

親和的情報空間における人物認識の研究

2008年3月

福添孝明

目次

第 1 章	親和的情報空間における人物認識の研究	1
1.1	親和的情報空間	1
1.2	人物状況認識を実現するための機能群	2
第 2 章	人物追跡エージェント	4
2.1	従来研究との比較	4
2.2	処理の流れ	5
2.2.1	処理の概要	5
2.2.2	選択式テンプレート画像の学習	5
2.2.3	テンプレートマッチング	7
2.2.4	人物状態の推定	8
2.2.5	テンプレート画像の更新	11
2.2.6	アクティブカメラ制御	11
2.2.7	探索範囲の制御	11
2.2.8	カルマンフィルタによる一期先テンプレート位置予測	13
2.3	評価実験	14
2.3.1	実験条件	14
2.3.2	実験環境	15
2.3.3	処理結果について	15
2.3.4	実験結果	17
2.4	考察	21
2.4.1	誤認識について	21
2.4.2	状態推定について	21
2.4.3	2 階と 3 階での実験における追跡成功率の差について	22
第 3 章	人物同定エージェント	32
3.1	まえがき	32

3.2	従来研究とその課題	33
3.3	人物同定に用いる特徴の検討	34
3.4	処理の流れ	36
3.4.1	全体の概要	36
3.4.2	人物領域取得	37
3.4.3	頭領域取得	38
3.4.4	人物特徴取得	39
3.4.5	ベイズの定理に基づく事後確率値算出	39
3.4.6	人物同定の結果確定	41
3.5	評価実験・考察	41
3.5.1	各評価実験の概要	41
3.5.2	複数特徴統合に基づく人物同定手法の評価実験	41
3.5.3	MAP 推定と順位点蓄積手法の比較評価実験	47
3.5.4	行動習慣性特徴を用いた場合の評価実験	56
3.5.5	在席習慣性特徴を用いた場合の評価実験	56
3.5.6	滞留習慣性特徴を用いた場合の評価実験	63
3.5.7	総合実験	67
第4章	人物状態認識エージェント	76
4.1	従来手法	76
4.2	目の状態認識による内部状態の推定	76
4.3	提案手法の流れ	77
4.3.1	概要	77
4.3.2	目領域検出	77
4.3.3	瞳部分検出方法	78
4.3.4	エッジ処理	78
4.3.5	Hough 変換	80
4.3.6	閉目・半瞬目の判断	81
4.4	評価実験	82
4.4.1	実験条件	82
4.4.2	脳波計測実験結果	83
4.4.3	CV 実験結果	84
4.4.4	全瞬目認識評価	85
4.4.5	半瞬目状態認識評価	87
4.5	考察	88

4.5.1	脳波との比較	88
4.5.2	CV 性能について	88
第 5 章	総括	92
5.1	各エージェントにおける研究成果の要約	92
5.2	エージェント連携と親和的情報空間の構築	92
5.3	まとめ	93
付録 A	人物状態認識の応用研究 1	97
A.1	序論	97
A.1.1	研究の背景	97
A.1.2	ファカルティ・ディベロップメント	97
A.1.3	学生による授業評価の問題点	98
A.1.4	授業評価システムの従来研究	98
A.2	提案手法の説明	99
A.2.1	概要	99
A.2.2	着席と顔向き判定	99
A.2.3	統合判断	101
A.3	評価実験	102
A.3.1	実験環境	102
A.3.2	実験結果	102
A.4	考察	103
A.5	結論	104
付録 B	人物状態認識の応用研究 2	112
B.1	研究の背景と目的	112
B.2	提案手法の説明	112
B.2.1	概要	112
B.2.2	大津の二値化による自動閾値選定	113
B.2.3	エッジ抽出	114
B.2.4	Hough 変換による鞆検出	114
B.2.5	監視領域内処理	115
B.3	評価実験	115
B.3.1	実験環境	115
B.3.2	実験結果	116
B.3.3	他形状鞆を用いた実験	116

B.4	考察	117
B.5	結論	117
付録 C	動画像処理システム	123
C.1	はじめに	123
C.2	開発環境について	123
C.2.1	Visual Studio 2005 について	123
C.2.2	開発環境のカスタマイズ	123
C.2.3	開発環境の説明	124
C.2.4	Debug と Release	125
C.3	動画像処理の実装方法	126
C.4	MILMIL の解説	127
C.4.1	概要	127
C.4.2	設定項目について	127
C.4.3	処理関数	129
C.4.4	定義項目について	130
C.4.5	画像処理の方法	130
C.4.6	AVI ファイルを入力とする方法	131
C.4.7	従来の MILMIL シリーズとの相違点について	131
C.4.8	処理の流れ	132
C.5	MILMIL シリーズの開発履歴	133
参考文献		135
謝辞		138

目次

2.1	人物追跡手法の流れ図	6
2.2	テンプレート取得位置	7
2.3	人物進入時	8
2.4	テンプレート領域の配置	9
2.5	選択されたテンプレートの例	10
2.6	テンプレートの移動における制約	12
2.7	天井に設置したアクティブカメラ (VC-C4R)	15
2.8	2階実験室全景	16
2.9	2階天井カメラ位置	16
2.10	3階居室全景	17
2.11	3階天井カメラ位置	18
2.12	処理結果画像の説明	19
2.13	状態推定結果の表示色	19
2.14	追跡中における処理結果 1 (実験 1)	23
2.15	追跡中における処理結果 2 (実験 1)	24
2.16	追跡中における処理結果 3 (実験 1)	25
2.17	追跡中における処理結果 1 (実験 2)	26
2.18	追跡中における処理結果 2 (実験 2)	27
2.19	追跡中における処理結果 3 (実験 2)	28
2.20	追跡中における処理結果 1 (実験 3)	29
2.21	追跡中における処理結果 2 (実験 3)	30
2.22	追跡中における処理結果 3 (実験 3)	31
3.1	滞留習慣性の説明図	36
3.2	提案手法の処理の流れ	37
3.3	人物領域抽出の説明図	38

3.4	頭領域抽出の説明図	38
3.5	人物特徴の定義	39
3.6	実験室（左）・研究室（右）	42
3.7	学習対象人物	43
3.8	頭幅と人物領域の高さとの比率	45
3.9	頭領域の明度	45
3.10	人物認証成功例	46
3.11	各人物特徴を削除したときの認証精度推移	46
3.12	被験者（左上から右へ Person01 ~ 14）	48
3.13	被認証者の動き	49
3.14	人物認証結果 1	50
3.15	人物認証結果グラフ 1	50
3.16	人物認証結果 2	51
3.17	人物認証結果グラフ 2	51
3.18	Person12 の頭領域検出失敗例	54
3.19	形状特徴のみ利用した場合の人物同定精度	57
3.20	行動習慣性特徴を加えた場合の人物同定精度	57
3.21	実験環境	58
3.22	実験環境の模式図	59
3.23	在席習慣性利用手法別の同定成功の割合	61
3.24	順位点蓄積の様子（成功例）	62
3.25	順位点蓄積の様子（失敗例）	62
3.26	実験環境と処理例（73号講義室）	63
3.27	実験環境と処理例（71号講義室）	63
3.28	実験環境と処理例（実験室）	64
3.29	順位点蓄積の結果（上：成功例，下：失敗例）	66
3.30	実験環境（基礎実験）	69
3.31	時間方向に統合した事後確率の推移（基礎実験）	70
3.32	実験状況の説明図（有用性評価実験）	71
3.33	実験環境（有用性評価実験）	72
3.34	人物特徴 F2 の頻度分布例（有用性評価実験）	72
4.1	処理の流れ図	78
4.2	目領域設定処理の流れ図	79
4.3	瞳部分検出の流れ図	80

4.4	瞳部分検出の流れ図	80
4.5	開眼状態	81
4.6	半瞬目状態	81
4.7	閉眼状態	82
4.8	AD 変換装置 (CEDmicro1401)	82
4.9	日本光電社製脳波計 (EEG-4418)	83
4.10	PC(PowerMacintosh8100AV)	83
4.11	エレクトロキャップ	84
4.12	エレクトロキャップ電極位置	84
4.13	タスク最初 15 分間における脳波解析結果	85
4.14	タスク最後 13 分間における脳波解析結果	86
4.15	開眼時 (左:顔画像, 右上:目領域, 右下:円描画)	89
4.16	誤認識 1(左:顔画像, 右上:目領域, 右下:円描画)	90
4.17	半瞬目時 (左:顔画像, 右上:目領域, 右下:円描画)	90
4.18	誤認識 2(左:顔画像, 右上:目領域, 右下:円描画)	90
4.19	瞬目時 (左:顔画像, 右上:目領域, 右下:円描画)	91
4.20	誤認識 3(左:顔画像, 右上:目領域, 右下:円描画)	91
4.21	誤認識 4(左:顔画像, 右上:目領域, 右下:円描画)	91
5.1	親和的情報空間の物理的配置図	94
5.2	人物追跡エージェント	95
5.3	人物同定エージェント	95
5.4	人物状態認識エージェント	96
A.1	着席判定におけるフローチャート	99
A.2	アクティブ判定の処理の流れ	100
A.3	アクティブ判定フローチャート	100
A.4	顔向き判定処理の流れ	101
A.5	受講者集中力評価システム処理の流れ	101
A.6	DV カメラの設置	105
A.7	DV 出力画像 (鹿児島大学情報工学科棟 72 号講義室)	106
A.8	選択した 12 名	106
A.9	着席判定の処理例	107
A.10	人物動作判定	108
A.11	人物顔向き判定	109

A.12	ターゲット 1 の統合判定結果	110
A.13	ターゲット 2 の統合判定結果	110
A.14	ターゲット全員の統合判断処理結果	111
B.1	全体の処理の流れ	113
B.2	$x - y$ 平面と $\rho - \theta$ 平面	115
B.3	監視領域の設定方法	116
B.4	人物追跡の処理例	118
B.5	不審行動判定の処理例 1	119
B.6	不審行動判定の処理例 2	120
B.7	不審行動判定の処理例 3	121
B.8	不審行動判定の処理例 4	122
B.9	その他の鞆検出例 1	122
B.10	その他の鞆検出例 2	122
C.1	設定のインポート	124
C.2	インポート手順	125
C.3	開発環境	126

表目次

2.1	各人物状態の定義	10
2.2	評価実験に用いた計算機の性能	15
2.3	実験 1 における追跡成功率と状態推定成功率	20
2.4	実験 2 における追跡成功率	20
3.1	流動環境下で人物同定に利用可能な人物特徴	36
3.2	人物特徴の有効範囲	42
3.3	学習時間と頭位置検出率	44
3.4	人物認証結果	44
3.5	頭幅取得成功率	52
3.6	人物特徴の有効性評価	52
3.7	時系列認証手法別認識成功率	53
3.8	人物認証成功率 1	53
3.9	人物認証成功率 2	54
3.10	人物認証成功率 3	55
3.11	各人物の累積順位点における順位結果	60
3.12	順位点蓄積手法	61
3.13	在席習慣性利用手法別の同定成功の割合	61
3.14	実験環境の状態	64
3.15	人物同定に成功した人数	65
3.16	人物特徴と人体寸法データベースの対応関係	68
3.17	被験者体型（基礎実験）	69
3.18	実験結果（基礎実験）	70
3.19	被験者体型（有用性評価実験）	73
3.20	実験結果（有用性評価実験）	74
4.1	目の状態と内部状態との定性的関係	77

4.2	評価実験に用いた計算機の性能	85
4.3	タスク最初 15 分間における結果 (全瞬目)	86
4.4	タスク最後 13 分間における結果 (全瞬目)	87
4.5	タスク最初 15 分間における結果 (半瞬目)	87
4.6	タスク最後 13 分間における結果 (半瞬目)	87
A.1	統合判断方法	102
A.2	実験環境	102
B.1	実験環境	116

第 1 章

親和的情報空間における人物認識の研究

1.1 親和的情報空間

親和的情報空間とは，ユーザーが意識的にサービスを要求する様な従来の情報空間とは異なり，自発的に情報を取得し適切なサービスを提供する情報空間と定義する．筆者の所属研究室はこの実現を長期目標としている．以下に親和的情報空間の詳細な定義を引用する．

（引用開始）

近年，PC や民生用ビデオカメラの普及に伴う小型化，低価格化，性能向上が急速に進行し，従来から研究成果が蓄積されてきた画像処理・認識技術（コンピュータビジョン，以下CV）が様々な領域で実用に供される状況が整ってきた．一方で，21世紀は少子化・高齢化・自動化の傾向が更に高まり，セキュリティ，通信，娯楽，教育，福祉，介護，など様々な領域で，人間にできるだけ負担を掛けない状況における，人間・計算機システム間の高度に親和的なコミュニケーションが強く要求されていくと予想される．

従来のマンマシンインタフェース関連研究においては，例えばジェスチャーの認識や視線によるポインティングなど，計算機システムへ意識的・意図的に利用者の意思を伝えるという観点からのものが主流であり，利用者にシステムの存在を意識させずにその意図を読みとって適切な時点で適切な応答を行うための技術蓄積は未だ不十分である．複数の人間が所在・居住する環境下で上記の要請を実施するためには，リアルタイムで3次元，動きの情報を抽出し，この情報に基づき環境内に発生している状況を認識することが必要となる．

しかしながら，CV 分野における従来の画像認識は，静止環境下で単独の静止物体（剛体）を対象に，1 視点，形状ベースで行うものが主体であり，動的に変化し且つ家具などの遮蔽物の存在により広い視野が確保できない環境下での，人物（非剛体）・状況の認識においては必ずしも有効に働かないという問題点があった．

そこで本研究プロジェクトでは，研究代表者がこれまで行ってきた，従来の形状ベースに代わる動きベースの認識法の提案，合目的な知能システム構築を目的とした複数処理モジュール統合方式の提案，居住空間での人物状態認識プロトタイプの開発の基盤の上に立ち，並列分散画像認識の枠組下で，形状のみならず動きの情報も利用して複数視点の画像を解析することによる，人間志向型の高度なコミュニケーション実現を目的とし，居住環境および所在者の状態認識に基づき“気づき”・“気づき”を自動的に呈示・発現するシステムの構築を目指した．

このシステムの総体を親和的情報空間（Friendly Informative Cyber Space: FICS）と称し，人間の側が計算機システムに合わせ働きかけねばならない従来の閉鎖的情報空間とは異なる，人物の所在する 3 次元空間，および自分自身の認知空間とシームレスに結合可能な情報空間を構築することを目的とした．

（引用終了）

この親和的情報空間は大きく分けて，次の二つの要素で構成される．

- 情報空間内の人物状況認識
- 人物状況や知識に基づく適切なサービスの推測とその実施

筆者は，主に前者の「情報空間内の人物状況認識」の実現を目標として研究を行ってきた．本書はその研究成果をまとめた物である．

1.2 人物状況認識を実現するための機能群

空間内の人物に対して必要なサービスを提供するとき，そのサービスを推測するために人物に関する様々な情報が必要になる．例えば，情報空間内に登録者が進入したとき，息が上がっている，汗をかいているなどの外見に表れる特徴から，人物が疲労していることを知ったとする．該当人物のスケジュールを確認するために人物同定を行い，その結果，前の時間まで出張に出ていることが分かったとする．そこで，飲料提供が適切なサービスであると判断した場合，それを提供するために人物位置を知る必要がある．

人物の状況は数多く存在するが，その中でも人物の位置取得や人物同定，その状態推定が最も基本となる機能になると考えた．そこで，各機能ごとに個別の研究目標を設定し，その研究によって得られた知見から「情報空間内の人物状況認識」を構築するのに必要な

事項を考察することにした。

第 2 章では、パソコンで姿勢制御可能なカメラにて得られた画像から、情報空間内に進入した人物を追跡する研究を説明する。第 3 章では、人物が自由に流動している状態にて人物を同定する手法を説明する。また、人物の状態推定については扱う対象が広いため、疲れ状態に限定した。その研究結果を第 4 章にて説明する。最後の第 5 章では、各機能ごとの研究から得られた知見を用いて「情報空間内の人物状態認識」を実現するための考察を行い、本論文をまとめる。

なお、A 章と B 章では、人物状態認識エージェントの研究にて行った、親和的情報空間との関係性は薄いですが、有益な研究成果を紹介する。C 章は、これらの研究を遂行するに当たって、開発の基盤となったプログラムの解説書である。研究室の備忘録を目的として掲載した。

第 2 章

人物追跡エージェント

2.1 従来研究との比較

動画像中から特定の対象を認識・追跡する手法が数多く提案されている [1]。その代表的なものに追跡対象内の特徴的な矩形領域をテンプレートとしマッチングにより実現する方法や、オプティカルフローを用いて背景と対象とのフローの違いを用いて対象を切り出し追跡を行う方法 [2][3] がある。

更に、Condensation の手法 [4] を用いて追跡対象の輪郭形状をパラメータ空間における確信度の分布で表し、確信度分布の伝播と再構成を繰り返すことにより輪郭形状を推定しながら対象を追跡する手法 [5] がある。また、K-means トラッキングのように対象と非対象の両方の情報を保持し、k-means クラスタリングによって各画素をターゲットと非ターゲットに分けながら追跡を行う手法 [6] がある。

オプティカルフローを用いた研究では動きの情報を利用しているため、複雑背景下においても移動物体の動きが正しく抽出されれば安定に追跡できるという利点がある。しかし、一旦人物が静止してしまうとフローが検出されないため、追跡が不安定になる。また一方、Condensation を利用した方法では観測値が非ガウス型になる状況でも過去の観測値に基づいた予測値と現在の観測値から最適な推定値を求めることができる。

K-means トラッキングによる方法では追跡対象に背景が入り混じる状況下でも追跡を行うことができる。しかし、Condensation を用いる方法や K-means トラッキングでは初期の対象情報の学習に追跡すべき対象の事前情報が与えられないため、出現後短時間で学習しなければならない状況下では適応が困難であると考えられる。

一方、テンプレートマッチングとは入力画像と参照画像（テンプレート画像）の類似度により追跡対象を判別する手法であり、参照画像の学習が容易に行えるという利点がある。類似度を評価する尺度として、SSD (Sum of Squared Differences), SAD (Sum of Absolute Differences), CC (Cross Correlation) などがある。

単一のテンプレートを用いた場合の追跡では対象の見え方の変化やオクルージョンに対して弱いため、テンプレートの数を増やし複数のテンプレートを用いたマルチテンプレートマッチングによる追跡手法が提案されている [7]。この手法では背景差分により 2 値化した画像から人物領域を外接四角形で囲み、その中を背景領域か人物領域かのどちらかになるまで 4 分割を行い、人物領域になったブロックをテンプレート画像とし追跡を行っている。しかし、対象の見え方の変化に伴うテンプレート画像の更新が行われていないため、複数の人物が自由に流動している状況下で特定人物を追跡し続けることは困難であると考えられる。

こうしたオブジェクトの抽出・追跡等を目的とするだけでなく、シーン中に生じる状態にラベルを付ける研究 [8] が盛んになりつつある。文献 [9] では階層的注視点制御の有効性を示すため複数人物の追跡問題を取り上げ、複数テンプレートの集合から人物を取り扱い、人物間の距離や重なり方などから状況に応じた処理を行っている。しかし、実際に対処できる状況が人物同士の遮蔽にのみに限られ環境物による遮蔽は考えられていない。

2.2 処理の流れ

2.2.1 処理の概要

本提案手法の流れを図 2.1 に示す。始めにテンプレートの学習を行い、その特徴から人物の行動状態を推定する。推定結果を用いて、テンプレートの更新や人物の行動予測を行う。

2.2.2 選択式テンプレート画像の学習

追跡に用いるテンプレート画像は非拘束状態で自動学習する。図 2.2 における領域 A と B (A, B は 3×10 ピクセルの長方形) の間を人物が通過したときにテンプレート画像学習領域 C (16×16 ピクセルの正方形が 12 行 \times 2 列) の中から、移動後の姿勢変化が比較的少なく、安定に追跡できると思われる胴体領域を取り出し、さらにその中で差異のできるだけ大きいものを選択しテンプレート画像とする。テンプレート画像の学習を選択式とした理由は、テンプレートが同一領域をテンプレートとすることを防ぐためであり、パターンの異なる特徴的な画像をテンプレートとして利用するためである。テンプレート画像の学習は次に示す手順で行う。

1. プログラム起動時に領域 A, B の背景画像を学習し、その後入力画像との背景差分により領域 A, B 共に変化が検出されたとき、テンプレート画像学習領域にある画像を 24 個のテンプレート画像候補とする。

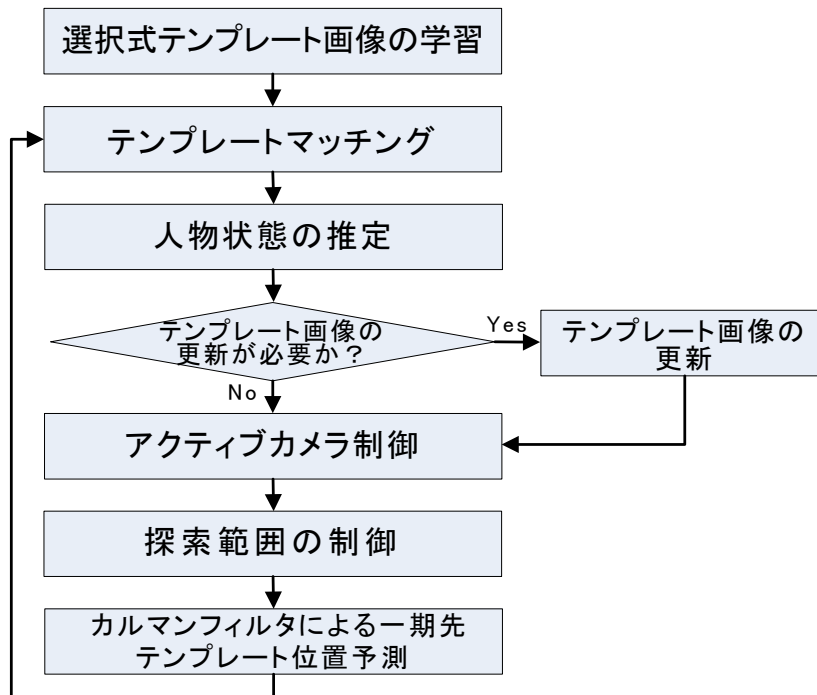


図 2.1 人物追跡手法の流れ図

2. テンプレート画像候補を上から背景差分により変化を検出していき、初めて変化のあった場所を頭部上端として検出する。
3. 頭部上端として検出されたテンプレートを含めて下に3段分のテンプレート（計6個）を頭部領域として排除し、その下4段目から9段目までにあるテンプレート（計12個）を胴体領域として更に絞ったテンプレート画像候補とする。
4. 胴体領域における12個それぞれのテンプレート画像候補のヒストグラムを作成し、各テンプレート間でのヒストグラムの差が大きい候補画像から順に実際にマッチングで用いるテンプレート画像（計8個）としていく。

図 2.3 に領域 A,B が共に変化検出された直後の図を示す。図 2.4 はテンプレート領域における各領域の配置を示す。図 2.5 に選択されたテンプレートの例を示す。

2 や 3 において頭部上端を検出した理由は、被験者の身長の高さを考慮しなくても適切に胴体領域からのテンプレート画像の取得を行えるからである。1 や 2 における変化の検出は照明の揺らぎを考慮し用意した基準値を用いた。4 においてヒストグラムを用いてテンプレート間の差異を求めた理由は、テンプレート画像学習後すぐに追跡を開始しなければならないため、処理の高速化を図るためである。

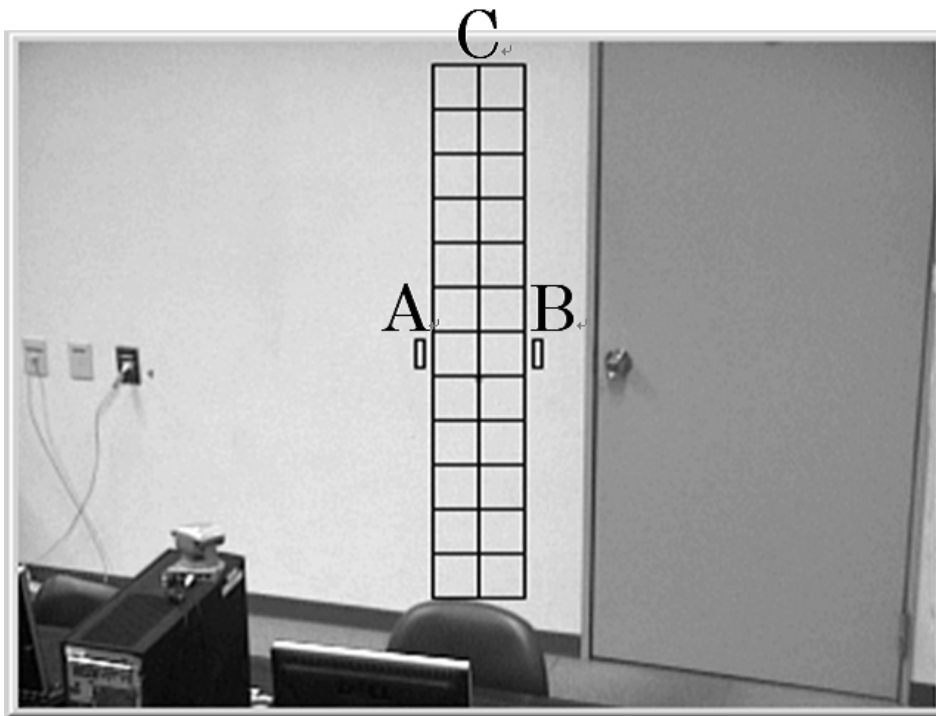


図 2.2 テンプレート取得位置

2.2.3 テンプレートマッチング

1 つ前のフレームからカルマンフィルタによって求められた一期先テンプレート位置を探索範囲の中心として、テンプレートマッチングを行い人物位置を求める。このとき評価指標として相関関数を用いる。一般に相関係数はテンプレート画像及び入力画像の濃度分散に関して正規化が行われるため、濃淡階調の線形変換の下では変化が少なく、コントラストや明度が大きく異なる場合でも用いることができるという利点がある。

テンプレートマッチングは二段階の構成で行う。一段階として、探索範囲内のテンプレート画像と同じ大きさの入力画像とテンプレート画像を全画素において比較し、その明度差分の和を求める。その差分値が基準値以上なら探索範囲内のその部分には人物領域は含まれないとして次の領域の差分を計算する。基準値以下なら人物領域の可能性があるので、相関関数によって類似度を計算する。探索範囲内を 2 ピクセルずつずらしながら差分の計算及び相関係数を求める。探索範囲内で求められた各位置の相関係数をソートし、最も相関値の高い位置をそのテンプレートの類似度の高い位置とする。

この計算を 8 個のテンプレートに対して行い、求められた 8 つの最大相関位置の重心（以下、全テンプレート重心位置と呼ぶ）を人物位置とする。2.2.2 項で学習した $m \times n$ 画素のテンプレート画像 $T(i,j)$ を入力画像 $I(i,j)$ 内の探索範囲内で走査させ (2.1) 式に

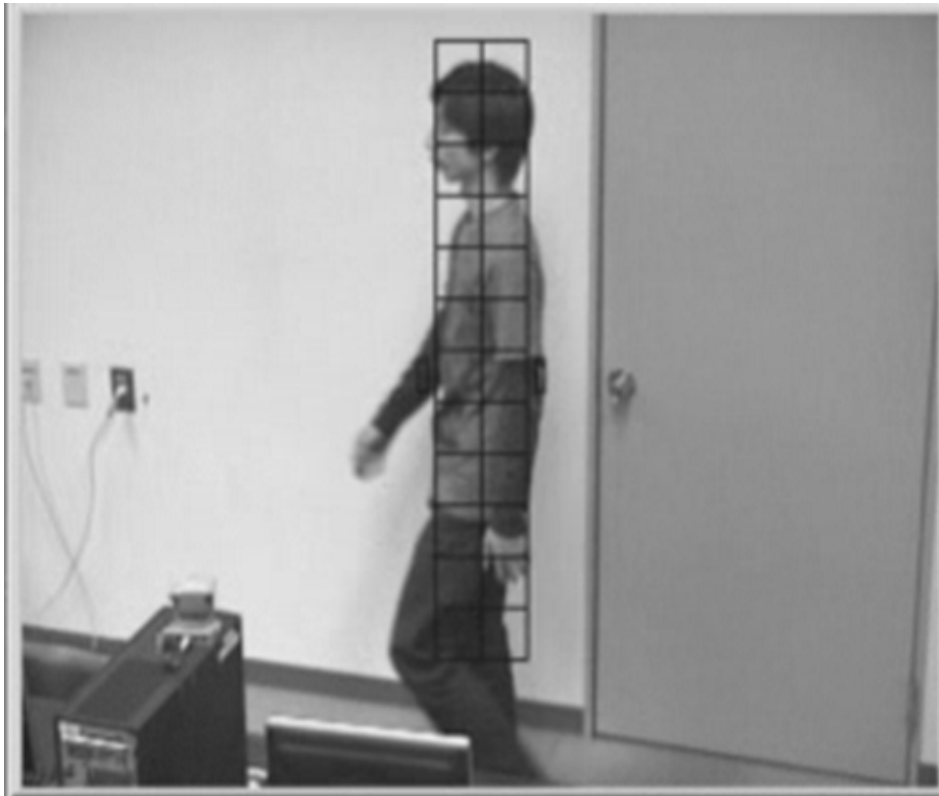


図 2.3 人物進入時

より相関係数 $R(a,b)$ を求める .

$$R(a,b) = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \{I(i,j) - \bar{I}\} \{T(i,j) - \bar{T}\}}{\sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \{I(i,j) - \bar{I}\}^2} \sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \{T(i,j) - \bar{T}\}^2}} \quad (2.1)$$

$$\bar{I} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} I(i,j), \bar{T} = \frac{1}{mn} \sum_{j=0}^{n-1} T(i,j)$$

2.2.4 人物状態の推定

追跡中に起こりうる人物の状態を5つに分類することにより, 人物の状態認識やテンプレート画像の更新制御, さらに遮蔽状態における探索範囲の拡大化を行う. 各状態の分類には個々のテンプレートの挙動や2.2.3項で示した入力画像とテンプレート画像との明度差分の結果を用いる.

テンプレートの挙動として一つめの条件として, テンプレート全体の移動量を参照する. 移動量がほぼ0ならば人物は止まっていると考えられ, 移動量があるならば人物は動いていると考えられる.

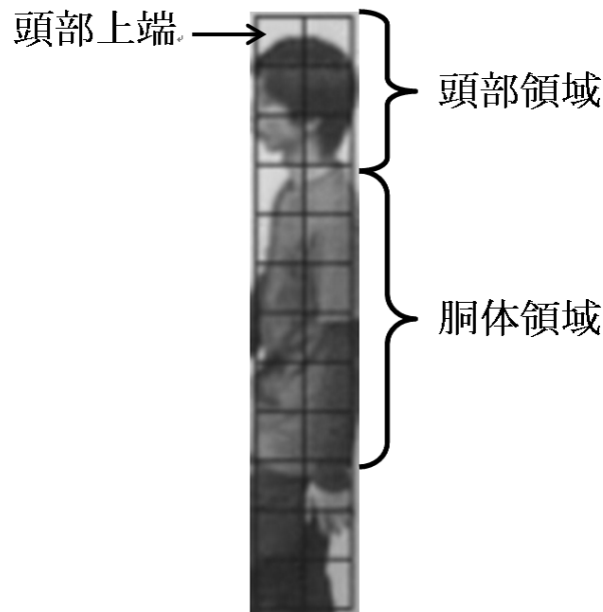


図 2.4 テンプレート領域の配置

二つめの条件として 8 個のテンプレートの重心と個々のテンプレートとの間の距離の総和（以下，テンプレート間距離和と呼ぶ）を人物状態の推定指標に用いる．人物状態の分類を行うもう 1 つの指標として，2.2.3 項で示した入力画像とテンプレート画像との明度差分の結果を用いる．

2.2.3 項におけるテンプレートマッチングにおいて一段階の差分値が基準値以上となることが探索範囲内の全ての位置において起こった場合，その探索範囲内には人物領域が存在しないものとなる．この状態のテンプレートを「デッドテンプレート」と呼ぶ．このデッドテンプレートの数を，人物状態を分類する三つ目の指標とする．

追跡中に起こりうる人物の状態を 5 つに分類しその定義を示す．人物同士のすれ違いや人物の一部にオクルージョンが発生する状態を「部分的遮蔽」，人物が部屋から退室したり，環境物などによって完全に遮蔽されたりする状態を「完全遮蔽」とする．また，人物が追跡中に立ち止まる状態を「立ち止まり静止」，人物が室内にある椅子に着席した状態を「着席」とする．更に，これらのどれにも属さない状態を「移動中」とする．表 2.1 に各状態とそれらを分類する指標との関係を示す．

次に表 2.1 についての詳細な説明を行う．テンプレートの移動量において，「常に移動量がある」とは個々のテンプレートが常に移動している状態であり，「移動量 0」とは個々のテンプレートの移動が数フレームに渡って停止した状態である．またテンプレート間距離和における基準値 1,2 は実験的に設定した．更に，デッドテンプレートの数について「移動中」や「立ち止まり静止」，「着席」の状態では人物領域内に全てのテンプレートが

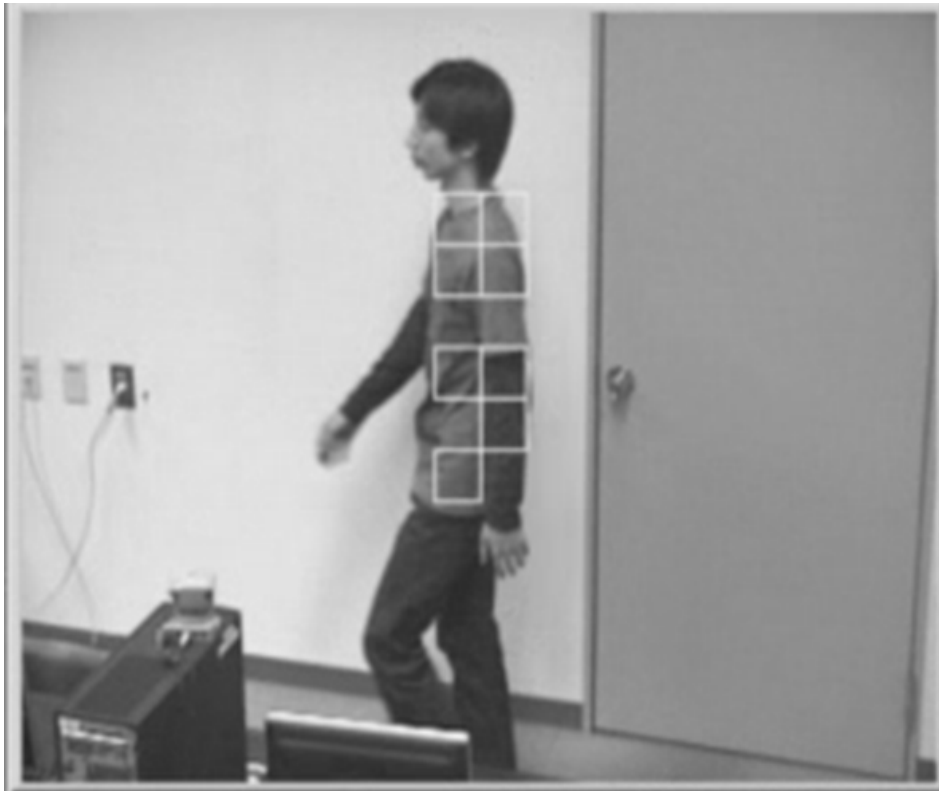


図 2.5 選択されたテンプレートの例

表 2.1 各人物状態の定義

	テンプレート移動量	テンプレート間距離和	デッドテンプレート数
移動中	常に移動量がある	基準値 1 以上	0
部分的遮蔽		基準値 1 以下	0~7
完全遮蔽			8
立ち止まり静止	移動量 0	基準値 2 以上	0
着席	移動量 0	基準値 2 以下	0

存在することになり、デッドテンプレートは発生しない。「部分的遮蔽」の状態では0~7個のテンプレートが遮蔽物によって遮蔽されることが考えられる。「完全遮蔽」においては人物が入力画像中に存在しないため全てのテンプレートがデッドテンプレートとなる。

2.2.5 テンプレート画像の更新

追跡中において人物が移動しても見え方の変化が少ない胴体領域からテンプレート画像を学習しているが、移動が大きくなった場合にはテンプレート画像を更新した方がより安定して追跡が行えるため、テンプレート画像を更新する。

しかし、更新するタイミングや場所をどこにするかという問題がある。そこで、本研究では2.2.4項により求められた人物状態の推定結果を用いて人物が遮蔽されておらず、かつ移動していて見え方に変化が生じている「移動中」の状態においてテンプレート画像を更新する。さらに、更新する場所は1フレーム前において最大相関位置であった場所とする。

上記に示すテンプレート画像更新のタイミングと場所で8個のテンプレート画像を更新していく。マッチングの結果8個それぞれのテンプレートの最大相関値の中で最も高い相関値を持つテンプレート画像を更新する。更新は対象の見え方の変化が顕著になる数フレームおいた間隔で行う。

2.2.6 アクティブカメラ制御

追跡対象が入力画像内から逸脱しないようカメラを制御する。具体的には、2.2.3項により求められた全テンプレート重心位置を、入力画像の中心に来ようカメラを制御する。入力画像中における人物の大きさをできるだけ一定に保ち、テンプレートマッチングを安定して行うため、カメラのズーム機能を利用する。全テンプレート重心位置へのカメラの光軸と、カメラからの鉛直方向下向きとのなす角度（カメラのチルト角）を利用しカメラと人物との距離を概算しズーム値を定める。

2.2.7 探索範囲の制御

探索範囲の可変性

テンプレートマッチングにおける探索範囲を状況に応じて変化させ、追跡精度と処理速度の両方の向上を図る。追跡開始時の数フレーム間の人物移動量によって探索範囲の基準サイズを決定する。それに1つ前のフレームにおけるカルマンフィルタによって求められた予測位置と実際にテンプレートマッチングによって求められた最大相関位置との誤差を足し合わせたものを、次のフレームにおけるテンプレートマッチングの探索範囲とする。誤差が小さい場合には探索範囲も小さくなり、処理速度も向上する。人物状態の推定結果からも探索範囲の制御を行う。人物状態が「部分的遮蔽」や「完全遮蔽」の状態においてデッドテンプレートが発生した場合、探索範囲を一時的に拡大させる。こうすることでい

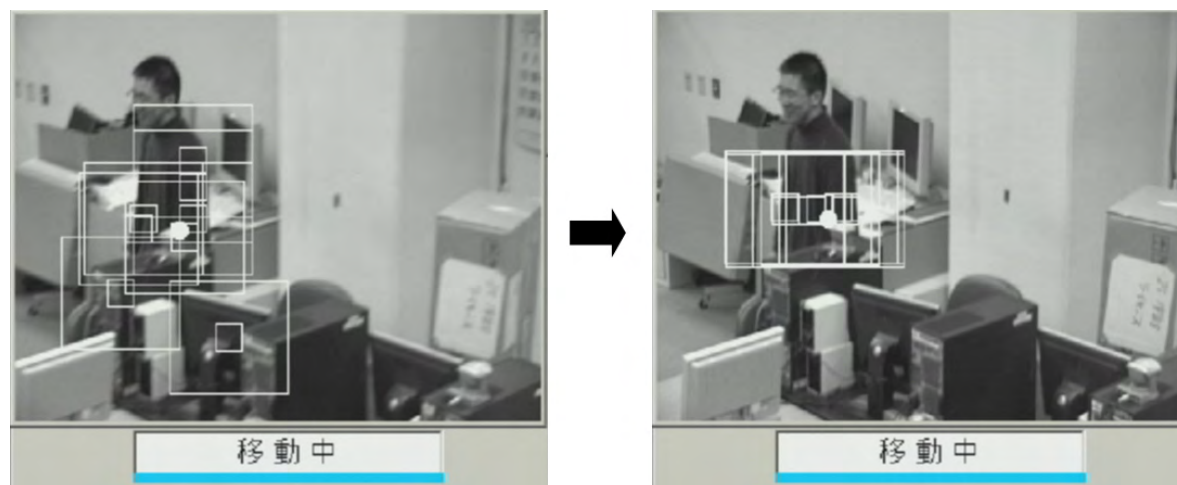


図 2.6 テンプレートの移動における制約

ち早く人物領域にマッチングを復帰させることができ、不安定な時間を短縮することができる。

テンプレートの移動における制約

テンプレート同士が重なったり離れ過ぎたりすると追跡が安定に行えないため、テンプレートの移動に制約を設けテンプレート学習時の配列をできるだけ保ちながら追跡を行えるよう制御する。テンプレートの X 軸方向の移動、Y 軸方向の移動に対してそれぞれ別々の制約を設ける。

