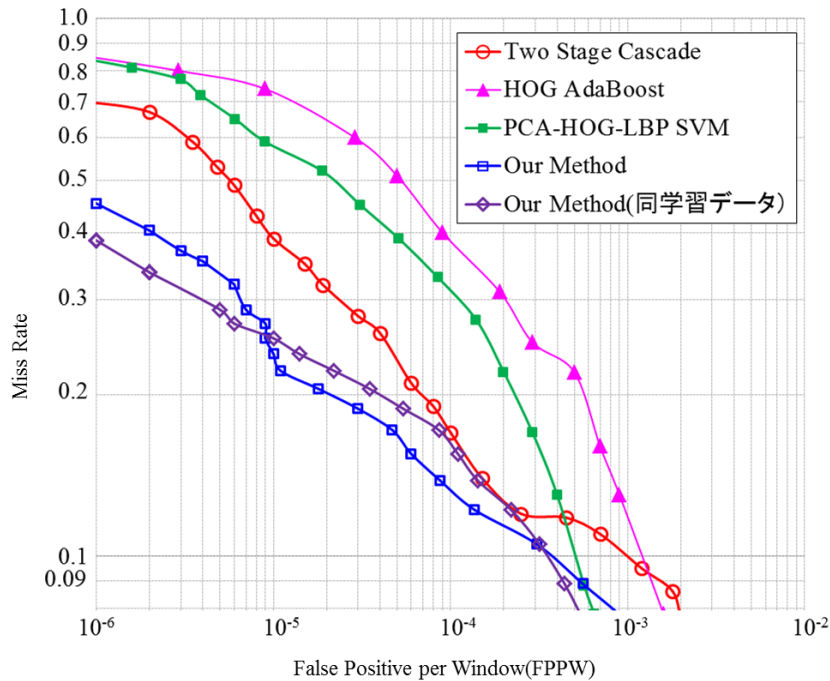


図2. 22: データセット : NLPR-HSの評価結果 (DETカーブ)

図2. 22のTwo Stage Cascadeは, R.Huらによるもの[18]である. HOG-AdaBoostは, M Liらによるもの[16]で, PCA-HOG-LBP SVMはC.Zengらによる研究をR.Huらがインプリメントした結果[18]である. Our Method(同学習データ)は, R.Huらによるもの[18]と同じ学習データを用いて学習しなおしたときの結果である. この学習データは, ポジティブサンプル画像が3510枚, ネガティブサンプル画像が399枚である. 実運用で用いられると考えられるMissRateが小さい範囲 (0.2以下) では, 本研究で収集した学習データによる結果のほうが, 精度が高い傾向がある. この差は, 本研究で収集したネガティブサンプル数 (8395枚) との差が表れたものと考えられる. 各手法と比較すると, 本手法による結果が最も精度が高い結果となった.

2.6.4 処理時間

処理時間については, 設定パラメータ, 特にサーチサイズに依存するところが大きい. R.Huらによるもの[18]では, 320画素×240画素の動画像に対し, ノートPC(Intel Core i5 CPU, 4GB)で, 6fps (frame per second) を実現している. 本アルゴリズムを同解像度の動画像を入力し, 同スペックのノートPCで実行した. 尚, サーチサイズは頭中心から腰中心までのサイズを30~100画素と設定した. その結果, 平均3fpsであった. しかし, 1スレッドでの実行であり, マルチスレッド化やGPUなど使用することで, 高速化は可能であると考えられる. 尚, 本アルゴリズムは, OS:Windows7, コンパイラ: VisualC++2008にて実装した.

2.6.5 評価結果まとめと，本人物検出アルゴリズムについての考察

本章では，公開データベースに対する人物検出精度実験を行った．データセット：USCでは，各特徴量の統合での効果を確認した．データセット：CAVIARやデータセット：NLPR-HSの評価実験結果からは，従来研究の精度と比較し，本アルゴリズムの有効性を確認することが出来た．処理時間については，複数の特徴量を用いているため従来研究と比較すると，長くなっている．しかし，マルチスレッド化やハード的な並列処理の実装など，改善の余地は大いにあると考えられる．

本アルゴリズムは2.4で述べたように，異なる特徴量を組みあわせることで，1つの特徴量では検出しにくい人物画像でも，異なる特徴量で補うことで精度向上を狙ったものである．実際に3つの特徴量を組み合わせることで精度が向上することは2.6.1で示した．このとき，どの要素が精度向上に寄与しているのか，その結果を詳細に分析し，考察を行った．

2.6.1で行った実験と同様に，USCのデータセットに対し，人物検出アルゴリズムを適用した．そしてその結果を正検出（人物部）と誤検出（背景）に分けた．そして，最終評価値を求めるときの各特徴量の評価値も同時に抽出し，それぞれの特徴量の評価値間の相関係数を求めた．その結果を表2.2で示す．

表2.2: 各特徴量の評価値間の相関係数

相関係数	HOG-ガボール	ガボール-ソベルハール	HOG-ソベルハール
人物部	0.52	0.57	0.46
背景	0.17	0.16	0.20

表2.2から，人物部については，すべての特徴量の組み合わせにおいて，相関係数は約0.5であり，1つの入力画像に対して特徴量の評価値に差がある場合もあればない場合もある．しかし，特徴量の評価値に差がある画像をピックアップしてみると，ある特徴量では検出するのが苦手なパターンを他の特徴量では検出できていることが観測できた．各特徴量間での差の大きかった画像例を図2.23に示す．



図2.23: 特徴量の評価値間で差が大きかった例

- (a) ガボールの評価値が大きく，HOGの評価値が小さい
- (b) ガボールの評価値が大きく，ソベルハールの評価値が小さい
- (c) HOGの評価値が大きく，ガボールの評価値が小さい

図2. 23の(a)でHOGの値が小さくなった理由は、服装の模様が縞模様（赤点線楕円）になっており、低周波成分のエッジの方向があらゆる方向に散らばってしまったことが原因だと考えられる。図2. 23の(b)は、暗い画像であり、エッジの強度が小さいことが原因である。ソベルハール特徴は、エッジ強度を用いるので評価値が小さくなったと考えられる。図2. 23の(c)は、背景模様で7画素程度の低周波の縞模様のようなテクスチャが観測でき、ガボール特徴も7画素×7画素で1周期であるので、その影響を受けたと考えられる。このように精度が向上した要因の1つは、異なる特徴量間で苦手とする画像パターンを補っていることが考えられる。

背景部については、表2.2から、各特徴量間の相関係数は0.2以下である。ほとんど相関はないと言える。そのため、ある1つの特徴量の評価値が大きい場合でも、その他の特徴量の評価値は小さい可能性が高い。その結果最終的に求める評価値も小さくなる。背景部の各特徴量間での差の大きかった画像例を図2. 24に示す。

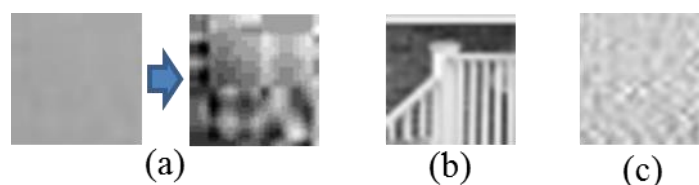


図2. 24: 特徴量の評価値間で差が大きかった例

- (a) ガボールの評価値が大きく、ソベルハールの評価値が小さい
- (b) ソベルハールの評価値が大きく、ガボールの評価値が小さい
- (c) HOGの評価値が大きく、ソベルハールの評価値が小さい

背景部において、ガボール特徴の評価値が高くなるのは、図2. 24の(a)の左図のように、ほとんど無地のような部分が多い。図2. 24の(a)の右図は、コントラストを引き伸ばし調整したものであるが、このようにノイズ画像になっている。ガボールはこのようなノイズ成分の評価値が高くなる傾向があるが、ソベルハール特徴量の評価値が十分小さいため、カバーできている。逆に図2. 24の(b)はソベルハールの評価値が大きかったが、ガボールの評価値が小さくなった。図2. 24の(c)はHOGの評価値が大きく、ソベルハールの評価値が小さかった。ガボールやHOGは各方向のエッジをその強度で正規化する。それはコントラストが小さい画像（暗い画像等）でも検出する目的のためであるが、ノイズ成分の評価値も高くなる可能性がある。本アルゴリズムでは、それを違う周波数の特徴量や、強度を主とした特徴量も加えることで、リジェクトできるようになっていると考えられる。

このように3つの特徴量を組み合わせることが、人物部と背景部どちらにも有効に働き、精度が向上したと考えられる。

さらなる精度向上については、さらに特徴量を取捨選択することや、学習手法の改善が考

えられる。特徴量については、従来法ではHOG特徴やそれに類似する特徴がよく用いられているが、本研究で示したように、その特徴と相関が少なく、その特徴が苦手とするパターンをカバーできる特徴を組み合わせることで、精度向上の可能性がある。学習手法については、画像から何かを検索する用途には、その処理時間の観点からブースティングに関連する学習手法が一般的には用いられることが多い。これらは、多数の弱識別器の繋がりで構成され、判定途中で棄却できる仕組みを作ることが出来る。画像全体を走査するときそのメリットを生かして、ある画素の判定中に明らかに探索しているものと違う場合はすぐに棄却し、次の画素に処理を移すことで、高速な処理が可能であるからである。本研究でも、LDAArray[19]だけでなく、AdaBoost[28]、LogitBoost[29]、RealAdaBoost[30]などのブースティングアルゴリズムを試したが、結果としてはLDAArray[19]が最も精度が高かった。学習手法自体の改善が精度向上につながることは勿論考えられる。その他には、本研究では特徴量ごとに識別器を作成したが、すべての特徴量を合わせて1つの識別器を作成すると、精度が向上する可能性がある。それは各特徴量の最小単位での組み合わせが可能であるからである。例えばガボールの上方向のエッジとHOGの下方向などである。ただし、処理が複雑化するため、処理時間は増える可能性がある。

2.7 飛行ロボットが取得した画像での人物検出実験

飛行ロボットが撮影する画像に対し、人物検出実験を行った。用いた飛行ロボットについては、第3章の認証実験において述べる。人物認証のための飛行ロボットの動きとしては、本研究では汎用的な飛行ロボットに適用することができるように、飛行ロボットの運動機能に依存した複雑な動きは想定していない。よって、出来るだけ高度3(m)を維持しながら、ロール角とピッチ角は 0° ，すなわち地面に対して平行な姿勢を基準として維持しつつ、ヨー角(向き)のみ可変で移動するようにした。本実験では飛行ロボットを手動で操作した。また、飛行ロボットに搭載したカメラは、上空から地上の方向に向かって見下ろすように撮影するため、俯角をつけて搭載する。またそのカメラは、俯角のみ可変であるが、飛行中は固定した。本実験では3種類の俯角(15° ， 30° ， 50°)で撮影を行った。そのため、撮影された人物は、透視投影変換の結果、画像上ではその俯角と画像上の位置に依存して、回転した状態となる。図2. 25にそのサンプル画像を示す。尚、屋内環境での撮影なので、風の影響はほとんどないと考えられるが、飛行ロボット自身は人物の歩行速度(5km/h)程度で移動しているため、上記基準の姿勢を出来るだけ維持するように操作を行ったが、飛行ロボットの姿勢の変動は生じていると考えられる。

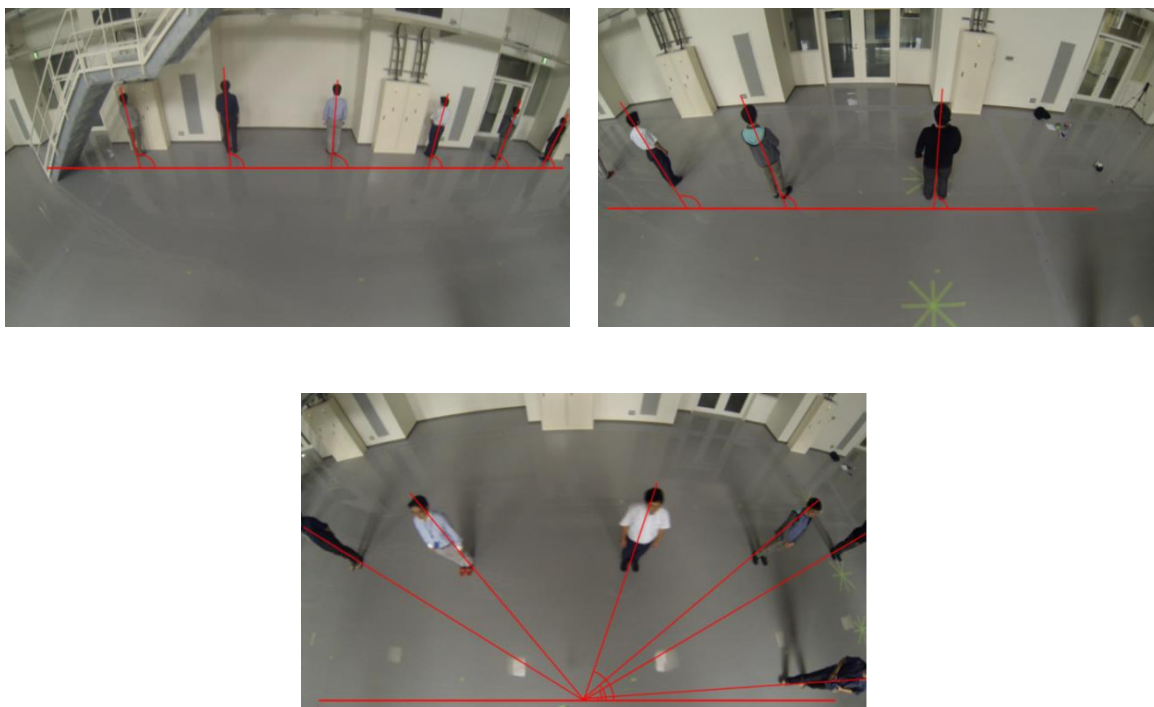


図2. 25: 飛行ロボットが撮影したデータサンプル (自ら撮影)
(左上: 俯角 15° ，右上: 俯角 30° ，下中央, 俯角 50°)

2.7.1 人物検出アルゴリズムのための画像変換

本実験で使用する飛行ロボットは、先に述べたように、高度を維持しながら出来るだけ地面に対して姿勢を平行にしながら移動する。その状態を維持しながら、カメラで地面を見下ろすように設定した俯角で撮影を行う。そのため図2. 25のように、その俯角によって画像に写った人物は回転する。この回転した人物を検出する手法は2通りある。1つは、回転した人物を検出できるように、人物検出アルゴリズムを改良する方法である。具体的には 360° 対応したテンプレートを作成するか、 360° を例えば 10° ごとに区切り、36個のマルチテンプレートを作成し検出処理を行うかである。 360° 対応したテンプレートは、画像としての変動が大きすぎ、作成は困難である。また、36個のマルチテンプレートによる処理も、特徴量抽出以外の演算処理は基本的には線形倍かかるので、処理時間が増大し、現実的ではない。もう1つの方法は、入力画像を人物検出アルゴリズムが適用できるように変換することである。本研究では、飛行ロボットに搭載しているカメラの俯角が出来るだけ変動しないように撮影するため、設定した俯角により透視投影変換から画像に写った人物の回転角は演算できる。人物検出アルゴリズムは、人物の角度が垂直であれば、検出を行うことが出来る。そのため、本実験では極座標変換により画像を変換し、その変換画像に対し、人物検出アルゴリズムに適用することとした。図2. 26にその変換例を示す。尚、認証実験時には、極座標画像上の検出位置を、元画像での位置に変換し、回転角に対応しつつ上半身の切り出しを行った。

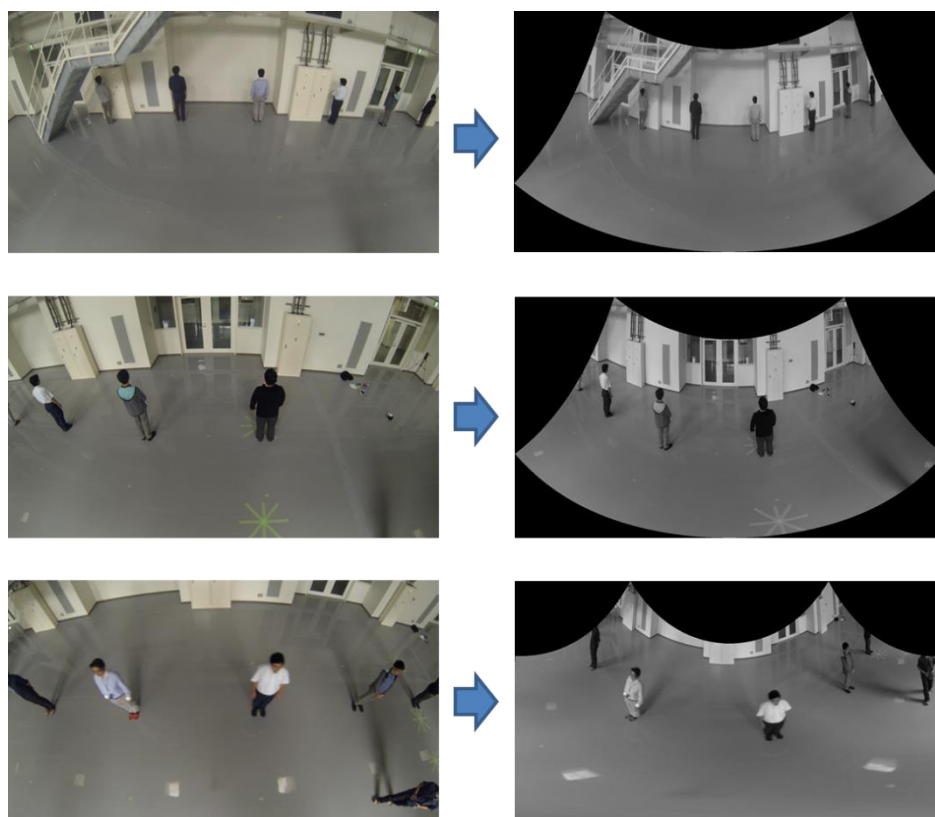


図2. 26: 人物検出のための画像変換（上から順に、俯角 15° , 30° , 50°)

2.7.2 人物検出実験

飛行ロボットによる実験対象者6人の撮影を合計15回行った。各データパターンの詳細については、第3章の「3.5 飛行ロボットが撮影した画像からの人物検索実験」にて述べる。本節では、その人物認証精度確認実験用のデータに対して、人物検出アルゴリズムを適用し、求めた人物検出精度を示す。飛行ロボットに搭載したカメラの俯角と、人物の向き（水平角）のバリエーションの組み合わせにより、合計15個の撮影パターンを作成した。全パターンで共通する要素として、飛行ロボットは高度約3(m)を保ちながら撮影を行った。また、撮影は4fps(frame per seconds)で、各パターンにおいて約10数秒間の飛行である。入力画像サイズは1280×720画素である。また、人物検出パラメータは表2.1と同じである。

尚、検出精度を求めるための答え（画像上の人物位置）は、全フレーム、人の目視で入力を行った。そのときの条件として、目視確認できる全人物を入力した。そのため、人物の仰角が50° 以上のものであっても、答えとして入力されている。

2.7.2.1 人物が静止状態での飛行ロボットの撮影による検出実験

図2.25で示したように、実験対象者6人は、2.5(m)の間隔を開けて直線状に並び、直立静止状態になった。それに対して、飛行ロボットは時速約5(km/h)で、平行に移動しながら撮影を行った。また撮影パターンとして、飛行ロボットに対する人物の向き（水平角）を90° ずつ4方向撮影した（0° , 90° , 180° , 270° ）。その各向きに対し、飛行ロボットに搭載したカメラの俯角を3パターン（15° , 30° , 50° ）変え、撮影を行った。よって人物が静止状態のパターンとして、合計12パターンの撮影を行った。図2.27に1つのパターン：俯角30° , 人物の向き0° での入力画像群の例を示す。

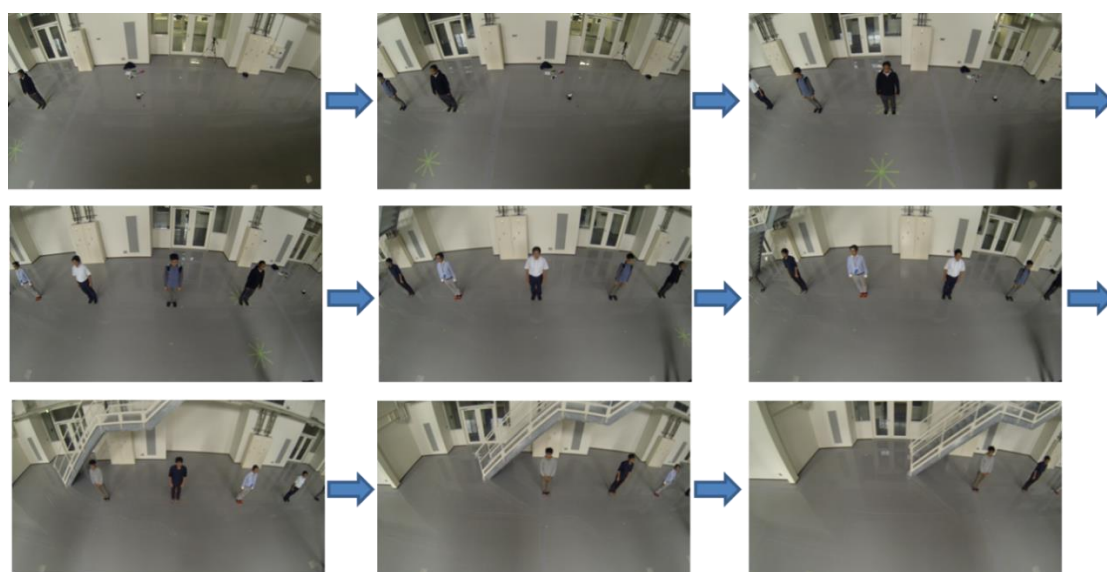


図2.27: 飛行ロボットが撮影した入力画像サンプル（俯角30° , 人物の向き0° ）

図2.28, 図2.29, 図2.30にそれぞれ 15° , 30° , 50° の検出実験結果を示す。

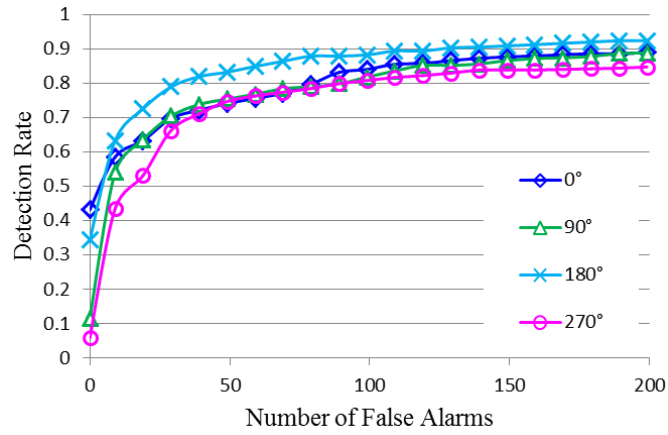


図2.28: 俯角 15° の評価結果 (ROCカーブ)

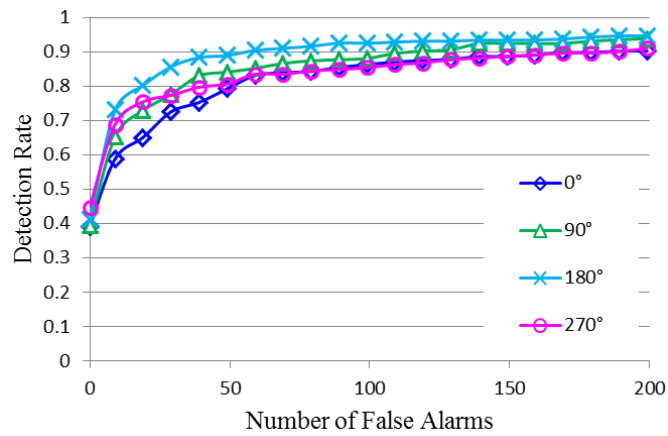


図2.29: 俯角 30° の評価結果 (ROCカーブ)

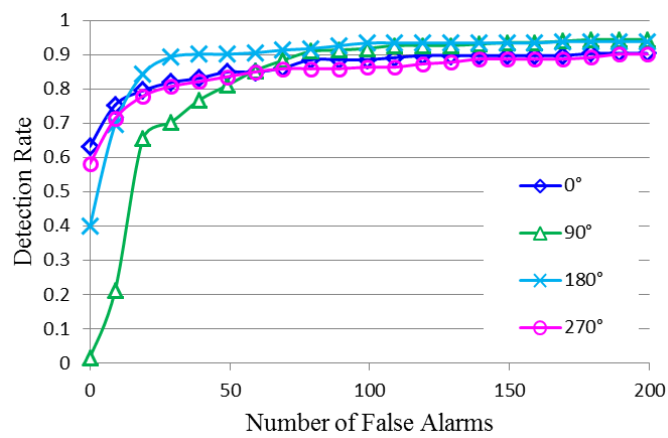


図2.30: 俯角 50° の評価結果 (ROCカーブ)

図2. 28-図2. 30より，どの俯角でも人物の向きが 180° のときの検出率は，他の人物の向きの検出率と比較して最も高い．誤アクセプト数が100のときの， 180° の検出率を他の向きの検出率と比較すると，平均10%高いことが分かる．また，誤アクセプト数 (Number of False Alarms) が200のときでも平均5%高い．これは， 180° のときは後ろ向きであるが，そのときの人物の Ω 型のラインが最も出やすかったことが影響しているのではないかと考えられる．横向き (90° , 270°) は肩のラインが出なくなり，特徴が減少するため検出率が落ちる．正面向き (0°) は，頭のラインのエッジが後ろ向き (180°) と比較して出にくいこと影響していると考えられる．さらに，顔や服装のテクスチャが余分なエッジとして表れ，検出率に影響を与えていることも考えられる．俯角による精度差としては，俯角 15° の検出率が他の俯角の検出率と比較して約5~10%低い．これについては背景模様の違いが影響している．俯角が大きい場合，背景としては大半が地面であり，本実験の環境では地面は無地であるため，誤検出数が少なくなった．一方，俯角が小さくなるとドアや窓などが写るため，そのエッジの影響で誤検出数が増えた．

図2. 28, 図2. 29, 図2. 30で示した結果をすべてまとめて，FAR (False Accept Rate), FRR (False Reject Rate)のグラフで表した．図2. 31にそのグラフを示す．

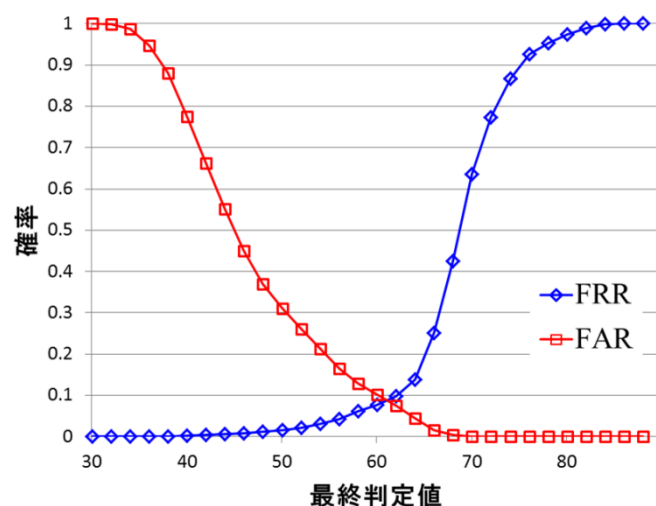


図2. 31: 集計結果 (FRR-FARグラフ)

図2. 31から，具体的に閾値を設定したときのFRRとFARが分かる．FRRは間違えて人物をリジェクトしてしまう確率で，FARは間違えて人物以外を受け入れてしまう確率である．この値は後の処理に対する影響を考慮して設定するとよい．本研究では後の処理が人物認証である．このときに，例えば，誤検出画像が入ってもよいから，人物画像を誤リジェクトしたくないのであれば，この閾値を小さくすればよい．逆に，誤検出画像を出来るだけ入力したくない場合は，閾値を高く設定する必要がある．これは，人物認証アルゴリズムの性質と，それを用いたアプリケーションとしての動作を考慮して設定すればよい．

仮に閾値をEER (Equal Error Rate)が得られる61に設定したときの検出結果例を図2. 32~図2.

34に示す. 緑枠が正しく人物を検出できた場合で, 赤枠が背景を誤検出した場合である.

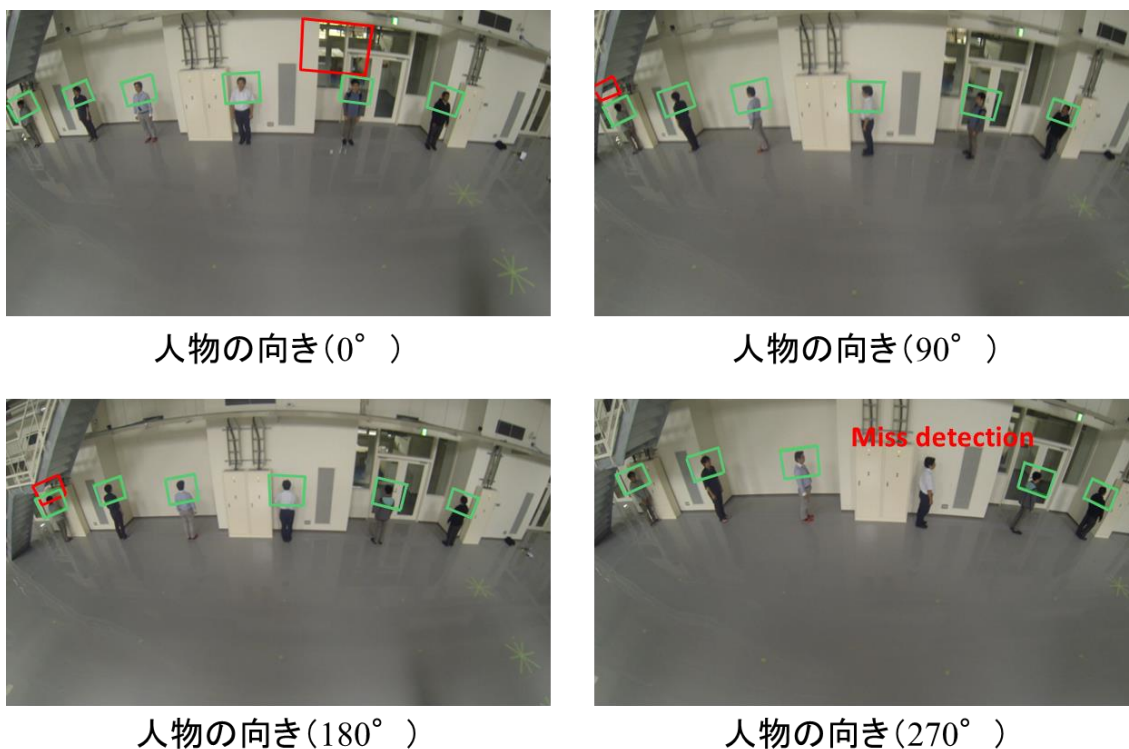


図2. 32: 俯角15° の検出結果画像

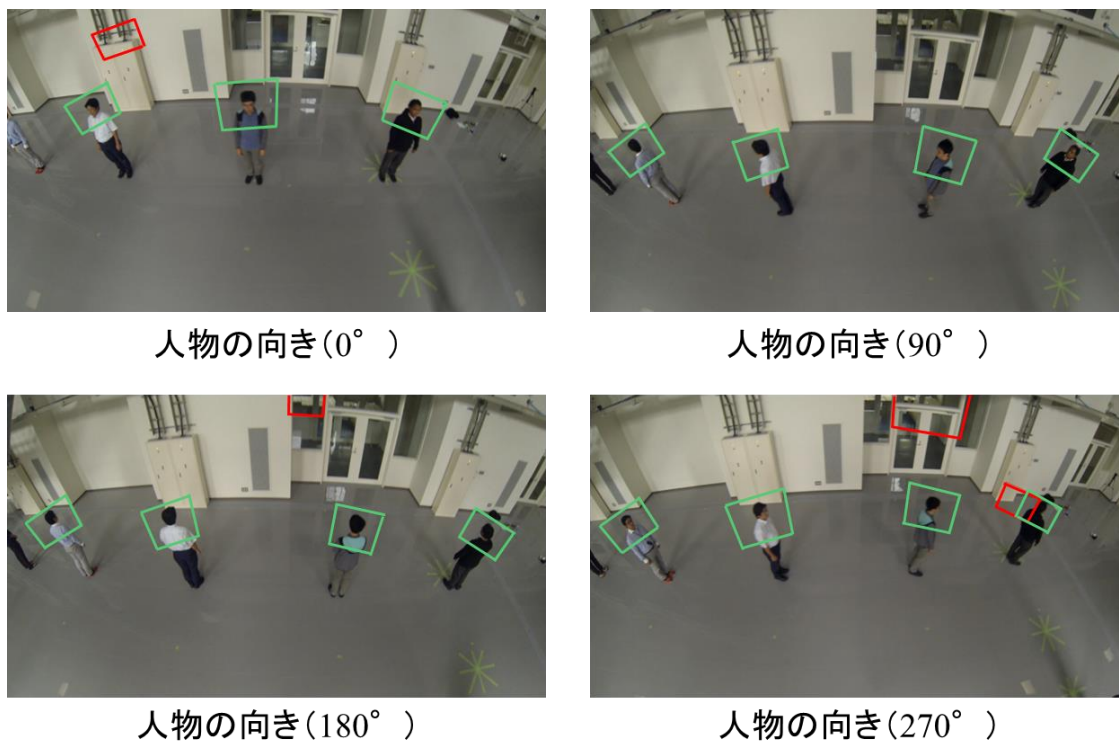


図2. 33: 俯角30° の検出結果画像

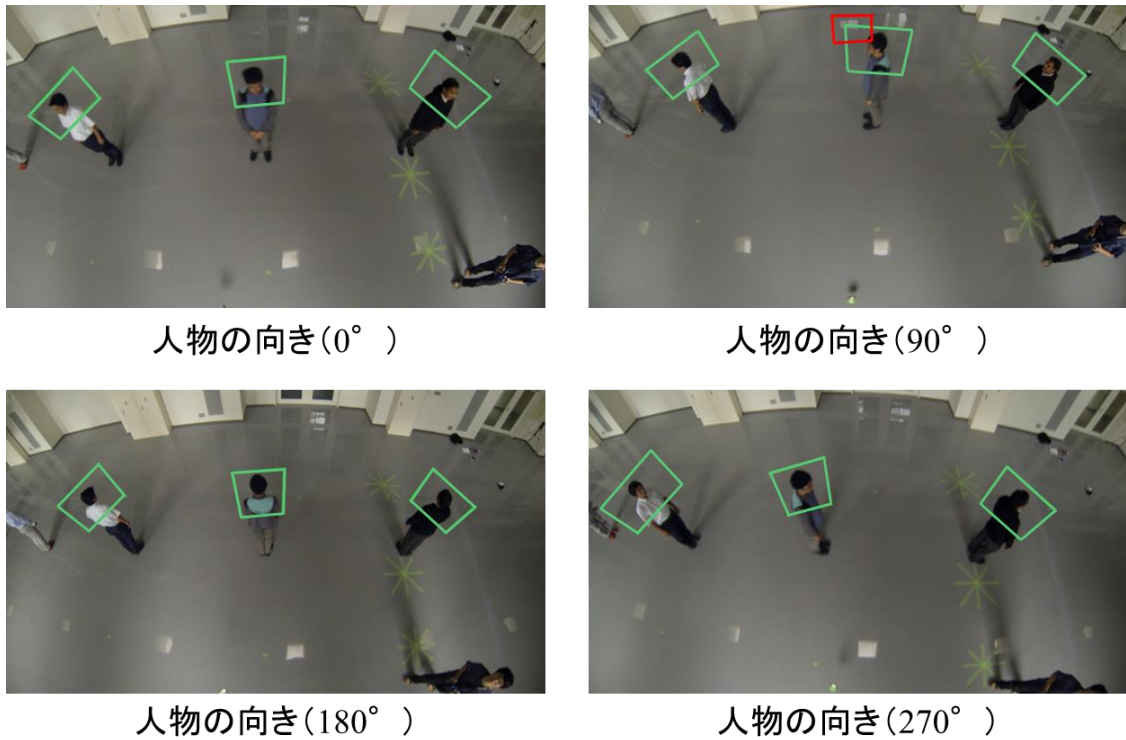


図2. 34: 俯角50° の検出結果画像

2.7.2.2 人物が移動状態での飛行ロボットの撮影による検出実験

飛行ロボットに搭載したカメラは俯角30° で取り付けた．本節では，人物が移動している状態で飛行ロボットによる撮影を行った映像に対し，人物検出アルゴリズムを適用した結果について述べる．合計3つの撮影パターンで撮影を行った．撮影パターンの詳細については，第3章の人物認証実験で述べる．パターンの概略としては，第1のパターンは，6人の実験対象者は通路を同じ方向に歩行し，飛行ロボットは実験対象者の正面から撮影した．飛行ロボットは人物群を上空で通過した後，振り返り，人物の背面を撮影しようとした．第2のパターンは，第1のパターンと同様に，6人の実験対象者は通路を同じ方向に歩行する．飛行ロボットは，その人物群の手前に位置し，同じ方向に移動しながら，人物の正面方向を撮影した．つまり，飛行ロボットは人物群の前方を後ろ向きに移動しながら人物を撮影した．第3のパターンは，6人の実験対象者にはランダムな方向に歩行してもらった．それに対し，飛行ロボットは全員が万遍なく写るように移動しながら撮影した．それぞれのパターンの検出実験結果のROCカーブのグラフを図2. 35に示す．また，検出結果例を図2. 36～図2. 38に，検出結果のFRR-FARグラフを図2. 39に示す．尚，検出結果例（図2. 36～図2. 38）の閾値は，2.7.2.1と同様に61である．

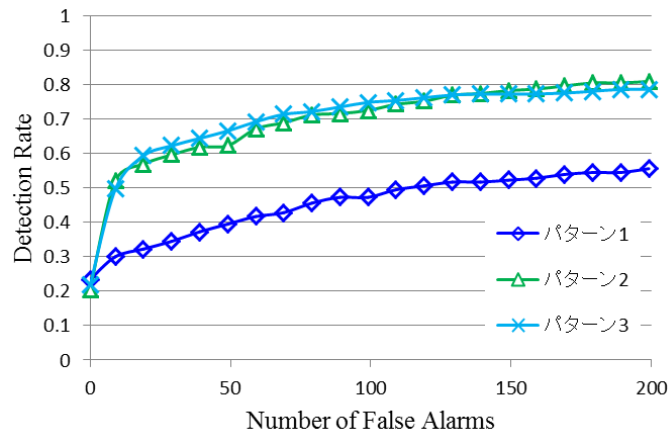


図2.35: 各移動パターンでの評価結果 (ROCカーブ)



図2.36: パターン1の検出結果画像



図2.37: パターン2の検出結果画像

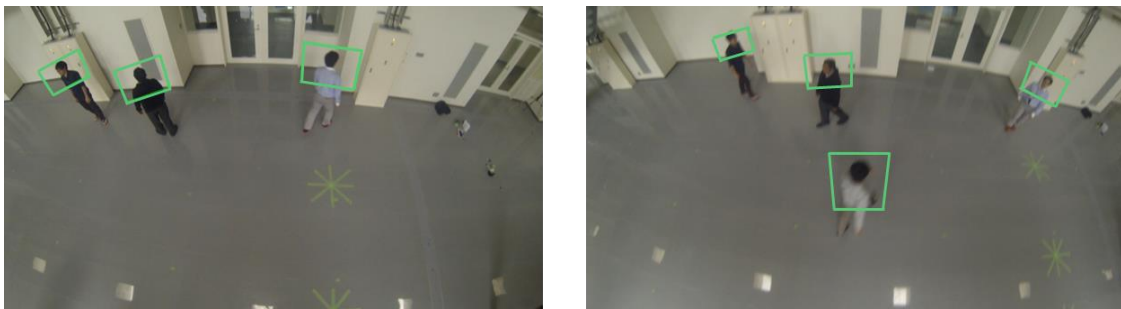


図2.38: パターン3の検出結果画像

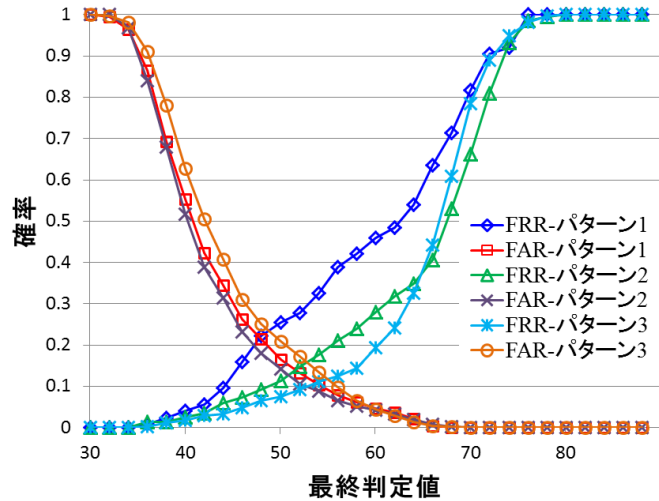


図2. 39:人物が移動中の3パターンの評価結果 (FRR-FARグラフ)

