

全方位カメラとPTZカメラを用いた人物監視 システムの構築

水本, 貴弘 / MIZUMOTO, Takahiro

(発行年 / Year)

2012-03-24

(学位授与年月日 / Date of Granted)

2012-03-24

(学位名 / Degree Name)

修士(工学)

(学位授与機関 / Degree Grantor)

法政大学 (Hosei University)

2011年度 修士論文

全方位カメラとPTZカメラを用いた
人物監視システムの構築

指導教授 赤松 茂 教授

法政大学大学院工学研究科
システム工学専攻修士課程

10R6131

ミズモト タカヒロ
水本 貴弘

概要

本研究では 1 台の全方位カメラと 2 台の Pan/Tilt/Zoom カメラ(以後, PTZ カメラ)を連携させ, 広範囲における人物監視システムを構築した.

このシステムでは, 周囲 360 度を一度に撮影可能な全方位カメラによって視覚情報を一元的に管理し, その情報を基に首振りカメラである PTZ カメラが人物の追跡をしているので, 追跡性能に安定性をもたらすことに成功している.

本システムの追跡性能は, 実験によって評価した.

キーワード: 人物追跡, パーティクルフィルタ, Bhattacharyya 係数, オクルージョン

Abstract

In this study, We developed a system for achieving automatic, stable, and accurate tracking of each person among those walking around an observed space by integrating the operations of an omni-directional camera that provides a 360° view and PTZ cameras whose orientation of the visual axis and the scale of the zooming magnification are controlled by a host computer.

The stability of the tracking performance was investigated by experiments.

Keywords: motion tracking, particle filter, Bhattacharyya coefficient, occlusion

目次

Abstract

第 1 章	はじめに	1
第 2 章	システム構成	
2.1	システム概要	3
2.2	ハードウェア構成	
2.2.1	全方位カメラ	4
2.2.2	PTZ カメラ	8
2.3	ソフトウェア構成	
2.3.1	画像処理ソフト	13
2.3.2	画像変換ソフト	14
2.3.3	開発環境	15
第 3 章	移動物体検出方法	
3.1	パーティクルフィルタ	16
3.2	フレーム間差分法	18
第 4 章	移動物体追跡方法	20
第 5 章	Zoom について	21

第6章 オクルージョン問題

6.1 自己オクルージョン	22
-------------------------	----

6.2 相互オクルージョン	22
-------------------------	----

第7章 対応付け

7.1 Bhattacharyya 係数	24
--------------------------------	----

7.2 座標軌跡からの予測	25
-------------------------	----

7.3 対応付けの手順	26
-----------------------	----

第8章 フローチャート 27

第9章 評価実験

9.1 追跡性能および Zoom の有効性の評価

9.1.1 実験内容	28
----------------------	----

9.1.2 評価方法	29
----------------------	----

9.1.3 実験結果	30
----------------------	----

9.1.4 考察	36
--------------------	----

9.2 人物の顔の追跡について

9.2.1 実験内容	37
----------------------	----

9.2.2 評価方法	38
----------------------	----

9.2.3 実験結果	39
----------------------	----

9.2.4 考察	46
--------------------	----

第10章 結び	47
謝辞	48
参考文献	49

第1章 はじめに

近年，人々の生活を脅かす犯罪を減らすための「防犯」，消費者の求めている商品やサービスを調査し流通を円滑にするための「マーケティング」，迷子や指名手配犯を迅速に発見するための「搜索」，その他多くの場面で人物を監視することが必要とされている[1]．

特に，2001年9月に起きたアメリカ同時多発テロ事件以降、各国の防犯に対する意識は急激に高まり，空港や商店など様々な場所の警備が強化された．

また，市場争いが年々激化しているため、消費者の情報を得るために各企業がマーケティングの重要性を再認識して力を入れている．

これらの監視活動や調査活動は，人間が行うことも可能ではあるが，それらに割り当てられる人間の活動時間および労働力は限られたものであるため，人間が常に行い続けることは現実的ではない．

そのため，カメラを用いた自動監視の研究が盛んに行われている[2][3][4]．

しかし，これらの監視は固定された単一カメラによるものが多く、監視領域が非常に限定されてしまう．

監視領域が限定されてしまうということは，検出する機会を少なくするのと同義であり，犯罪や顧客を見逃してしまう可能性が高くなってしまう．特に，犯罪者および犯罪行為の見逃しは，致命的な事柄に発展しかねないので，絶対にあってはならないことである．

警視庁では，犯罪が発生するがい然性の極めて高い繁華街等の防犯対策の一環として，「街頭防犯カメラシステム」というものを導入している．これは，犯罪の予防と被害の未然防止を図るため，苦境空間に防犯カメラを設置し，設営した映像を常時にモニター画面に映し出し，これを録画するものである．

平成23年には，警視庁本部において録画した294件の映像データを警察署長に提供し，犯人の検挙活動等に活用され，うち190件が事件の証拠資料として立件されたという実例もある[5]．

この問題を解決するために，多数のカメラを死角が出来ないように配置する方法や1台で周囲の全範囲を監視可能な全方位カメラを設置する方法がある[6]．

しかしながら，前者においては複数のカメラを設置する場所の確保，費用，そして各カメラの情報を一元的に把握するための統合作業が複雑となってしまう．

後者においては，移動物体がどの範囲に存在しているかを一元的に把握することは出来るが，移動物体の詳細な特徴を掴むといった用途に適していない．

そこで，本研究では，広範囲の撮影かつ情報を一元的に管理出来る全方位カメラと首振りカメラである PTZ カメラを連動させることによって，広範囲における人物監視システムを構築した．

また，複数の人物を追跡出来るように，各カメラと各追跡対象を Bhattacharyya 係数と追跡対象の座標軌跡を利用して対応付けするようにしている．

第2章 システム構成

2.1 システム概要

本研究では，1 台の全方位カメラと 2 台の PTZ カメラを組み合わせ，周囲 360 度に映る人物を撮影し，追跡するシステムを構築した．システムの基本情報を下記に示す．

■ハードウェア

OS	: Windows XP Professional SP3	
CPU	: Pentium4 HT 3.00GHz	
メモリ	: 2GB	
画像入力ボード	: Picport Color (Letron Vision)	×3
全方位カメラ	: CV-S3200 + HYPER70	×1
PTZ カメラ	: EVI-D100	×2

■ソフトウェア

画像処理	: HALCON (ver.7.1)	
パノラマ変換	: Hyper Omni Image SDK	
システム実装	: Microsoft Visual C++ 2008 Express Edition	

2.2 ハードウェア

2.2.1 全方位カメラ

今回使用した全方位カメラは、一般の CCD カメラに全方位ミラーを取り付けたものである。天井に向けて設置した CCD カメラの先端に全方位ミラーを取り付けることで、ミラーに映った周囲 360 度の画像を取得することができる。ただし、この場合得られる画像は中心付近が欠損した円画像（ドーナツ状の画像）となる[7]。

全方位カメラの構成を表 2-1 に示す。

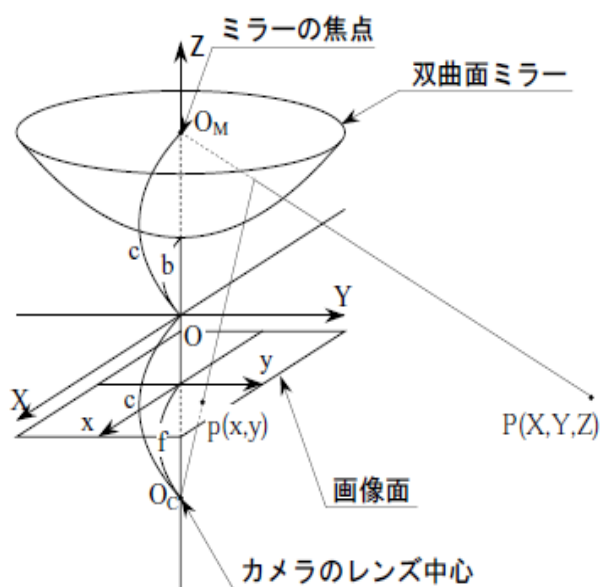
表 2-1 全方位カメラの構成

カメラ	CV-S3200(jAi)
ミラー	HYPER70(映蔵)
レンズ	HF9HA-1B(FUJINON)
アダプタ	RA3727P5 F37-M27mm(raynox)
	RA5237A F52-M37mm(raynox)

次に、全方位カメラの全体図と、その撮影原理について述べる。



(a) 全体図



(b) 視覚系の構成

図 2-1 全方位カメラ

図 2-1 より，ミラーの焦点とカメラのレンズ中心は，それぞれ対称の位置 $(0,0,c)$ ， $(0,0,-c)$ にある．画像面は XY 平面に平行で，カメラのレンズ中心からカメラの焦点距離 f だけ離れた平面とする．この場合，ミラー面は次式で表すことができる．

$$\frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \quad (2.1)$$

図 2-1 から，使用した全方位ミラーである **HYPER70** のミラーパラメータと焦点距離の各値を当てはめると，表 2-2 のように表すことができる．

表 2-2 ミラーパラメータ

a	32.799	[mm]
b	47.408	[mm]
c	55.207	[mm]
f	9	[mm]