全テンプレート重心位置の X 座標に、X 軸方向である程度の幅を加えた位置を通る Y 軸に平行な直線を考えたとき、その 2 本の直線の間を人物が存在する領域とする。その領域を外れたテンプレートは領域内に残っている全てのテンプレートから求めた重心位置へと戻される。

図 2.6 に X 軸方向における制約を受け、テンプレートの移動が起こった様子を示す。背景に残ったテンプレートが再び人物領域へ戻っていることが確認できる。テンプレート学習時における選択された直後のテンプレートの Y 座標をそれぞれ保持しておき、その後の追跡ではこの Y 座標からある決められた範囲内しか Y 軸方向に移動できないように制約を設ける。このようにすることにより、できるだけテンプレートを人物領域内で広げることができる。

2.2.8 カルマンフィルタによる一期先テンプレート位置予測

追跡におけるテンプレートの移動を線形予測フィルタの 1 つであるカルマンフィルタ [9] を用いて予測する。次のフレームにおけるテンプレートの位置を予測し、その予測した位置を中心にテンプレートマッチングを行う探索範囲を設ける。また一期先テンプレート位置予測は 8 個のテンプレートそれぞれ独立に行う。

一期先予測を行う理由は、次のフレームから人物を探索する際、過去のフレームにおける追跡結果を利用し人物位置を予測してマッチングを行う方が事前知識のない場合に比べ安定して追跡が行えると考えるからである。

本研究ではアクティブカメラを用いているためテンプレート位置予測を行うには座標系の変換が必要となる。座標系の利用は画像座標系からこのカメラ独自のカメラ内部パラメータ（パン・チルト角に基づくパラメータ）を使って世界座標系へ変換したものをを用いる。変換式を式 2.2 に示す。

$$\begin{aligned} X &= X_now + 0.727(x - 160) \\ Y &= Y_now + 0.727(120 - y) \end{aligned} \quad (2.2)$$

ここで X, Y は世界座標系の座標、 X_now, Y_now はカメラ内部パラメータによるカメラ光軸上の世界座標系の座標、 x, y は画像座標系の座標を表す。カルマンフィルタへの入力は式 2.3 に示す状態ベクトル X_t とする。ここで $X_t, Y_t, \dot{X}_t, \dot{Y}_t$ はそれぞれ世界座標系における時刻 t の人物の X 軸方向の位置、 Y 軸方向の位置、 X 軸方向の移動速度、 Y 軸方向の移動速度を示す。

$$X_t = (X_t \quad Y_t \quad \dot{X}_t \quad \dot{Y}_t)^T \quad (2.3)$$

以下カルマンフィルタの構成について述べる。テンプレートの移動は式 2.4 の線形離散時間確率システムの状態方程式に従うものとする。

$$\begin{aligned} X_{t+1} &= F X_t + w_t \\ Z_t &= H X_t + v_t \end{aligned} \quad (2.4)$$

ここで、 Z_T は観測ベクトル、 w_t は系に加わる平均値 0、共分散が Q_t の 4 次元外乱ベクトル、 v_t は平均値 0、共分散 R_t の 4 次元観測誤差ベクトル、 H は 4×4 の観測行列、 F は 4×4 の状態遷移行列である。本システムでは、 H を 4×4 の単位行列、 F は式 2.5 に示す行列とする。

$$F = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.5)$$

w_t, y_t が共にガウス性白色雑音で、互いに、また状態量 X_t とともに無相関である場合、本システムにおけるカルマンフィルタのアルゴリズムは次のように表される。

フィルタリング

$$\hat{X}_t = \bar{X}_t + K_t(X_t - H\hat{X}_t) \quad (2.6)$$

$$K_t = \hat{P}_t H R_t^{-1} \quad (2.7)$$

$$\hat{P}_t = (\bar{P}_t^{-1} - H^T R_t^{-1} H)^{-1} \quad (2.8)$$

一期先予測

$$\bar{X}_{t+1} = F\hat{X}_t \quad (2.9)$$

$$\bar{P}_{t+1} = F\hat{P}_t F^T + Q_t \quad (2.10)$$

初期条件

$$\hat{X}_0 = X_0$$

$$\hat{P}_0 = P_0$$

各テンプレートの状態ベクトルを式 2.6 から式 2.8 により最適な推定値更新を行う。そして式 2.9 から式 2.10 により一期先予測値へと変換される。最後に世界座標系から画像座標系へと変換され次のフレームにおいて、求められた予測値を中心とした探索範囲でテンプレートマッチングが行われる。

2.3 評価実験

前章で述べた方式に基づき人物の状態推定を利用した人物追跡システムを汎用パソコン(PC)上のソフトウェアとして試作し、提案手法の妥当性を確認する実験を行った。

2.3.1 実験条件

表 2.2 に今回の実験で用いたパソコンのスペックを示す。なお入力画像における解像度は 320×240 ピクセルとした。図 2.7 に使用したアクティブカメラの写真を示す。



図 2.7 天井に設置したアクティブカメラ (VC-C4R)

表 2.2 評価実験に用いた計算機の性能

CPU	Pentium4-3.0GHz (2 階実験室) Pentium4-3.8GHz (3 階居室)
メモリ	1GB
グラフィック	Matrox Millennium G550
キャプチャー	Matrox Meteor
カメラ	Canon VC-C4R (2 階実験室) Canon VC-C50iR (3 階居室)
OS	Microsoft WindowsXP professional
開発言語	Microsoft Visual C++6.0

2.3.2 実験環境

図 2.8 に 2 階実験室において天井カメラ付近から室内を撮影した写真を 2 枚パノラマ合成して示す。図 2.9 に 2 階天井カメラの位置を示す。図 2.10 に 3 階居室において天井カメラ付近から室内を撮影した写真を、図 2.11 に 3 階天井カメラの位置を示す。

2.3.3 処理結果について

処理結果の表示については、テンプレートの学習が行われるまでは初期位置を表示し、テンプレートの学習が終わると前章で述べた方式により求められた人物位置を表示する。個々のテンプレートの最大相関位置を 16×16 ピクセルの正方形で表し、またそれを囲むように 2.2.7 項で求められた探索範囲を同じく正方形で表し合わせて表示した。また、全テンプレート重心位置を直径 10 ピクセルの円で表示した。これら正方形や円の色は白と



図 2.8 2階実験室全景



図 2.9 2階天井カメラ位置

した．図 2.12 にその例を示す．2.2.4 項により求められた状態推定結果をその処理結果画面の下に表示した．状態推定結果は文字だけでなく，視覚的にわかりやすくするため色を付け表示した．状態と色との関係を図 2.13 に示す．



図 2.10 3 階居室全景

2.3.4 実験結果

この試作プログラムを用いて、情報工学科 2 階における実験室内で人物を移動させ、そこで起こりうる状態を推定しながら追跡を行う実験を行った。また、複雑環境における応用実験として情報工学科 3 階学生居室での実験と複数人が混在する状況下での実験をそれぞれ行った。各実験とも処理速度は 15FPS (Frame Per Second) であった。

実験 1 (2 階実験室)

以下に実験条件を示す。ただし、最低 2 回 5 つ全ての状態を起こしてもらう。

被験者数	10 名
試行回数	5 回
時間	120 秒間
状態	状態に関する指示なし
場所	情報工学科 2 階実験室

1 回の実験において追跡開始から終了するまでを 120 秒間とした。追跡の途中で人物を見失い追跡に失敗した場合は、追跡に失敗した最後の状態までを状態推定結果の評価に用いる。処理結果を図 2.14 から図 2.16 の 1~36 に示す。



図 2.11 3 階天井カメラ位置

実験 2 (3 階居室)

複雑環境における応用実験として 3 階学生居室においても追跡実験を行った。以下に実験の詳細な内容を示す。

被験者数	3 名
試行回数	10 回
時間	120 秒間
状態	状態に関する指示なし
場所	情報工学科 3 階学生居室

処理結果を図 2.17 から図 2.19 の 1~36 に示す。

実験 3 (2 階実験室)

追跡対象 1 名, 非追跡対象 2 名が存在する状況下ですれ違いが何度も起こる環境を設定し追跡実験を行った。処理結果を図 2.20 から図 2.22 の 1~36 に示す。

追跡結果の定量的評価

追跡結果の定量的評価について 2 つ示す。

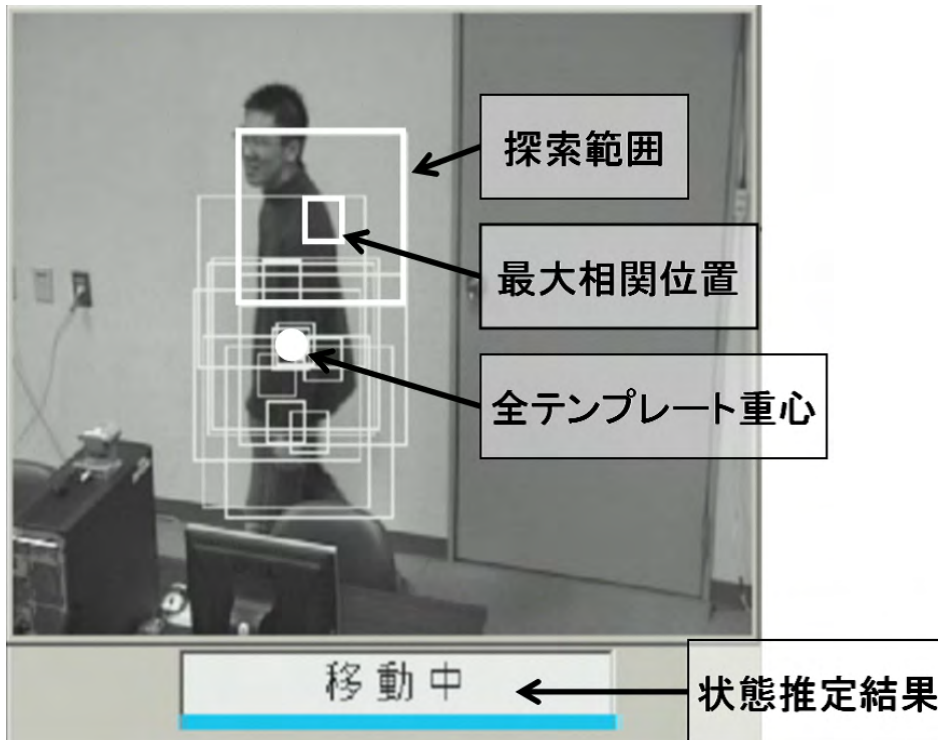


図 2.12 処理結果画像の説明

	移動中	部分的遮蔽	完全遮蔽	立ち止まり 静止	着席
色	水色	橙	赤	青	緑

図 2.13 状態推定結果の表示色

追跡成功率 追跡成功率を秒単位で計算する．1 回 120 秒間の追跡実験において見失わず追跡ができていた時間（秒）を 5 回の実験で合計し，分母を 600 秒（120 秒 × 5 回）とし追跡成功率を計算する．追跡成功の基準は 8 個のテンプレートの探索範囲のどこかに人物領域が含まれているものとする．

状態推定成功率 目視による人物の状態とプログラムにおける状態推定結果がどれくらい正しいかを示す．状態推定における正誤の単位は 1 回のイベント（「部分的遮蔽」や「着席」などの状態）ごととする．目視において 120 秒間の追跡中に何回イベントが起こったかを算出し，そのイベントごとに正しく推定されているかを計算する．イベントの推定結果はあるイベントが起こっている最中に最も多く推定された状態とする．ただし，「移動中」のイベントはそれ以外のイベントが終了してから次のイベントが開始までの間とす

表 2.3 実験 1 における追跡成功率と状態推定成功率

	追跡成功率	状態推定成功率				
		移動中	部分的遮蔽	完全遮蔽	立ち止まり	着席
被験者 A	53.5	98.1	90.0	81.8	90.0	33.3
被験者 B	85.8	93.0	87.0	90.9	55.6	72.7
被験者 C	100	97.8	70.0	100.0	63.6	63.2
被験者 D	74.5	80.9	94.7	100.0	14.3	45.5
被験者 E	83.8	93.5	85.7	100.0	69.6	66.7
被験者 F	100.0	97.6	72.7	100.0	68.4	36.8
被験者 G	71.2	93.3	85.7	100.0	100	60.0
被験者 H	81.0	95.7	70.8	100.0	70.6	22.2
被験者 I	78.2	87.8	71.0	100.0	92.0	70.0
被験者 J	13.0	100.0	57.1	33.3	0.0	0.0
合計 (10 人)	74.1	93.8	78.6	94.5	68.5	49.2

表 2.4 実験 2 における追跡成功率

	追跡成功率
被験者 A	54.9 % (659/1200)
被験者 F	70.6 % (847/1200)
被験者 G	31.7 % (380/1200)
合計 (3 人)	52.4 % (1886/3600)

る。つまり、ある「移動中」以外のイベントとイベントの間が「移動中」のイベントとなる。状態推定成功率はその状態が確認された回数における実際に正しく推定された回数の割合を百分率で表したものとする。表 2.3、表 2.4 にそれぞれ実験 1 と実験 2 における追跡成功率と状態推定成功率を示す。

2.4 考察

2.4.1 誤認識について

室内における人物追跡の実験を行い 2.3.4 項に定義した追跡成功率，状態推定成功率を示した．追跡成功率における追跡失敗の原因は背景との誤認識であるが，さらに詳しくこの誤認識の理由を調べると 2 つある．1 つは単純な背景領域との誤認識である．追跡中に人物後方の背景や手前にある机・パソコン等の環境物と取得したテンプレート画像とがパターンの似通っている場合，その部分がテンプレートマッチングにおける探索範囲内に含まれてしまった際には人物領域と誤認識することがある．2.2.7 項に示したテンプレート制約にて同時に 4 個のテンプレートまでなら人物領域内に戻されることがあるが，それ以上の数のテンプレートが誤認識を起こしてしまった場合は逆に背景の方に（誤認識の方に）人物領域内のテンプレートが戻され追跡に失敗する．

もう 1 つの誤認識の原因は，ダンボールによって部分的に遮蔽される「部分的遮蔽」状態において探索範囲の拡大を行った際に，遮蔽されていない部分や遮蔽物を通り過ぎた人物を捕らえるのではなく，拡大によって誤認識の恐れのある背景（主に机やパソコン）をその探索範囲内に含めてしまうため，背景と誤認識を起こしてしまい追跡に失敗することである．2.2.8 項に示したカルマンフィルタによるテンプレート移動の一期先予測機能により，ある程度の速度以上で通過すればダンボールを通り過ぎた人物を捕らえることができるが，予測が十分できない速度で通り過ぎると上記のような誤認識を起こしてしまう．

この誤認識による追跡の失敗を防ぐ方法として，「部分的遮蔽」状態において唯一確認できる人物領域である頭部に着目し，「部分的遮蔽」状態を認識すれば直ちに遮蔽されたテンプレート上部付近から頭部を検出し，遮蔽されている間はその頭部を追跡対象とする方法が考えられる．そして遮蔽物を通り過ぎた後はまたテンプレートマッチングにより追跡を再開する．

2.4.2 状態推定について

状態推定成功率について「立ち止まり静止」状態と「着席」状態では，立っていれば胴体領域が大きく確認でき，座っていれば机等に遮蔽され上半身部分しか確認できないことを利用し，人物領域内でのテンプレートの散らばりをテンプレート間距離和という形で表している．そのため，座った状態でも十分に机等に遮蔽されなかったり，数個のテンプレートが背景等と誤認識をしてしまったりすると，テンプレート間距離和が基準値を超え「着席」状態でも「立ち止まり静止」状態と推定してしまう．またその逆で，「立ち止まり

静止」状態でも十分に人物領域内でテンプレートが散らばらないことがあると「着席」状態と推定してしまうことがある。これがこの2状態の状態推定成功率が他の状態より低い原因である。

2.4.3 2階と3階での実験における追跡成功率の差について

3階居室での実験の方が2階実験室での実験より21.7%の追跡成功率低下が認められる。同条件での実験ではないとしても、大きな差が生じていると思われる。追跡成功率の低下の原因として2つ考えられる。

1つ目は単純に背景が複雑になったことによる誤認識の増加であると考えられる。処理結果を確認すると、背景と誤認識してしまったテンプレートが人物領域に戻る(2.2.7項に示した制約による)ことがたびたび確認できた。

2つ目は3階居室での実験は2階実験室と比べカメラと人物との距離が大きく変わることが考えられる。2.2.6項に示してあるように入力画像中における人物の大きさをできるだけ一定に保つようズーム機能を取り入れているが、移動しても人物領域内で位置の変わらない定まった1点(例えば頭部中心など)が求められていないため正確な距離が求められず概算での距離となりズーム機能が十分ではなかった。そのため入力画像中における人物の大きさがテンプレートを学習した初期の状態と異なりテンプレートマッチングの結果が不安定になったと考えられる。

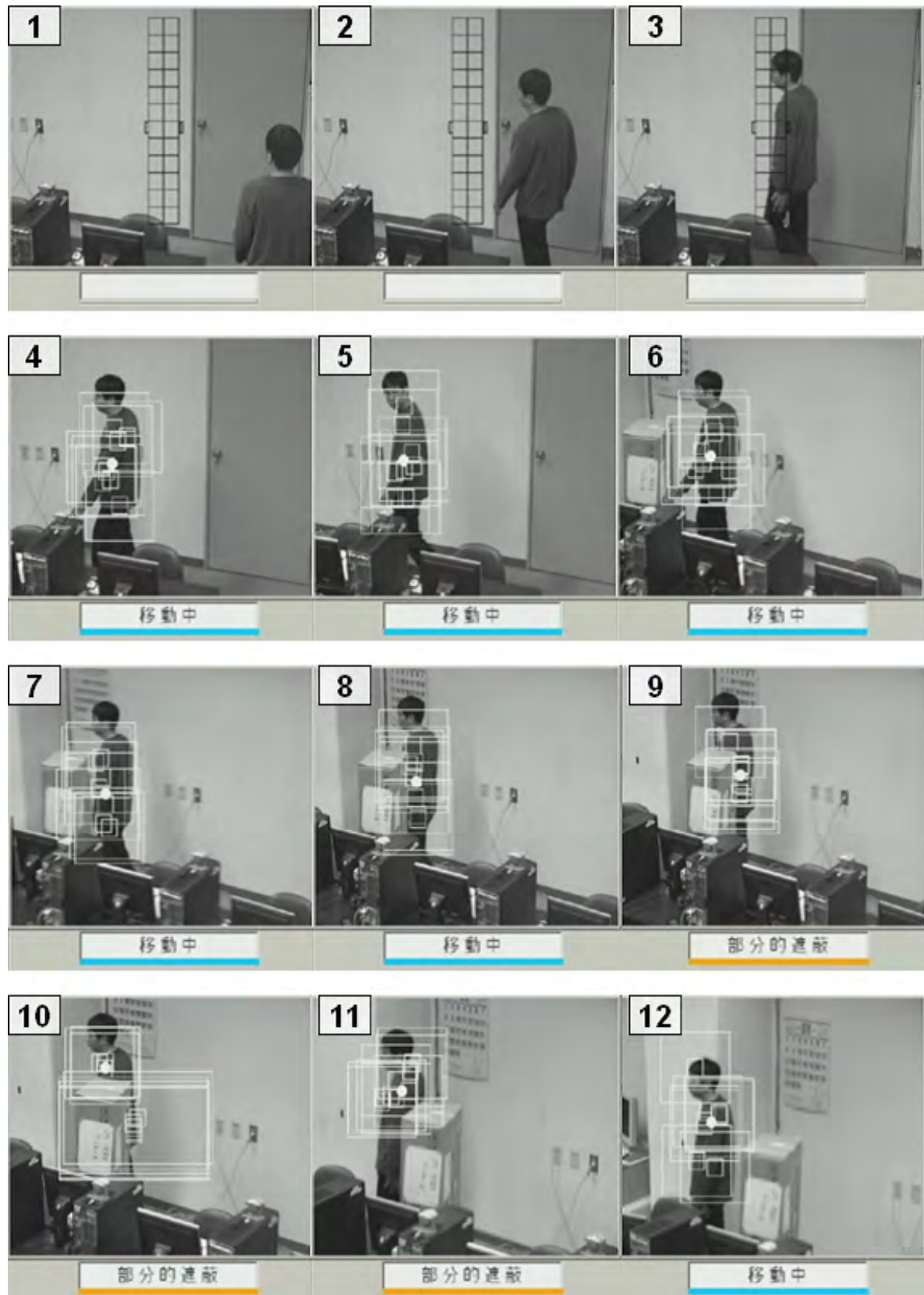


図 2.14 追跡中における処理結果 1 (実験 1)



図 2.15 追跡中における処理結果 2 (実験 1)

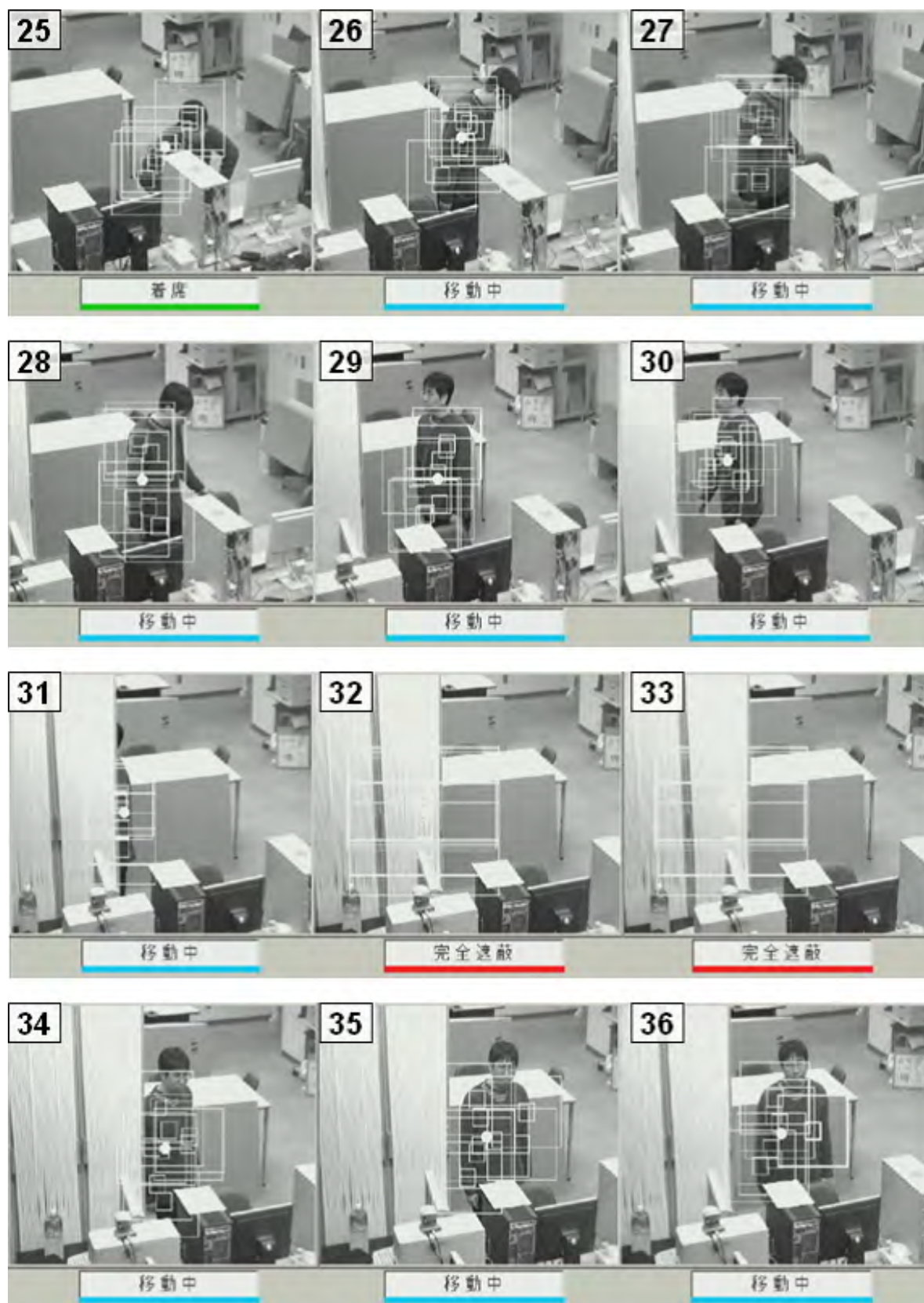


図 2.16 追跡中における処理結果 3 (実験 1)

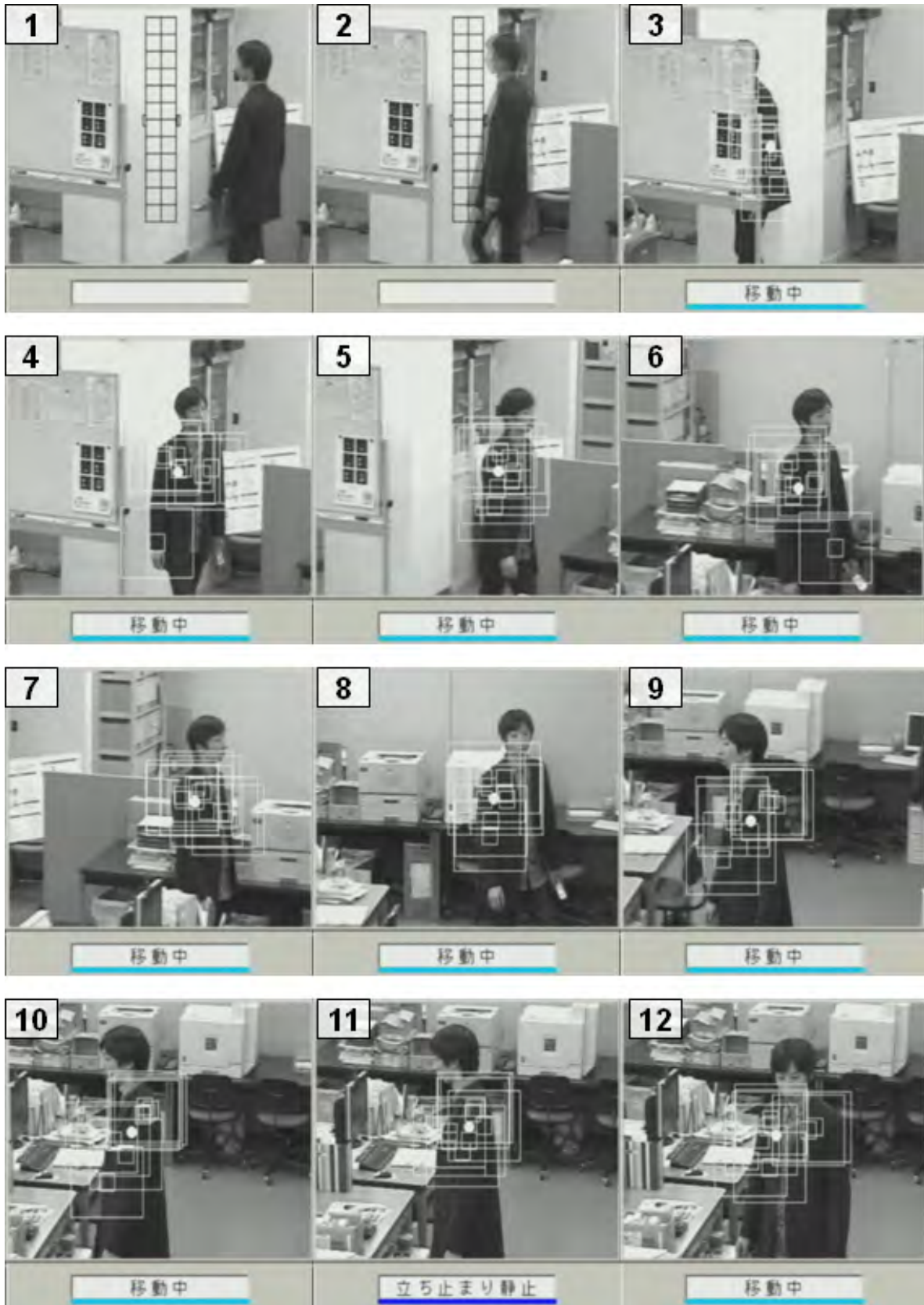


図 2.17 追跡中における処理結果 1 (実験 2)



図 2.18 追跡中における処理結果 2 (実験 2)



図 2.19 追跡中における処理結果 3 (実験 2)

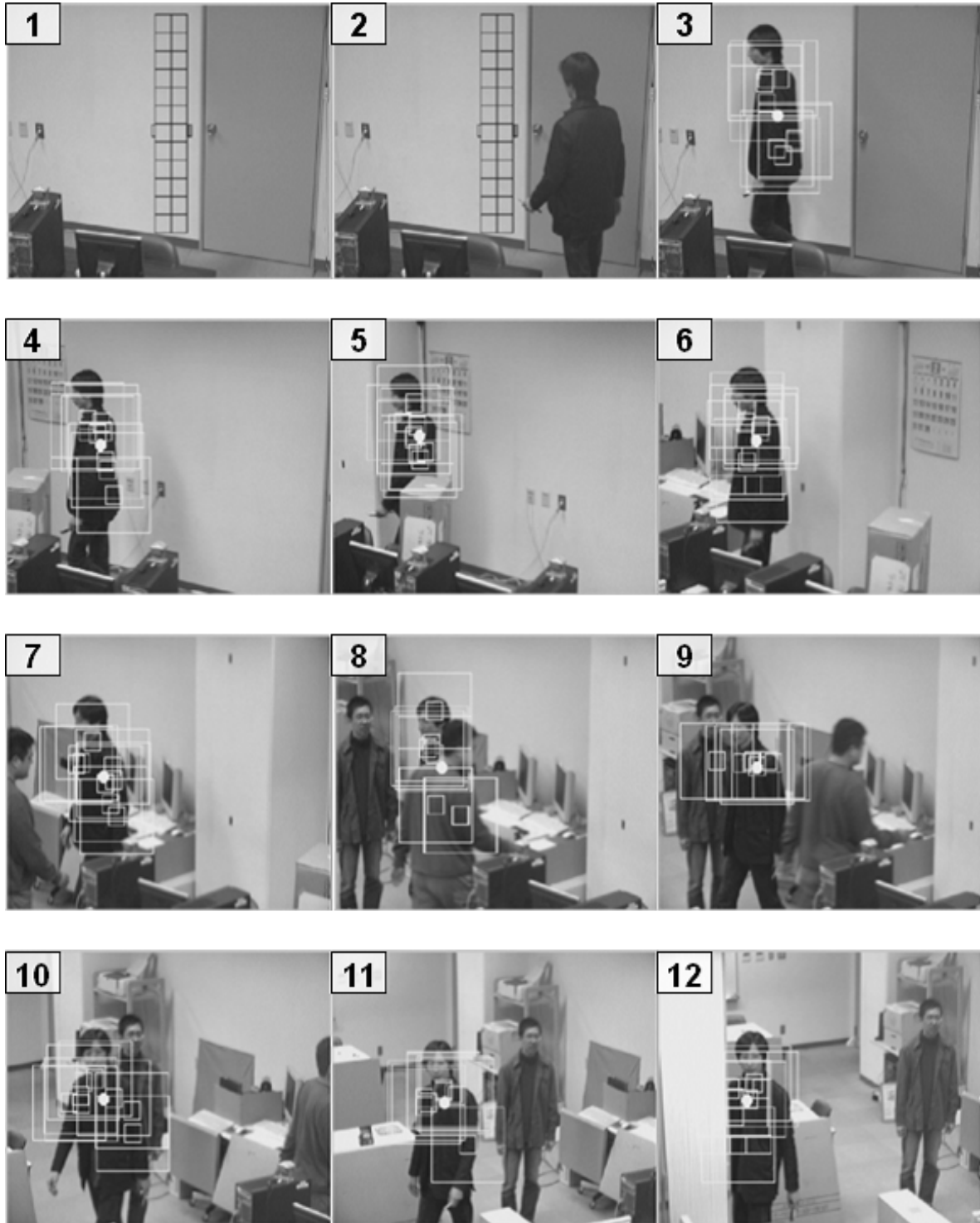


図 2.20 追跡中における処理結果 1 (実験 3)

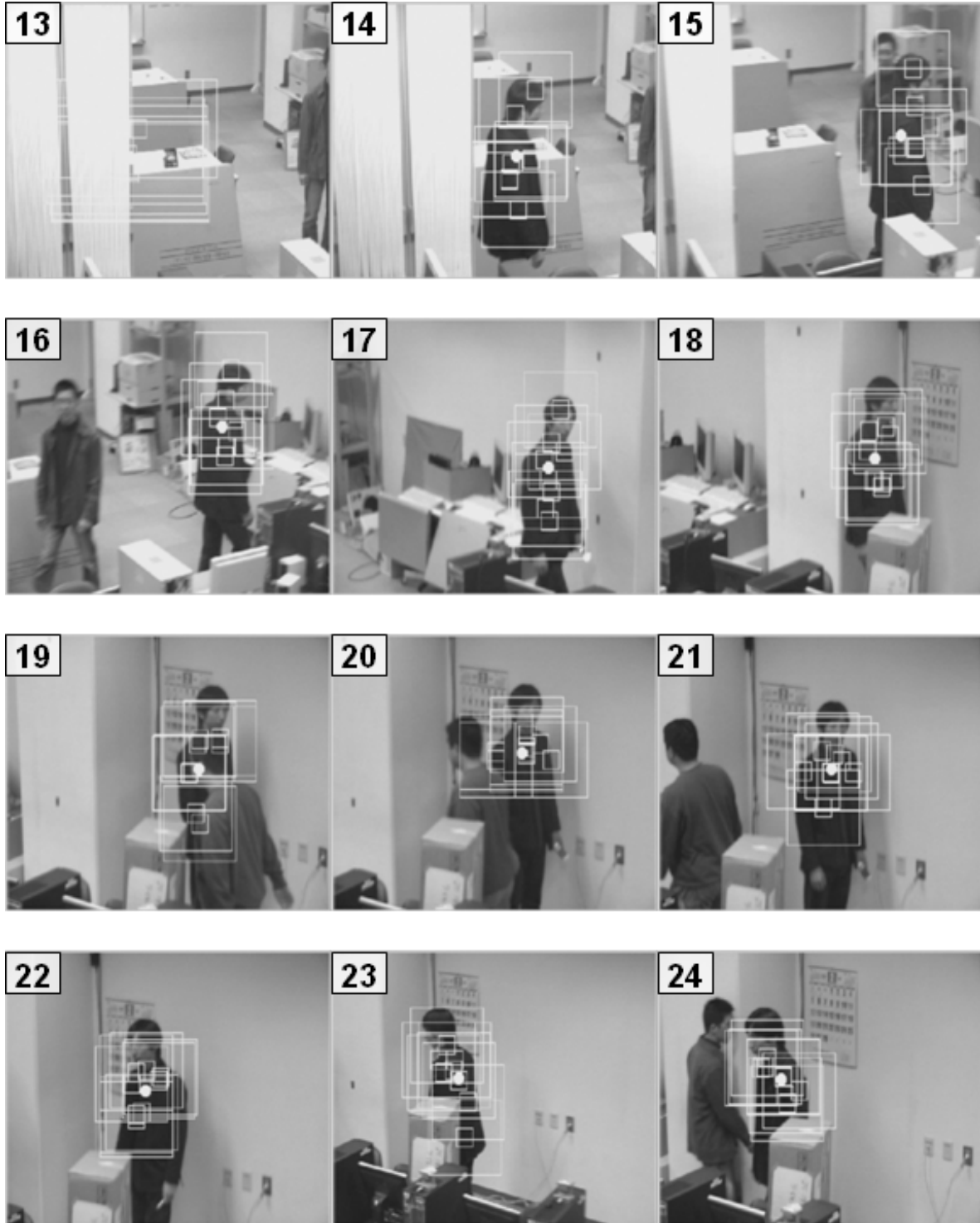


図 2.21 追跡中における処理結果 2 (実験 3)



図 2.22 追跡中における処理結果 3 (実験 3)

第3章

人物同定エージェント

3.1 まえがき

生体情報を利用して人物同定を行う手法は、拘束状態にて行うものが一般的である。例えば指紋や静脈を用いた手法では、人物が情報取得装置に接触する必要がある。また顔や網膜虹彩などの画像パターン認識を用いる手法は、人物の顔をカメラの前に一時的に固定しなければならない。これらの手法は高い精度を得ることが可能となるが、被同定者は物理的かつ心理的に拘束される。

こうした拘束型の手法は、例えば原子力発電施設や組織における機密情報を扱う部屋など、高いセキュリティ性能を必要とする場を主に想定して開発されてきた。しかし人物同定が必要となるのはセキュリティ応用のみならず、例えば研究室や講義における出欠管理、コンビニエンスストアにおけるリピーター検出など様々な応用の場が考えられる。

これらの用途においては人物を入力機器の所に拘束したり、物理的・心理的負担を強いることなく、人物が自由に流動する環境にて自然な状態で得た情報を用いることにより、非拘束状態にて個人同定を行うことが望まれる。一方、同定する機会は出現時に限られることは無く、人物が当該環境に滞留している期間内に同定することができれば、有用となる場合が多い。また、講義における出欠管理では受講者以外の人物が入室する可能性が低いことや、リピーター検出では学習した人物以外を同定する必要はないことなど、同定可能条件を緩和できる点も存在する。

そこで我々は、被同定者に意識させない位置に設置したテレビカメラから得ることが可能な情報のみを利用することで、物理的かつ心理的に拘束することなく、人物を同定可能とする手法について研究を行ってきた [11][12][13][14]。

人物同定の手がかりとなる体型や習慣的な挙動を確率分布の形で学習しておき、ベイズの定理に基づく計算式で各時点における各人物の事後確率を算出する。更に時間方向におけるベイズ統合計算結果に基づいて、当該人物の同定を行う。

本論文では、「非拘束状態における人物同定」という研究目標に取り組む第一段階として、少人数、準固定メンバー、在席・流動状態の強い再現性を仮定した比較的簡単な状況下で有効に働く手法を提案する。具体的には以下の4つの制約条件を設定した。

- 対象人数が少ない
- 定まった対象者以外の人物が混入する可能性が低い
- 対象者が同じ座席に座る可能性が高い
- 移動空間が狭く、対象者が毎回同じ経路をたどる可能性が高い

この具体的な場として、学部講義に比べて出席者数が少なく、出席者の範囲が履修登録者として定まっており、通路の狭い小教室で実施されることが多い大学院講義を想定した実験条件を設定し、有用性評価実験に用いた。

以下、従来研究とその課題、人物同定に用いる特徴、提案手法、評価実験、考察について順に述べる。

3.2 従来研究とその課題

人物が自由に流動する環境（以下、流動環境）では、人物同定を行うのに必要十分な情報量を持つ特徴を取得することが困難である。したがって、不十分な情報量を有する特徴を複数統合することで、全体としての情報量を高める必要がある。複数特徴を統合する枠組は、次の三つに分類できる [15]。

- アンサンブルモデル
- マルチモーダルモデル
- マルチサンプルモデル

アンサンブルモデルは、一つの特徴に対して複数の手法を用いて性能向上を行うモデルである。マルチモーダルモデルは、複数の特徴を統合判定して性能向上を行うモデルである。マルチサンプルモデルは、一つの特徴に対して同じ手法で複数回取得することで、ノイズや学習データの変動等の悪影響を軽減するモデルである。

アンサンブルモデルを用いる場合、例えば顔パターンの様に高い情報量を持つ特徴であれば、固有顔による特徴抽出 [16] やグラフマッチングによる方法 [17] など様々な手法が提案されており、各手法の結果を統合することで精度向上を期待することができる。しかし、流動環境にて取得可能な特徴に関しては、特徴抽出手法自体が制約や限定されるために、このモデルを積極的に用いる理由が存在しない。

マルチモーダルモデルを用いる場合、用いる特徴の選択が重要になると考えられる。このモデルを用いた従来研究として、杉江ら [18] や Brunelli ら [19]、Choudhury ら [20] は

顔画像識別と音声特徴認識の結果を統合する手法を、前田ら [21] は歩行特徴を、Fox ら [22] は唇の動きを加えて認識する手法を提案している。これらの研究では、高いセキュリティ性能を必要とする場を主に想定して開発されてきたために、流動環境で取得することが困難な特徴が多い。例えば、音声情報を取得するためにはマイクの前で発声する必要があるため、被同定者はマイクの前に一時的に拘束される。