図2. 35より，パターン1の人物検出率は，パターン2や3と比較して，平均して約0.3も低いことが分かる．図2. 39を参照すると，各パターンのFAR曲線の差は小さい．しかし，FRR曲線に差が確認できる．パターン1のFRRが高くなっており，最終判定値が50～60の範囲では，他のパターンと比較して，FRRにおいて約0.15～0.20の差がある．つまり，誤検出の数や誤検出画像の最終判定値に差があるわけではなく，パターン1の人物のリジェクト率が高くなっていることが，検出率低下の原因である．パターン1の画像を分析してみると，図2. 36の右図に示したように，誤リジェクトが多く発生していることが分かった．パターン1は，最初に人物に正対して数秒進み，その後振り返って人物の背面の画像を取得しようとしている．そのため飛行ロボットの移動方向は，人物と逆方向であるため相対速度は大きくなり撮影した画像がブレた．さらに，飛行ロボットの振り返り動作は，約1秒間で行っており，高速な回転により，そのときに撮影した画像も大きくブレた．そのため，そのブレの影響により人物のΩ型のエッジが出にくくなり，人物検出アルゴリズムで検出できなかった人物画像が他のパターンよりも多く存在したため検出率が低下したと考えられる．ブレの影響により検出できなかったと考えられる検出失敗画像例を図2. 40に示す．

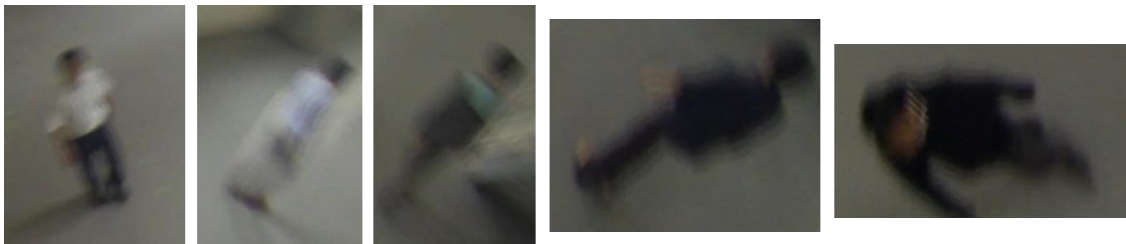


図2. 40: パターン1の検出失敗画像例

2.7.3 飛行ロボットが撮影した画像での人物検出実験まとめと考察

飛行ロボットが撮影する画像での人物検出実験では、第1に、撮影される画像内の人物の角度が回転した状態になることが、検出を行う上で対応すべき課題であった。飛行ロボットは上空から見下ろすように撮影するため、透視投影変換の結果、画像上ではそのカメラの俯角と画像上での位置に依存して、撮影される人物画像は回転した状態となる。これについては、飛行ロボットに搭載するカメラの俯角は固定で、飛行ロボットは地面に対して傾かずに移動することを前提として、入力画像を極座標変換することで、人物検出アルゴリズムに適用できるようにした。人物が静止状態の検出実験では、 15° ~ 50° の俯角の違いや、人物の向きの違いによる検出率の差がないことを確認し、すべてを集計した結果、EER10%の検出率であった。人物が動作状態では、ブレの影響によると考えられる検出率の低下はあるものの、検出可能であることを確認した。

飛行ロボットと人物との相対速度が大きくなると、人物画像のブレが大きくなり、人物検出を失敗する。しかし、飛行ロボットのように、カメラ自体が動きながら撮影する場合は、カメラの性能によって大小あるものの、ブレの問題は必ず残ると考えられる。そのため、人物を撮影するときに、出来るだけ画像がブレないような飛行ロボットの動きの工夫も必要であると考えられる。例えば、人物と併走することで相対速度を小さくすることや、必要以上に動き回らずに最大限の画像を取得する軌道上を飛行するなどである。

2.8 まとめ

本章では、人物検出アルゴリズムについて述べた。本研究により作成した人物検出アルゴリズムは、異なる3つの種類の特徴量を用いたアルゴリズムである。それは、ガボールエッジ特徴[25] (7×7)、ソベルエッジ特徴、ソベル+ハール特徴である。ガボール特徴 (7×7) で低周波のエッジ特徴を、ソベルエッジ特徴で高周波のエッジ特徴を、ソベル+ハール特徴で、エッジの集まりや繋がりを取得できると考えた。また、それらの統合方法にも工夫を施した。それぞれの特徴量の識別器はLDAArray法[19]を用いて作成し、それらの統合には、線形判別分析を適用した。その結果、精度は従来手法と比較して高精度であることが確認できた。処理時間は、従来手法よりも多くかかっているが、ソフトウェア、ハードウェアの両面で改善の余地が十分であることを述べた。また、飛行ロボットが撮影した画像に人物検出アルゴリズムを適用した。飛行ロボットが撮影した画像でも人物検出が可能であることと、画像ブレが大きくなると検出率が低下する問題点を確認した。画像ブレについては、カメラの性能、飛行ロボットの動きの改善で対応できる可能性がある。

本アルゴリズムによって検出できる人物画像の条件は、人物の頭から肩のライン (Ω 型)のエッジが見えることである。そのため、屈んでいる、寝転んでいる等の姿勢や、頭に手を置いている、大きな帽子を被っている等の Ω 型が崩れるときには検出率は低下する。同様に飛行ロボットが撮影した場合に起こり得ることは、飛行ロボットと人物の位置関係によって

は、人物の仰角が大きくなり（ 50° 以上）， Ω 型が出なくなることである．この場合も検出率は低下すると考えられる．これは人物検出アルゴリズムが Ω 型を検索のキーとして用いているためである．このような人物画像を検出するためには，その画像でのテンプレートの学習は最低限必要であると考えられる．その他，顔検出アルゴリズムや，動体追跡アルゴリズムとの併用，組み合わせにより検出精度を上げられる可能性はある．

第3章 飛行ロボットによる人物認証

3.1 はじめに

本章では，飛行ロボットが撮影した画像から対象人物を探索する手法を提案する．図3.1に想定する飛行ロボットによる人物探索のイメージ図を示す．

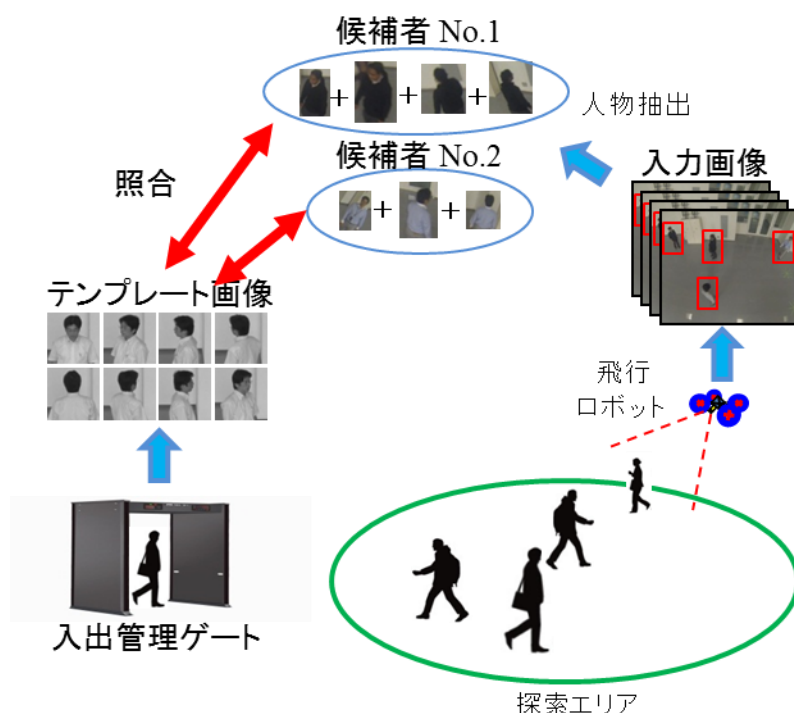


図3.1: 飛行ロボットによる人物探索のイメージ図

本研究で想定する飛行ロボットによる人物探索の適用シーンは，遊園地やショッピングモールなどの閉鎖空間で迷子や犯人の探索である．図3.1で示したようなシステムを想定している．入出管理ゲートで撮影した人物画像と，飛行ロボットが探索エリア内を移動しながら撮影した人物画像を照らし合わせることで人物探索を行う．本章では，その人物画像を照らし合わせる手法である，人物認証アルゴリズムを提案する．それは，飛行ロボットが撮影した飛行ロボット特有の入力画像のメリットを生かすことが出来る，累積アピランンス照合アルゴリズムである．そのアルゴリズムを用いて特定人物の探索を行う．そのアルゴリズムを用いれば，飛行ロボットは1人1人の人物に接近せずに，複数人の人物間を移動しながら撮影し

た画像をもとに、人物検索を行うことが出来る。移動しながら撮影した複数視点からの低解像度の画像の認証を重ねることで、認証精度を高めることが出来るからである。

ここで、本システムの具体的な飛行ロボットの使用方法について述べる。迷子や犯人などのターゲット探索の要請を受ければ、飛行ロボットが探索エリアへ飛び立ち、上空を移動しながら映像を撮影する。映像内から認証アルゴリズムによりターゲット探索を行い、ターゲット候補を見つけ出す。最終的には探索を依頼した人が、そのターゲット候補の確認行為（映像を確認する、飛行ロボットを通じてターゲットと会話する等）を行うことで探索終了となる。よって本手法で目指す認証精度は、指紋認証や虹彩認証のように、その人物を確定し、入出管理の開閉キーに使用するような高い認証精度ではない。認証アルゴリズムの精度を高めることで、この映像の確認行為の回数を出来るだけ減らし、素早くターゲットを見つけ出すことが目的である。よって本研究では、50～100人の中から1人の候補に絞り込める認証精度を目標とする。

3.2 関連研究

飛行ロボットが上空から撮影する人物画像は、人物同士の重なりで、足部はオクルージョンになりやすい。また、人物の安全面やその行動を阻害しないことを考慮すると、飛行ロボットは出来るだけ遠方から人物を撮影することが望ましい。そのため、取得する画像は低解像度が想定される。さらに、通路などを移動中の人物を飛行ロボットが移動しながら撮影するので、撮影される人物の角度はさまざまである。よって、本研究では、上半身の低解像度の人物画像で、さまざまな角度の情報を利用できる方式を考えた。

カメラを用いた人物認証手法は種々あるが、飛行ロボットが撮影した画像を認証する場合は、人物の風貌特徴を認証するアピアランス認証が適している。その他の人物認証手法としては、顔認証[1]、歩容認証[2]、指紋認証、虹彩認証、静脈認証などが挙げられる。指紋認証、虹彩認証、静脈認証については、カメラに対して指や手をかざす、カメラに目を合わせるなど、認証される側の動作を伴った撮影を行い認証するので、今回想定する飛行ロボットによる人物検索には適していない。顔認証は、遠方から撮影した人物の認証手法として最も有望であり精度も高い。顔認証は、飛行ロボットが人物の正面付近に位置する場合には用いることが可能であるが、位置によっては顔画像を撮影できない場合もある。また顔認証を行うためには、目間のサイズが数十画素以上必要である。そのため適用機会が制限される。歩容認証は、一定歩数の間の全身の動きを特徴量として認証する手法である。上空からの映像では下半身が映らない可能性があり、この手法もあまり適していない。このように、飛行ロボットによる認証を行うためには、課題や制限が多いため適用可能な認証手法も限られる。また、一般的な個人認証技術には、RFIDなどの機器を用いたものもあるが、認証される人がICタグなどの機器を所持しておく必要がある。犯人捜索ではそもそも機器などは所持しておらず、迷子捜索では、認証され

る人が機器を紛失する可能性があるため向いていない。そのため本研究では風貌による認証を行うことが出来るアピアランス認証が最も適していると考えた。

アピアランス認証手法の最新手法としては、色特徴とその他の特徴の組み合わせによるものが多い[3,4,5,6,7]。これらは安定した色特徴を取得できる場面では有効である。しかし、飛行ロボットによる認証は、照明変動が大きい場面も想定される。さらに、テンプレート撮影用カメラと、飛行ロボットに搭載しているカメラが異なる場合がある。その場合、異なるカメラ間の色校正も影響するため、色特徴は用いにくい。よって、我々はエッジ特徴のみで人物検索を行う。エッジ特徴を主として用いている研究としては、人物の全身の模様特徴を用いた研究がある[8,9]。著者らは、飛行ロボットが撮影する特有の画像として、低解像度の上半身画像という厳しい条件下で認証できる方式にチャレンジした。その場合、頭や肩のラインや、服装の大まかなエッジがその人物を見分けるためのロバストな特徴の1つとなる。画像1枚では認証精度は低いが、飛行ロボットの利点を生かし、異なる方向から撮影した複数枚の人物画像の情報を蓄積させることで、精度向上を図った。

監視カメラのような固定カメラでの複数枚テンプレート、複数枚入力画像の照合では、一般的には全テンプレート画像と全入力画像の組み合わせ数をすべて照合した後、その最大値を照合値とする手法がある。その他の伝統的な手法としては、Kittlerらは、複数回の照合値を平均で判定した[10]。Maedaらは、テンプレート画像群や入力画像群を部分空間で表し、部分空間どうしの距離で評価した[11,12]。さらに従来研究としては、Nandakumarらは複数回の認証結果や複数種類の認証結果から得られる照合値の分布を、事前に取得した照合値の確率分布と比較し判定することで認証精度を向上させる、確率的融合判定を用いた[13]。これらの方法は、通路に設置された監視カメラのように、人物の移動方向や入力画像枚数がある程度予測できるものであれば、有効である。複数回事前に歩行テストをすることで、照合結果の確率分布も取得可能だからである。さらに、Bakらは、平均リーマン共分散行列に基づく手法で複数フレームの情報を統合する手法を提案した[14]。Bazzaniらは、認証を行う前に、複数フレーム画像を取捨選択する手法を提案した[15]。飛行ロボットが撮影した画像から作成した入力画像は、背景を誤検出した画像や重複を含むあらゆる方向から撮影された順不同な人物画像である。このような入力画像に対応する必要がある。さらに、画像1枚の照合結果を累積する仕組みにしておきたい。なぜなら、将来、入力画像の解像度が改善され、照合アルゴリズムを改変しようとした場合、画像1枚単位の処理と統合処理を分離しておく、変更が容易になるメリットがあるためである。また、画像1枚の入力ごとに照合値を出力できるため、応答性も最速となる。著者らはこのような条件を満たすアピアランス認証アルゴリズムを開発した。

ロボットによる人の認証に関する研究も行われている。Morbidi[16]らは、人の肩の動きや足の動きをトラッキングすることによる歩容認証を提案している。Bellotto[17]らは、服装情報だけでなく顔情報を組み合わせた人物認証を提案している。Satake[18]らは、服の模様から得たSIFT特徴を用いて、人物追跡ロボットによる認証を提案している。著者らはさらに、複数人の認証対象者間を移動しながら撮影した画像をもとに、複数人の認証が行えるアピアランス認証アルゴリズムを開発した。

3.3 飛行ロボットによる人物搜索手法

3.3.1 飛行ロボットによる人物搜索システム

本論文で想定する人物探索システムは、入出管理ゲートにより人の出入りを管理している場所において、特定の人物の探索を行う。このような入出の管理は、大型遊園地[19]ではすでに導入されている。人物認証を行うためのテンプレートは、入出管理ゲートを通過する際に様々な角度から撮影した複数枚の上半身の画像である（図3. 1）。人物搜索時には、飛行ロボットは搜索エリアに移動し、飛行しながらさまざまな視点から複数枚の人物画像を撮影する。その撮影した画像とテンプレートとを照合することで、人物搜索を行う。図3. 2に飛行ロボットによる人物認証システムのフローチャートを示す。

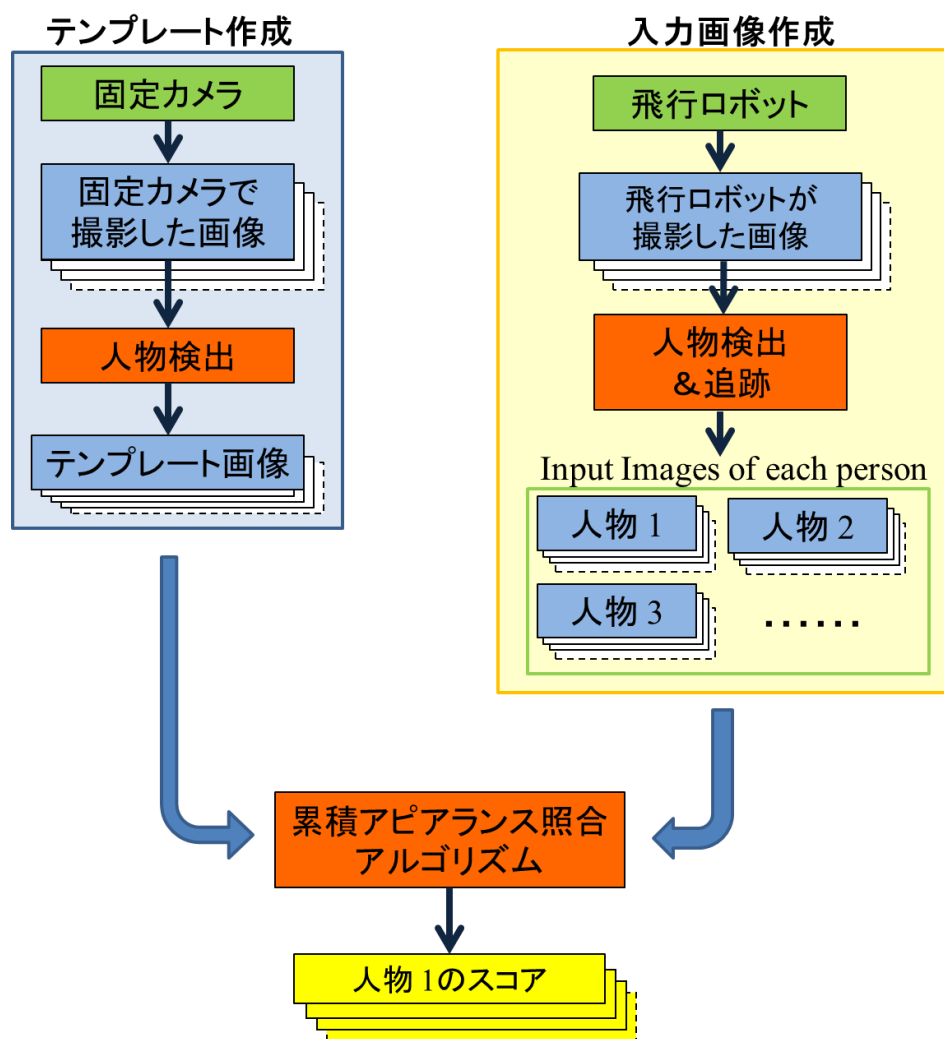


図3. 2: 飛行ロボットによる人物認証システムのフローチャート

上で述べたように、図 3.2 のテンプレート作成では、入出管理ゲートで撮影した複数枚の異なる視点から撮影した画像を用いる。それらの画像から人物の上半身部分を抽出し、テンプレート画像を作成する。入力画像作成でも、飛行ロボットが撮影した画像から人物部分を抽出する。これらの人物画像抽出には、第 2 章で述べた人物検出アルゴリズムを適用することができる。また、入力画像には、複数人の人物が複数フレームに渡って写っており、照合するためには、各人物ごとの画像群にまとめる必要がある。これは従来の追跡手法[20]で行うことが出来る。

このようにして作成したテンプレート画像群と各人物の入力画像群に、照合アルゴリズムを適用し、照合値を出力することで人物認証を行う。その値が設定した閾値以上となれば検索対象人物の発見である。

このようなシステムによる撮影された画像から人物認証を行うためには、次のようなアルゴリズムの特性が必要である。そのアルゴリズムは、画像 1 枚では他人と区別がつかない低解像度の上半身画像でも、異なる複数の視点から撮影した画像の照合結果を蓄積することで、人物認証を行うことが出来る。また、一見情報がなさそうな人物の背面画像でも、照合に用いることが出来る。さらに、人物検出アルゴリズムで背景を誤検出する可能性があるため、誤検出画像の入力に対してもロバストである必要がある。本研究ではこれらを満たすアルゴリズムとして、累積アピアランス照合アルゴリズムを提案する。それは、1 枚画像のアピアランス照合アルゴリズムを複数枚のテンプレート画像と入力画像間で繰り返し、その結果を累積させる手法である。その入力画像の順番によって照合値は変化しない。そのため、複数の視点から人物画像を取得しようとしたとき、そのための飛行ロボットの経路の自由度は高くなる。

尚、本システムでは基本的には服装が変わらないときを対象としている。

3.3.2 累積アピアランス照合アルゴリズム

3.3.2.1 1枚画像のアピアランス照合アルゴリズム

人物の頭や肩のライン (Ω 型) を個人特徴として捉え、人物認証を行う。そのため上半身画像からエッジ検出を行い、その強度や方向特徴を取得する。そのエッジ検出には、第 2 章の人物検出アルゴリズムでも用いた、指定した方向のエッジ検出を行うことが可能であるガボール特徴[21]を用いた。本手法の特徴量は、各画素における 4 方向 (0° , 45° , 90° , 135°) のエッジの方向成分とその強度の合計値である。各方向のエッジ強度の合計値を求め、その値で各方向のエッジ強度を正規化することで、各画素における方向成分を求める。これは、照明変動によりエッジの強度が変化した場合でもその方向成分は変化しにくいと考えられるためである。しかし、方向成分だけでは、エッジがない場所のノイズを捉える可能性があるため、その強度の合計値も特徴量として加えた。

飛行ロボットが取得した画像から人物の上半身の特徴を得るため、最初に画像内から人物検出を行う。図 3.3 に、飛行ロボットの位置とカメラ角度を想定し撮影した画像と、その画像か

ら人物画像 ($N \text{ pixel} \times N \text{ pixel}$) を抽出し、拡大縮小した照合用上半身画像を示す。目視確認で本人を見分けられる程度の低解像度画像の設定として、本論文では $N=46\text{pixel}$ とした。

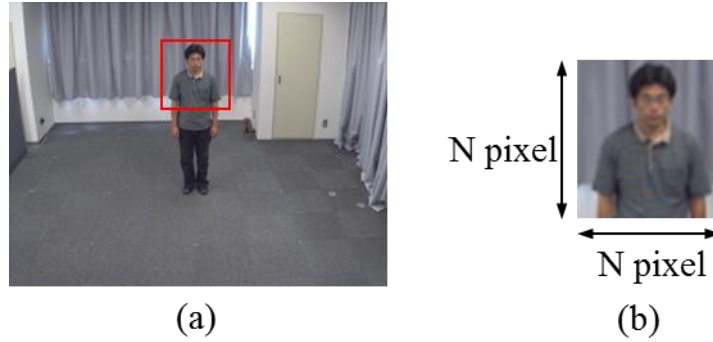


図3. 3: 人物認証システムのための入力画像：
 (a)飛行ロボットの高度と同じ位置に設置した固定カメラから撮影した画像
 (b)人物検出アルゴリズムにより抽出した認証用入力画像

最初に入力画像 I_{in} から認証に利用する上半身が映っている画像 I_{cut} を抽出する。この関数は式(3. 1)で表される。

$$I_{cut} = f_{cut}(I_{in}) \quad (3. 1)$$

抽出した画像 I_{cut} を照合に用いる特徴量に変換する。その特徴量はガボール特徴を適用した。グレイ変換した入力画像に2次元ガボールウェーブレット変換を行い、4方向 ($0^\circ, 45^\circ, 90^\circ, 135^\circ$) のエッジ強度を画素ごとに求める。各方向のエッジ強度を求める関数を $f_{gabor,i}$ ($i=0,45,90,135$), k ($k \in \{1, \dots, N \times N\}$) を画素番号の添え字とすると、取得するエッジ強度は式(3. 2)で表される。

$$g_{i,k} = f_{gabor,i}(I_{cut}, k) \quad (3. 2)$$

4方向のエッジ強度の合計値と、そのエッジ強度の合計値で正規化した4方向のエッジ特徴を組み合わせ、1画素あたり5次元の特徴量を作成する。照合に用いる各特徴量 $G_{0,k} \sim G_{4,k}$ は、式(3.3)(3.4)で表される。

$$G_{0,k} = g_{0,k} + g_{45,k} + g_{90,k} + g_{135,k} \quad (3. 3)$$

$$G_{1,k} = \frac{g_{0,k}}{G_{0,k}} \quad G_{2,k} = \frac{g_{45,k}}{G_{0,k}} \quad G_{3,k} = \frac{g_{90,k}}{G_{0,k}} \quad G_{4,k} = \frac{g_{135,k}}{G_{0,k}} \quad (3. 4)$$

テンプレート用画像と照合用画像 に対して式(3.1)～式(3.4)を適用し，照合に用いる特徴量を得る．図 3.4 にエッジ特徴量のイメージ図を示す．

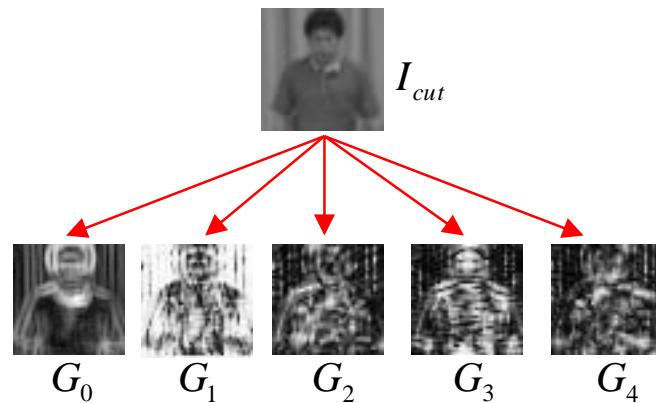


図 3.4: エッジ特徴量

得られた特徴量を用いて照合を行う．その照合手法は正規化相関法を適用した．ここで，人物の動きや姿勢の変動を吸収できるようにするため，特徴量を一定の大きさでブロック分割し，位置的なずれを吸収できるように，ブロック単位でずらしマッチングを行った．その様子を図 3.5 で示す．

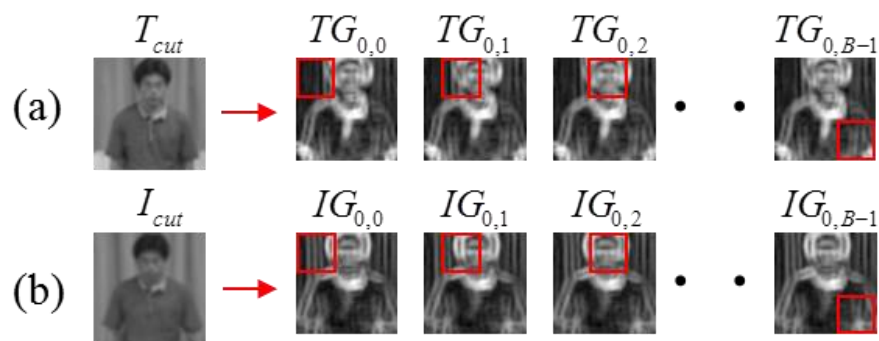


図 3.5: 特徴量ブロック：

(a) テンプレート画像の特徴量ブロック

(b) 入力画像の特徴量ブロック

この処理は 5 次元のエッジ特徴それぞれについて行い，その 5 つの結果を合計したものが，ブロック単位の照合値 S_b となる．その関数は式(3.5)で表される．

$$S_b = \frac{1}{5} \sum_{v=0}^4 \max_m [f_{cor}(TG_{v,b}, IG_{v,b,m})] \quad (3.5)$$

テンプレート特徴量 $TG_{0,b} \sim TG_{4,b}$ と入力特徴量 $IG_{0,b} \sim IG_{4,b}$ は一定の大きさ (本論文では $N/4$ pixel \times $N/4$ pixel) で重なりを持ちながらブロック分割したものである。 b はブロック番号で、 $TG_{0,b} \sim TG_{4,b}$ 、 $IG_{0,b} \sim IG_{4,b}$ は画素ごとの特徴量をブロック範囲でまとめたものである。 f_{cor} は正規化相関関数である。 ブロック分割後は、位置的に対応する各ブロック単位で正規化相関 f_{cor} を行う。 このとき入力特徴量の切り出し範囲はテンプレートと比べて範囲 m の大きさだけ大きく切り出す。 その様子を図 3. 6 に示す。 そして、範囲 m 内で位置をずらしながら照合を繰り返し、その最大値を求める。 それを各ブロックの照合値 S_b とする。

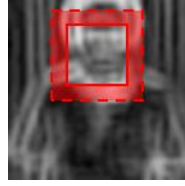


図 3. 6: 入力画像の特徴量ブロックのマージン m (赤い範囲)

最終的に求めたい画像 1 枚単位の照合値を S とすると、 S はブロック単位の照合値 S_b の平均値で式(3. 6)で表される。 B はブロック数である。

$$S = \frac{1}{B} \sum_{b=0}^{B-1} S_b \quad (3.6)$$

3.3.2.2 累積アピアランス照合アルゴリズムの概要

本アルゴリズムは、飛行ロボットが複数の異なる視点から撮影した人物画像を照合し、その結果を累積することで認証精度を高める。 その累積照合の評価値を式(3. 7)で定義した。

$$V = \sum_{i=1}^N (\alpha_i \cdot G_i) \quad (3.7)$$

V が累積照合値であり、この値で人物認証を行う。 α_i は重み係数である。 評価値 G_i ($i \in \{1, \dots, N$: N はテンプレート画像数}) は、入力画像 I_k ($k \in \{1, \dots, M$: M は入力画像数}) が得られた際の各テンプレート画像 T_i の評価値の最大値である。 図 3. 7 にテンプレート画像例 T_i と入力画像例 I_k を示す。 各テンプレート画像 T_i に対応する画像が入力されると、その評価値 G_i は高くなる。 最終的に求める累積照合値 V は評価値 G_i の合計値である。 そのため、異なる角度の入力画像が増えれば各評価値 G_i が高くなり、累積照合値 V も高くなる。 累積照合値 V が設定した閾値以上となれば、認証完了である。

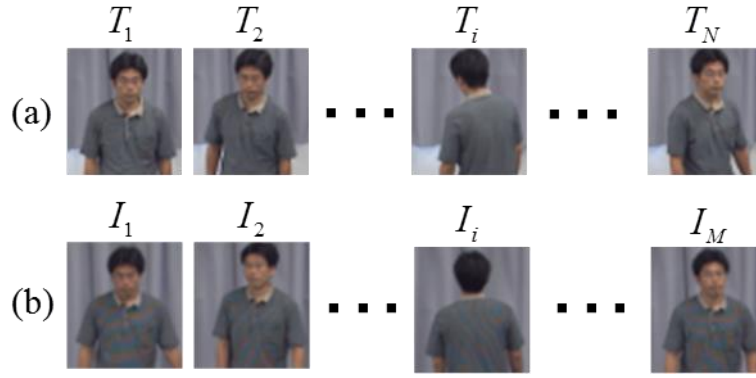


図 3.7: テンプレート画像と入力画像 (複数)

(a) テンプレート画像 : $T_i (i \in \{1, \dots, N\})$

(b) 入力画像 : $I_k (k \in \{1, \dots, M\})$

評価値 G_i は, 式(3.8)で表される.

$$G_i = \max_k (S'_{i,k}) \quad (3.8)$$

$$S'_{i,k} = \frac{S_{i,k} - \bar{m}_i}{\sigma_i} \quad (3.9)$$

評価値 G_i は各テンプレート画像 T_i と, すべての入力画像 $I_k (k \in \{1, \dots, M\})$ を照合した結果の最大値である. $S'_{i,k}$ は $S_{i,k}$ を正規化した値である. 式(3.9)の $S_{i,k}$ は1枚画像のアピランス照合結果であり, 式(3.6)の照合を T_i と I_k で行った結果である. \bar{m}_i , σ_i は, テンプレート画像 T_i と他人画像群を照合したときの平均値と標準偏差である. 他人分布はあらかじめ同環境で撮影した複数人の他人画像とテンプレート画像 T_i とを画像照合することで得られる. このように画像照合結果を正規化することで, 個々のテンプレート画像の他人との分離のしやすさを反映している.

重み係数 α_i は, 式(3.10)で表される.

$$\alpha_i = \frac{1}{\sum_{j=1}^N R_{i,j}} \quad (3.10)$$

$$R_{i,j} = f_{\text{verify}}(T_i, T_j) \quad (3.11)$$

式(3.11)の f_{verify} は1枚画像のアピランス照合関数であり, $R_{i,j}$ は同テンプレート内の画像間の評価値である. α_i はすべてのテンプレート画像 $T_j (j \in \{1, \dots, N\})$ に対する, 各テンプレート画像 T_i の類似性を表している. 同テンプレート画像群中で類似する画像が多ければ重みは小さ

くなる．逆に類似する画像が少ない，すなわち特異な画像であれば，重みは大きくなる．同じ角度のテンプレートが複数枚存在する場合，そのまま合算するアルゴリズムでは，その角度の入力画像が入力されると，無条件に累積照合値は高くなってしまうことである．重み係数 α_i を適用することでこの問題を解決した．

本アルゴリズムは，このような演算を行うことで，異なる角度の入力画像の照合値を累積させることができる．また，テンプレート画像の他人との分離しやすさと，そのテンプレート画像内での特異性を考慮している．

また，本アルゴリズムの流れであれば，入力画像の時系列パターンは全く影響しない．式(3.7)の評価値 G_i は式(3.8)で示したように，入力画像群の照合値の最大値になるため，入力画像群の順番が変わっても，同じ値が得られるからである．

3.3.2.3 累積アピアランス照合アルゴリズムの詳細説明

時系列的に入力画像が順に入力される際の，照合値算出の流れについて述べる．入力画像を1枚(I_1)を入力すると，すべてのテンプレート画像 $T_i(i \in \{1, \dots, N\})$ と入力画像 I_1 で画像照合関数 f_{verify} を行い，その結果を評価値 $G_i(i \in \{1, \dots, N\})$ に代入する．次に，2枚目の入力画像 I_2 が入ってきた場合は，同様にすべてのテンプレート画像 $T_i(i \in \{1, \dots, N\})$ と画像照合を行う．その結果を $G_i(i \in \{1, \dots, N\})$ に代入するときに，そのとき持っている G_i の値（この場合は1枚目の照合結果）と比較し，大きい値の場合のみを代入する．基本的には，これをすべての入力画像 $I_k(k \in \{1, \dots, M\})$ に対し行い，評価値 G_i を求める．その結果を重み付き合算することで，最終的な累積照合値 V が得られる．

誤検出の入力に対する対応策の1つとしては，テンプレート画像の持つスコア G_i の初期値を0とした．それ以下のスコアでは更新されない．これは私たちのアピアランス照合は，類似性が高いときはプラスになるが，少しでも異なると，マイナスになるという特徴があるからである．

誤検出の入力に対するもう1つの対策として，この評価値 G_i を更新するときに，もう1つ制限を加えている．1つの入力画像 I_k が入力されたときに，全テンプレート画像 $T_i(i \in \{1, \dots, N\})$ を照合した結果の最大値を得るテンプレート画像 T_{index_i} を求める．そのテンプレート画像 T_{index_i} と類似度が高いテンプレート画像 T_i の評価値 G_i のみを更新するという制限である． $index_i$ は式(3.12)で表される．

$$index_i = \arg \max_i (S_{i,k}) \quad (3.12)$$

テンプレート画像同士の類似性は式(3.10)で表され，その類似性の高さの閾値を T_{self} と定義する． $R_{index_i,i}$ がこの値 T_{self} より大きいテンプレート画像 T_i に対応する評価値 G_i のみを更新する．2つのテンプレート画像間の評価値が低く，入力画像とこの2つのテンプレート画像との評価値が

どちらも高いことは、基本的にはありえないためである。その場合はイレギュラーな画像、例えば誤検出した画像が入力されたものとして、その影響を少なくするためにこの制限を加えている。誤検出した入力画像は大抵の場合は、画像照合結果が0以下となりその場合は評価値 G_i の初期値が0なので影響はない。しかし、誤検出した入力画像の模様によっては、各テンプレート画像に対する評価値 $S_{i,k}$ ($i \in \{1, \dots, N\}$)が全体的に低い正の値になることがある。この制限がないと、評価値 G_i がこの値ですべて更新され、式(3.7)により全体的に低い値を合算することになり、少数枚の人物入力画像に対する合算値との差が小さくなってしまう。累積アピアランス照合アルゴリズムのアルゴリズムフローを表3.1に示す。

表3.1: 累積アピアランス照合アルゴリズムフロー

<ul style="list-style-type: none"> • Given example images T_i ($i \in \{1, \dots, N\}$) for template, I_k ($k \in \{1, \dots, M\}$) for input • Initialize score $G_i = 0$ ($i \in \{1, \dots, N\}$) • Determine in advance the value of $R_{i,j}$, \bar{m}_i, σ_i • For $k = 1, \dots, M$: <ul style="list-style-type: none"> For $i = 1, \dots, N$: <ol style="list-style-type: none"> 1. Calculate score $S'_{i,k}$ $S'_{i,k} = \frac{S_{i,k} - \bar{m}_i}{\sigma_i}$ end 2. Find index for maximum of $S'_{i,k}$ $index_i = \arg \max_i (S'_{i,k})$ For $i = 1, \dots, N$: <ol style="list-style-type: none"> 3. Update score G_i $if (S'_{i,k} > G_i \text{ and } R_{index_i,i} > T_{self})$ $G_i = S'_{i,k}$ end end • The final cumulative score V is: $V = \sum_{i=1}^N (\alpha_i \cdot G_i) \quad , \quad \alpha_i = \frac{1}{\sum_{j=1}^N R_{i,j}}$

3.3.2.4 さまざまな入力画像パターンに対する照合値変化

作成したアルゴリズムに対し，異なるパターンで入力画像を入力した場合に，照合結果がどのように変化するかを検証した．同じような変化の少ない画像が複数枚入力されたときや，異なる角度の入力画像が入力されたとき，他人の画像が入力されたとき，誤検出した画像が含まれたときの照合結果を求めた．これにより，これまでに述べたような意図した照合値変化が起るかを確認した．式(3.7)の $\alpha_i \cdot G_i$ のヒストグラムと共に示す．図3.8に，用いたテンプレート画像を示す．本実験で用いたテンプレート画像は，人物の周囲1回転の画像で，合計28枚である．

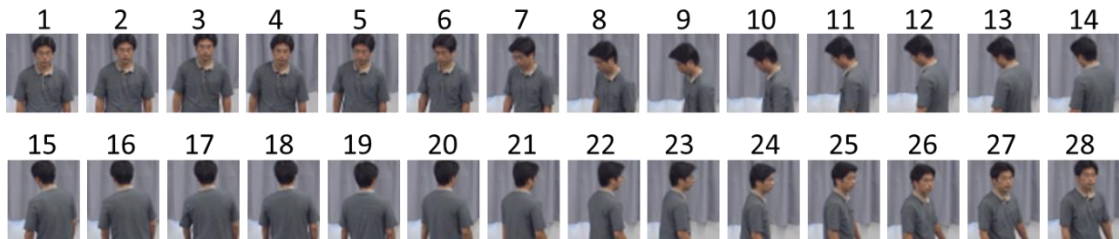


図3.8: テンプレート画像

まず第1に，(a)：比較のために本人の入力画像1枚のときのヒストグラムを求めた．(b)：次に(a)と比較して変化の少ない本人の入力画像を1枚加えた．(c)：さらに同様の画像を1枚加え，結果を求めた．(d)：(a)に対して異なる角度の本人の入力画像を1枚加えた．(e)：さらに異なる角度の入力画像を1枚加えた．(f)：すべての本人の入力画像（全29枚）を入力した．(g)：他人の入力画像（全48枚）を入力したときの結果を求めた．(h)：比較のために，テンプレート画像を入力した場合の結果を求めた．(i)：(c)に対して，誤検出した画像3枚を加えたときの結果を求めた．(a)-(i)の結果を図3.7～図3.17に示す．