ここで，図 2-1 を水平方向と鉛直方向に分けたものを，図 2-2，図 2-3 に示す．

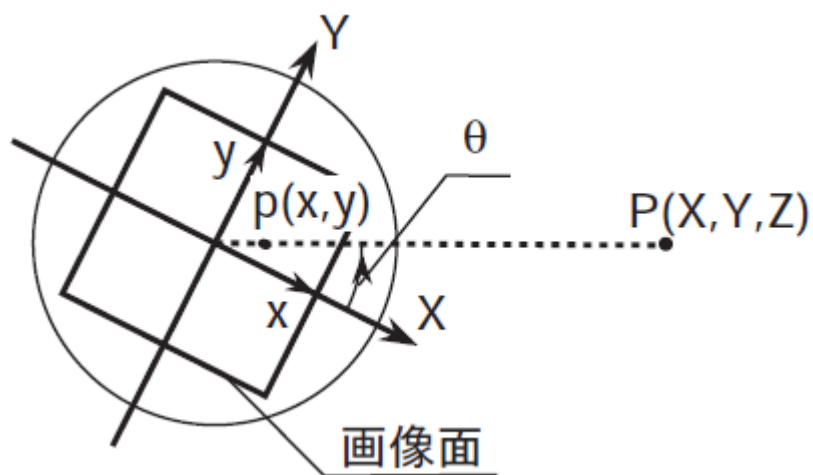


図 2-2 点の射影水平方向

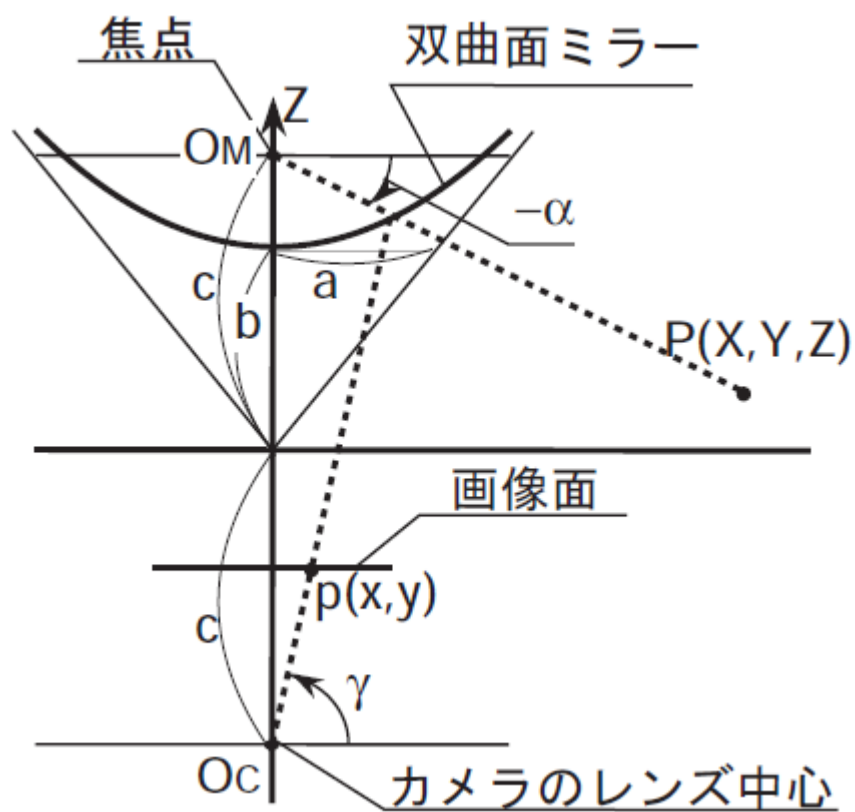


図 2-3 点の射影鉛直方向

図 2-2 より，空間内の任意点 $P(X,Y,Z)$ と画像上の写像点 $p(x,y)$ との関係について考えると，点 P の方位角 θ は次式で表すことができる．

$$\tan \theta = \frac{Y}{X} = \frac{y}{x} \quad (2.2)$$

すなわち，空間内の対称物体の方位角 θ が，そのまま画像上の写像の方位として直接現れる．また，図 2-3 に示した鉛直断面を考えると，点 P と点 p の間には次式の関係が成り立つことが分かる．

$$Z = \sqrt{X^2 + Y^2} \tan \alpha + c \quad (2.3)$$

$$\alpha = \tan^{-1} \frac{(b^2 + c^2) \sin \gamma - 2bc}{(b^2 - c^2) \cos \gamma} \quad (2.4)$$

$$\gamma = \tan^{-1} \frac{f}{\sqrt{x^2 + y^2}} \quad (2.5)$$

ミラーの焦点からの点 P の方位角 θ および伏角 α は，カメラのレンズ中心を双曲線の焦点位置にすることで，写像点 p より一意に求まる．このとき，ミラーの焦点は固定であるため，ミラーの焦点から見たカメラを鉛直軸周りに回転して得られる画像や，一般のカメラ画像へと入力画像を変換することができる．

また，上式を変形すると，次式が求まる．

$$x = X \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \quad (2.6)$$

$$y = Y \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \quad (2.7)$$

これにより，空間内の任意点 $P(X,Y,Z)$ に対応する写像点 $p(x,y)$ を求めることができる．

2.2.2 PTZ カメラ

PTZ カメラは，SONY EVI-D100 を 2 台使用した．このカメラは，高速パン・チルト機能を備え，広角ズームレンズを内蔵した回転台付のカメラである．EVI-D100 には SONY が開発した制御プロトコル VISCA 端子があり，RS-232C 接続することによって，シリアル通信によりホスト側から様々な制御を行うことが可能である．

また，VISCA 端子を利用して複数台の EVI-D100 同士を接続し，一律に制御することもできる[7]．



図 2-4 SONY EVI-D100

EVI-D100 の仕様は表 2-3 に示すとおりである．

表 2-3 PTZ カメラの仕様

映像信号	NTSCカラー JEITA標準方式
映像素子	1/4インチ カラーCCD (総画素数:約41万画素, 有効画素数約38万画素)
レンズ	光学10倍, デジタル40倍 f=3.1~31mm, F1.8~F2.9 水平画角:6.6度~65度
最至近撮影距離	100mm(WIDE端), 600mm(TELE端)
シャッタースピード	1/4~1/10000秒
パン・チルト機能	水平±100度, 最大速度:300度/秒 垂直±25度, 最大速度:125度/秒

VISCA ではコンピュータなどコマンドを出すホスト側をコントローラ、EVI-D100 などコマンドを受ける側を周辺機器と呼ぶ。VISCA による通信は RS-232C に準拠しており、1 台のコントローラに 7 台までの周辺機器を接続することができる。

RS-232C のパラメータは次のとおりである。

- ・通信速度 : 9600bps
- ・データ長 : 8 ビット
- ・スタートビット : 1 ビット
- ・ストップビット : 1 ビット
- ・パリティ : なし

複数台の周辺機器を用いる場合、それぞれをデジチェーン状に接続する必要がある。このとき、実際の内部の接続は図 2-5 のように一方通行のリングになっている。また、ネットワーク上の各機器にはアドレスがついており、コントローラのアドレスは 0 に固定されている。周辺機器のアドレスはコントローラから近い側から順に 1, 2, 3 とついていく。これにより、どの周辺機器を制御するかを区別することができる。

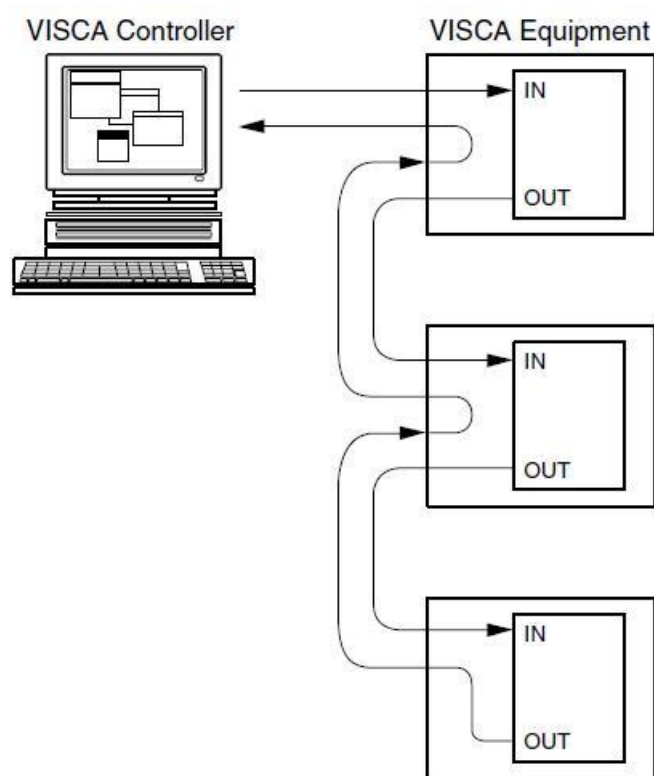


図 2-5 VISCA ネットワーク構造

コントローラから周辺機器を制御する場合、その通信に用いられる基本単位をパケットと呼ぶ。パケットの最初のバイトをヘッダと呼び、差し出しと宛先のアドレスが入っている。例えば、アドレス 0 のコントローラからアドレス 1 の周辺機器へ送るパケットのヘッダは 16 進数で 81H となる。アドレス 2 の周辺機器へ送るパケットは 82H となる。

一部の設定用コマンドは、一度にすべての周辺機器に対して送ることができる（ブロードキャスト）。ブロードキャストの場合は、ヘッダを 16 進数で 88H とする。

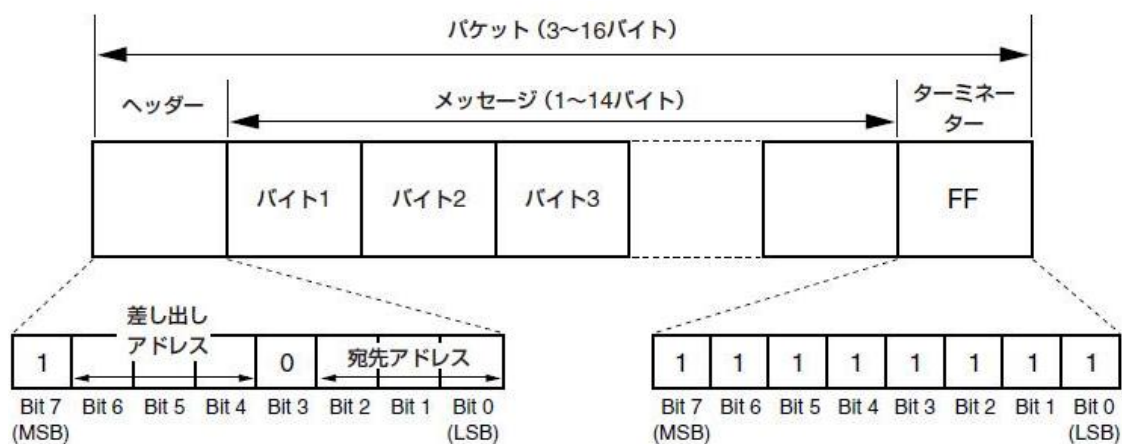


図 2-6 パケットの構造

EVI-D100 の制御を始める前には、必ず Address コマンドと IF_Clear コマンドをブロードキャストで送る必要がある。

アドレスを設定するコマンドには、次の 2 つが存在する。

- Address コマンド

周辺機器のアドレスを設定することができる。ネットワークを初期化するときや、下記のネットワークチェンジメッセージを受け取ったときに使用する。

- Network Change コマンド

ネットワーク内の機器が取り外されたり追加されたりしたときに、周辺機器からコントローラに送られる。このメッセージを受け取ったときは、アドレスを再設定する必要がある。

実際にアドレス設定をする場合には、次のようなパケットを送信する。

表 2-4 アドレス設定パケット

	Packet
Address	88 30 01 FF
Network Change	X0 38 FF
Note	X = 9～F: EVI-D100のアドレス+8

次に、EVI-D100 内のコマンドバッファをクリアし、実行中の命令を中断するには表 2-5 のようなパケットを送信する。

表 2-5 コマンドクリアパケット

	Packet
IF_Clear	8X 01 00 01 FF
IF_Clear(broadcast)	88 01 00 01 FF
Note	X = 1～7: EVI-D100のアドレス