マルチサンプルモデルを用いた場合、取得回数を増やすことで精度が高くなることが期待できるが、一方で人物を拘束する時間が長くなる問題がある。よって、一般的には目的に応じて両者のトレードオフを検討することになる。しかし、人物が視野内に滞留しているときに取得可能な特徴であれば、人物を拘束する時間は問題とならない。

これらの検討結果から、流動環境で人物同定をするためには、マルチモーダルモデルとマルチサンプルモデルを組み合わせる方式が有効であると考えられる。しかし、マルチモーダルモデルに基づく従来研究では、流動環境にて人物同定に有用となりうる特徴についての検討が十分に行われてこなかったため、次章では、その特徴について検討していく。

3.3 人物同定に用いる特徴の検討

流動環境にて取得でき、かつ人物同定に有用と考えられる特徴として以下の候補が考えられる。

- 人物の明度パターン
- 人物の体型特徴
- 手足の動き特性
- 習慣性

人物の明度パターンとして代表的な顔画像は、人物同定を行う上で高い情報量を持つ。しかし、カメラと対象の向きによっては人物同定に有用な正面顔画像が取得不可能な場面も考えられる。この問題に対処するために、部屋内に複数のカメラを設置して顔領域を取得する手法が提案されているが [23]、廊下を通過することを前提に正面顔画像が取得できることを暗黙に利用しているため、自由空間内で人物が流動する場合にはそのまま適用するのは困難である。

次に体型特徴について検討する。人物が流動または静止している状態にて取得可能な、人物全体の体型特徴として次のものが考えられる。

- 人物領域の高さ
- 人物領域の幅
- 人物領域の面積

- 顔領域の高さ
- 顔領域の幅

人物の体型特徴として、安定して取得可能な特徴は人物領域と頭領域と考えられる。人物領域は人物全体で最も大きい体型特徴であるために完全遮蔽が起こりにくい。また頭領域は人物全体で最も上に位置しているが、一般的な環境では床に設置された遮蔽物が多い傾向にあることから、体の他の部分に比べると隠れる可能性が低い。画像中の個々の特徴値は、カメラから人物までの距離によって変化するため、二つの体型特徴の比をとり正規化する必要がある。このとき、人物領域の幅はこれらの体型特徴の中で最もオクルージョンが発生する可能性が低いことが想定されるため、これとの比率値にするのが最適であると考える。

人物の短期的な行動特性としては、手足の振り周期が考えられる。特に足の行動周期は歩容認証 [24][25] や年齢層推定 [26] として研究が進められている。しかし、流動環境内には様々な障害物があるため、振り周期が乱れることが想定される。したがって、手足の振り周期などを行動特性として利用することは適切でない。

一方、頻繁に訪れる場所における人物の挙動には習慣性があると考えられており、行動の異常検知 [27][28] の研究例がある。これらの研究では、全方位カメラで撮影した画像から人物領域を認識し、行動モデルを自動作成することで非日常状態を検出する。しかし、それらの特徴が人物同定に有用な情報源であるかについては十分に検証されていない。

我々は、これまでに人物同定に有用となり得る習慣性特徴として、行動習慣性 [13] や在席習慣性 [14] を検討してきた。行動習慣性とは空間内の人物の流動軌跡を特徴として用いるために、特定の観測点と人物の重心との距離を頻度分布とするものである。また、在席習慣性とは空間内の着席可能位置に観測領域を設定し、背景差分による着席検出によって着席位置の頻度分布を特徴とするものである。しかし、これらの習慣性特徴は観測点を環境毎に実験的に設定せねばならず、在席習慣性に関しては着席する席のない空間では適用できない問題があった。そこで、新たに滞留習慣性を考案した。

滞留習慣性とは、空間をある一定の面積で分割し、その各領域における頻度分布を習慣性特徴とするものである。図 3.1 は、B 地点から通路を通り A 席に座る人物の滞留習慣性の例を示しており、濃度は頻度の高さを示す。移動速度の遅い人物であれば通路領域の濃度は高くなり、複数の席を好む人物であれば複数の領域に濃度が表われることになる。この特徴化によって人物の挙動を定量的に表現することが可能となる。

これらの検討結果より、表 3.1 に示す各特徴を、流動環境下で人物同定に有効な特徴として選定した。体型特徴値はカメラと姿勢の関係で変動するが、今回設けた制約条件を満たす環境においては、通路内を直進通行することが主であるため、人物の取り得る姿勢が複数個に制約される。これに応じて体型特徴値も複数個に離散化されるため、移動環境内

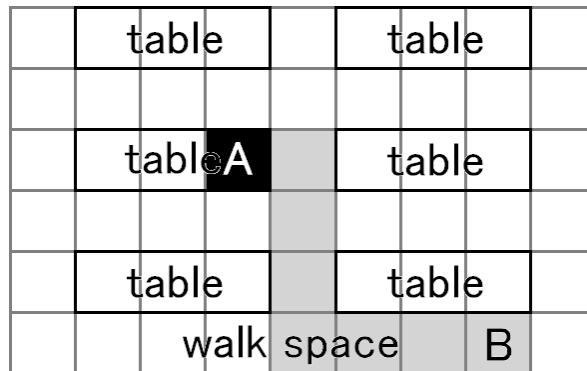


図 3.1 滞留習慣性の説明図

表 3.1 流動環境下で人物同定に利用可能な人物特徴

ID	Kind of feature
F1	Ratio of body height to body width
F2	Ratio of head width to body width
F3	Ratio of head height to body width
F4	Ratio of body area to squared body width
F5	Distribution of habitual behavior histogram

で十分な学習を行なうことにより，各人特有の体型値変化を確率分布として学習し，認識に用いる．

3.4 処理の流れ

3.4.1 全体の概要

本提案手法の流れを図 3.2 に示す．まず背景差分を基に人物領域の検出を行い，その領域から頭領域の検出を行う．次に表 3.1 に示した特徴群を抽出する．学習時にはその特徴群から頻度分布を作成し，これに基づき特徴毎の事前確率分布や条件付き確率分布を求める．同定時には，各時点毎の各人物の事後確率を計算する．更に時間方向におけるベイズ統合計算結果に基づいて，当該人物を同定する．

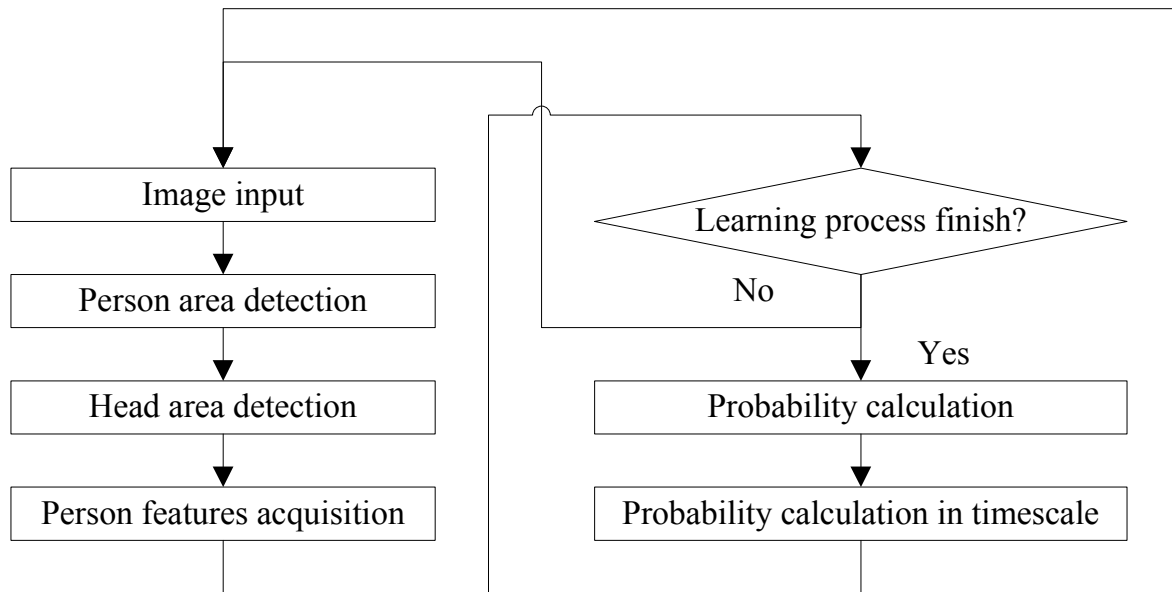


図 3.2 提案手法の処理の流れ

3.4.2 人物領域取得

人物領域を取得するために背景差分法を用いる．この手法は背景画像を複数枚取得し，式のパラメータに基づいて背景学習を行うことから，例えば照明のちらつきや空調によるブラインドの変動など，周期性を持った環境変動に対してロバストである．

背景の更新式を式 (3.1) に示す．このとき， B_t は時刻 t の背景画像， I_t は時刻 t の画像， d は背景更新を行うか否かを判別する式の値， λ は d の閾値， n は学習に用いる画像数である．初期入力画像 I_0 を背景画像 B_0 として取得する．逐次取得した画像から，各画像点にて d を求め， $d \leq \lambda$ となる点は背景を更新し， $d > \lambda$ となる点は人物候補点とする．

$$B_t = \begin{cases} B_{t-1} & (d > \lambda) \\ I_t + \frac{B_{t-1} * (n-1)}{n} & (d \leq \lambda) \end{cases} \quad (3.1)$$

$$d = \left| \frac{I_t - B_{t-1}}{n} \right|, B_0 = n * I_0, B_1 = n * I_1$$

次に背景差分結果から人物領域を検出する．図 3.3 は人物領域抽出の説明図である．人物候補点を縦軸方向へ投射し，そのヒストグラムから最大値を検索する．そこから左右に探索を行い，ヒストグラム値が 0 となる位置を人物領域の幅として決定する．検出済みのヒストグラムは削除し，この処理を繰り返すことで複数人の検出を行う (A と B)．人物領域の高さは，検出済みの人物領域の幅領域のみを横軸方向へ投射し (A' と B')，同様の

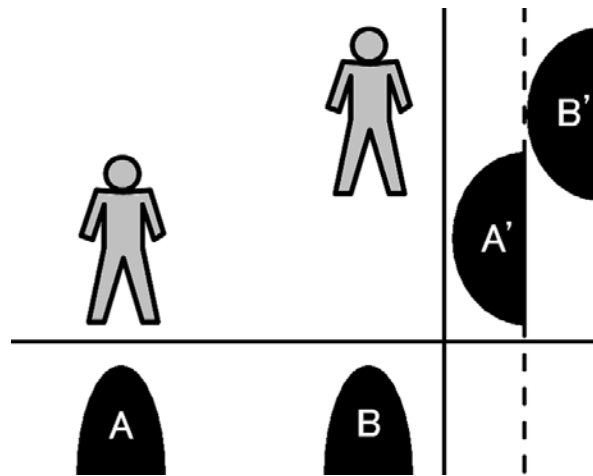


図 3.3 人物領域抽出の説明図

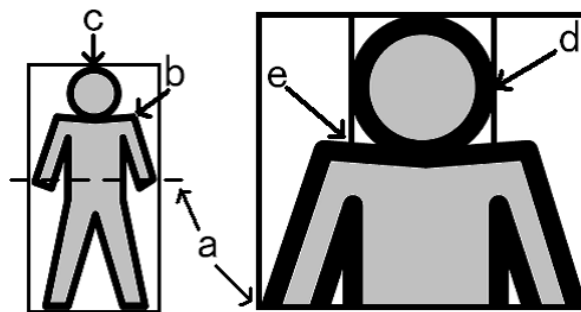


図 3.4 頭領域抽出の説明図

処理を行うことで検出する．人物領域の面積は，人物候補点の総和とする．

なお複数の人物領域が検出された場合は，前フレームの人物領域と対応付けを行う必要があるため，画像間の中心距離が最も近い領域を同一人物領域として対応づける．

3.4.3 頭領域取得

先に取得した人物領域から頭領域を検出する．図 3.4 は頭領域抽出の説明図である．まず人物領域を上下に分割する線上を探索開始線とする (a)．その地点の左端から逐次上方に向かって明度値が低い値へ変化する点を体上部の境界点とする (b)．得られた上部境界点の集合において，画像座標位置の最も上に位置する点を頭頂部とする (c)．この位置から左右方向へ上部境界点集合の微分値を計算し，最大となる位置を頭の左右位置とする (d)．また，各最大値座標よりも外側の点の縦軸位置を頭下部位置とする (e)．

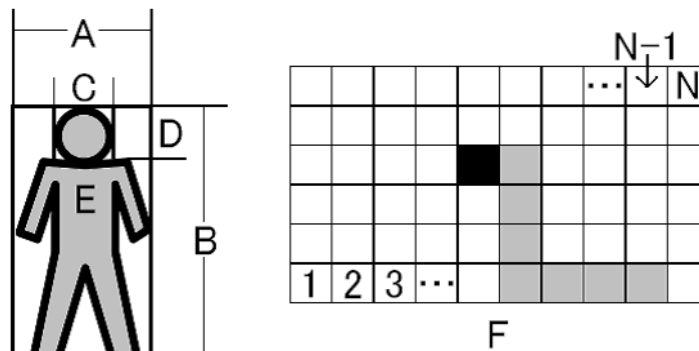


図 3.5 人物特徴の定義

3.4.4 人物特徴取得

人物領域と頭領域の検出結果から，表 3.1 に示した体型特徴の取得を行う．特徴取得式を式 (3.2) に示す．ここで， $F1$ から $F5$ は表 3.1 に示す特徴の特徴値， $k1$ から $k4$ は後述する正規化係数， A から F は図 3.5 で示す特徴値を表す．

$$\begin{aligned} F1 &= k1 * \frac{B}{A}, & F2 &= k2 * \frac{C}{A} \\ F3 &= k3 * \frac{D}{A}, & F4 &= k4 * \frac{E}{A^2}, & F5 &= F \end{aligned} \quad (3.2)$$

学習データは取得した特徴値を頻度分布として配列に蓄積加算していく．体型特徴は実数であるが，配列の添え字は整数で表現されるため，各特徴の最小値が配列要素の 0 番目に，最大値が配列要素の最終要素となるように，整数へ正規化する．この正規化に用いる値を正規化係数とし，式 (3.2) の $k1$ から $k4$ がこれに該当する．

3.4.5 ベイズの定理に基づく事後確率値算出

概要

人物同定時に取得した特徴値と前節で取得した離散化された各特徴の頻度分布から，各人物の事後確率を算出する．更にマルチサンプルモデルに基づき，時間方向における各人物の事後確率を算出する．

事前確率

事前確率は学習データとして蓄積保存した特徴の頻度分布から作成する．事前確率の計算式を式 (3.3) に示す．ここで， P_{pri} は事前確率， p は人物の識別子， P は学習人物数， f は特徴の識別子， F は人物同定に用いる特徴数， n は f が示す特徴における特徴値，

N は n の最大値を表す .

$$P_{pri}(f, n) = \frac{\sum_{p=1}^P x(p, f, n)}{\sum_{p=1}^P \sum_{n=1}^N x(p, f, n)} \quad (3.3)$$

条件付確率

条件付確率の計算式を式 (3.4) に示す . ここで P_{con} は条件付確率を表す .

$$P_{con}(f, n, p) = \frac{x(p, f, n)}{\sum_{p=1}^P \sum_{n=1}^N x(p, f, n)} \quad (3.4)$$

事後確率

ベイズの定理に基づく事後確率の計算式を式 (3.5) に示す . ここで , P_{post} は事後確率 , ν は人物同定時において f が示す特徴における特徴値を表す .

$$P_{post}(p) = \frac{\sum_{f=1}^F (P_{pri}(f) \cdot P_{con}(f, p))}{F} \quad (3.5)$$

$$P_{pri}(f) = \begin{cases} P_{pri}(f, n) & n = \nu \\ 0 & otherwise \end{cases}$$

$$P_{con}(f, p) = \begin{cases} P_{con}(f, n, p) & n = \nu \\ 0 & otherwise \end{cases}$$

時間方向への統合

各時点における事後確率値は , 特徴抽出の失敗などにより必ずしも常に人物同定に有効な値となっているとは限らない . よって信頼性を向上させるために , マルチサンプルモデルの概念を用いて , 時間方向に統合を行う .

時間方向の事後確率式を式 (3.6) に示す . ここで , \bar{P}_{post} は時間方向の事後確率 , \bar{P}_{pri} は時間方向の事前確率 , \bar{P}_{con} は時間方向の条件付確率 , $P_{post}(p, t)$ は t 時点における $P_{post}(p)$, t は t 番目のフレーム , T はフレームの最大値を表す . 時間方向の事前確率 \bar{P}_{pri} は , 各時点における事後確率値の順位に応じた確率となるように与える .

$$\bar{P}_{post}(p) = \frac{\sum_{t=1}^T (\bar{P}_{pri}(t) \cdot \bar{P}_{con}(p, t))}{T} \quad (3.6)$$

$$\bar{P}_{pri}(t) = \begin{cases} \frac{P}{\frac{P(P-1)}{2}} = \frac{2}{P-1} & P_{post} : 1st \\ \frac{P-1}{\frac{P(P-1)}{2}} = \frac{2}{P} & P_{post} : 2nd \\ \vdots \\ \frac{1}{\frac{P(P-1)}{2}} = \frac{2}{P(P-1)} & P_{post} : last \end{cases}$$

$$\bar{P}_{con}(p, t) = P_{post}(p, t)$$

3.4.6 人物同定の結果確定

次のいずれかの条件を満たしたとき，人物同定を完了する．条件文に示した「一定の差」は，目的とするアプリケーションによって設定する．

- 被同定者が視野内から消失，もしくは座席にて静止したとき
- 最も高い人物の \bar{P}_{post} と次点の人物の \bar{P}_{post} に一定の差が生じたとき

3.5 評価実験・考察

3.5.1 各評価実験の概要

3.5.2 節は，複数特徴統合によって人物同定が可能であるかを評価するための実験である [11]．3.5.3 節は，複数特徴統合後の事後確率の扱いについて，MAP 推定法と順位点蓄積手法との比較を行う実験である [12]．3.5.4 節は，習慣性特徴として行動習慣性特徴を用いた場合の評価実験 [13]，3.5.5 節は，習慣性特徴として在席習慣性特徴を用いた場合の評価実験 [14]，3.5.6 節は，習慣性特徴として滞留習慣性特徴を用いた場合の評価実験である．最後の 3.5.7 節は，これまでに行った実験を再評価した総合実験である．

3.5.2 複数特徴統合に基づく人物同定手法の評価実験

実験環境

CPU:2.6GHz の計算機にて，640 × 480 の解像度能力を持つ Matrox 社製 Meteor キャプチャーカードを使用した．開発環境は，処理速度を高めるために C++ 言語を基本にした Microsoft 社製 Visual C++ 6.0 を用いた．実験ではカメラの前に人物がいない状態，もしくは人物が一人だけいる状態のどちらかであるとしている．つまり複数人が同



図 3.6 実験室（左）・研究室（右）

表 3.2 人物特徴の有効範囲

	上限	下限
人物領域の縦横比	1.0	3.0
頭幅と人物領域の高さとの比率	1.0	9.0
頭幅と人物領域の幅との比率	1.0	3.0
頭幅と人物領域の面積	0.0	20.0
頭領域の明度ヒストグラム	0.0	4.0
人物領域のエッジ量	0.0	30.0

時に存在しない状況とした。これは分散協調型認証の有効性実験を簡便化するために設定した条件であり、個々の人物の検出対応付け機能を組み込むことで複数人対応とすることも可能であると考える。

単純背景環境として実験室（図 3.6 左）、複雑背景環境として研究室（図 3.6 右）に行った。被験者は部屋に侵入した後、歩いて数回周った後に、侵入した場所より退出する。まず人物特徴の取得範囲を確認するための予備実験を行った。次に認証システムに研究室の人間 10 人（図 3.7）を学習させ、その後人物を再び進入させ人物認証を行った。

人物特徴の有効範囲取得人物特徴の分布を取得する予備実験を行った。全学習者の特徴分布から下限値・上限値を取得し、それを特徴の有効範囲とする。認証時に、人物の隠れなどに伴い取得データが異常であることを判断する材料となる。

図 3.8 と図 3.9 は 6 つの人物特徴のうち代表として 2 つの特徴分布を示す。横軸は時間、縦軸は特徴量を示している。表 3.2 に取得した各人物特徴の有効範囲を示す。



図 3.7 学習対象人物

人物の学習時間と頭位置設定成功率 表 3.3 は人物の学習時において、学習した時間と頭位置設定の成功率を示している。Time は学習した時間、Frame は学習した画像枚数、Rate は頭領域の検出に成功した確率となる。なお Time はプログラムの仕様上、直接求めることが出来ないため、Frame とプログラムの平均処理速度から計算して求めている。

人物認証精度評価 図 3.10 は人物認証の時系列認証結果の成功例を示している。横軸は時間、縦軸は各画像で認証した結果の最高確率になった人物の回数の頻度を示している。左上に示す人物が入室したときの人物認証結果である。同人物を認証した結果を示す線が最も上に位置することが分かる。2 番目の候補を示す線との距離が十分離れていることより、人物認証の信頼性が高いといえる。表 3.4 は人物認証の結果を示す。1st は最終的に最上位に位置する人物を示す。例えば、Person1-1st の項目は、1 番目の人物が進入したときの結果であるので「1」であれば認証に成功したことになる。2nd と 3rd は 2 番目と 3 番目の認証候補を示している。最終行の Rate は人物認証の成功割合を示している。10 名の人物に対し実験室においては 80%、より複雑な環境である研究室では 50% の認識率が得られた。

表 3.3 学習時間と頭位置検出率

	Time(s)	Frame	Rate(%)
Person1	25.6	384	93.7
Person2	26.3	394	85.7
Person3	29.1	437	94.8
Person4	24.7	370	95.4
Person5	27.9	418	84.4
Person6	24.5	368	95.1
Person7	22.1	331	96.2
Person8	23.2	348	93.8
Person9	24.5	368	91.1
Person10	25.1	377	98.2

表 3.4 人物認証結果

	実験室			研究室		
	1st	2nd	3rd	1st	2nd	3rd
Person1	1	2	10	9	10	2
Person2	2	3	10	2	10	9
Person3	3	5	2	10	3	2
Person4	5	8	3	4	10	2
Person5	5	3	10	10	9	5
Person6	3	10	8	7	4	9
Person7	7	10	9	7	2	5
Person8	8	3	10	9	10	2
Person9	9	10	7	9	2	10
Person10	10	5	3	10	2	9
Rate(%)	80	80	80	50	60	70

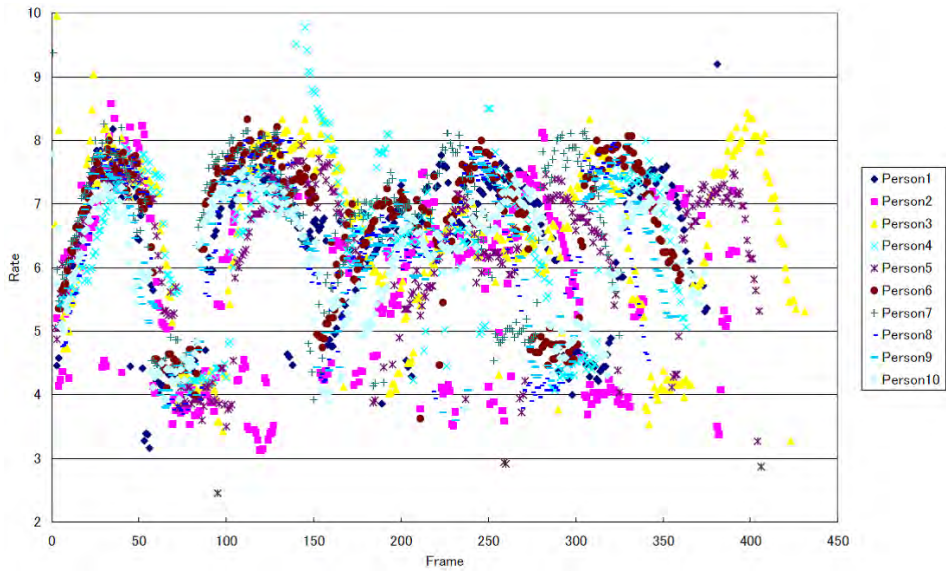


図 3.8 頭幅と人物領域の高さとの比率

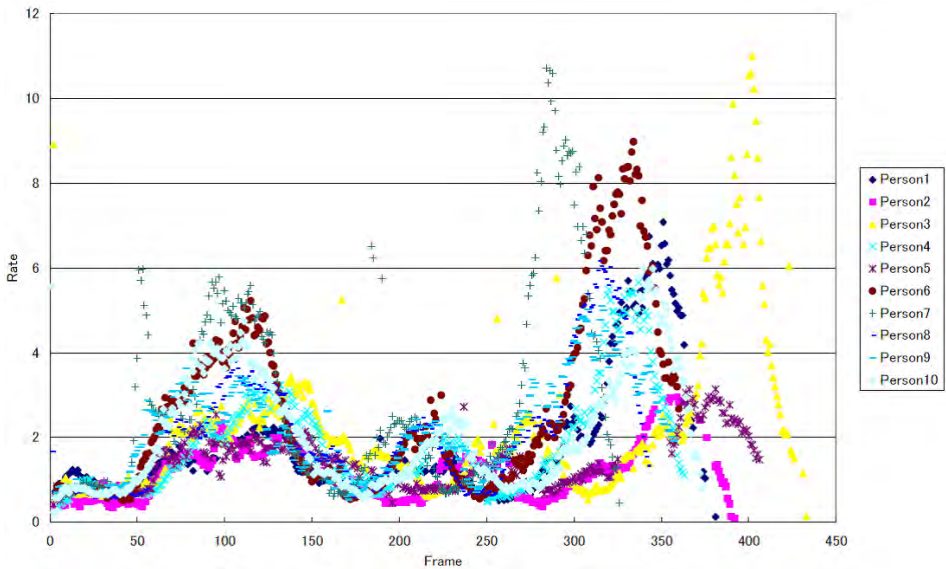


図 3.9 頭領域の明度

各人物特徴の有効性分析 図 3.10 に示す人物認証が成功した事例において，各人物特徴の有効性について分析を行った．図 3.11 では人物特徴を各々無効にしたときの 1 位と 2 位の距離変化を示している．中央線より下を推移している人物特徴は，それを無効にしたことにより 1 位と 2 位が逆転している，つまり結果が悪くなっていることから，その人物特徴は必要であると考えられる．図 3.11 では頭幅と人物領域の面積が最も強い影響を与えていることが判る．

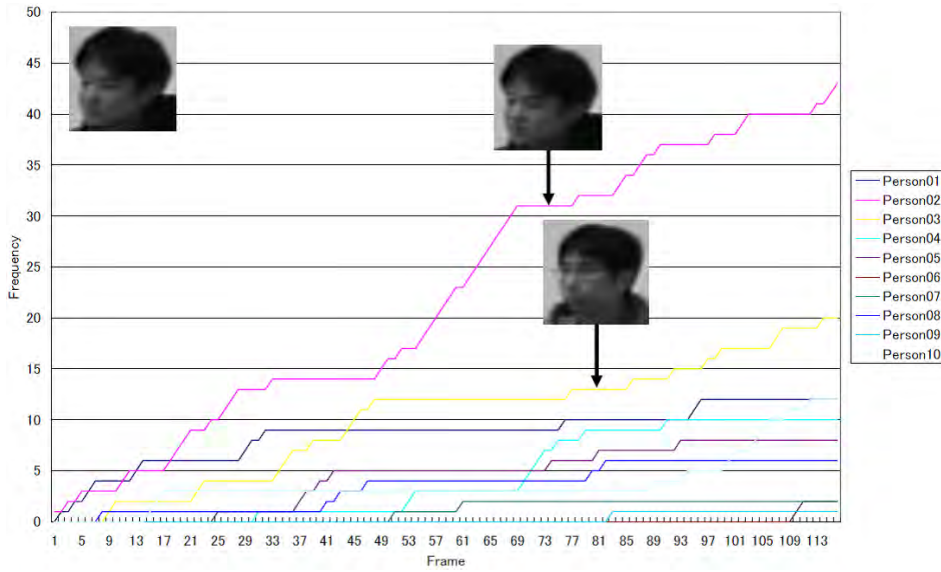


図 3.10 人物認証成功例

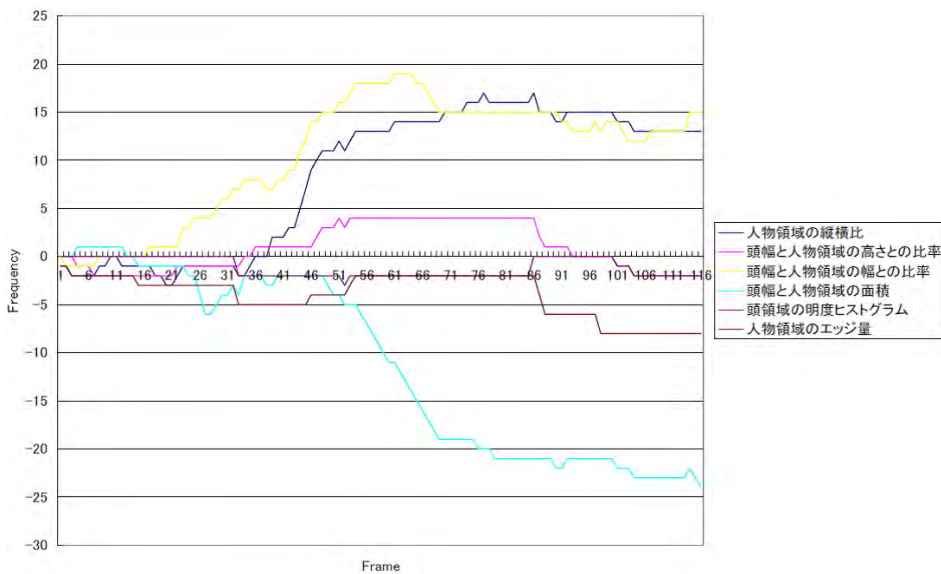


図 3.11 各人物特徴を削除したときの認証精度推移

考察

全ての実験において処理速度は概ね 15fps(frame persecond) となった．この値はキャプチャボードの性能限界値であり，ハードウェアの性能向上によって更なる処理速度の向上が期待出来るが，現時点においてもリアルタイム処理にて人物認証が実現出来た．

人物特徴の有効範囲を測定する実験結果から，人物特徴の識別能力が伺える．図 3.8 の

「頭幅と人物領域の高さとの比率」においても人物による差異は認められるが、図 3.9 の「頭領域の明度」ではより顕著に人物毎の差が表れている。記載できなかった他の 4 特徴も図 3.8 の分布に類する傾向を示した。より人物認証精度を高めるためには「頭領域の明度」のような傾向を示す特徴が有用であると考えられる。

表 3.3 に示した結果より頭の検出率は概ね 90% 以上であることが分かる。頭検出が失敗した要因は、頭部が複雑な背景位置にあり、上部輪郭の検出が不安定になったことが考えられる。また、前屈みになるなどして、頭両端の微分成分が明確に検出できなかったこと等も考えられる。しかし、検出に失敗した場合でも、他の画像にて頭検出が行えれば大きな問題とはならない。短時間の学習では、検出失敗によって人物特徴が十分に取得できない可能性があるが、一定時間継続的に学習すれば、頭位置の検出に失敗したこと影響を低減することが可能である。

認証精度の実験結果は必ずしも高くないが、まだ特徴の数が少数であり、非拘束状態のために認証に決定的な情報を取得できないことを考慮すると、本方式の基本的な有効性は示せたと考えられる。図 3.11 に示した結果より、頭幅と人物領域の面積が最も有効であったが、これは体の向きなどを考慮していない概略的な情報であるので、更に改良を進めることによって認識率が向上することが期待できる。

実験室環境に比べ、研究室環境では認証精度の低下が見られた。その要因は人物の移動可能範囲の広さと考えられる。実験室環境では、カメラから人物までの距離が 3~6m 程度であったが、研究室では 1m~6m と変化が大きく、同一人物であっても、近くで取得した特徴量と遠くで取得した特徴量に変化が生じたものと考えられる。これは人物領域の面積等によって、人物までの距離が推測可能であるので、距離による変化を防ぐ補正処理が必要であると考えられる。

3.5.3 MAP 推定と順位点蓄積手法の比較評価実験

実験条件

CPU:Pentium4-3.0GHz の計算機にて、640 × 480 の解像度能力を持つ Matrox 社製 Meteor キャプチャーカードを使用した。開発環境は、処理速度を高めるために C++ 言語を基本にした Microsoft 社製 Visual C++.net を用いた。実験は室内環境にて図 3.12 に示す 14 人の被験者を対象とした。被験者は全て研究室に属する人間で作為的に選んだものではない。被験者は図 3.13 に示したように、右奥から進入し中央の通路を通り左手前から退出する。この動作を 3 回行ってもらい、その様子を DV カメラに録画する。録画済み動画像から認証システムにて 3 つの学習データを作成し、認証精度の測定を行った。

実験結果



図 3.12 被験者 (左上から右へ Person01 ~ 14)

頭幅検出成功率 人物の頭幅取得の成功率を表 3.5 に示す。行は対象人物ごとの成功率，列は動画像別の成功率を示している。人物領域の両端に達するまでに頭部を検出できた場合を成功としている。なお頭幅の取得手法については 3.4.3 章にて説明している。この実験結果より，頭幅の取得は平均 90.1% の成功率で取得できていることが分かる。

人物特徴の有効性 人物特徴の有効性を評価するために，各人物特徴を利用しないときの人物認証率変化を実験した。実験結果を表 3.6 に示す。行は該当する人物特徴を利用しない場合の人物認証率，列は動画像別の成功率である。全ての実験において，全特徴を利用したときよりも，各特徴を用いないときの方が低下していることが分かる。

人物認証の成功例 人物認証の成功例を図 3.14 と図 3.16 に示す。また，このときの時系列認証における人物認証結果の推移グラフを図 3.15 と図 3.17 に示す。図 3.14 と図 3.16 では，右上に示す各画像の認証結果は進入した人物 (Person03) と違う人物の顔 (Person10) が表示されているため認証結果は失敗である。しかし，右下に示す時系列認証の結果では進入人物の顔が表示されており認証に成功している。図 3.15 と図 3.17 では，Person01 が進入したときの結果であるが，各画像の認証と時系列認証の両方とも成功している。なお，認証システムは常に 30FPS の速度で動作した。

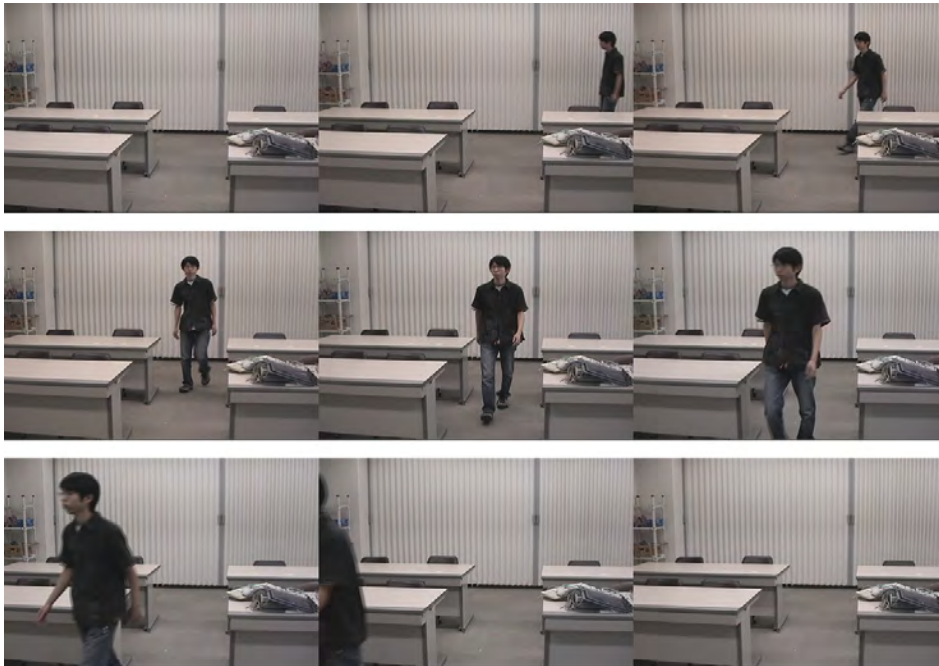


図 3.13 被認証者の動き

時系列認証手法の比較 時系列認証手法別の成功率を表 3.7 にしめす。行は学習に用いた動画像番号 (S1~S3) と認証手法の違いを示し、列は認証に用いた動画像番号を示す。「多数決」は各画像の認証結果の 1 位を、「時系列 3」は 3 位までを、「時系列 A」は全ての順位を時系列フィルタリングに適用した場合の結果である。「時系列 3」の認証結果は全て「多数決」よりも良いことが分かる。「時系列 A」に関しては、「多数決」より概ね結果が良いが、悪くなる結果もあった。

人物認証結果の詳細 時系列認証手法の比較実験で最も成績が良かった「時系列 3」における認証結果の詳細を表 3.8~3.10 に示す。各表の違いは学習に用いた動画像が異なる点である。行は行題に示す人物 (P01 は図 3.12 の Person01 を示す) が進入したことを、列は対象とした動画像番号と認証結果の順位を、表中の値は認証された人物の番号を示す。行題に示す人物番号と認識結果の人物番号が一致すれば人物認証に成功したことになる。最下行の Rate はその順位における人物認証成功率を示している。

考察

頭領域の検出 表 3.5 に示すように、頭領域取得の成功率は平均 90% 前後であり、高い確率で成功している。しかし、Person12 のみ著しく結果が悪くなっている。ここで図 3.18 に Person12 の頭領域検出の処理失敗例を示す。Person12 は今回対象とした人物の中で唯一ストレートの髪型であり、動画像中では肩まで髪が伸びることがあった。それに