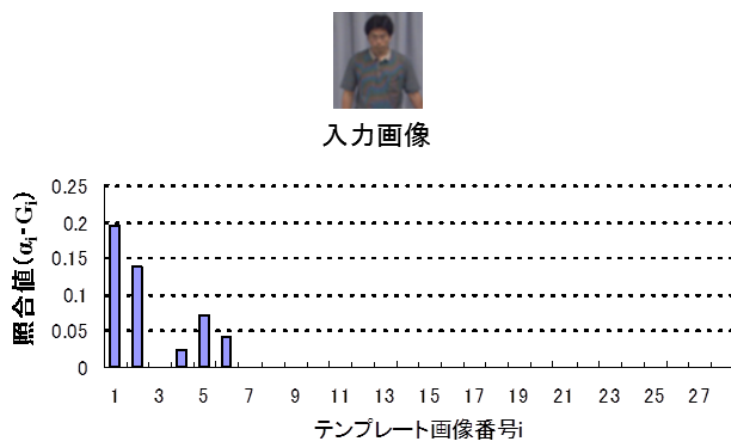


図3.9: (a):本人の入力画像が1枚入力されたときの結果 ($V = 0.46$)

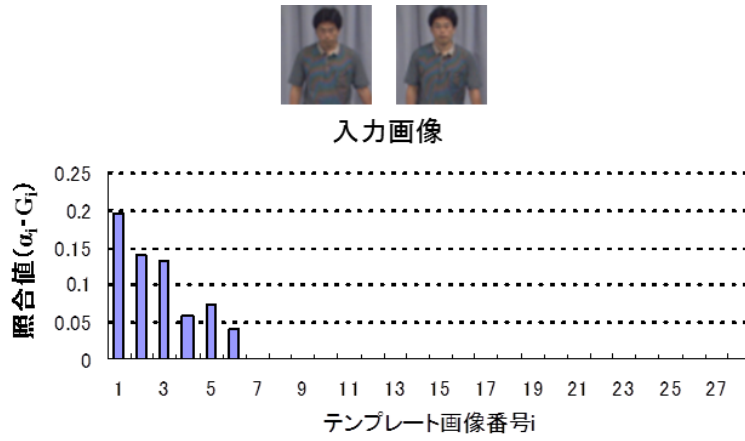


図3.10: (b):変化の少ない本人の入力画像が2枚入力されたときの結果 ($V = 0.63$)

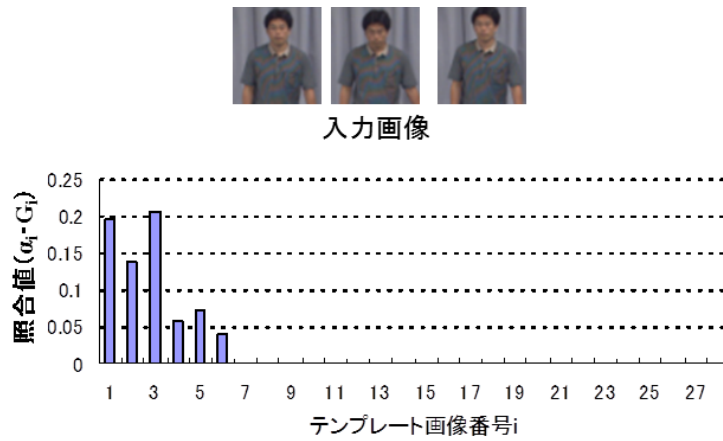


図3.11: (c):変化の少ない本人の入力画像が3枚入力されたときの結果 ($V = 0.71$)

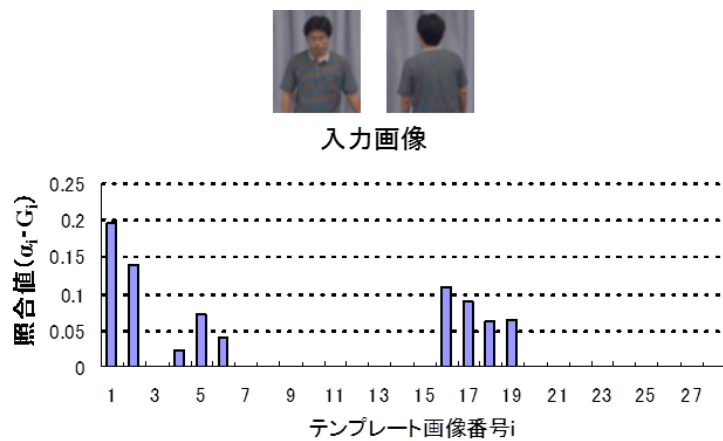


図3.12: (d):異なる角度の本人の入力画像が2枚入力されたときの結果 ($V = 0.79$)

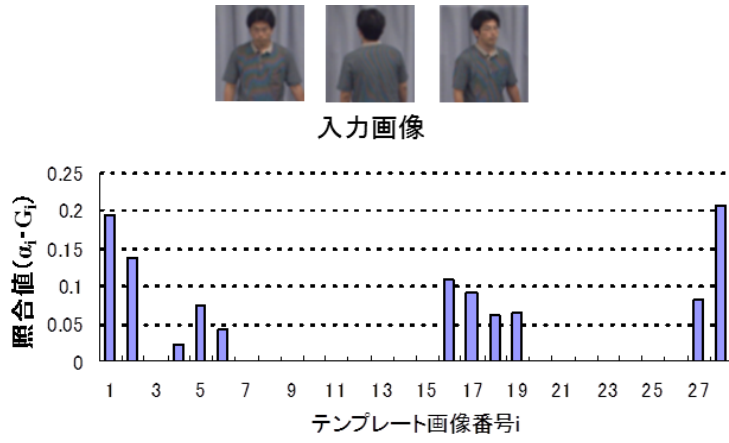


図3.13: (e):異なる角度の本人の入力画像が3枚入力されたときの結果 ($V = 1.08$)

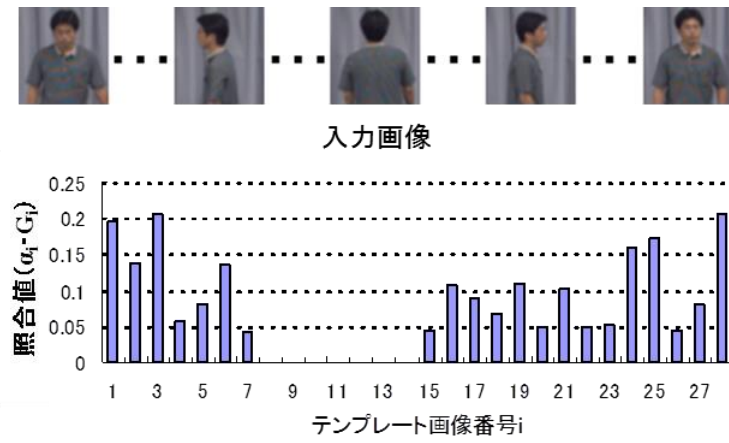


図3.14: (f):すべての本人の入力画像(29枚)が入力されたときの結果 ($V = 2.18$)

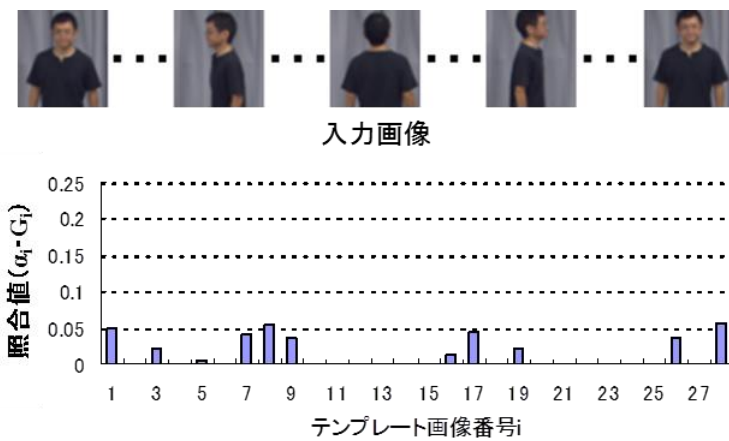


図3.15: (g):他人の入力画像(48枚)が入力されたときの結果 ($V = 0.37$)

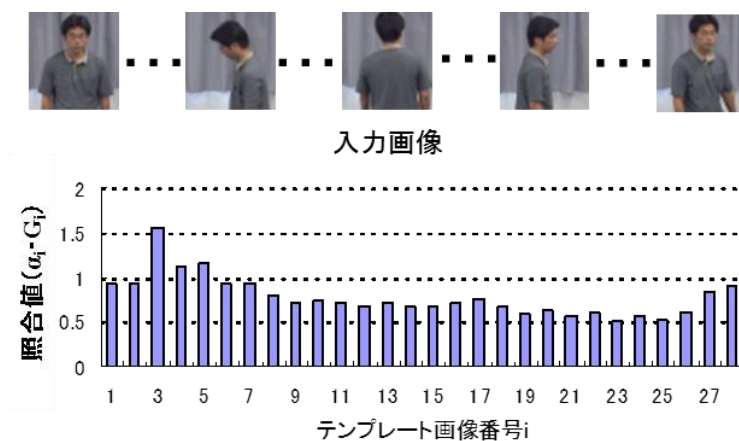


図3.16: (h):すべてのテンプレート画像(28枚)が入力されたときの結果 ($V=21.71$)

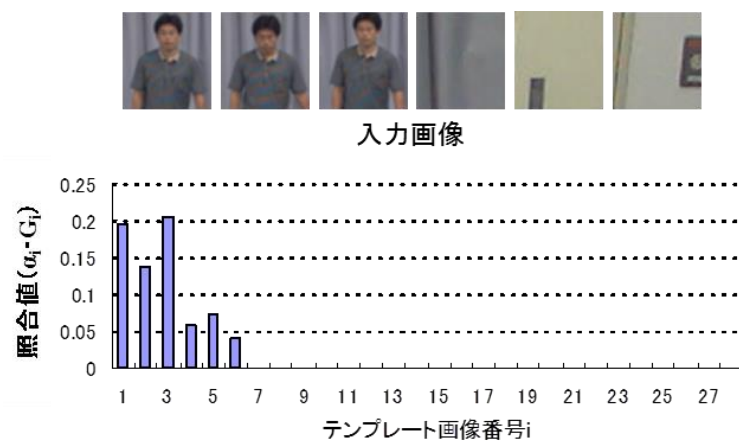


図3.17: (i):誤検出画像が3枚含んだときの結果 ($V=0.71$)

図3.9～図3.11より、変化の少ない入力画像を入力すると照合値は上昇するが、図3.12, 図3.13より、異なる角度の画像を入力したほうが、より大きく照合値は高くなることが分かった。変化の小さい入力画像を複数枚入力するよりも、同じ枚数入力するなら変化の大きい画像を入力するほうが、より照合値が大きくなるのが確認できた。これは、一般的に人が写真を見て見分けるときと同じように、複数の角度からの情報を得ることで情報量は増え、より見分けやすくなったからであると考えられる。図3.14, 図3.15から分かるように、他人を入力すると、角度によっては照合値が出力されているが、本人との差は大きい。また図3.17から分かるように、誤検出が入力されたとしても、照合値には変化はない。これらのことから、本アルゴリズムは、異なる角度の画像が入力されると照合値が上昇し、誤検出画像の入力に対して照合値は上昇しないことが確認できた。

3.4 累積アピランス照合アルゴリズムの基本性能評価

3.4.1 累積アピランス照合アルゴリズムの評価方法概要

アピランス認証の精度は、入力画像に写る人物の大きさや角度、解像度、ブレ等によって変わる。飛行ロボットのブレ等によって生じる誤差を無視した、アピランス認証方法の性能の確認と、精度の評価を試みる。

飛行機ロボットが飛行する高度に固定したカメラで撮影した映像を用いて、アルゴリズムの性能を評価した。こうすることで、飛行ロボットの飛行が認証に与える影響を無視した評価を行うことが出来る。カラーカメラ (AXIS 233M) を高さ3(m)、俯角15度に設置し、屋内環境で人物画像の撮影を行い、評価データを作成した。実験対象者は、合計38人 (男性35人、女性3人) である。1人あたり周囲1回転の画像を撮影したが、撮影枚数や角度幅については、人物ごとに異なる。(30~60枚) テンプレート用画像と入力用画像を2回に分けて撮影したが、そのとき服装変化はない。

服装の種類の違いによる照合率の変化を確認するため、上記撮影条件で、夏服と冬服、制服の3セット撮影した。撮影したサンプル画像を図3. 18に示す。図3. 19に38人の後姿のサンプルを示す。このときのテンプレート画像と入力画像の解像度は 46×46 pixelである。



図3.18: 3人の実験対象者のサンプルデータ: (a)夏服 (b)冬服 (c)制服

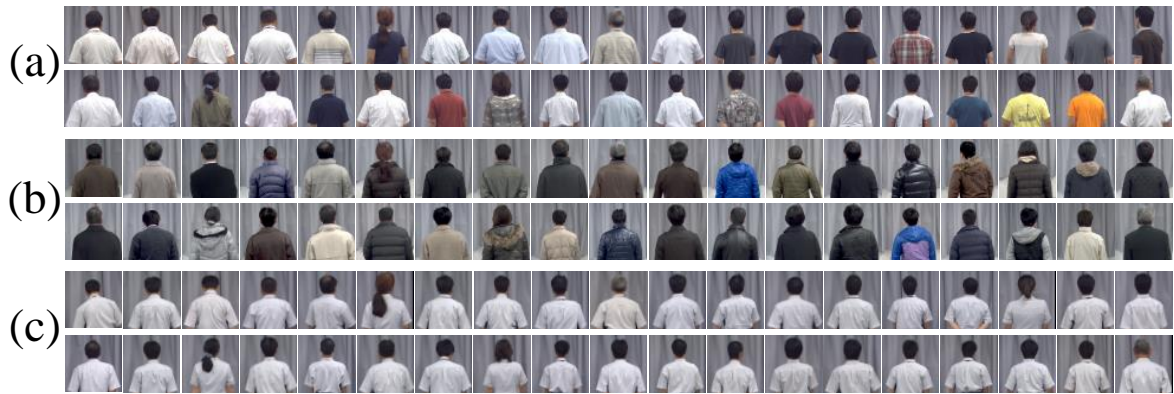


図3. 19: 38人の実験対象者の後ろ姿: (a)夏服 (b)冬服 (c)制服

3.4.2 服装の違いによる認証精度の評価

本研究では、私服を着用し照合時にテンプレート画像と入力画像で服装が変わらないシーンを想定している。ここでは、実験対象者全員が異なる服装を着ている場合と、スーツの様に同じ服装を着ている場合を想定して評価を行った。スーツのように同じ服装を着ている場合でも本アルゴリズムが適用可能であるか調査した。

本評価は、図 3. 18 に示す飛行ロボットを想定した高さで撮影した 38 人の画像を用いて累積アピランス認証の照合結果を計算した。図 3. 20 に夏の私服、冬の私服、制服(仕事着)の照合結果を示す。本結果は、FRR(False Reject Rate), FAR(False Accept Rate)のグラフである。

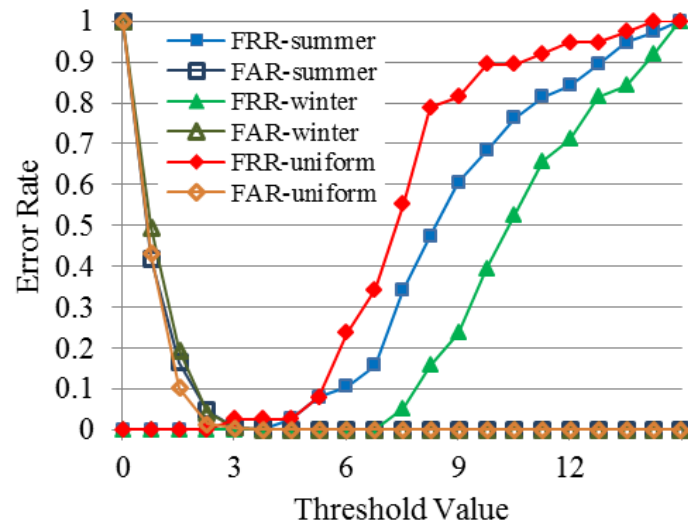


図3. 20: テンプレート画像と入力画像の服装が同じ場合の照合結果 (夏私服, 冬私服, 制服)

図 3.20 より、3 種類の服装の EER(Equal Error Rate)は 0 で、100%本人と他人を分離することが出来た。各 FRR は、閾値の値によって差はあるが、閾値の値が 8 のときには、制服の FRR は約 0.8、夏の私服の FRR は約 0.5、冬の私服の FRR は約 0.2 であり、各服装の種類で約 30% ずつ差が出た。これは服の模様が照合値に影響を与えていると考えられる。制服は全員同じ模様なので、体型（頭、肩のライン）による差が識別結果として表れている。夏の私服はTシャツやカッターシャツなど似ている服を着た人が複数人存在する。冬の私服はコートやジャンパーなど人物ごとに異なる。その差が結果に表れたと考えられる。この結果より、服装の模様が照合値に影響を与えることも確認でき、且つ、服装の模様が同じであっても、本人と他人を分離できることを確認した。

3.4.3 テンプレート画像と入力画像の服装が異なる場合の認証精度の評価

本アルゴリズムは服装が変わらない前提で作成したものであるが、手法の限界についても評価するため、テンプレート画像と入力画像の服装が異なる場合についても評価を行った。

本評価は、3.4.2 で用いた画像と同じ画像を用いて行った。図 3.21 に 3 パターンの服装が異なる場合の照合結果を示す。その 3 パターンは、テンプレート：制服、入力：夏の私服、テンプレート：制服、入力：冬の私服、テンプレート：夏の私服、入力：冬の私服である。

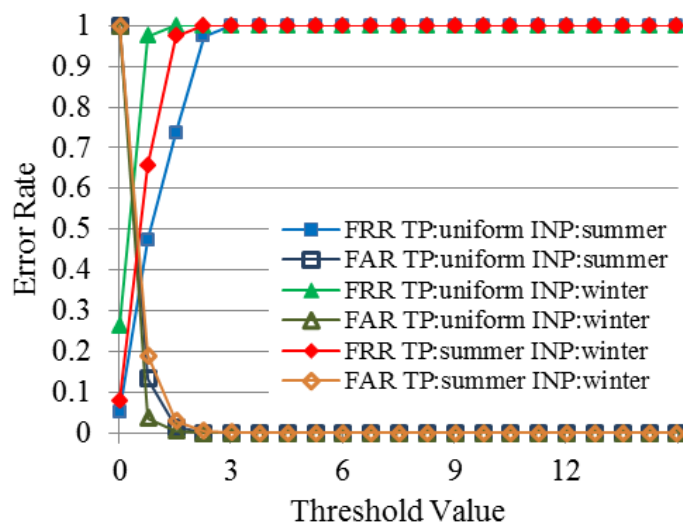


図3.21: テンプレート画像と入力画像の服装が異なる場合の照合結果
 (テンプレート画像:制服 入力画像:夏私服
 テンプレート画像:制服 入力画像:冬私服
 テンプレート画像:夏私服 入力画像:冬私服)

図 3.21 より，異なる服装の照合結果は，本人照合値が大きく落ち，識別困難であることが分かった．3 パターンの中では，制服と夏服の結果が最も精度が高かった．今回撮影した制服は夏に着る制服であり，薄着なので体のラインが出やすい．夏の私服も薄着であり，体のラインが出やすいため，制服と夏の私服による体のラインの一致が照合値を高めることになったと考えられる．このことから，服の模様だけでなく，体のラインが照合値に寄与していることが分かる．逆に夏服と冬服のように，服の模様と体のラインが違う場合は識別できない結果となった．

3.4.4 解像度の違いによる評価

飛行ロボットによって撮影される画像は低解像度が想定される．本節では入力画像の解像度を変化させて照合結果を比較した．アルゴリズムの比較のために，総当たり max 照合アルゴリズム，総当たり平均アルゴリズムについても同様に結果を取得した．解像度は縮小なし(46×46pixel)と，低解像度での評価結果を得るために3つの倍率(50%，33%，25%)で縮小し，照合実験を行った．図 3.22 にその各倍率画像のサンプルを示す．各倍率の入力画像は，線形補間による拡大アルゴリズムで拡大した後，照合を行った．図 3.23，図 3.24，図 3.25 に，夏私服，冬私服，制服の照合実験結果を示す．

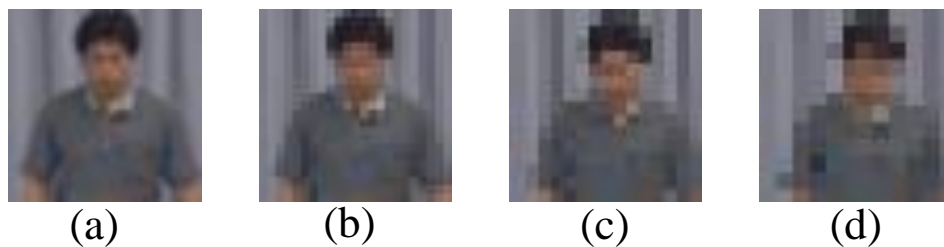


図3.22: 異なるサイズの入力画像例

(a) 46×46 pixels(100%) (b) 23×23 pixels(50%)

(a) 15×15 pixels(33%) (b) 12×12 pixels(25%)

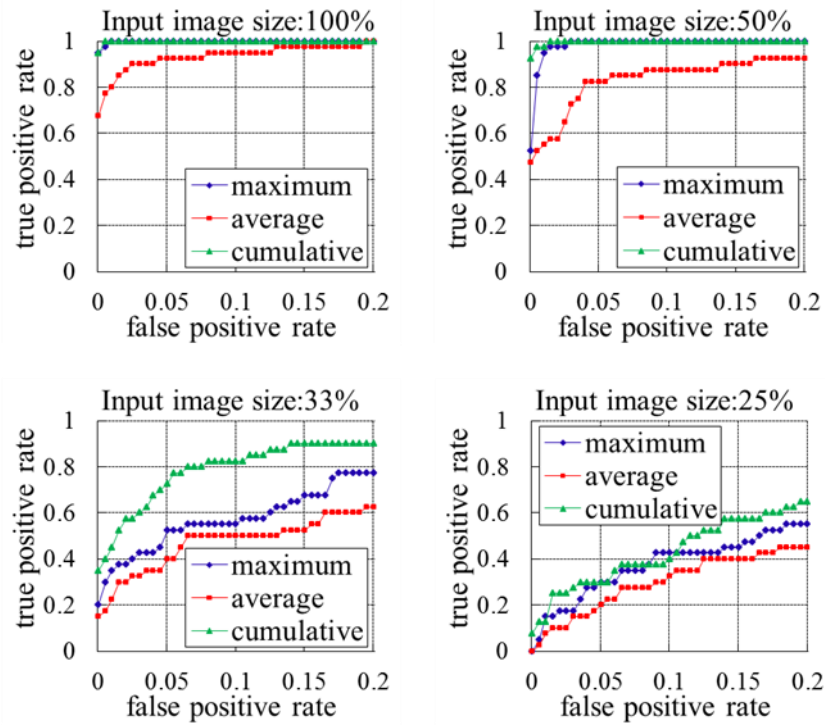


図3. 23: 夏私服-人物認証結果 (ROCカーブ)

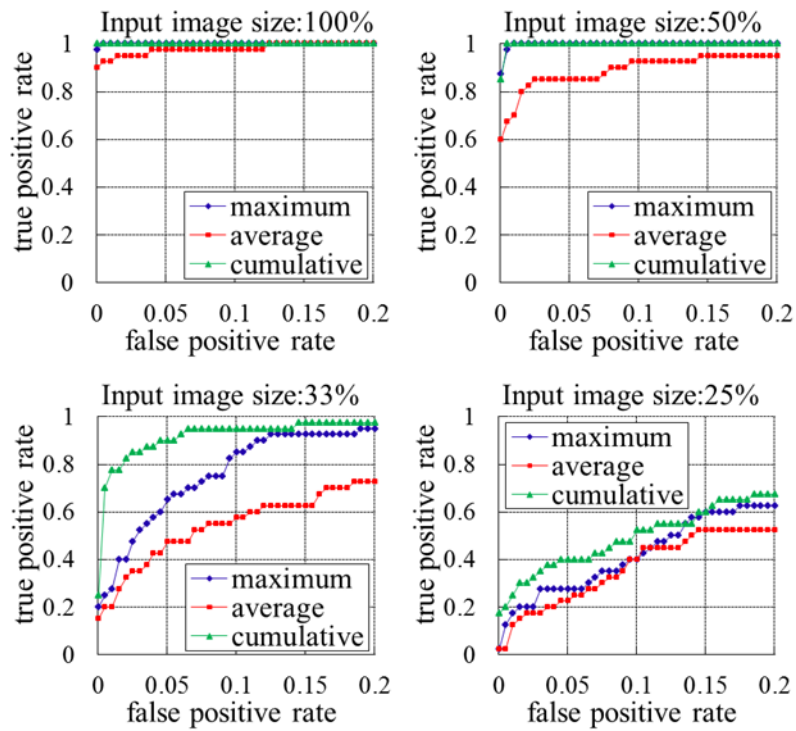


図3. 24: 冬私服-人物認証結果 (ROCカーブ)

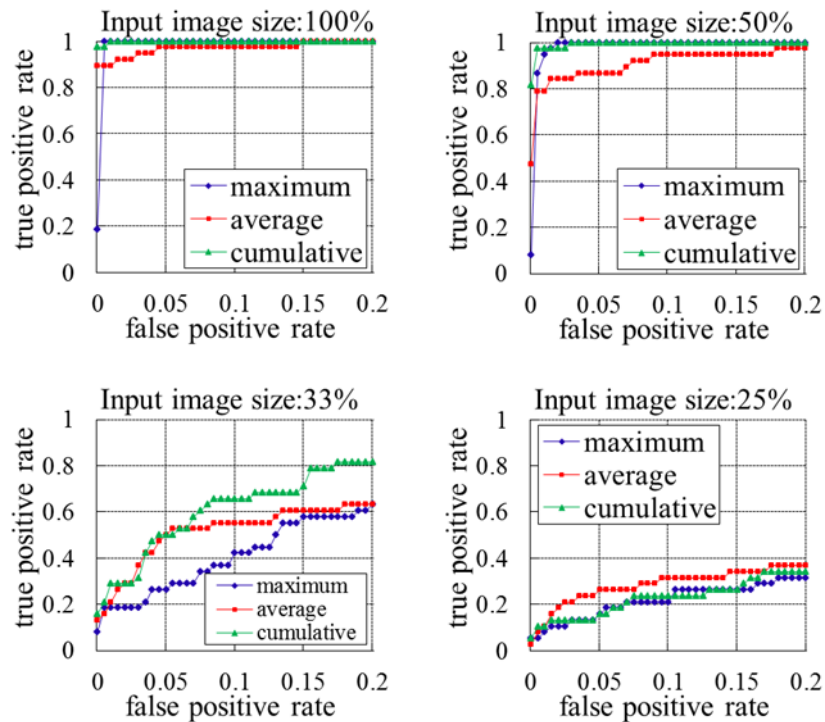


図3.25: 制服-人物認証結果 (ROCカーブ)

図3.23より、本手法は、倍率100%と50%のときは、ほぼ100%識別することができ、倍率が33%、25%では識別精度が低下した。しかし、図3.22 (c)のように、画像1枚では目視では識別しにくい倍率33%の画像でも、図3.23、図3.24より、誤認識率が0.05のときに夏と冬の私服で平均約85%認証率であった。このことから、提案手法は、画像1枚では目視で識別しにくい低解像度の画像でも、複数枚、複数角度の画像の情報を重ねることで、識別可能であるといえる。また他の手法と比較すると、50%、33%、25%と低解像度になるに従って、提案手法の精度が良いことが確認出来た。制服の結果(図3.25)でも、100%、50%、33%において、他の手法よりも精度は高いことが確認できた。

3.4.5 人物の角度バリエーションの変化による評価結果

飛行ロボットから得られる入力画像に写っている人物の角度と枚数のパターンは、飛行ロボットの軌道と人物の動きによって、さまざまである。人物角度のバリエーションと画像枚数の違いによる照合精度を比較する。3.2.2.4では、異なる角度の入力画像が増えると照合値が高くなることは確認したが、本節では認証精度の変化を確認する。第1に比較のために、入力画像1枚(0°)のときの結果を求めた。第2に同じ角度の入力画像が複数枚あるパターンとして、入力画像5枚(0°近辺)のときの結果を求めた。第3に角度バリエーションが2つ増えたときとして、入力画像3枚(0°, 180°, -45°)の結果を求めた。最後に、さらに角度バリエー

ションを2枚増やして、入力画像5枚 (0° , 180° , -45° , 45° , 90°) のときの結果を求めた。これらの全4パターンは実験対象者40人分作成し、累積アピラランス照合アルゴリズムに適用した。その4パターンの入力画像図のサンプルを図3.26に示す。その照合結果を図3.27に示す。

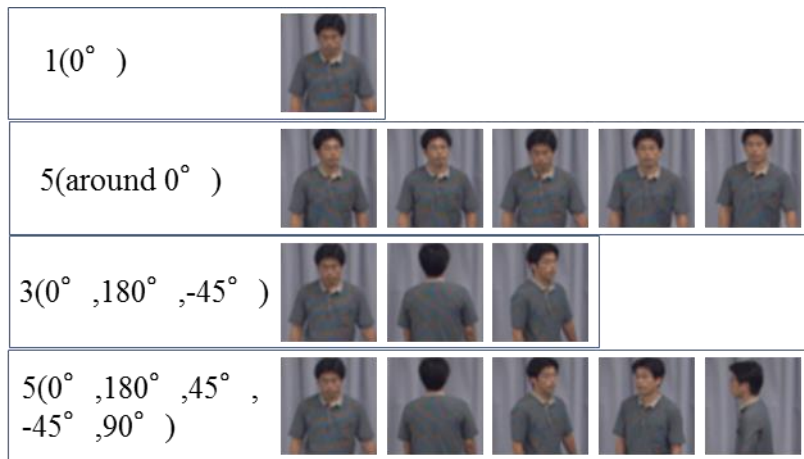


図3.26: 4パターンの入力画像図

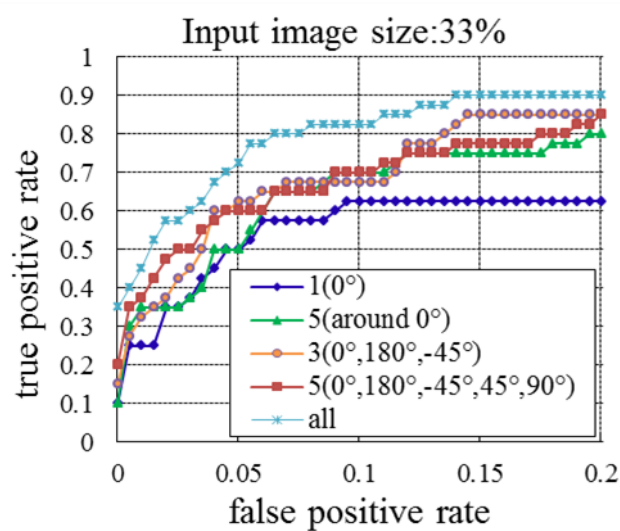


図3.27: 4パターンの照合結果 (ROCカーブ)

図3.27より、1枚の画像を入力するよりも、同じような角度でも複数枚入力するほうが精度は上がることが分かる。これは、1つの角度のテンプレートに対し、複数枚の同じような入力画像の中からより似ているものが選択され照合値が上がり、その結果照合精度が上がっているからである。しかし、同じような角度を5枚入力するよりも、異なる角度を3枚入力するほうが精度は高い。また異なる角度を5枚に増やすと、さらに精度が高くなることもわかる。これは入力画像間で類似度の低い画像を入力すればするほど、より精度は向上しやすいことを示

している。この結果は飛行ロボットの軌道を考えると都合がよい。複数の認証する人物間の移動のための軌道と、異なる角度の画像を取得するための軌道を組み合わせることで、効率よく探索を行うことができる。

3.4.6 誤検出画像の入力の影響

飛行ロボットが取得した画像から、人物検出アルゴリズムにより人物部を抽出する。そのとき、人物検出アルゴリズムの精度によっては誤検出した画像が入力されることが考えられる。誤検出画像のサンプルイメージを図3. 28に示す。誤検出した画像を入力したときに、照合値に大きな影響が出るようであれば、人物検出アルゴリズムも精度向上の工夫をする必要が出てくる。そのため、どの程度照合値に影響が出るかを調べた。実験対象者40人の異なる5つの角度の入力画像に対し、それぞれ10枚のランダムな誤検出画像を追加し、累積アピアランス照合結果を求めた。その照合結果グラフを図3. 29に示す。



図3. 28: 10枚のランダムな誤検出画像を含んだ入力画像例



図3. 29: 10枚のランダムな誤検出画像を含んだ入力画像の照合結果 (ROCカーブ)

図3. 29より、誤検出画像を入力しても精度はほとんど変化しないことが分かる。人物画像5枚に対し誤検出画像を10枚で、誤検出画像のほうが2倍多いが、1~2%の照合率の低下に留まっている。このことから誤検出画像に対してロバストであるといえる。誤検出画像の影響を受けにくい照合アルゴリズムなので、人物検出アルゴリズムは、誤検出画像が多少増えても人物画像をリジェクトしないように閾値をシフトできる。人物画像が増えることで照合率も向上する。

3.4.7 基本性能評価のまとめと考察

固定カメラを用いて、飛行ロボットが撮影すると想定できる入力画像を作成し、その画像に対して人物認証アルゴリズムを適用し精度評価を行った。

服装については、テンプレートと入力と同じ服装であれば、例え制服のように複数人が同じ服装を着ていても、EER(Equal Error Rate)は0で、100%本人と他人を分離することが出来た。テンプレートと入力異なる服装の場合は、最も精度がよい制服と夏服の照合結果でもEERは40%であり、識別できないことも確認した。

解像度については、1枚では目視でも識別しにくい約15×15画素で写った人物の上半身の低解像度画像でも異なる角度の入力画像を複数累積させることで、私服では平均85%の認証率(誤認率5%)を得ることが出来た。飛行ロボットによる人物搜索のような、搜索対象候補を挙げ、最終的に人間が画像等を確認するシステムに適用することは、十分可能である。

また、入力される人物画像の角度バリエーションが多いほど、照合精度が高くなることも分かった。これは認証を効率的、且つ高精度に行うための飛行ロボットの軌道計画を作成するときに、重要なポイントとなる。

誤検出画像の入力に対しては、1~2%の認証率の低下に留まった。飛行ロボットはあらゆる場所を飛行しあらゆる背景画像を取得することから、固定監視カメラと比較すると、誤検出画像は増えると考えられる。そのため、本認証アルゴリズムが誤検出画像に対してロバストであることは、大きなメリットであるといえる。さらに、処理時間については考慮する必要があるが、認証率を上げるために、人物検出の閾値を下げ、本人画像の誤リジェクトを減らすことも可能である。このとき、同時に誤検出画像が増えるが、認証への影響度合いは少ない。

本アルゴリズムは、全員が同じ服装(制服)を着ていても38人を識別することが出来た。その識別を可能とした要因について考察を行う。図3. 30に示したテンプレート画像に対して、3人の人物の入力画像を本認証アルゴリズムに適用した結果の式(3. 7)の $\alpha_i \cdot G_i$ のヒストグラムを図3. 31示す。その3人は、本人と、他人の中で最も照合値が高かった人物、無作為に選んだ照合値が低い人物である。

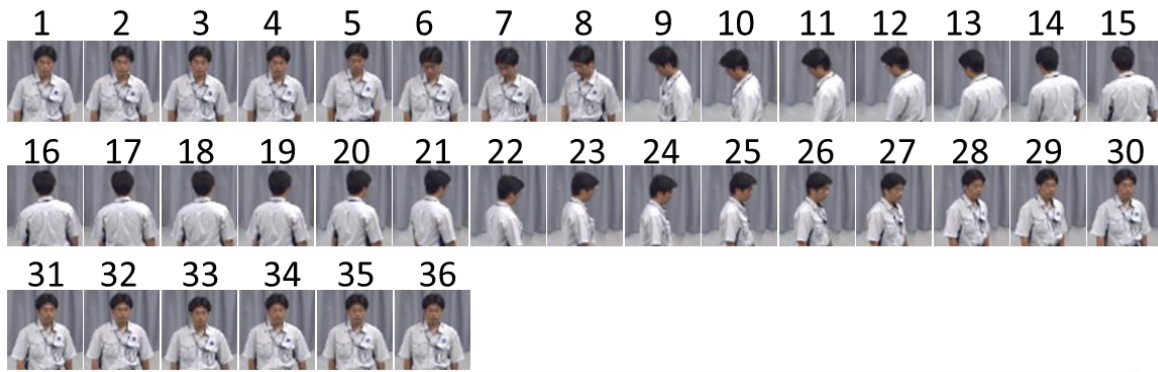


図3.30: 制服テンプレート画像例

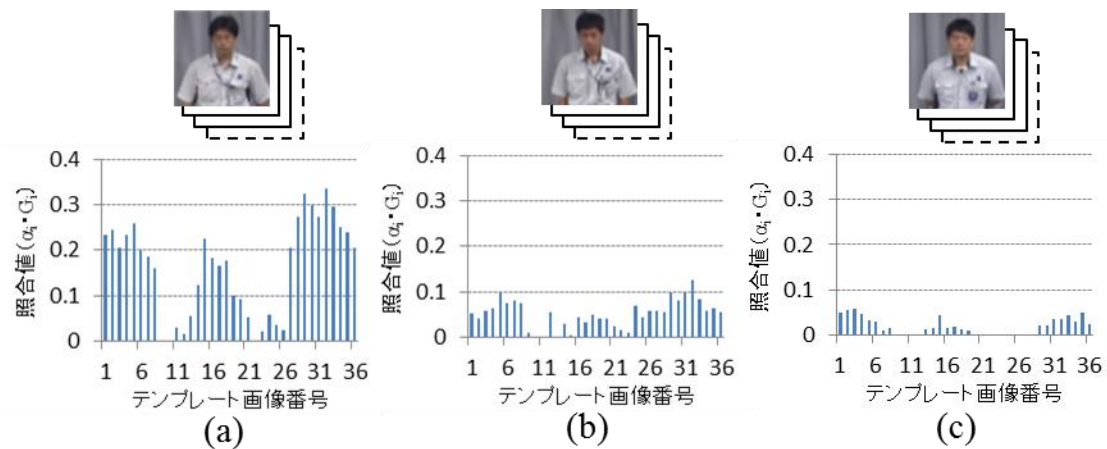


図3.31: 3人の入力画像例と照合結果 (ヒストグラム)
 (a)本人, (b)他人の中で最も照合値が高かった人物, (c)他人の代表例

図3.31の(a)と(b), (c)を比較すると, ヒストグラムは全体的に小さくなっており, 他人画像が入力された場合は, ある特定のテンプレート画像の照合値が低下しているわけではなく, すべてのテンプレート画像に対応する照合値が平均的に低下していることが分かった. よって, 本アルゴリズムを用いた場合は, 特定の角度で本人と他人を分離しているわけではないことが分かった.