上記のアドレス設定とコマンドクリアにより、EVI-D100 の制御準備が整う。実際の制御においては、表 2-6 のコマンドリストを使用する。

表 2-6 コマンドリスト

Command Set	Command	Command Packet	Comments
Pan-tiltDrive	Up	8x 01 06 01 VV WW 03 01 FF	VV: Pan 速度設定 01 (低速) ～ 18 (高速)
	Down	8x 01 06 01 VV WW 03 02 FF	WW: Tilt 速度設定 01 (低速) ～ 14 (高速)
	Left	8x 01 06 01 VV WW 01 03 FF	YYYY: Pan Position FA60～05A0 (center 0000)
	Right	8x 01 06 01 VV WW 02 03 FF	ZZZZ: Tilt Position FE98～0168 (center 0000)
	UpLeft	8x 01 06 01 VV WW 01 01 FF	
	UpRight	8x 01 06 01 VV WW 02 01 FF	
	DownLeft	8x 01 06 01 VV WW 01 02 FF	
	DownRight	8x 01 06 01 VV WW 02 02 FF	
	Stop	8x 01 06 01 VV WW 03 03 FF	
	AbsolutePosition	8x 01 06 02 VV WW 0Y 0Y 0Y 0Y 0Z 0Z 0Z 0Z FF	
	RelativePosition	8x 01 06 03 VV WW 0Y 0Y 0Y 0Y 0Z 0Z 0Z 0Z FF	
	Home	8x 01 06 04 FF	
	Reset	8x 01 06 05 FF	
Pan-tiltLimitSet	LimitSet	8x 01 06 07 00 0W 0Y 0Y 0Y 0Y 0Z 0Z 0Z 0Z FF	W: 1 UpRight, 0 DownLeft YYYY: Pan Limit Position FA60～05A0 (center 0000)
	LimitClear	8x 01 06 07 01 0W 07 0F 0F 0F 07 0F 0F 0F FF	ZZZZ: Tilt Limit Position FE98～0168 (center 0000)

表 2-6 において、特にカメラを任意の位置へ向けたいときには、Absolute Position（絶対位置）のコマンドを使用する。

表 2-7 絶対位置コマンド

AbsolutePosition	8x 01 06 02 VV WW 0Y 0Y 0Y 0Y 0Z 0Z 0Z 0Z FF
VV(Pan速度設定)	01(低速)～18(高速)
WW(Tilt速度設定)	01(低速)～14(高速)
YYYY(Pan Position)	FA60～05A0(center 0000)
ZZZZ(Tilt Position)	FE98～0168(center 0000)

コマンドでは、Pan Position および Tilt Position はそれぞれ 16 進数で書かれている。
Pan Position では、カメラが正面を向いている位置をセンターとし、左端が FA60（-100 度）、右端が 05A0（+100 度）となっている。Tilt Position においても同様である。

これより、Pan Position の絶対角度は表 2-8 のようになる。

表 2-8 絶対角度表

位置	-100度	-95度	-90度	-85度	-80度	-75度	-70度	-65度	-60度	-55度
16進数	FA60	FAA8	FAF0	FB38	FB80	FBC8	FC10	FC58	FCA0	FCE8
数値	64096	64168	64240	64312	64384	64456	64528	64600	64672	64744
位置	-50度	-45度	-40度	-35度	-30度	-25度	-20度	-15度	-10度	-5度
16進数	FD30	FD78	FDC0	FE08	FE50	FE98	FEE0	FF28	FF70	FFB8
数値	64816	64888	64960	65032	65104	65176	65248	65320	65392	65464

位置	+5度	+10度	+15度	+20度	+25度	+30度	+35度	+40度	+45度	+50度
16進数	0048	0090	00D8	0120	0168	01B0	01F8	0240	0288	02D0
数値	72	144	216	288	360	432	504	576	648	720
位置	+55度	+60度	+65度	+70度	+75度	+80度	+85度	+90度	+95度	+100度
16進数	0318	0360	03A8	03F0	0438	0480	04C8	0510	0558	05A0
数値	792	864	936	1008	1080	1152	1224	1296	1368	1440

表 2-8 を用いることで、任意の方向へカメラを向けることができる。表に無い角度にした場合は、数値より角度を逆算すればよい。

2.3 ソフトウェア

2.3.1 画像処理ソフト

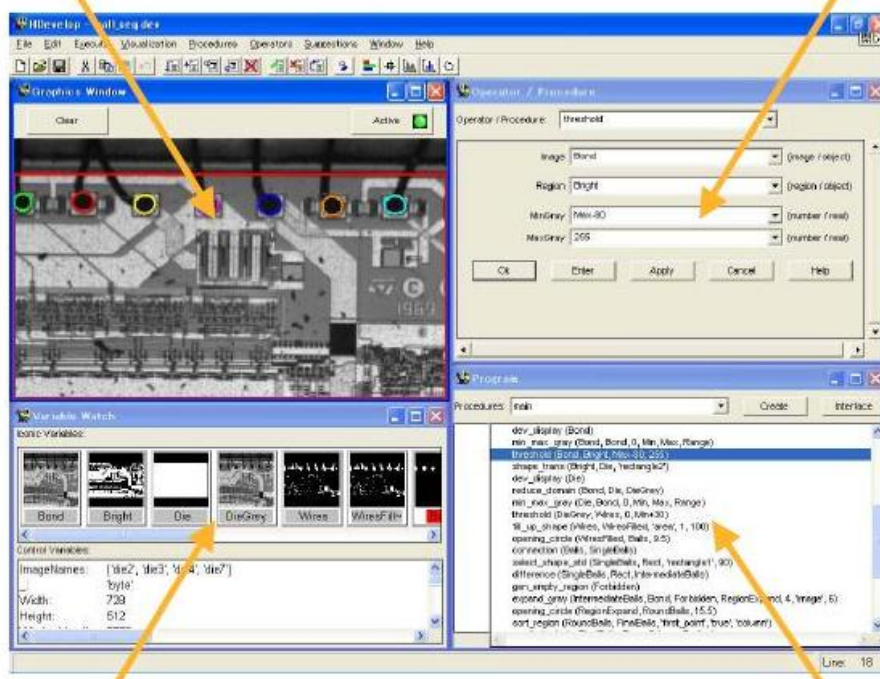
画像処理ソフト「HALCON」は、1300 を越える画像処理ライブラリオペレータと、それをベースにマシンビジョンシステムを効率的に構築できる HDevelop から構成されている。マシンビジョンシステムとは、従来目視によって行われていた各種検査業務を、コンピュータやカメラなどのハードウェアと画像処理ソフトウェアをベースにした画像処理システムに置き換えたものである。その適用範囲は、計測検査（直線検出、円弧状のエッジ検出等）、パターンマッチング、バーコードリーダー、文字認識・文字照合など様々な分野に及ぶ。

この HDevelop を利用し、必要な画像処理ライブラリに対応する画像処理用オペレータを選択して並べることで、数 100 行のコードを要していた複雑なプログラムもたったの数行で構築できる。また、こうして作成したプログラムを Visual Basic, Visual C++, Visual C# といった汎用コードに変換することもできる。

本研究では、様々な移動物体検出方法を試すためにこの画像処理ソフト HALCON を利用し、完成させたプログラムを Visual C++ に変換して実装した[7]。

Graphics ウィンドウ

Operator ウィンドウ



Variable Watch ウィンドウ

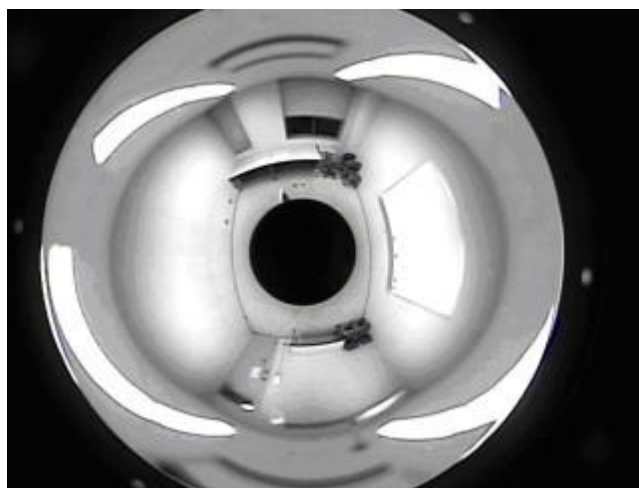
Program ウィンドウ

図 2-7 HALCON 画面

2.3.2 画像変換ソフト

全方位カメラで撮影した画像は，図 2-8 (a) に示すようにドーナツ状の画像となる．この画像を展開し，パノラマ画像を生成するために，画像変換ソフト「HyperOmni Image SDK」を用いた．HyperOmni Image SDK は全方位カメラで撮影したドーナツ状の画像をパノラマ変換するためのライブラリであり，これを用いて画像変換プログラムを作成した [7]．

(a) 入力画像



(b) 出力画像



図 2-8 パノラマ変換

使用した全方位ミラー固有のパラメータと，円中心座標および外円半径の数値を変えることで，より正確なパノラマ画像を生成することができる．

2.3.3 開発環境

本研究では、Microsoft C++ 2008 を統合開発環境として使用している。

統合開発環境とは、コンパイラ、テキストエディタ、デバッガなどを一つの対話型操作環境から利用出来るようにしたものである。

Microsoft C++ 2008 はその中でも、C, C++, C++/CLI用統合開発環境であり、HALCONで作成したプログラムをVisual C++に変換し、実装している。

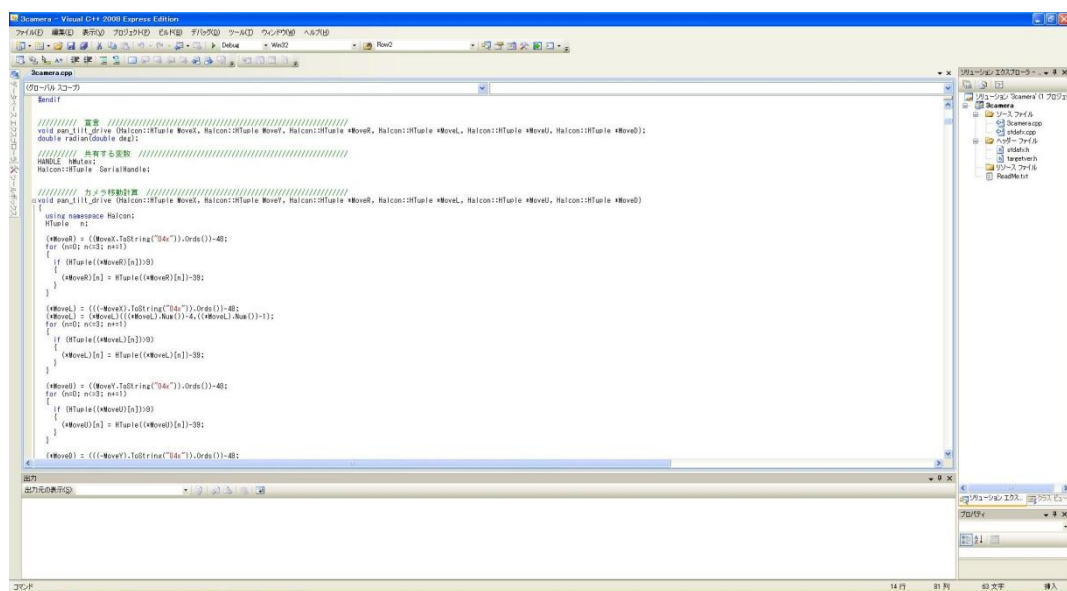


図 2-9 Microsoft C++ 2008 画面

第3章 移動物体検出方法

本システムでは、パーティクルフィルタとフレーム間差分法という 2 種類の手法を用いて、人物の検出をしている。

3.1 パーティクルフィルタ

パーティクルフィルタは、追跡対象の前の状態および現在の観測情報からの予測によって現在の状態を推定する手法である。[8][9]

この手法は、重みに基づいてパーティクルを撒く「サンプリング」、各パーティクルを状態方程式に基づいて移動させる「予測」、現在の観測と状態を基に尤度を計算する「重み付け」、パーティクルの分布から追跡対象の現在の状態を推定する「状態推定」という 4 つのアルゴリズムから成り立っている[10]。具体的な手順は以下の通りである。

1. 尤度が高いパーティクルの周辺にはパーティクルを多く撒き、尤度が低いパーティクルは消滅させる。
(初期状態では一様に撒く)
2. 状態方程式に基づいてパーティクルを移動させる。
3. それぞれのパーティクルの尤度を再計算する。
4. パーティクルの分布から追跡対象の位置を予測する。

なお、本研究では 3.2 で解説するフレーム間差分法によって得られた差分値を尤度とした。

図 3-1 がパーティクルフィルタを適用した場合の実際の図である。パーティクルが多く存在している場所に人物がいると判断し、検出をしている。

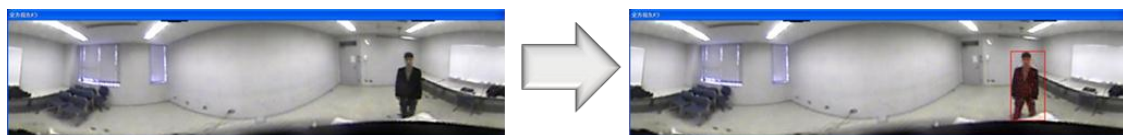


図 3-1 パーティクルフィルタ適用図

数式を用いて表すと下記のようなになる.