図 3.14 人物認証結果 1

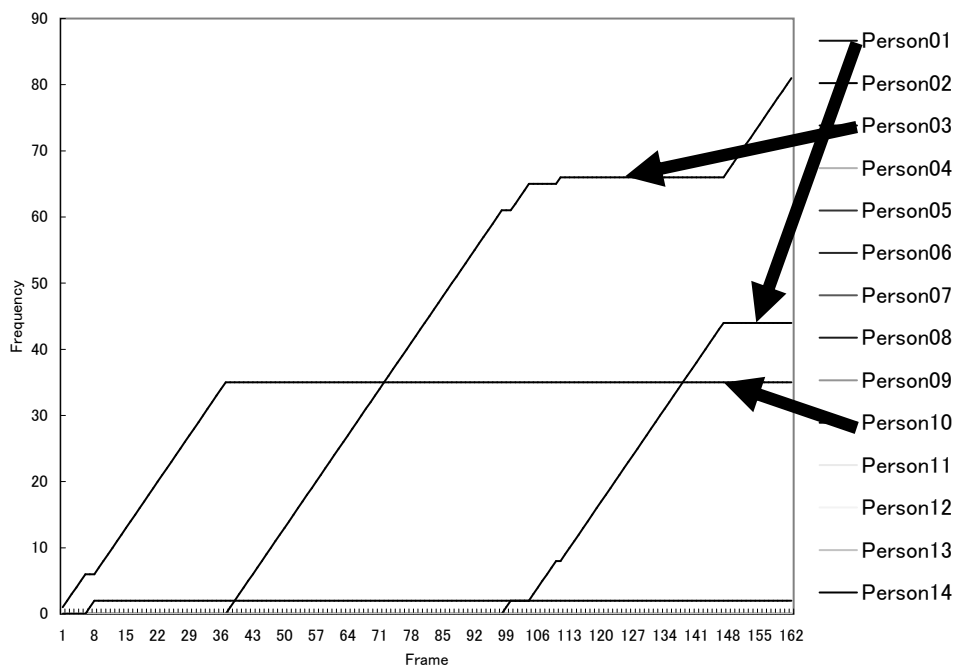


図 3.15 人物認証結果グラフ 1

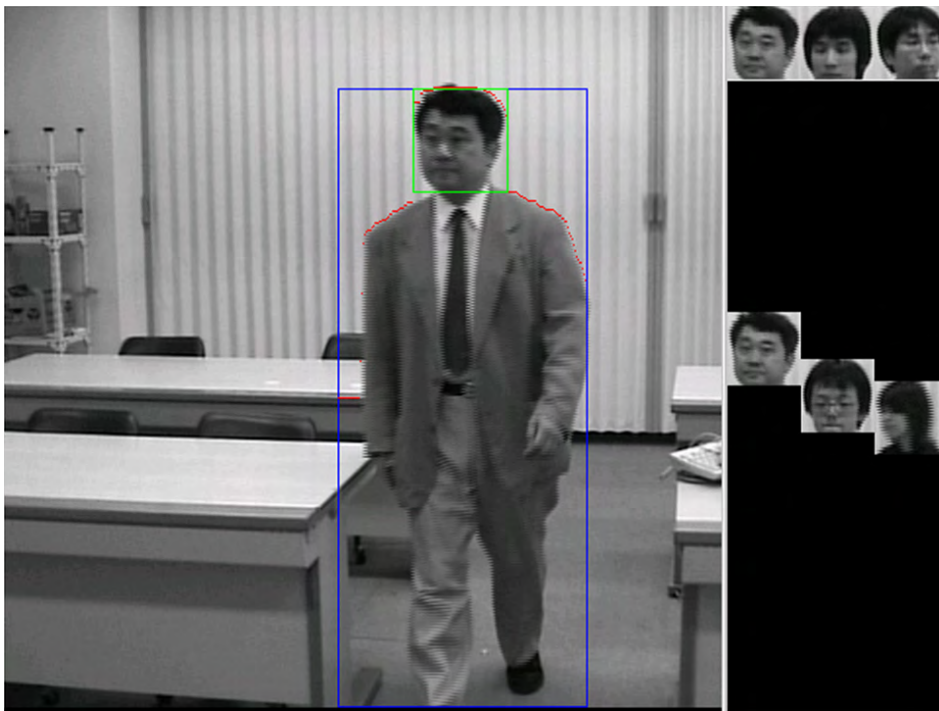


図 3.16 人物認証結果 2

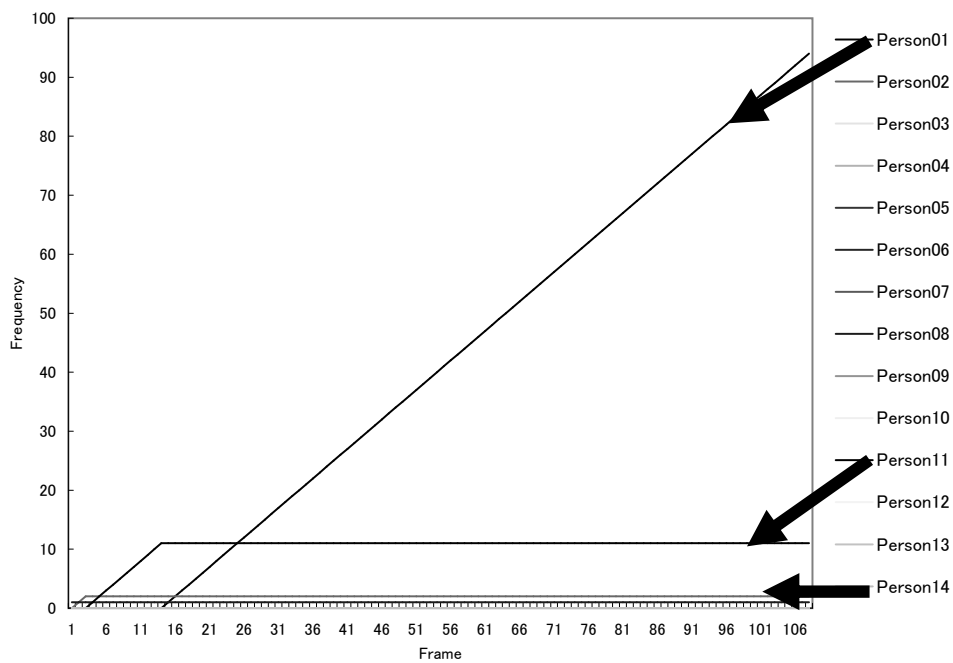


図 3.17 人物認証結果グラフ 2

表 3.5 頭幅取得成功率

	S1	S2	S3	Ave
Person01	97.8	97.1	94.8	96.6
Person02	97.1	97.1	93.9	96.0
Person03	95.8	95.8	94.9	95.5
Person04	85.3	95.2	80.3	86.9
Person05	82.0	89.2	81.0	84.1
Person06	88.6	87.2	94.9	90.3
Person07	97.5	96.4	98.2	97.4
Person08	97.6	95.1	96.1	96.3
Person09	95.3	93.0	94.5	94.3
Person10	97.7	98.2	96.8	97.6
Person11	97.4	98.5	95.1	97.0
Person12	46.9	56.4	39.5	47.6
Person13	93.7	97.1	97.1	95.9
Person14	92.3	88.7	76.8	85.9
Average	90.4	91.8	88.1	90.1

	S1	S2	S3	Ave
全特徴利用	85.7	64.3	71.4	73.8
人物領域の幅	71.4	42.9	57.1	57.1
人物領域の高さ	64.3	50.0	50.0	54.8
人物領域の面積	64.3	64.3	64.3	64.3
頭領域の高さ	85.7	64.3	57.1	69.0

表 3.6 人物特徴の有効性評価

伴い，体上部輪郭の微分成分を抽出する場合に顕著な差が出なかったことが失敗の要因である．これに対処する手法としては，微分成分の極大値のみで判断するのではなく，輪郭曲線の形状から髪型を推定するなどの手法が必要ではないかと考える．

人物特徴の有効性 表 3.6 に人物特徴の有効性を示した．各特徴が有効であれば，全ての特徴を利用したときの人物認証成功率が最も高くなるはずである．実際に，実験では全ての動画像において，利用する特徴を減らすことによる人物認証率の低下を確認すること

表 3.7 時系列認証手法別認識成功率

学習	認証	S1	S2	S3	Ave.
S1	多数決	92.9	71.4	85.7	83.3
	時系列 3	100.0	85.7	78.6	88.1
	時系列 A	92.9	92.9	85.7	90.5
S2	多数決	64.3	100.0	71.4	78.6
	時系列 3	85.7	100.0	85.7	90.5
	時系列 A	78.6	100.0	85.7	88.1
S3	多数決	71.4	78.6	92.9	81.0
	時系列 3	78.6	71.4	100.0	83.3
	時系列 A	78.6	71.4	78.6	76.2

表 3.8 人物認証成功率 1

	S1			S2			S3		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
P01	1	10	6	1	6	10	1	10	6
P02	2	11	1	2	10	11	9	7	4
P03	10	3	2	10	3	1	10	3	1
P04	4	9	1	9	4	7	4	9	1
P05	5	11	6	5	6	11	11	1	14
P06	6	1	10	10	1	2	6	1	10
P07	6	7	1	7	2	10	7	10	9
P08	9	8	4	8	9	4	8	9	4
P09	9	1	4	9	4	8	9	4	1
P10	10	1	2	10	2	1	10	1	11
P11	11	1	2	11	2	1	11	1	10
P12	12	14	5	12	14	5	12	14	5
P13	10	13	3	10	3	2	10	3	2
P14	14	11	2	11	14	2	14	11	2
Rate	71.4	100.0	100.0	64.3	85.7	85.7	71.4	78.6	78.6

表 3.9 人物認証成功率 2

	S1			S2			S3		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
P01	1	6	11	1	6	11	1	11	6
P02	11	2	1	2	11	10	9	4	1
P03	3	6	10	3	6	10	3	10	6
P04	1	9	4	9	4	11	1	9	4
P05	1	5	11	5	1	11	1	11	14
P06	1	6	11	6	1	3	1	6	11
P07	1	7	6	7	11	6	11	7	9
P08	9	8	4	8	9	1	8	9	4
P09	1	9	4	9	4	1	9	4	1
P10	3	1	6	10	2	6	10	1	6
P11	11	1	6	11	1	2	11	1	6
P12	12	14	5	12	14	5	12	14	5
P13	3	6	1	13	3	2	3	13	2
P14	11	14	1	14	11	1	14	11	2
Rate	28.6	78.6	85.7	92.9	100.0	100.0	57.1	78.6	85.7

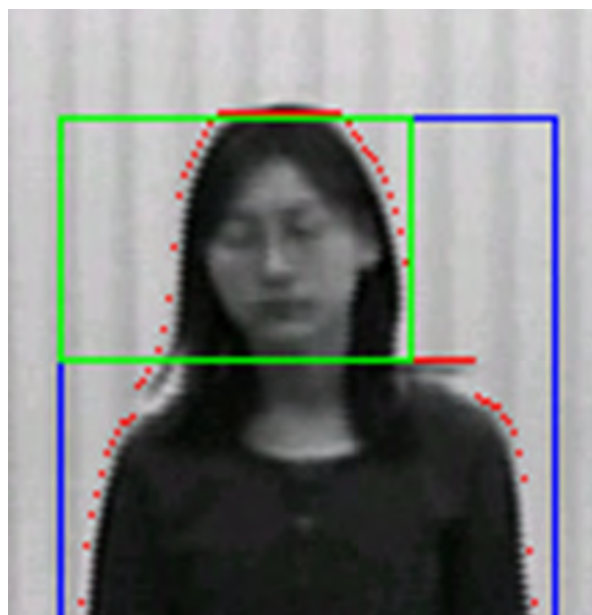


図 3.18 Person12 の頭領域検出失敗例

表 3.10 人物認証成功率 3

	S1			S2			S3		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
P01	1	10	6	1	6	10	1	10	6
P02	2	11	1	2	10	11	9	7	4
P03	10	3	2	10	3	1	10	3	1
P04	4	9	1	9	4	7	4	9	1
P05	5	11	6	5	6	11	11	1	14
P06	6	1	10	10	1	2	6	1	10
P07	6	7	1	7	2	10	7	10	9
P08	9	8	4	8	9	4	8	9	4
P09	9	1	4	9	4	8	9	4	1
P10	10	1	2	10	2	1	10	1	11
P11	11	1	2	11	2	1	11	1	10
P12	12	14	5	12	14	5	12	14	5
P13	10	13	3	10	3	2	10	3	2
P14	14	11	2	11	14	2	14	11	2
Rate	57.1	64.3	78.6	64.3	71.4	71.4	92.9	100.0	100.0

が出来た。したがって、本実験で用いた人物特徴は全て有効であると言える。今回は実装していない、頭髪量や動特性といった人物特徴においても、安定して取得できる手法を開発し適用することで人物認証の成功率向上を期待することが出来る。

時系列認証手法 表 3.7 に時系列認証手法の違いによる人物認証成功率の違いを示した。単純な多数決よりは、順位による時系列フィルタリングを用いた手法の方が良い結果となっている。しかし、全ての順位を反映させた結果では多数決の手法よりも精度が低下しているものが見られる。これは低い順位として認識された人物の結果は、進入人物とほとんど関係ない人物であるのにも関わらず、時系列認証へ関与させたために悪影響をもたらしたと考えられる。実際に、3位までのみを反映させた「時系列3」の結果が、最も安定してよい結果になっていることから推測できる。今回は、各画像の人物認証結果として3位までの人物候補を出力している関係から、時系列フィルタリングに3位までを適用させる手法を試みたが、どの順位までを反映させると最も結果が良くなるかを調べる必要がある。

人物認証結果 表 3.8 から 3.10 に各画像の認証結果の 3 位までを反映させた時系列フィルタリング手法を用いたときの人物認証結果を示した。学習に用いた動画像と認証に用いた動画像が一致するときは概ね 100% に近い結果となっている。一方、そうでない場合は 60%~70% 程度の結果となっている。認証精度が低下した原因としては、人物がカメラへ近づく方向へ動くとき、取得する人物特徴の値が変化することが考えられる。人物特徴は比率を用いており、遠方と近傍でも同じ特徴量が得られることを期待したが、ある程度は変化し認証結果に悪い影響を与えていることが考えられる。

3.5.4 行動習慣性特徴を用いた場合の評価実験

実験条件

CPU:Pentium4-3.0GHz の計算機にて、640 × 480 の解像度能力を持つ Matrox 社製 Meteor キャプチャーカードを使用した。プログラムは、C++ 言語を基本にした Microsoft 社製 Visual C++.net 2003 にて開発した。実験は室内環境にて 8 人の被験者を対象とした。カメラは研究室広域を撮影するように設置した。被験者には、その視野内に進入してもらい自分の机に座り、一定時間後に退出するという動作を行ってもらう。その様子を DV ビデオカメラにて撮影し記録した。評価実験では、条件を同一にするために DV ビデオカメラの映像を計算機に入力し評価を行った。なお今回の実験では、人物の行動習慣性を測定するための観測点は各机の中心を基準に 6 点設定して行っている。

実験結果

図 3.19 と図 3.20 に実験結果を示す。図 3.19 は従来使用していた 4 特徴のみの結果、図 3.20 は今回採用した行動習慣性を追加した 5 特徴の結果となっている。縦軸は順位重み付けによる点数であり、高いほど良い結果である。黄色は 1 位の点数、赤紫色は 2 位の点数、青色は 3 位の点数を示している。1 位の点数に注目すると、8 人中 6 人は点数が増えている。また、1 位と 2 位の点数の差は 8 人中 5 人が大きくなっている。この実験結果より、行動習慣性を新たに人物特徴として追加することにより、人物認証の精度が改善されていることが確認できた。

3.5.5 在席習慣性特徴を用いた場合の評価実験

実験条件

実験は図 3.21 に示すような屋内環境にて行った。カメラは東芝製 IK-TF5C、レンズはキヤノン製 JK-Z06P を装着し、絞りや焦点距離などを固定した。映像は Pentium4-3.8GHz の計算機に 640 × 480 の解像度能力を持つ Matrox 社製 MeteorII-CameraLink を

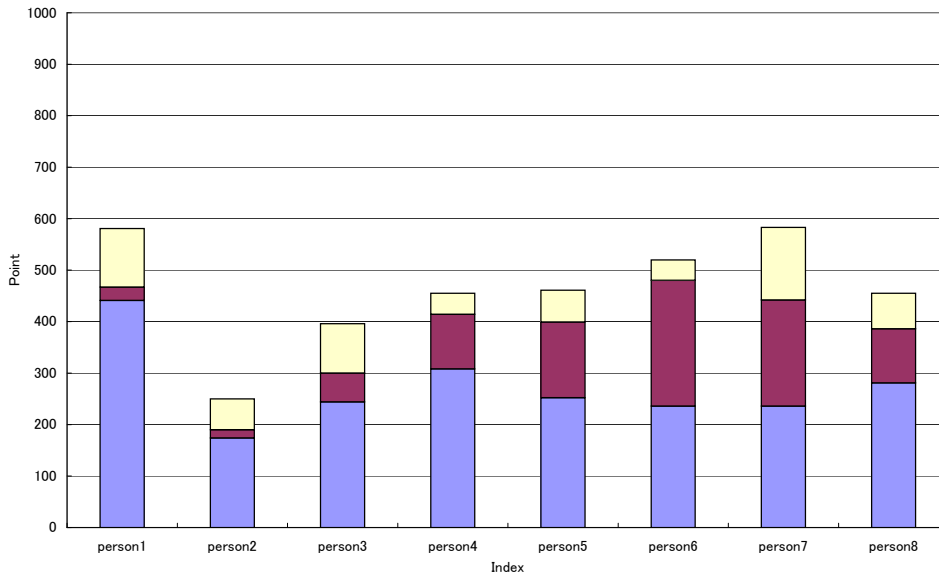


図 3.19 形状特徴のみ利用した場合の人物同定精度

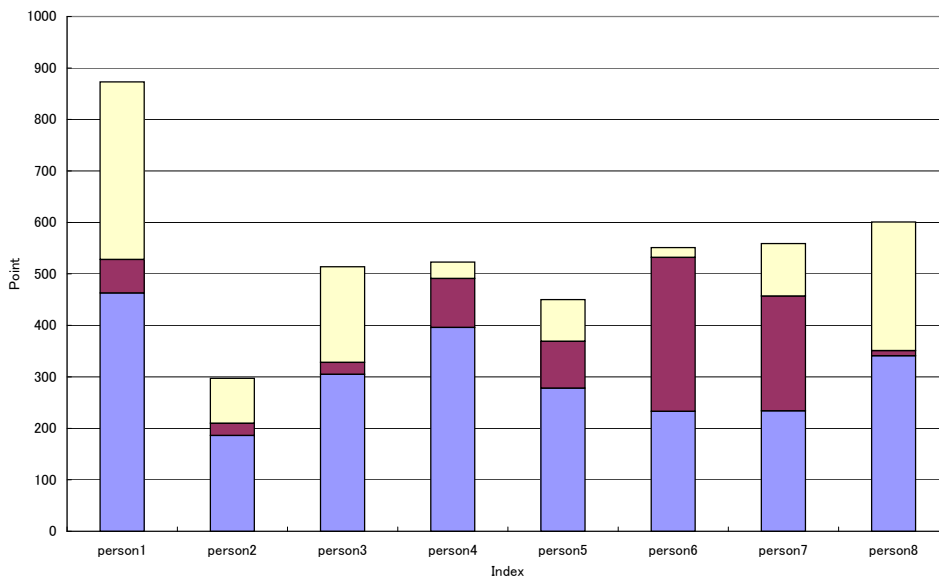


図 3.20 行動習慣性特徴を加えた場合の人物同定精度

装着し，無圧縮 AVI ファイルとして保存した．保存したファイルから評価実験用プログラムに動画像を入力して評価を行う．評価実験用プログラムは処理速度を高めるために C++ 言語を基本にした Microsoft 製 Visual C++ 2005 を用いた．被験者は図 3.21 に示す実験室を普段から利用している 9 人を対象とした．

図 3.22 は実験環境の模式図である．図中の椅子の箇所は人物が在席可能である作業用パソコンの場所を示している．今回は在席習慣性を得るための在席観測点は，この 6 つの

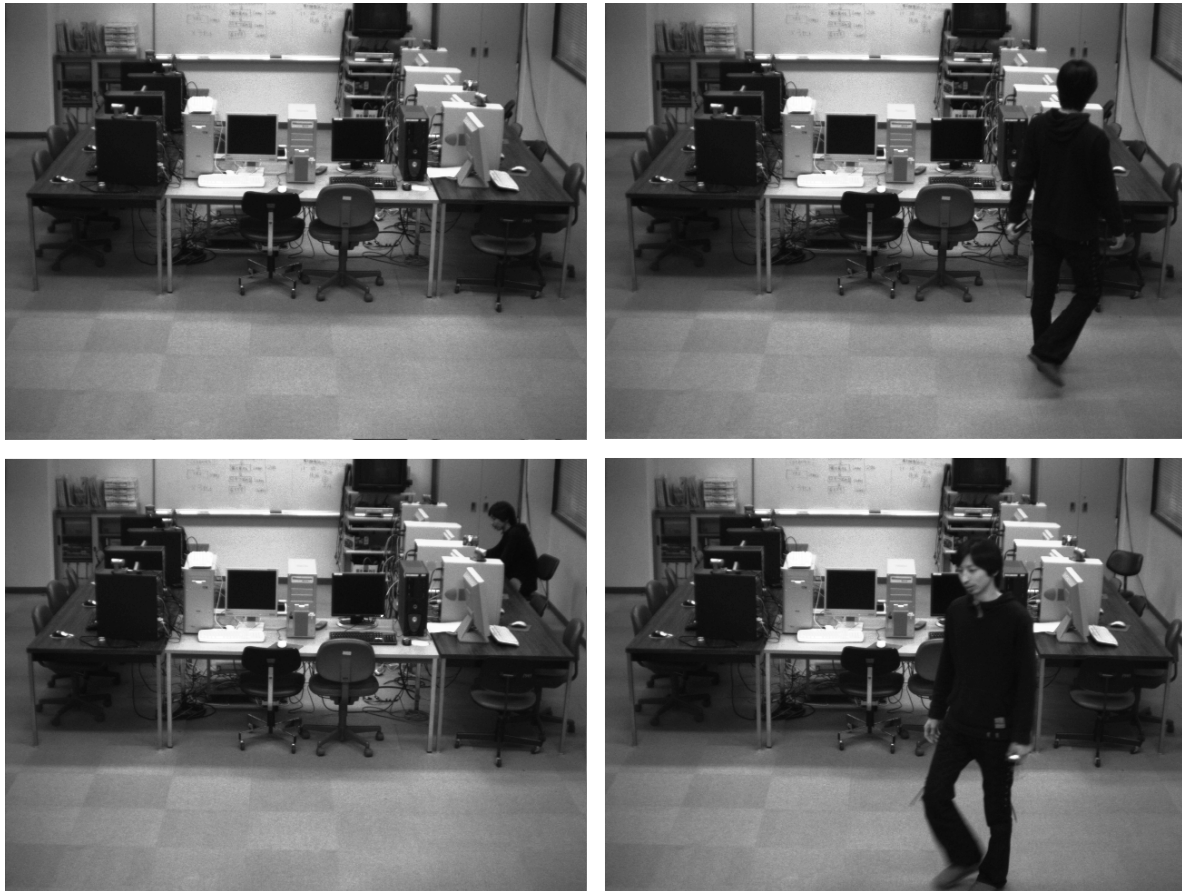


図 3.21 実験環境

椅子の場所へ設定した。被験者は図中の点線で示す経路を通り、在席観測点に至る。カメラは被験者が意識しないように、図中の C で示す位置の天井に近い箇所へ取り付けられた。この箇所で得られる画像は図 3.21 の左上の画像である。各被験者は実験室にて研究を行うために図 3.22 の左上に示す入り口から部屋へ入る（図 3.21 右上）。被験者は研究や嗜好によってよく使う作業用パソコンがあり、そのパソコンにて作業を行う（図 3.21 左下）。作業後は入り口の方へ部屋を退出していく（図 3.21 右下）。今回は上記流れの動画を各被験者ごとに 5 回分撮影した。この 9 人分の 5 回、つまり 45 セットを学習データとして蓄積し、各セットにて認証率の評価実験を行った。

実験結果

表 3.11 は実験結果を示す。横軸の P1~P9 は人物ごとの、縦軸の Method1~4 は表 3.12 に示した手法ごとの違いである。表中の値は実際に進入した人物が得た蓄積順位点の順位を示しており、「1」と表記されていればその人物が最も高い順位点を得たことを示している。つまり、「1」に近い数が多いほど同定成功の割合が高いということになる。

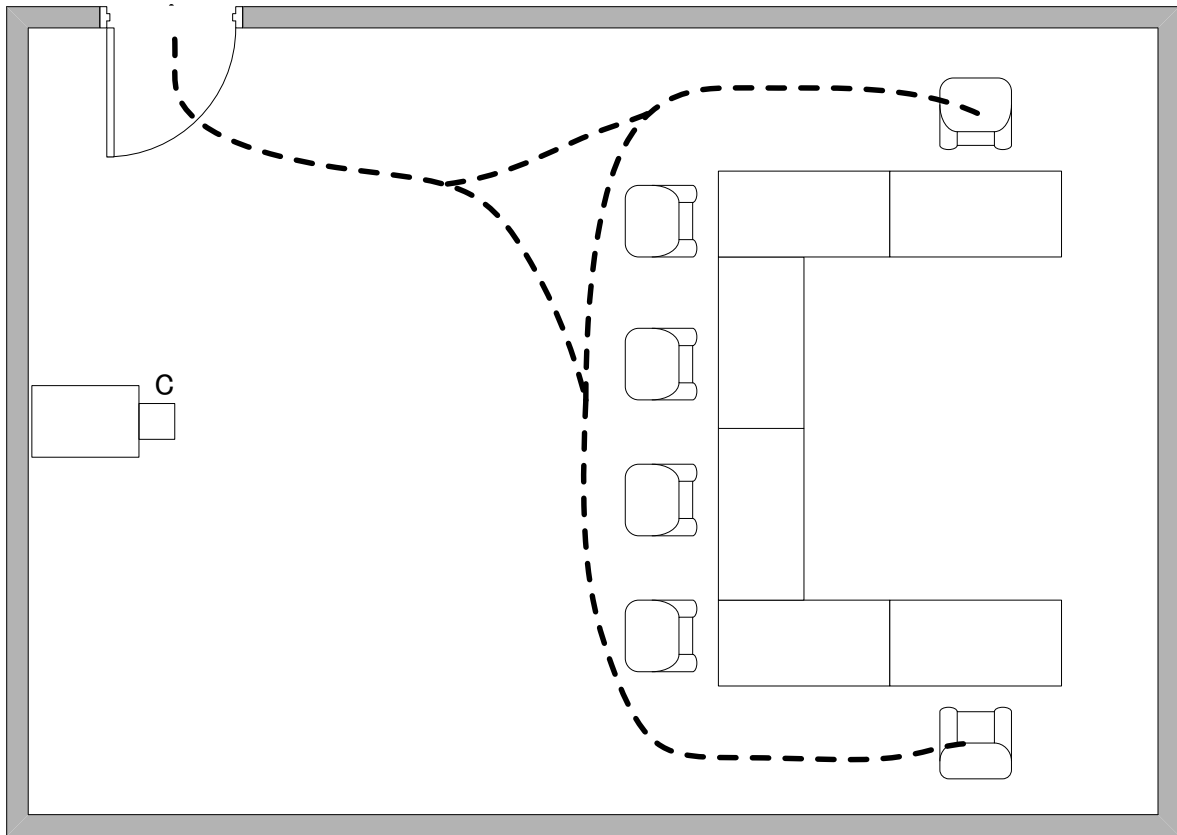


図 3.22 実験環境の模式図

表 3.13 は在席習慣性利用手法別の人物同定に成功した割合を示している．図 3.23 はその表をグラフにしたものである．1st は被験者の累積順位点が 1 位になった割合，2nd はその累積順位点が 1～2 位になった割合，3rd はその累積順位点が 1～3 位になった割合である．1st の結果を見ると Method2 が最も高く，2nd と 3rd の結果では Method2 と Method4 が高い結果となった．一方，在席習慣性を唯一用いていない Method1 は全てにおいて最も低い結果となった．

図 3.24 に順位点蓄積の成功例として人物 P3 が進入したときの様子を示す．始めのフレームでは人物の順位点が絡み合っているが，次第に順位点の蓄積によって安定していく様子が確認できる．図 3.25 に順位点蓄積の失敗例として人物 P5 が進入したときの様子を示す．人物 P5 は最終的に 3 位と判断されており，異なる人物を同定した結果となった．

考察

特徴抽出について 本報告では在席習慣性を導入するために実際に普段から利用している実験室にカメラを設置し撮影を行った．実験室は複雑な環境であり，背景差分での安定した人物抽出が困難であった．例えば，後ろ向きの頭部は明度的に暗いため，背景の明度が

表 3.11 各人物の累積順位点における順位結果

	P1	P2	P3	P4	P5	P6	P7	P8	P9
Method1	2	1	3	1	4	2	2	1	5
	2	1	3	1	1	2	1	2	2
	2	1	2	1	4	4	1	1	3
	2	1	1	1	3	1	1	1	3
	2	4	3	1	3	1	1	1	2
Method2	2	2	1	1	2	1	1	1	1
	2	2	1	1	2	1	1	1	1
	1	2	1	1	2	3	1	1	1
	2	2	1	1	2	1	1	1	2
	2	5	1	1	2	1	1	1	1
Method3	2	2	1	2	2	1	1	1	1
	2	2	1	5	2	1	1	1	1
	1	2	1	3	2	3	1	1	1
	2	2	1	2	2	1	1	1	2
	2	4	1	2	2	1	1	1	1
Method4	2	2	1	1	2	1	1	1	1
	2	2	1	1	2	1	1	1	1
	2	2	1	1	2	3	1	1	1
	2	2	1	1	2	1	1	1	2
	2	4	1	1	2	1	1	1	1

暗い領域と重なるときに差が無くなり，検出できなくなる．人物領域検出は局所射影することによって，背景差分の乱れが吸収され比較的安定していたが，面積の小さい頭領域は，取得に失敗することがあった．

これに対処するためには人体モデルを導入し，体領域など比較的安定して抽出することの出来る領域からモデルの当てはめを行い，頭部領域を推測するといった枠組みが必要であると考えられる．

在席習慣性の導入について 在席習慣性導入の成果を見るために，表 3.12 に示した 4 つの方法を試した．まず，在席習慣性を全く用いない Method1 が全ての事例において成功した割合が低かったことから，在席習慣性導入による同定成功割合の改善が認められた．

Method3 が Method2,4 に比べて若干であるが悪い結果となったのは，在席習慣性の重

表 3.12 順位点蓄積手法

Method1	在席習慣性を除く 4 つの形状特徴の事後確率を用いる
Method2	4 つの形状特徴と在席習慣性を同等に扱う
Method3	4 つの形状特徴と重みを 4 倍にした在席習慣性を用いる
Method4	在席習慣性がある場合は 5 特徴，それ以外は 4 特徴

表 3.13 在席習慣性利用手法別の同定成功の割合

	1st	2nd	3rd
Method1	0.47	0.73	0.89
Method2	0.64	0.96	0.98
Method3	0.53	0.91	0.96
Method4	0.62	0.96	0.98

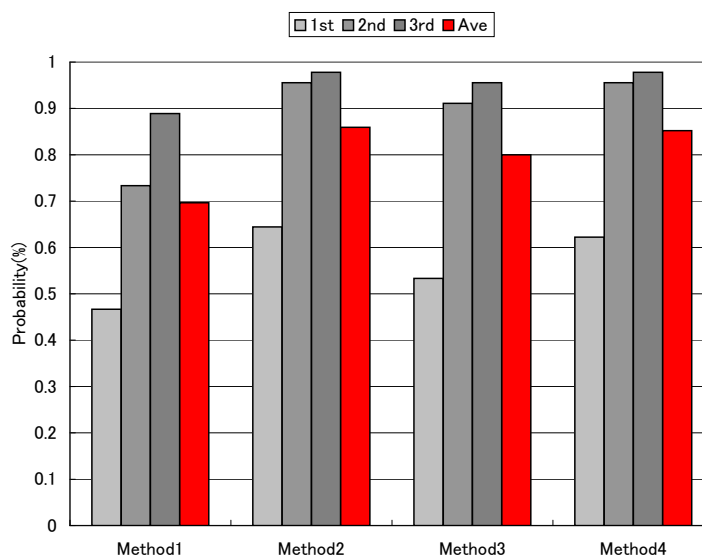


図 3.23 在席習慣性利用手法別の同定成功の割合

みを大きくしたことが原因であると考えられる。その理由として、本提案手法では複数の特徴を取得することで、一つの特徴が不安定になった場合でも他の特徴でその影響が抑えられることが期待出来るためである。

Method3 では種類の違う形状特徴と在席習慣性を同じ影響力にするのは良くないと考え、在席習慣性の重みを大きくし両特徴の影響力を同じに調整した。しかし、結果が良く

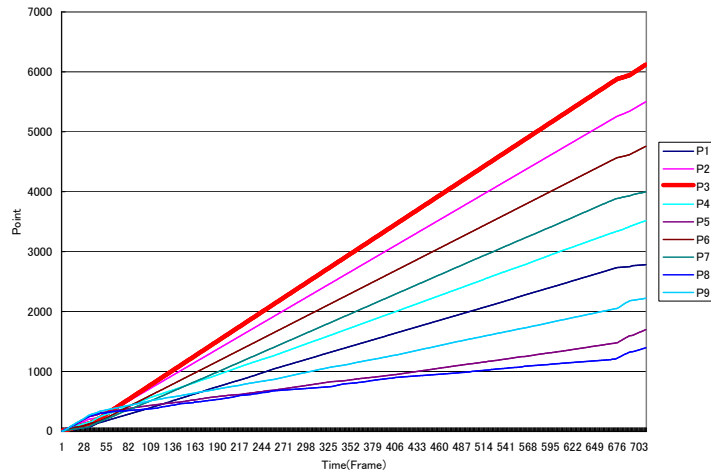


図 3.24 順位点蓄積の様子 (成功例)

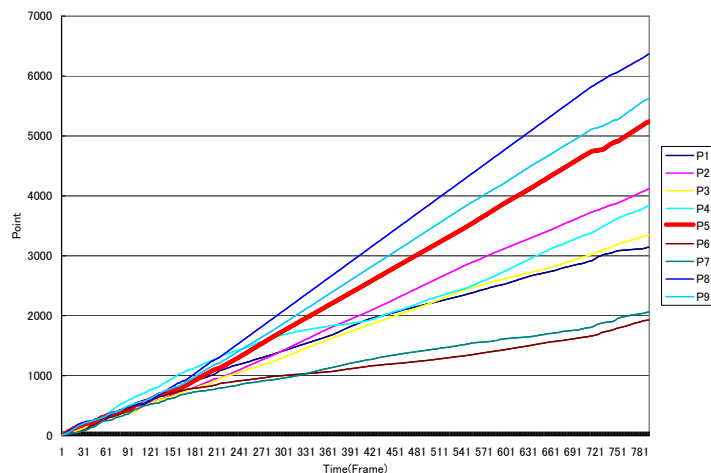


図 3.25 順位点蓄積の様子 (失敗例)

なかったことから，現在利用している一つのベイズ識別器においては，ある特定の特徴に重みを与えることは良くないということを示すものであると考える．この識別手法を発展させるために，アンサンブル学習法など，複数の識別器を用いる枠組みを導入し，評価していく必要があると考える．

n 位同定成功の割合について 図 3.23 に示した実験結果から，どの手法においても進出した人物が 2 位から 3 位までに含まれる確率は平均して 92.1% と高い確率であるのに対し，1 位と判断された確率は平均して 56.5% という結果となった．図 3.24 や図 3.25 で示したとおり，最終フレームでは人物の累積順位点は安定している．したがって，この状態から成功の割合を向上させるには，現在利用している特徴抽出の精度を高めることが必要であることを示していると考えられる．

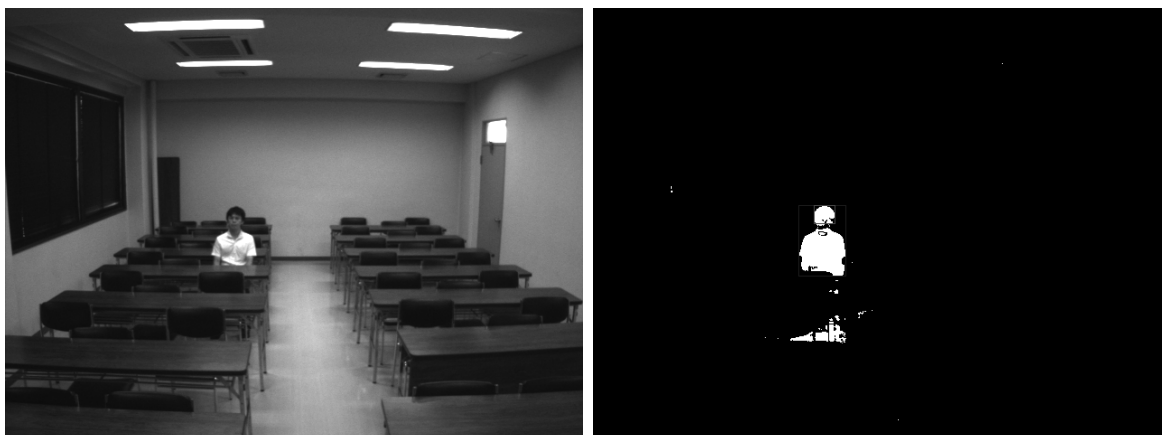


図 3.26 実験環境と処理例 (73 号講義室)

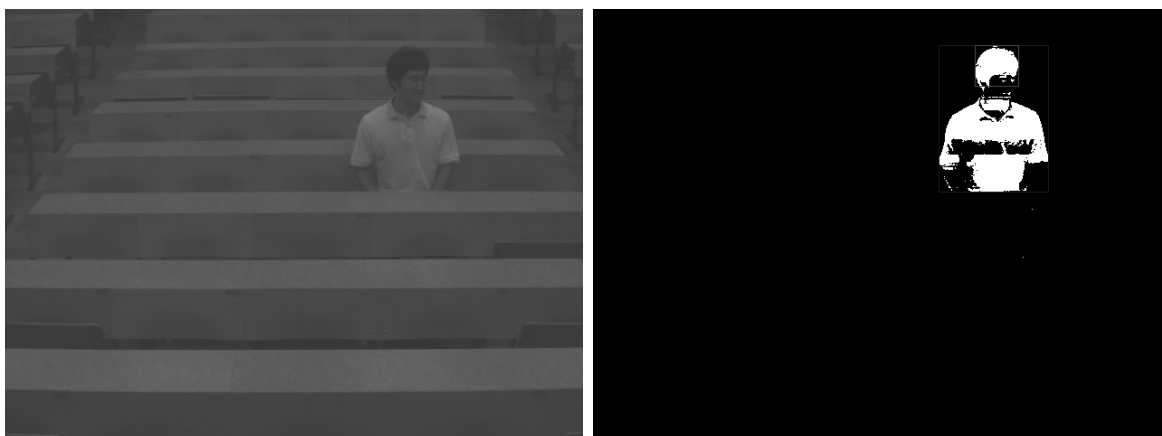


図 3.27 実験環境と処理例 (71 号講義室)

3.5.6 滞留習慣性特徴を用いた場合の評価実験

実験環境

図 3.26 ~ 3.28 に示す 3 つの環境に 1 台のカメラを設置して映像の取得を行った。実験環境の状態を表 3.14 に示す。カメラはデジタル伝送が可能な Camera Link カメラである東芝 IK-TF5C を、レンズはキヤノン JK-Z06P (焦点距離 8 ~ 48mm, 実験では 8mm 固定) や東芝 JK-L04TF (焦点距離 4mm) を装着し、明るさが飽和しない絞り位置で固定して設置した。

カメラの映像は Pentium4-3.8GHz の処理能力を持つ計算機に VGA 解像度能力を持つ Matrox MeteorII Camera Link を装着して取得し、無圧縮 AVI ファイルとして保存した。提案手法を Microsoft Visual C++ 2005 にてプログラム実装し、上記ファイルから



図 3.28 実験環境と処理例（実験室）

表 3.14 実験環境の状態

環境名	73号講義室	71号講義室	実験室
レンズ	JK-L04TF	JK-Z06P	JK-L04TF
n の設定値	500	500	500
λ の設定値	15	10	10
学習人数	9人	9人	6人

映像を入力して評価を行った。表 3.14 中の n と λ の設定値とは、3.4.2 章の動的背景抽出法におけるパラメータ類であり、環境ノイズに反応しない値を設定した。人物特徴の頻度分布として作成して順位点の蓄積は人物が視野内に進入し退出した時点まで行っている。全てのセットから頻度分布を作成し、その一部のセットで人物同定の成功人数を評価した。

各実験環境は次に挙げる項目が評価対象である。73号講義室は、机に移動するための通路が中央にあるため、移動時と着席時の形状特徴の取得が可能であり、滞留習慣性の影響が大きいと考えられる環境である。被験者は中央の通路を通り、普段好んで座る席に座るように指示を与えた。着席中は前方で行っている講義に注目してもらい、数十秒後には部屋を退席してもらおう。この一連の流れを1セットとして、6セット分の映像を取得した。なお、滞留習慣性の影響力を評価するために、6セット目のみ違う席に座るように指示を与えた。71号講義室は、視野内に机のみが写っている状態であり、73号講義室と比較して着席時の形状特徴しか取得出来ないという違いがある。他の条件は73号講義室と同等である。実験室は、固定された席がない環境である。ここでは、被験者は実験室の左方から視野内に進入し、常に視野内にて移動した状態を保つように指示を与えた。つま

表 3.15 人物同定に成功した人数

	成功人数	2位以内
73号講義室-1	9/9	9/9
73号講義室-2	9/9	9/9
71号講義室-1	8/9	9/9
71号講義室-2	7/9	8/9
実験室-1	5/6	6/6
実験室-2	5/6	6/6

り、滞留習慣性による判別能力を意図的に落とし、形状特徴のみによる判別能力を評価するものである。また、身体のおクルージョンの発生による差を評価するために、その発生しない前方領域と発生する後方領域とに分けて撮影した。各領域ごとに3セットの映像を取得した。なお、これらの実験では、各部屋の撮影時に時間の都合の良かった人物に実験協力をお願いして撮影しており、意図的に人物を指定してはいない。

実験結果

表 3.15 に人物同定に成功した人数の数を示す。各項目の分子が成功した人物の数、分母が学習人物の数である。成功人数の列は該当者の蓄積順位点が最も高くなった人物の数であり、2位以内の列は蓄積順位点が2位以内に含まれる人物の数である。

「73号講義室-1」は普段座る席に座ったときの結果、「73号講義室-2」は意図的に座る席を変更して滞留習慣性の影響を評価するための結果である。「71号講義室-1」と「71号講義室-2」も同様である。「実験室-1」は身体領域のおクルージョンが発生しないときの結果、「実験室-2」はそれが発生するときの結果である。