さらに, 各テンプレート画像において, 画像1枚間の照合でどのように差が表れているのかを調べた. 本人入力画像と他人入力画像で最も照合値に差があったのが, テンプレート番号30であった. テンプレート番号30に対して, 各人物でヒットした入力画像 (最も照合値が高かった入力画像) を求め, 1枚画像照合の各ブロックにおける照合値を画像化した照合値分布画像を求めた. さらに本人の照合値分布画像との差を求めた. これにより, 画像内のどの部分が本人と差が出ているのかが分かる. 図3.32にその流れを示す. 尚, 照合値分布の格子の

サイズと実際に照合しているエリアは一致していない。照合エリアは3.2.2.1で述べたように重なりを持ちながらブロック分割した。図3. 32のように、照合値分布の左上の格子（水色実線）の照合エリアは水色点線で示したエリアである。

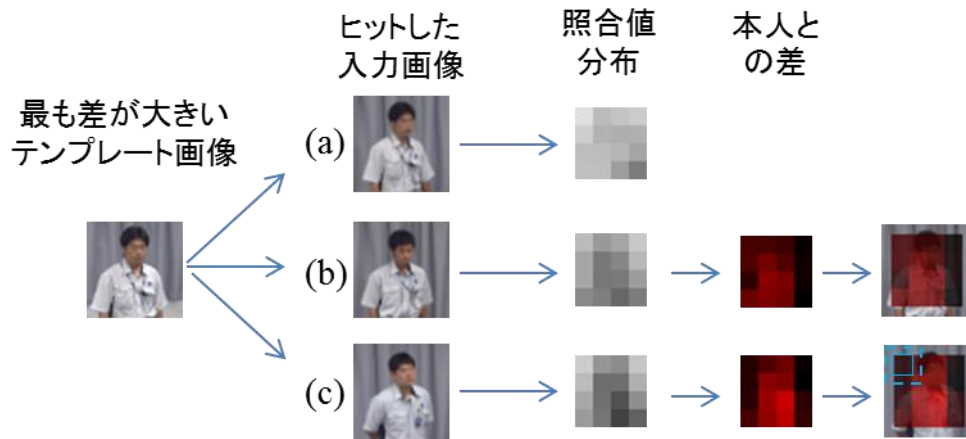


図3. 32: 1枚画像照合の分布分析

図3. 32(b)の本人との差を見てみると、特定の部位に目立って大きく差があるわけではなく、頭部から肩も含めて全体的に差がでている。その中で最も差があるのは胸の部分である。服装は同じなのでしわの出方が異なることが考えられる。同じ服装でも体型による差が表れたものと考えられる。図3. 32(c)では、左胸部分が大きく異なっていることが分かる。テンプレート画像や(a), (b)では、カードキーを胸ポケットに入れているが、(c)は、胸にぶら下げている。画像としては数画素の差であるが、その差が検出されていると考えられる。

本アルゴリズムは、さまざまな角度の人物画像において、同じ服装でも体形による写り方の差や、付属物などのわずかな違いを捉え、それらのわずかな差を累積して識別していると考えられる。

3.5 飛行ロボットが撮影した画像からの人物検索実験

3.5.1 飛行ロボットが撮影した画像からの人物検索実験概要

飛行ロボットが実際に飛行して撮影した映像から人物検索を行った。飛行ロボットの動きに加えて、人物の動きと向きが認証の結果に影響を与えると考えられる。よって、第1に、人物が静止した状態で飛行ロボットを飛行させて撮影し、認証実験を行った。第2に人の向きを変えて同様の実験を行った。第1、第2の実験結果を集計し、人物検索するための閾値を設定した。そして、最後に、人が歩いている状態で飛行ロボットが飛行し撮影した映像に対して認証実験を行った。尚、本実験では6人の人間と、20 m × 40 m 程度の広さで行った。飛行ロボットは約10~15秒飛行した間に人物の撮影を行った。

本実験の人物抽出については、目視確認を行い、画像1枚ごとに手動で行った。また人物ID付与についても手動で行った。そのため、以後述べる各シーンにおいて、各人物ごとの画像は誤検出画像は含まれず、全フレームの画像がまとまっている。人物検出アルゴリズムを適用した結果を用いたときの一連の流れによる精度検証については、第4章で述べる。本章では、人物検出アルゴリズムによる影響がない場合の人物認証精度について検証する。

撮影環境は屋内だが、窓があり直射日光が当たる環境で、場所によって照明変動がある。図3.33にその環境にて飛行ロボットによる撮影実験をしている様子と撮影に用いた飛行ロボットを示す。

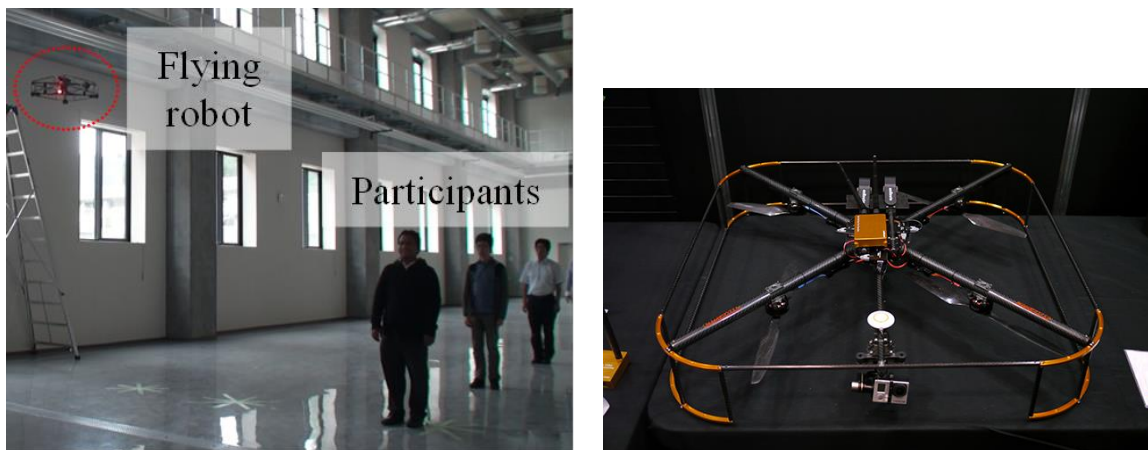


図3.33: 飛行ロボットによる撮影実験の様子と撮影に用いた飛行ロボット (ZionPG560)

用いた飛行ロボットは enRoute 社製 Zion PG 560 であり、飛行ロボットに搭載したカメラは GoProHero3 である。以下に述べる実験では、飛行ロボットの操縦は手動で行い、出来るだけ

風の影響を少なくするため、屋内で撮影を行った。実験対象者は6人(ID1-ID6)である。撮影は4FPSで行った。また、本実験での飛行ロボットの高度は約3(m)である。

また、6人の実験対象者のテンプレート画像を作成した。そのサンプル画像(ID3)を図3.34に示す。このテンプレート画像は、飛行ロボットに搭載したカメラと同じカメラを高さ3(m)の位置に固定した状態で撮影した。水平角は人物の周囲8方向(45°ずつ)であり、人物の仰角は、15°、30°、50°の3パターンである。(合計24画像)

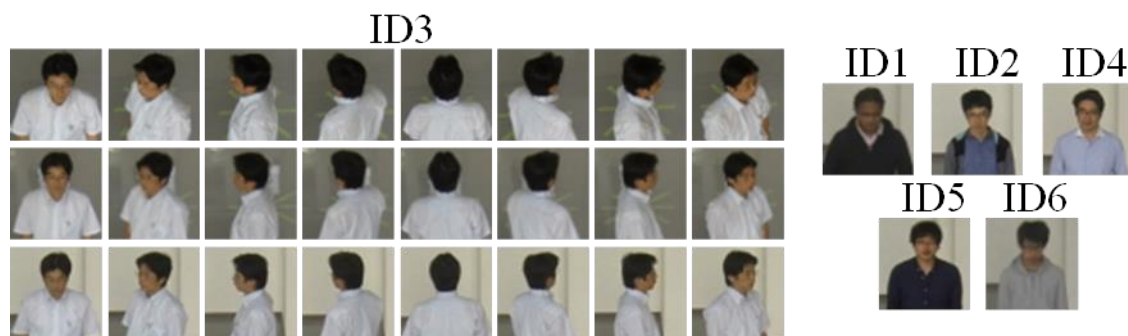


図3.34: ID3のテンプレート画像(8×3角度)とID1~ID6のテンプレート画像(1角度)

3.5.2 静止した複数人物を飛行ロボットで撮影した場合の認証

6人の実験対象者の中から1人の人物を捜索する実験を行なった。認証対象者6人は2.5m間隔で直線状に並び、直立静止した。飛行ロボットはその直線に対して撮影する人物の仰角ごとに異なる一定の距離D(m)を保ちながら平行に移動した。飛行ロボットに搭載したカメラの向きは、その直線に対して垂直方向であり、人物を撮影できる向きとした。その模式図を図3.35に示す。また、そのとき撮影された画像のサンプルと人物抽出結果を図3.36に示す。

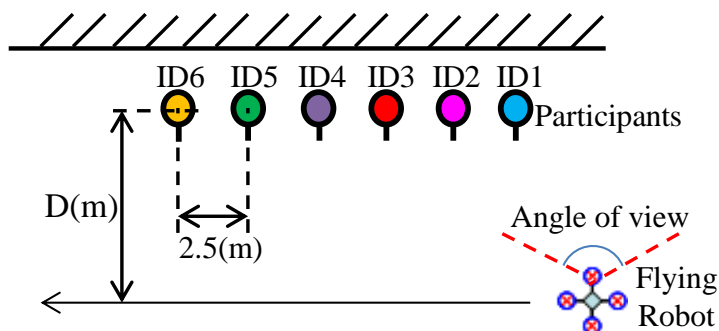


図3.35: 飛行ロボットの動きと人物位置模式図

このように、撮影された画像の中から人物抽出を行い、人物認証を撮影フレームごとに行っていく。その撮影データは、移動しながら撮影しているため、画像ブレが大きい。また、カメラと人物との距離も変化するため、人物がカメラ中央に位置した時は、上半身で 50 画素程度の解像度はあるが、中央から外れるにつれて小さくなり、ピントも合わなくなる。飛行ロボットの高度も $\pm 0.5(\text{m})$ 程度は上下するため、写っている人物画像の仰角方向の角度も変化している。

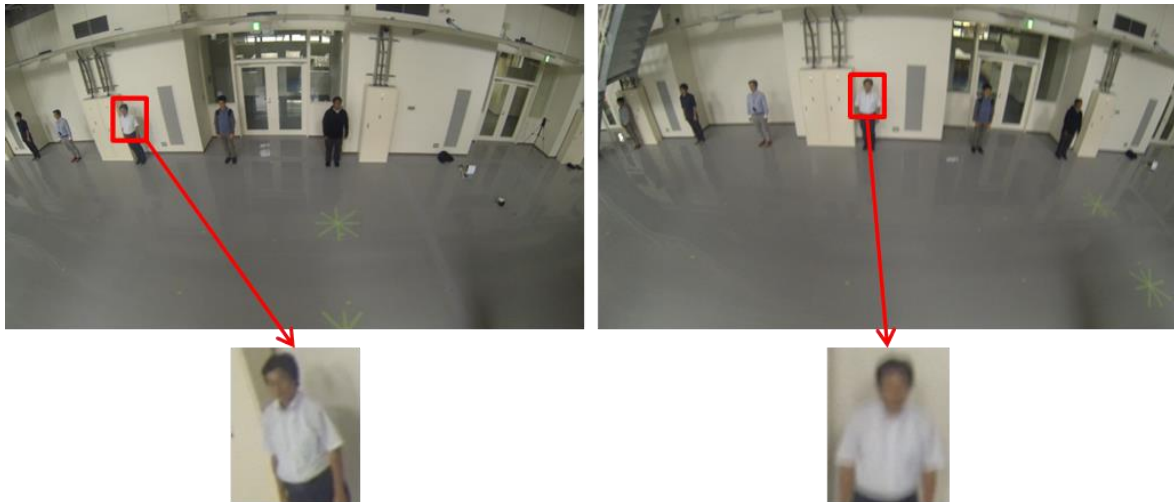


図3. 36: 画像サンプルと人物抽出結果

本実験では、人物の抽出と異なるフレーム間の人物の対応づけについては、目視による手動入力とした。これは、人物抽出、人物追跡の精度と、人物認証の精度を分けて示すためである。

3.5.2.1 人物の見えの角度の違いによる認証結果の比較

人物の向きが 0° （正面方向）で、人物の仰角(体の中央付近：高さ約 1m)が約 30° のときの人物検索実験を行なった。検索対象は ID3（図 3. 34）とした。このとき $D = 3.5 (\text{m})$ とした。このときの時系列での照合値の変化と画像入力を図 3. 37 に示す。ID3 の入力画像が入力し始める 7.0 秒から徐々に累積照合値が上昇していることが確認できた。その間、他人の入力画像も入力されるが、照合値の上昇度合いは小さい。飛行ロボットで撮影した映像に、異なる角度のデータが含まれているときは本人の照合値が増えることを確認した。

次に、人物の向いている方向を 90° ずつ回転させて、異なる方向での認証実験を行なった。その結果を 0° の結果も含めて図 3. 38 に示す。また、そのときの ID3 の入力画像(3 画像ごとに抜粋)を図 3. 39 に示す。

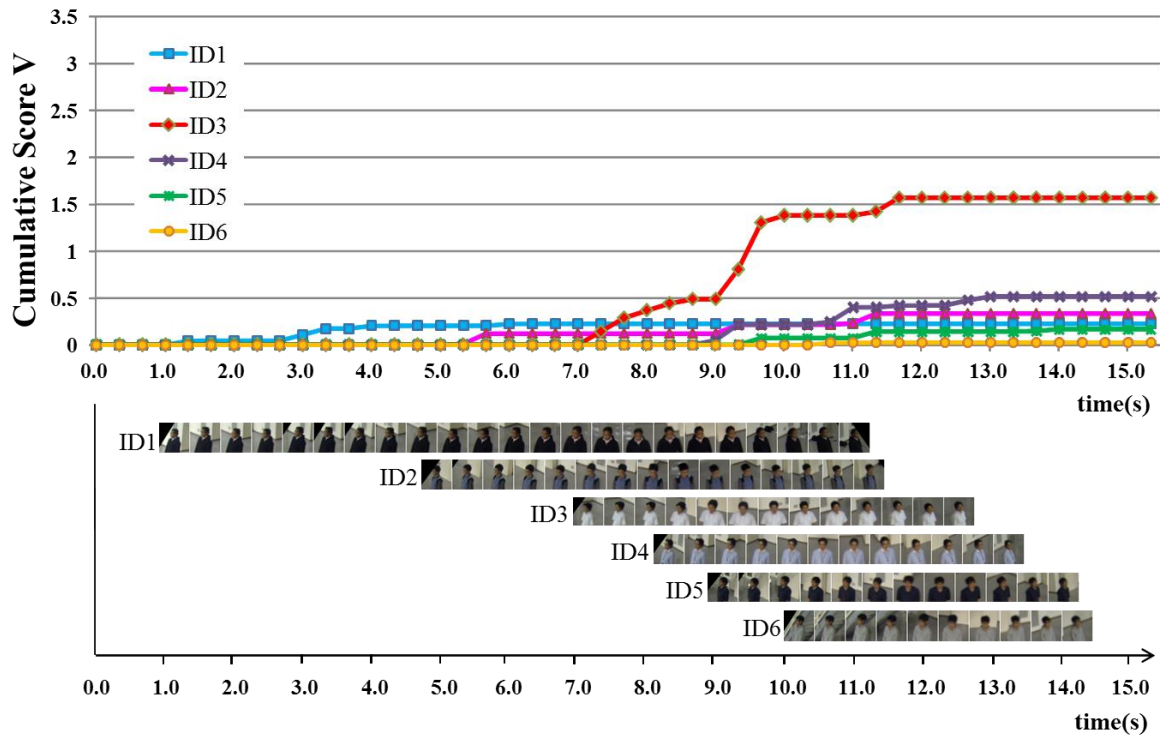


図3.37: 時系列での画像入力と照合値の変化
(スペースの都合で、画像は3フレームごとに抜粋して表示している.)

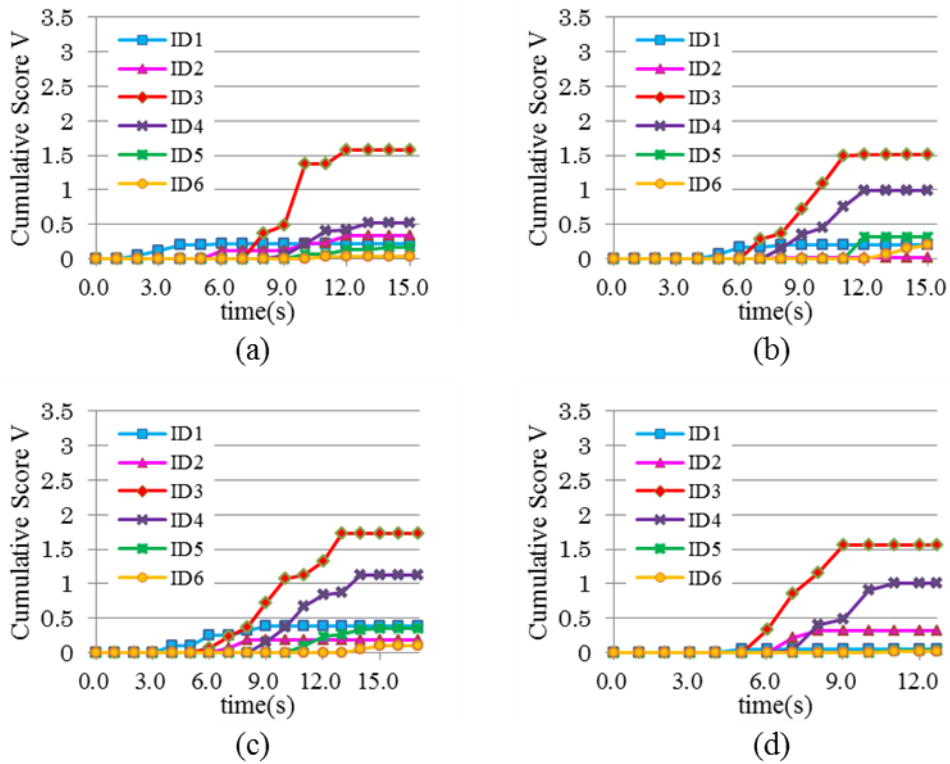


図3.38: 時系列での照合値の変化 ((a)0°, (b)90°, (c)180°, (d)270°)

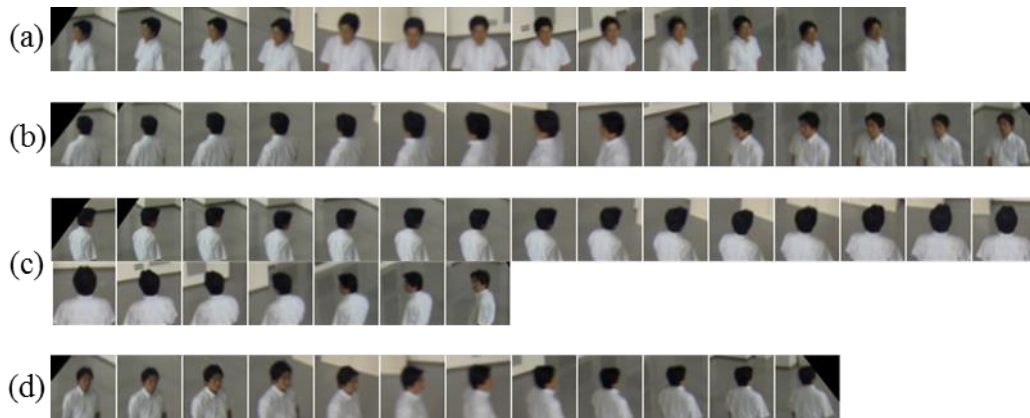


図3. 39: 各向きでのID3の入力画像 (3画像ごとに抜粋)
 ((a)0°, (b)90°, (c)180°, (d)270°)

図 3. 38 より，人物の向きがどの角度であっても，最終的な照合値は約 1.5 に収束し他人と分離できている．また，ID4 の照合値が 90°，180°，270°のときは約 1.0 まで上昇したが，0°は約 0.5 に留まった．これは，ID3 のテンプレートの正面向き(0°)に，他人と分離できる特徴が他の向きよりも多いからだと考えられる．図 3. 39 より，180°の入力画像枚数が他の入力画像枚数の約 1.5 倍程度多いが，照合値に対する影響は軽微である．同じ角度の画像の枚数が増えても，照合値に対する影響は小さいという本アルゴリズムの特徴が表れている．尚，撮影枚数については，出来る限り一定速度になるように飛行ロボットの操作を行ったが，速度に誤差が生じ，撮影枚数も一定にはならなかった．飛行ロボットは風や空気の流れの影響を受けやすく，操作を自動操作に切り替えても一定速度，一定方向に飛ぶことは難しい．よって，これは実運用時でも考えられることである．

3.5.2.2 飛行ロボットの高さ（仰角）の違いによる認証結果の比較

A の実験に加えて，人物の向きが 0°（正面方向）のときに，人物の仰角が 15°と 50°の場合の認証実験を行なった．撮影される人物画像が画面中央になるように，15° のとき， $D = 7.5$ (m)，50° のとき， $D = 1.7$ (m)とした．その照合結果を図 3. 40 に示す．そのときの ID3 の入力画像を図 3. 41 に示す．

図 3. 40 より，人物の仰角が 15°，50°のどちらも他人と分離できている．図 3. 38，図 3. 40 より，15°，30°，50°と仰角が大きくなるにつれて，最終的な照合値は 1.7，1.5，1.2 と，小さくなる傾向が見受けられる．これは人間の見た目の感覚とも一致している．しかし，50°程度の仰角であれば，認証に寄与できるといえる．

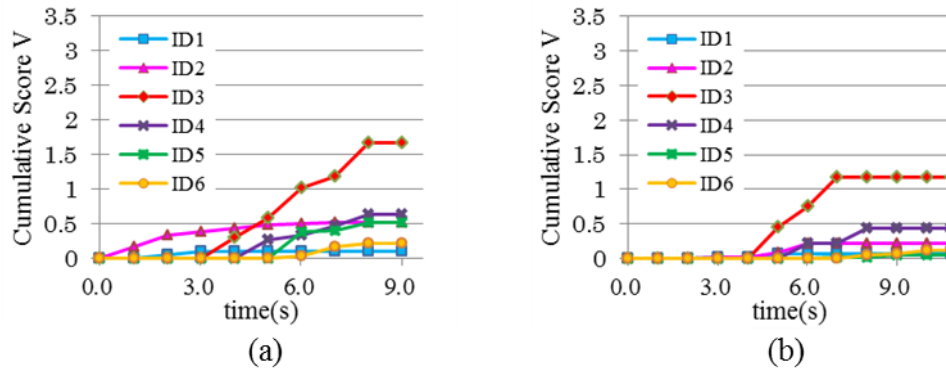


図3.40: 時系列での照合値の変化 ((a)15°, (b)50°)

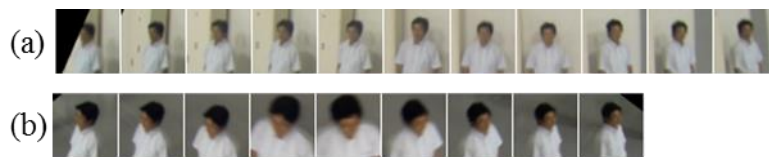


図3.41: 各向きでのID3の入力画像 (3画像ごとに抜粋) ((a)15°, (b)50°)

3.5.2.3 検索対象者 (テンプレート) の違いによる認証結果の比較

3.5.2.1の実験を他の検索対象者 (ID1-ID6) でも行った. その結果を図 3.42 に示す. (ID3の結果も含む)

図 3.42 から分かるように, 検索対象者によって照合値の上がり方, 最終照合値に差が生じている. 入力画像枚数は順に, (ID1-ID6: 67,45,38,36,35,30)であるが, 最終照合値との相関性は小さい. その他の撮影条件は同じことから, テンプレート画像そのものによる他人との分離しやすさが表れていると考えられる. 最終照合値は, どの人物も約 1.5 以上であり, 他人は約 1.0 に留まっている. どのテンプレートも他人と分離可能であるといえる.

3.5.2.4 認証精度まとめと, 人物捜索用閾値設定

3.5.2.1~3.5.2.3の実験で示した結果は, どの条件においても, 本人を間違えることなく他人と分離することが出来た. しかし, 示した結果は全パターンの一部である. ここで, 全パターン (人物の仰角 3 パターン×人物の向き 4 パターン×6 人のテンプレート) の照合結果を本人の照合値と他人の照合値に分けて集計する. その結果を図 3.43 に, FRR(False Reject Rate), FAR(False Accept Rate)のグラフで示す.

図3.43から, 照合値が0.8のときに, EER(Equal Error Rate)が約10%となった. 仮に0.8を閾

値として設定すると、誤認率は約10%となるが、本研究の目的である検索対象の候補を挙げ、最終的に目視確認を行うことは可能であると考えられる。よって、次節の人物検索実験は、閾値0.8と設定し行った。

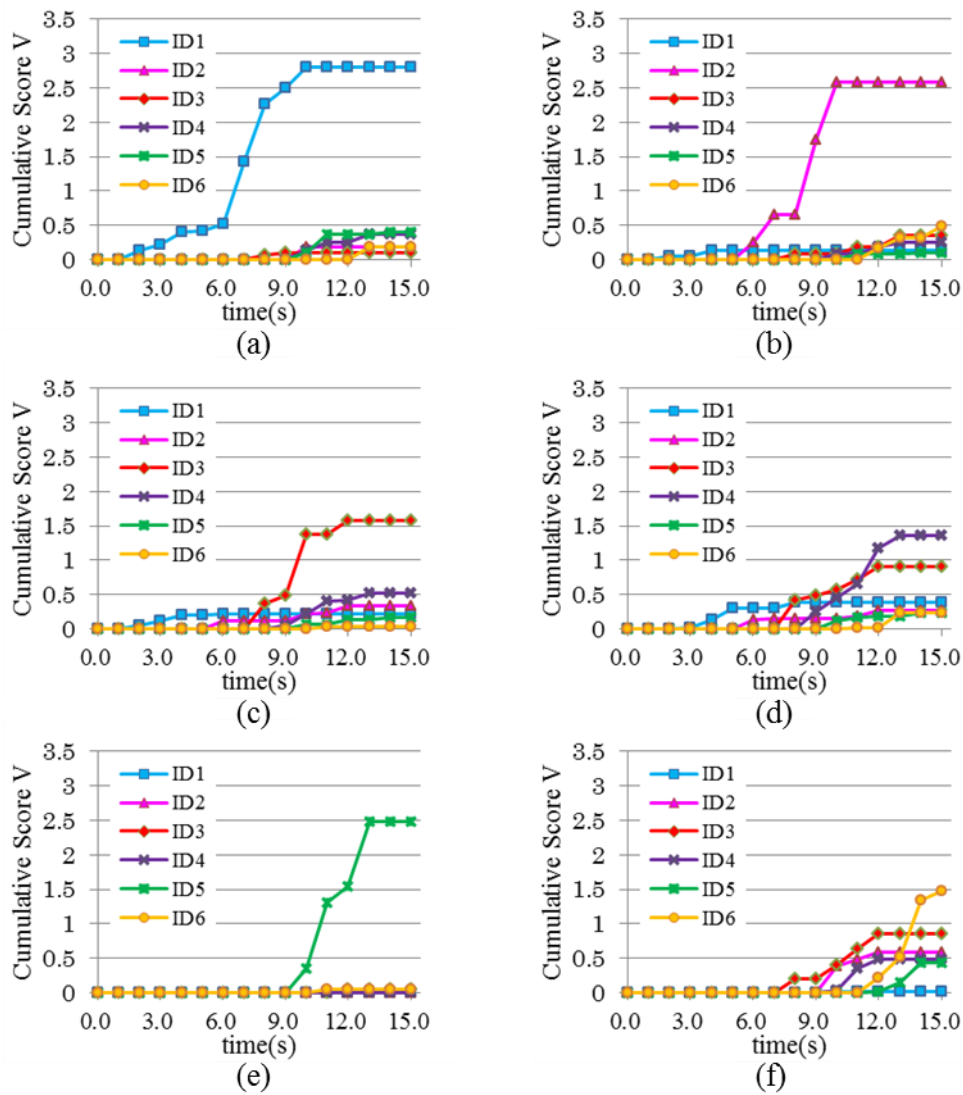


図3.42: 各テンプレートの時系列での照合値の変化
((a)ID1, (b)ID2, (c)ID3, (d)ID4, (e)ID5, (f)ID6)

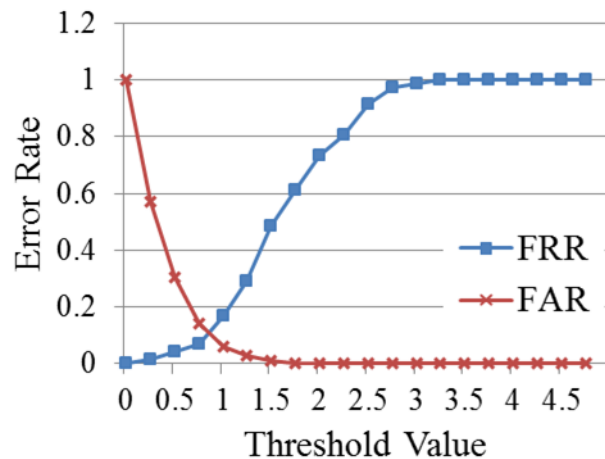


図3.43: 人物が静止状態での全パターンのFAR,FRRグラフ

3.5.3 複数の歩行中の人間を飛行ロボットで撮影した場合の認証実験

人物が通路などを歩行している状態で飛行ロボットが探索するシーンを想定し、人物検索が可能か3パターンの実験を行った。パターン1は、通路を歩行中の人物に対して、人物の正面方向を撮影した映像である。パターン2は、通路を歩行中の人物に対して、同じ方向に移動しながら、人物の正面方向を撮影した映像であり、パターン3は、ランダムな方向に移動している人物に対して、移動しながら万遍なく撮影した映像である。本パターンは人物検出実験の2.7.2.2のパターン番号と一致している。

3.5.3.1 通路を歩行中の人物に対して、人物の正面方向を撮影した映像を認証した結果

人物は通路を全員同じ方向に歩き、飛行ロボットは人物の正面方向に向きを併せ前進する。数人が通りすぎた後、飛行ロボットが振り返り、人物の背面方向を撮影しようとする。その模式図と画像サンプルを図3.44に示す。また、各テンプレートを検索対象とした場合の照合結果を図3.45に、そのときの入力画像を図3.46に示す。

図3.45に示したように、最終照合値が閾値0.8以上になったのは、6人中1人だけ（ID1）であった。他人の照合値も上昇せず、誤認証は0であった。これは、入力画像に原因があると考えられる。取得画像(図3.43)を見ると、図3.37、図3.39、図3.41と比較してブレている画像が多い。入力画像の半数以上が大きくブレている。3.4.5の実験と比較しても、人物角度のバリエーションには問題がなく、入力画像枚数は同程度である。よってこの画像ブレが、照合値が上昇しない原因と考えられる。本パターンでは、飛行ロボットは人物に対して対面しすれ違うため、画像ブレが大きくなったと考えられる。また、飛行ロボットの振り返りも高速であったため画像ブレが起こったと考えられる。しかし、このようなブレた画像に対しては、他人の照合値も上昇せず、誤認証を起こさないことが分かった。

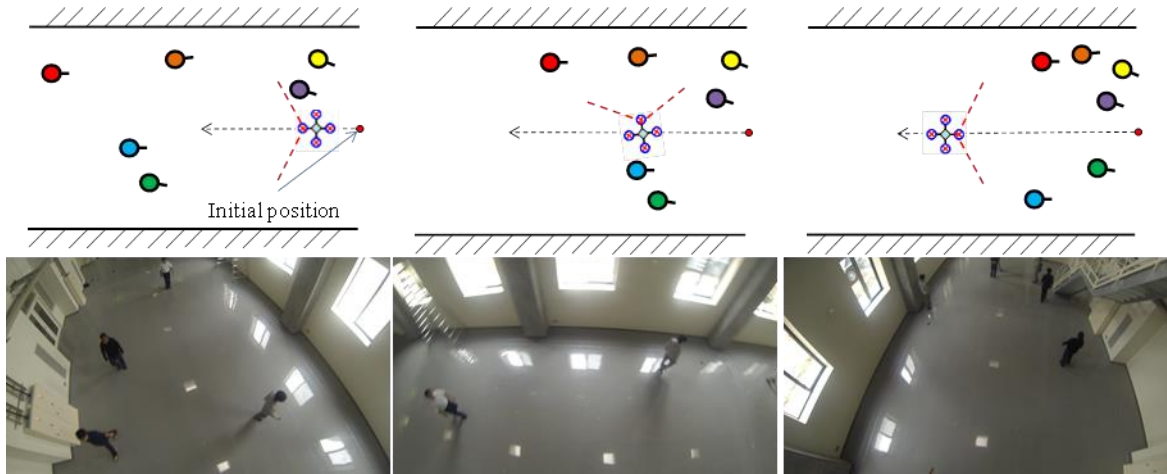


図3.44: 飛行パターンの模式図と取得画像抜粋

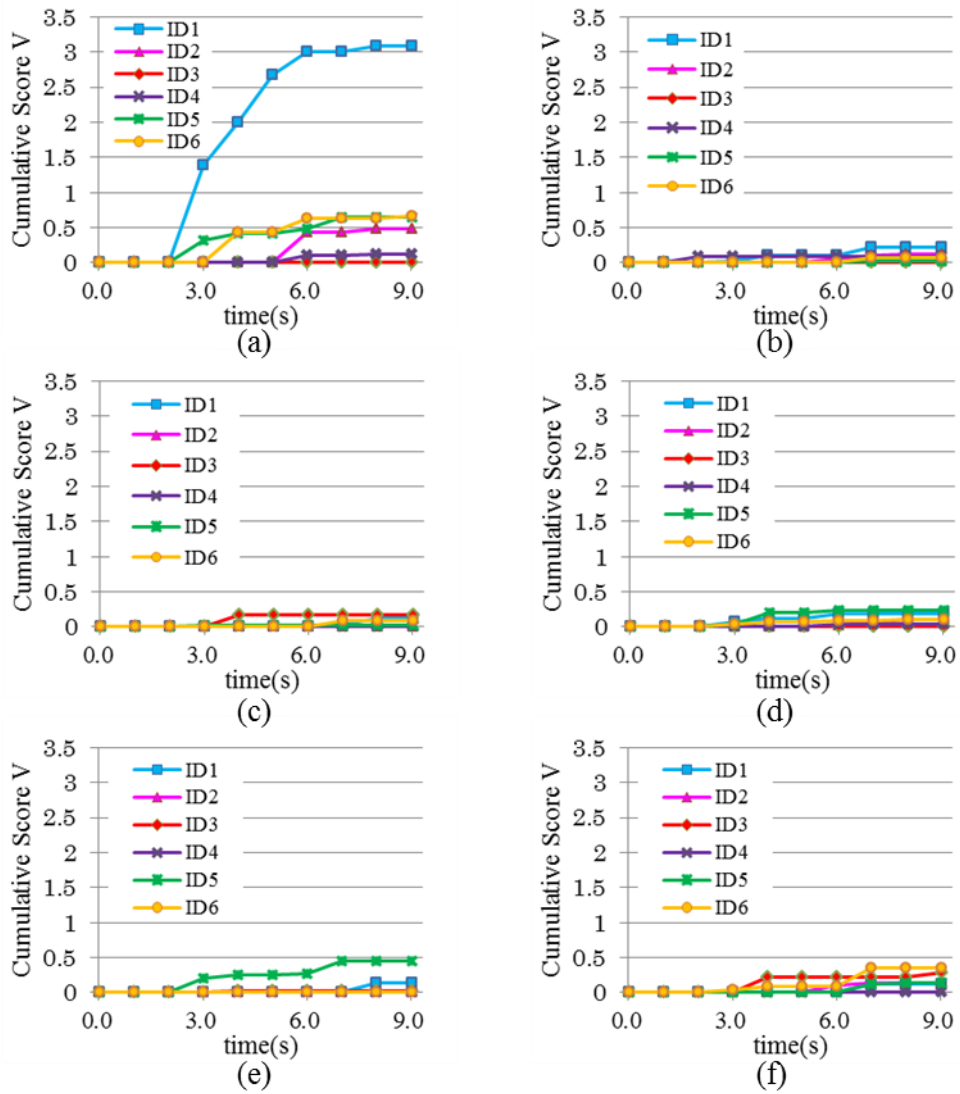


図3.45: 各テンプレートの時系列での照合値の変化
 ((a)ID1, (b)ID2, (c)ID3, (d)ID4, (e)ID5, (f)ID6)

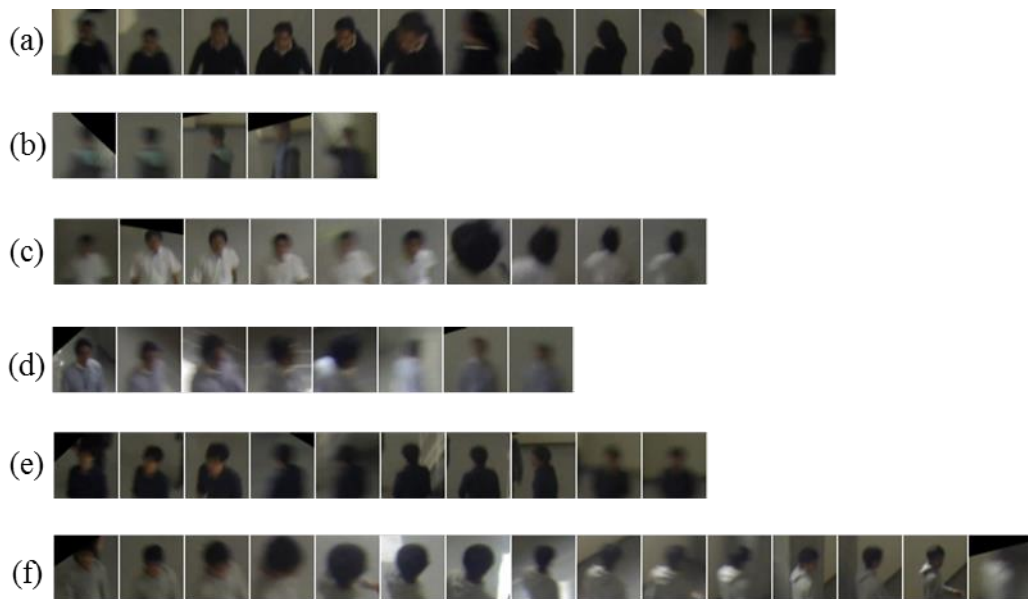


図3.46: 各実験対象者の入力画像 (3画像ごとに抜粋)
 ((a)ID1, (b)ID2, (c)ID3, (d)ID4, (e)ID5, (f)ID6)

3.5.3.2 通路を歩行中の人物に対して、同じ方向に移動しながら、人物の正面方向を撮影した映像を認証した結果

人物は通路を全員同じ方向に歩き、飛行ロボットは人物方向を向きながら人物の前方を人物と同じ方向に進む。すなわち人物の前方を後ろ向きに進む。その模式図と画像サンプルを図3.47に示す。また、各テンプレートを検索対象とした場合の照合結果を図3.48に、そのときの入力画像を図3.49に示す。

図3.48に示したように、本実験では、6人中2人 (ID1, ID5) の最終照合値が閾値0.8以上となった。誤認証は0であった。図3.49から、飛行ロボットと人物が同じ向きに進んだため、パターン1と比較するとブレの少ない画像が取得できた。しかし、角度バリエーションは少なくなった。図3.45の(a)では、ID1の照合値は約3.0まで上昇しているが、図3.48の(a)では、ID1の照合値は約2.0に留まっている。一方、図3.46, 図3.49から、入力画像数はパターン2のほうが約2倍多い。この結果からも、角度バリエーションの多さが照合値上昇に寄与していることが分かる。