時刻 t における対象の状態量を x^t , 画像による観測を Z^t とし, 時刻 t までに得られる観測を $Z_t = (z_1, \dots, z_t)$ とする. このとき, 時刻 t における対象の事前確率 $P(x_t|Z_{t-1})$ は時刻 $t-1$ における事後確率 $P(x_{t-1}|Z_{t-1})$ と事後確率 $t-1$ から t への状態遷移確率 $P(x_t|x_{t-1})$ を用いて式 (3.1) のように表せる.

$$P(x_t|Z_{t-1}) = \int P(x_t|x_{t-1})P(x_{t-1}|Z_{t-1})dx_{t-1} \quad (3.1)$$

ここで, $P(x_t|Z_{t-1})$ を一定とすると, 時刻 t における事後確率 $P(x_t|Z_t)$ は, ベイズの法則より, 時刻 t における尤 $P(z_t|x_t)$ と事前確率 $P(x_t|Z_{t-1})$ により式 (3.2) のように表せる.

$$P(x_t|Z_t) \propto P(z_t|x_t)P(x_t|Z_{t-1}) \quad (3.2)$$

3.2 フレーム間差分法

移動物体がない理想的な背景画像を得ることは現実では難しい．そのような場合，移動物体を撮影した異なる時刻の 3 枚の画像を用いて移動物体領域を取り出すことができる．これがフレーム間差分法である．図 3-2 に示すように，A と B，B と C の差分画像 AB と BC を作成し，2 値化処理を施して 2 値画像を得る．次に，2 値画像 AB と BC の論理積処理（AND 処理）を行い，AB と BC の共通領域を得ることができる．このような処理を繰り返すことで動画中から移動物体を検出することができる．

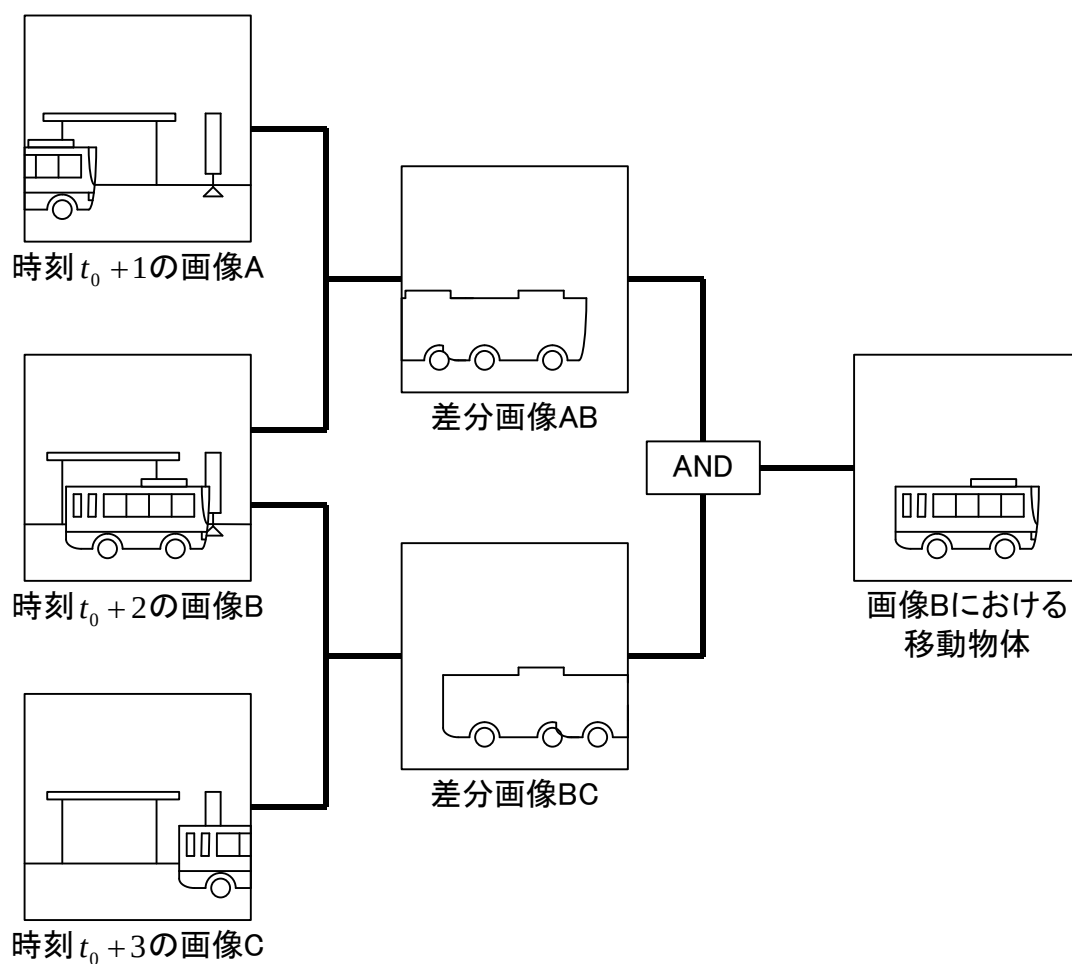


図 3-2 フレーム間差分法による移動物体検出

フレーム間差分法による移動物体検出では，対象に動きがない場合には検出が難しいが，背景を常に更新するため，環境の変化に強い面がある．

数式を用いて表すと下記のようなになる.

フレーム間差分法は, カメラによって取り込んだ時間 t の画像 $f_t(m, n)$ と時間 $t+1$ の画像 $f_{t+1}(m, n)$ の差の絶対値を求め, 差分画像 $s_t(m, n)$ を得る. $s_t(m, n)$ は以下の式(3.3)で求められる[11].

$$s_t(m, n) = |f_{t+1}(m, n) - f_t(m, n)|, \quad (3.3)$$

m : 行座標 n : 列座標

第4章 移動物体追跡方法

本研究の移動物体の追跡については，第3章で解説したパーティクルフィルタおよびフレーム間差分法を利用している．

まず，全方位カメラによって撮影された全方位画像をパノラマ画像に変換し，パーティクルフィルタおよびフレーム間差分法で移動物体を検出する．

その後，パノラマ画像上で検出された人物の領域の中心列座標に対応した角度に PTZ カメラを向けさせることによって，人物の追跡をしている．

カメラの視線方向は-100~+100度を5度ずつに区切った41段階に分けられている．また，パノラマ画像の横の大きさは640pixel，縦の大きさは142pixelであり，2台のPTZカメラはそれぞれ321pixel分の領域を担当している．1台目のカメラが0~320pixel，2台目のカメラが320~640pixelの領域を担当することによって，周囲360度全ての範囲をPTZカメラで監視することも可能である(図4-1)．

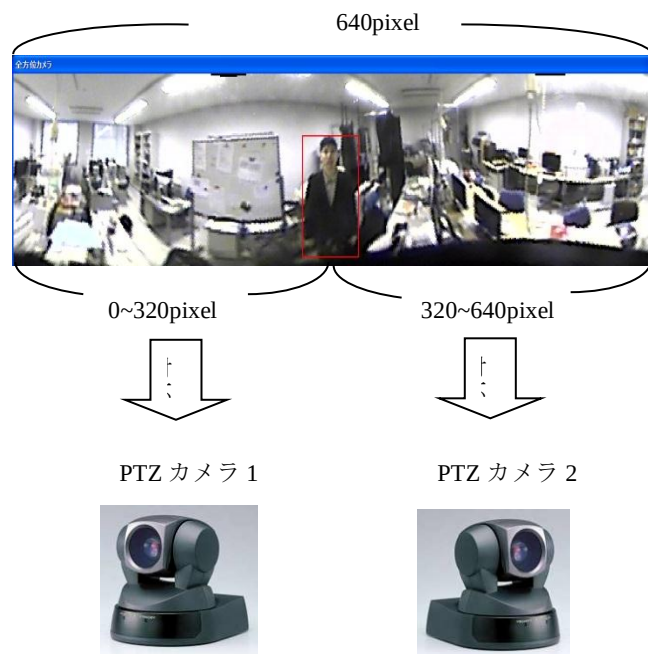


図 4-1 PTZ カメラの担当領域の一例

第5章 Zoom 制御

画像上の人物の大きさは、マーケティングや検索などに利用することを考えると、一定以上の大きさでほぼ同じ大きさになることが望ましい。そこで本研究では、PTZ カメラの Zoom 機能を用いて、画像上の人物の大きさを調整できるようにしてある(図 5-1)。



図 5-1 Zoom 適用前(左図)と Zoom 適用後(右図)の比較画像

Zoom 倍率はパノラマ画像上で検出された人物の領域の中心行座標によって決めている。行座標の値はカメラから人物までの距離によって変動するので、本研究ではこの値を利用して Zoom 倍率を求めている。具体的な手順は以下の通りである。

1. ある行座標 row に人物がいる時の Zoom 倍率を mag とする。(この Zoom 倍率は、使用者の利用目的によって異なるため、予め求めておく任意の倍率となる)
2. 人物が行座標 row' に居るときの Zoom 倍率 mag' は以下の式(5.1)によって決定される。(142 という数値は、パノラマ画像の縦の pixel 数である。)

$$mag \times 142 - row' / row = mag' \quad (5.1)$$

以上の式で求められた Zoom 倍率 mag' を PTZ カメラに伝えることにより、大きさの調整をする。

第6章 オクルージョン問題

カメラを用いて複数の人物を追跡する場合には、必ずオクルージョンが発生する[12]

オクルージョンとは、カメラと追跡対象の位置関係により物体の見え方が変化してしまう現象であり、自己オクルージョンと相互オクルージョンの2種が存在する[13].

6.1 自己オクルージョン

自己オクルージョンとは、カメラに対する物体の姿勢変化により、映像内の物体形状が変化してしまうことである.

これにより、形状を利用して人物の追跡や認証などを行うシステムは、その役割を果たせなくなってしまうことがある.

6.2 相互オクルージョン

相互オクルージョンとは、追跡対象が人や物陰に入ってしまう、カメラから遮断されてしまう状態になることである.

