

NAIST-IS-MT9951007

修士論文

高解像度全方位画像センサを用いた遠隔監視システム のためのイベント検出

市川 徹

2001年2月9日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

市川 徹

審査委員： 横矢 直和 教授
小笠原 司 教授
竹村 治雄 助教授

高解像度全方位画像センサを用いた遠隔監視システム のためのイベント検出*

市川 徹

内容梗概

遠隔地の監視を目的とした遠隔監視システムの開発は従来より盛んに行なわれてきた。監視システムへの主な要求は監視者の負担軽減やセキュリティ確保であり、そのために広範囲な情景を詳細に獲得する必要がある。さらに、近年では、環境内を撮像、記録するだけでなく、その中で起こるイベントの検出や行動の解析といった高度な機能も要求されるようになっている。

本研究では高解像度で撮像可能な全方位画像センサを使用し、監視環境内におけるイベント検出を行なう手法を提案する。提案手法では、処理毎に解像度を変更して対象物の検出・追跡を行ない、対象物の動き解析の結果に基づいたイベント検出を行なう。この手法は特殊な背景知識や前提条件を必要とせずイベント検出を行える。また、常に広範囲の情景を高解像度で監視することが可能であり、環境内の多地点で起こる複数のイベントを同時に検出・提示できる。

提案したイベント検出手法による実験では、監視環境内に存在する2人の人物を1/3sの更新間隔で、監視・追跡すると同時に各人物の動きを解析し、発生するイベントの検出、提示をすることができた。この結果より、本手法は監視ビデオの要約や編集への適用が可能であり、監視システムの応用のための有効な手段であると考えられる。

キーワード

監視システム, 全方位画像センサ, 行動解析, イベント検出, 透視投影画像

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT9951007, 2001年2月9日.

Event Detection for a Video Surveillance System Using a High-resolution Omnidirectional Sensor*

Toru Ichikawa

Abstract

For the purpose of remote surveillance, many video surveillance systems has been developed. A main goal of the indoor surveillance is to reduce the burden of surveyor and maintain the security. To make up these conditions, It is essential to acquire the whole environmental video streams in detail. In addition to that, it is required to detect the events or summarize the activity taking place in an environment.

This paper presents a method of event detection using a high-resolution omnidirectional sensor. The proposed method can detect and track objects by changing resolution and can detect general events without particular information or assumptions based on objects' motion. It has the advantage of acquiring omnidirectional video streams of whole environment in high-resolution and detecting multiple events taking place in different places.

Experiments have proved that proposed method can survey and monitor 2 objects with the speed of 3 frames per second, and can detect and display their events. These results show the possibility of applying this method to summarize or edit the video streams.

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT9951007, February 9, 2001.

Keywords:

surveillance system, omnidirectional image sensor, behavior analysis, event detection, perspective image

目次

1. はじめに	1
1.1 遠隔監視システム	1
1.2 イベント検出の従来手法	3
1.3 本研究の概要	4
2. 高解像度全方位画像センサ	6
2.1 高解像度全方位画像センサの構成	6
2.2 高解像度全方位画像センサの特徴	8
3. 対象物の検出と追跡	12
3.1 対象物の検出	12
3.1.1 背景差分による対象物領域の抽出	13
3.1.2 領域情報の取得	15
3.2 対象物の追跡	15
3.2.1 重心間距離による対応付け	17
3.2.2 カラーヒストグラムを用いた対応付け	18
3.2.3 時系列間での対象物領域の同定	19
3.2.4 監視者への提示	21
4. 対象物の動き解析とイベント検出	26
4.1 対象物の動き解析	26
4.2 イベント検出	27
5. イベント検出実験	31
5.1 実験設定	31
5.2 実験結果	33
5.3 考察	34
6. おわりに	39