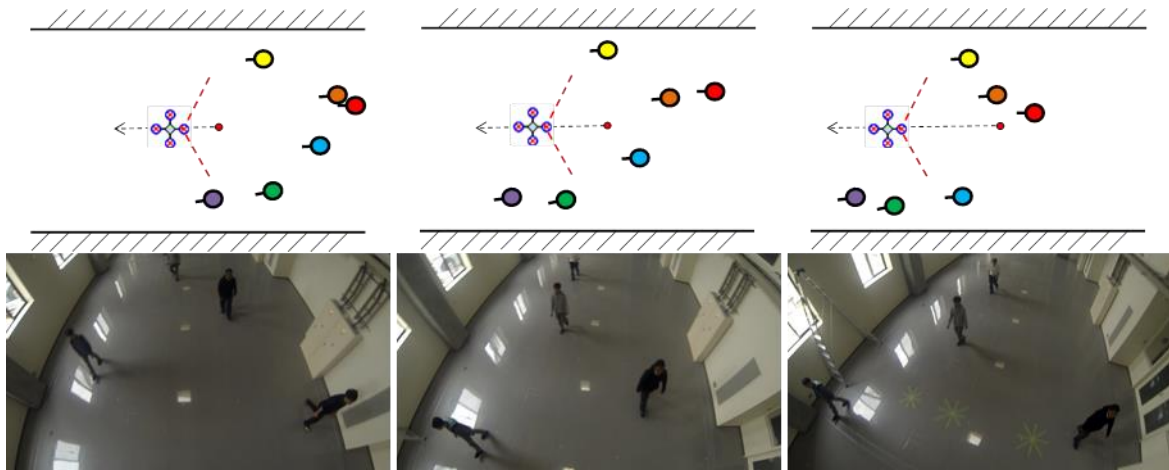


図3.47: 飛行パターンの模式図と取得画像抜粋

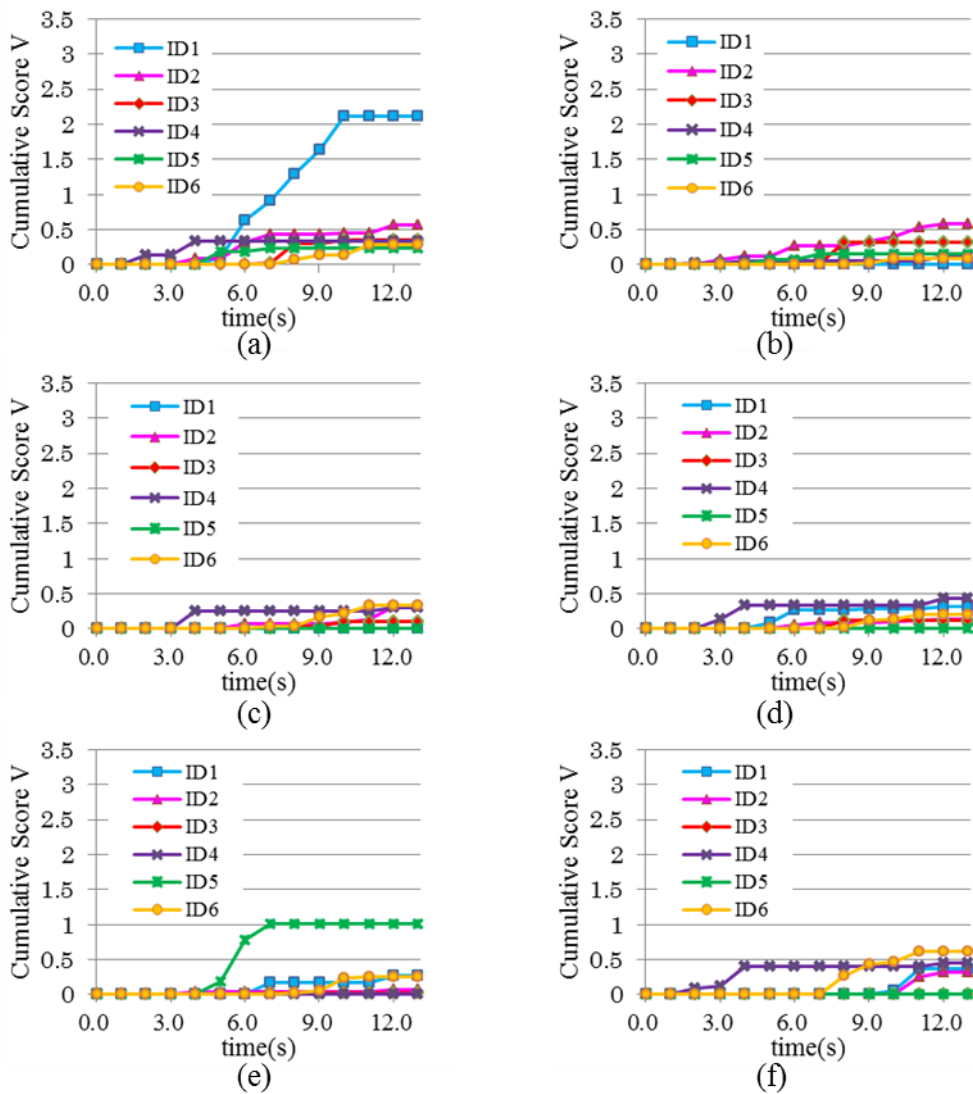


図3.48: 各テンプレートの時系列での照合値の変化
 ((a)ID1, (b)ID2, (c)ID3, (d)ID4, (e)ID5, (f)ID6)



図3.49: 各実験対象者の入力画像 (3画像ごとに抜粋)
((a)ID1, (b)ID2, (c)ID3, (d)ID4, (e)ID5, (f)ID6)

3.5.3.3 ランダムな方向に移動している人物に対して、移動しながら万遍なく撮影した映像を認証した結果

人物はランダムな方向を歩き、方向転換も任意のタイミングでランダムに行ってもらった。これは、イベント会場などの広場を想定した。飛行ロボットは人物全員が映るように向きを変えながら、通り過ぎていく。その模式図と画像サンプルを図3.50に示す。また、照合結果を図3.51に、取得できた画像を図3.52に示す。

図3.51に示したように、本実験では、実験対象者6人中6人とも照合値が閾値0.8を上回る結果となった。ただし、誤認証がID3のときに1人起こった。取得できた画像(図3.52)から、さまざまな角度の人物画像を取得できていることがわかる。人物自身がさまざまな方向に向きを変え移動し、その様子を飛行ロボットが撮影しているため、さまざまな方向の人物画像が、ブレが少なく取得できている。ブレが少ない、角度バリエーションが多いという好条件が揃ったため、どのテンプレートの場合も照合値が上昇し、閾値0.8以上となったと考えられる。図3.51の(c)の誤認証については、本人の照合値も閾値0.8以上となっている。2人以上が閾値を越えた場合には、リジェクトする、再度撮影しなおし認証する、2人とも候補として挙げる等の、その後の行為を決めることで対応可能である。

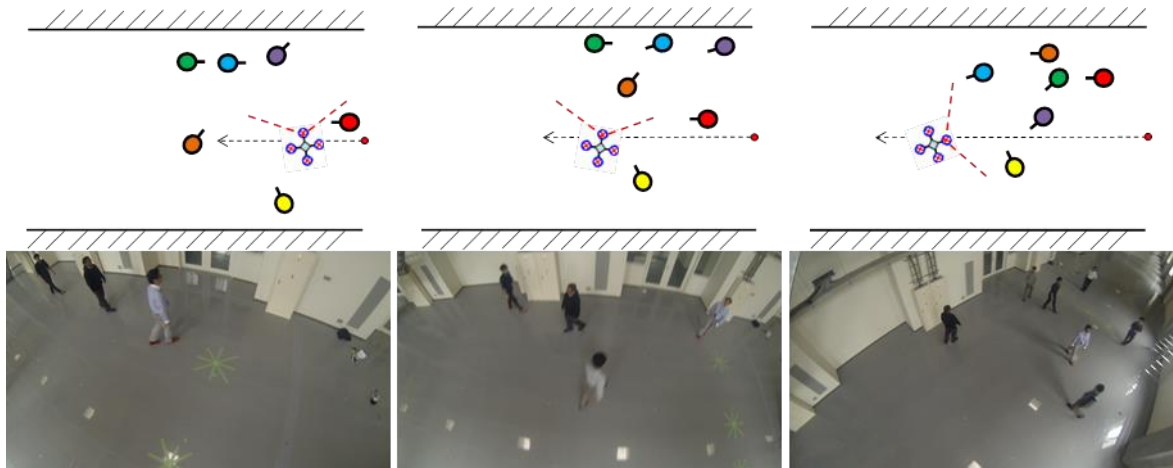


図3.50: 飛行パターンの模式図と取得画像抜粋

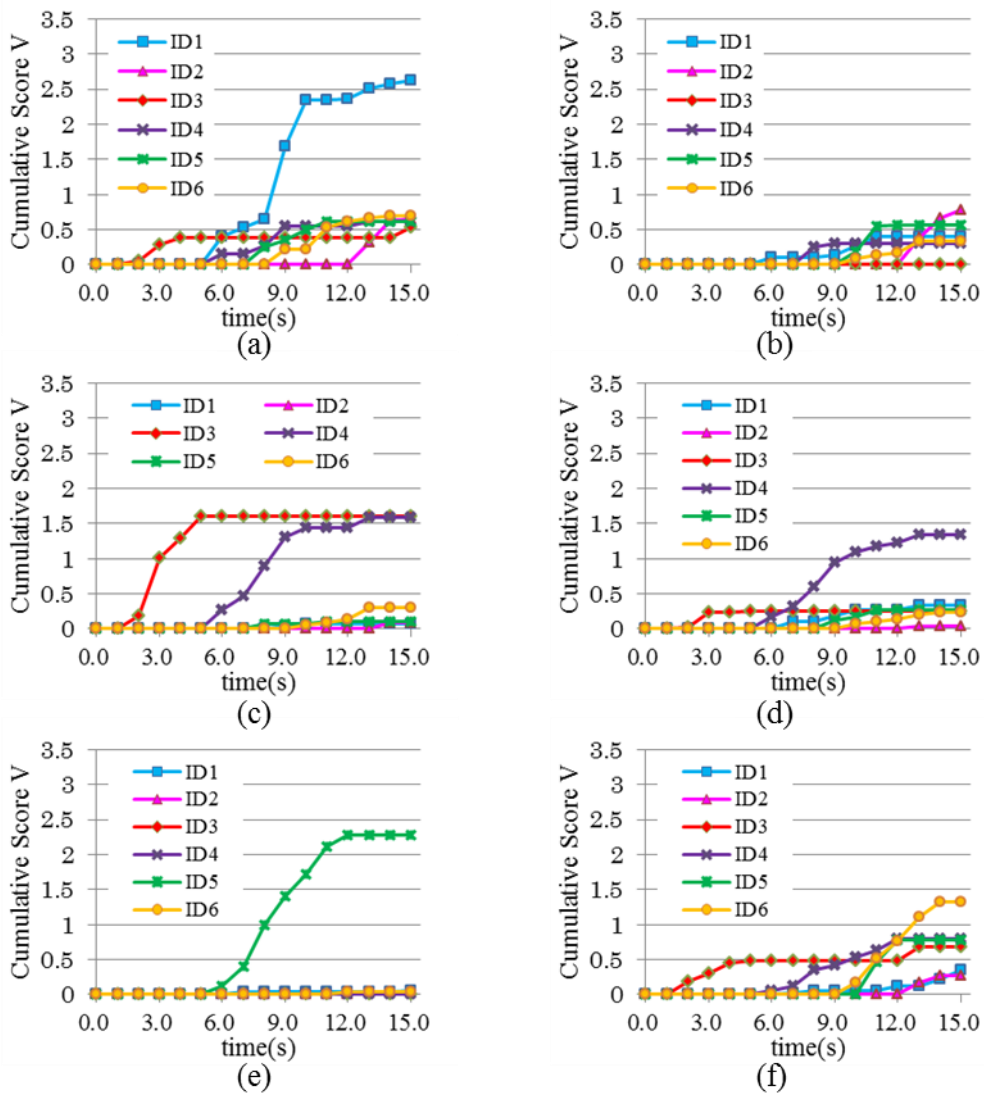


図3.51: 各テンプレートの時系列での照合値の変化
 ((a)ID1, (b)ID2, (c)ID3, (d)ID4, (e)ID5, (f)ID6)

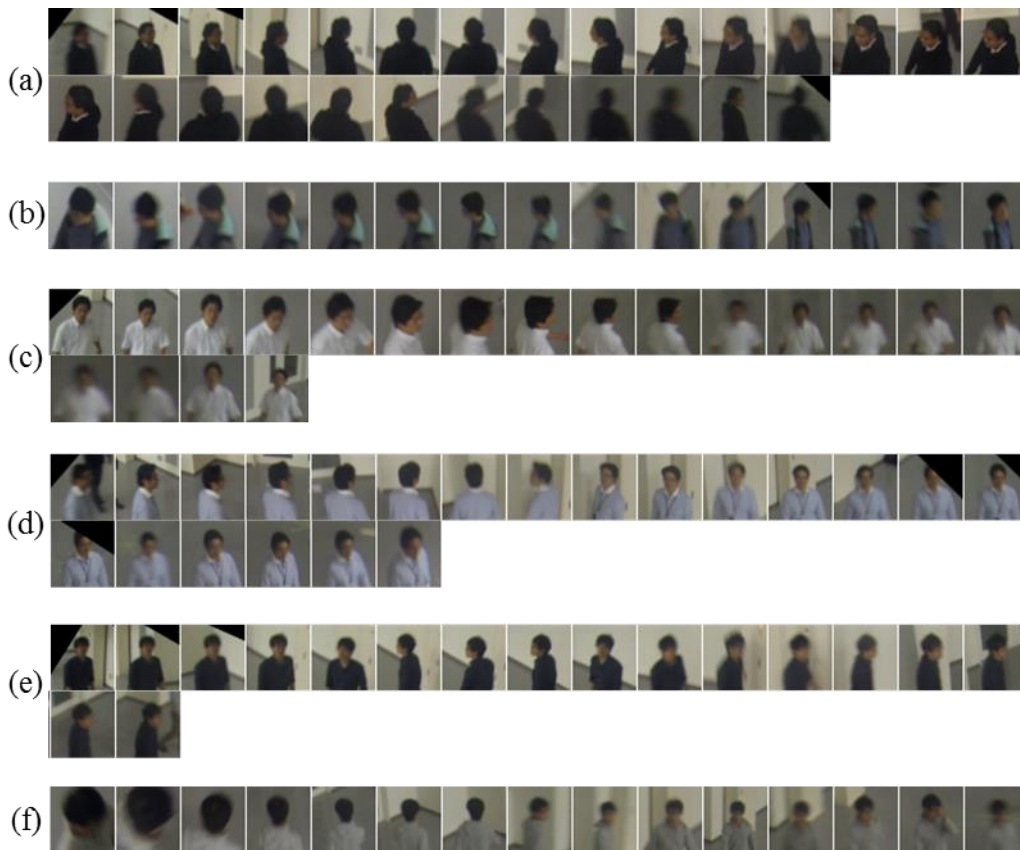


図3. 52: 各実験対象者の入力画像 (3画像ごとに抜粋)
 ((a)ID1, (b)ID2, (c)ID3, (d)ID4, (e)ID5, (f)ID6)

3.5.4 飛行ロボットによる人物搜索実験のまとめと考察

本実験の結果は、手動による人物を行った結果を用いて人物認証を行った結果である。

実際の飛行ロボットが撮影した画像を用いた認証実験では、認証対象者が静止状態で、認証対象者がどの方向を向いていても、飛行ロボットが認証対象者から数メートル離れた位置を通過すれば、実験対象者6人の中から搜索対象候補者を見つけることが出来た。認証対象者が移動状態では、飛行ロボットの軌道と、認証対象者の移動方向、位置関係によって差はあるが、実験対象者6人の中から搜索対象候補者を見つけることが出来るシーンを確認できた。

飛行ロボットでの実験結果(3.5)より、移動するカメラと認証対象者との相対関係が認証に影響を与えていることが分かる。認証精度を上げるポイントは2つある。第1のポイントは、画像ブレを少なくすることである。飛行ロボットが認証対象者との相対速度を小さくする軌道を飛行すれば、認証精度は高くなると考えられる。ブレの影響は、固定カメラでの実験結果(3.4)からも推察できる。テンプレート画像や入力画像の違いから直接比較できないが、ブレのない画像では、同程度の解像度で38人の人物をほぼ100%認証できている。第2のポイントは、入力人物画像の角度のバリエーションを増やすことである。認証対象者の移動方向

と飛行ロボットの軌道を工夫する必要がある。また、複数台の飛行ロボットを用いることが出来るならば、複数の位置から異なる角度の認証対象者の画像を取得する手段も考えられる。認証対象者と飛行ロボットとの相対速度を小さくし、撮影される認証対象者の角度バリエーションを増やすような軌道を飛行させることが出来れば、認証精度は高くなる。

本手法では、固定カメラでのアピランス照合の結果以上の精度は出ない。そのため、本研究のように本人候補を挙げ、人が目視確認を行うシステムであれば有用であるが、街中で犯人を見つけ、自動発報するようなシステムにそのまま適用することは難しい。本人候補を挙げたあと、別途精度の高い認証手法（例えば顔認証）と組み合わせることで実現できる可能性はある。

3.6 まとめ

本章では、飛行ロボットによる人物検索を行うための手法を提案した。飛行ロボットにカメラを搭載し、撮影した画像に対し人物認証を行うことで、人物検索を行う。その人物認証アルゴリズムとして、累積アピアランス照合アルゴリズムを提案した。そのアルゴリズムは、上半身のみ、低解像度、エッジ特徴のみ（カラーは使わない）という認証するには悪条件ではあるが、あらゆる角度の人物画像の情報を蓄積することで認証することが出来るものである。そのアルゴリズムの精度について検証を行った。解像度においては、私服を着た被験者38人による認証実験の結果、 46×46 画素、 23×23 画素の画像においては、ほぼ100%の認証率であった。 15×15 画素の画像では、平均85%の認証率（誤認率5%）を得ることが出来た。また、誤検出画像の入力にもロバストであることを示した。さらに、私服以外で、実環境で起こりうる服装パターンでの精度も明らかにした。まず、スーツのような認証対象者全員が同じ服装を着ている場合（制服）について認証実験を行った。私服よりも精度は落ちるが認証可能であることを示した。また、テンプレートと入力で服装が異なる場合について実験を行った結果、認証困難であることが分かった。このように、実環境において生じると考えられる解像度変化や服装パターン変化等での精度を明らかにし、適用可能性、適用限界地点を示した。

また、映像からの人物抽出が手動の場合において、飛行ロボットを実際に飛行させ撮影したデータに対して、人物検索実験を行い、課題はあるが人物検索可能であることを示した。認証対象者が静止状態の場合は、認証対象者がどの方向を向いていても、飛行ロボットが認証対象者から数メートル離れた位置を通過すれば、実験対象者6人の中から検索対象候補者を見つけることが出来た。認証対象者が移動状態では、飛行ロボットの軌道と、認証対象者の移動方向、位置関係によって差はあるが、実験対象者6人の中から検索対象候補者を見つけることが出来るシーンを確認できた。

第4章 人物検出と人物認証の統合による 人物検索

4.1 はじめに

第2章、第3章では、人物検出アルゴリズムと人物認証アルゴリズムのそれぞれについて述べた。しかし、人物検索の過程では、映像取得から、人物検出、人物認証と演算負荷が高いと考えられる処理を一連の流れで行う必要がある。本章では、その実現性の検証や、人物検出アルゴリズムと人物認証アルゴリズムを組み合わせたときの人物認証精度について述べる。

飛行ロボットが撮影した画像から人物検出アルゴリズムによる人物抽出を行い、その結果を用いて人物認証を行ったときの精度検証を行う。人物検出アルゴリズムの結果を用いて人物認証アルゴリズムに適用するには、複数のフレームに渡って写っている各人物の画像をまとめたシーケンスデータを作成する必要がある。図4.1に3フレームでのシーケンスデータ作成例を示す。図4.1は、各フレームで検出された人物画像を、各人物ごとにまとめている様子を示している。シーケンスデータの作成については、本研究では簡易なアルゴリズムを適用し作成した。本章ではその結果についても述べる。また、作成されたシーケンスデータを人物認証アルゴリズムに適用したときの精度検証も行った。

さらに、人物検出から人物認証までの一連の流れをシステムとして作成し、現実的な処理時間で結果が返せるか検証を行った。

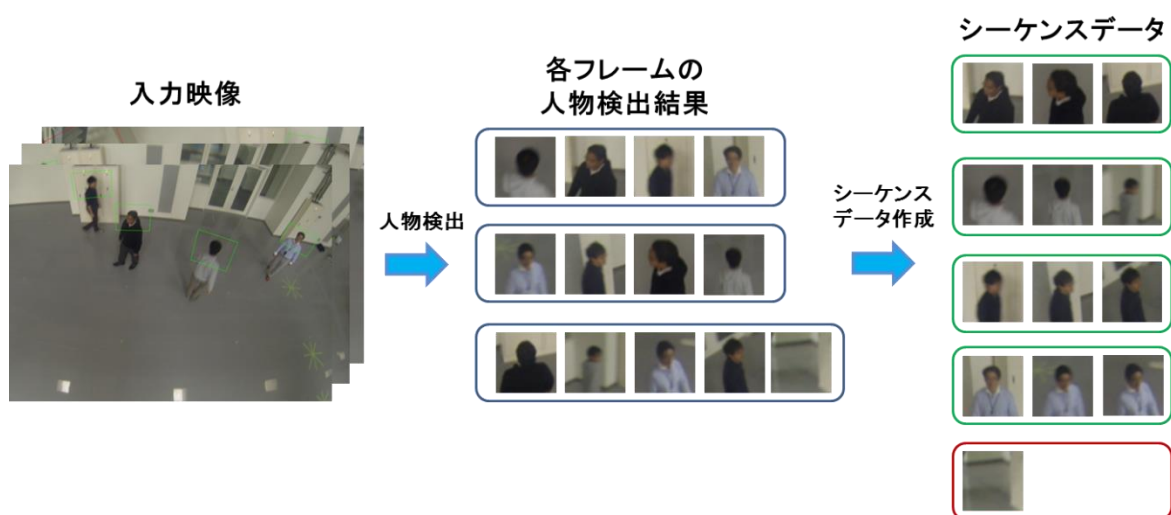


図4.1: 3フレームでのシーケンスデータ作成
(赤枠：誤検出シーケンスデータ)

4.2 複数フレーム間の人物画像の統合

4.2.1 複数フレーム間の人物画像の統合手法

人物検出アルゴリズムを用いて検出された各フレームの人物抽出結果を統合する手法について述べる。人物検出アルゴリズムは、画像1枚に対して人物検出を行い、その人物位置とその大きさを出力する。そのため、隣り合うフレーム間でも、出力された人物が同じ人物であるかどうかは判断していない。よって、本研究では、簡易なフレーム間追跡アルゴリズムを適用し、複数フレーム間で同じ人物をまとめ、人物ごとのシーケンスデータを作成することにした。その追跡アルゴリズムは、画像内の人物位置、サイズ、そして検出画像間の類似度によって追跡を行う。

シーケンスデータを作成するアルゴリズムの詳細は次の通りである。(a) : n フレーム目の画像に人物検出アルゴリズムを適用し、設定した閾値以上の評価値を得る人物位置候補（以後、追跡元オブジェクト）を求める。(b) : $n+1$ フレーム目の画像に人物検出アルゴリズムを適用し、設定した閾値以上の評価値を得る人物位置候補（以後、ターゲットオブジェクト）を求める。(c) : 追跡元オブジェクトの人物検出の評価値の最も高いものを選択する。(d) : 選択した追跡元オブジェクトと求めた複数のターゲットオブジェクトとを照らし合わせ、距離とサイズの閾値で候補を絞る。(e) : 追跡元オブジェクトと残った複数のターゲットオブジェクトとの間で、画像の正規化相関を行う。(f) : 正規化相関結果が設定した閾値以上で、最大値を得るターゲットオブジェクトを求める。(g) : 求めたターゲットオブジェクトと追跡元オブジェクトを1つのシーケンスデータとして保存する。これを全ての追跡元オブジェクト数分繰り返す。ターゲットオブジェクトが求められなかった追跡元オブジェクトは追跡終了である。(h) : フレーム番号 n を更新($n++$)し、(b)~(g)を繰り返す。

4.2.2 人物シーケンスデータの作成

第3章では、手動による人物画像のシーケンスデータを作成し、人物認証結果を求めた。よって、各シーンのシーケンスデータは各人物ごとに全フレームが連結している。また、誤検出画像は含んでいない。人物検出とシーケンスデータの作成までが理想的に行えたときの結果と言える。本節では3.5で用いた飛行ロボットが撮影した映像に対して人物検出アルゴリズム、フレーム間追跡アルゴリズムを適用し、シーケンスデータを作成した。そのため、それらのシーケンスデータは、誤検出画像が含まれる場合や、誤検出画像のみの場合、人物画像が途中で途切れ分割される場合もある。それらの例を図4. 2に示す。また、表4. 1に人物検出と人物追跡を行ったときの設定パラメータを示す。

表4.1 人物検出パラメータと，人物追跡パラメータ

人物検出パラメータ	サーチ個数	20
	サーチサイズ	15~150
	大分類閾値	20
	最終照合値閾値	61
人物追跡パラメータ	距離閾値(画素)	100
	面積比閾値	0.5
	相関値閾値	0.7

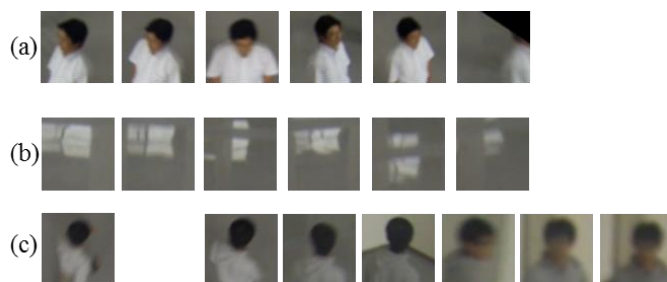


図4.2: シーケンスデータ例

(a)誤検出含む (b)誤検出画像のみ (c)シーケンス分割

作成された各シーケンスデータの精度検証のための正解IDは，そのシーケンスデータに含んでいる人物画像が最も多いIDの番号とした．全く人物画像が含まれていない場合は，誤検出シーケンスデータとして扱った．各シーンのシーケンスデータ数の表を表4.2に示す．

表4.2: 各シーンのシーケンスデータ数

撮影シーン		ID1	ID2	ID3	ID4	ID5	ID6	誤検出
静止	水平角0° 仰角50°	5	4	1	1	4	1	13
	水平角90° 仰角50°	1	2	4	3	4	3	23
	水平角180° 仰角50°	3	3	2	1	4	1	25
	水平角270° 仰角50°	1	3	3	2	3	2	24
	水平角0° 仰角30°	11	8	3	2	1	4	70
	水平角90° 仰角30°	11	4	4	5	4	4	93
	水平角180° 仰角30°	9	6	6	2	6	7	66
	水平角270° 仰角30°	8	7	3	1	2	5	74
	水平角0° 仰角15°	16	15	3	3	4	7	124
	水平角90° 仰角15°	18	17	5	1	3	6	134
	水平角180° 仰角15°	14	14	5	4	8	9	138
水平角270° 仰角15°	20	18	4	3	3	8	81	
歩行	パターン1	3	0	7	1	4	7	34
	パターン2	6	8	2	2	1	2	47
	パターン3	6	5	4	9	13	6	62

表4.2から、人物データは複数個のシーケンスデータが作成され、誤検出データが多数作成されたことが分かる。人物データが複数個に分割されている理由は3つ考えられる。1つめは各シーンの中で人物検出を失敗しているフレームがあり、追跡が途切れたことによるものである。2つめは、追跡アルゴリズムでの誤リジェクトである。そして3つめは、これは分割ではないが、人物検出時に1人の人物に対して、複数個の検出結果が出力された場合である。図4.3に複数検出の例を示す。この場合は、1人の人物に対してこの検出結果を含んだ複数のシーケンスが作成されることになる。しかし、本研究の目的は人物検索であることから、どちらのシーケンスの照合値が高くなってよい。よって、人物検出結果で複数の重なりがある場合もシビアに選択し一方をリジェクトすることはせず、どちらも結果として出力することにした。

また、表4.2の仰角 15° のID1とID2（黄色部分）は、他に比べてシーケンス数が多くなっている。この原因は、図4.4に示すように、ID1とID2の間に窓ガラスがあり、この部分で多くの人物誤検出を出力し、その影響を受けていることである。

表4.2から、誤検出シーケンスが多量に発生していることが分かるが、誤検出データに対し人物認証アルゴリズムがロバストであることは、第3章で述べた。本人画像のリジェクト数を減少させることを優先し、表4.1に示したように、人物検出の閾値を61(人物が静止しているシーンで、FRR,FARが約10%)に設定したため、誤検出データが多数作成された。

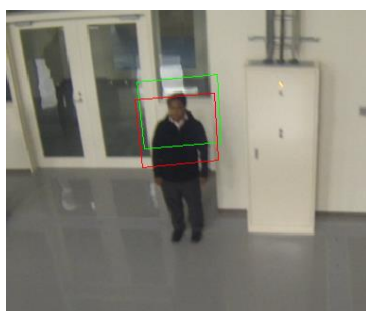


図4.3: 1人の人物に対して複数個の検出の例



図4.4: ドア部の誤検出多数の例

表4.3: 手動入力シーケンスに対する画像数の割合

撮影シーン		ID1	ID2	ID3	ID4	ID5	ID6
静止	水平角0° 仰角50°	1	0.9	1	0.9	0.9	1
	水平角90° 仰角50°	1	1	1	0.5	1	1
	水平角180° 仰角50°	0.5	0.9	1	1	0.9	1
	水平角270° 仰角50°	1	0.5	0.6	0.5	0.6	0.5
	水平角0° 仰角30°	0.6	1	1	0.9	1	1
	水平角90° 仰角30°	1	1	1	0.6	1	1
	水平角180° 仰角30°	1	1	1	1	1	1
	水平角270° 仰角30°	1	1	1	1	1	1
	水平角0° 仰角15°	0.6	0.6	0.9	1	1	1
	水平角90° 仰角15°	0.4	0.7	1	1	1	1
	水平角180° 仰角15°	1	0.7	1	1	1	1
	水平角270° 仰角15°	0.6	0.6	1	1	1	0.9
歩行	歩行パターン1	0.5	0	0.3	0.1	0.3	0.2
	歩行パターン2	0.6	0.3	0.6	0.6	0.8	0.9
	歩行パターン3	0.9	0.6	0.5	0.5	0.4	0.8

表4.2からシーケンスデータが複数作成されることが分かることは前述した。しかし、照合値はそのシーケンスデータがどのように分割されているかに依存する。そこで、出力された各シーケンスに含まれている対象人物画像の数を求め、手動で求めたシーケンスデータの数に対する割合を求めた。そして、各シーン、各IDにおいてその最大値を求めた結果を、表4.3に示す。例えば、この値が0.5であれば、アルゴリズムで作成された複数のシーケンスの中には、手動入力で作成したシーケンスに対して、最大で半分の長さのシーケンスが存在することになる。この値が1であれば、手動入力と同じ長さのシーケンスが作成されていることになる。ただし、手動入力に対して、検出位置や大きさのズレは含まれる。人物が静止状態のシーンでは、この値の平均は0.89であり、約90%のシーケンスデータは手動入力と同等のものが作成されていると言える。しかし、歩行状態では平均0.49であり、手動入力と比較すると約半分の長さのシーケンスデータとなった。これは、2つの原因が考えられる。1つは、第2章(2.7.2.2)で述べたように、人物検出アルゴリズムによる人物画像の誤リジェクトである。人物画像自体がなくなるので、その前後のシーケンスは分離する。もう1つは、人物と飛行ロボットの双方が動くことで、フレーム間での人物画像の見え方の差が大きくなり、そのことによりフレーム間統合時の相関係数が小さくなって統合失敗していることが考えられる。

これらのシーケンスデータを用いて、人物認証実験を行った。

4.3 作成したシーケンスデータによる人物認証実験

4.3.1 静止した複数人物を飛行ロボットで撮影した場合の認証

3.5.2で扱った静止した複数人物を飛行ロボットで撮影したデータについて、評価実験を行った。図3.43と同様に、すべてをまとめてFRR,FAR曲線を求めた。図4.5に、図3.43の結果も含めたグラフを示す。尚、各シーン、各IDにおいて複数のシーケンスがある場合はその中の最大値を出力値とした。

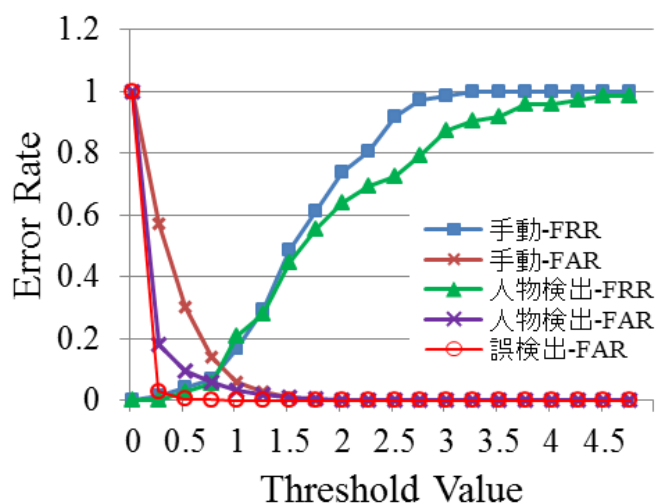


図4.5: 静止状態の全シーケンスのFAR,FRRグラフ

図4.5より、手動入力と比較すると、人物検出アルゴリズムから作成したシーケンスのFRRは照合値が2.5~3.5の間で低下していることが分かる。これは、シーケンスの分割による影響が表れていると考えられる。しかし、照合値が1.5以下の範囲ではほとんど差がないことから、比較的長いシーケンスが分割され照合値は落ちたが、閾値近辺(0.8)まで情報量が落ちることはなかったと考えられる。また、人物検出アルゴリズムから作成したシーケンスのFARが低下していることも分かる。この原因も同様で、シーケンス分割により、他人の照合値も低下したと考えられる。これらの結果を合わせると、EERについては、人物検出アルゴリズムは0.7で、手動入力は0.1であることから、むしろ小さくなる結果となった。誤検出のFARは閾値(0.8)以上は0であり、誤検出した画像を人物と間違えることは全くなかった。

これらのことから、静止した人物を飛行ロボットが撮影したデータについては、照合値が高いシーケンスについては低下したが、他人と見分ける照合精度については、影響がなかったと言える。

4.3.2 複数の歩行中の人間を飛行ロボットで撮影した場合の認証実験

3.5.3で扱った複数の歩行中の人間を飛行ロボットで撮影したデータについて、人物検索実験を行った。3.5.3と同様に、人物検索閾値は0.8としキャッチ数を求めた。シーケンスの全画像を入力した結果得られた本人照合値と他人の照合値の最大値を表にまとめた、表4.4に手動入力シーケンスによる結果を、表4.5に人物検出アルゴリズムを用いて作成したシーケンスによる結果を示す。

表4.4: 手動シーケンスによる照合結果
(黄色: キャッチ成功 赤: 誤キャッチ)

手動	テンプレートID1		テンプレートID2		テンプレートID3		テンプレートID4		テンプレートID5		テンプレートID6	
	本人値	他人値	本人値	他人値	本人値	他人値	本人値	他人値	本人値	他人値	本人値	他人値
歩行パターン1	3.09	0.65	0.12	0.21	0.18	0.12	0.04	0.24	0.45	0.13	0.35	0.29
歩行パターン2	2.12	0.56	0.59	0.32	0.11	0.34	0.44	0.33	1.01	0.26	0.61	0.46
歩行パターン3	2.63	0.7	0.8	0.57	1.6	1.59	1.35	0.34	2.27	0.05	1.32	0.79

表4.5: 人物検出アルゴリズムのシーケンスによる照合結果

人物検出 アルゴリズム	テンプレートID1		テンプレートID2		テンプレートID3		テンプレートID4		テンプレートID5		テンプレートID6	
	本人値	他人値	本人値	他人値	本人値	他人値	本人値	他人値	本人値	他人値	本人値	他人値
歩行パターン1	1.04	0.49	0	0.07	0.16	0.08	0.05	0.25	0.87	0.62	0.39	0.63
歩行パターン2	2.18	0.63	1.04	0.41	0.13	0.24	0.49	0.27	1.09	0.86	1	0.55
歩行パターン3	2.96	0.7	0.31	0.43	1.25	0.44	0.53	0.43	1.15	0.84	0.95	0.77

表4.4, 表4.5の結果から、手動と人物検出アルゴリズムでは、照合値に差があることが分かる。各シーン、各IDの照合値については、シーケンスの長さや、手動入力による切り出し位置、切り出しの大きさと、人物検出アルゴリズムによる切り出し位置、切り出しの大きさの差によっても影響を受け変動する。しかし、手動と人物検出アルゴリズムによる結果を集計すると、合計キャッチ数は同じで、誤キャッチ数は、人物検出アルゴリズムのほうが1つ多くなるに留まった。本人照合値の平均値は手動シーケンスでは1.06で、人物検出アルゴリズムのシーケンスでは0.87である。表4.3から、人物検出アルゴリズムのシーケンス画像数は、手動によるシーケンスと比較して、歩行パターンでは平均0.49の割合に減っているが、本人照合値はその割合で低下しているわけではないことが分かった。つまり1つのシーケンス内の画像それぞれが完全に無関連な画像であれば、同割合で低下すると考えられるが、フレーム間で連続している画像なので、ある程度相関があると考えられる。1つのシーケンスが2つに分離したとしても、情報量が完全に2つに分かれるわけではなく、分離したどちらのシーケンスにも共通する情報量が存在し、その情報量により照合値の低下が1.06から0.87に留まったと考えられる。その結果、キャッチ数や誤キャッチ数は、手動と人物検出アルゴリズムで、同等の結果を得ることができたと考えられる。

この情報量の差による照合値変化の例を次に示す. 図4. 6にテンプレートID 1 の歩行パターン1 の時系列変化の例を示す. このパターンでは, 人物検出アルゴリズムによるシーケンス作成が2つに分かれた. 表4. 4, 表4. 5に示したように, そのときの照合値は3.09から1.04に低下している. その原因はシーケンス画像の分かれ方にあると考えられる. 図4. 6のように, このパターンでは, 人物の前向き画像群と後ろ向きの画像群で半分に分かれている. 人物の前向き画像と後ろ向き画像では相関が小さいと考えられ, 情報量が分離してしまったため, 照合値が低下したと考えられる.