これにより、複数の人物を追跡する際に各カメラと各追跡対象の対応付けが困難となり、検出・追跡失敗の可能性を高めてしまう[14].



図 6-1 相互オクルージョン

この相互オクルージョンへの対策を、第7章にて解説する.

第7章 対応付け

オクルージョン問題への対作としては、各カメラと各追跡対象を対応付けしておくという方法と3次元空間をコンピュータ上に作りだしオクルージョンの影響を受けないようにその仮想空間上で追跡するという方法がある。

しかしながら、2次元画像から3次元空間を作成することは困難であり、多くの事前準備が必要となってしまう[9]。

本研究では汎用性が高く、誰にでも使えるシステムの作成を目指し、Bhattacharyya 係数による類似度判定[15]と座標軌跡からの人物の移動先を予測することで、各カメラと各追跡対象を対応付けし、オクルージョンへの対策をすることにした。

この手法では、Bhattacharyya 係数と座標軌跡からの予測という2つの結果を総合的に判断し最終的な対応付けを決定しているので、信頼性の高い対応付けが可能となっている。

7.1 Bhattacharyya 係数

Bhattacharyya 係数 ρ は 2 つの画像の類似度を判定出来る数式であり，以下の式(7.1)で表せる[16].

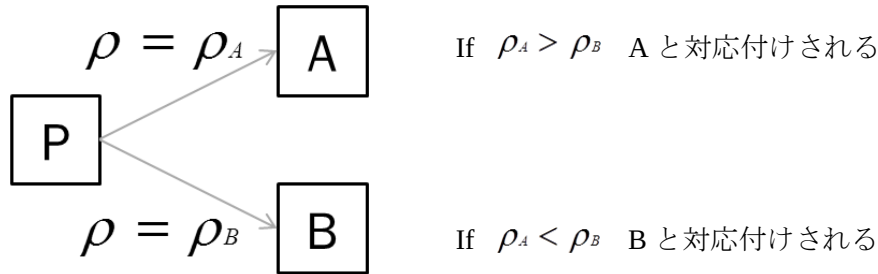
このとき， p_u と q_u は正規化された色相分布であり，以下の式(7.2)のように表せる.

また， m は色相の成分総数であり，0~255 を 4 ずつに区切った 64 段階となっている． u は色相の成分番号である．

$$\rho = \sum_{u=1}^m \sqrt{p_u q_u} \quad (7.1)$$

$$\sum_{u=1}^m p_u = 1 \quad , \quad \sum_{u=1}^m q_u = 1 \quad (7.2)$$

これを利用して，追跡中の人物の色相分布と検出された領域の色相分布の類似度を判定し， ρ の値がもっとも大きくなる領域と対応付けするという仮説(図 7-1)を立てる．



P: 追跡対象

A, B: 検出領域

図 7-1 Bhattacharyya 係数を利用した対応付け

7.2 座標軌跡からの予測

各フレームでの人物の検出座標を比較することにより，人物の進行方向および進行速度を求めることが出来る．

それを利用することにより，次フレームでの人物の検出位置を絞ることが可能となる．具体的な手順は，以下の通りである．

1. 検出された人物の座標を保存する．
2. 前フレーム $f_{t-1}(m,n)$ と前々フレーム $f_{t-2}(m,n)$ の座標を比較し，移動距離 c と進行方向を求める． c は以下の式(7.3)で求めている．また，進行方向は a の符号によってパノラマ画像上で左右のどちらに進んでいるかを判断している．

$$\sqrt{a^2 + b^2} = c \quad (7.3)$$
$$a = (f_{t-1}(m) - f_{t-2}(m))$$
$$b = (f_{t-1}(n) - f_{t-2}(n))$$

m : 行座標 n : 列座標

進行方向と移動距離から移動可能範囲(図 7-2)を導きだし，現フレームで検出されたどの領域に過去フレームで検出されたどの人物が対応するかの仮説を立てる．ただし，予測領域内に複数の検出領域がある場合は，仮説は立てないものとする．

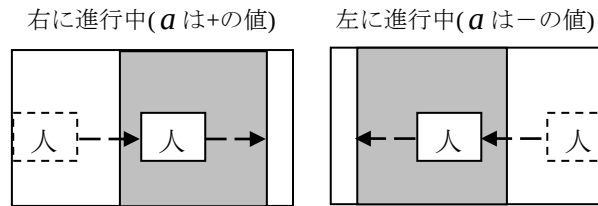


図 7-2 人物の移動可能範囲の予測

7.3 対応付けの手順

この節では、対応付けをする必要が発生する最小人数である 2 人のときを例として Bhattacharyya 係数と座標軌跡からの予測を用いた対応付けの手順を説明する.

対象となる人物を A と B, 候補となる領域を X と Y と定義したとき, 以下のような判断基準のもと, 対応付けをする.

なお, このときの数式の意味は以下の通りである.

$H^{(h)}(X \rightarrow A)=1$: Bhattacharyya 係数により, X が A に対応するという仮説が立っている.

$H^{(h)}(Y \rightarrow A)=1$: Bhattacharyya 係数により, Y が A に対応するという仮説が立っている.

$H^{(h)}(Y \rightarrow B)=1$: Bhattacharyya 係数により, Y が B に対応するという仮説が立っている.

$H^{(t)}(X \rightarrow A)=1$: 座標軌跡からの予測により, X が A に対応するという仮説が立っている.

$H^{(t)}(X \rightarrow B)=1$: 座標軌跡からの予測により, X が B に対応するという仮説が立っている.

$H^{(t)}(X \rightarrow A)=0$: 座標軌跡からの予測で仮説を立てられていない.

1. $H^{(h)}(X \rightarrow A)=1$ かつ $H^{(t)}(X \rightarrow A)=1$ のとき

X は A に, Y は B に対応付けされる.

(Bhattacharyya 係数と座標軌跡からの予測が同一の仮説を立てたとき, 上記と同様の結果となる)

2. $H^{(h)}(X \rightarrow A)=1$ かつ $H^{(t)}(X \rightarrow B)=1$ のとき

$H^{(h)}(Y \rightarrow A)=1$ ならば

X は B に, Y は A に対応付けされる.

$H^{(h)}(Y \rightarrow B)=1$ ならば

X は A に, Y は B に対応付けされる.

(Bhattacharyya 係数と座標軌跡からの予測が異なる仮説を立てたとき, 上記と同様の結果となる)

3. $H^{(t)}(X \rightarrow A)=0$ のとき

$H^{(h)}(X \rightarrow A)=1$ かつ $H^{(h)}(Y \rightarrow B)=1$ のとき

X は A に, Y は B に対応付けされる.

$H^{(h)}(X \rightarrow A)=1$ かつ $H^{(h)}(Y \rightarrow A)=1$ のとき

対応付けは行わず、次フレームに繰り越す.

第8章 フローチャート

本研究で使われているシステムの処理フローは以下の通りになる。

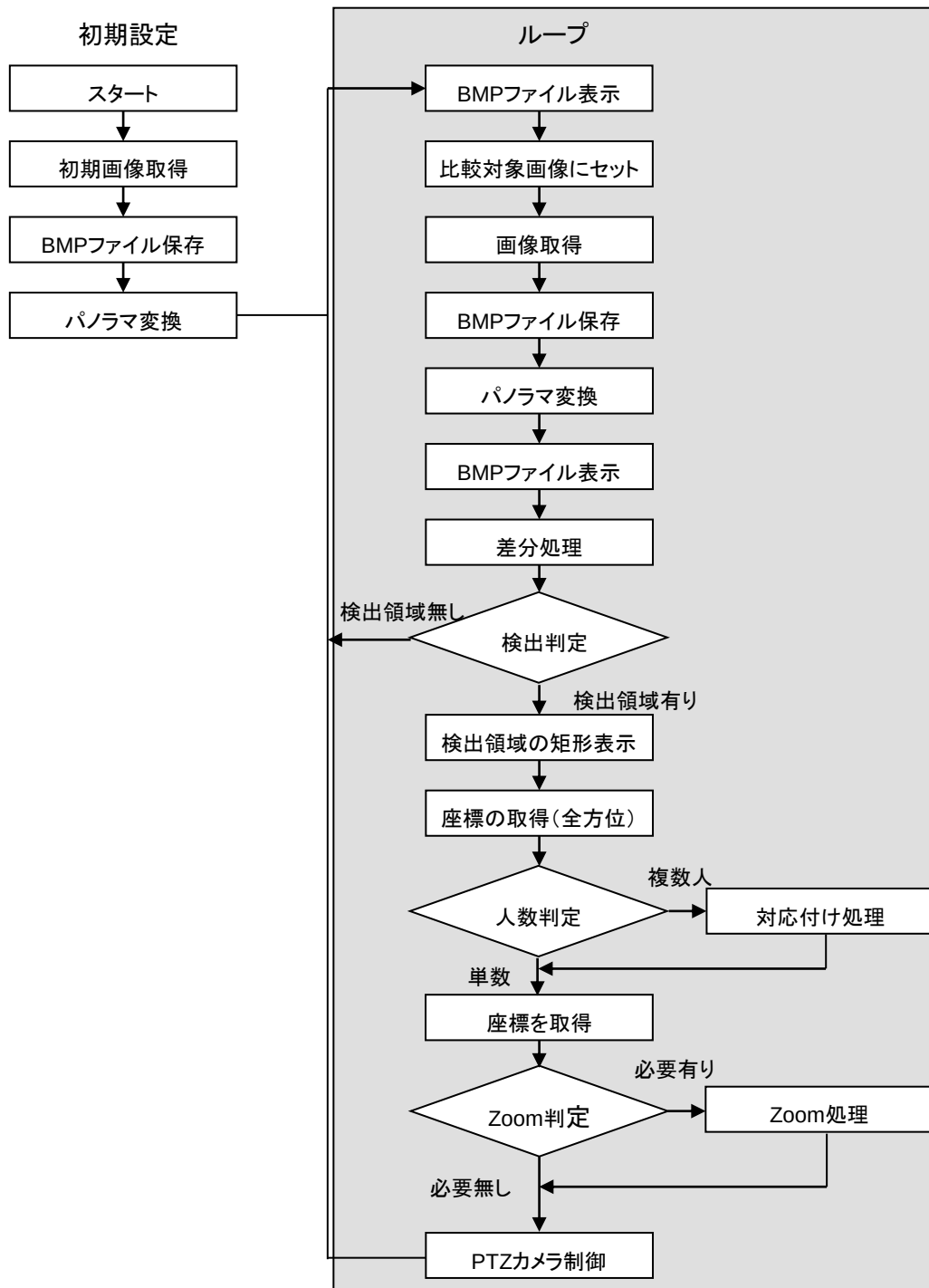


図 8-1 フローチャート

第9章 評価実験

本システムの有効性を検証するために、9.1 および 9.2 に示す 2 つの評価実験を行った。それぞれの実験について、実験内容、評価方法、実験結果、実験風景、考察の 5 つの項目でまとめている。

9.1 追跡性能および Zoom の有効性の評価

9.1.1 実験内容

この節では、本システムの追跡性能および Zoom の有効性を評価する実験について記載する。

実験の内容は、以下の通りである。

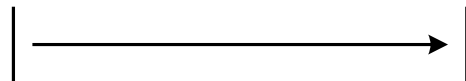
室内を 1 人の人物が「通常歩行」「早歩き」「一時停止」「U ターン」の 4 種類の歩行パターン(図 9-1)で歩き、それを追跡する。