謝辭 41

參考文獻 42

目 次

1	全方位画像センサとパン/チルト/ズームカメラによるシステム	3
2	高解像度全方位画像センサ	8
3	高解像度全方位画像センサの視覚系の構成	9
4	高解像度全方位画像センサによる全方位画像と透視投影画像	10
5	HyperOmni Vision による全方位画像と透視投影画像	10
6	異なるズームにおける全方位画像の違い	11
7	背景差分による領域抽出	14
8	ラベリング処理と面積による切捨て処理	15
9	同系色による対象物の分割	16
10	フレーム間での対象物の同定(なし 単一対象物領域)	22
11	フレーム間での対象物領域の同定(単一対象物領域 単一対象物領域)	22
12	フレーム間での対象物領域の同定(単一対象物領域 なし)	23
13	フレーム間での対象物領域の同定(複数の対象物領域 単一対象物領域)	23
14	フレーム間での対象物領域の同定(単一対象物領域 複数対象物領域)	24
15	領域間の八近傍距離による分離・分割の判定	24
16	全方位画像と透視投影画像の対応点計算	25
17	監視者への提示画像	25
18	動き解析の処理の流れ	29
19	システム構成	32
20	監視環境	33
21	監視環境の全方位画像	34
22	対象物が侵入したときの入力画像と提示画像の一例	35
23	イベントの検出例(出現イベント)	36
24	イベントの検出例(退出イベント)	36
25	イベントの検出例(静止イベント)	37

26	イベントの検出例 (作業イベント)	37
27	イベントの検出例 (オブジェクトイベント)	38

表 目 次

1	高解像度全方位画像センサの仕様	7
2	検出された対象物の領域情報	17
3	イベントの設定	28
4	検出したイベント数の比較	38

1. はじめに

遠隔地の環境をビデオカメラなどの画像センサで撮影し、環境内の情景を監視者に提示したり、記録する遠隔監視システムの開発は従来より盛んに行なわれている [1, 2, 3]. こうした遠隔監視システムにおいて重要なことは、監視環境全体のセキュリティをどのように確保するか、また、遠隔地の情報をいかに効率良く監視者に伝えてその負担を軽減するかである。そのためシステムに要求される基本的な要求として、(1) 常時、広範囲の監視 (2) 対象物の自動検出といったことが挙げられる。しかし、制御すべきカメラの台数が増加したり、環境内の注視すべき対象物が増加すると、それらに適切に対処することは難しくなり、たとえ上述した2つの要求を満たしたシステムであっても、環境内の対象物をすべて監視することは監視者への負担が大きい。また、記録時間が長くなれば、膨大なビデオストリームの中から重要な情報を効率良く得ることも困難となる。

そこで、近年では、画像センサによって得られる監視環境内の情報を解析し、環境内に存在する人物の行動解析や環境内で発生するイベントの自動検出によって、必要な情報だけを監視者に提示する機能や、重要な情報のみを記録・編集したビデオストリーム要約の自動作成といった高度な機能への要求が高まっている。

まず、1.1節では、従来の遠隔監視システムとその問題点について述べ、1.2節では行動解析やストリーム要約のためのイベント検出に関する研究について述べた後、1.3節で本研究で提案する監視システムのためのイベント検出手法の概要について述べる。

1.1 遠隔監視システム

銀行やオフィスでの防犯や、エレベーターの監視、また、駐車場などの入出庫管理など、ある環境内の状況を遠隔地において監視する遠隔監視システムでは、ビデオカメラなどの画像センサで取得した環境内の様子を遠隔地の監視者に提示するシステムが一般的である。こうした遠隔監視システムの現在の主流は、通常の CCD カメラを用いて環境を監視する方法である。しかし、通常の CCD カメラに使用されるレンズは、通常、画角が 30 – 40 度程度であり視野に制限を受ける。そ

のため、1 台で広範囲な環境全体を一度に監視することは不可能であり、この問題を解決するために、カメラを移動・回転させる方法、もしくは、カメラを複数台設置する方法が考えられる。

カメラを移動・回転させる方法では、監視者が手動でカメラの制御を行ない環境全体を監視するものと、対象物の動きを予測してカメラを自動で制御するもの [1, 2, 3] が存在する。カメラを手動で制御する場合、監視環境全体を把握するために常時カメラを動かす必要があり、監視者には負担となる。また、自動でカメラを制御する場合、ある対象物を追従している間は環境全体を監視することができなくなる。さらに、いずれの場合も環境内に複数の対象物が存在する場合には、各対象物を同時に追跡することは不可能であり、カメラの移動中に生じる死角も問題となる。