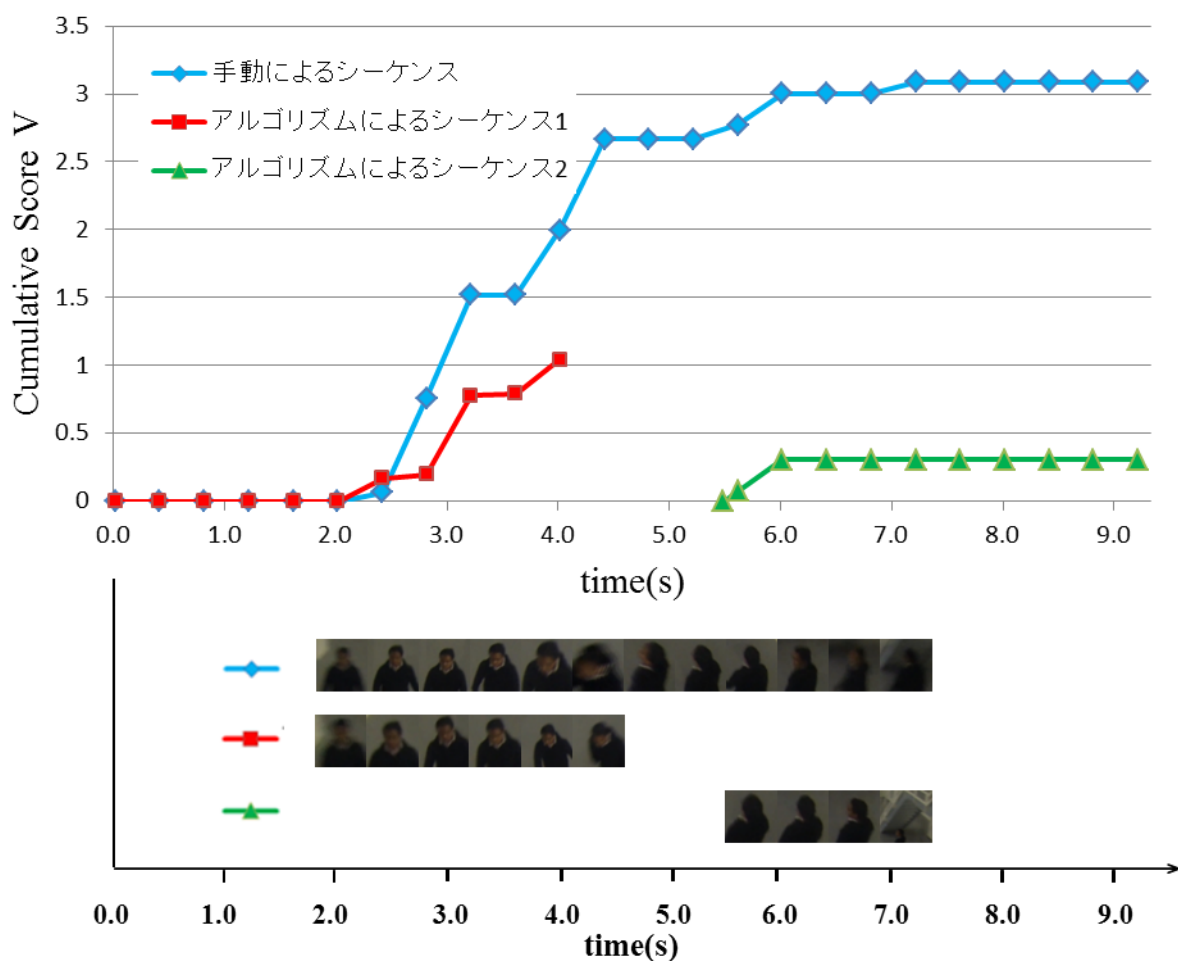


図4. 6: 歩行パターン1, ID1の, 各シーケンスによる照合値の変化とそのときの入力画像 (入力画像は紙面表示の都合上, 4画像ごとに抜粋)

次に、図4.7に、テンプレートID 1の歩行パターン2の時系列変化の例を示す。こちらは図4.6とは対照的に、手動による照合値は2.12で、人物検出アルゴリズムによる結果は2.18であり、ほとんど変化はなかった。撮影された入力画像は正面付近の画像が大半で、人物検出アルゴリズムにより、分割はされたが、情報量は分割されなかったためと考えられる。

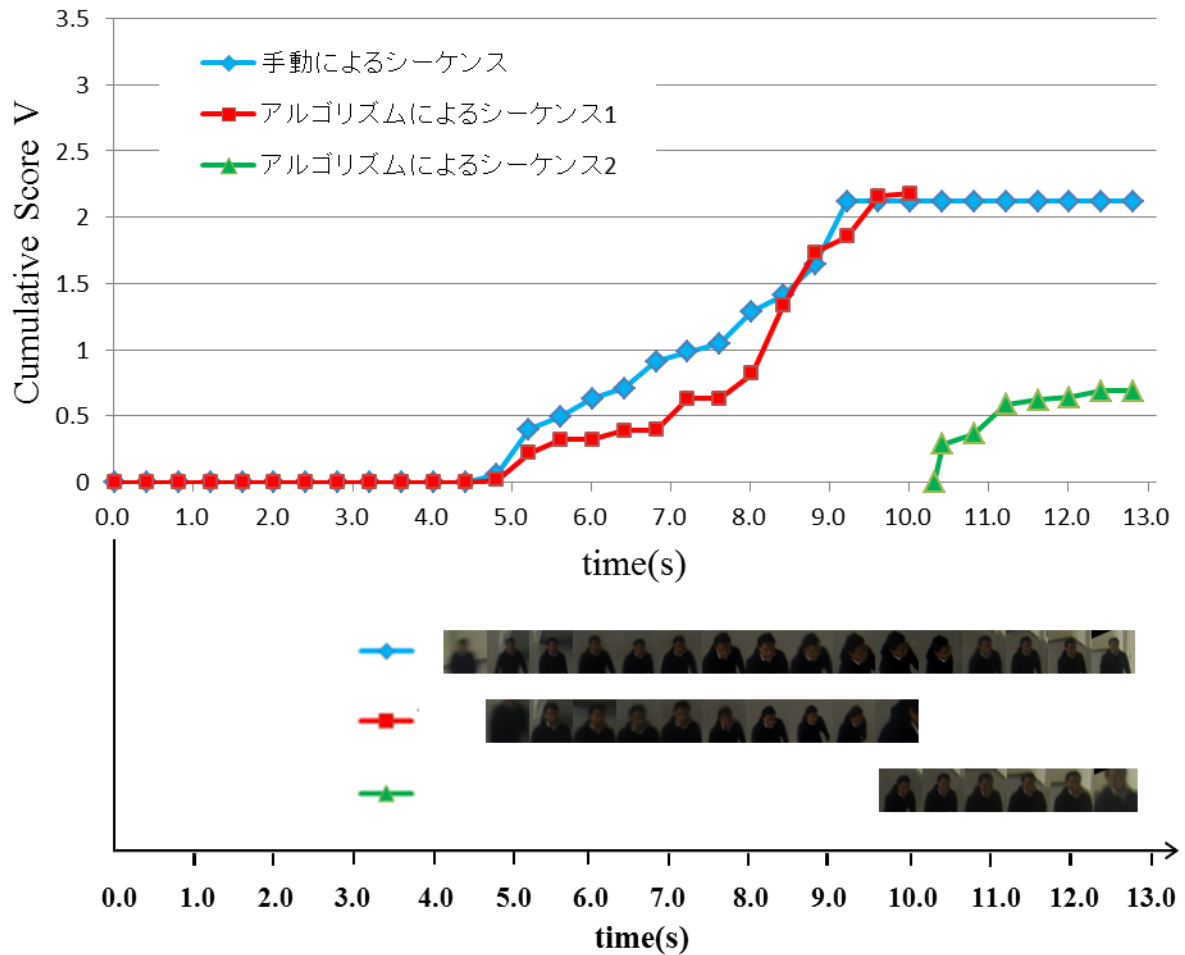


図4.7: 歩行パターン2, ID1の, 各シーケンスによる照合値の変化とそのときの入力画像 (入力画像は紙面表示の都合上, 4画像ごとに抜粋)

4.4 リアルタイム処理実験

本研究の目的である人物検索（迷子探索や犯人検索）においては、取得した映像を処理し認証結果を出力するまでのタイムラグは短ければ短いほどよい。なぜならば、映像を取得し認証処理している間にも人物は移動するため、タイムラグが大きければ、飛行ロボットが発見の通報をしたときにはその人物は既にその場所にいない可能性がある。逆にリアルタイムに処理できれば、発見と同時にそのまま追跡も可能である。よって、カメラから映像を取得しながら、その映像を逐次認証アルゴリズムに適用し、認証結果を出力するリアルタイム処理アプリケーションを作成した。そのアプリケーション画面を図4.8に示す。

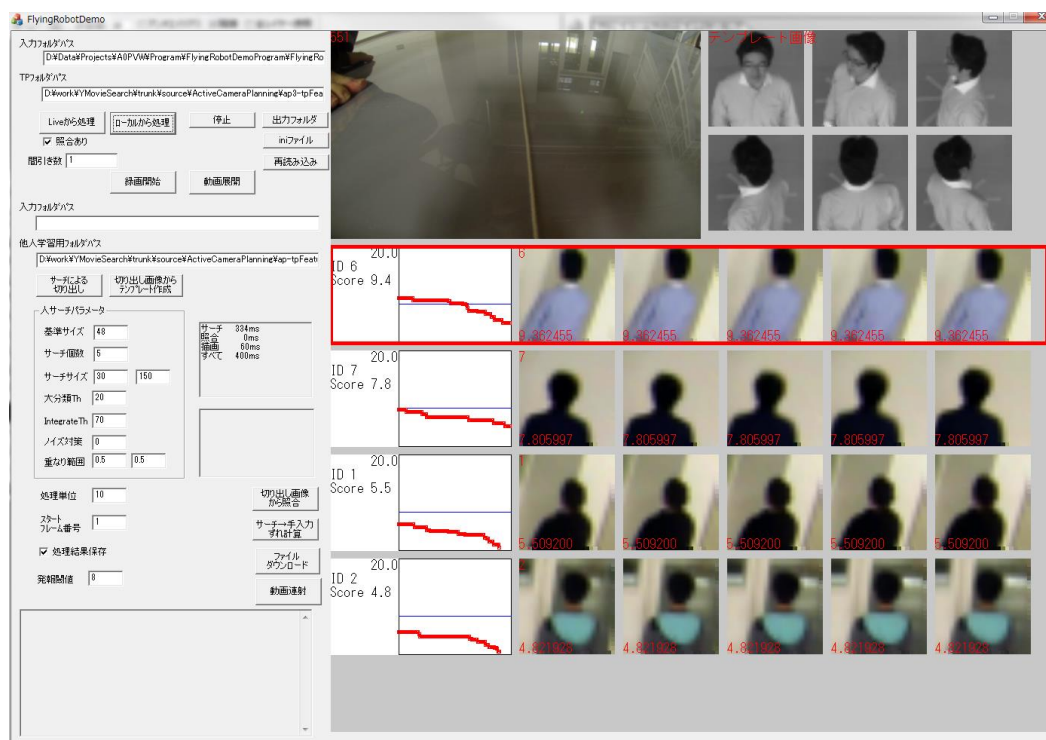


図4.8: リアルタイム処理アプリケーション画面

処理時間に関わる主な要素は、人物検出アルゴリズムと人物照合アルゴリズムであり、それにシステム関連の処理（カメラからPCへの画像転送等）が加わる。アルゴリズムの処理時間は、PCの処理能力や設定したパラメータによって大きく変化する。本実験では、人物検出アルゴリズムのマルチスレッド化と照合アルゴリズムのパイプライン処理化を行い、処理時間の短縮を行った。リアルタイムで処理できるかどうかは、人物検出アルゴリズムのパラメータ、照合アルゴリズムのテンプレート画像数、PCの処理能力など多数の要素が関係する。表4.6にリアルタイムで処理可能であったパラメータ等の項目表を示す。

表4.6: 使用PC, アルゴリズム設定パラメータ等

PC	機種	Panasonic CF-SX3
	CPU	Intel Corei5 1.6GHz
	スレッド数	4
画像	入力画像サイズ	320 × 240
	フレームレート	3
	写っている人物数	2~3人
人物検出パラメータ	サーチ個数	5
	サーチサイズ	30~100
	大分類閾値	20
	最終照合値閾値	70
人物認証	テンプレート画像数	24

1スレッドでの人物検出アルゴリズムの処理時間については、2.6.4で述べたように、表4.6と同パラメータで約330ms (3fps) である。人物認証アルゴリズムの処理時間は、テンプレート24画像、入力1画像あたり、約100msである。人物検出アルゴリズムのマルチスレッド化と人物認証アルゴリズムのパイプライン処理化を行うことで、全体として3fpsで処理することが出来た。尚、本アプリケーションを実装したPCとカメラに依存したことではあるが、カメラとの無線通信処理により、入力画像を認証アルゴリズムに入力するまでに、1~2秒程度の遅延がある。その結果、全体としての出力も1~2秒遅延した。

アルゴリズム部分で、最も処理負荷が大きいのが、人物検出アルゴリズムである。さらに高フレームレートでのリアルタイム処理を行うならば、人物検出アルゴリズムを高速化する必要がある。また、入力フレームすべてに対して人物検出は行わず、追跡のみを行うことで高フレームレート処理が可能になる。しかし、認証精度は追跡の精度に影響を受ける。処理時間の確認として、10fpsの入力映像で、3フレームごとに人物検出を行い、その他のフレームは追跡のみを行うことで、リアルタイムでの処理が可能であることは確認した。

4.5 まとめと考察

飛行ロボットが撮影したデータに対して、人物検出アルゴリズムと追跡アルゴリズムを適応し、シーケンスデータを作成した。そのシーケンスデータは、誤検出シーケンスデータや、1人の人物に対して重複、分割したシーケンスデータが含まれたものである。そのシーケンスデータを用いて、人物検索実験を行った。人物が静止状態の映像に対する結果は、手動によるシーケンスの結果と同等であることが分かった。また、人物が移動状態の映像に対しても、キャッチ数や誤キャッチ数は、ほぼ同等の結果を得ることができた。さらに、カメラから映像を入力し、アルゴリズムの一連の流れ（人物検出アルゴリズム、追跡アルゴリズム、認証アルゴリズム）を行うリアルタイムアプリケーションを作成し、3fpsで処理できることを確認した。よって、飛行ロボットが画像を撮影し、人物認証を行う部分については、システムとしての実現性を示すことが出来た。

人物検出アルゴリズムを用いたシーケンスデータによる人物認証精度面について考察する。1人の人物のシーケンスデータの分割に関しては、本実験データにおいては認証精度への影響度合いは小さかったが、認証精度が低下する要因の1つである。分割しないための方策としては、追跡アルゴリズムの精度を上げること、フレームレートを上げることが考えられる。追跡アルゴリズムの精度については、本実験では基本的な追跡アルゴリズムを適用したが、追跡アルゴリズムについては古くから数多くの研究[1]があり、それらを適用することで精度向上の可能性がある。フレームレートを上げることについては、フレームレートが上がることで人物画像の変化度合いが小さくなり、追跡しやすくなることを狙ったものである。しかし、フレームレートを上げると、それを処理するための処理量も増える。人物検出、追跡等のバランス調整が必要と考えられる。誤検出のシーケンスデータに関しては、本認証アルゴリズムの認証精度に関してはほとんど影響がなかった。よって、精度面においては、1人の人物の画像を1つのシーケンスデータとしてまとめることが最も重要であるといえる。且つ、異なる角度の人物画像を多く含むことが望ましい。そのためには追跡アルゴリズムの精度向上に加えて、高フレームレートの処理が必要になると考えられる。

また、本実験環境は、想定する遊園地やショッピングモールと比較すると、地面にほとんど模様がない点が異なると考えられる。この違いによって受ける影響は、人物検出時に誤検出が増えることである。しかし、図4.5の結果から分かるように、人物認証には誤検出による影響はほとんどない。そのため、異なる環境でも同程度の認証精度が得られると考えられる。

処理時間については、リアルタイムに結果が出力できるアプリケーションを作成したが、本実験ではCPUのマルチスレッド機能を利用するに留まった。人物検出アルゴリズムは画素間で処理が独立していることから、並列処理が得意なGPUを用いれば、さらに高速化が可能と考えられる。また、本実験では、背景を誤検出したシーケンスが数多く出力された。これらは認証精度には影響しなかったが、認証の処理時間はその分掛かるので、誤検出データは少ないほうがシステムとしての負担は小さくなる。そのため、やはり人物検出アルゴリズムについても高精度であることが望ましい。

第5章 人物認証のための飛行ロボットの軌道生成

5.1 はじめに

本章では、飛行ロボットによる人物検索を行うための軌道生成について述べる。本研究では、上空から人物映像を撮影し、アピアランス認証アルゴリズムに適用することで人物検索を行うことが目的である。第2章、第3章、第4章では、その飛行ロボットが撮影した映像から人物認証を行う手法、実験結果について述べた。本章では、その人物認証アルゴリズムを用いたときに、飛行ロボットがどのように人物画像を撮影すれば、効率よく、且つ、高精度に認証を行えるかについて述べる。飛行ロボットはホロノミックな移動体であるので、すべての方向に移動することができる。そのため、その軌道は無限大に存在する。効率的な人物認証を行うために、無限大に存在する軌道の中から累積アピアランス認証アルゴリズムに適した軌道を見つけ出すことが本章の目的である。複数の認証対象者を効率よく、且つ、精度よく認証するための軌道生成手法を提案する。累積アピアランス認証アルゴリズムはあらゆる角度から撮影された人物画像を用いて認証を行い、その角度や組み合わせによって認証精度が変わる。本手法はその特性を考慮した軌道生成を行うことで、認証精度が高くなる軌道を生成する。人物認証のための飛行ロボットの複数軌道のイメージ図を図5.1に示す。

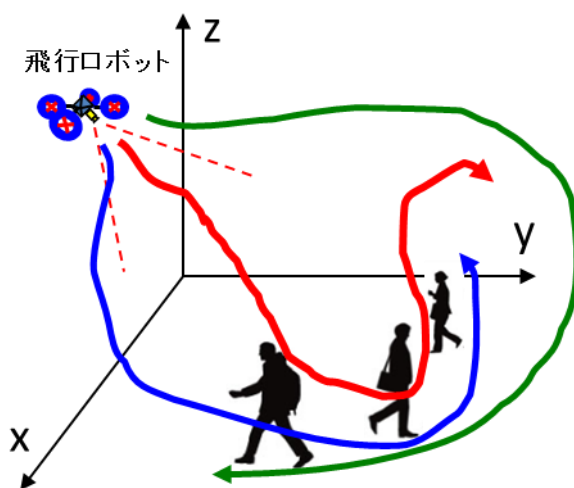


図5.1: 人物認証のための飛行ロボットの複数の軌道

5.2 関連研究

ロボットによる人物捜索に関連する研究として、Royらによるもの[1]、Stifflerらによるもの[2]、Mottaghiらによるもの[3]などがある。これらは人物の存在確率密度分布をもとにロボットが人を捜す問題を解いている。また人物捜索と類似する人の見守りの研究に、Takemura[4]によるものがある。環境中にいる人を見続けるための複数のアクティブカメラの視線を計画している。これらの研究では人物の存在確率に基づいてロボットの経路やカメラの動きの計画を行っている。本研究でも複数人の存在確率に基づいて軌道を作成する。このとき、人物の認証精度を考慮した方法を提案する。特に、人物の動きを予測し、同じ経路で複数人を異なる角度で撮影し照合値を高くする軌道を生成する。人物の動きとロボットの軌道に関する研究については、人物をトラッキングする軌道を求める研究として、レーザースキャナーによるものが数多く研究されている[5][6]。RFIDを用いたものとしてはGermaらによるもの[7]がある。また、移動ロボットが歩行者正面に回り込む軌道を生成する研究として、筆者ら[8]や関根[9]によるものがある。これらにより人物の背面を追跡する、または正面に回り込む軌道を生成することは可能である。しかし、著者らは飛行ロボットによる複数人物を効率よく認証するための軌道を求めることが目的である。そのため人物認証アルゴリズムの特性を生かした軌道を生成する必要がある。

著者らはその累積アピランス照合アルゴリズムを用いたときに、精度よく照合を行うことが出来る軌道を生成する手法を提案する。その特徴は、複数人の認証対象者を効率よく認証できる軌道である。また、捜索対象者の画像、すなわちテンプレート画像の角度による認証しやすさの違いを考慮した軌道である。認証しやすさとは、画像による他人との見分けやすさのことであり、一般的には正面方向から撮影した画像が最も認証しやすいと考えられる。しかし、背面に特徴がある模様の服を着ている場合には、背面方向のほうが他人と分離しやすい可能性もある。

5.3 飛行ロボットによる自動人物捜索システム

本節では、自動的に人物捜索を行うための飛行ロボットの構成図を述べる。飛行ロボットによる自動人物認証のイメージ図を図5.2に示す。本研究では、自動的に人物捜索を行うためには、俯瞰用ロボットと認証用ロボットの2台が必要であると考え、俯瞰用ロボットが撮影した映像から、捜索対象エリアにいる複数人の人物の情報（位置、速度、向き）を把握し、その情報を認証用ロボットに渡す。認証用ロボットはこの情報にもとづき、認証対象者全員の認証率を高める軌道を生成する。その軌道に沿って数秒間撮影した後、人物照合アルゴリズムを適用する。このとき人物の位置や向きは時間経過で変化する。その人物の動きに対応するために、数秒間撮影後（本論文の実験では5秒間）、俯瞰用ロボットは再度人物の情報を取得し、

認証用ロボットが軌道を生成する．これを繰り返すことで人物認証を完了させる．図 5.3 に人物検索フローを示す．

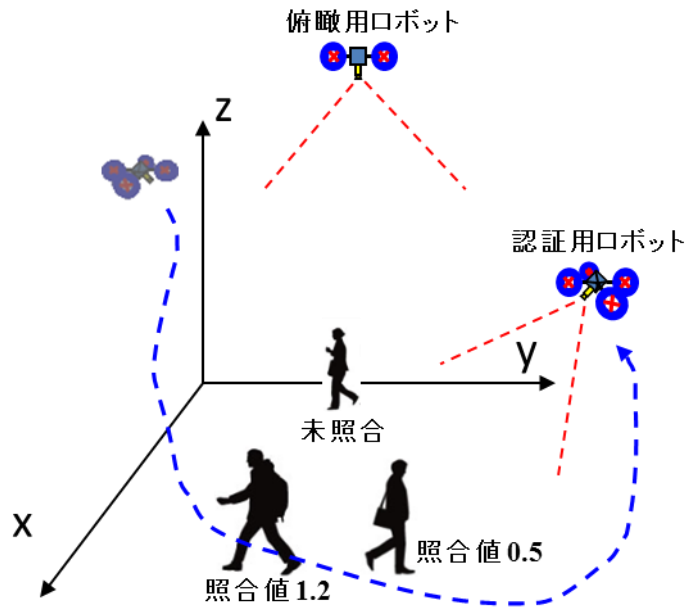


図5.2: 飛行ロボットによる自動人物認証

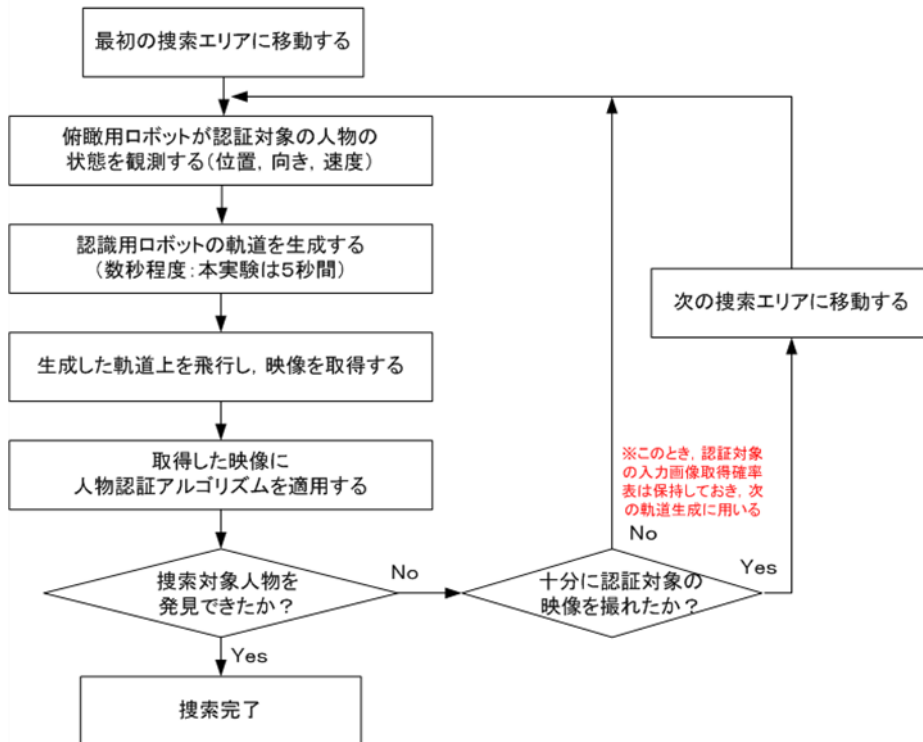


図5.3: 飛行ロボットによる人物検索フロー

俯瞰用ロボットの役割は、検索対象エリアの人物の位置の把握と追跡、認証結果の結びつけである。認証用ロボット単体では、一度認証した人物がフレームアウトすると、次に画像内に写った時にその人物は認証行為が終了しているのか判断できず、検索対象エリア内の認証をし続けることとなる。俯瞰用ロボットは、認証用ロボットによる認証結果と人物を結びつけ、検索対象エリア内の全員の人物の認証終了を判断する。尚、各処理の演算については、各ロボットが行う必要はなく、各ロボットは映像のみを取得し、映像を演算用 PC に送信し、演算用 PC で演算してもよい。

俯瞰用ロボットによる人物の情報（位置、向き、速度）の取得については従来技術を適用することで実現可能である。俯瞰用ロボットによる人物の位置検出は、第2章で述べた人物検出システムや、従来技術のHog特徴量による人物検出アルゴリズム[10] を上空から撮影した人物画像で学習したものが適用できる。また、赤外カメラを用いた従来技術[11]も研究されている。認証用ロボットが撮影した映像内からの人物位置検出についても同様である。また、向きや速度についても、従来技術を用いて、取得した数フレーム間の映像からトラッキング[12]を行い取得することができる。よって本研究では、認証用ロボットの、人物認証のための軌道生成手法に注力する。

尚、本論文で対象とする認証対象者は人物照合アルゴリズムの性質上、直立歩行、もしくは静止状態を対象としている。また、テンプレート画像の服装と入力画像の服装が同じときを対象としている。

5.4 軌道生成手法-概要

5.4.1 軌道生成手法の考え方

本研究では、複数の人物の人物認証精度を最も高める軌道を生成する。本節では、その人物認証精度に影響を与える要素について述べる。最も影響を与える要素は、飛行ロボットが撮影する入力画像と比較する検索対象のテンプレート画像である。テンプレート画像は第N章で述べたように、複数角度から撮影された複数枚の人物画像である。本研究の累積アピランス照合アルゴリズムでは、それぞれの角度によって、照合値の上がり方が異なる。一般的には、人物画像の背面方向よりも正面方向に他人と分ける特徴があると考えられる。その場合は、飛行ロボットは正面方向に回り込み、照合を行えばよいと考えられる。しかし、背面方向の服に特異な模様があったとした場合、正面方向よりも背面方向のほうが他人と分けるための特徴が多いこともあり得る。そのため、第3章で述べたように、累積アピランス照合アルゴリズムでは、各テンプレート画像に重み係数を掛け、正規化した演算を行うことで、各テンプレート画像の他人との分離しやすさに対応した。この他人との分離しやすさ、すなわちテンプレート画像に対応する評価値 $\alpha_i \cdot G_i$ の上がりやすさは、検索対象となるテンプレートによって異なる。式(5. 1)～式(5. 6)に、第3章 (3.3.2) で述べた累積アピランス照合アルゴリズムの評価式を再掲する。

$$V = \sum_{i=1}^N (\alpha_i \cdot G_i) \quad (5.1)$$

$$G_i = \max_k (S'_{i,k}) \quad (5.2)$$

$$S'_{i,k} = \frac{S_{i,k} - \bar{m}_i}{\sigma_i} \quad (5.3)$$

$$S_{i,k} = f_{\text{verify}}(T_i, I_k) \quad (5.4)$$

$$\alpha_i = \frac{1}{\sum_{j=1}^N R_{i,j}} \quad (5.5)$$

$$R_{i,j} = f_{\text{verify}}(T_i, T_j) \quad (5.6)$$

式(5.1)で示したように、 V が累積照合値であり、評価値 $\alpha_i \cdot G_i$ の合計値である。重み係数 α_i 、評価値 G_i のどちらも、対応するテンプレート画像 T_i により決められる。2人の実験対象者のテンプレート画像を用いて、式(5.4)の画像照合関数 $f_{\text{verify}}(T_i, I_k)$ に本人が入力された場合の平均的な照合値0.75を代入したときの各 $\alpha_i \cdot G_i$ をヒストグラムで表したものを図5.4、図5.5に示す。

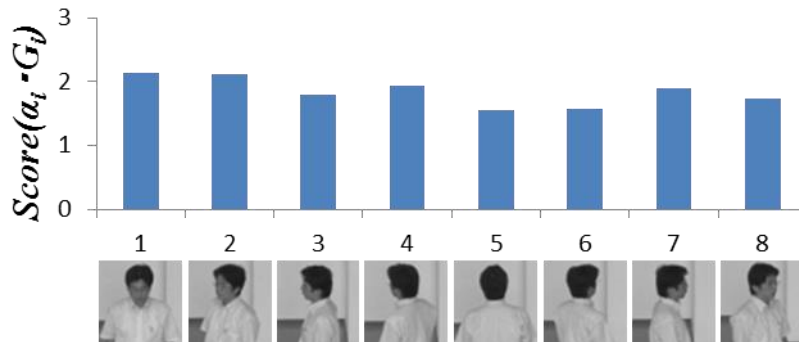


図5.4: テンプレート画像と照合値の上がりやすさを示したヒストグラム
(実験対象者A)

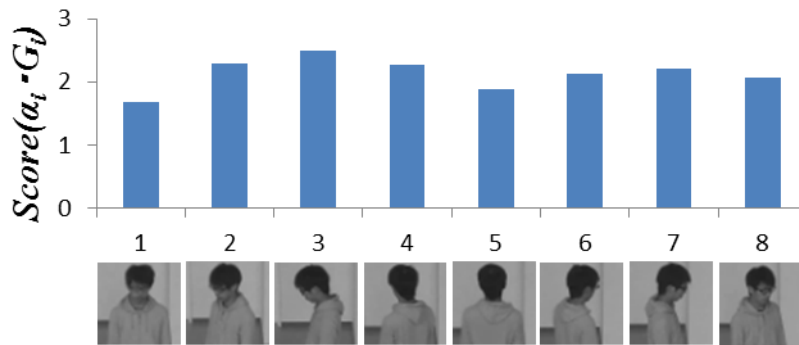


図5.5: テンプレート画像と照合値の上がりやすさを示したヒストグラム
(実験対象者B)

図 5.4, 図 5.5 で示した値は, 各テンプレート画像に対応する入力画像が得られたときの, 照合値の上がりやすさを示している. 図 5.4 の実験対象者 A は, 背面方向の入力画像よりも, 正面方向の入力画像を入力したほうが, 照合値が上がりやすい. 一方実験対象者 B は, 背面と正面は同等である. 検索対象となるテンプレートの人物画像の角度によって, 照合値の上がりやすさは異なる. 本論文で述べる軌道生成手法は, このテンプレートごとの照合値のあがりやすさを考慮した軌道生成手法である.

よって, 検索対象者 (テンプレート) を決めれば, どの方向の検索対象者の人物画像を取得すれば照合値が上がりやすいかが分かる. 飛行ロボットはその照合値の上がりやすさを参照しながら, テンプレート画像の人物角度に対応した人物画像を取得できるように移動すればよい.

ある軌道上を飛行ロボットが飛行したときに, 俯瞰用ロボットが取得した検索エリア内の人物の位置や向きから, 得られる入力画像内にどのような角度とサイズで人物画像が写っているかが推定できる. それが推定できると, その入力人物画像群が検索対象者だったとした場合に, テンプレート画像と照らし合わせることで, 得られる照合値も推定できる.

複数の軌道を生成し, この推定照合値を求め, 最も高い推定照合値が得られる軌道が求める軌道である. ただし, 認証対象者が複数存在する場合は, その推定照合値が平均的に高いものが求める軌道である. 次節以降で, その詳細について述べる.

5.4.2 軌道生成手法概要

本軌道生成手法は, 複数の軌道候補 $c_j (j \in \{1, \dots, J\})$ を作成し, そのすべての軌道候補を軌道評価関数 f_{es} に適用し, その評価値の最大値を得る軌道を求める軌道 R とした. 人物の位置や向きは時間経過で変化することに対応するため, 数秒間隔でこの軌道生成を行う. 作成する軌道は数秒間であるので, 複雑な動きを想定せず, ロボットの初期位置から移動可能な複数の軌道候補群を放射上に作成した. 求める軌道 R は以下の式で作成した.

$$R = \arg \max(f_{es}(c_j)) \quad (5.7)$$

式(5.7)の軌道評価関数 f_{es} は、入力軌道 c_j を通ったときに、認証対象者全員が搜索対象だと想定し、取得できると推定される照合値の平均的な高さを評価する関数である。こうすることで、搜索対象者が認証対象者の中のどの人であっても、搜索できる確率が上がる。その評価関数は式(5.8)で表される。

$$f_{es}(c_j) = \frac{\overline{E(\hat{V}_u)}}{\sigma_{E(\hat{V}_u)} + 1} \quad (5.8)$$

$\overline{E(\hat{V}_u)}$ は各認証対象者 h_u の推定照合値 \hat{V}_u の期待値 $E(\hat{V}_u)$ の平均値であり、 $\sigma_{E(\hat{V}_u)}$ はその標準偏差である。推定照合値 \hat{V}_u とは、飛行ロボットが軌道 c_j 上を飛行し照合を行うと仮定したときに、人物 h_u が搜索対象人物だった場合に得られると推定される人物 h_u に対する照合値である。本手法においては、推定照合値 \hat{V}_u は確率分布として求められるため、その期待値 $E(\hat{V}_u)$ を評価関数に用いる。次にその推定照合値とその期待値について述べるが、分かりやすさのため、認証対象者を1人とした場合について説明を行う。

推定照合値の期待値は式(5.9)により求める。

$$E(\hat{V}) = \sum_{l=1}^L (P_{pattern,l} \cdot \hat{V}_l) \quad (5.9)$$

$P_{pattern,l}$ は l 番目の画像パターンが得られる確率である。 \hat{V}_l は l 番目の画像パターンの推定照合値である。画像パターンとは、飛行ロボットが取得する複数枚の画像の人物角度と人物画像サイズの組み合わせパターンのことである。飛行ロボットが軌道 c_j を飛行したときに取得する人物画像の角度やサイズは一意には決まらない。なぜなら、時間 t における認証対象者 h の位置や向きが不明だからである。そこで本手法では、時間 t における認証対象者 h の位置と向きを予測し、存在確率分布 P_h として表した。こうすることで、指定した画像パターンを取得する確率を求めることができる。飛行ロボットが軌道 c_j 上を飛行したときに、取得する可能性のある画像パターンを求め、その画像パターンごとに推定照合値と確率を計算する。式(5.9)の $E(\hat{V}_u)$ はそれらの期待値となる。

画像パターン群 A を以下に定義する。

$$A = \{A_1, A_2, \dots, A_l, \dots, A_L\} \quad (5.10)$$

$$A_l = \{a_{l,1}, a_{l,2}, \dots, a_{l,k}, \dots, a_{l,F}\} \quad (5.11)$$

各画像パターン A_l は番号の配列で表され、各番号 $a_{l,k}$ は入力画像の状態番号である。状態番号とは、入力画像に対応するテンプレート画像 T_i の番号と、人物画像のサイズを表す番号を組

み合わせた番号である．式(5.9)，式(5.10)の L はそのパターン数である．式(5.11)の F は取得フレーム数である．

画像パターン A_l の取得確率 $P_{pattern,l}$ は式(5.12)により求める．

$$P_{pattern,l} = \prod_{k=1}^F P_{image,k,a_{l,k}} \quad (5.12)$$

式(5.12)の P_{image} は人物画像取得確率である．撮影番号 k に対応する時間において，画像パターン番号 $a_{l,k}$ に対応する人物角度とサイズの画像を取得できる確率を示している． P_{image} は軌道 c_j と人物の位置と角度の確率分布 P_h から求めることができる．詳細は後に述べる．

推定照合値 \hat{v}_l は，画像パターン A_l に対応した人物角度とサイズの本人画像が入力されると想定し，本人画像が入力されたときの平均的な照合値を式(5.1)，式(5.2)，式(5.3)に代入し，照合値を求めることで得られる．具体的には，画像パターン番号 $a_{l,k}$ に対応するテンプレート画像を T_q とすると，式(5.15)に示すように，式(3)に本人画像が入力されたときの画像照合関数 f_{verify} の平均的な結果 \bar{s} を入力する．

$$\hat{V} = \sum_{i=1}^N (\alpha_i \cdot \bar{G}_i) \quad (5.13)$$

$$\bar{G}_i = \max_k (\bar{S}'_{i,k}) \quad (5.14)$$

$$\bar{S}'_{i,k} = \frac{\bar{S} \cdot \beta_{q,i} - \bar{m}_i}{\sigma_i} \quad (5.15)$$

このとき，テンプレート画像 T_q に類似するその他のテンプレート画像に対する照合値を考慮し，式(5.6)の重み $\beta_{q,i}$ を掛けて，全テンプレート画像 $T_i (i \in \{1, \dots, N\})$ に対し式(5.15)を行う．これをすべての画像パターン A_l の数 ($k \in \{1, \dots, F\}$) だけ繰り返し，式(5.14)，式(5.13)により求めた結果が，推定照合値 \hat{v}_l である．尚， \bar{S} はあらかじめ同一環境にて複数人のデータを取得し，演算することで求めておく．

以上のようにして，期待値 $E(\hat{V})$ は求められる．この演算を各認証対象者 h_u の人物数分繰り返し，期待値 $E(\hat{V}_u)$ を求め，式(5.8)に適用することで，軌道 R を求めることができる．

5.5 軌道生成手法-詳細

軌道生成の流れについて述べる。図5.6に軌道生成のフローを示す。

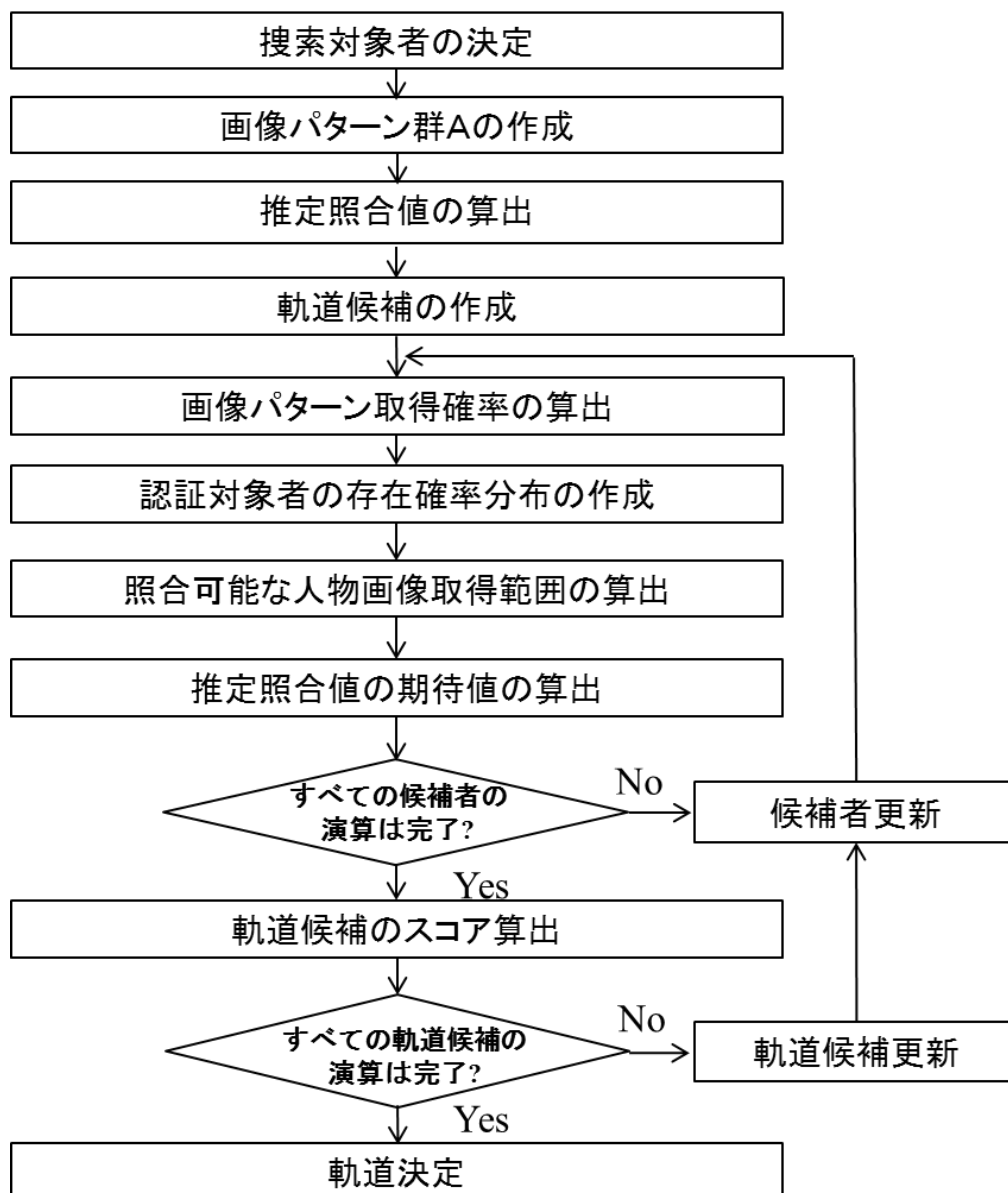


図5.6: 軌道生成フロー

最初に検索対象を指定し、テンプレート画像 T を決定する。テンプレート画像が決まれば、照合に影響する入力画像の人物角度が決まるため、画像パターン群 A を作成する。画像パターン群 A が作成できれば、各画像パターン A_l から推定照合値 \hat{V}_l も求められるので、軌道生成前にあらかじめ求めておく。画像パターン群 A とそれに対応する推定照合値を求めた後、軌道生成を開始する。軌道生成は、まず軌道候補 c_j を作成する。その後、各軌道候補 c_j に軌道評価関数 f_{es} を適用するために必要な演算を行う。尚、各軌道候補 c_j に掛かる時間はすべて同じであるとする。そのためすべての軌道候補 c_j において画像取得数 F が同じとなる。次に、各軌道候補 c_j 、各認証対象者 h_u ごとに期待値 $E(\hat{V})$ を求める。そのために、各画像パターン A_l ごとに画像パターン取得確率 $P_{pattern,l}$ と推定照合値 \hat{V}_l を求める。画像パターン取得確率 $P_{pattern,l}$ は式(5.12)に示したように、撮影番号 k に対応した時間ごとに人物画像取得確率 P_{image} を求めることで得られる。人物画像取得確率 P_{image} は、撮影番号 k に対応した時間における認証対象者の存在確率分布 P_h と、その時間の軌道候補 c_j の撮影範囲を求め、それを組み合わせることで取得する。このようにして、軌道生成を行う。これらの演算の詳細について以下の5.4.1～5.4.6で述べる。