図 3.29 は、71号講義室の評価実験における順位点蓄積の結果であり、上は成功例、下は失敗例である。横軸は時間（フレーム数）、縦軸は蓄積順位点であり、各データは各人の蓄積順位点の推移の様子を示している。初期段階では各人物の蓄積順位点の差は認められないが、滞留中に各人物の順位点が蓄積され、最終的には該当者が最も高い蓄積順位点となったことが分かる。

考察

提案手法が想定する環境での評価実験は、従来手法が利用しているようなデータベースが利用できないため、自ら評価用のデータ取得を行う必要があった。そのため、実験の評価は学習データの一部を利用せざるを得なかった。しかし、実験結果から十分な学習を行

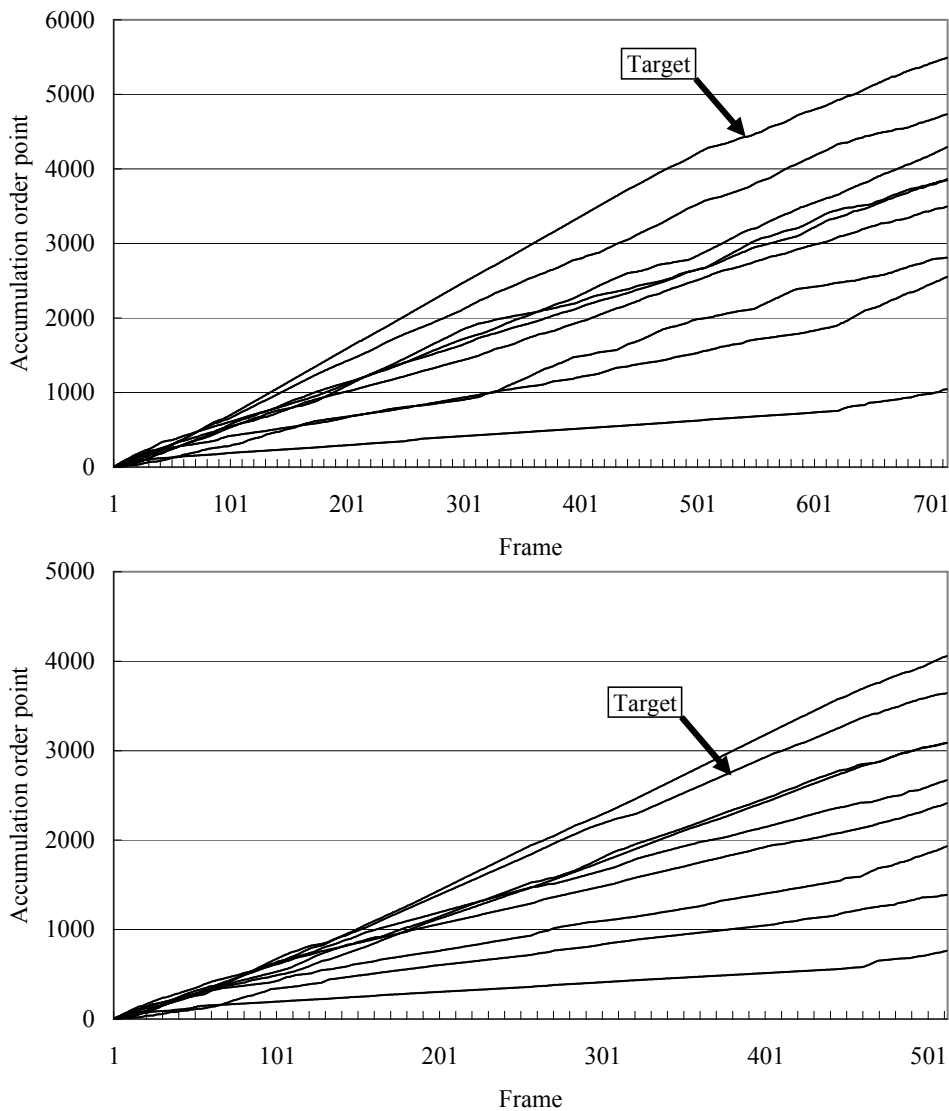


図 3.29 順位点蓄積の結果（上：成功例，下：失敗例）

うことで高い精度で人物同定を行うことが出来ることが示唆された。

図 3.29 の失敗例は，71 号講義室で人物同定に失敗した人物の順位点蓄積の様子を示している．この失敗した人物は女性で髪型がストレートであった．この場合，頭領域の微分値が緩やかになってしまい，頭領域の検出失敗の要因となった．頭頂部から微分検出の距離を制限するなどの対処が必要だと考えられる．

3 つの環境にて評価実験を行ったが，人物同定に失敗したのは 1~2 人であり，高い精度で人物同定を行うことが出来ることを確認した．特に実験室では，意図的に滞留習慣性の有効性が働かないような状態で行ったにも関わらず，他の環境と同様の精度であった．また，身体のオクルージョンの発生による精度低下が認められないことも評価すべき点である．この実験室での実験は，従来手法の適用が最も困難な環境であり，本提案手法の有

効性を示せたものである。

3.5.7 総合実験

実験概要

始めに提案手法の基本性能を確認する実験を行い、今回選択した体型特徴の有効性を評価した。次にその実験結果を基に大学院講義における出欠管理を想定した有用性評価実験を行った。カメラは Camera Link 接続カメラである東芝 IK-TF5C を、レンズは東芝 JK-L04TF (焦点距離 4mm) を装着し、明るさが飽和しない絞り位置で固定し、実験環境の天井部付近から床面に向けて斜め下方に傾けて設置した。カメラは VGA(640 × 480) 解像度能力を持つ Matrox MeteorII-Camera Link に接続し、無圧縮 AVI ファイルとして 30FPS の速度で蓄積保存した。Pentium4-3.8GHz の処理能力を持つ計算機上で、Microsoft Visual C++ 2005 にて本手法のプログラムを実装し評価を行った。なお情報取得にかかる処理時間は、平均 10.1ms であった。

基礎実験

提案手法の基本性能を評価するための実験を行った。図 3.30 に実験環境を示す。予備実験により、環境ノイズによる変化が起きない値として式 (3.1) の $n = 100$, $\lambda = 25$, 各特徴値を十分に記録できる値として式 (3.2), 式 (3.3), 式 (3.4) の $N = 256$, $k_1 = 60$, $k_2 = N$, $k_3 = 2 * N$, $k_4 = 0.5 * N$ とした。また、3.4.6 節に示す終了条件の一定の差は 0.5 に設定した。なお、この実験では体型特徴の有効性を評価することを目的としているため、滞留習慣性を除く F1 から F4 の 4 つの特徴を用いた。

被験者の体型について人体寸法データベース [30] との比較検討を行った。図 3.5 に示す特徴 A から F と人体寸法データベースの特徴との対応関係を表 3.16 に示す。例えば、Feature 列の A 行は人物特徴 A はデータベースの D3-Maximum body breadth の箇所に対応させたことを示す。身長 (人物特徴 B) 以外の特徴値記載について同意が得られなかった 1 名を含む、8 名の被験者体型を表 3.17 に示す。各特徴の単位は mm である。Ave. は被験者の平均値、D.A. はデータベースの青年男性の平均値を示す。これらのデータについて平均値の差の検定を行った。被験者の特徴 (サンプル数: 7 から 8) とデータベースの青年男性の特徴 (サンプル数: 216 から 217) に対して T 検定 (棄却域 5%) を行い、全ての特徴において平均値の差を認めないという結果が出た。このことから、今回の被験者の体型は概して標準的であり、一つの体型特徴のみで有効な同定が行える可能性は低いと考えられる。

被験者はカメラ視野の右上から中央を通り左下へ抜けていく。この流れを各被験者ごとに 6 回撮影した。全ての被験者において、同定結果が確定したのは、3.4.6 節に記載した

表 3.16 人物特徴と人体寸法データベースの対応関係

Feature	ID	Name
A	D3	Maximum body breadth
B	B1	Height
C	A4	Ear to ear breadth
D	A36	Total head height
E		No data
F		No data

第1条件「被同定者が視野内から消失」を満たしたことによるものであった。このとき、被験者が視野内に出現してから同定結果が確定するまでの時間は平均 9.2s であった。

表 3.18 に 6 分割交差検定 (cross validation, CV) による実験結果を示す。Scene 列の S* は人物同定に用いた動画像の識別子である。例えば S1 の場合は S2 から S6 の 5 回の動画像列から得られた特徴値を学習データとし、S1 のデータにて人物同定の成功率を評価した結果である。表中の値は人物同定の成功率を示している。

D*列は各体型特徴の有効性を評価するために、特定の体型特徴を用いないときの結果である。例えば、D1 は表 3.1 の F1 以外の F2, F3, F4 のみを、D0 は F1, F2, F3, F4 の全てを用いたときの結果である。この実験では F2, F3, F4 の体型特徴を用いたときが最も良い結果となった。

次に時間方向の事前確率として順位に基づく確率を導入したことが有効であることを確認するための実験を行った。表 3.18 の D1N 列は、時間方向の事前確率を全ての人物で同一となるように設定したときの結果である。順位に基づく確率を導入したときの結果である D1 と比較すると、D1 の方が良い結果となっている。

図 3.31 に時間方向に統合した事後確率推移の一例を示す。人物 p1 が進入してきたときの例を示しており、矢印で示す線が該当人物の \bar{P}_{post} の推移を示している。前半の時間帯では各人物の確率に顕著な差が認められないが、後半では p1 の確率が最も高くなっている。

大学院講義における出欠管理を想定した有用性評価実験

基礎実験の結果を基に選定した体型特徴 F2, F3, F4 及び滞留習慣性特徴 F5 を用いて、大学院の講義にて本提案手法を適用することを想定し、有用性評価実験を行った。鹿児島大学大学院情報工学専攻では、講義の種類によって受講者数に差があるが、平均して 15 名程度であることから評価実験の被験者数も同じ人数とした。



図 3.30 実験環境（基礎実験）

表 3.17 被験者体型（基礎実験）

ID	A	B	C	D
p1	518	1813	218	230
p2	452	1749	218	315
p3	455	1720	190	240
p4	409	1783	180	214
p5	462	1658	191	219
p6	550	1755	212	378
p7	-	1600	-	-
p8	365	1640	179	231
Ave.	459	1715	198	261
D.A.	470	1714	191	239

予備実験の結果から，環境ノイズによる変化が起きない値として式 (3.1) の $n = 100$, $\lambda = 10$ ，各特徴値を十分に記録できる値として式 (3.2)，式 (3.3)，式 (3.4) の $N = 250$, $k_2 = 200$, $k_3 = 200$, $k_4 = 0.5 * N$ とした．また，3.4.6 節に示す終了条件の一定の差は 0.5 に設定した．

被験者の体型について人体寸法データベース [30] との比較検討を行った．図 3.5 に示す特徴 A から F と人体寸法データベースの特徴との対応関係を表 3.16 に示す．身長（特徴 B）以外の特徴値記載について同意が得られなかった 2 名を含む，15 名の被験者体型

表 3.18 実験結果 (基礎実験)

Scene	D0	D1	D2	D3	D4	D1N
S1	87.5	87.5	87.5	87.5	75.0	75.0
S2	100.0	100.0	75.0	87.5	100.0	100.0
S3	87.5	100.0	87.5	87.5	100.0	87.5
S4	87.5	87.5	87.5	87.5	75.0	87.5
S5	75.0	87.5	75.0	87.5	75.0	87.5
S6	87.5	87.5	87.5	75.0	75.0	75.0
Ave.	87.5	91.7	83.3	85.4	83.3	85.4

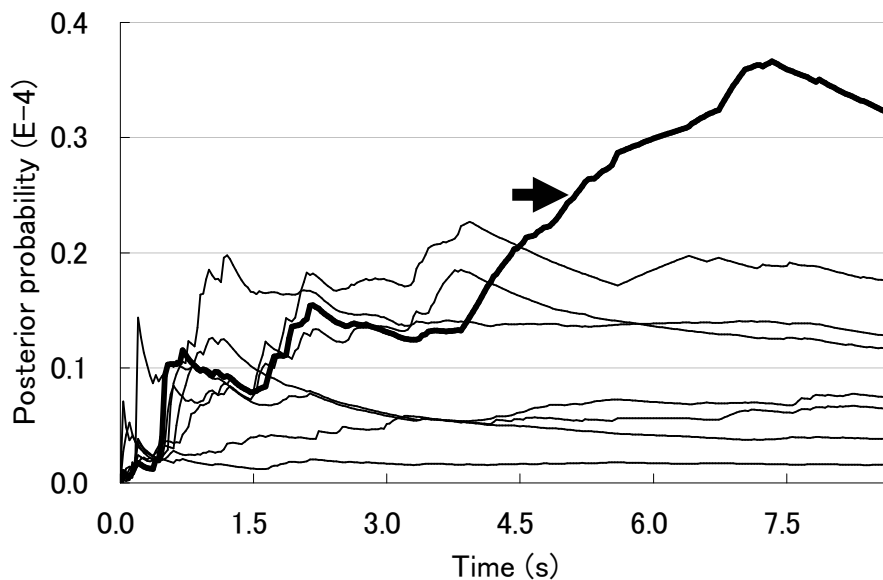


図 3.31 時間方向に統合した事後確率の推移 (基礎実験)

を表 3.19 に示す。各特徴の単位は mm である。Ave. は被験者の平均値，D.A. はデータベースの青年男性の平均値を示す。これらのデータについて平均値の差の検定を行った。被験者の特徴 (サンプル数: 13 から 15) とデータベースの青年男性の特徴 (サンプル数: 216 から 217) に対して T 検定 (棄却域 5%) を行い、全ての特徴において平均値の差を認めないという結果が出た。このことから、今回の被験者の体型は概して標準的であり、体型特徴のみで有効な同定が行える可能性は低いと考えられる。

実験は図 3.32 に示す構成の教室で行い、A の位置にカメラを設置して撮影を行った。図 3.33 はカメラから得られた画像の例である。被験者はカメラ視野奥から順に座っていく。このとき各人の滞留時間の分散を小さくするために、前方の列に座る人物は $B \rightarrow C \rightarrow D$

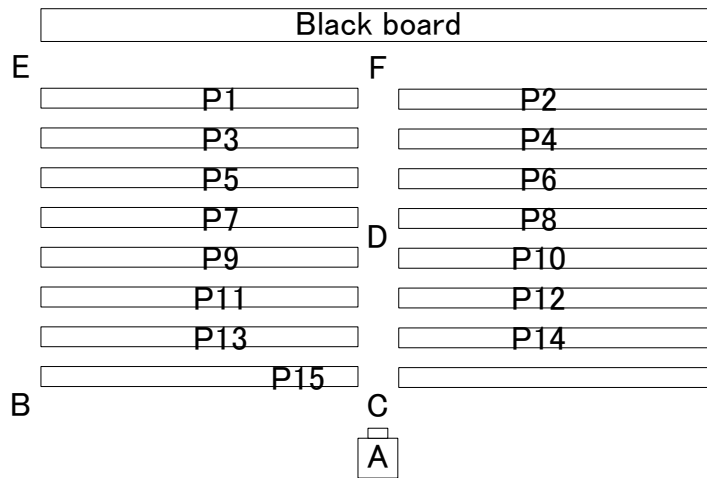


図 3.32 実験状況の説明図（有用性評価実験）

の経路で、後方の列に座る人物は $E \rightarrow F \rightarrow D$ の経路にて指定の席に着席する．その後は移動せずに着席状態を維持する．この一連の流れを 6 回撮影した．全ての被験者において、同定結果が確定したのは、3.4.6 節に記載した第 1 条件「被同定者が視野内から消失」を満たしたことによるものであった．このとき、被験者が視野内に出現してから同定結果が確定するまでの時間は平均 16.7s であった．

表 3.20 に 6 分割交差検定 (cross validation, CV) による実験結果を示す．人物同定成功率は平均 96.7% であり、失敗した人物は P9 が 1 回、P10 が 2 回であった．この実験結果に関して検定を行った．実験回数が少ないため、小標本を対象とした Student の t 分布を用いた区間推定を選択した．6 つの実験結果から求めた 95% 信頼区間は $96.7 \pm 3.9\%$ であった．また、実験を 90 回の試行とし、同定成功試行の成功率を 100%、同定失敗試行を 0% としたとき、95% 信頼区間は $96.7 \pm 3.8\%$ であった．15 人を対象に人物同定を行なう場合、1 名のみ失敗する際の平均成功率は 93.3% であり、上記の信頼区間に含まれている．一方、2 名の失敗が発生する際の平均成功率は 86.7% であり、信頼区間外となる．以上のことより、15 人を対象とした今回の実験において、提案方式を用いることにより、1 名の認識失敗に留められるという結果が得られたと結論した．

図 3.34 に取得した人物特徴 F2 の頻度分布例を示す．この例では、着席時に蓄積されたと思われる高い山が一つ (a)、移動中に蓄積されたと考えられる小さな山が三つ (b)、見受けられる．

考察

基本性能 人物領域検出手法として、背景差分法の一つである動的背景抽出法と射影による方法を用いた．背景差分法は環境の明度変化に弱い、動的背景抽出法では背景画像



図 3.33 実験環境（有用性評価実験）

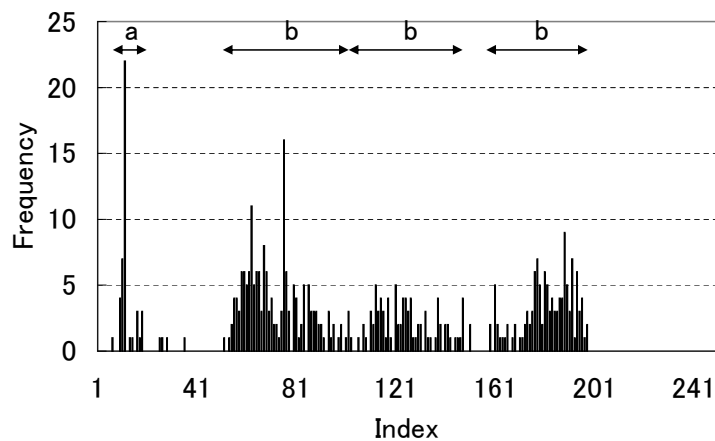


図 3.34 人物特徴 F2 の頻度分布例（有用性評価実験）

の揺らぎを学習することで対処している．今回行った基礎実験では，画像処理に失敗しているのはごく一部のフレームのみであり，人物同定への影響は少なかった．

提案手法の処理時間について，特徴取得の処理時間は平均 10.1ms，画像取得速度は 30FPS(33.3ms)であった．これは人物特徴の取得が比較的簡便な手法で構成されていることが要因であり，実時間認識が可能となる見通しが得られたと考える．また，同定可能となる時間について，基礎実験では平均 9.2s，有用性評価実験では平均 16.7sであった．一般に，人物同定にかかる時間が長くなることに比例して被同定者の負担が高くなるが，本提案手法では人物を拘束していないため，その問題が軽減される点が従来手法と大きく異なる．

3.4.4 節に示した人物特徴の離散化幅 N や 3.5.7 節に示した画像解像度は，今回は実験

表 3.19 被験者体型（有用性評価実験）

ID	A	B	C	D
P1	462	1658	191	219
P2	452	1749	218	315
P3	518	1813	218	230
P4	409	1783	180	214
P5	-	1670	-	-
P6	400	1685	185	235
P7	405	1660	170	237
P8	567	1552	346	314
P9	365	1640	179	231
P10	455	1720	190	240
P11	495	1715	200	250
P12	-	1600	-	-
P13	462	1735	208	245
P14	481	1745	190	210
P15	550	1755	212	378
Ave.	463	1699	207	255
D. Ave.	470	1714	191	239

的に設定した。離散化幅を広くし過ぎると識別能力が低下し、狭くし過ぎると同一人物と同定する性能が低下する傾向にあるが、この自動決定方式は未だ見出せていない。今後、MDL 基準を用いることによる離散化幅の自動決定法の検討を行っていく。また、適切な画面解像度について、低い解像度では人物同定成功率が落ちることが予想され、高い解像度では必要以上の情報を処理することで、不必要に処理速度が低下することが考えられる。この適切な解像度に対する検討も今後の課題である。

各モデルに基づく統合の有効性 本提案手法では、マルチモーダルモデルとマルチサンプルモデルを組み合わせた方式とし、評価実験でその有効性を評価した。

表 3.18 にて示したとおり、D1 の試行、つまり F1 以外の体型特徴を用いたときが最も高い結果となった。特徴 F1 が有効でなかった理由として、ベイズの定理では各特徴の独立性が求められるが、この特徴のみ、一つの人物領域から得られる二つの長さの比率にしておき、独立性条件が満たされなかったためと考えられる。

表 3.20 実験結果（有用性評価実験）

Scene	Rate	Fail ID
S1	93.3	P10
S2	100.0	
S3	100.0	
S4	93.3	P9
S5	100.0	
S6	93.3	P10
Ave.	96.7	

有用性評価実験では、同じ経路を通り同じ席に座る試行を6回撮影し評価を行ったが、この実験に引き続き、学習データは変えずに人物の着席位置を一組交換し、滞留習慣性の寄与を意図的に低下させた実験を行った。この場合、有効性評価実験で同定に2回失敗しているP10を正しく同定できたものの、人物同定成功率は平均86.7%となり、有効性評価実験(平均96.7%)に比べて低下する結果となった。このことから、体型特徴と滞留習慣性特徴の有効性は人物毎に異なっており、両者を組み合わせて同定に用いるマルチモーダルモデルを参考にした提案手法は有効であったと考える。

また、時間方向の事前確率を設定する方法として、各人物の事前確率を同一とする方法と、前提知識を用いて一枚の画像から算出した事後確率に基づく順位に応じた事前確率を設定する方法を考案し導入したが、表3.18にて示した通りD1NよりD1の試行、つまり順位に応じた事前確率を設定する方法が良い結果となった。このことから、マルチサンプルモデルに基づいた時間方向の統合において、順位に基づく事前確率を導入したことは有効であったといえる。

多人数の対応可能性 有用性評価実験では大学院講義へ適用することを想定し、15名を学習し人物同定の評価を行った。しかし、例えば高校の授業や大学の講義といった約50名規模を対象に適用するには、更に識別能力の向上が必要となる。

体型特徴は、極端に体型の異なる人物が対象者とならない限りは、人数が増加するに伴って識別能力が下がる。滞留習慣性特徴は、カメラの解像度を高めることや、人物のオクルージョンができるだけ発生しないような位置、例えば講義室内通路の延長線上の壁側などへ設置することで、更に識別能力を上げることが期待できる。また、コンビニ等におけるリピーター検出に応用する場合は、入り口やレジ付近にカメラを向け、全身が視野内に入る程度の距離に置くと特徴の取得が容易になる。ただし、外光による変化や人物のす

れ違い等も考えられるため，外光変化にロバストな背景差分や人物追跡のアルゴリズムなどが必要になる．

今回想定した比較的簡単な状況では，常時安定に取得できる，体型特徴と滞留習慣性の組み合わせで有効なシステムが構築できたが，例えば学部講義や高校講義のように，人数や流動性に関する制約が緩和された状況では，体型特徴のみでは不十分であり，頭部周辺の明度特徴や顔パターン情報，手足の動き特性など，常時安定に取得することが困難な画像特徴の追加導入を検討していく必要があるが，これらは今後の課題である．

第4章

人物状態認識エージェント

4.1 従来手法

テンプレートマッチング手法を用いた瞳検出や眼の輪郭線抽出の研究 [31][32] や色情報を用いた瞳領域の切り出しと二次元ハフ変換を用いた瞳検出 [33] など多数行われている。特に内面状態の推定に結びつくものとしてニューラルネット手法を用いた瞬き計測によるドライバの運転意識低下レベルの推定 [34][35][36][37] が提案されている。またテンプレートマッチング手法にて抽出した顔を構成する目や鼻・口・眉などのパーツの位置関係や形状を認識することで表情を推察する方式 [38] などが挙げられる人物の瞬目の開閉時間の変化と意識低下の相関は高いが、疲労状態・覚醒低下状態の推定には人物の全瞬目だけでは不十分で、半瞬目状態の推移的な情報の認識も必要である。また人間の主観的な解析結果だけであり、生理的指標との対照実験が十分に行われていない。

この章では二次元ハフ変換を用いて自動検出した目領域に含まれる瞳部分の面積・まぶたの形状変化を連続計測することにより求めた瞬目回数の変化と半瞬目状態との推移を、目視検出結果と生理的指標である脳波解析結果と比較した結果を中心にまとめる。

4.2 目の状態認識による内部状態の推定

心理的な動揺や怒り、疲労など人物の内部状態が外面に顕れる仕草・部位として、顔の局所の変形・表情変化、目線の動き、瞬目回数の変化、頭・手など身体一部の動き、貧乏揺すりなど身体全体の動き・姿勢変化、等が挙げられる。特に表 4.1 に示すように、視線、目の状態、瞬目など、目に関する仕草は、非常に多くの情報量を有している [39]。

本研究では画像計測により求めた瞬目の回数・間隔に基づいて疲労度の推定を行った。完全に目を閉じた状態となる通常の瞬目（以下、「全瞬目」）に加え、眠気を生じ目がとろんとした状態（以下、「半瞬目」）も同時に測定するための画像処理手法を新たに開発した。

表 4.1 目の状態と内部状態との定性的関係

仕草・部位	外部への顕れ	内部状態
視線	意識的に相手に向ける	目で主張
	無意識的に相手に向ける	好意を持つ
	笑いながら相手からそらす	否定的気持
	無意識的に相手からそらす	好きでない, 緊張
	視線を全くそらさない	緊張, 自意識過剰, 放心状態
	ゆらゆら・ぼんやりする	興味・関心が無い
	横目・斜め下に向ける	嫌悪, 軽蔑
目の状態	きょろきょろ落ち着かない	周囲が気になる, 自意識過剰
	目がうつろ	何かに心を奪われ虚脱状態
	目が潤む	感動, 悔しさ, 悲しみ
	目配せ	親しいコミュニケーション
瞬目	回数が多い	強く感情が動き, 落ち着こうとしている
	回数が少ない	落ち着きいていて, 真剣に思考中
	目を強くつむる	下向き: 我慢中
		上向き: 想起中
目を軽くつむる	疲れた一呼吸おく	

4.3 提案手法の流れ

4.3.1 概要

目の状態認識のための手順を説明する。処理の流れを図 4.1 に示す。始めに撮影された顔画像に対して目の領域を検出する。次に円の二次元ハフ変換により瞳部分の検出を行う。最後に抽出した瞳領域の面積から目の状態認識を行う。

4.3.2 目領域検出

顔面のグレイスケール画像の連続フレーム間差分により、顔面内部において瞬目等による明度変化が大きくなる点を累積し、最も変化量が大きくかつ画素が黒領域である点を瞳の一部分とし、隣接しあう画素の黒画素を探索し目領域を抽出していく。目領域の検出処理の流れ図を図 4.2 に示す。

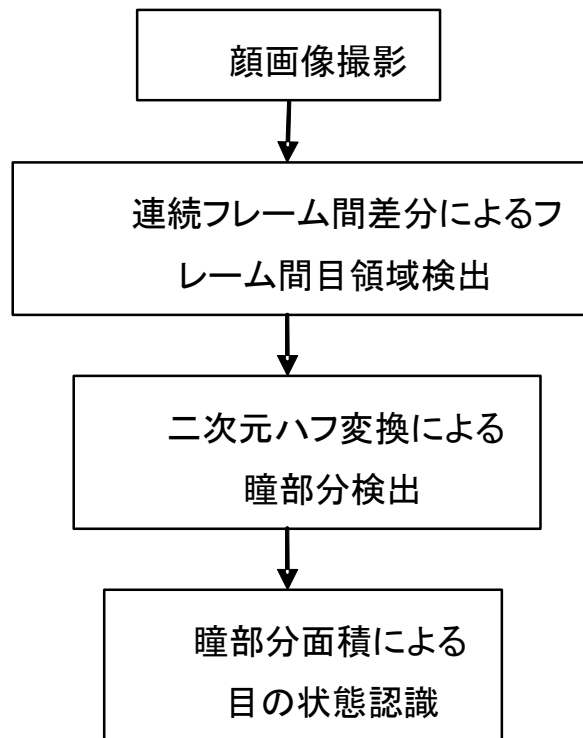


図 4.1 処理の流れ図

4.3.3 瞳部分検出方法

目領域から瞳位置を検出する。瞳は個人差によらずほぼ円形である事に注目し、二次元ハフ変換による手法を用いた。先に抽出されている眼領域にエッジ処理を加え、二次元ハフ変換を行う事によって瞳の円を検出する。同時に瞳の半径 r 、及び瞳の中心座標 (X_{center}, Y_{center}) の情報を得ることができる。

二次元ハフ変換を用いた理由は、円の一部に隠れが生じた場合でも画像情報を大局的にとらえることにより検出でき、開眼状態に加え半瞬目状態といった瞳がまぶたに隠れている状態においても瞳の検出が可能であると考えられる点である。また計算時間・メモリ容量を削減するために、目領域をあらかじめ設定し探索範囲を絞り込む事に加え、抽出された目領域の画素値を二値化する事で余分なエッジ点群を排除した。図 4.3 に二次元ハフ変換により瞳の円検出を行う処理手順を示す。

4.3.4 エッジ処理

モノクロ画像の濃度差を利用したエッジ検出には低い周波数成分を除く微分フィルタを利用する。デジタル画像に対する微分は単純に濃度差だけで求められ、 (i, j) 点の微分

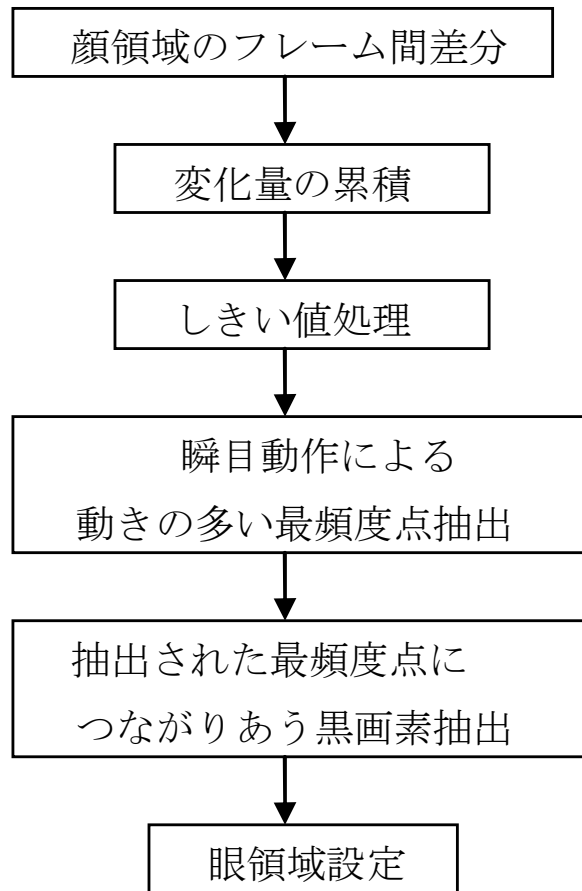


図 4.2 目領域設定処理の流れ図

値を求める式を 4.1 に示す .

$$\begin{aligned}\Delta x f &= f[i+1, j] - f[i-1, j] \\ \Delta y f &= f[i, j+1] - f[i, j-1]\end{aligned}\tag{4.1}$$

しかし, この式では 2 個の画素間の濃度差を用いており, 雑音に弱いという欠点があるため, Prewitt Edge Filter を使用した . これは X, Y 座標の明度変化を微分するオペレータである . 一次微分の空間フィルタは座標 (i, j) でのエッジ強度 $f(i, j)$ は 4.2 式となる .

$$\begin{aligned}f_x(i, j) &= -1 * f(i-1, j) + 0 * f(i, j) + f(i+1, j) \\ f_y(i, j) &= -1 * f(i, j-1) + 0 * f(i, j) + f(i, j+1)\end{aligned}\tag{4.2}$$

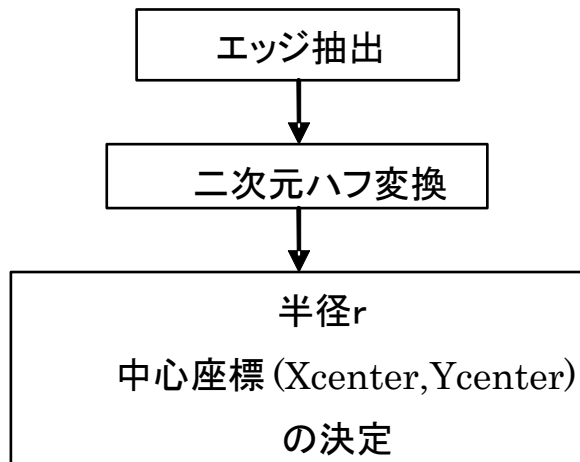


図 4.3 瞳部分検出の流れ図

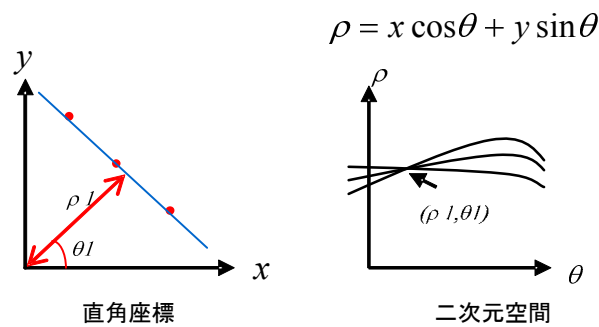


図 4.4 瞳部分検出の流れ図

4.3.5 Hough 変換

Hough 変換の概念図を図 4.4 に示す．たとえば，直線の検出の場合，元になる直角座標上の点 (x, y) を角度 θ と距離 ρ の二次元空間に変換し，角度 θ と距離 ρ 毎に，その個数をメモリ上に加算してゆく (Accumulator Cell などと呼ばれることがある)．個数が最大になった組み合わせ (角度 θ と距離 ρ の) を元の直角座標に戻したのが，最も直線らしい点の集まりである．個数を下げてゆくと，次の候補が順次得られる．角度 θ と距離 ρ を細かく分けると，精度が上がるが，計算時間が長くなり，メモリ容量も増えるのが欠点である．円の検出の場合には，元になる直角座標上の点 (x, y) を，円の中心点 (x_p, y_p) と半径 r の三次元空間に変換し，同様の処理を行う．

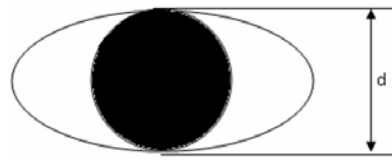


図 4.5 開眼状態

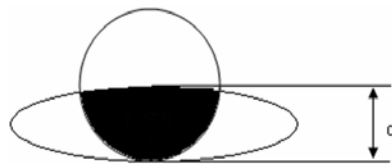


図 4.6 半瞬目状態

円の検出

半径が r で、中心点が $P(x_p, y_p)$ である円上に、座標上の点 (x_i, y_i) が存在するとすると、式 4.3 の関係が成り立つ。

$$(x_i - x_p)^2 + (y_i - y_p)^2 = r^2 \quad (4.3)$$

上の式を用いて、直角座標上の点 (x, y) を新しい三次元空間 (x_p, y_p, r) 上に変換すると、直角座標上の一点は $x_p - y_p - r$ 空間上の 1 枚の面に対応する。直角座標上の点が多数あると、 $x_p - y_p - r$ 空間上に多数の曲面が得られる。それらの曲面が共有する点があれば、それは元の $x - y$ 直角座標上では一個の円上に並ぶことになる。なお逆変換は、数式を用いるまでもなく、点 (x_p, y_p) を中心に半径 r の円を描けばよい。

4.3.6 閉目・半瞬目の判断

検出された瞳より、開眼状態・半瞬目状態・閉眼状態の目状態認識を行う。被験者の通常時の開眼幅 l として、100 フレーム間の目領域幅の最大値を取り、開眼幅 l と二次元ハフ変換によって求められた円の中心座標 (X_{center}, Y_{center}) によって得られる。 X_{center} 座標上における黒画素数をカウントする事で求めた瞳縦幅 d との比率により瞳の三状態 (開眼・半瞬目・閉眼) の判断を行った。検出すべき円と黒領域の関係を図 4.5 から図 4.7 に示す。また、開眼度 $Open_{eye}$ を式 4.4 として定義し、50 超過のときを開眼、20 以上 50 以下のときを半瞬目、20 未満のときを閉眼と判別した。

$$Open_{eye} = \frac{d}{l} * 100 \quad (4.4)$$

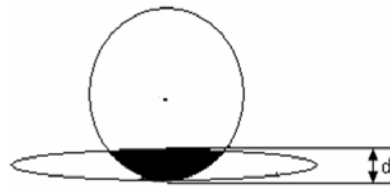


図 4.7 閉眼状態



図 4.8 AD 変換装置 (CEDmicro1401)

半瞬目状態の判断については通常の瞬目による瞬時的な半瞬目状態と眠気が生じていると推定される連続的な半瞬目状態との区別をつけるために、0.5 秒以上 (5 フレーム以上) 半瞬目状態が続いた際に半瞬目状態とした。

4.4 評価実験

4.4.1 実験条件

脳波計測実験機器

生理的指標を得るために脳波の検出を行う。脳波変化脳波計測システムとして AD 変換装置 (図 4.8), 日本光電社製脳波計 (図 4.9), 計算機 (図 4.10) を用いる。被験者にエレクトロキャップ (図 4.11) を取り付け右耳, 左耳, A1, A2, Cz, O1, O2, P3, P4, Pz (図 4.12) の電極を用いて, 電位変化の測定を行う。

CV 実験機器

本研究は表 4.2 に示す計算機にて開発, 実験を行った。



図 4.9 日本光電社製脳波計 (EEG-4418)



図 4.10 PC(PowerMacintosh8100AV)

4.4.2 脳波計測実験結果

心的操作に伴って発生する電位変化である事象関連電位の計測実験を行った。被験者に対するタスクとしてオドボール課題 (1 秒毎に発生する音を低音と高音かに聞き分け, 出現頻度の低い高音の出現回数を記憶する課題) を 30 分の事前練習の後, 15 分を 3 セット・13 分を 1 セット課した。各セットの間隔は約 10 秒である。各セットの高音時 (以下, target) 及び低音時 (以下, 「non-target」) の反応に対する脳波を比較することで, 生理的指標として用いる。結果をタスクの最初の 15 分間の結果を図 4.13 に, 最後の 13 分間の結果を図 4.14 に示す。



図 4.11 エレクトロキャップ

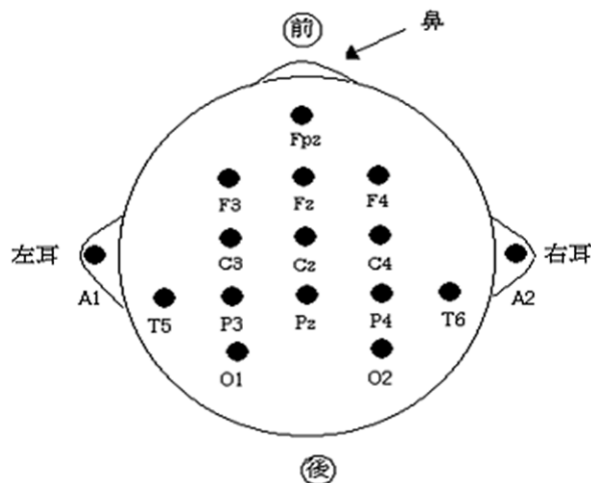


図 4.12 エレクトロキャップ電極位置

4.4.3 CV 実験結果

予備実験により被験者の瞳の半径は 10~11 ピクセルであったため、二次元ハフ変換のパラメータ空間である半径の値を 8 から 13 ピクセルとし評価実験を行った。時間ごとの状態変化との関連性を関連付けるために脳波データを取る実験と並行して行った。最初の 1 セット目 (15 分) と最後の 4 セット目 (13 分) の実験結果に対してビデオの映像を目視で、開眼状態・半瞬目状態・閉眼状態の三状態に分類し、実験後システムが判断した三状態と照らし合わせることで評価を行った。各実験の処理速度は 10FPS(Frame Per Second) であった。

表 4.2 評価実験に用いた計算機の性能

CPU	Intel Xeon 3.6GHz
RAM	2G
OS	Windows XP Professional
Graphic	Matrox Millennium G550
Capture	Matrox Meteor
Camera	Sony DCR-TRV70



図 4.13 タスク最初 15 分間における脳波解析結果

4.4.4 全瞬目認識評価

閉眼に対する性能の評価を行った．タスクの最初の 15 分間の結果を表 4.3 に，最後の 13 分間の結果を表 4.4 に示す．ここで表の $x1$ は目視による瞬目のイベント数， $y1$ は CV

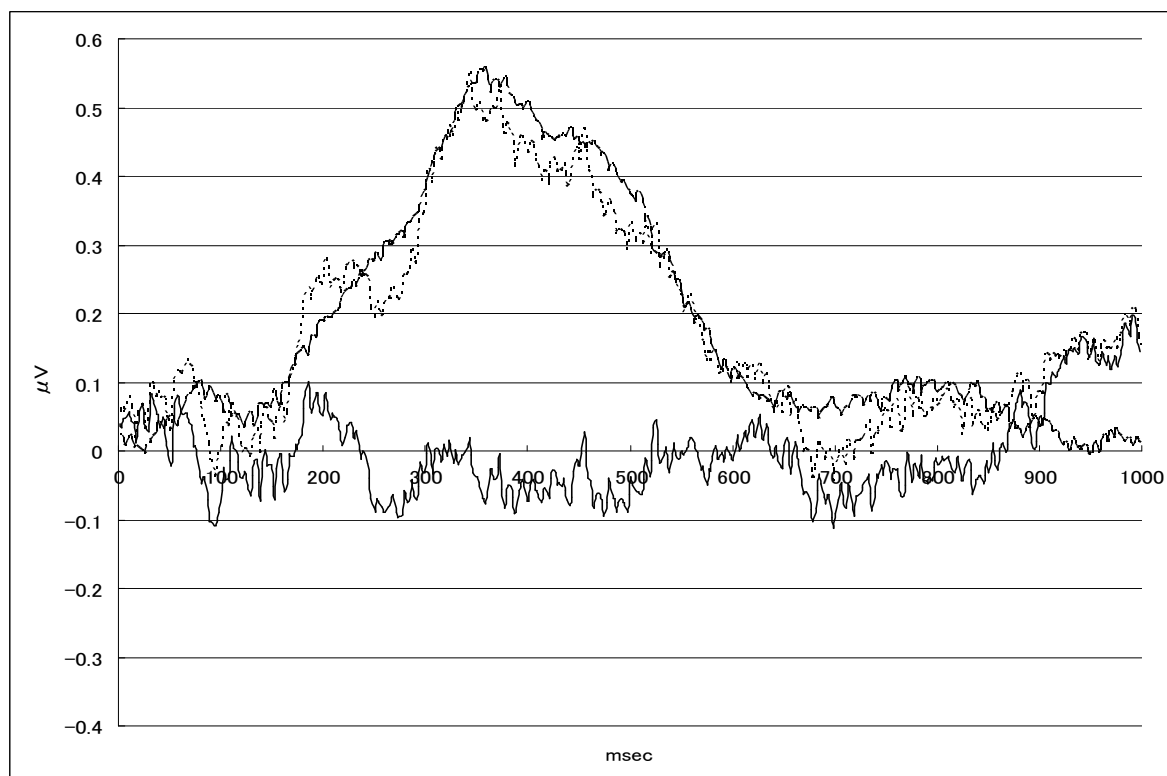


図 4.14 タスク最後 13 分間における脳波解析結果

表 4.3 タスク最初 15 分間における結果 (全瞬目)

Frame	x1	y1	z1	r1
0-2300	201	199	2	41
2301-4600	240	236	4	40
4601-end	224	199	25	34
Sum	665	634	31	115

および目視両方で瞬目と判断されたイベント数, $z1$ は目視で閉眼だったが CV では認識されなかったイベント数, $r1$ は目視では瞬目・半瞬目状態だったが, CV において誤認識で瞬目とされたイベント数を示す. 認識率 ($\frac{y1}{x1}$) は最初 15 分間が 95.3%, 最後 13 分が 82.1% となった. またそれぞれの FRR, FAR は最初 15 分間が 4.7%, 17.3%, 最後 13 分が 19.6%, 18.4% となった.