実験回数は Zoom をしなかった場合と Zoom をした場合をそれぞれ 100 回ずつとし、1 つの歩行パターンに対して計 200 回の追跡を行なった。PTZ カメラから追跡中の人物が完全にフレームアウトした場合を失敗とし、成功率についてまとめた。

また、Zoom の有効性を評価するために、追跡結果だけでなく、撮影された人物の顔の大きさ(200 個分)を調べ、散布図としてまとめた。

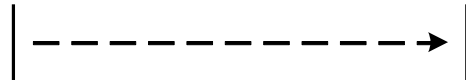
<通常歩行>

任意の速度で普段どおりに歩く。



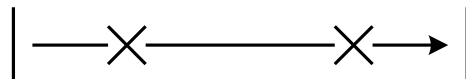
<早歩き>

通常歩行よりも早い速度で歩く。



<一時停止>

2 地点間中の 2 ヶ所で一時停止する。



<U ターン>

2 地点間の 1/2 で折り返す。



図 9-1 歩行パターン

9.1.2 評価方法

評価実験を行うにあたり，実験の成功・失敗条件をあらかじめ決め，これをもとに評価を行った．

「実験の成功条件」

PTZ カメラによる移動物体追跡が一部分でも追跡し続けられたとき．

「実験の失敗条件」

PTZ カメラによる移動物体追跡が一度でも完全にフレームアウトしてしまったとき

1 回の試行において PTZ カメラによる移動物体追跡が最後まで行えれば成功，それ以外(追跡途中で途切れる等)は失敗という条件で評価を行った．

全方位カメラによる移動物体追跡について考慮に入っていないのは，カメラの特性上移動物体を検知出来ないことがまず無いためである．

9.1.3 実験結果

実験の成功率を表 9-1 から 9-2 および図 9-2 から 9-5 にまとめた。

また、図 9-6 および 9-7 は顔の大きさを調べたものの結果であり、図 9-8 および 9-9 は実験の成功時と失敗時の画像である。

成功率は、各歩行パターンを 10 測定毎に表記し、最終的な平均成功率を最下段にまとめている。

表 9-1 Zoom 機能未使用時の成功率

実験回数	通常歩行	早歩き	一時停止	U ターン
1～10	90%	80%	90%	80%
11～20	80%	90%	90%	90%
21～30	90%	70%	90%	90%
31～40	100%	70%	90%	80%
41～50	90%	80%	80%	80%
51～60	100%	90%	100%	90%
61～70	80%	80%	90%	80%
71～80	90%	70%	90%	90%
81～90	90%	70%	80%	90%
91～100	90%	80%	80%	80%
平均	90%	78%	88%	85%

表 9-2 Zoom 機能使用時の成功率

実験回数	通常歩行	早歩き	一時停止	U ターン
1～10	100%	80%	100%	90%
11～20	100%	90%	90%	80%
21～30	90%	80%	100%	100%
31～40	100%	80%	90%	80%
41～50	90%	90%	90%	90%
51～60	100%	80%	100%	90%
61～70	90%	70%	90%	90%
71～80	90%	80%	90%	90%
81～90	100%	90%	90%	90%
91～100	90%	80%	90%	80%
平均	95%	82%	93%	88%

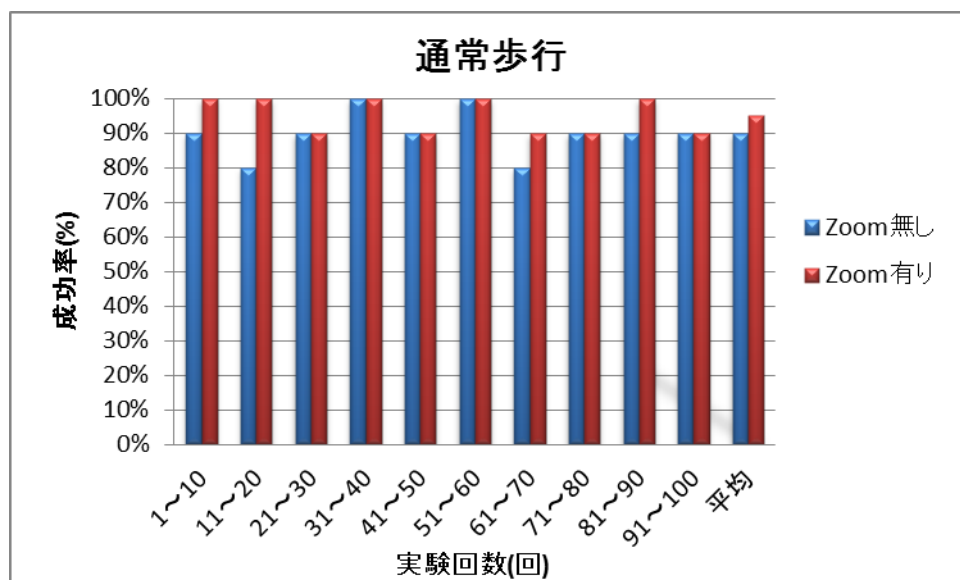


図 9-2 通常歩行の成功率

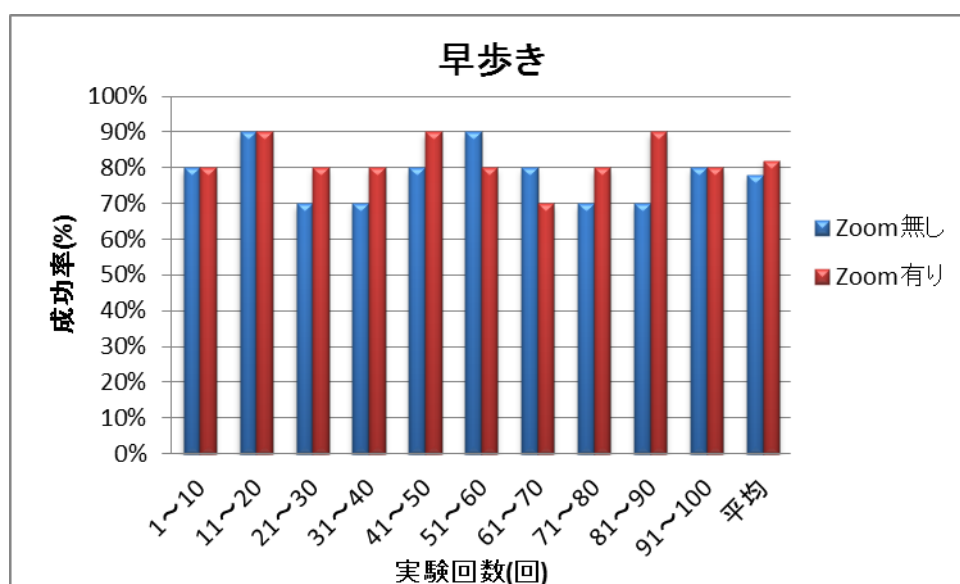


図 9-3 早歩きの成功率

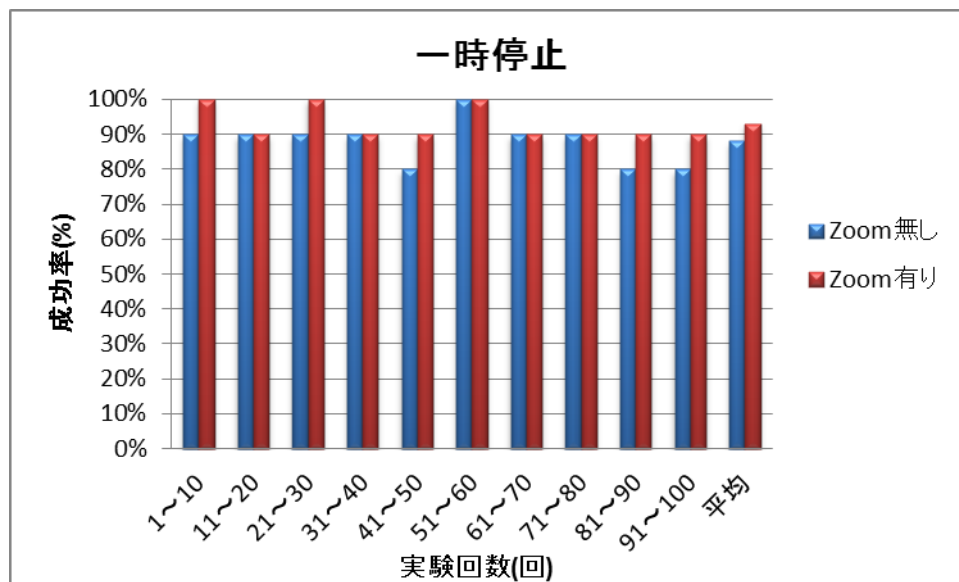


図 9-4 一時停止の成功率

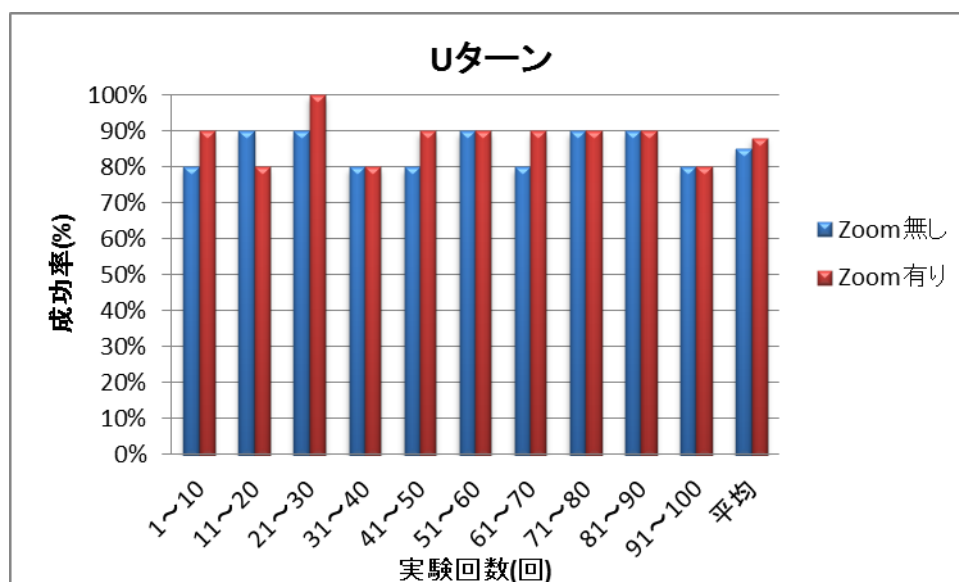


図 9-5 Uターンの成功率

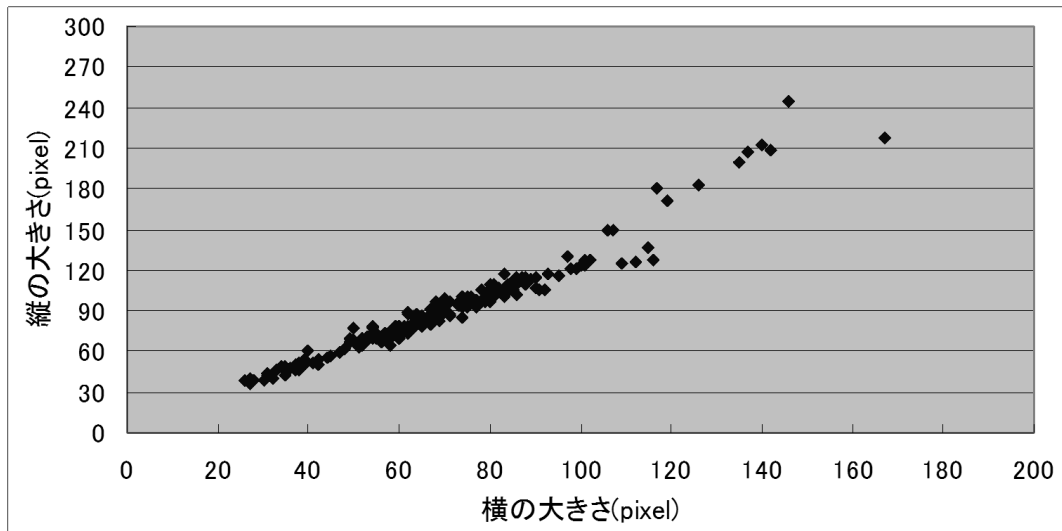


図 9-6 顔の大きさのばらつき (Zoom 無し)