5.5.1 画像パターン群 A の作成

飛行ロボットが取得する入力画像は、さまざまな角度とサイズの人物である可能性があり、また、写っていない可能性もある。著者らはそれらの入力画像の状態を番号化した。1つの要素は、入力画像の人物角度に対応するテンプレート画像番号である。もう1つは、得られる人物画像サイズである。累積アピランス照合アルゴリズムの照合値は、入力画像の人物サイズによって変化する。よって、サイズを一定の範囲で分割した。テンプレート画像数を N 、サイズ分割数を S とすると、入力画像の状態番号の総数 L は、

$$L = N \cdot S + 1 \quad (5.16)$$

である。入力画像に人物が写っていないことや、入力画像に対応するテンプレート画像がないことも入力画像の状態の1つとした。先に述べたように、どの軌道候補 c_j でも取得フレーム数が F なので、 L の F 乗個の入力画像取得パターンが存在する。そのパターン番号を l とし、式(5.10)、式(5.11)で示したように、画像パターン群 A 、各画像パターンを A_l と表した。

5.5.2 推定照合値の算出

画像パターン群 A を作成すると、各画像パターン A_l に対応する推定照合値を求める。(式(5.13)、式(5.14)、式(5.15)) 本人が入力画像である場合の画像照合関数 f_{verify} の平均的な結果 \bar{S} は、入力される人物画像サイズで異なる。そのため、本論文では、人物画像の上半身のサイズ

を次のように4つのサイズで分割した。サイズ番号を s とすると, $s=1$ のとき 46pixel 以上, $s=2$ のとき $45\sim 23\text{pixel}$, $s=3$ のとき $22\sim 15\text{pixel}$, $s=4$ のとき $14\sim 10\text{pixel}$ である。そして, 各人物画像サイズに平均的な結果 \bar{S}_s ($s=1,2,3,4$) を求め, 以下のように設定した。

$$\bar{S}_1 = 0.75, \bar{S}_2 = 0.72, \bar{S}_3 = 0.68, \bar{S}_4 = 0.65 \quad (5.17)$$

この値は, 本論文の実験環境において, あらかじめ複数人の照合結果を求めその平均値から算出した。

5.5.3 軌道候補の作成

軌道候補 c_j の求め方について説明する。本研究では, あらゆる機動性能の飛行ロボットに適用できるように複雑な軌道候補を作成せず, 認証用ロボットの初期位置から直線, または円弧で放射上に複数の候補を作成した。また, 安全性の面から, 軌道候補の中で人物に $3(\text{m})$ 以内に接近するものは候補から削除した。図5.7にそのイメージ図を示す。図5.7は上空から (Z 軸方向) 見下ろした図であり, 軌道候補ごとにランダムな色で表示した。軌道候補については, 用いる飛行ロボットの機動性能によっては, より複雑な軌道候補も設定可能である。

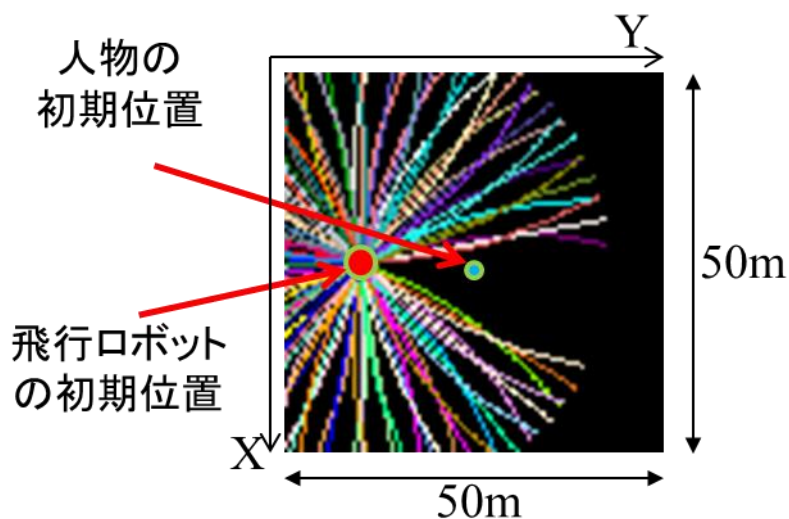


図5.7: 軌道候補

5.5.4 画像パターン取得確率の算出

画像パターン取得確率は各撮影番号 k に対応する時間 t における人物画像取得確率 P_{image} のすべての積から求める。人物画像取得確率 P_{image} は式(5.18)から求める。

$$P_{image,k,a_{i,k}} = P_{I,f_{ime}(k),f_{template}(a_{i,k}),f_{size}(a_{i,k})} = P_{I,t,i,s} \quad (5.18)$$

$$t = f_{ime}(k), i = f_{template}(a_{i,k}), s = f_{size}(a_{i,k}) \quad (5.19)$$

$$P_{I,t,i,s} = \int_{\theta'_i} \left(\int_{E_{i,s,t}} P_h(h_t) dx dy \right) d\theta \quad (5.20)$$

$$h_t = (x, y, \theta) \quad (5.21)$$

$$\theta'_i = f_{dir}(r_t, \theta_i) \quad (5.22)$$

$$\theta_i : \theta_{tp_azi,i} - \theta_{margin_azi} \leq \theta_i \leq \theta_{tp_azi,i} + \theta_{margin_azi} \quad (5.23)$$

$$r_t = \{ rx, ry, rz, roll, pitch, yaw \} \quad (5.24)$$

$P_{I,t,i,s}$ は、時間 t に、テンプレート番号 i に対応するサイズ s の人物画像を撮影する確率を示している。 f_{ime} は撮影番号 k に対応する時間 t を返す関数で、 $f_{template}$ は状態番号 $a_{i,k}$ から対応するテンプレート画像番号 i を返す関数で、 f_{size} は状態番号 $a_{i,k}$ から対応するサイズ番号 s を返す関数である。 $P_{I,t,i,s}$ は式(5.20)により求められる。 h_t は式(5.21)に示した人物の持つパラメータで、 $P_h(h_t)$ は認証対象者 h が時間 t に、座標 x, y , 方位角 θ の向きで存在する確率であり、詳細は次節で述べる。 $E_{i,s,t}$ は x, y 平面上の時間 t において、テンプレート番号 i , サイズ s の状態の入力画像を撮影できる範囲を示している。これは、テンプレート画像 T_i に写っている人物の仰角とロボットの状態 r_t から求める。 θ'_i はテンプレート画像 T_i と照合できる入力画像の人物の角度範囲（方位角）を示している。これは、テンプレート画像 T_i を照合できる範囲 θ_i (式(5.23))を、関数 f_{dir} によってフィールド上でのテンプレート画像 T_i を取得できる角度範囲に変更したものである。これは飛行ロボットの向き yaw 角(式(5.24))によって変わるからである。式(5.23)の $\theta_{tp_azi,i}$ はテンプレート画像 T_i の人物角度であり、 θ_{margin_azi} は照合できる角度の範囲である。式(5.24)は飛行ロボットの持つ座標 (rx, ry, rz) と方向 $(roll, pitch, yaw)$ のパラメータである。このように、フィールド上での、認証対象者の存在確率分布 P_h と、照合可能な人物画像の取得範囲 E から人物画像取得確率 P_{image} を求めることが出来る。人物画像取得確率 P_{image} を表に表したサンプルを図5.8に示す。

	$s=1$	$s=2$	$s=3$	$s=4$	other
$i=1$	0.02	0.05	▪	0.002	0.35
$i=2$	0.1	▪	▪	▪	
▪	▪	▪	▪	▪	
▪	▪	▪	▪	▪	
▪	▪	▪	▪	▪	
$i=N$	0.001	▪	▪	0.05	

図5. 8: 人物画像取得確率表

5.5.5 認証対象者の存在確率分布の作成

時間 t において、認証対象者 h が、ポジション x,y に存在し、水平角度 θ の方向を向き、認証できる正常状態（直立）である確率を求めたい。 $t=0$ （初期位置）のとき、俯瞰用ロボットから認証対象者 h の位置 x_0, y_0 、速度 v_{x_0}, v_{y_0} は得られるものとする。また、人物の体の向きは進行方向を向いているものとする。認証対象者 h の存在確率 P_h は以下の式で表す。

$$P_h(h_t) = P_p(h_t) \cdot P_o(h_t) \cdot P_s(h_t) \quad (5.25)$$

存在確率 P_h は、地図情報 P_o と認証対象者 P_s の状態と、人物位置と角度の確率 P_p で表されると考えた(式(5.25))。 h_t は式(5.21)で示したのと同じである。地図情報 P_o と認証対象者の状態 P_s は以下の式で表す。

$$P_o(h_t) = \{1:(passage), 0:(obstacle)\} \quad (5.26)$$

$$P_s(h_t) = \{1:(normal), 0:(irregular)\} \quad (5.27)$$

地図情報 P_o は、通路:1 や障害物:0 の通行可能性で表した。認証対象者の状態 P_s は、直立歩行状態のような認証できる状態:1 か、座っている、寝転んでいる等、認証出来ない状態:0 かで表した。

人物位置と角度の確率 P_p は以下の式で表す。

$$P_p(h_t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{len_p^2(h_t) + len_\theta^2(h_t)}{2}\right) \quad (5.28)$$

$$len_p(h_t) = \frac{\sqrt{(x - x_{center})^2 + (y - y_{center})^2}}{\sigma_{p,t}} \quad (5.29)$$

$$x_{center} = x_0 + vx_0 \cdot t \quad (5.30)$$

$$y_{center} = y_0 + vy_0 \cdot t \quad (5.31)$$

$$\sigma_{p,t} = \sqrt{\alpha^2 \cdot t} \quad (5.32)$$

$$len_\theta(h_t) = \frac{|\theta - f_\theta(x, y)|}{\sigma_{\theta,t}} \quad (5.33)$$

$$f_\theta(x, y) = \arctan\left(\frac{y - y_0}{x - x_0}\right) \quad (5.34)$$

$$\sigma_{\theta,t} = \sqrt{\beta^2 \cdot t} \quad (5.35)$$

時間 t における存在確率 P_p は、人物の位置と方向を変数とするガウス分布で表されると想定した。位置の確率については、基本的には人物は等速直線歩行をするものとし、等速直線歩行した位置を中心にガウス上にばらつくと考えた。角度の確率も同様に、進行方向の角度を中心にガウス上にばらつくと考えた。進行方向とは、人物の初期位置と入力座標 x, y を結んだ方向のことである。式(5.29)の len_p は等速直線歩行した場合の人物位置からの距離を $\sigma_{p,t}$ で正規化したものである。式(5.30)、式(5.31)の x_{center}, y_{center} は等速直線歩行したときの位置を表し、式(5.32)の $\sigma_{p,t}$ は時間 t における位置のばらつきである。 α は1秒あたりの位置のばらつき係数である。これは、時間経過と共にばらつきも大きくなると思ったからである。式(5.33)の len_θ は進行方向と入力角度 θ の角度差を σ_θ で正規化したものである。式(5.34)は進行方向を求める関数である。式(5.35)の $\sigma_{\theta,t}$ は時間 t における角度のばらつきであり、 β は1秒あたりの角度ばらつき係数である。図5.9に人物位置の存在確率のイメージ図を示す。

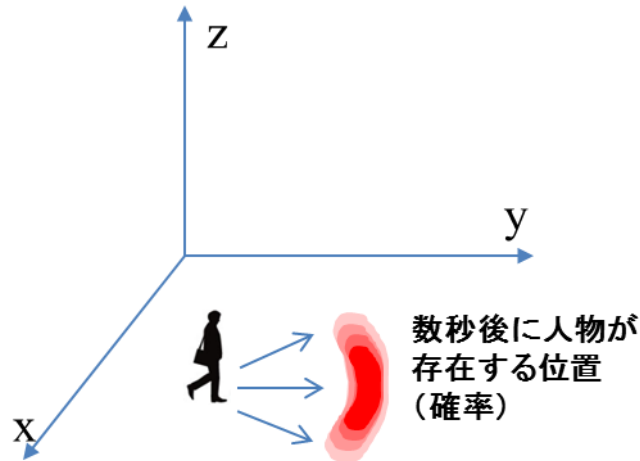


図5.9: 人物存在確率のイメージ図

5.5.6 照合可能な人物画像取得範囲

照合可能な xy 平面上のエリア E を求める。エリア E は3つのエリアの組み合わせから求める。1つ目は、カメラ撮影範囲で、これは認証用ロボットに搭載されたカメラで撮影できる範囲である。2つ目はテンプレート画像に映っている人物画像と同じ角度の人物画像を取得できる範囲である。3つ目は、照合可能な人物サイズが撮影できる範囲である。

5.5.6.1 カメラ撮影範囲

カメラ撮影範囲について、以下に説明する。時間 t において、認証用ロボットのカメラが撮影している範囲は、カメラの持っているイメージプレーンを透視投影変換して、 xy 平面上に射影することで求めることができる。認証用ロボットに搭載されたカメラの持つパラメータを以下に示す。

$$c_{param} = \{ccd_{wid}, ccd_{hei}, fd, image_{wid}, image_{hei}, dep\} \quad (5.36)$$

ccd_{wid} , ccd_{hei} はカメラのイメージプレーンの幅と高さであり、 fd は焦点距離である。 $image_{wid}$, $image_{hei}$ は取得画像サイズの幅と高さであり、 dep はカメラの俯角である。図5.10にその模式図を示す。

焦点位置を r_t 、カメラのイメージプレーンの4頂点を $im_{k,t}$ ($k=0,1,2,3$)、透視投影変換後の4頂点を $pro_{k,t}$ ($k=0,1,2,3$)とする。 $pro_{k,t}$ は一般的な透視投影変換で求めることが出来る。この4頂点 $pro_{k,t}$ に囲まれた範囲が求めるカメラ撮影である。この範囲を $E_{camera,t}$ とする。

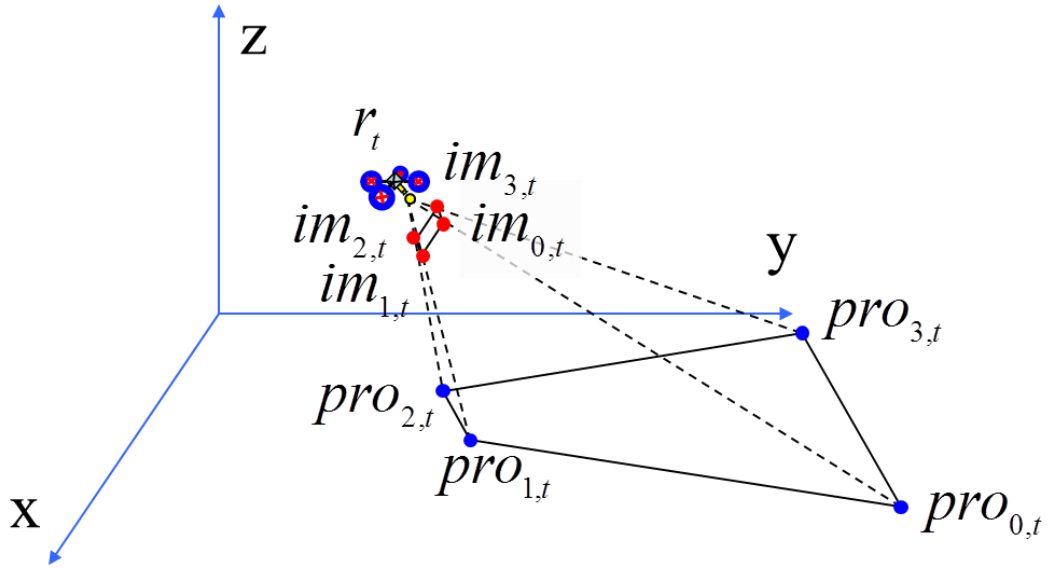


図5.10: カメラ撮影範囲のイメージ図

透視投影変換後の4頂点 $pro_{k,t}$ は以下の式で求めることが出来る．焦点位置の初期位置 r_0 と、イメージプレーンの初期位置 $im_{k,0}(k=0,1,2,3)$ は、以下のように設定する．

$$r_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (5.37)$$

$$im_{0,0} = \begin{pmatrix} im_{x,0,0} \\ im_{y,0,0} \\ im_{z,0,0} \\ 1 \end{pmatrix} = \begin{pmatrix} ccd_{wid} / 2 \\ fd \\ ccd_{hei} / 2 \\ 1 \end{pmatrix} \quad (5.38)$$

$$im_{1,0} = \begin{pmatrix} im_{x,1,0} \\ im_{y,1,0} \\ im_{z,1,0} \\ 1 \end{pmatrix} = \begin{pmatrix} ccd_{wid} / 2 \\ fd \\ -ccd_{hei} / 2 \\ 1 \end{pmatrix} \quad (5.39)$$

$$im_{2,0} = \begin{pmatrix} im_{x,2,0} \\ im_{y,2,0} \\ im_{z,2,0} \\ 1 \end{pmatrix} = \begin{pmatrix} -ccd_{wid}/2 \\ fd \\ -ccd_{hei}/2 \\ 1 \end{pmatrix} \quad (5.40)$$

$$im_{3,0} = \begin{pmatrix} im_{x,3,0} \\ im_{y,3,0} \\ im_{z,3,0} \\ 1 \end{pmatrix} = \begin{pmatrix} -ccd_{wid}/2 \\ fd \\ ccd_{hei}/2 \\ 1 \end{pmatrix} \quad (5.41)$$

時間 t における焦点位置 r_t は, 式(5.24)より,

$$r_t = M_t \cdot R_{\theta_{x,t}} \cdot R_{\theta_{y,t}} \cdot R_{\theta_{z,t}} \cdot r_0 \quad (5.42)$$

$$M_t = \begin{pmatrix} 1 & 0 & 0 & rx_t \\ 0 & 1 & 0 & ry_t \\ 0 & 0 & 1 & rz_t \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.43)$$

$$R_{\theta_{x,t}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta_{x,t} & -\sin \theta_{x,t} & 0 \\ 0 & \sin \theta_{x,t} & \cos \theta_{x,t} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.44)$$

$$R_{\theta_{y,t}} = \begin{pmatrix} \cos \theta_{y,t} & 0 & \sin \theta_{y,t} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta_{y,t} & 0 & \cos \theta_{y,t} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.45)$$

$$R_{\theta_{z,t}} = \begin{pmatrix} \cos \theta_{z,t} & -\sin \theta_{z,t} & 0 & 0 \\ \sin \theta_{z,t} & \cos \theta_{z,t} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.46)$$

$$\theta_{x,t} = pitch_t + dep \quad (5.47)$$

$$\theta_{y,t} = roll_t \quad (5.48)$$

$$\theta_{z,t} = yaw_t \quad (5.49)$$

と表される.

時間 t におけるイメージプレーンの4つの頂点位置 $im_{k,t}$ ($k=0,1,2,3$)も同様に,

$$im_{k,t} = M_t \cdot R_{\alpha,t} \cdot R_{\beta,t} \cdot R_{\gamma,t} \cdot im_{k,0} \quad (5.50)$$

と表される.

焦点位置 r_t とイメージプレーンの各頂点位置 $im_{k,t}$ ($k=0,1,2,3$)を結ぶ単位ベクトル $v_{k,t}$ を式(5.51)で表す.

$$v_{k,t} = \begin{pmatrix} v_{x,k,t} \\ v_{y,k,t} \\ v_{z,k,t} \end{pmatrix} \quad (5.51)$$

求める射影点 $pro_{k,t}$ は, ベクトル $v_{k,t}$ と xy 平面 ($z=0$) が交わる点である. よって,

$$pro_{k,t} = \begin{pmatrix} pro_{x,k,t} \\ pro_{y,k,t} \\ 0 \end{pmatrix} \quad (5.52)$$

$$pro_{x,k,t} = \frac{-r_{z,t} \cdot v_{x,k,t}}{v_{z,k,t}} + r_{x,t} \quad (5.53)$$

$$pro_{y,k,t} = \frac{-r_{z,t} \cdot v_{y,k,t}}{v_{z,k,t}} + r_{y,t} \quad (5.54)$$

となる. 求めた射影点が xy 平面上で交わらない場合は, シミュレーション上は演算できる最遠点を射影点と置く. 求めた射影点で囲われたエリアが認識用ロボットのカメラが撮影している範囲となる. この範囲を $E_{camera,t}$ とする.

5.5.6.2 テンプレート画像と同じ角度の人物画像を取得できる範囲

次にテンプレート画像に映っている人物画像と同じ角度の人物画像を取得できる範囲について以下に述べる．図 5. 11 にそのイメージ図を示す．テンプレート画像 $T_i(i \in \{1, \dots, N\})$ を取得したときの俯角を $\theta_{tp_dep,i}$ とし，そのテンプレート画像と照合可能な角度の範囲をマージン $\theta_{mar_dep_min}$, $\theta_{mar_dep_max}$ と設定すると，照合可能な俯角 θ の範囲は，以下の式になる．

$$\theta_{tp_dep,i} - \theta_{mar_dep_min} < \theta < \theta_{tp_dep,i} + \theta_{mar_dep_max} \quad (5.55)$$

$$\theta = \arctan \left(\frac{rz - height}{\sqrt{(x - rx)^2 + (y - ry)^2}} \right) \quad (5.56)$$

$height$ は検索対象の身長であり，テンプレート取得時に取得する．各テンプレート画像 T_i に映っている人物角度（仰角）は異なるので，式(5.58)を満たす範囲は，各テンプレート画像で作成される．その範囲を $E_{dep,it}$ とする．

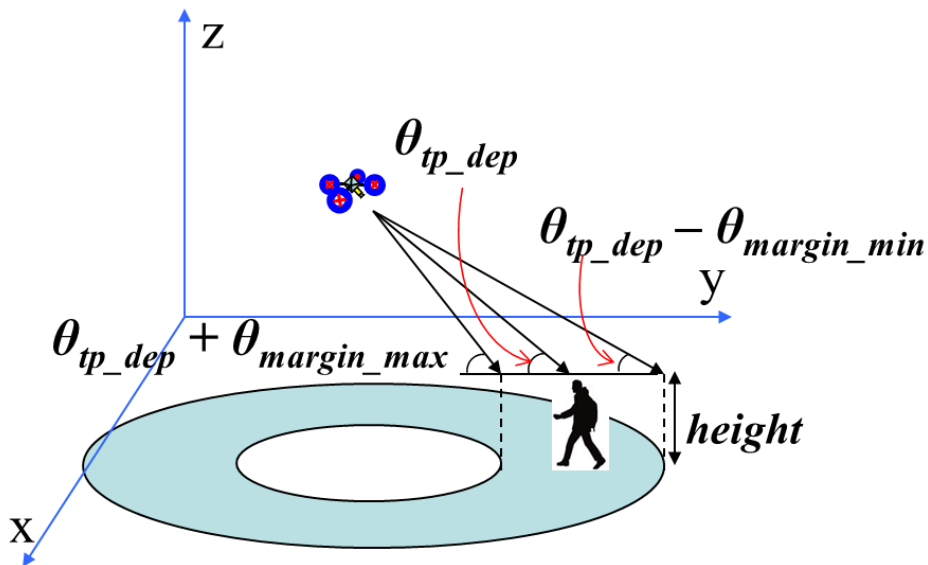


図5.11: テンプレート画像の人物画像と同じ角度の人物画像を取得できる範囲

5.5.6.3 照合可能な人物画像が撮影できる範囲

照合可能な人物画像が撮影できる範囲を求める。その範囲とは、入力画像の人物の上半身のサイズが照合可能な範囲であること、その上半身が入力画像内に収まっていることである。

第1に入力画像の人物の上半身のサイズが照合可能である範囲を求める。時間 t における認証用ロボットの状態 r_t で、イメージプレーンに映った人物画像のサイズが、照合できる範囲内に収めることができる xy 平面上の範囲を求める。そのイメージ図を図 5.12 に示す。

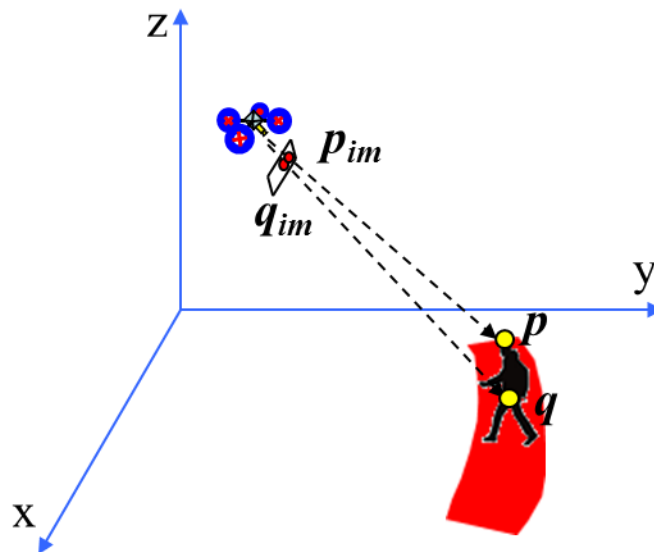


図5.12: 照合可能な人物サイズが撮影できる範囲

人物の頭部を点 p 、人物の腹部を q と設定する。 p 、 q をイメージプレーンに投影した点を p_{im} 、 q_{im} とすると、イメージプレーン上での人物の上半身のサイズ H_{im} は以下の式となる。

$$H_{im} = |p_{im} - q_{im}| \quad (5.57)$$

H_{im} が照合できるサイズの範囲内にあればその地点は照合可能な範囲内にあると判断できる。2点 p 、 q を以下のように設定する。

$$p = \begin{pmatrix} p_x \\ p_y \\ height \end{pmatrix} \quad (5.58)$$

$$q = \begin{pmatrix} p_x \\ p_y \\ height/2 \end{pmatrix} \quad (5.59)$$

ここで、式(5.54)で表したイメージプレーンの4頂点を含む平面は、以下の式で表される

$$ax + by + cz + d = 0 \quad (5.60)$$

$$a = (im_{y,2,t} - im_{y,1,t}) \cdot (im_{z,3,t} - im_{z,1,t}) - (im_{y,3,t} - im_{y,1,t}) \cdot (im_{z,2,t} - im_{z,1,t}) \quad (5.61)$$

$$b = (im_{z,2,t} - im_{z,1,t}) \cdot (im_{x,3,t} - im_{x,1,t}) - (im_{z,3,t} - im_{z,1,t}) \cdot (im_{x,2,t} - im_{x,1,t}) \quad (5.62)$$

$$c = (im_{x,2,t} - im_{x,1,t}) \cdot (im_{y,3,t} - im_{y,1,t}) - (im_{x,3,t} - im_{x,1,t}) \cdot (im_{y,2,t} - im_{y,1,t}) \quad (5.63)$$

$$d = -(a \cdot im_{x,1,t} + b \cdot im_{y,1,t} + c \cdot im_{z,1,t}) \quad (5.64)$$

焦点 r_i と人物の頭部の点 p を結ぶ単位ベクトルを $v_{p,t}$ 、焦点 r_i と人物の腹部の点 q を結ぶ単位ベクトルを $v_{q,t}$ とする。

$$v_{p,t} = \begin{pmatrix} v_{p x,t} \\ v_{p y,t} \\ v_{p z,t} \end{pmatrix} \quad (5.65)$$

$$v_{q,t} = \begin{pmatrix} v_{q x,t} \\ v_{q y,t} \\ v_{q z,t} \end{pmatrix} \quad (5.66)$$

求める p_{im} 、 q_{im} は、

$$p_{im} = \begin{pmatrix} p_{im x} \\ p_{im y} \\ p_{im z} \end{pmatrix} \quad (5.67)$$

$$q_{im} = \begin{pmatrix} q_{im x} \\ q_{im y} \\ q_{im z} \end{pmatrix} \quad (5.68)$$

$$p_{imx} = v_{px,t} \cdot \left(\frac{-(a \cdot r_{x,t} + b \cdot r_{y,t} + c \cdot r_{z,t} + d)}{a \cdot v_{px,t} + b \cdot v_{py,t} + c \cdot v_{pz,t}} \right) + r_{x,t} \quad (5.69)$$

$$p_{imy} = v_{py,t} \cdot \left(\frac{-(a \cdot r_{x,t} + b \cdot r_{y,t} + c \cdot r_{z,t} + d)}{a \cdot v_{px,t} + b \cdot v_{py,t} + c \cdot v_{pz,t}} \right) + r_{y,t} \quad (5.70)$$

$$p_{imz} = v_{pz,t} \cdot \left(\frac{-(a \cdot r_{x,t} + b \cdot r_{y,t} + c \cdot r_{z,t} + d)}{a \cdot v_{px,t} + b \cdot v_{py,t} + c \cdot v_{pz,t}} \right) + r_{z,t} \quad (5.71)$$

$$q_{imx} = v_{qx,t} \cdot \left(\frac{-(a \cdot r_{x,t} + b \cdot r_{y,t} + c \cdot r_{z,t} + d)}{a \cdot v_{qx,t} + b \cdot v_{qy,t} + c \cdot v_{qz,t}} \right) + r_{x,t} \quad (5.72)$$

$$q_{imy} = v_{qy,t} \cdot \left(\frac{-(a \cdot r_{x,t} + b \cdot r_{y,t} + c \cdot r_{z,t} + d)}{a \cdot v_{qx,t} + b \cdot v_{qy,t} + c \cdot v_{qz,t}} \right) + r_{y,t} \quad (5.73)$$

$$q_{imz} = v_{qz,t} \cdot \left(\frac{-(a \cdot r_{x,t} + b \cdot r_{y,t} + c \cdot r_{z,t} + d)}{a \cdot v_{qx,t} + b \cdot v_{qy,t} + c \cdot v_{qz,t}} \right) + r_{z,t} \quad (5.74)$$

p_{im} , q_{im} は以上のように求めることができる。これによりイメージプレーン上での人物の上半身のサイズ H_{im} を式(5.57)により求めることが出来る。そのサイズ H_{im} が照合できる各サイズ s 内である範囲を $E_{H,s,t}$ と表す。5.5.2と同様に、 s はサイズの範囲の添え字である($s=1,2,3,4$)。サイズの範囲は、 $s=1$ のとき46pixel以上、 $s=2$ のとき45～23pixel、 $s=3$ のとき22～15pixel、 $s=4$ のとき14～10pixelと設定した。

第2に人物の上半身すべてが入力画像内にあることをチェックする。求めた頭部の点 p_{im} 、と足部の点 q_{im} が、カメラの撮影範囲内に入っていればよい。イメージプレーンの4頂点は式(5.50)で求めているので、その範囲内にあるかをチェックすればよい。その範囲を $E_{pq,t}$ と表す。

5.5.6.4 各範囲の統合

画像取得範囲 E は、これまで述べたカメラ撮影範囲、テンプレート画像の人物画像と同じ角度の人物画像を取得できる範囲、照合可能な人物画像が撮影できる範囲（サイズ、上半身が画像内にある）を統合したエリアとなる。よって、

$$E_{i,s,t} = E_{camera,t} \cap E_{dep,i,t} \cap E_{H,s,t} \cap E_{pq,t} \quad (5.75)$$

と表すことが出来る。 i はテンプレート画像の添え字である。

5.6 軌道生成シミュレーション

5.6.1 照合用テンプレート画像

これまでに述べた手法の有効性を確かめるため、軌道生成シミュレーションを作成した。シミュレーションに用いたテンプレート画像は、探索エリアの入場ゲートで人物の周囲1回転が撮影できると想定し、図5.4で示したものをを用いた。このテンプレート画像の撮影は屋内環境で行った。カメラに対する人物の頭部の角度が、俯角30度の位置になるように撮影した。カメラに対する人物の回転角度は、 45° ずつ8角度を撮影した。撮影した画像からテンプレート用画像を作成した。

5.6.2 軌道生成パラメータ

シミュレーションに用いた軌道生成を行うためのパラメータについて述べる。飛行ロボットの速度は等速(8.33(m/s))とした。飛行ロボットの向きについては、roll角、pitch角は固定で 0° とした。yaw角は各軌道を飛行中は固定とした。軌道候補生成について以下に述べる。ロボットの初期位置から直線で進む軌道を作成した。その直線はロボットの初期位置から放射上に 10° ずつ36パターン分作成した。また、各直線上から大きく離れない範囲で、各直線を円弧に変形させた軌道を2種類作成した。さらに、各軌道においてロボットのyaw角を 10° ずつ回転させた軌道を作成した。そのため、軌道候補数は約3800となった。

飛行ロボットに搭載するカメラパラメータについては、一般的なカメラを想定し、式(5.36)の値を次のように設定した。 $ccd_{wid}=0.036mm$, $ccd_{hei}=0.024mm$, $fd=0.022mm$, $image_{wid}=640$, $image_{hei}=480$, $dep=30$ である。式(5.57)の照合可能な俯角の角度マージンは、テンプレートと比較して、カメラに対する人物の俯角が変わっても照合値の低下しない範囲を設定した($\theta_{mar_dep_min}=10$, $\theta_{mar_dep_max}=10$)。また、本シミュレーションでは、1fps (frame per second)での撮影を想定し、5秒間の軌道を生成した。よって、画像撮影数 F は5である。

5.6.3 画像取得範囲と人物存在確率作成シミュレーション

5.5.6で述べた画像取得範囲をシミュレーションにより求めていく様子を示す。図5.13に(a)カメラ撮影範囲、(b)テンプレート画像の人物画像と同じ角度(仰角)の人物画像を取得できる範囲、(c)照合可能な人物画像が撮影できる範囲(サイズ)を示す。それらの範囲を合成し、照合可能な人物画像が撮影できる範囲(上半身が画像内)をチェックした最終結果を図5.14に示す。図5.13では、ピンク色が飛行ロボットの位置であり、飛行ロボットの位置を $x=75$, $y=75$, $z=4$ (単位はm)と設定した。方向は $deg=0$ (水平右方向)を向いている。また、図5.15に時間経過と共に人物存在確率が移り変わっていく様子を示す。認証対象者のパラメータは、位置($x=25, y=25$)、向き($deg=270$:シミュレーション図の垂直下方向)、速度($speed=1.1(m/s)$)である。時間と共に存在確率分布が広がっているのが分かる。

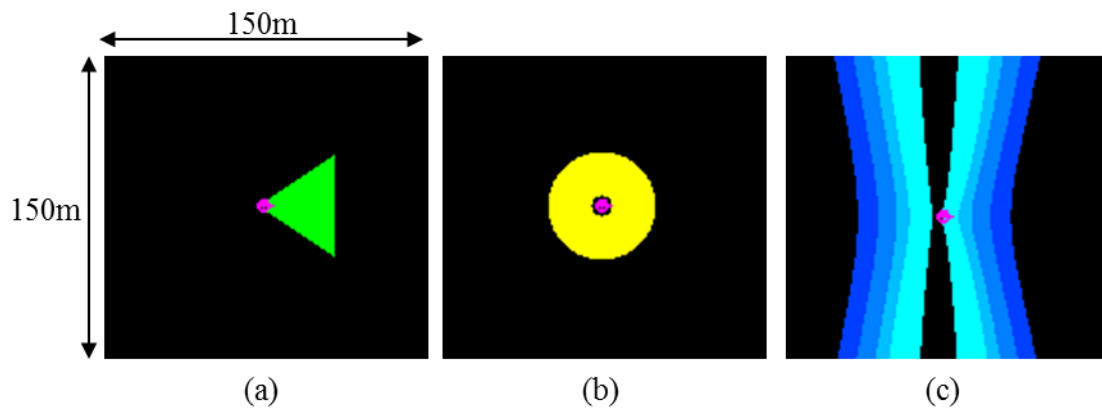
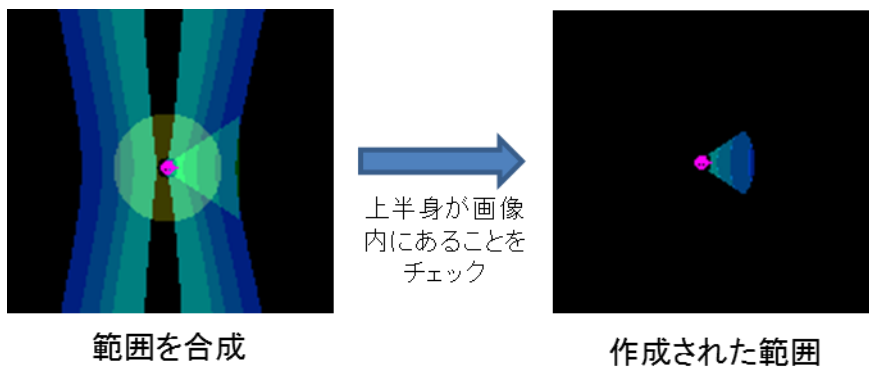


図5.13:画像取得範囲シミュレーション

(a): カメラ撮影範囲

(b): テンプレート画像の人物画像と同じ角度（仰角）の人物画像を取得できる範囲

(c): 照合可能な人物サイズが撮影できる範囲（ $s=1\sim 4$ に対応して水色～青色）



範囲を合成

作成された範囲

図5.14: 合成範囲

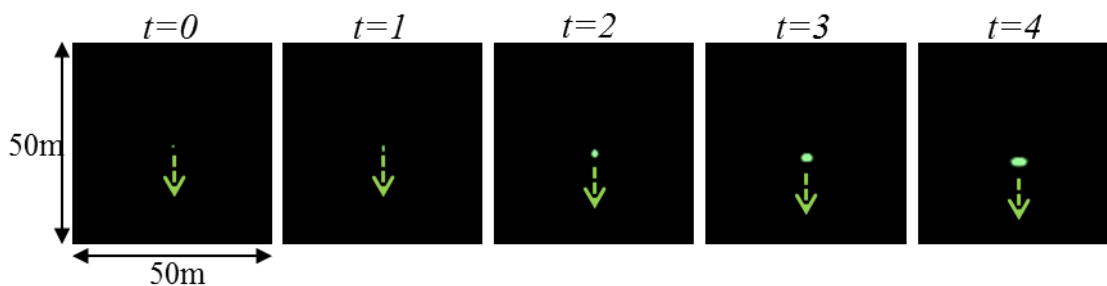


図5.15: 人物存在確率分布（緑色が存在確率の高い部分，矢印は進行方向）

5.6.4 認証対象者1人の軌道生成シミュレーション結果

飛行ロボットの初期位置を $x=25$, $y=10$, $z=4$ (単位はm)と設定し, 認証対象者のパラメータは, 位置($x=25,y=25$), 向き(deg=270:シミュレーション図の垂直下方向), 速度(speed=1.1(m/s))とした. 用いたテンプレートは実験対象者Aである. (図5.4) また式(5.26), 式(5.27)で表される地図情報 P_o と認証対象者 P_s の確率については, 本シミュレーションでは常に1とした. このときの軌道候補と式(5.8)で示した軌道評価値を図5.16に, 選ばれた軌道の飛行シミュレーション図を図5.17に示す.

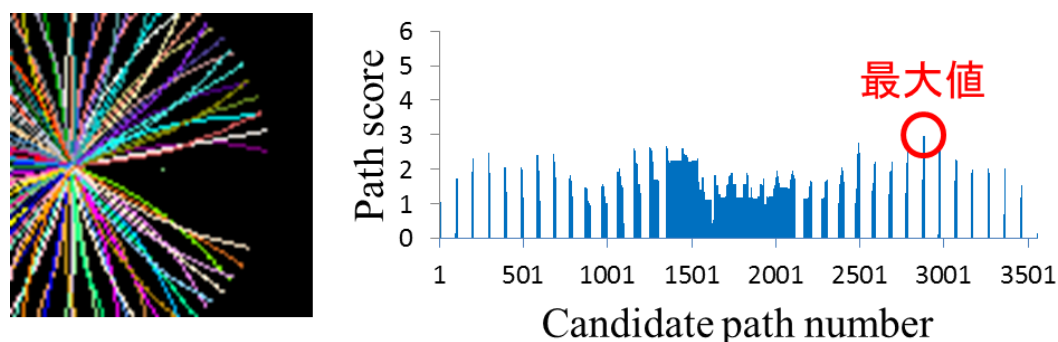


図5.16: 軌道候補と軌道評価値
(認証対象者1人, テンプレートは実験対象者A)

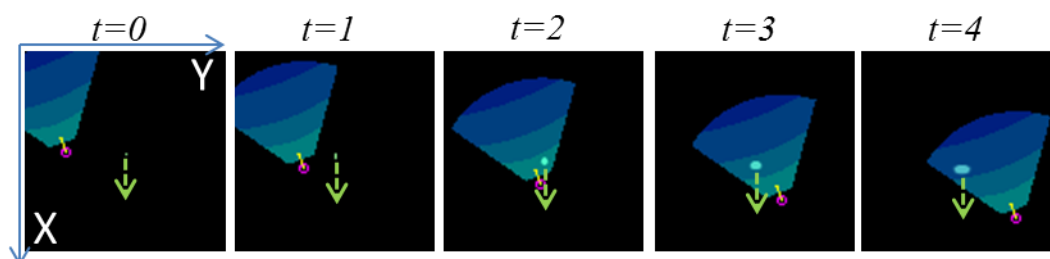


図5.17: 飛行シミュレーション図 (軌道番号2878)

図5.16に示した各軌道候補の軌道評価値の最大値を得る軌道は軌道番号2878となった. 図5.17は時間 t における人物画像取得範囲と, 認証対象者の存在確率分布を示した. 人物の存在確率分布は, 時間が経過するにしたがって, X 方向(画像垂直下方向)に移動しながら分布が広がっていることが分かる. 飛行ロボットはこの人物の動きに対し, 認証対象者のおよそ正面方向の画像を取得する軌道を描いていることが確認できる.