表 4.4 タスク最後 13 分間における結果 (全瞬目)

Frame	x1	y1	z1	r1
0-2300	270	237	43	45
2301-4600	217	156	61	42
4601-end	84	76	8	18
sum	571	469	297	105

表 4.5 タスク最初 15 分間における結果 (半瞬目)

Frame	x2	y2	z2	r2
0-3000	9	1	8	4
3001-6000	14	3	11	13
6001-end	7	1	6	5
sum	30	5	25	22

表 4.6 タスク最後 13 分間における結果 (半瞬目)

Frame	x2	y2	z2	r2
0-3000	9	5	4	13
3001-6000	30	25	12	15
6001-end	30	15	15	2
sum	69	45	31	30

4.4.5 半瞬目状態認識評価

瞬目と同様に半瞬目に対する認識評価を行った。タスクの最初の 15 分間の結果を表 4.5 に、最後 13 分間の結果を表 4.6 に示す。ここで表の x_2 は目視による半瞬目のイベント数、 y_2 は CV および目視両方で半瞬目と判断されたイベント数、 z_2 は目視で半瞬目だったが CV では認識されなかったイベント数、 r_2 は目視で開眼・半瞬目状態だったが CV において誤認識で半瞬目とされたイベント数を示す。認識率 ($\frac{y_2}{x_2}$) は最初 15 分間が 16.66%、最後 13 分が 65.22% となった。またそれぞれの FRR, FAR は最初 15 分間が 83.33%, 73.33%, 最後 13 分が 44.92%, 43.47% となった。

4.5 考察

4.5.1 脳波との比較

図 4.13 の脳波解析結果より，タスク最初 15 分間音が鳴ったと認識する波形の変化が音の発生後 100msec ~ 200msec の区間に見られる．その後，target と non-target の判断を行い，target と判断すると脳内に回数をカウントするタスクに対応する特徴的な波形が音の発生後 300msec 付近に検出されるため，target と non-target の差分が大きくなる^{*1}．これより，被験者は音の聞き分けができていないと推測される．

図 4.14 の脳波解析結果では，target，non-target の波形とも音の発生後 100msec から 200msec までの区間に見られる音の認識動作に対する波形は顕著に現れているが，その後ほぼ同じ推移をたどっており，target と non-target の波形の差分に特徴的な部分が現れていない．このことから，被験者はタスクを行えない疲労状態・意識の虚脱状態であると推測される．全瞬目自動検出結果 (表 4.3，表 4.4)，半瞬目認識評価結果 (表 4.5，表 4.6) より，瞬目イベント数 (x1) からタスク最初の 15 分間では 665 回，タスク最後 13 分間では 571 回確認されている．半瞬目イベント数 (x2) からタスクの最初の 15 分間では 30 回，最後の 13 分間では 69 回確認されている．タスク最後の 13 分における 6001 フレーム以降を $\frac{5}{3}$ 倍し，15 分に換算すると瞬目は約 680 回，半瞬目は約 89 回であったと推測される．

瞬目のイベント数に変化は見られないが，最後 13 分では半瞬目状態が増加していることから，目がとろんとしている半瞬目の状態からの瞬目において，数回にわたる連続的なしばたきが多く含まれていると推測される．よって今回提案した瞬目・半瞬目の認識手法に，瞬目の周期を加えた状態の認識を考えていく必要があると考えられる．

4.5.2 CV 性能について

閉眼認識評価結果 (表 4.3 と表 4.4) より，全閉の場合は検出率は始めは非常に高いが，最後は半分程度に低下している．始めは FAR は多いが，最後は減っている．始めは FRR は少ないが，最後は増えている．以上の 3 点が挙げられる．半瞬目認識評価結果 (表 4.5 と表 4.6) では，半閉の場合は検出率は始めは低いが，最後は良くなっている．FAR・FRR 共に，始めは高いが，最後は改善されている．以上の 3 点が挙げられる．開眼に対する検出成功例と失敗例を図 4.15，図 4.16 に示す．誤認識の代表的な原因として，大きく見開いた眼に対し，まぶたが小さくなってきた際に円周がまぶたのエッジに影響されて

^{*1} 耳に聞こえた音に対する反応の遅れを P300 と呼ぶ

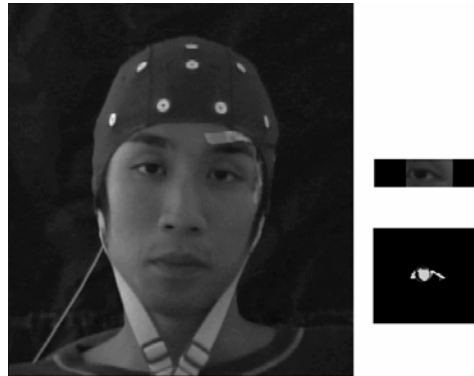


図 4.15 開眼時 (左:顔画像, 右上:目領域, 右下:円描画)

しまい、瞳部分が誤検出され判定結果に影響が出てしまっている。図 4.16 は開眼時に瞬目と判定された例である。

半瞬目に対する検出成功例と失敗例を図 4.17 と図 4.18 に示す。誤認識の代表的な原因として、図 4.18 に示すように目視において半瞬目と判定した場合において CV では閾値に達しておらず開眼と判断された例が挙げられる。目視における半瞬目の閾値による正確な判定が難しいためである。しかし、表 4.5 と表 4.6 における目視による半瞬目のイベント数 (X_2)・CV による半瞬目と判定されたイベント数 (Y_2) は、最初の 15 分と最後の 13 分では、共に大きく増加しており、疲労状態・意識の虚脱状態であると推測された脳波解析結果と一致するため、瞬目の検出に半瞬目の検出を加えることで人物の疲労状態推定に有効であると考察される。

瞬目に対する検出成功例と失敗例を図 4.19, 図 4.20, 図 4.21 に示す。誤認識の代表的な原因として、瞬目時は瞳がほとんど隠れているためにまぶたの影の影響され、円が誤検出されてしまう場合があったためである (図 4.20)。また図 4.21 は瞬目時の影の影響で黒領域の面積が大きくなり、開眼度 $Open_{eye}$ の閾値により半瞬目と判断されてしまった例である。これらの開眼・半瞬目・瞬目に対する誤認識の事例の対応策として、円の誤検出により一時的に円の中心座標に変化があった場合、誤検出の可能性があると前フレームの円の中心座標を考慮して次の円の候補を検索し誤検出を防ぎ、さらに候補が見当たらなければ処理をしないという対処が考えられる。

また二次元ハフ変換による加え、テンプレートマッチング手法を用いることによりまぶたの形状からの開眼・瞬目・半瞬目の 3 状態の推定を加える事で誤検出を防ぐ対処が考えられる。目視による半瞬目の状態認識の難易点の対応策については、半瞬目状態の判断として目視による判断だけではなく、さらに 波などの別の生理的指標を用いた判断を行う事が考えられる。

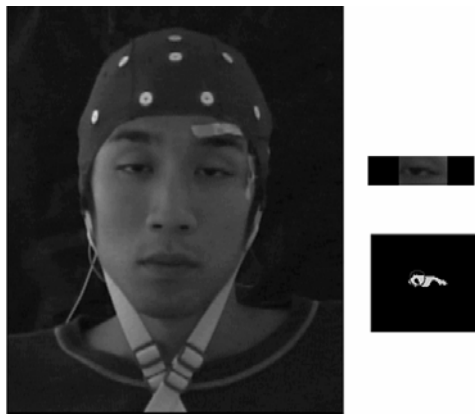


図 4.16 誤認識 1(左:顔画像, 右上:目領域, 右下:円描画)

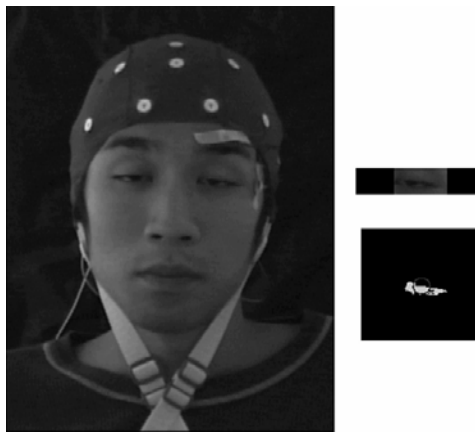


図 4.17 半瞬目時(左:顔画像, 右上:目領域, 右下:円描画)

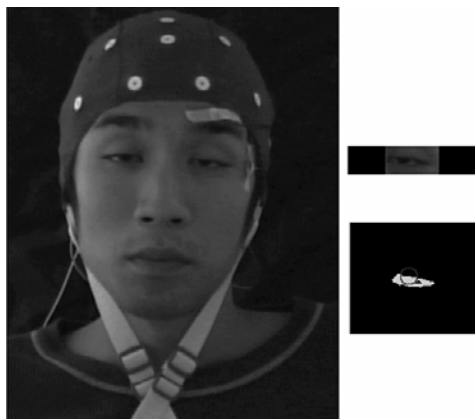


図 4.18 誤認識 2(左:顔画像, 右上:目領域, 右下:円描画)

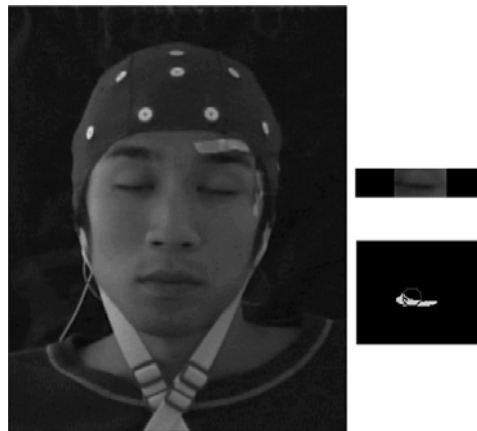


図 4.19 瞬目時 (左:顔画像, 右上:目領域, 右下:円描画)

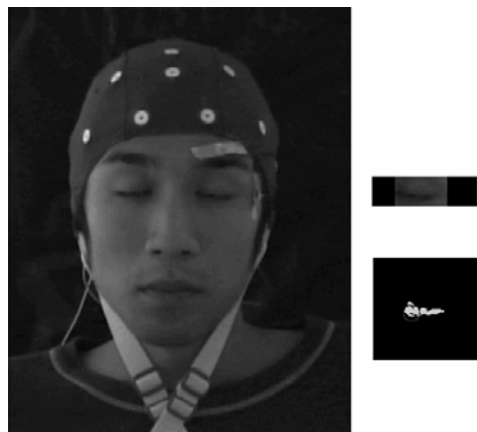


図 4.20 誤認識 3(左:顔画像, 右上:目領域, 右下:円描画)

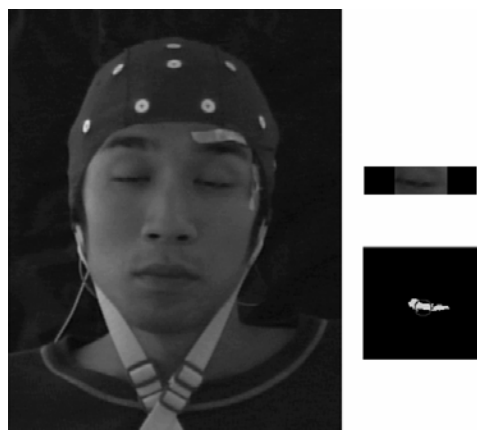


図 4.21 誤認識 4(左:顔画像, 右上:目領域, 右下:円描画)

第 5 章

総括

5.1 各エージェントにおける研究成果の要約

第 2 章では、人物追跡エージェントに関する研究成果を記述した。この研究では、テンプレート間距離やテンプレートデッド数の情報を用いて人物の移動状況を推定することで、人物の行動予測の最適化を行った。この人物移動状態推定は人物追跡に役立ったが、対象人物の行動軌跡や歩行速度等は人物同定エージェントにとって有益な情報と成り得る。

第 3 章では、人物同定エージェントに関する研究成果を記述した。人物を拘束しない状態で同定を行うために、複数の形状特徴や空間的習慣性を確率的手法を用いて、各人物の確率を算出して同定を行う手法を考案した。このエージェントでは、特徴を安定して取得するために背景差分を基礎とした画像処理を行っているため、複数の人物が進入したときの対応付け問題の対処が困難である。しかし、人物追跡エージェントの追跡情報を用いることで、その問題を比較的容易に解ける可能性がある。

第 4 章では、人物状態認識エージェントに関する研究成果を記述した。人物に精神的負荷をかけ、CV 処理にて目の開閉度を測定することにより、疲労時には半瞬目状態が増加することを発見した。人物の状態認識は一般に普遍的な特徴を用いるが、各人固有のくせというものも存在するため、その情報を用いることで新たなる状態認識が可能になると考えられる。しかし、その情報を用いる場合は、人物を同定する必要がある。

5.2 エージェント連携と親和的情報空間の構築

ここまでで紹介した研究成果は個々のエージェントに設定した目標を実現するために得られた成果であるが、改めて「情報空間内の人物状況認識」という当初の目標を実現させることを考えたとき、各エージェントで得られた情報はその他のエージェントに有用な情

報を持っており、エージェント連携を用いることが有用であると導かれる。各エージェントは前提条件に基づく制約から取得困難な情報が存在するが、その情報は他のエージェントでは比較的容易に取得できる情報もある。

これらのことを考慮すると、親和的情報空間を設置する場合に必要な物理的環境は図 5.1 に示したものとなる。部屋の隅には各エージェントを動かすための PC が必要である。一台の PC の処理能力にてリアルタイム処理が困難な場合は、クラスタリング処理にて対処するために複数台の PC が必要になる。但し、動画像データは情報量が多いため、複数台構成の場合は画像処理部と情報取得部の適切な配置が要求される。人物追跡エージェントは、雲台制御可能なカメラが必要である。人物同定エージェントは、画角の広い固定台カメラが必要である。人物状態認識エージェントは、対象人物に対して解像度が高いことが望ましいため、雲台制御ならびに望遠制御可能なカメラが必要である。

次に、情報空間内の人物状況認識における各エージェントの活動状態の一例を図 5.2 から図 5.4 に示す。図 5.2 では、人物追跡エージェントが雲台制御可能なカメラで対象者の追跡を行っている例である。このとき、追跡のために人物をテンプレート学習するが、そのテンプレートパターンは人物同定に有用と成りうる情報であるので、人物同定エージェントへ渡す。また、テンプレートの揺らぎが疲労度の高さを示す場合は、その情報を人物状態認識エージェントに送る。図 5.3 では、人物同定エージェントが固定カメラで侵入者の同定を行っている例である。人物追跡エージェントから対象者が着席したことが伝えられると、それまでに取得した情報群から、対象者が人物 A であることを同定する。図 5.4 では、人物状態認識エージェントが対象者の状態を推定する例である。まず、人物の疲労度を測定するためにカメラの視野を対象人物の位置に拡大する必要がある。そのため、空間内の位置情報を人物追跡エージェントから取得する。これによって対象者の目を計測することが可能になるため、開閉度を測定し半瞬目状態が増えたことから対象者の疲れが高いことを推定する。

これらの活動で人物へのサービスを考えるために必要な情報は、必要十分な量はあると考えられる。この後は実際にどのようなサービスが最も適しているかを推測し、それを実行する機能が必要となってくるが、それは親和的情報空間の「人物状況や知識に基づく適切なサービスの推測とその実施」が担当する部分であり、今後の研究に期待される。

5.3 まとめ

親和的情報空間を構築するための一要素である「情報空間内の人物状況認識」について、最低限必要と考えられる機能に細分化した。そして、その機能を実現させるエージェントごとに研究目標を定め、各目標ごとに評価実験を行い考察を行った。また、これらの研究成果からエージェント連携が「情報空間内の人物状況認識」の機能を高めると考察

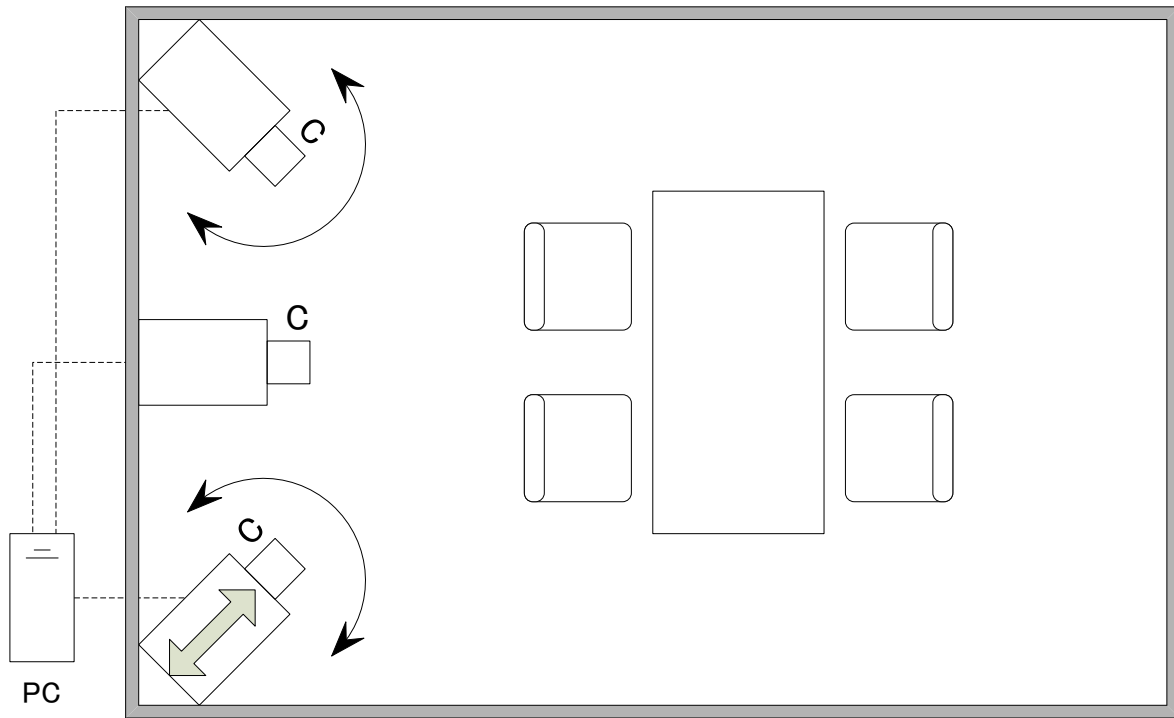
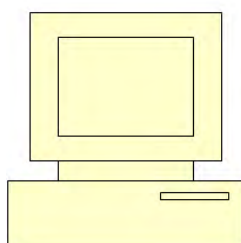


図 5.1 親和的情報空間の物理的配置図

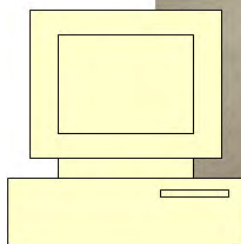
し、研究成果の知見に基づいた、親和的情報空間に必要な環境や動作時の一例を示した。

以上のことから、親和的情報空間における人物認識の研究は、一定の成果を得たと結論する。



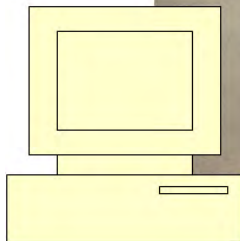
追跡	進入時にテンプレート学習し追跡
同定	テンプレートパターン(追跡)からAさんである確率が上昇
認識	テンプレートの揺らぎ(追跡)から対象者は疲労度が高いことを予想

図 5.2 人物追跡エージェント



追跡	着席を検知する
同定	着席位置(追跡)の人物から形状を取得、滞留習慣性(追跡)の情報から、Aさんである確率が最も高い
認識	待機中

図 5.3 人物同定エージェント



追跡	人物の空間位置を伝える
同定	待機中
認識	疲労度が高い(追跡)と虚ろ目が多いことから 疲労度がかなり高いことを推測

図 5.4 人物状態認識エージェント

付録 A

人物状態認識の応用研究 1

A.1 序論

A.1.1 研究の背景

文部科学省の掲げている大学改革の一つの傾向として、学生による授業評価が広まりつつある。本章では大学の授業評価に対する姿勢、多項目アンケート形式における学生による評価活動の問題点を挙げ、新しい観点で受講者の集中度を「客観的」に「リアルタイム」で時系列表示させる自動認識システムを提案する。教員は自らの豊富な専門知識や経験に基づいて授業を実施するものであるが、授業とは学生のために行われるものである。受講者が授業内容を実際にどのように受け止めたかを教員は常に把握しておく必要がある。それにより常に充実した質の良い講義を行い、受講者にフィードバックしていくことができる。この取り組みをファカルティ・ディベロップメント (Faculty Development) という。

A.1.2 ファカルティ・ディベロップメント

ファカルティ・ディベロップメント (以下 FD) とは、教員が授業内容・方法を改善し、向上させるための組織的な取組の総称である。FD を実施している大学は、年々増加しており、平成 13 年度 (2001 年度) の調査によれば 409 校の大学 (61%) で実施されている。具体的な例としては、新任教員のための研修会の開催、教員相互の授業参観の実施、センター等の設置などを挙げることができる。

現在、大学にあっては、自己点検・自己評価の一環として学生の授業評価が広く行われるようになってきた。この学生による授業評価は、授業態度や授業内容の理解度などに対する自己評価と教授法や教師の授業態度などに対する教授法評価という 2 つの授業評価内容で行われることが多い。教員は、この授業評価結果をもとに授業の改善を試みるものである。

「学生による授業評価」は、程度はさまざまではあるが、文部科学省高等教育局の調査によれば、2001 年度現在、513 大学（76%）が何らかの授業評価を実施している。近年、注目されているテーマは、その結果を使ってどのように授業を改善してゆくか、ということにある。授業評価の結果を改革に反映するための組織的取り組みが行われていると答えたのは、国立 48 大学では 50%、公立 14 大学では 30%、私立 120 大学では 32% となっている。組織開発のためには、教員のみならず職員の開発・育成も大事である。FD の取り組み方は各教育機関により様々であるが、このように FD が近年いかに注目されているかがわかる。

A.1.3 学生による授業評価の問題点

学生による授業評価は、現状では学生の多項目アンケートにより調査をおこなっている。このことにより、授業の実像を明らかにし、その改善につなげるための指標として、授業の受け手である学生の意識を把握できるという点では、たいへん有用である。しかし、以下の問題点が考えられる。

- 授業を数値で評価するのは困難
- 学生評価の客観性
- 評価が教育技術面に偏っている

これらの問題点からアンケートの判断だけではなく、さらに複眼的な講義状況把握が必要である。

A.1.4 授業評価システムの従来研究

従来手法では、授業評価アンケートをWEB 上で行い、集計が容易にできるシステムが主流である。この方法であれば、毎時間アンケートを行えば時間当たりの授業評価の把握が容易になり、その結果に対する学生へのフィードバックがその場で可能であるが、学生が評価者という面においては客観性に欠ける。

「複数の講義室俯瞰画像を用いた受講者注目対象の推定」という研究では、顔領域の中において、受講者の顔向きによる視線方向により判断している。この手法では授業の評価としては顔向きによる注目対象のみを対象とするために、下を向いてメモと取る行動などは、講義を聞いていないと判断されてしまう可能性がある。

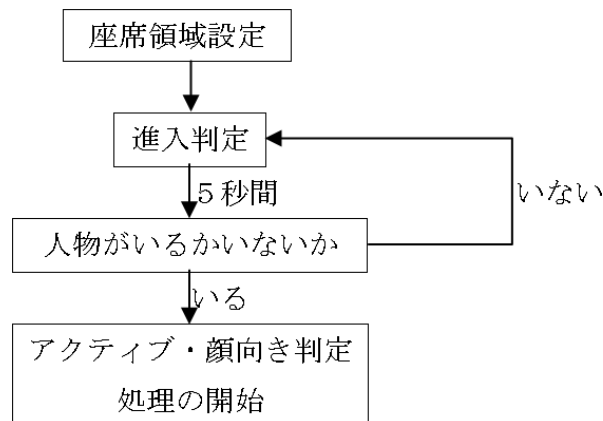


図 A.1 着席判定におけるフローチャート

A.2 提案手法の説明

A.2.1 概要

本提案手法では顔向きによる注目方向だけではなく、時系列変化における動きを導入した。カメラにとらえた人物の動きと顔の向きの2点を総合的に判断することにより、受講者を非拘束的にさらにリアルタイムで状況の把握が可能になる。

A.2.2 着席と顔向きの判定

あらかじめ座席領域を設定しておき、背景差分によって領域内の人物の進入を検知する。通過者との誤認知を防ぐために、侵入者が5秒間にわたってその場に存在しているとき、その席に着席したものととして座席領域内のターゲットのアクティブ度や顔向き判定の処理を開始する。図 A.1 にフローチャートを示す。

着席後の受講者の動きによるアクティブ判定は、人物領域内の時系列画素変化を用いた。10 フレームごとの毎フレームの総和が閾値に達したとき、ターゲットに何らかの動きがあったとしてアクティブと判定する。図 A.2 に処理の流れを示す。

また、ターゲットの半数以上が同時にアクティブであると判定された時には、教科書のページをめくったなど、講義において何らかの大きな動きがあったものと予想できる。そこで、その動きに対応できなかった者は授業を聞いていなかったと判定する。図 A.3 にフローチャートを示す。

明度ヒストグラムとは、対象者領域の明度分布状態を示したものである。顔画像中で顔部分が大きく映し出されている場合、肌色部分が大きく出るため、ヒストグラムは数値の

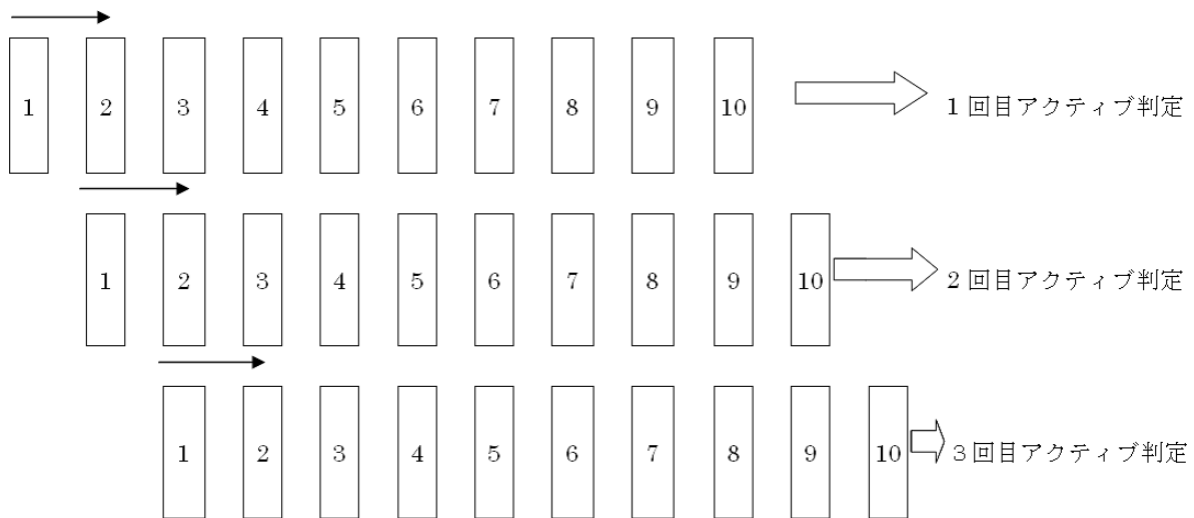


図 A.2 アクティブ判定の処理の流れ

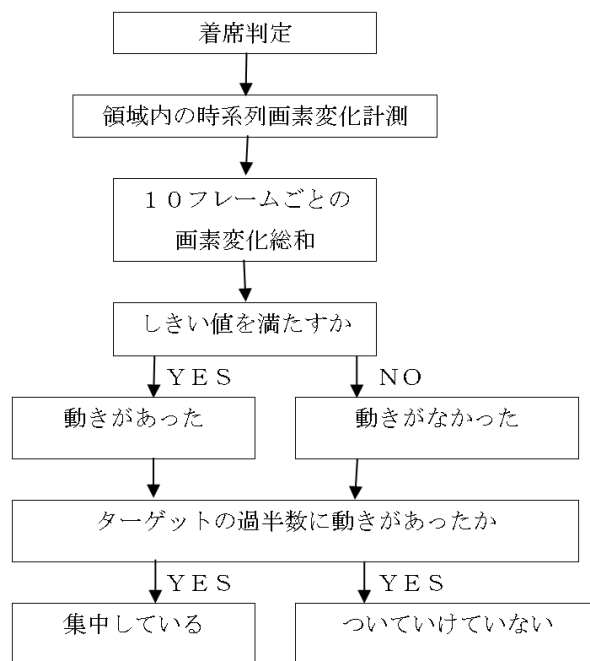


図 A.3 アクティブ判定フローチャート

低い値に集中する．逆に下を向いていると髪が大きく映し出されるため，ヒストグラムの値は比較的大きい値に集中する．

着席時には人物は前を向いていると仮定し，人物が座席に着席したという判定後，時系列差分の変化が減少したときに人物が着席したと判断する．そこから差分が大きく変化するまで，明度をヒストグラムとして保存する．各階調ごとに最大値と最小値を計測することで，顔領域のデータとした．ヒストグラムの学習終了後，入力画像から得られた顔領域

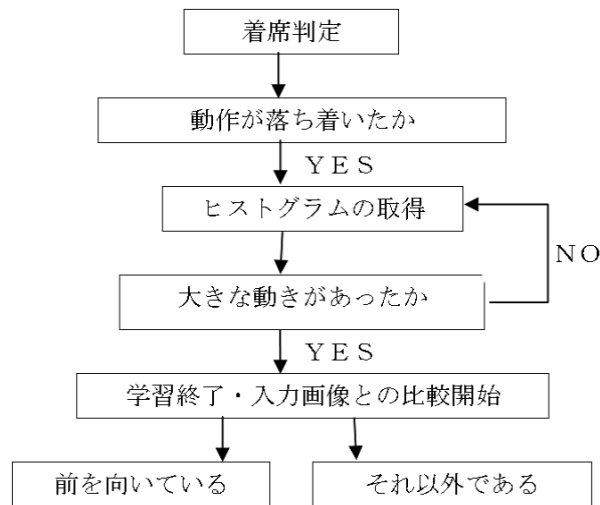


図 A.4 顔向き判定処理の流れ

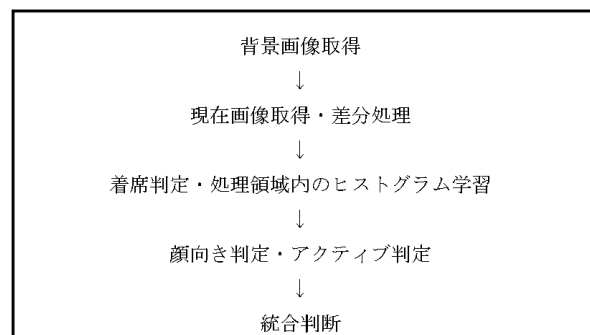


図 A.5 受講者集中力評価システム処理の流れ

のヒストグラムと学習時に作成したによって得られたデータとの比較により、顔向きの判定を行った。判定は前を向いているか、それ以外の 2 種類である。図 A.4 に処理の流れを示す。

人物の動きと顔向きの 2 つの判定結果を元に統合判定を行い、人物がアクティブであるか否かの判断を行う。これを用いて、受講者の集中力を評価する。処理の流れを図 A.5 に示す。

A.2.3 統合判断

ヒストグラムによる顔向きに加え、人物の動作によるアクティブ判定ターゲットの人物の統合判断を行う。判断項目は集中している・寝ている・状態未知の 3 項目である。統合判断方法を表 A.1 に示す。

ここで判断項目に、“それ以外”という項目を付け加えた。この項目がない場合、判断が

表 A.1 統合判断方法

		大	小
顔向き	前	集中している	
	後	状態未知	寝ている

表 A.2 実験環境

PC	CPU Pentium4-1.7G mem 512M
カメラ	SONY Handycam DCR-TRV30
グラフィック	Matrox Millennium G450
キャプチャー	Matrox Meteor
OS	Microsoft Windows 2000
開発言語	Micorosoft VisualC ++ 6.0

閾値に近いときにどちらかの項目を選択するしかなく，誤検出の要因となり得る．よって，判断不可能状態も選択肢として入れることで，信頼性の向上を図った．

A.3 評価実験

A.3.1 実験環境

本研究では，Microsoft Visual C++ 6.0 を利用して，アプリケーションのシステムの開発を行った．鹿児島大学情報工学科棟 72 号講義室天井に DV (SONY Handycam DCR-TRV30) を設置した (図 A.6)．収録された画像を Matrox 社のキャプチャーボードを経由して映像を取り込む (図 A.7)．画像処理を行ったシステムの性能を表 A.2 に示す．

A.3.2 実験結果

実験において，渡邊睦助教授の人工知能基礎の講義を用いた．実験対象者はカメラ画像に写っている 17 名のうち，前の座席の人物に隠れてしまう人物を省き 12 名を選択した (図 A.8)．以下に着席判定・人物動作判定・顔向き判定及び統合判定の実験結果を説明していく．

着席判定の処理例

図 A.9 に人物 1 が着席する一連の動作を示す。着席後しばらくして着席判定が行われ、ターゲット 1 への動作判定がスムーズに行われた。

人物動作判定

各座席において着席判定が行われた後における人物動作判定の実験結果を図 A.10 に示す。各ターゲットごとに動きがあったときは赤・動きが検出されなかったときは青で表示し、ターゲットの過半数に動きが検出されたときには緑で表示をさせた。

この結果から 3 番と 7 番のターゲットは全体的に動作が少なく、過半数の受講者の動作が検出されても動きがないため、この 3 番と 7 番のターゲットは講義についていっていないと判断できる。しかし 11 番のターゲットも全体的には動作は少ないが、過半数の受講者の動きにはしっかりとした反応が見られるため、寝ているとはいえない。

人物顔向き判定

以下にターゲットが着席しヒストグラムの学習・顔向き判定を行う一連の動作を示す(図 A.11)。着席判定後、動作がおちつくまで学習を待機し、次の大きな動作があるまでヒストグラムの学習を繰り返し、前を向いている時のヒストグラムの学習を繰り返す。

ヒストグラム学習後の前を向いている基準となるヒストグラムの範囲は、上から 2 番目ヒストグラムであり、それに対する現在画像のヒストグラムとの比較により、顔向き判定を行っている。前を向いていると判定した場合には顔部分に白い四角の枠を表示させた。顔向き判定においては、処理領域を顔が大きくはみ出したときに失敗したが、それ以外は良好な結果を得ることができた。

A.4 考察

この統合判定処理を 60 分連続して処理を行ったところ、以下の結果を得た。ターゲット 1 (図 A.12) の場合、NON-ACTIVE が 12% に対し ACTIVE が 54% と大部分の状態でアクティブと判定された。また状態未知判定が 34% となっているが、これは顔向きは上ではないが、動作が検出されたことを意味するため、ノートをとるなどの動作も含まれていると考えられる。つまりこのターゲットは集中していたと推測できる。一方、ターゲット 2 (図 A.13) は NON-ACTIVE が 55% と大多数を占めている。この結果からこのターゲットは寝ている、もしくは講義についていけなかったと推測できる。

図 A.14 に 12 人全員における統合結果を示す。ACTIVE が 40%、状態未知が 34% という結果を得た。NON-ACTIVE が 26% という結果が出ているが、今実験で用いた講義

は PowerPoint を用いた，プレゼン形式の講義であるため，受講者が前を向いて教師の話聞き，それに対するメモ等をとっていると考察できる．ACTIVE 及び状態未知を加えると 7 割が講義に関連した動作と考えられるため，この講義は大多数の受講者を引きつけていたと評価する．今後の課題として，この状態未知を更に細分化し，より適切な結果が得られるようにしたい．

A.5 結論

本研究は受講者の集中度を客観的にリアルタイムで時系列表示させる自動認識システムの構築を行った．明度ヒストグラムによる顔向き判別に加え，人物の動きによるアクティブ判定を加えることで，人物の状態判別への精度・信頼性が向上したと考える．