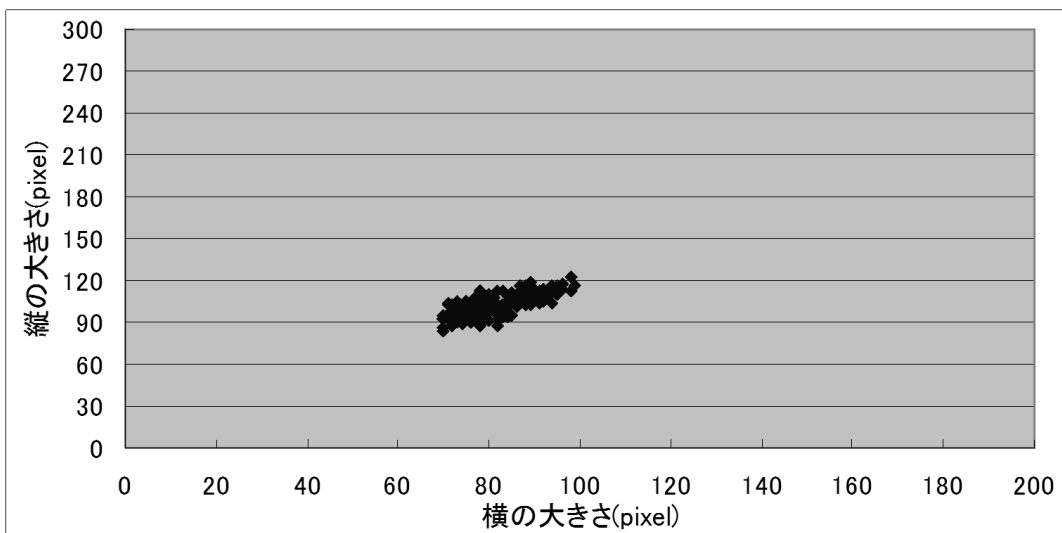


図 9-7 顔の大きさのばらつき (Zoom 有り)



図 9-8 追跡成功時の画像



図 9-9 追跡失敗時の画像

9.1.4 考察

まず図 9-6 および 9-7 から本システムは Zoom 機能を使用し、追跡対象の見た目の大きさをほぼ一定の大きさに出来ていることが分かる。

そして、これらの実験結果より Zoom 機能を利用し見た目の大きさを調整したことで、追跡精度の向上に繋がることが分かった。

これは、見た目の大きさを調整し、PTZ カメラに対する追跡対象の見た目の速度(追跡対象がカメラに近ければ近いほど、カメラ上での人物の大きさは大きくなり見た目の速度は速くなる)を一定化したことにより、フレームアウトする確率を下げた成果であると考えられる。

しかしながら、通常歩行、一時停止、U ターンと比べ、早歩きは追跡精度が低いので、まだ追跡対象の速度への対策が不十分であることが分かる。

この追跡失敗の原因の多くは、全方位カメラ上での追跡は正しく行えていたが、PTZ カメラの首振り速度が間に合わずフレームアウトしてしまうものであった。

ゆえに、解決策としてはプログラムを改良しコンピュータから PTZ カメラへの指令を早くする、PTZ カメラの設置台数を増やし 1 台毎の監視領域の幅を狭め負担を軽減するなどを考えている。

9.2 オクルージョンに対する頑健性の評価

9.2.1 実験内容

この節では、本システムのオクルージョンに対する頑健性を評価する実験について記載する。

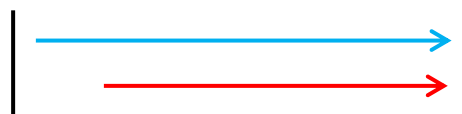
実験の内容は、以下の通りである。

室内を2人の人物が「通常歩行」「Uターン」「交差」「追い越し」の4種類の歩行パターン(図9-10)で歩き、それを追跡する。なお、Uターンをするタイミングは、カメラに映る2人の人物が重なった時とした。

実験回数は、7章で示したオクルージョンへの対策として対応付けを行った場合と対応付けを行わなかった場合をそれぞれ100回ずつとし、1つの歩行パターンに対して計200回の追跡を行ない、成功率についてまとめた。

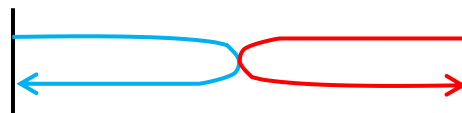
<通常歩行>

二人の人物が重ならないように、
任意の速度で普段どおりに歩く。



<Uターン>

両端から歩き出し、
2地点間の1/2で折り返す。



<交差>

両端から歩き出し、
もう片方の端を目指す。



<追い越し>

部屋の中央付近で、先に歩いている人物を
もう一人の人物が追い越す。

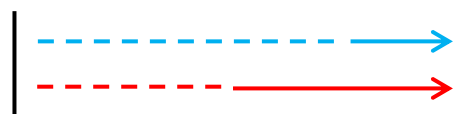
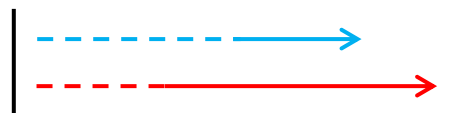


図9-10 歩行パターン

9.2.2 評価方法

評価実験を行うにあたり，実験の成功・失敗条件をあらかじめ決め，これをもとに評価を行った．

「実験の成功条件」

各 PTZ カメラと各追跡対象との対応付けに誤りがないまま，追跡対象の一部分でも追跡し続けられたとき．

「実験の失敗条件」

PTZ カメラによる移動物体追跡が一度でも完全にフレームアウトしてしまったとき．

各 PTZ カメラと各追跡対象との対応付けに誤りが生じてしまった場合．

1 回の試行において PTZ カメラによる移動物体追跡が最後まで行え，各 PTZ カメラの対応付けに誤りがなかった場合を成功とし，それ以外(追跡途中で途切れる，対応付けに誤りが生じる等)は失敗という条件で評価を行った．

全方位カメラによる移動物体追跡について考慮に入れていないのは，カメラの特性上移動物体を検知出来ないことがまず無いためである．

9.2.3 実験結果

実験の成功率を表 9-3 から 9-4 および図 9-11 から 9-14 にまとめた。

また，図 9-15 から 9-18 は実験の成功時と失敗時の画像である。

成功率は，各歩行パターンを 10 測定毎に表記し，最終的な平均成功率を最下段にまとめている。

表 9-3 オクルージョン対策無しの成功率

実験回数	通常歩行	U ターン	交差	追い越し
1～10	100%	40%	30%	50%
11～20	90%	40%	40%	40%
21～30	90%	30%	40%	50%
31～40	80%	50%	30%	50%
41～50	90%	50%	30%	60%
51～60	90%	30%	30%	50%
61～70	100%	50%	30%	40%
71～80	90%	40%	30%	50%
81～90	100%	40%	20%	50%
91～100	90%	50%	30%	50%
平均	92%	42%	31%	49%

表 9-4 オクルージョン対策有りの成功率

実験回数	通常歩行	U ターン	交差	追い越し
1～10	100%	90%	90%	90%
11～20	100%	80%	90%	100%
21～30	90%	90%	70%	80%
31～40	90%	90%	80%	90%
41～50	90%	80%	70%	80%
51～60	80%	90%	80%	80%
61～70	90%	70%	90%	90%
71～80	100%	90%	80%	90%
81～90	90%	80%	80%	90%
91～100	90%	70%	90%	80%
平均	92%	83%	82%	87%