また、本軌道生成に掛かった時間は約1分となった(Intel Core i7 CPU 1スレッド)。本シミュレーションのうちの90%以上の処理時間が掛かっている部分は、画像パターン A_l の取得確率 $P_{pattern,l}$ を求める部分であった(式(5.12))。

次に、パラメータは同じで、探索対象者(テンプレート)を実験対象者B(図5.5)に変えてシミュレーションを行った。軌道評価値を図5.18に、求めた飛行シミュレーション図を図5.19に示す。

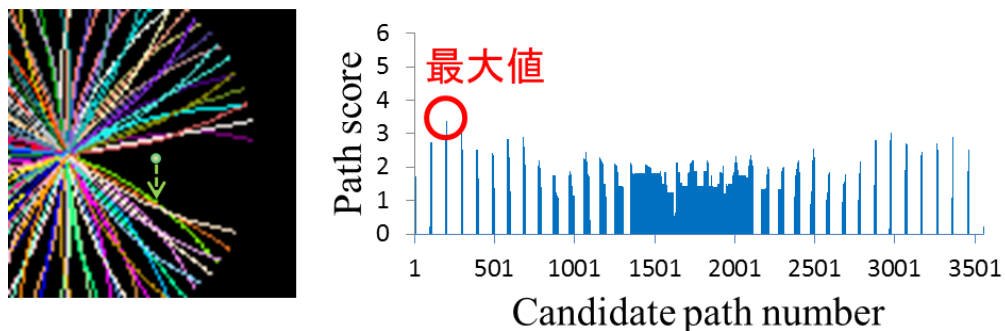


図5.18: 軌道候補と軌道評価値
(認証対象者1人, テンプレートは実験対象者B)

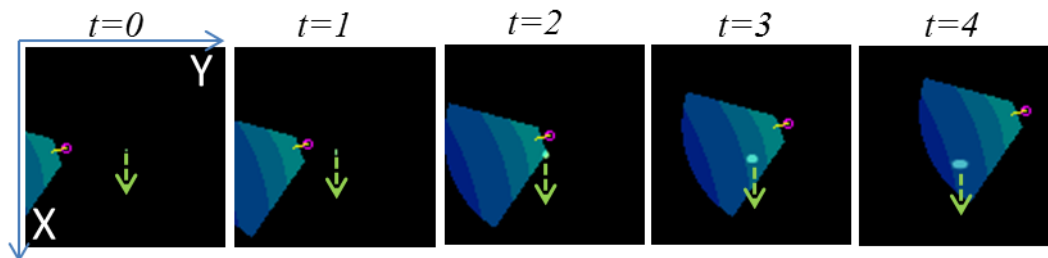


図5.19: 飛行シミュレーション図(軌道番号198)

用いたテンプレートは、正面方向の照合の重みが小さく他の方向の重みが大きいため、正面方向を撮影する軌道よりも、他の方向を撮影する軌道が選択されたと考えられる。このように、本手法では、テンプレート画像の角度やその照合重みの違いに対応した軌道を生成していることが分かる。

5.6.5 複数人の認証対象者の軌道生成シミュレーション

5.6.4の設定に対して、認証対象者を2人加えたときの軌道生成を行った。加えた認証対象者のパラメータについては、1人目は位置(x=15,y=35), 向き(deg=30), 速度(speed = 1.1(m/s)), 2人目は位置(x = 15,y = 30), 向き(deg = 135), 速度(speed = 1.1(m/s))とした。軌道候補と軌道評価値を図5. 20に, 飛行シミュレーション図を図5. 21に示す。テンプレートは実験対象者Aである。

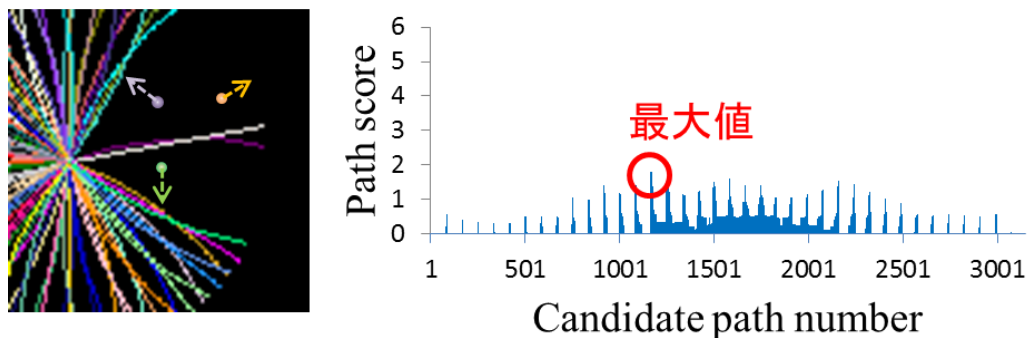


図5. 20: 軌道候補と軌道評価値
(認証対象者3人, テンプレートは実験対象者A)

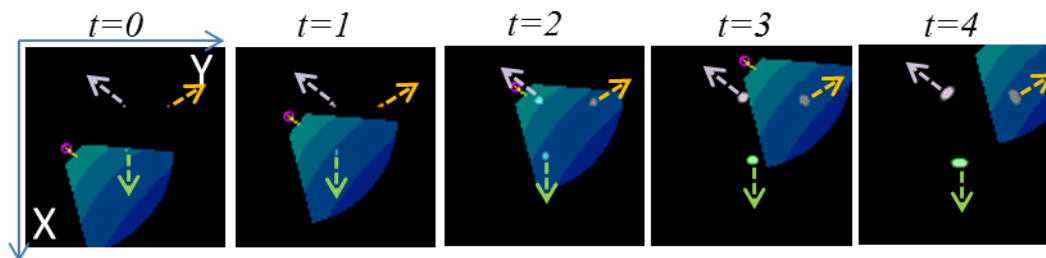


図5. 21: 飛行シミュレーション図 (軌道番号1166)

図5. 20, 図5. 21から, 認証対象者全員の画像を取得する軌道が選ばれていることが確認できる。

5.6.6 事前知識がある場合の軌道生成シミュレーション

1試行の飛行で認証できない場合は、それまでに得た人物映像を考慮し、認証精度が高くなる軌道を再計画し、再試行する。このような事前知識がある場合の軌道生成についてシミュレーションを行った。軌道の再計画時はそれまでに求めた式(5.18)で示した人物画像取得確率を引き継ぎ、それまでに求めた軌道と再計画時の軌道を組み合わせたときの人物画像取得確率を求め、それをもとに軌道評価値を求める。本実験では、5.6.4で求めた実験対象者Aの軌道を事前知識の軌道とした。再計画時の軌道の初期位置は5.6.4で求めた実験対象者Aの軌道の終了位置とした。軌道候補と軌道評価値を図5.22に、求めた軌道と飛行シミュレーション結果を図5.23に示す。また比較のために、同条件で、事前知識がない場合のシミュレーション結果を図5.24に示す。

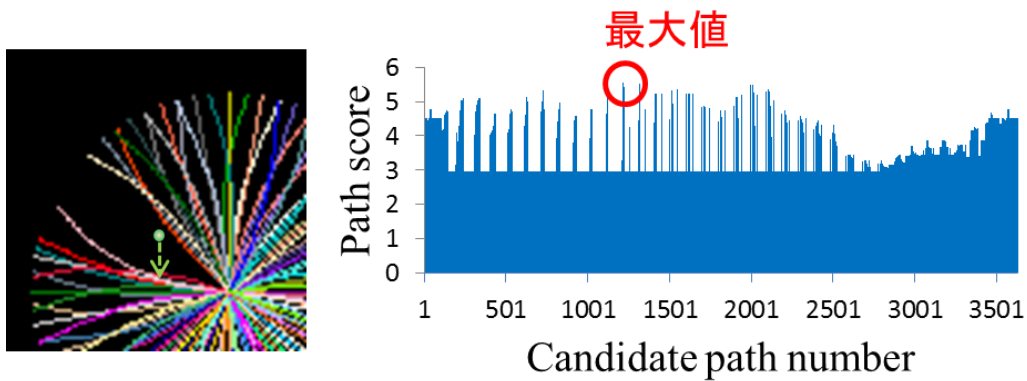


図5.22: 事前知識ありの軌道候補と軌道評価値
(認証対象者1人、テンプレートは実験対象者A)

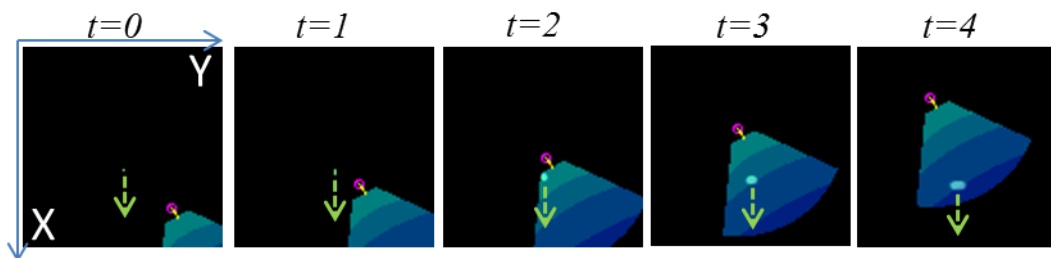


図5.23: 飛行シミュレーション図 (軌道番号1215)

図5.22に示したように、事前知識がある場合は、軌道評価値の最小値は、事前知識で得られた軌道評価値となる。図5.23より、認証対象者は垂直下方向に進んでおり、認証対象者の背

面方向を取得する軌道になっていることが分かる。これは、再計画前に、正面方向の画像を取得しているため、異なる方向の人物画像を取得する軌道のほうが、軌道評価値が高くなったためである。それは図5.24からも確認できる。事前知識がない場合は、正面方向の画像を取得する軌道が選択された。

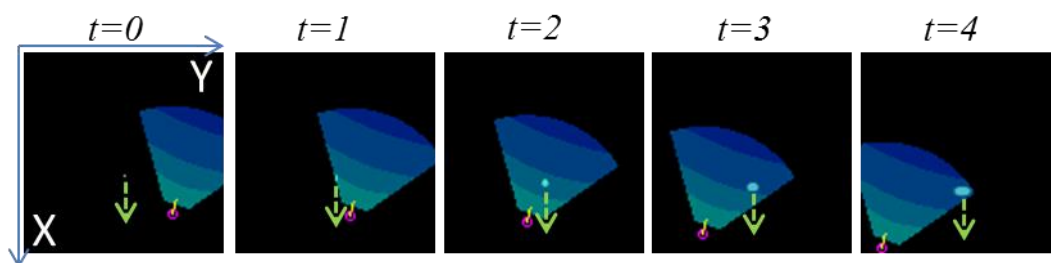


図5.24: 事前知識なしの場合の飛行シミュレーション結果図 (軌道番号189)

5.6.7 より複雑な軌道作成が可能な場合の演算例

これまでの軌道候補作成は、あらゆる飛行ロボットに適用可能なように、認証用ロボットの初期位置から直線、または円弧で放射上に複数の候補を作成した。本節では、より高い機動性能を持つ飛行ロボットにも適用可能であることを示すために、手動で作成した図5.25に示す軌道を描いたときのシミュレーションを行った。この軌道は3人の認証対象者の周囲をぐるりと回りながら撮影する軌道である。そのシミュレーション結果図を図5.26に示す。このときの軌道評価値は3.43であり、図5.20と比較しても最も高い軌道評価値となった。

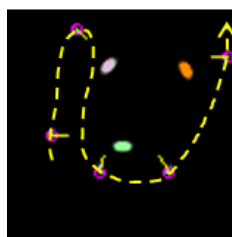


図5.25: 手動で設定した軌道

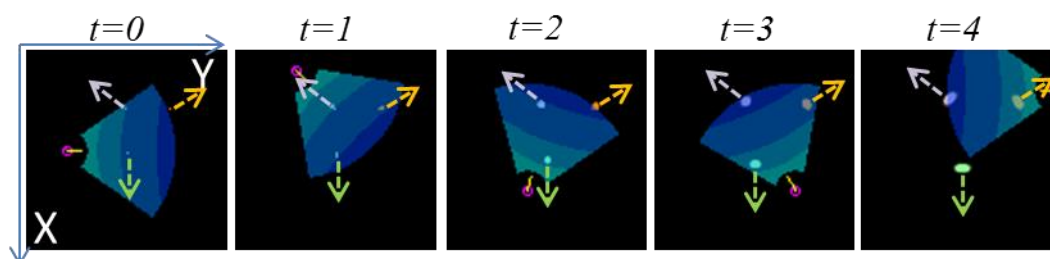


図5.26: 飛行シミュレーション図 (手動で設定した軌道) : 軌道評価値3.43

5.6.8 軌道生成シミュレーションまとめと考察

軌道候補生成，軌道評価，求めた軌道結果の図示化を行うことができるシミュレーションを作成した．本シミュレーションを用いて，複数のパターンの軌道生成を行い，その結果を確認することが出来た．そのパターンは，認証対象者が1人のときの軌道生成，同じセッティングで検索対象の人物を変更したときの軌道生成，3人の認証対象者がいる場合の軌道生成，事前知識がある場合の軌道生成である．

また，本軌道生成シミュレーションに掛かった時間は約1分であった(Intel Core i7 CPU 1スレッド)．本シミュレーションに掛かった処理時間のうち，90%以上が画像パターン A_l の取得確率 $P_{pattern,l}$ を求める部分であった(式(5.12))．

軌道候補生成については，本シミュレーションでは，低スペックな飛行ロボットでも動作できることを考慮し，複雑な軌道候補は作成せず単純な候補作成のみ行った．もし，高い軌道能力を持つ飛行ロボットの軌道生成を行いたい場合は，その軌道能力に合わせた軌道候補生成を行えばよい．その後の軌道評価演算処理は同じである．そのときのシミュレーション例は5.6.7で示した．

処理時間については，画像パターン群 A の数が膨大であることが最も影響している．式(5.16)より，テンプレート画像数 N とサイズ分割数 S の掛け算+1の数の入力画像の状態番号数 L が存在する．本シミュレーションでは， $N=8$ ， $S=4$ なので， $L=8 \times 4 + 1 = 33$ である．そして，取得フレーム数 $F=5$ なので，画像パターンの総数は， $L^F = 33^5$ となる．1つの軌道候補の評価値を求めるために，それだけの数の演算が必要である．本シミュレーションでも演算処理の工夫を行うことで，高速化を図った．すべての組み合わせ演算を行う前に図5.8に示した人物画像取得確率表をソートしておき，その確率が高い組み合わせ順から演算を行い，その確率の合計値が閾値(本シミュレーションでは0.9999)以上になれば演算を終了する．このことにより本節で示したパターンの処理時間は平均1/10倍以下となったが，設定したシミュレーションのパターンによりその処理時間は変動する．よって，今後リアルタイム化を目指すためには，理論的に演算処理を減らす工夫が必要であると考えられる．

5.7 まとめ

本章では，飛行ロボットによる人物検索のための軌道生成手法を提案した．その軌道生成手法は，人物認証アルゴリズムとして累積アピランズ照合アルゴリズムを適用したときに，複数人の認証対象者の照合値が高くなる軌道を生成するものである．ある軌道を与えると，その軌道を飛行ロボットが通過しながら撮影したときに，取得できると推定できる認証対象者の照合値を返す関数を作成した．その関数に複数の軌道候補を入力し，その評価値の中から最大値を得る軌道が求める軌道である．その関数は，人物の位置と向きの確率分布と，幾何学的に演算した画像の取得範囲から，テンプレート画像に対応する入力人物画像の取得確

率を求める。そして、その人物が搜索対象人物であると想定し、その入力人物画像の取得確率から推定照合値を返すものである。このようにして、ある軌道を通ったときにどれぐらいの照合値が得られるかを推定し、それが最も高くなる軌道が求める軌道となる。

本手法の有効性を確かめるためシミュレーションを作成した。搜索対象の人物の違い（テンプレート画像の違い）による軌道変化や、複数人の認証対象者がいる場合での軌道生成、事前知識があるときの軌道生成結果をシミュレーションで示すことが出来た。また、軌道候補生成と軌道評価関数は分離した処理であるため、高性能な飛行ロボットに適用するためには、それに対応した軌道候補生成を行えば、その後の処理は同じであることを述べた。

本軌道生成シミュレーションに掛かる処理時間は約1分であった(Intel Core i7 CPU 1スレッド)。処理時間が多く掛かっている原因についても言及した。画像取得パターン数が膨大であることが原因である。高速化が今後の課題である。

第6章 結論

6.1 結論

本論文では、飛行ロボットによる人物検索システムを提案した。そのシステムは、大型の遊園地やショッピングモール等の閉鎖空間において、迷子や犯人検索を行うものである。カメラを搭載した飛行ロボットが、検索エリア内を飛行しながら人物映像を撮影し、その映像内に写った人物画像を認証することで人物検索を行う。そのシステムを実現するための技術のポイントである、飛行ロボットが人物検索を行うための、人物検出、人物認証、そして、飛行ロボットの軌道生成について、それぞれ手法を提案し、精度検証実験、実環境での実証実験を行い、その実現性を示した。

飛行ロボットは上空から人物画像を撮影するため、写った人物画像の足元はオクルージョンとなりやすい。また、固定された防犯カメラで撮影した映像とは違い、その画像に含まれる背景はさまざまな場所を写している。そのため、固定カメラで撮影した映像と比較して、背景を誤検出する可能性が高い。よって、従来法より高精度な上半身による人物検出アルゴリズムを提案した。公開されているデータベースを用いて精度検証を行い、従来法よりも高精度であることを確認した。さらに飛行ロボットが実際に飛行しながら撮影した画像から人物検出を行い、実環境においても適用可能であることを示した。

飛行ロボットが撮影する人物画像は、目視では顔を確認できないぐらい低解像度の上半身画像である。さらに照明変動やカラー変化も考えられる。そのため、本研究で提案した人物認証アルゴリズムで用いた特徴量は、低解像度の上半身のエッジ特徴のみである。認証するにはシビアな条件であるが、飛行ロボットのメリットであるあらゆる視点から撮影できることを生かし、複数枚の複数方向から撮影された人物画像の情報を蓄積することで精度向上を図った。38人の実験対象者による評価を行った結果、 15×15 画素の上半身の画像から夏私服と冬私服で平均85%の認証率（誤認率5%）を得ることが出来た。全員が同じ服装（制服）を着た場合でも、識別可能であることを示した。テンプレートと入力で服装が異なる場合の実験も行い、本アルゴリズムの限界点も示した。さらに、飛行ロボットが実際に飛行しながら撮影した画像を用いて、人物検索実験を行い、6人の実験対象者の中から複数の撮影パターンで人物検索を行えることを示した。

作成した人物検出アルゴリズムと人物認証アルゴリズムを統合し、一連の流れの中で人物検索が可能であることを示した。手入力による人物検出結果と比較しても精度的に同等であることも示した。さらにリアルタイムに実行できるアプリケーションを作成し、映像を撮影しながら人物検出、人物認証処理が行えるシステムの1例も示した。

これらのことにより、検索エリア内に飛行ロボットを飛行させ、人物検索が行えるようになったと言える。飛行ロボットを手動によって操作する、もしくは、自動巡回飛行の機能を用いることで、飛行ロボットによる人物検索が可能である。

さらに本研究では、効率的かつ高精度な飛行ロボットによる人物認証を行うために、飛行ロボットの軌道生成手法についても提案した。飛行ロボットの軌道は無限大に存在する。本軌道生成手法は、その中から、本研究で提案した人物認証アルゴリズムを用いたときに、複数の認証対象者を効率よく、且つ、高精度に認証するものである。その軌道生成シミュレーションを作成し、複数のパターンでその効果を確認した。人物（テンプレート画像）の違いによる軌道変化や、2試行目の飛行など事前知識がある場合についても、シミュレーションにより確認することができた。

6.2 展望

人物認証アルゴリズムについては、将来、人物の頭上を飛行しても安全な飛行ロボットが作られると想定すると、より鋭角（人物の頭が写る方向）な画像に対しても、認証できるか検証する必要がある。人物検出アルゴリズムについても、提案手法は Ω 型検出を基本としているため、頭上からの人物画像については、再検討の必要がある。

軌道生成については、シミュレーションにより軌道生成できることは確認した。今後は、この軌道を用いたときの実環境での認証精度の検証が必要である。さらに、本軌道生成システムを実環境に適用するには、俯瞰用ロボットによる人物検出と追跡を実装する必要がある。また、処理時間についても、1回の軌道生成に1分程度かかっており、演算処理の並列化等では解決できない可能性もある。実現性を優先した高速化処理、第2案の検討も行いたい。

将来的には、複数の飛行ロボットによる連携が考えられる。同時刻の複数視点からの映像が取得できると、それらを組み合わせることで、人物検出や人物認証の精度を向上させられる可能性がある。さらに役割分担できる。最終判定をより高精度に行うために、アピアランス認証アルゴリズムだけでなく、顔認証を行う飛行ロボットがいてもよい。アピアランス認証で閾値以上となった人物に対してのみ、その人物の正面に回り込み、顔認証を行えば、より高精度に人物認証が可能となる。このような複数台のロボットによる連携システムも、次の研究トピックとして考えられる。

参考文献

第1章

- [1] 「安全・安心な社会の構築に資する科学技術政策に関する懇談会」報告書, 2004.
http://www.mext.go.jp/a_menu/kagaku/anzen/houkoku/04042302.htm (2015/12/11 アクセス)
- [2] 警視庁万世橋警察署 生活安全課課長代理 (防犯・少年担当) 警部 清水克修, 「イギリス・ロンドン市における防犯カメラの現状について」, 2008.
<http://www.secu354.co.jp/contents/seminar/08/seminar-081010-4.htm> (2015/12/11 アクセス)
- [3] 千代田区議会議員 小林やすお 活動報告, 2013
<http://yasuo101.sakurasv.com/report/index.php/2013/10/02/2949/> (2016/2/10 アクセス)
- [4] 西原博史, 『監視カメラとプライバシー』, パースペクティヴズ, 2009, 96pp.
- [5] N. Otsu, "Towards Flexible and Intelligent Vision Systems--From Thresholding to CHLAC", MVA.2005, pp. 430-439, 2005.
- [6] 国立研究開発法人 産業技術総合研究所, 「カメラ映像から異常動作をリアルタイムで自動検出するソフトを開発」, 2007.
http://www.aist.go.jp/aist_j/press_release/pr2007/pr20071016/pr20071016.html#a
(2015/12/11 アクセス)
- [7] W. Hu, T. Tan, L. Wang, S. Maybank, "A survey on visual surveillance of object motion and behaviors.", Systems Man and Cybernetics, Part C: Applications and Reviews, IEEE Transactions, vol. 34, no. 3, pp. 334-352, 2004.
- [8] 坂野鋭, “生体認証技術の最近の動向”, 日本法科学技術学会誌, vol. 12, no. 1, pp. 1-12, 2007.
- [9] A. Briod, P. Kornatowski, J. C. Zufferey, D. Floreano, "A Collision - resilient Flying Robot", Journal of Field Robotics, vol. 31, no. 4, pp. 496-509, 2014.
- [10] R. Wood, N. Radhika, G. Y. Wei, "Flight of the Robobees.", Scientific American, vol. 308, no. 3, pp. 60-65, 2013.
- [11] NEC: Facial recognition: Gate System
<http://jpn.nec.com/ad/usj/entry.html> (2015/12/11 アクセス)

第2章

- [1] 内村圭一, 岩崎洋一郎, 松島宏典, 『画像処理入門』, 培風館, 2010

- [2] 鈴木秀和, 見浪護, “車載カメラを用いた複数歩行者の実時間認識”, 電子情報通信学会論文誌 A, vol. 88, no. 2, pp. 131-141, 2005.
- [3] P. Viola, J. Michael, "Rapid object detection using a boosted cascade of simple features", Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference, vol. 1, pp. 511-518, 2001
- [4] Y. Freund, RE. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of computer and system sciences, vol. 55, no. 1, pp. 119-139, 1997.
- [5] N. Dalal, T. Bill, "Histograms of oriented gradients for human detection", Computer Vision and Pattern Recognition, CVPR 2005. IEEE Computer Society Conference, vol. 1, pp. 886-893, 2005.
- [6] K. Levi, W. Yair, "Learning object detection from a small number of examples: the importance of good features", Computer Vision and Pattern Recognition, CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference, vol. 2, pp. 53-60, 2004.
- [7] B. Wu, R. Nevatia, "Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors", Computer Vision, ICCV 2005. Tenth IEEE International Conference, vol. 1, pp. 90-97, 2005.
- [8] T. Mita, T. Kaneko, B. Stenger, "Discriminative feature co-occurrence selection for object detection", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 30, no. 7, pp. 1257-1269, 2008.
- [9] P. Sabzmejdani, G. Mori, "Detecting pedestrians by learning shapelet features", Computer Vision and Pattern Recognition, CVPR'07. IEEE Conference, pp. 1-8, 2007.
- [10] T. Mitsui, Y. Yuji, F. Hironobu, "Object detection by two-stage boosting with joint features", The Institute of Electronics, Information and Communication Engineers Transactions on Information and Systems, vol. J92-D, no. 9, pp. 1591-1601, 2009.
- [11] O. Tuzel, F. Porikli, P. Meer, "Pedestrian detection via classification on riemannian manifolds.", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 30, no. 10, pp. 1713-1727, 2008.
- [12] N. Dalal, B. Triggs, C. Schmid, "Human detection using oriented histograms of flow and appearance", Computer Vision–ECCV 2006. Springer Berlin Heidelberg, pp. 428-441, 2006.
- [13] 村井泰裕, 藤吉弘亘, 数井誠人, "時空間特徴に基づくエスカレーターシーンにおける人の異常行動検知", 信学技報, PRMU2008, pp.42-50, 2008.
- [14] Y. Yamauchi, H. Fujiyoshi, BW. Hwang, "People detection based on co-occurrence of appearance and spatiotemporal features", National Institute of Informatics Transactions on Progress in Informatics, vol. 1, no. 7, pp. 33-42, 2010.
- [15] M. Li, Z. Zhang, K. Huang, T. Tan, "Estimating the number of people in crowded scenes by mid based foreground segmentation and head-shoulder detection", Pattern Recognition, ICPR 2008. 19th International Conference, pp. 1-4, 2008.

- [16] M. Li, Z. Zhang, K. Huang, T. Tan, "Rapid and robust human detection and tracking based on omega-shape features", Image Processing (ICIP), 2009 16th IEEE International Conference, pp. 2545-2548, 2009.
- [17] C. Zeng, H. Ma, "Robust head-shoulder detection by pca-based multilevel hog-lbp detector for people counting", Pattern Recognition (ICPR), 2010 20th International Conference, pp. 2069-2072, 2010.
- [18] R. Hu, R. Wang, S. Shan, X. Chen, "Robust head-shoulder detection using a two-stage cascade framework", Pattern Recognition (ICPR), 2014 22nd International Conference, pp. 2796-2801, 2014.
- [19] グローリー株式会社, 米澤亨, "被写体識別方法, 被写体識別プログラムおよび被写体識別装置", 特許第5290401号, 2013-06-14.
- [20] 池村翔, 藤吉弘亘, "距離情報に基づく局所特徴量によるリアルタイム人検出", 電子情報通信学会論文誌 D, vol. 93, no. 3, pp. 355-364, 2010.
- [21] 中村克行, 趙卉菁, 柴崎亮介, 坂本圭司, 大鋸朋生, 鈴木尚毅, "複数のレーザレンジスキャナを用いた歩行者トラッキングとその信頼性評価", 電子情報通信学会論文誌 D, vol. 88, no. 7, pp. 1143-1152, 2005.
- [22] 小林貴訓, 杉村大輔, 関真規人, 平澤宏祐, 鈴木直彦, 鹿毛裕史, 杉本晃宏, "分散カメラとレーザ測域センサの統合によるエリア内人物追跡", 画像の認識理解シンポジウム, pp. IS-2-23, 2007.
- [23] G. Gate, A. Breheret, F. Nashashibi, "Fast pedestrian detection in dense environment with a laser scanner and a camera", Vehicular Technology Conference, VTC Spring 2009. IEEE 69th, pp. 1-6, 2009.
- [24] N. Bellotto, H. Hu, "Multisensor-based human detection and tracking for mobile service robots", Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions, vol. 39, no. 1, pp. 167-181, 2009.
- [25] 中野宏毅, 吉田靖夫, 山本鎮男, 『ウェブレットによる信号処理と画像処理』, 共立出版, 1999.
- [26] CAVIAR: Context Aware Vision using Image-based Active Recognition.
<http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>
- [27] A. Martin, G. Doddington, T. Kamm, M. Ordowski, M. Przybocki, "The DET curve in assessment of detection task performance", NATIONAL INST OF STANDARDS AND TECHNOLOGY GAITHERSBURG MD, 1997.
- [28] Y. Freund, RE. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of computer and system sciences, vol. 55, no. 1, pp. 119-139, 1997.
- [29] J. Friedman, T. Hastie, R. Tibshirani, "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)", The annals of statistics, vol. 28, no. 2, pp. 337-407, 2000.
- [30] RE. Schapire, Y. Singer, "Improved boosting algorithms using confidence-rated predictions", Machine learning, vol. 37, no. 3, pp. 297-336, 1999.

第3章

- [1] M. A. Turk and A. P. Pentland, "Face recognition using Eigenfaces", Computer Vision and Pattern Recognition, IEEE Computer Society Conference, pp. 586-591, 1991.
- [2] J. Han and B. Bhanu, "Individual recognition using gait energy image", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 2, pp. 316-322, 2006.
- [3] B. Prosser, W. Zheng, S. Gong, T. Xiang, "Person Re-Identification by Support Vector Ranking", BMVC, vol. 2, no. 5, 2010.
- [4] W. Zheng, S. Gong, T. Xiang, "Person re-identification by probabilistic relative distance comparison", Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference, pp. 649-656, 2011.
- [5] A. Mignon, F. Jurie, "Pcca: A new approach for distance learning from sparse pairwise constraints", Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference, pp. 2666-2672, 2012.
- [6] D. Gray, T. Hai, "Viewpoint invariant pedestrian recognition with an ensemble of localized features", Computer Vision—ECCV 2008. Springer Berlin Heidelberg, pp. 262-275, 2008.
- [7] C. Liu, S. Gong, CC. Loy, X. Lin, "Person re-identification: What features are important?", Computer Vision—ECCV 2012. Workshops and Demonstrations. Springer Berlin Heidelberg, pp. 391-401, 2012.
- [8] S. Bak, E. Corvee, F. Brémont and M. Thonnat, "Person Re-identification Using Haar-based and DCD-based Signature", AMCSS, pp.1-8, 2010.
- [9] X. Wang, G. Doretto, T. Sebastian, J. Rittscher and P. Tu, "Shape and Appearance Context Modeling", Proc. of ICCV, pp.1-8, 2007.
- [10] J. Kittler, J. Matas, K. Jonsson and M.U. R. Sanchez, "Combining evidence in personal identity verification systems", Pattern Recognition Letters, vol. 18, no. 9, pp. 845-852, 1997.
- [11] K. Maeda, S. Watanabe, "A Pattern Matching Method with Local Structure", IEICE Trans. D, vol. J68-D, no. 3, pp.345-352, 1985.
- [12] K. Fukui, O. Yamaguchi, K. Suzuki, K. Maeda, "Face Recognition under Variable Lighting Condition with Constrained Mutual Subspace Method—Learning of Constraint Subspace to Reduce Influence of Lighting Changes", IEICE Trans. D-II, vol. J82, no. 4, pp.613-620, 1999.
- [13] K. Nandakumar, Y. Chen, S. C. Dass and A. K. Jain, "Likelihood Ratio Based Biometric Score Fusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp.342-347, 2008.
- [14] S. Bak, E. Corvee, F. Brémont and M. Thonnat, "Multiple-shot human re-identification by mean 131iemannian covariance grid", Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference, pp. 179-184, 2011.
- [15] L. Bazzani, M. Cristani, A. Perina, V. Murino, "Multiple-shot person re-identification by chromatic and epitomic analyses", Pattern Recognition Letters, vol. 33, no. 7, pp. 898-903, 2012.

- [16] F. Morbidi, C. Ray, G. L. Mariottini, “Cooperative active target tracking for heterogeneous robots with application to gait monitoring”, 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.3608-3613, 2011.
- [17] N. Bellotto, H. Hu, “A bank of unscented Kalman filters for multimodal human perception with mobile service robots. International Journal of Social Robotics”, vol. 2, no. 2, pp. 121-136, 2010.
- [18] J. Satake, C. Masaya, J. Miura, “Visual person identification using a distance-dependent appearance model for a person following robot”, International Journal of Automation and Computing, vol. 10, no. 5, pp. 438-446, 2013.
- [19] NEC: Facial recognition: Gate System
<http://jpn.nec.com/ad/usj/entry.html> (2015/12/11 アクセス)
- [20] ZH. Khan, IYH. Gu, AG. Backhouse, “Robust visual object tracking using multi-mode anisotropic mean shift and particle filters”, Circuits and Systems for Video Technology, IEEE Transactions, vol. 21, no. 1, pp. 74-87, 2011.
- [21] 中野宏毅, 吉田靖夫, 山本鎮男, 『ウェーブレットによる信号処理と画像処理』, 共立出版, 1999.

第4章

- [1] ZH. Khan, IYH. Gu, AG. Backhouse, “Robust visual object tracking using multi-mode anisotropic mean shift and particle filters”, Circuits and Systems for Video Technology, IEEE Transactions, vol. 21, no. 1, pp. 74-87, 2011.

第5章

- [1] N. Roy, G. Gordon, and S. Thrun, “Finding Approximate POMDP Solutions Through Belief Compression”, Journal of Artificial Intelligence Research 23, pp.1-40, 2005.
- [2] N. M. Stiffler and J. M. O’Kane, “Visibility-Based Pursuit-Evasion with Probabilistic Evader Models”, 2011 IEEE International Conference on Robotics and Automation, pp.4254-4259, 2011.
- [3] R. Mottaghi, R. Vaughan, “An Integrated Particle Filter & Potential Field Method for Cooperative Robot Target Tracking”, Proceedings of the 2006 IEEE International Conference on Robotics and Automation, pp.1342-1347, 2006.
- [4] N. Takemura and J. Miura, “View Planning of Multiple Active Cameras for Wide Area Surveillance”, Proceedings of the 2006 IEEE International Conference on Robotics and Automation, pp.3173-3179, 2007.
- [5] 橋本雅文, 緒方聡, 大場史憲, 岡田三郎, “車載レーザレンジセンサによる複数移動物体の検出・追跡法.” 日本機械学会論文集 C 編, vol. 72, no. 717, pp. 1566-1573, 2006

- [6] M. Saito, K. Yamazaki, N. Hatao, R. Hanai, K. Okada, M. Inaba, "Pedestrian detection using a LRF and a small omni-view camera for outdoor personal mobility robot", Robotics and Biomimetics (ROBIO), 2010 IEEE International Conference, pp. 155-160, 2010.
- [7] T. Germa, F. Lerasle, N. Ouadah, V. Cadenat, "Vision and RFID data fusion for tracking people in crowds by a mobile robot", Computer Vision and Image Understanding, vol. 114, no. 6, pp. 641-651, 2010.
- [8] K. Tanaka, E. Takeuchi, K. Ohno, S. Tadokoro, T. Yonezawa, "Logical winnowing methods from multiple identification candidates using corresponding appearance identification results in time-series", Robotics and Automation (ICRA), 2012 IEEE International Conference, pp. 3631-3636, 2012.
- [9] 関根大輔, 川田浩彦, 大矢晃久, 油田信一, "移動ロボットの動作による人間の行動制御", ロボティクス・メカトロニクス講演会講演概要集, pp. ROMBUNNO. 1P1-B15, 2009.
- [10] N. Dalal, T. Bill, "Histograms of oriented gradients for human detection", Computer Vision and Pattern Recognition, CVPR 2005. IEEE Computer Society Conference, vol. 1, pp. 886-893, 2005.
- [11] M. Teutsch, T. Mueller, M. Huber, J. Beyerer, "Low resolution person detection with a moving thermal infrared camera by hot spot classification", Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference, pp. 209-216, 2014.
- [12] ZH. Khan, IYH. Gu, AG. Backhouse, "Robust visual object tracking using multi-mode anisotropic mean shift and particle filters", Circuits and Systems for Video Technology, IEEE Transactions, vol. 21, no. 1, pp. 74-87, 2011.

業績

論文

- [1] T. Yonezawa, E. Takeuchi, K. Ohno, S. Tadokoro, "Person Search Based on Cumulative Information from Low Resolution Person Images Captured by a Flying Robot", ROBOMECH Journal, 査読中.

国際会議

フルペーパー 査読あり

- [2] T. Yonezawa, E. Takeuchi, K. Ohno, S. Tadokoro, "Appearance-based Identification Algorithm to Search for Persons using a Flying Robot", Safety Security and Rescue Robotics (SSRR), 2015 IEEE International Symposium, 2015.
- [3] T. Yonezawa, E. Takeuchi, K. Ohno, S. Tadokoro, "Path-creation Method to Search for Persons using a Flying Robot", Robotics and Biomimetics (ROBIO), 2015 IEEE International Conference, pp. 1439-1446, 2015.
- [4] K. Tanaka, E. Takeuchi, K. Ohno, S. Tadokoro, T. Yonezawa, "Logical winnowing methods from multiple identification candidates using corresponding appearance identification results in time-series", Robotics and Automation (ICRA), 2012 IEEE International Conference, pp. 3631-3636, 2012.

Extended Abstract 査読あり

- [5] K. Tanaka, E. Takeuchi, K. Ohno, S. Tadokoro, T. Yonezawa, "Forereaching Motion Generation of Mobile Robots for Pedestrian Face Identification", SICE Annual Conference 2010, TB15.01, 2010.

国内会議

査読なし

- [6] 田中一志, 竹内栄二郎, 大野和則, 田所諭, 米澤亨, "複数人物見守りのための移動ロボットの観測計画", 日本機械学会2011年度年次大会, 2011.
- [7] 田中一志, 竹内栄二郎, 大野和則, 田所諭, 米澤亨, "移動ロボットの歩行者正面への回り込み動作生成", ロボティクス・メカトロニクス講演会2010, 2A2-C03, 2010.

謝辞

本研究はグローリー株式会社研究開発センター及び、東北大学大学院情報科学研究科，社会人博士後期課程において，多数の方々のご協力を得て行われたものです。

本研究を行うにあたり，入学当時から終始懇切なご指導，ご助言を賜りました田所諭 教授に厚くお礼申し上げます。ご多忙にもかかわらず，本論文の副査をしていただき，貴重なアドバイスを下さいました岡谷貴之 教授，永谷圭司 准教授，大野和則 准教授に感謝致します。また，本研究を進めるにあたり，多大なる助言と協力を頂きました，竹内栄二郎 特任准教授(現名古屋大学)に深く感謝致します。大野和則 准教授には，副査のみならず，入学当時から研究についての数多くの有益な助言と協力を頂きました。また海外での発表会等でも大変お世話になりました。深く感謝致します。秘書の西條由希さん，高橋みつるさんには，研究を進めるにあたり様々な事務作業をして頂き，研究を円滑に進められるようにサポートして頂きました。田中一志さん(博士課程前期)には，遠方から研究を進めるにあたって，研究開始当初からサポートして頂きました。深く感謝致します。

東北大学社会人博士後期課程での研究を進めるにあたり，仕事と研究の両立をサポートして頂き，本研究の進展のために並々ならぬご尽力を賜ったグローリー株式会社研究開発センターの亀山博文 センター長に厚くお礼申し上げます。また，本研究を始めるきっかけを作って頂き，研究推進をサポートして頂きました神瀬陽二郎 技師長に深く感謝致します。研究の進捗具合と仕事の調整をして頂きました安達和隆さん，藤原一雄グループマネージャーに感謝致します。また実験の推進，実験システムの開発においてご協力頂きました中嶋康博さんに感謝致します。最後に，研究活動に対する理解や日々の健康について気遣って頂いた家族にこの場を借りて深く感謝の意を表します。皆様本当にありがとうございました。