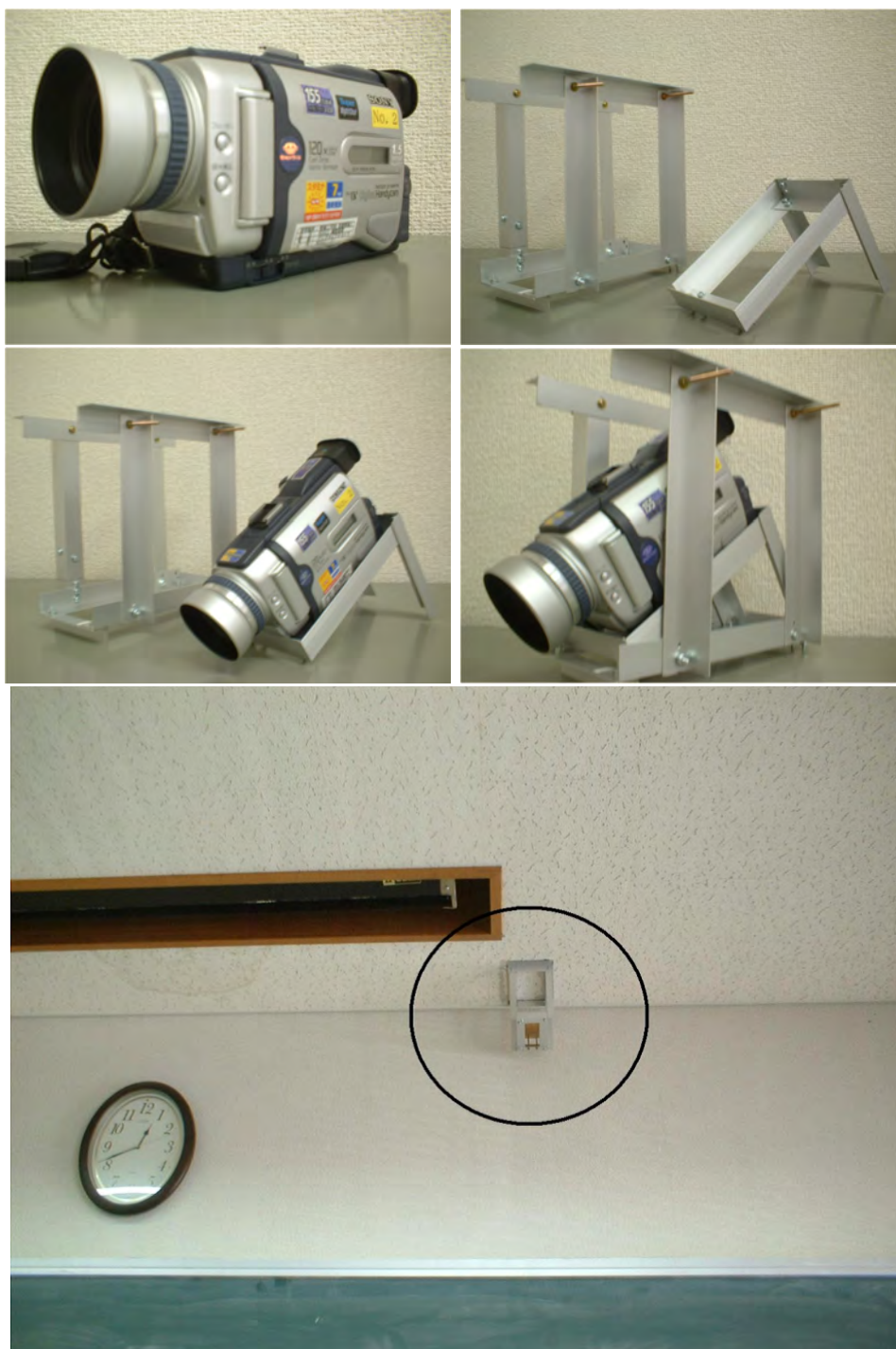


図 A.6 DV カメラの設置



図 A.7 DV 出力画像 (鹿児島大学情報工学科棟 72 号講義室)



図 A.8 選択した 12 名

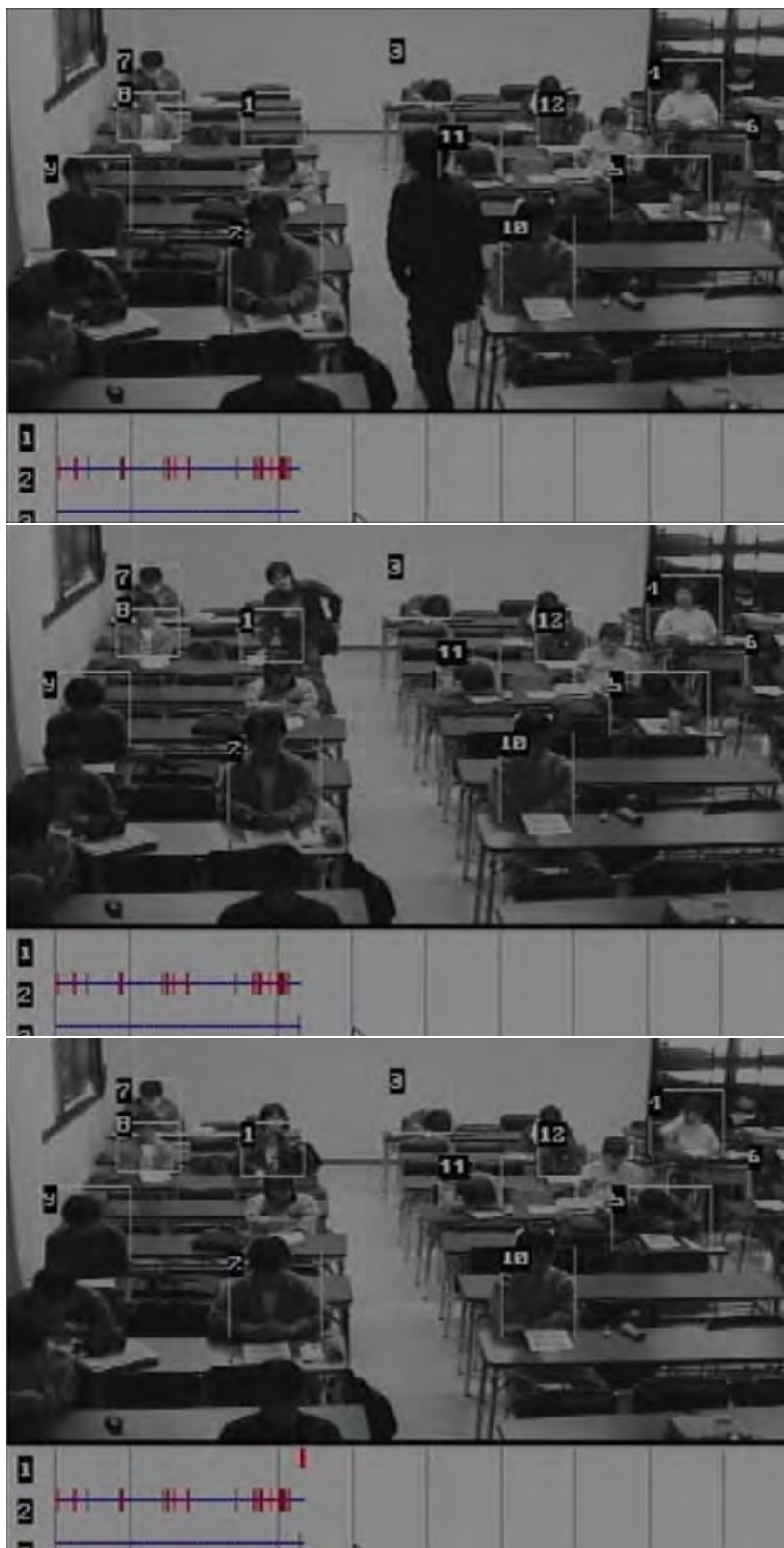


図 A.9 着席判定の処理例

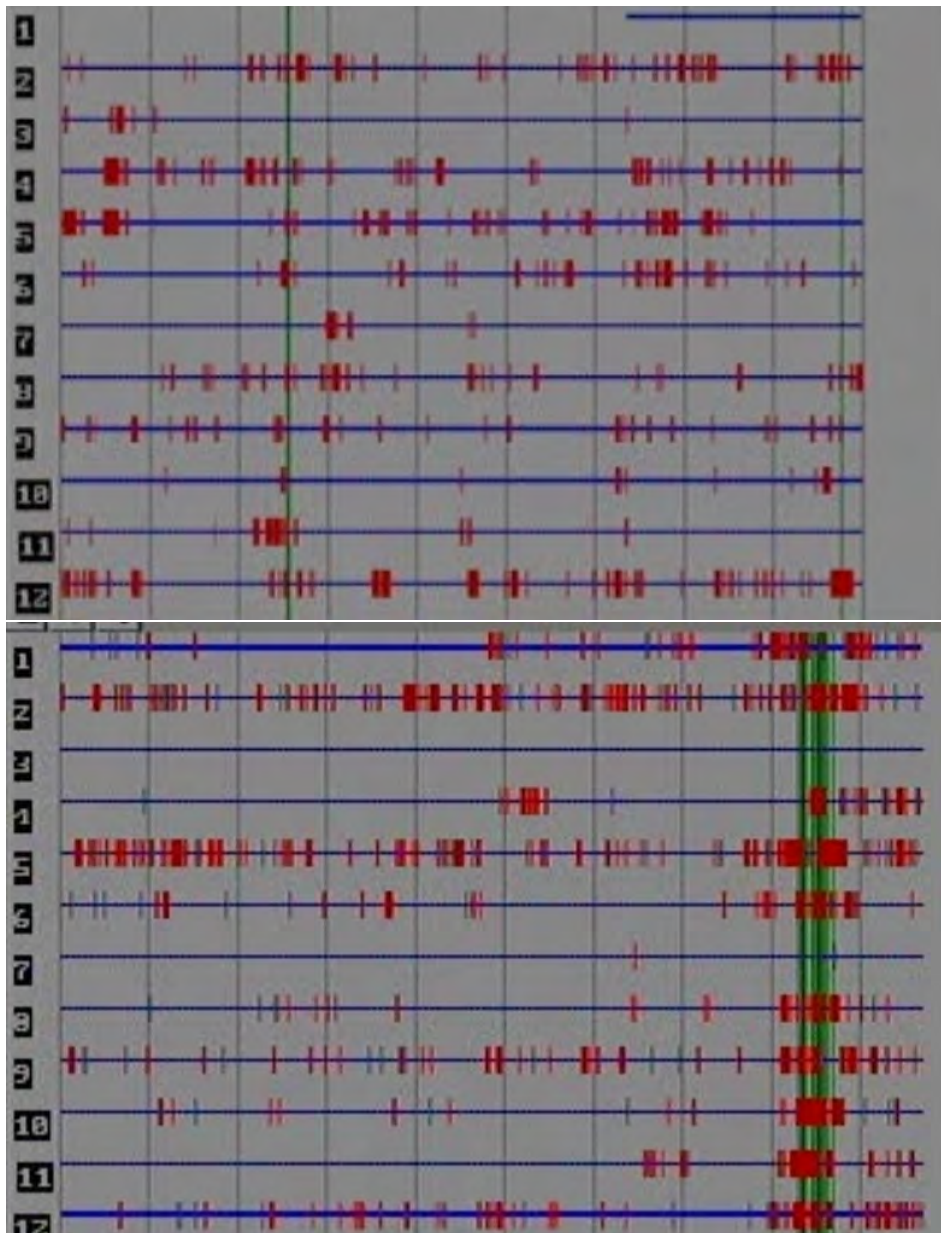


図 A.10 人物動作判定

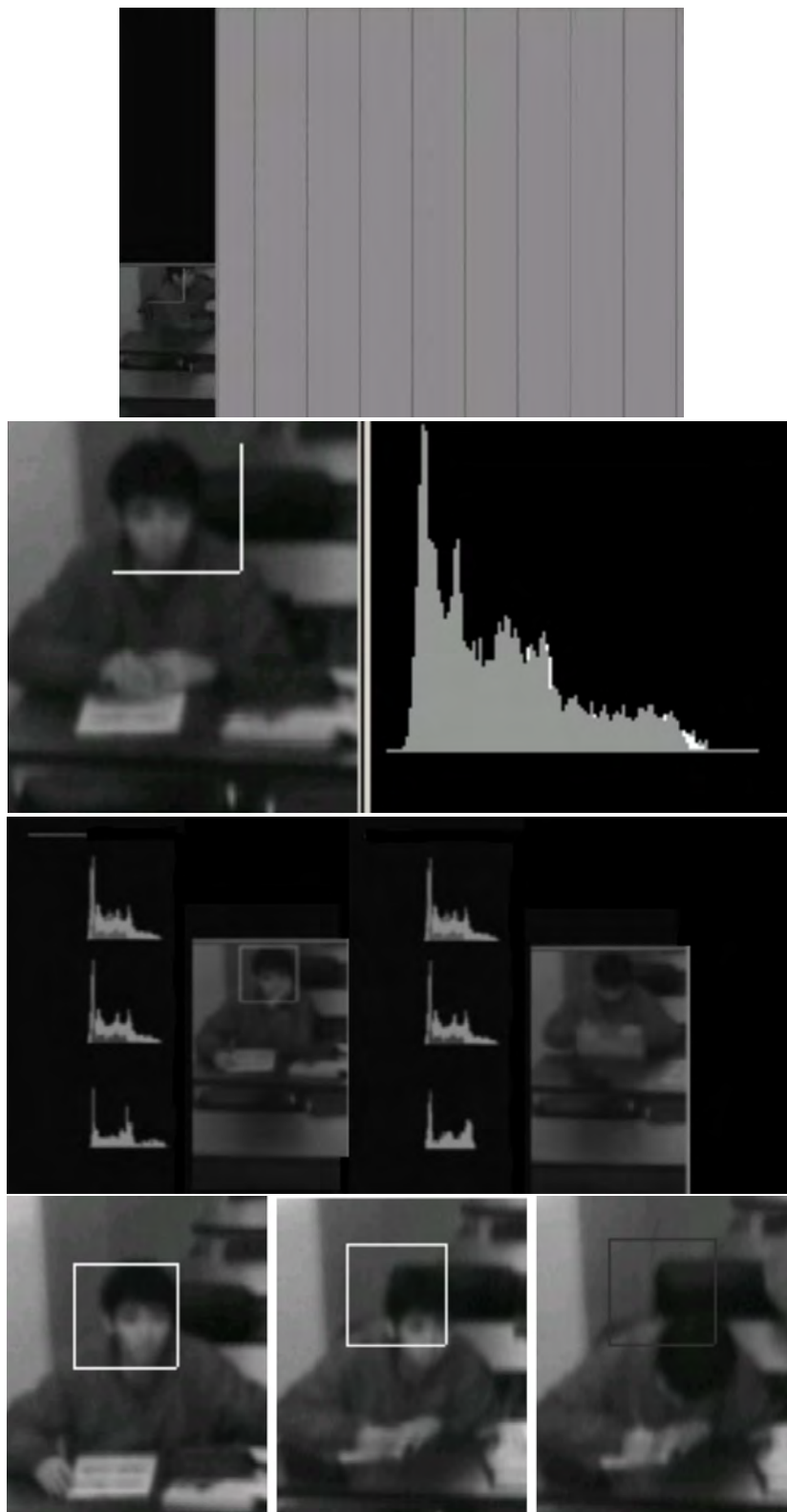


図 A.11 人物顔向き判定

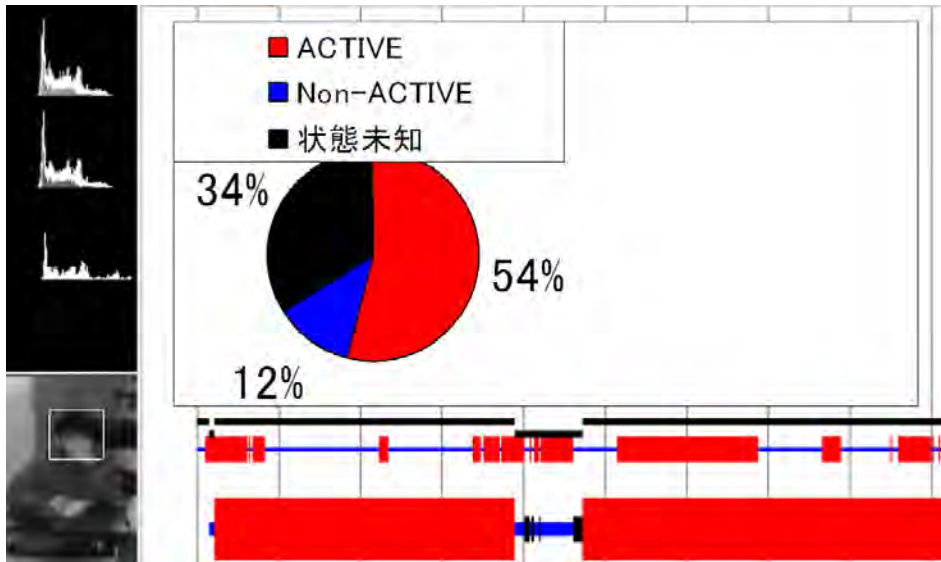


図 A.12 ターゲット 1 の統合判定結果

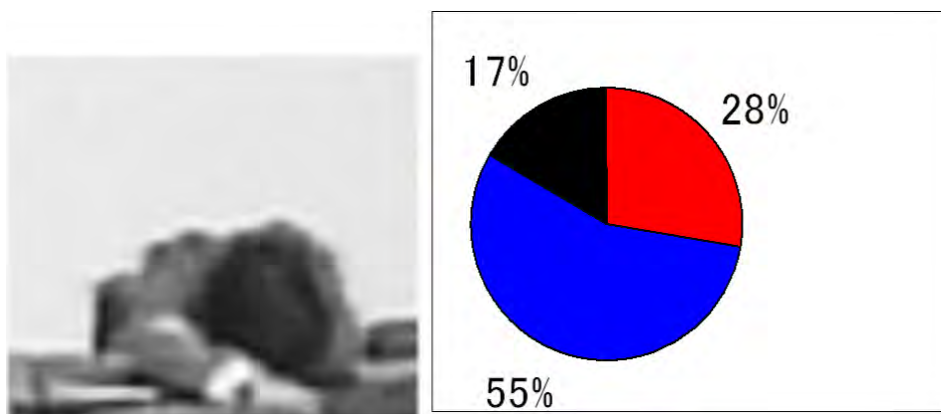


図 A.13 ターゲット 2 の統合判定結果

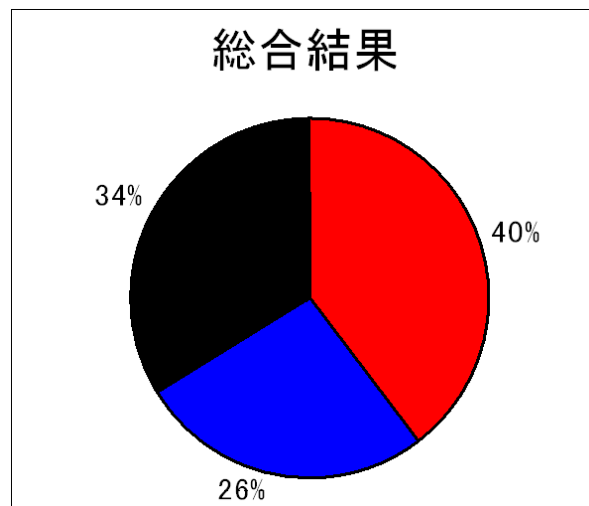


図 A.14 ターゲット全員の統合判断処理結果

付録 B

人物状態認識の応用研究 2

B.1 研究の背景と目的

近年，犯罪率が増加の傾向にあるといわれている．その一例として本の万引きに目を向けた場合，悲惨な状況にある．店舗側はこの被害に対して監視カメラをつけるなど対策をとっているが，従来のカメラの機能は追跡と状態記憶となっているため抑止効果はあるものの，最終的には人が記憶した映像をみて一連の動作を確認せねばならない．店員が監視カメラの映像を常時確認する，記憶した映像を日々から見直すということは，日常業務をこなしながらでは難しく，決定的瞬間をとらえていたとしても意味をなさない場合もある．

そこで本研究では，監視用アクティブカメラを用いて人物の追跡を行い，その後不審（万引き）と思われる行動をとったかを認識することにより，不審行動検出を行う．人物が不審行動をしたことを認識し，検出時にはアラーム等警報を出すことで，人的作業を軽減し，セキュリティコストを減少させることを目的とする．

B.2 提案手法の説明

B.2.1 概要

図 B.1 に全体の処理の流れを示す．ただし，「アクティブカメラを用いて人物を追跡」から「静止を確認したら画像をズーム」の処理は，本論文の 2 章の技術を用いているため，該当部分の説明は省略する．

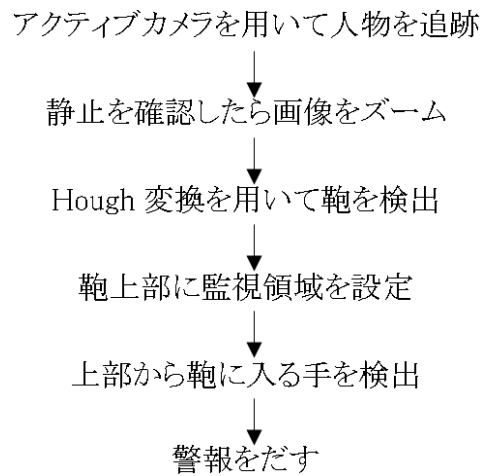


図 B.1 全体の処理の流れ

B.2.2 大津の二値化による自動閾値選定

Hough 変換時に処理の高速化をはかるため，前段階処理として画像を二値化する．全体で比較的最適となる 1 つの閾値を用いると，場所などの変化に対応できないため，大津の二値化 [40] を用いる．その手法は，濃度ヒストグラムの 0 次モーメントと 1 次モーメントの累積モーメントのみを用いた，簡単な構造で汎用性を持った閾値選定法である．

最適閾値 k は絶対分離点と呼ばれる確立 0 レベルか，閾値を境に分離される 2 クラスの平均レベルの midpoint となる．しかし，絶対分離点はその直前の正のレベルを閾値とすることであるから，最適閾値の条件からはずされる．画素を背景クラスに属するもの S_1 と対象クラスに属するもの S_2 に分類し，それぞれの濃度ヒストグラムの 0 次モーメントと 1 次モーメントの累積量から各クラスの平均レベル μ_1, μ_2 を求める．それぞれ i レベルの正規化ヒストグラム $i * P_i$ と各クラスの生起確率 ω_1, ω_2 を用いて，式 B.1 と式 B.2 で表される．更にこれらの式の平均より，最適閾値は $\frac{k=(\mu_1+\mu_2)}{2}$ となる．

$$\mu_1 = \sum_{i \in S_1} \frac{i * P_i}{\omega_1} \quad (\text{B.1})$$

$$\mu_2 = \sum_{i \in S_2} \frac{i * P_i}{\omega_2} \quad (\text{B.2})$$

B.2.3 エッジ抽出

連続性が求められる輪郭抽出には，Canny Edge Detector の使用が考えられる．しかし，その手法は計算に時間がかかり，動画処理の目的によっては適さないことがある．そこで，本提案手法ではエッジ抽出方法として Roberts Filter[41] を用いる．Roberts Filter は輪郭抽出に用いられる 1 次微分 (グラディエント) の基本式 (式 B.3 と式 B.4) を改良したものである．

$$f_x = f(x+1, y) - f(x, y) \quad (\text{B.3})$$

$$f_y = f(x, y+1) - f(x, y) \quad (\text{B.4})$$

グラディエントは簡単な構造をしているが，輪郭の中でもエッジの抽出が得意という性質がある．Roberts Filter は式 B.5 と式 B.6 で示される．

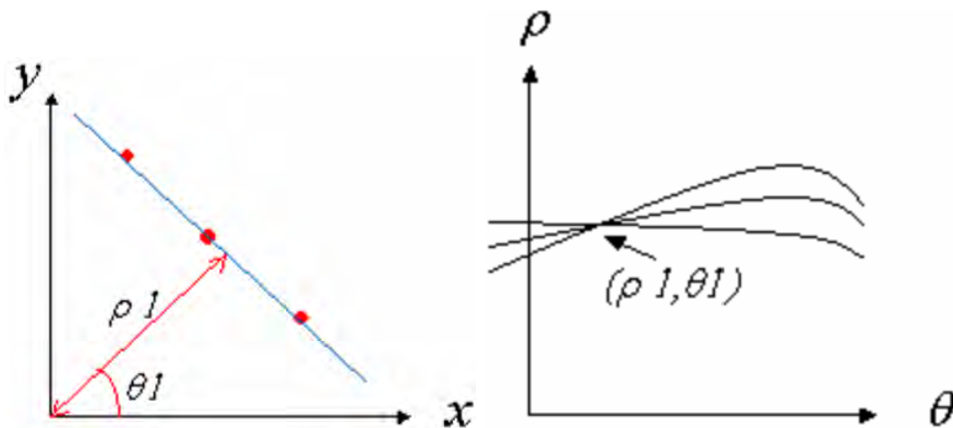
$$f_x = f(x, y) - f(x+1, y+1) \quad (\text{B.5})$$

$$f_y = f(x+1, y) - f(x, y+1) \quad (\text{B.6})$$

B.2.4 Hough 変換による靫検出

エッジ抽出された画像をもとに靫となる線を検出する．一般的に 1 次の直線の式は $y = ax + b$ と表現される．この式は傾きが垂直になると a が無限大になるため，計算が困難である．そこで， ρ, θ をパラメータとする $\rho = x \cos \theta + y \sin \theta$ を利用する． $x - y$ 空間での 1 点を通る直線は無限に存在するが，それぞれに対し ρ, θ は一義的に決まる．これを利用し， $\rho - \theta$ パラメータ空間において，曲線の交差する点が曲線に対応するそれぞれの点を共通に通る直線として表すことができる．図 B.2 のように， $\rho - \theta$ パラメータ空間に変換することを Hough 変換という．

ただし，Hough 変換は投票と多数決の原理によって直線や円を検出するという方法である．多少のエッジ抽出の失敗に耐え，直線等を検出することが可能であるが，角度制限はない．対象物がありながら，同方向のみにしか直線を得られないということが考えられるため，垂直方向・水平方向への角度制限を設ける．このとき， $0 \text{ or } \pi, \frac{\pi}{2} \text{ or } \frac{3\pi}{2}$ に限定するのではなく，多少の揺れに対処できるように幅をもたせる．また，制限角度内に直線が存在した場合でも，ほぼ同位置に存在する場合は先に検出されたものを優先させるものとし，後に検出されたものは無効とする．4 つの交点が出た場合のみ，重心を計算する．この重心をもとに監視領域を設定する．また，カメラのズーム後は人物の大きな移動はないものとして考え，この後の処理速度向上のために Hough 変換は停止する．

図 B.2 $x-y$ 平面と $\rho-\theta$ 平面

B.2.5 監視領域内処理

不審行動をしたかどうかを判断するため、鞆上部の監視領域内を処理する。本や CD、ビデオテープなどの四角の物を持ち、さらにそれを鞆の中に入れるということを不審行動の定義とする。四角の物を持っていない時や距離が遠い時は、注意すべき状態だとする。まず、アクティブカメラによる追跡処理画像から、四角の物があるかを判断するエッジ抽出画面と下向きの動きがあるかを判断する時間差分画面を作成する。このとき、手や物の距離が遠いとき”ATTENTION”を、四角の物が検出できていないとき”MOVING”を画面上に表示させる。図 B.3 に示すように、エッジ抽出画像のある y 座標の領域において、最左端と最右端をとる。長さでエッジの占有率で正規化を行い、四角形があるかどうかを判断する。時間差分画面では上下を比較し下向きの変化率をとり、最大となる部分を検出し、動きを判断する。下向きの動きがあり、位置が鞆に近く、エッジ抽出画面で四角の物を検出されている時には”CAUTION”を表示させる。

B.3 評価実験

B.3.1 実験環境

表 B.1 に示す PC 環境で実験を行った。実験場所は鹿児島大学工学部情報棟 2 階実験室である。

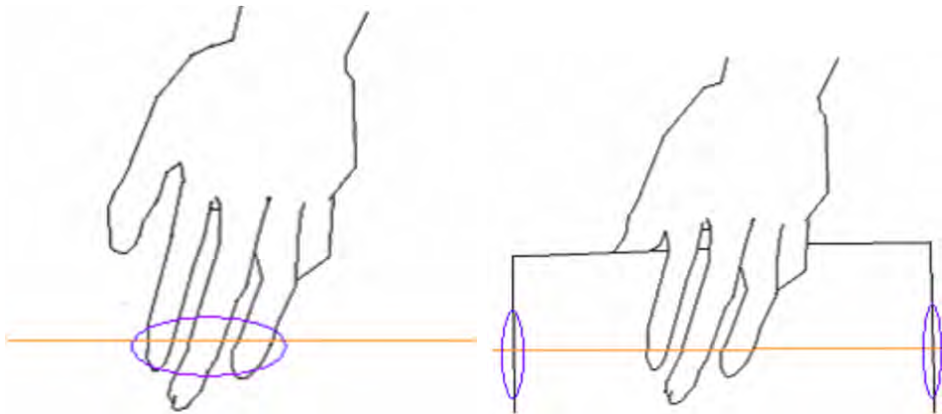


図 B.3 監視領域の設定方法

表 B.1 実験環境

PC	CPU Pentium4-3.0G mem 1.0GB
カメラ	Canon VC-C4R
グラフィック	Matrox Millennium G450
キャプチャー	Matrox Meteor
OS	Microsoft Windows XP
開発言語	Micorosoft VisualC ++ 6.0

B.3.2 実験結果

人物追跡から警報をならすまでの一連の処理の様子を示す．処理の様子を図 B.4 から図 B.8 に示す．1 から 6 までは人物追跡部，7 から 16 は対象人物を拡大して撮影した画像，7' ~ 16' はその画像に対して不審行動判定の処理を示している．7 から 10 の部分は Hough 変換による鞆位置検出を行い，11 ~ 16 は状況判定を行っている．11 から 13 が何も持っていない時の状態を示し，14 から 16 が物を持っている時の状態を示している．物を持っていない時に”MOVING”を表示し，距離が遠いときに”ATTENTION”を表示し，物を持っていて距離も近い時には”CAUTION”を表示している．

B.3.3 他形状鞆を用いた実験

前節で用いた四角形状の鞆以外に Hough 変換を用いて鞆検出が可能であるか実験を行った．その実験結果を図 B.9 と図 B.10 に示す．この実験では 2 種類の鞆で実験を行っ

た．一つめの鞆では鞆位置を正確にとらえることができたが，二つめの鞆では実際より少し大きめにとらえていた．ただし，監視領域の自動設定という観点でみた場合，監視領域を設定するうえで致命的な問題にはならないと考えられる．

B.4 考察

処理にかかる時間は Hough 変換時には平均 4.5FPS，監視領域処理時は平均 12.5FPS であった．Hough 変換時はフレームレートが低く，実時間で不審行動を監視するためには十分な速度であるとはいいがたい．しかし，鞆位置を決定した後 Hough 変換を終了させることで実用的な速度になると考えられる．今後の改良点として，今回は不審行動検知に時系列差分変化を用いたが，動きがより詳しく把握できるオプティカルフローの導入が考えられる．