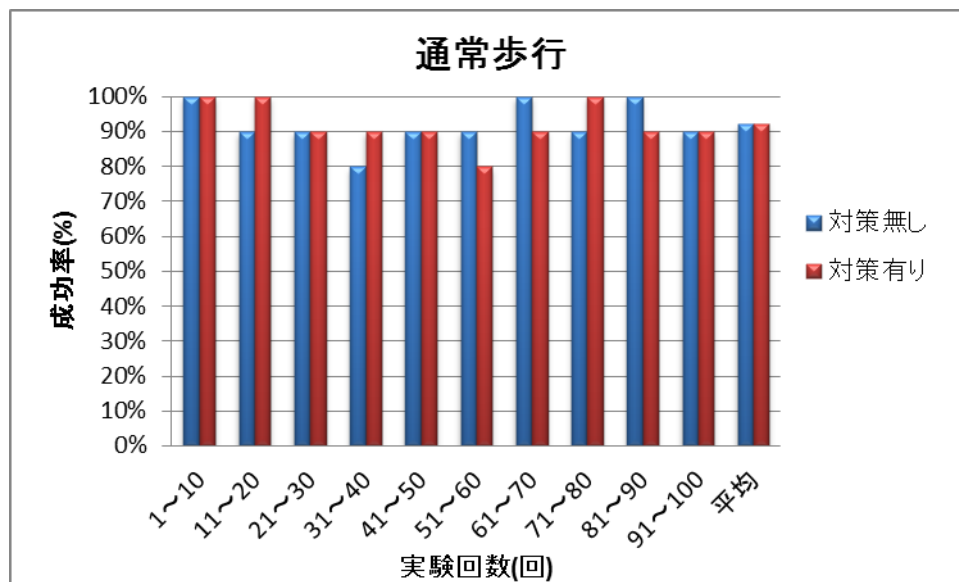


図 9-11 通常歩行の成功率

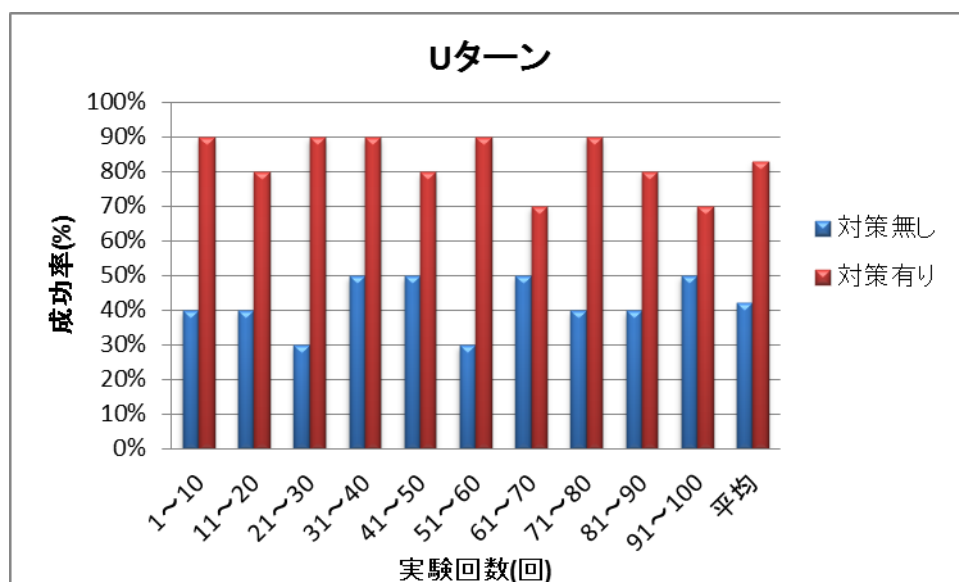


図 9-12 U ターンの成功率

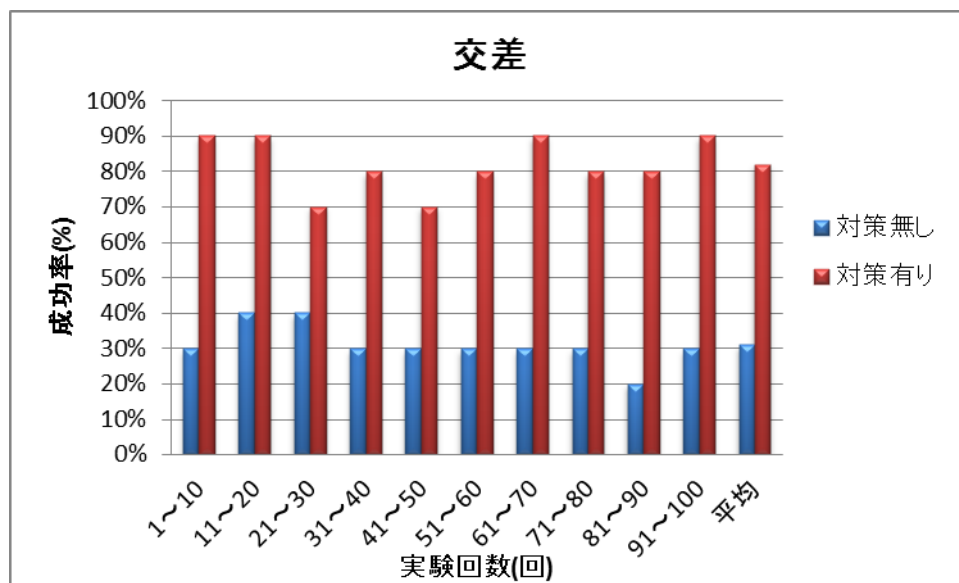


図 9-13 交差の成功率

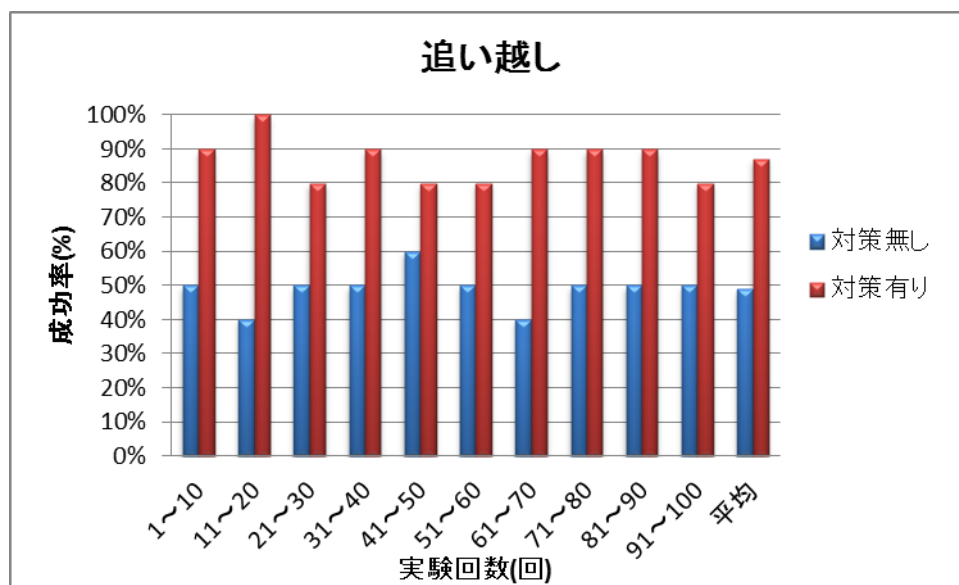


図 9-14 追い越しの成功率



図 9-15 追跡成功時の画像(オクルージョン無し)



図 9-16 追跡成功時の画像(オクルージョン有り)



図 9-17 追跡失敗時の画像(フレームアウト)



図 9-18 追跡失敗時の画像(対応付け失敗)

9.2.4 考察

実験結果より，オクルージョンが発生する場合，提案手法を用いることによって追跡精度を大幅に向上させられることが分かった．

オクルージョンへの対策をしなかった場合での追跡成功は，人物の位置関係に変化が少なかったため(カメラ上で左の人物は左のまま，右の人物は右のまま)，擬似的に各カメラと各追跡対象が対応付けされたためであると考えられる．そのため，位置関係が逆になる交差は，通常歩行，Uターン，追い越しと比べ，追跡精度が低くなってしまったと考えている．

また，通常歩行の場合，オクルージョンが発生しないためオクルージョンへの対策のあるなしに関わらず，同様の追跡精度となった．これにより，提案手法の実装はシステムの基本性能を低下させずに可能なことが分かった．

追跡失敗の原因の多くは，対応付けの誤りであった．提案手法により追跡精度の大幅な向上に成功はしたが，対応付けの主たる部分は色情報に頼っているので誤りが生じやすいと考えられる．

そのため，追跡対象が同系色の服装をしていた場合や，夜間や照明のない薄暗い室内などではより追跡精度が低下してしまうと考えられる．

この対応付けの誤りへの対策としては，形状や歩容など色情報以外の特徴での判定を加えることを考えている．

第10章 結び

本稿では、オクルージョンに対して頑健な人物の検出および追跡手法を提案し、1台の全方位カメラと2台のPTZカメラを連動させることにより、広範囲における自動追尾機能付きの人物監視システムの構築が出来た。

そして、オクルージョンへの対策をしている提案手法により、オクルージョン発生下での複数人物の同時追跡にも成功している。このことから、提案手法はオクルージョンに対して有効な対策であると考えられる。

また、撮影する追跡対象の見た目の大きさを一定化することによって、PTZカメラから追跡対象がフレームアウトしてしまう確率を下げることに成功している。

しかしながら、追跡対象の歩行速度が速い場合や、同系色の衣服を着用していた場合に、追跡性能が低下してしまう可能性が高いことも分かっているため、今後はより悪条件の実験を繰り返し、何事に対しても対応出来るシステムを作成することが課題となる。

謝辞

本研究を進めるにあたり，ご指導ならびに貴重なご意見をいただいた，法政大学理工学部応用情報工学科 赤松茂教授に心より感謝致します．

また，本研究の実験に協力していただいた法政大学赤松研究室の皆様に深く感謝申し上げます．

参考文献

- [1] 佐藤洋介, 橋本浩二, 柴田義孝: “全方位カメラを用いた追跡監視カメラシステム”, 情報処理学会研究報告. マルチメディア通信と分散処理研究会報告 2008(54), pp.99-104, 2008-06-12
- [2] 池田真也, 大久保寛, 田川憲男: “動画像での動的なラベル割当に基づく間接的な物体追跡”, 映像情報メディア学会技術報告, 32, pp.69-72(2008)
- [3] 山内悠嗣, 藤吉弘亘, HwangBoon-Woo, 金出武雄: “アピアランスと時空間特徴の共起に基づく人検出”, 画像の認識・理解シンポジウム(MIRU2007), pp1492-1497(2007)
- [4] 藤吉弘亘, 山下隆義: “画像特徴量 [V] -勾配に着目した画像特徴量と人物追跡-”, 電気情報通信学会誌, 94, 3, pp.239-244, 2011-03-01
- [5] 安全な暮らし・街頭防犯カメラシステム (警視庁ホームページ), <http://www.keishicho.metro.tokyo.jp/anzen/sub0.htm>
- [6] 濱島京子, 呂健, 石原浩二: “全方位視覚センサによる移動体存在領域検出手法” 産業安全研究所特別研究報告 NIIS-SRR, 33, pp.25-41(2005)
- [7] 濱山内葉: “全方位カメラと PTZ カメラの協調による移動物体追跡システムの構築” 法政大学工学部システム制御工学科 2008 年度卒業論文
- [8] Y. Kobayashi, D. Sugimura, Y. Sato, H. Hirasawa, N. Suzuki, H. Kage, and A. Sugimoto, "3D Head Tracking using the Particle Filter with Cascaded Classifiers," British Machine Vision Conference, pp.37-46, 2006.
- [9] Tatsuya. Osawa, Xiaojun. Wu, Kaoru. Wakabayashi, and Takayuki. Yasuno, "Human Tracking by Particle Filtering Using Full 3D Model of Both Target and Environment," 18th International Conference on Pattern Recognition, Vol2, pp.25-28, 2006.
- [10] 松村遼, 岡村健史郎: “パーティクルフィルタを用いた進入検出と物体追跡”, 国立高等学校機構大島商船高等専門学校紀要, 41, pp75-85, 2010-03-09
- [11] 藤吉弘亘, 小村剛史, 矢入(江口)郁子, 香山健太郎, 吉水宏: “歩行者 ITS のためのフレーム間差分による移動体検出法とその評価”, 情報処理学会論文志. コンピュータビジョンとイメージメディア, 45, pp.11-20(2004)
- [12] 岩城 徹哉, 武川直樹, 久保田吉彦: “オクルージョン推定に基づく領域分割ベースステレオビジョン -領域統合による奥行き傾斜の推定-, 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解 106(605), 115-120, 2007-03-08
- [13] 濱島京子, 呂健, 石原浩二: “全方位視覚センサによる移動物体存在領域検出手法”, Specific Research Reports of the National Institute of Industrial Safety, NIIS-SRR-NO.33, 2005

- [14] 齋藤哲平：“2 台のカメラを用いた複数人物を対象とした単体追跡システムの構築”，法政大学大学院工学研究科システム工学専攻修士課程 2010 年度修士論文
- [15] 赤塚久哉，高橋友和，井手一郎，村瀬洋：“不連続な撮影領域間の人物軌跡の対応付けに関する検討”，電子情報通信学会総合大会講演論文集 2006 年_情報・システム(2), 198, 2006-03-08
- [16] 星川佑磨，中西達也，橋本優希，寺林賢司，梅田和昇：“色情報と差分ステレオを組み合わせた人物トラッキング”，2009 年度精密工学会秋季大会学術講演論文集，2009-09