B.5 結論

一般的な鞆に対して，鞆位置検出と不審行動の検知を行うシステムを開発し，一定の成果が得られた．

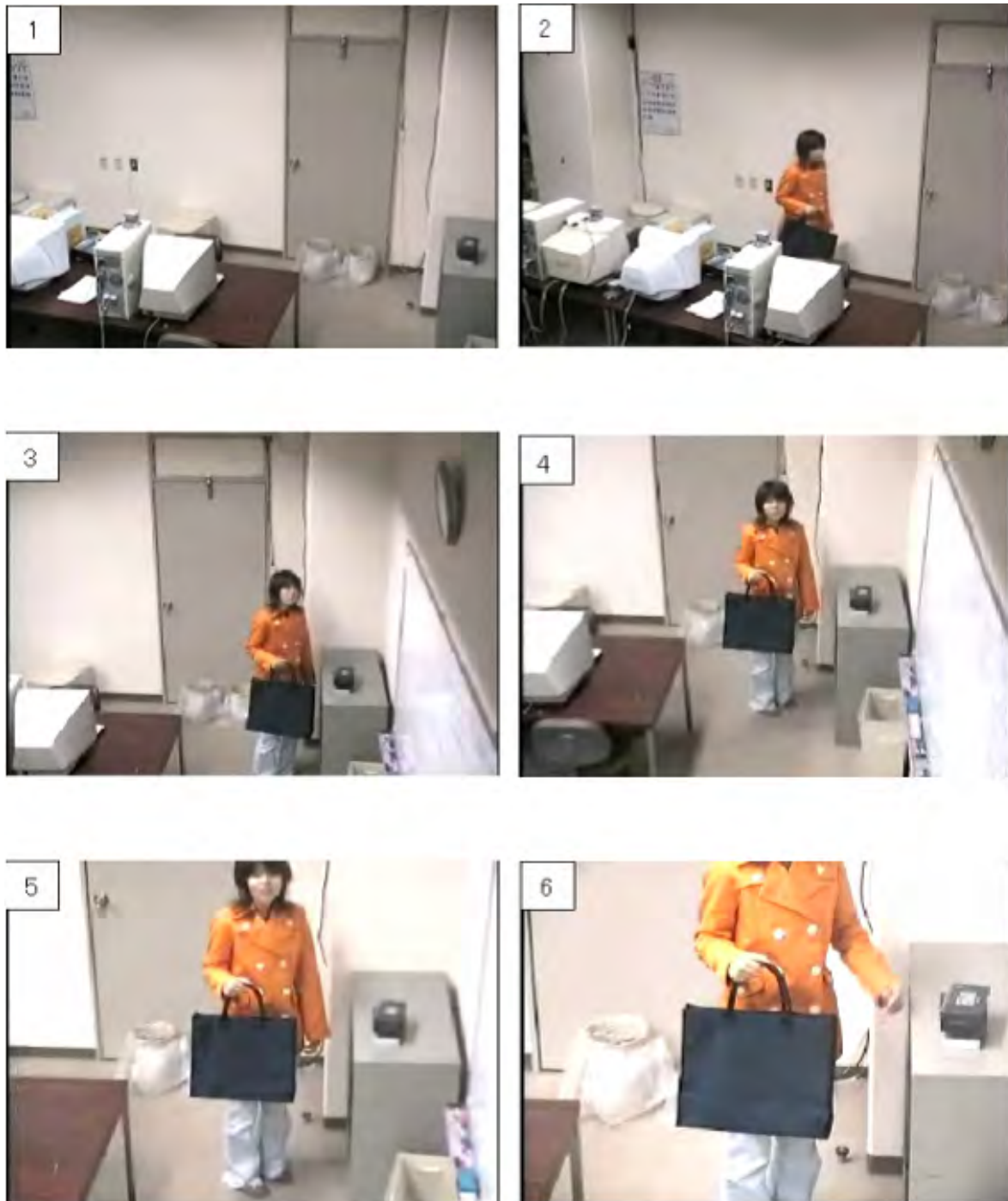


図 B.4 人物追跡の処理例

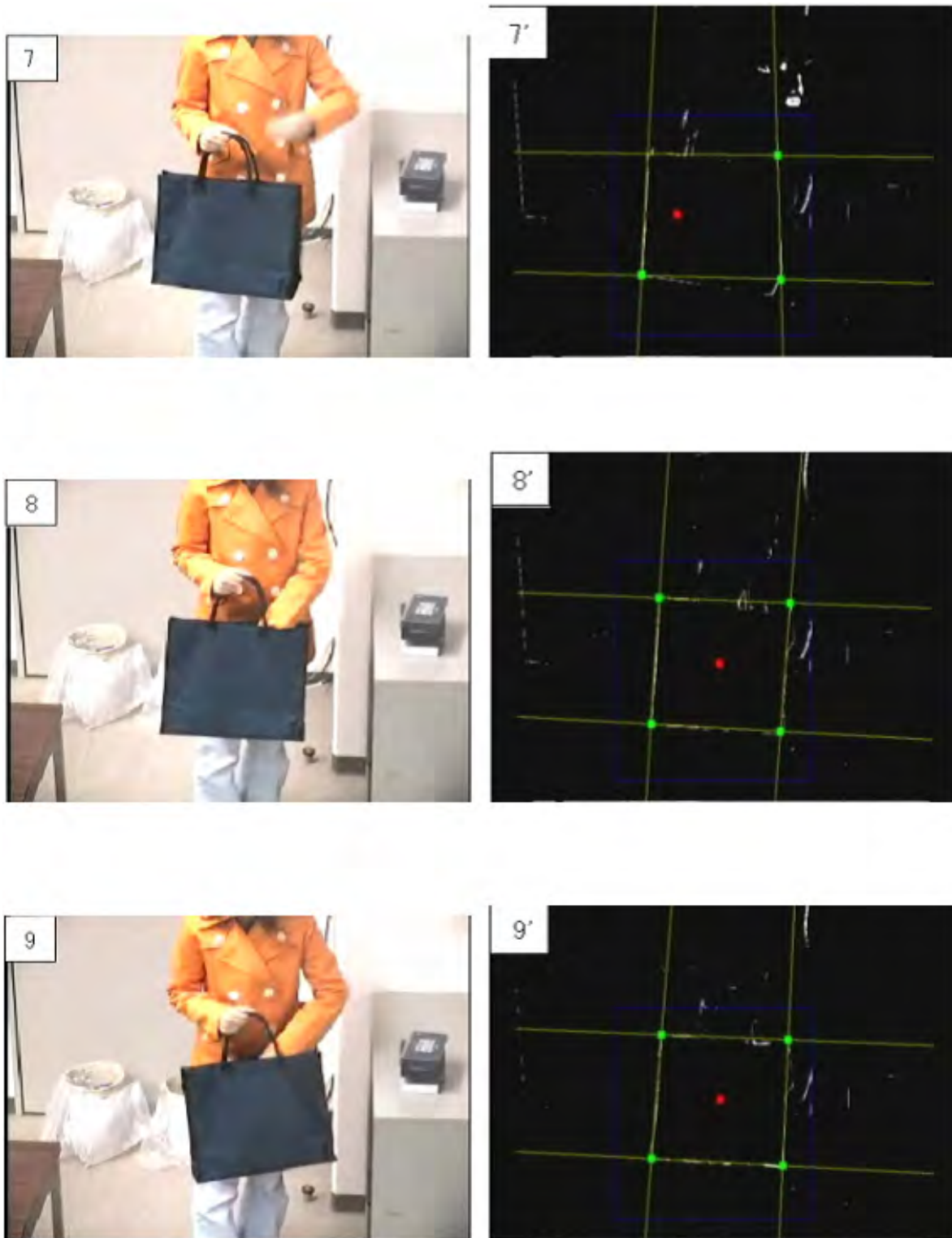


図 B.5 不審行動判定の処理例 1

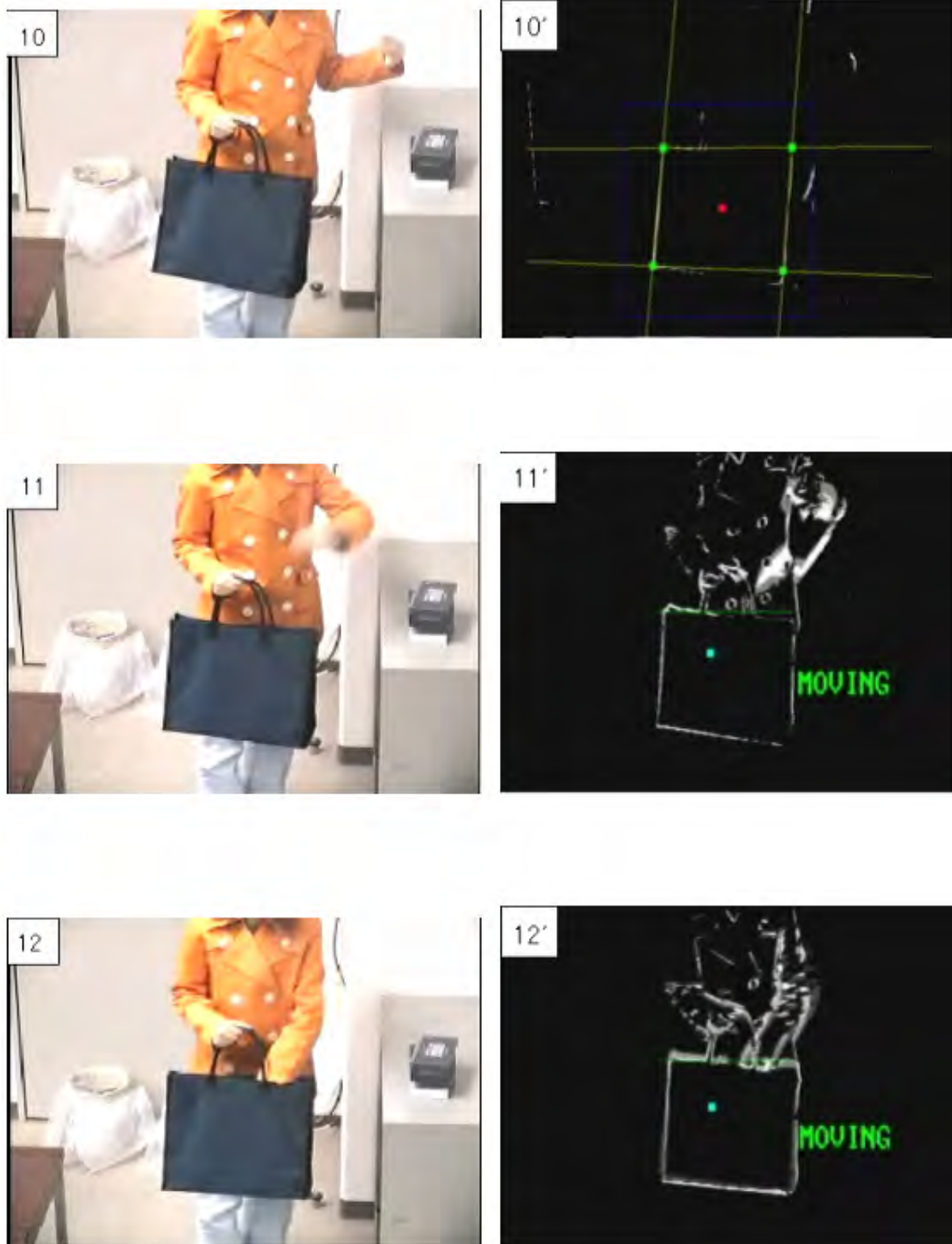


図 B.6 不審行動判定の処理例 2

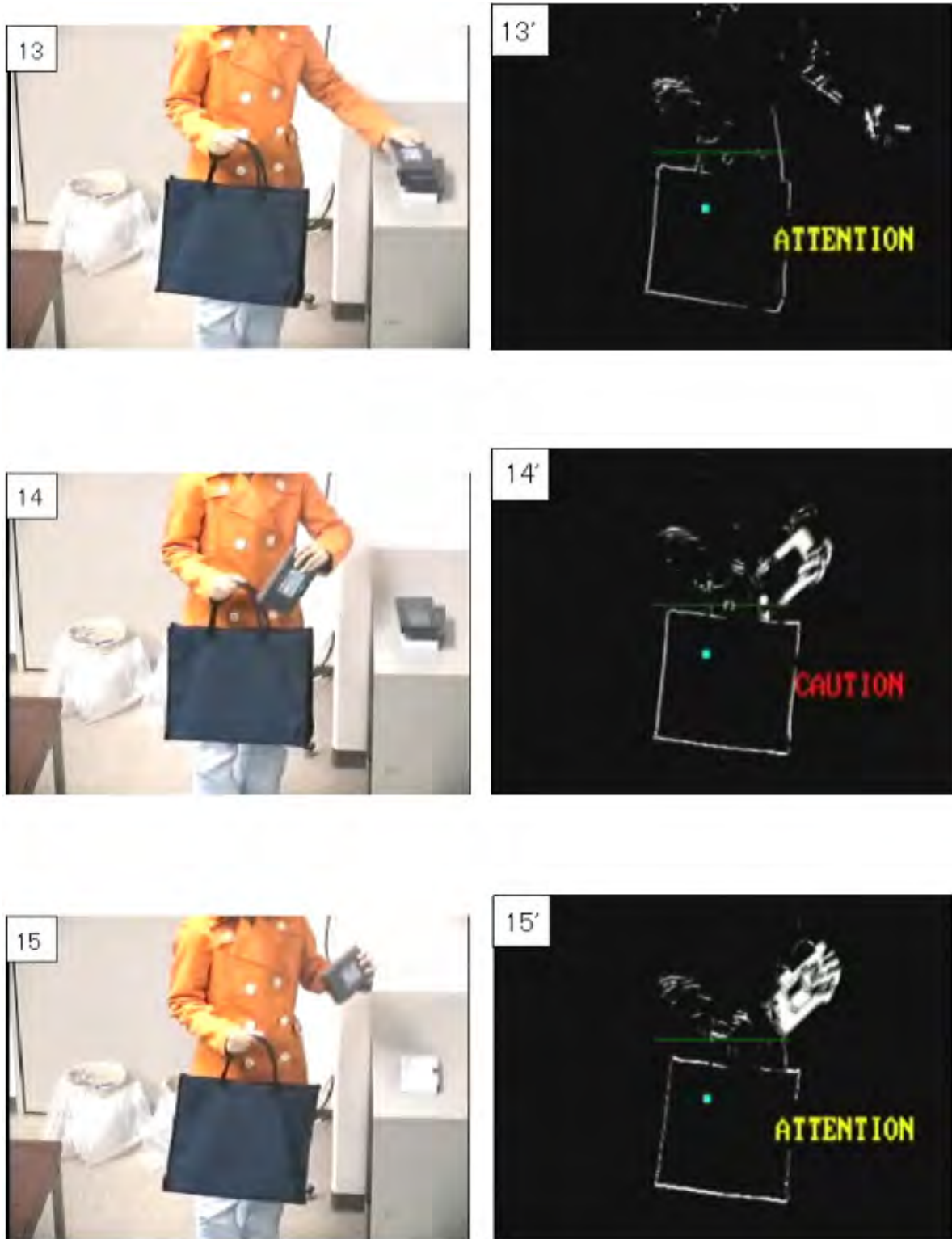


図 B.7 不審行動判定の処理例 3



図 B.8 不審行動判定の処理例 4

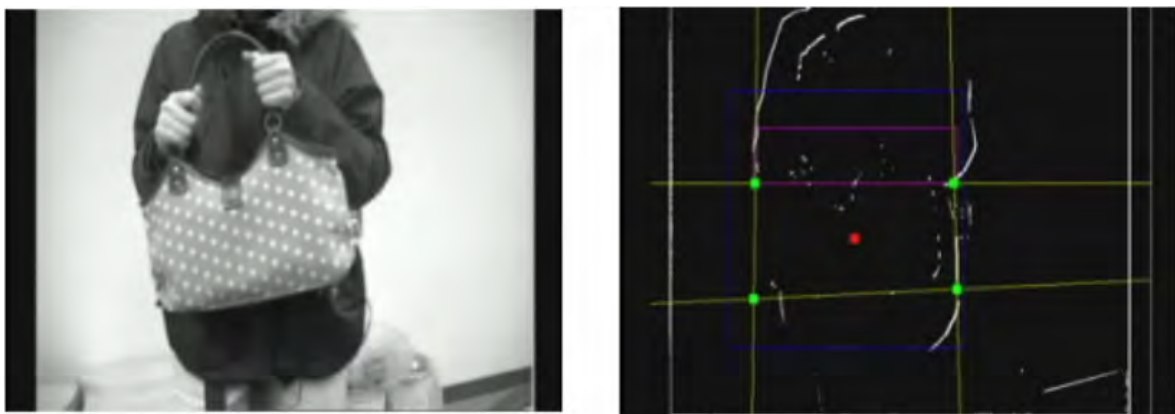


図 B.9 その他の鞆検出例 1

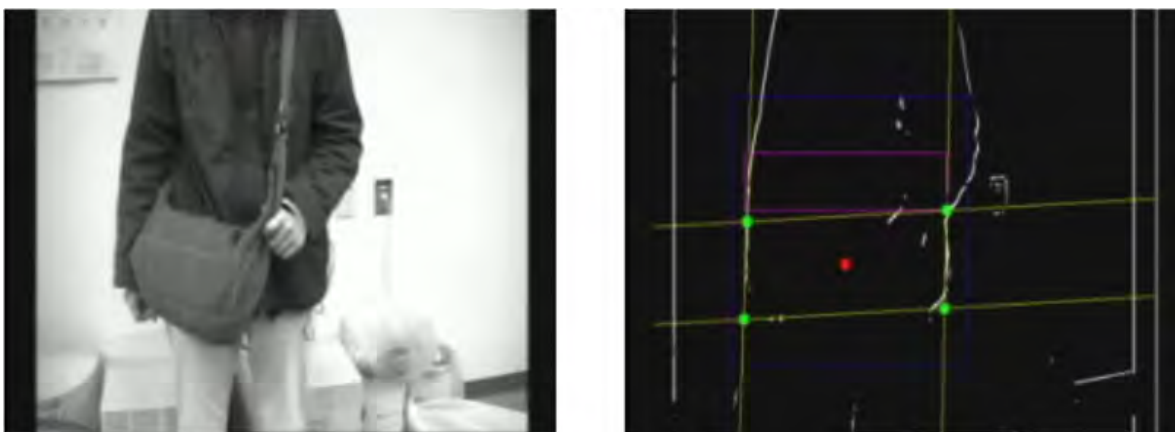


図 B.10 その他の鞆検出例 2

付録 C

動画像処理システム

C.1 始めに

本書では MILMIL シリーズの最終作品である MILMIL-forever (以下 MILMIL) の使用を前提に記述する。MILMIL は Visual Studio 2005 で開発したことにより、使用者は同じ開発環境を利用する必要がある。また、動画像ライブラリ MIL-Lite 8.0 の使用を前提としているため、開発機にそのライブラリがインストールされている必要がある。

使用者は第 C.2 節にて開発環境の利用方法を知る必要がある。動画像処理の実装方法は第 C.3 節に書かれている。MILMIL を詳しく知りたい場合は第 C.4 節を参照すること。第 C.5 節では MILMIL シリーズの開発の歴史について書かれている。

C.2 開発環境について

C.2.1 Visual Studio 2005 について

Visual Studio 2005 とは Microsoft 製の統合開発環境である。この開発環境では基本言語として Visual Basic 等も選択できるが、MILMIL シリーズは処理の高速性を求めることから VC++ を選択している。

C.2.2 開発環境のカスタマイズ

Visual Studio 2005 は開発環境のカスタマイズを行うことが出来る。基本的に選択した言語によって、窓の配置やショートカットキーが変更される。本書での開発環境の図やショートカットキーは、筆者が最適だと思われる設定を前提として書いている。この設定はファイルとして読み込み可能であるため、使用者は設定をインポートすることを推奨する。インポートするためには図 C.1 が示すメニューを選択する。そして図 C.2 に

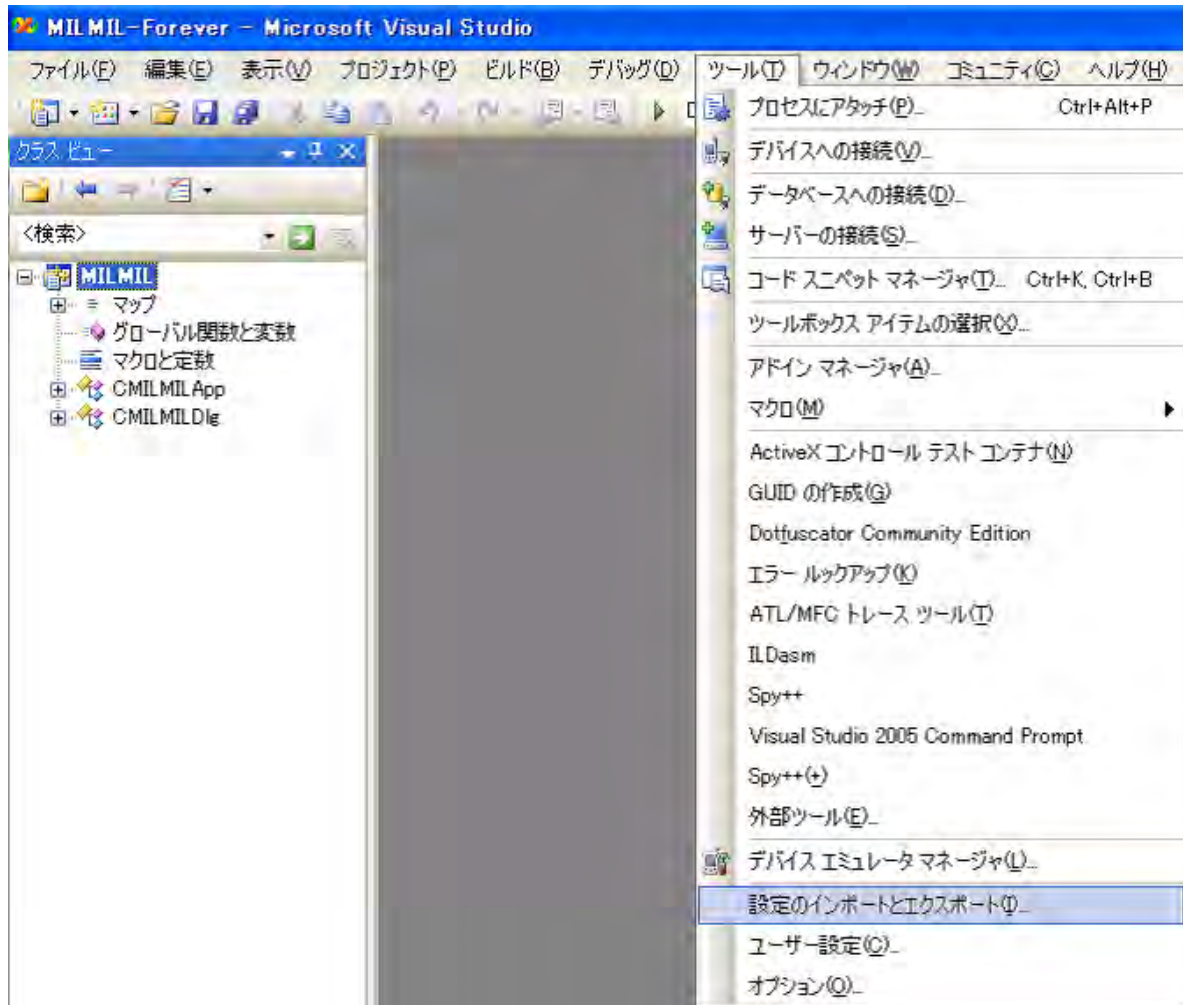


図 C.1 設定のインポート

示す手順に従う。図 C.2 の左下では設定ファイルを指定するために参照ボタンを押し「MILMIL.vssettings」ファイルを指定する。この手順が完了すると設定の読み込みが完了する。

C.2.3 開発環境の説明

使用者は主に図 C.3 に示すクラスビューとコードウィンドウにて開発を行う。図の左側はクラスビューであり上側にはクラス情報が、下側にはその選択したクラスのメンバ変数や関数が表示される。この関数をダブルクリックすると右側のコードウィンドウに該当する関数のコードを記述する箇所が表示される。必要なコードを記入し F5 キーを押すことで自動的にコンパイルとリンクが行われプログラムを実行することが出来る。



図 C.2 インポート手順

C.2.4 Debug と Release

図に示す箇所に Debug と Release を選択するウィンドウがある。通常の開発では Debug を選択する。これが選択されている場合は開発環境はプログラムコード内にデバッグ用のコードを埋め込む。これによってデバッガで変数の推移や処理の流れ等を追跡することが可能でバグ修正に大いに役立つ。しかし、これらの機能によって処理速度の低下が発生する。そのため、バグ取りが完了した場合は Release モードに変更しデバッグコードを除去する必要がある。また、このときは Ctrl+F5 キーで実行することで開発環境から独立してプログラムが実行出来るようになり、本来の性能を發揮できるようになる。

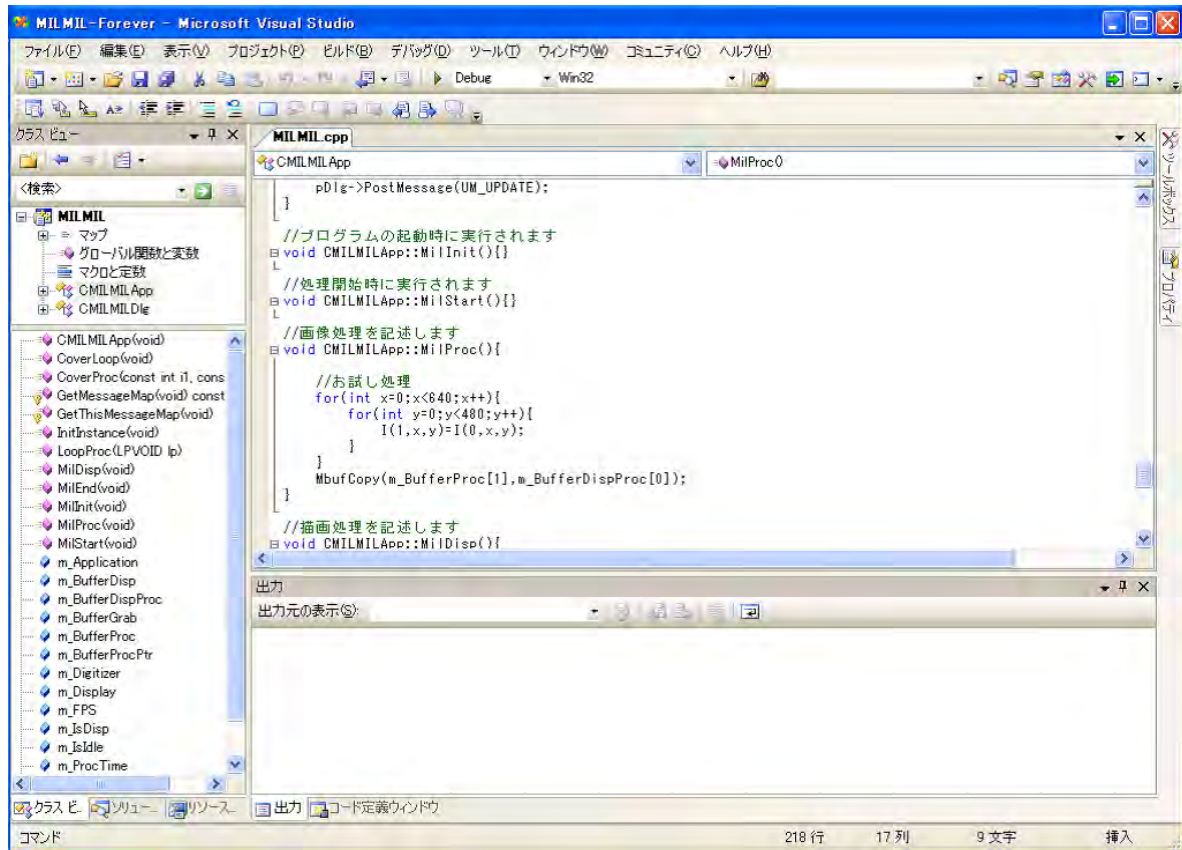


図 C.3 開発環境

る。ただし、Debug と Release ではコードが変更され Debug で発生しなかったバグが発生することがあるが、この場合はデバッガの利用に制限があるために、バグの発見が困難になる。

C.3 動画画像処理の実装方法

MILMIL を起動して処理開始ボタンを押すと MilProc 関数が定期的呼び出される。よって、この関数に静止画像処理を実装することで動画画像処理が実現できる。静止画像は MIL-Lite ライブラリ形式であるが、画素への直接アクセスを提供する関数はないため、I マクロを使用する。このマクロを使用すると、指定した画素の明度値を取得、もしくは書き換えることができる。

C.4 MILMIL の解説

C.4.1 概要

MILMIL では、C.4.5 章を読むだけで動画像処理を実装することが出来る。また、変数の初期化など処理以外のプログラムを記述する関数群は C.4.3 章に示す。MILMIL の設定を変更する場合は C.4.2 章を参考にすること。やむを得ない事情により MILMIL システム全体に手を加える必要がある場合は、C.4.8 章にて流れを理解してから行うこと。本解説書に記述されている箇所以外を編集すると予期しないトラブルに巻き込まれる可能性がある。

C.4.2 設定項目について

MILMIL.h の上部に設定項目を用意してある。

```
////MILMIL 設定項目////
```

```
//MILMIL をカラー対応にする
```

```
//#define ColorMode
```

デフォルトでは、上記定義をコメントアウトしてある。

カラー処理を行うときはこのコメントを外す。

この場合、画素アクセスマクロ I の引数が変更となるので注意すること。

未定義（モノクロ処理）の場合は引数が 3 つ。

I(処理画像番号, 座標 X, 座標 Y)

定義（カラー処理）の場合は引数が 4 つ。

I(処理画像番号, 色の指定, 座標 X, 座標 Y)

色の指定は 0 が赤, 1 が緑, 2 が青となる。

```
//入力用画像の解像度
```

```
static const int GX=640;
```

```
static const int GY=480;
```

キャプチャーボードの解像度。

処理速度を上げるために解像度を下げることが出来る。

キャプチャーボードの性能以上に解像度上げた場合の動作は未知数。

//編集用画像の解像度

```
static const int PX=640;
```

```
static const int PY=480;
```

編集用画像は複数枚維持できる。

ただし、全ての編集用画像の解像度はここで指定した値になる。

//編集用画像の数

```
static const int PN=2;
```

編集用画像の枚数を指定する。

画素アクセスマクロ使用時はこの値以上を指定するとエラーになる。

例えば PN=2 のとき、設定可能範囲は 0~1 であるので I(2,0,0) はエラーとなる。

//画面の数

```
#define DN 1
```

表示用画像の数を指定する。この数を 2 以上にした場合は下の設定項目を入力すること。

なお、配列の仕様で 0 は指定できない。

//表示用画像の解像度

```
static const int X[]={640};
```

```
static const int Y[]={480};
```

//表示用画像の表示位置（左上の位置を指定）

```
static const int IX[]={0};
```

```
static const int IY[]={0};
```

DN 値を変更したときは必ず追記すること。

例えば DN=2 とした場合、

```
static const int X[]={640,800};
```

```
static const int Y[]={480,800};
```

```
static const int IX[]={0,200};
```

```
static const int IY[]={0,200};
```

の様に設定する。

C.4.3 処理関数

プログラムはこれらの関数のみ追加することが望ましい。これらの関数は MILMILApp にて定義されている。よって、ダイアログにアクセスするときは pDlg マクロを使用すること。

```
void MilInit();
```

プログラムの起動時に一回だけ実行される関数。変数の初期化等に利用する。

この関数が実行されるときはダイアログが作られていないので pDlg マクロは利用できない。

ダイアログの初期化は MILMILDlg の OnInitDialog 関数で行うこと。

```
void MilStart();
```

処理開始ボタンを押したときに実行される関数。

背景差分の背景を保存するなどに用いる。

```
void MilProc();
```

処理中に常に呼び出される関数。

主に画像処理や画像からの情報収集を行う。

ここに描画関数も書けるがちらつきの原因となるため、MilDisp 関数で書くことを推奨する。

```
void MilDisp();
```

処理中に常に呼び出される関数。

必ず MilProc 関数の後に呼び出される。

MilProc 関数で得た情報を描画する。

```
void MilEnd();
```

処理停止ボタンを押したときに実行される関数。

取得した情報のファイル保存などに用いる。

C.4.4 定義項目について

MILMIL.h には有用な定義が書かれている。これらの定義は MILMIL システムに深く関わる定義であるので変更はしないこと。使用方法は次の通り。

```
#define pApp ((CMILMILApp*)&theApp)
```

pApp はダイアログクラスでアプリケーションクラスの変数関数を利用するとき用いる。

```
#define pDlg ((CMILMILDlg*)theApp.m_pMainWnd)
```

pDlg はアプリケーションクラスでダイアログクラスの変数関数を利用するとき用いる。

```
#define ZM(name) ZeroMemory(name, sizeof(name))
```

ZM は配列の初期化を行う。

引数に配列名を指定すると自動的にサイズを取得し配列の値を全て 0 で埋める。

```
#define L(name, index) name.GetAt(name.FindIndex(index))
```

L は Clist の要素へアクセスを容易にする。

Clist は本来、序数で要素にアクセスできないが、それを可能にする。

MILMIL-Forever では AVI ファイルのリスト管理に用いている。

```
I(n, c, x, y) *(m_BufferProcPtr[n][c]+P*(y)+x)
```

```
I(n, x, y) *(m_BufferProcPtr[n]+P*(y)+x)
```

I は画像の画素へのアクセスを行う。

カラー処理を有効にしているときは引数が 4 つ、無効にしているときは 3 つとなる。

C.4.5 画像処理の方法

この章で出てくるアスタリスクに囲まれた項目は C.4.2 を参照すること。MILMIL では画像取得用・画像編集用・画像表示用メモリを用意している。画像取得用は常に 2 枚用意されており、1 枚の画像メモリを編集しているときに別の画像メモリに次の画像を読み込むため、高速な取り込みを可能となっている。画像編集用は C.4.2 章に示す*PN*の値に応じて用意される。画像取得用・画像編集用メモリは*ColorMode*を定義することに

よりカラー対応メモリとなる。画像表示用メモリは常にカラー対応メモリとなっており、MgraLine 関数等を利用することで図形を色描画することが出来る。画像編集用メモリのみポインタアクセスするためのマクロ*I*が用意されている。I マクロはカラー非対応時は引数が 3 つ、カラー対応時は 4 つのなっている。

画像は画像入力用メモリに保存された後、MilProc 関数が始まる前に画像編集用メモリの 0 枚目にコピーされる。従って、0 枚目の画像編集用メモリに入力された処理に適切なプログラムを行うことになる。画像処理が終わり画面を表示するときは（続きは後日）

C.4.6 AVI ファイルを入力とする方法

MILMIL-Forever は起動時にはカメラ入力になっている。AVI ファイルから画像を読み込みたいときはダイアログにその AVI ファイルをドラッグする。AVI ファイルの対応形式は非圧縮形式のみである。

複数の AVI ファイルを読み込みたいときは MAL(MILMIL AVI LIST) 形式ファイルを作成する。MAL ファイルは普通のテキストで拡張子を .mal にすれば良い。テキストには読み込みたい AVI ファイルの名前（拡張子は必要ない）を記述する。必ず MAL ファイルと AVI ファイルは同じフォルダに置くこと。カメラ入力に戻りたいときはカメラボタンを押す。

C.4.7 従来の MILMIL シリーズとの相違点について

記述クラスの変更

MILMILNET までは画像処理プログラムはダイアログクラスに記述していた。しかし、ダイアログの表示に関する変数関数と画像処理に関する変数関数が混在し、開発の負担が増す原因になっていると考えられる。したがって、MILMIL-Final と MILMIL-Forever はアプリケーションクラスに画像処理に関する変数や関数を用意してある。よって、これ以前の MILMIL シリーズから移植する場合はこの点に注意する必要がある。

中間ファイルの扱い

MILMIL シリーズではコンパイル時に発生する中間ファイルが MILMIL フォルダ内に生成されるため、容量が無駄に大きくなる傾向にあった。MILMIL-forever は中間ファイルをパソコンの一時フォルダ内に生成するように変更したため、MILMIL のフォルダは従来のシリーズより小さくなる。

C.4.8 処理の流れ

MILMIL システムを変更するときは処理の流れを参考にすること。

< プログラム起動時 >

MILMILApp::CMILMILApp (変数の初期化)

MILMILApp::InitInstance (MIL ライブラリの初期化)

MILMILApp::MilInit (開発者が必要とする一度しか呼び出されない初期化処理)

MILMILDlg::OnInitDialog (ダイアログクラスに関する初期化)

< 処理開始ボタン >

MILMILDlg::OnBnClickedProc (開始処理ボタンを処理停止ボタンに変更)

MILMILApp::CoverLoop (処理関数に必要な初期化)

MILMILApp::MilStart (開発者が必要とする変数の初期化)

MILMILApp::LoopProc (高速処理のためのループ)

MILMILApp::CoverProc (キャプチャーボードもしくは AVI ファイルからの画像読み込み)

MILMILApp::MilProc (開発者が必要とする画像処理プログラム)

MILMILApp::MilDisp (開発者が必要とする画像表示プログラム)

MILMILApp::LoopProc へ戻る

< 処理停止ボタン >

MILMILDlg::OnBnClickedProc (開始停止ボタンを処理開始ボタンに変更)

MILMILApp::CoverLoop (MILMILApp::LoopProc を停止する)

MILMILApp::MilEnd (開発者が必要とする終了処理)

< 終了ボタン >

MILMILApp::MilFinish (開発者が必要とする一度しか呼び出されない終了処理)

MILMILApp::InitInstance (MIL ライブラリの解放)

C.5 MILMIL シリーズの開発履歴

筆者が配属された平成 12 年当時の渡邊研究室では、呼吸モニタリングソフトを改良して動画像処理プログラムを実装するのが主流であった。しかし、このソフトは(株)東芝にて数年かけて作られたものであり、機能の追加は極めて困難である。また、先輩の研究用プログラムは Visual Basic で作られており、開発環境の分裂化が起きようとしていた。そこで、誰でも簡単に開発が可能で共通に利用できる共通プログラムが必要であると感じたことが MILMIL の開発動機である。

当時は画像処理ライブラリである MIL-Lite のリファレンスが存在せず、先輩のプログラムから関数の動作を推測するなどの理由から、開発は困難を極め夏休みを全日費やすことになったが、同年 9 月に共通プログラムが完成した。この共通プログラムは使用しているライブラリ MIL-Lite を参考にして MILMIL と命名した。以後、同様の思想で開発した共通プログラムには MILMIL の名前が付き、MILMIL シリーズとして長年に渡って使われることになる。

MILMIL 開発直後に先輩の研究でカラー画像処理が必要になったことから、カラー処理可能な MILMIL の開発依頼があり、MILMILColor を開発した。また、同年 10 月より情報工学実験 の TA を担当することになり、実験内容として学生による動画像処理を立案したことから、OnTimer 関数のみを動画像処理が実装可能である MILMILExp を開発した。なお、これ以降も情報工学実験の TA を担当することになったが、実験の準備期間に最新版 MILMIL を MILMILExp に改造して用いたことから、MILMILExp は複数の版が存在する。同年 11 月には工学部公開用に複数の動画像処理を実装した MILMILV2 を開発した。この版では、初めて画像メモリへのポインタアクセスを実装した。しかし、当時はこの版は高機能で使いこなすことが困難であり、画像へのアクセスを配列方式に戻し MILMLILite を開発することになった。

研究室の開発環境は Visual Studio 6.0 から Visual Studio.net に変更になり、画像処理ライブラリが 6.2 から 7.5 にバージョンアップしたことから、MILMILNET を開発した。これ以前の版は OnTimer 呼び出し方式であったが、この方式は Timer の呼び出し精度が悪いことから十分な速度が出ない問題点があった。よって、この版では Hook 関数呼び出し方式(リリースモード時のみ)を採用した。これによって、キャプチャーボード本来の性能が出るようになった。

MILMIL-Final は最終版となることを想定して命名した。この版から呼び出し方式を ThreadLoop 方式に変更した。ThreadLoop 方式とは、処理用のスレッドを新規に作成しその中で処理関数を常にループさせる方式である。これによって呼び出しタイミングによる考慮する必要が無くなった。処理効率向上のために MILMILV2 で実装した Pointer ア

	Name	Development	Version	Buffer	Access	Call
1st	MILMIL	VC++ 6.0	Lite 6.2	Mono	Array	OnTimer
2nd	MILMILColor	VC++ 6.0	Lite 6.2	Color	Array	OnTimer
3rd	MILMILExp	VC++ 6.0	Lite 6.2	Mono	Array	OnTimer
4th	MILMILV2	VC++ 6.0	Lite 6.2	Color	Pointer	OnTimer
5th	MILMILLite	VC++ 6.0	Lite 6.2	Mono	Array	OnTimer
6th	MILMILNET	VC++.NET	Lite 7.5	Mono	Array	OnHook
7th	MILMIL-Final	VC++.NET	Lite 8.0	Mono	Macro	ThreadLoop
8th	MILMIL-Forever	VC++ 2005	Lite 8.0	Color	Macro	ThreadLoop

クセス方式をマクロ定義にて容易に利用できる用に改良して実装した。また、AVI ファイルの読み込みをサポートした。

なお、この後に MILMIL2006 を開発した。これは MIL-Lite8.0 から利用可能となった RingBuffer 方式を利用するものである。しかし、この方式では AVI 読み込みのサポートが困難であることが判明したため公開することは無かった。

平成 18 年 10 月、プログラムに完成は無いことやこれからも MILMIL シリーズが発展することを祈願して、MILMIL-Forever を開発した。開発環境は Visual Studio 2005 になった。基本性能は MILMIL-Final と同様であるが、カラー処理の Pointer 方式を初めて採用した点や AVI ファイルの連続読み込み、ディスプレイの複数サポートなどが新しい機能となる。

今後の MILMIL の開発余地として、マルチスレッドを用いた処理ルーチンの実装が挙げられる。MILMIL-Forever では、この処理を実装すると開発者の理解が困難になることが予想されたため実装しなかった。マルチコア CPU が十分に普及したときに開発を検討して頂きたい。

参考文献

- [1] 加藤丈和, 深尾隆則, 羽下哲司, "対象追跡 フレーム間の類似度に着目した手法から動きのモデルに着目した手法まで", "情処研報 CVIM, 2005-CVIM-150, pp.185-198, 2005.
- [2] 白井良明, 三浦純, "複雑背景における人の追跡", "情処論: コンピュータビジョンとイメージメディア (CVIM), vol.43, no.SIG4, 2002.
- [3] 羽下哲司, 鷲見和彦, 八木康史, "時間平均シルエットを用いた能動カメラによる人の追跡", "信学論 D-, Vol. J88-D- No.2 pp.291-301, 2005.
- [4] M. Isard, A. Blake, "Condensation - Conditional Density Propagation for Visual Tracking," International Journal of Computer Vision, vo.29, no.1, pp.5-28, 1998.
- [5] 鈴木達也, 岩崎慎介, 佐藤洋一, 杉本晃宏 "環境モデルの導入による人物追跡の安定化," 画像の認識理解シンポジウム (MIRU2004), pp.243-248, 2004.
- [6] 和田俊和, 濱塚俊明, 加藤丈和, "K-means トラッキング:背景混入に対して頑健な対象追跡法," 画像の認識理解シンポジウム (MIRU2004), pp.7-12, 2004.
- [7] 齋藤慎也, 佐治斉, "テンプレートの可変分割と統合による人物の動作追跡," 情処研報 CVIM, 2004-CVIM-145, pp.33-39, 2004.
- [8] 申熙卓, 渡辺俊典, 菅原研, "テンポラルテンプレートを用いた動画解析手法," 信学技法 vol.102, PRMU2002-111, 2002.
- [9] 佐竹純二, 尺長健, "階層的注視点制御による動画像上での複数人物追跡," 信学論 D-, Vol. J86-D- No.8 pp.1212-1221, 2003.
- [10] 片山徹, "応用カルマンフィルタ," 朝倉書店, 1983.
- [11] 福添孝明, 伊藤雅人, 水戸大輔, 渡邊睦, "分散協調型人物認証方式の研究開発," MIRU2004, pp.607-612, 2004.
- [12] 福添孝明, 伊藤雅人, 水戸大輔, 渡邊睦, "複数画像特徴の統合に基づく完全非拘束型人物認証方式の研究," MIRU2005, pp.747-752, 2005.
- [13] 福添孝明, 伊藤雅人, 水戸大輔, 渡邊睦, "複数画像特徴の統合に基づく完全非拘束型人物認証方式の研究," SI2005, pp.1101-1102, Dep. 2005.

- [14] 福添孝明, 伊藤雅人, 渡邊睦, "複数画像特徴の統合に基づく完全非拘束型人物同定方式の研究における在席習慣性特徴の導入," *MIRU2006*, pp.782–787, 2006.
- [15] 社団法人日本自動認識システム協会, "よくわかるバイオメトリクスの基礎," オーム社, 2005.
- [16] M.A. Turk, A.P. Pentland, "Face recognition using eigenfaces," *IEEE Proc. CVPR*, pp.586–591, Jun. 1991.
- [17] L. Wiskott, J.M. Fellous, N. Kruger, C. Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Trans. PAMI*, Vol.19, No.7, pp.775–779, Jul. 1997.
- [18] 杉江嘉昭, 小林哲則, "Dempster-Shafer 理論を用いた音声・画像情報の統合による個人認識システム," *信学技報 MVE*, Vol.101, No.425, pp.63–68, 2001.
- [19] R. Brunelli, D. Falavigna, "Person identification using multiple cues", *IEEE Trans. PAMI*, Vol.17, No.10, pp.955–966, Oct. 1995.
- [20] T. Choudhury, B. Clarkson, T. Jebara, A. Pentland, "Multimodal person recognition using unconstrained audio and video," *AVBPA*, pp.176–181, Mar. 1999.
- [21] 前田茂則, 岡本真明, 河原達也, 美濃導彦, 池田克夫, 堂下修司, "顔画像特徴, 歩行画像特徴および音声特徴の統合による個人識別", *信学論 (D)*, Vol.J79-D2, No.4, pp.600–607, Apr.1996.
- [22] N.A. Fox, R. Gross, P. Chazal, J.F. Cohn, R.B. Reilly, "Person identification using automatic integration of speech, lip, and face experts", *ACM Proc. BMA*, pp.25–32, 2003.
- [23] 西山正志, 湯浅真由美, 柴田智行, 若杉智和, 山口修, "顔画像の階層的な対応付けを用いた歩行者認識システム," *MIRU2006*, p.530, 2006.
- [24] 江本光晴, 林朗, 末松伸朗, 岩田一貴, "パーティクルフィルタを用いた歩き方による個人識別," *信学技報 PRMU*, Vol.106, No.230, pp.41–48, Sep. 2006.
- [25] C.Y. Yam, M.S. Nixon, J.N. Carter, "Automated person recognition by walking and running via model-based approaches," *Pattern Recognition*, Vol.37, No.5, pp.1057–1072, 2004.
- [26] 深山篤, 澤木美奈子, 村瀬洋, 萩田紀博, "歩行動作特性からの年齢層の推定," *信学論 (D)*, Vol.J84-D2, No.7, pp.1522–1526, Jul. 2001.
- [27] 小阪勇氣, 平山高嗣, 岩井儀雄, 谷内田正彦, "全方位画像を用いた歩行軌跡による確率的な人物行動認識システム," *情処学 CVIM 研報*, Vol.2005, No.38, pp.163–170, May. 2005.
- [28] 青木茂樹, 岩井嘉男, 大西正輝, 小島篤博, 福永邦雄, "人物の位置・姿勢に注目した行動パターンの学習・認識と非日常状態検出への応用," *信学論 (D)*, Vol.J87-D2,

- No.5, pp.1083–1093, May. 2004.
- [29] 島田竜也, 河口尚広, 加賀健太, 山田博三, 森晃徳, ” 屋内侵入者検知のための動的背景抽出法, ” 信学論 (D), Vol.J88-D2, No.10, pp.2054–2068, Oct. 2005.
- [30] 河内まき子, 持丸正明, ” 2005 AIST 人体寸法データベース, ” 産業技術総合研究所 H16PRO 287.
- [31] 池上祥一郎, 塩山忠義, ” 画像中における目の検出, ” PRMU2003-198, pp.49–54, 2004.
- [32] 小田政志, 舩曳信生, 東野輝夫, ” Deformable Template マッチング法による眼の輪郭線抽出に関する研究, ” PRMU2003-140-150, pp.1–6, 2003.
- [33] 加島英樹, 本郷仁志, 加藤邦人, 山本和彦, ” 顔の向きの変化に強い瞳検出方法の研究, ” 2001-CVIM-127, pp.139–144, 2001.
- [34] 山田浩史, 乗松有, 小塚一宏, 中野倫明, 山本新, ” ドライバの運転状態の検知 1-インナミラー内蔵カメラによる視線検出-, ”画像の認識・理解シンポジウム (MIRU2004), pp.63–68, 2004.
- [35] 足立和正, 井東道昌, 小林史和, 中野倫明, 山本新, ” ドライバの運転状態の検知 2-インナミラー内蔵カメラによる視線検出-, ”画像の認識・理解シンポジウム (MIRU2004), pp.69–74, 2004.
- [36] 白井了, 井東道昌, 乗松有, 中野倫明, 山本新, ” ドライバ運転状態検知のための視線方向の検出, ” PRMU2003-151-170, pp.67–72, 2003.
- [37] 井東道昌, 稲垣文治, 山本新, 森恵, ” ドライバ状態モニタ, ” 2005-CVIM-150, pp.47–54, 2005
- [38] 片岡祐介, 岩口浩章, 佐治斉, ” 目と口の追跡による笑顔の分類, ” 2001-CVIM-129, pp.109–116, 2001.
- [39] 関輝夫, ” しぐさでわかる相手の心理, ” 新星出版社, 2000 .
- [40] 大津展之, ” 判別および最小 2 乗基準に基づく自動しきい値選定法 ” 信学論 (D), Vol.J63-D, No.4, pp.349–356, 1980 .
- [41] 井上誠喜ら, ” C 言語で学ぶ実践画像処理 ”, オーム社, 1999 .

謝辞

本研究に当たり，熱心に御指導頂きました渡邊睦教授，内山博之教授，森邦彦教授に深く感謝の意を表します．また，人物認識の研究チームとして御討論頂きました伊藤雅人氏，植野栄一氏，梅原末世氏，田中宏樹氏，西小野政幸氏，濱田未来氏，別府恵美氏，堀之内あすか氏，水戸大輔氏にお礼を申し上げます．最後に研究活動を支えていただきました渡邊研究室の皆様感謝いたします．