カメラを複数台設置する方法 [4, 5] に関しては、環境全体の情景を一度に取得することは可能となるが対象物の追跡や提示に関しては各カメラからの画像を切替える必要があり、手動では監視者の負担が大きくなる。自動で各カメラの画像を切替えるとしても、監視者が各カメラの位置関係を把握していなければ環境内全体の状況を理解することが困難である。

近年、これらの問題を解決するために、全方位画像センサを監視システムへ応用することが提案されている [6, 7, 8, 9, 10, 11]。全方位画像センサとは周囲 360 度の情景を一度に取得できる画像センサであり [12]、このセンサを用いることで環境全体を常に監視することが可能となる。しかし、一般に全方位画像センサは 1 台のカメラで周囲 360 度の情景を取得するため、その分解能は十分とは言えず、提案されるシステムの多くは、図 1 に示されるように全方位画像センサと pan/tilt/zoom 機能のついたカメラ (以下、PTZ カメラ) の組合せで構成され、全方位画像センサで環境内から注視対象物とその位置を検出し、PTZ カメラによって対象物の追跡・監視を行なうものである [6, 7, 8]。従って、上で述べた複数の注視対象物の追跡や画像提示などの問題は解決されずに残ることになる。



図 1 全方位画像センサとパン/チルト/ズームカメラによるシステム

1.2 イベント検出の従来手法

監視環境内で起こるイベントを検出するためには、環境内に存在する対象物を検出してその動きを解析する必要がある。ビデオストリーム内の人物の行動解析や行動認識に関する研究は、コンピュータービジョンの分野において盛んに行なわれている [13, 14, 15]。しかし、行動解析や行動認識といった研究の多くは、ある特定の行動に焦点を当て、その行動内に現れる一連の動きパターンをあらかじめテンプレートとして持っておき、ストリーム内からそのテンプレートの並びに類似するパターンを見つけ出すというアプローチをとる。こうしたテンプレートを作成するためには、個々の行動毎に動きの特徴や種類といった複雑な背景知識が必要であり、行動内のキーとなる動作の選定や類似度の評価が非常に重要な要素となる。また、検出される行動の環境内での発生位置や向きといった撮影条件もある程度周知のものとして扱われる。しかし、監視システムのような環境内の任意地点で起こる数種類のイベントを検出する場合には、テンプレートの設定が複雑なうえ、膨大な数になるためリアルタイムでの処理は困難となる。

その他、対象物の行動を認識する方法として特殊なモデルによるマッチングで

はなく、対象物の位置とその移動経路に基づいて行なう方法 [16] や環境内の対象物の動き解析と環境情報を組み合わせてイベントを検出する方法 [17, 18] が提案されている。しかし、これらは監視を行なう環境の間取りや構成要素といった環境情報を事前に取得しておかなければならない。

また、イベント検出の応用としてストリームの要約や編集を考えた場合、ストリームからサブストリームを切り出す方法として、シーンの切り替わりを検出する方法 [19] やオプティカルフローに基づいた、ストリーム内の動きの不連続点を検出する方法 [20] などが挙げられる。しかし、いずれの方法も連続性のある一連の動きをサブストリームとして切り出しを行なうだけであり、切り出されたサブストリームの行動認識は行なわない。このため、切り出された各サブストリームを理解するためには、人手による行動認識処理や先に述べたテンプレートを用いた処理が必要となり、本研究のような監視目的のリアルタイム処理を併せて行なう場合には適さない。

1.3 本研究の概要

本研究では、画像センサによって遠隔地の環境を広範囲に渡って監視し、環境内で発生する重要な情報を監視者に提示したり、記録された監視ビデオのストリーム要約を行なうことを目的とする。そのために環境内の多地点で同時に発生する複数のイベントを検出する手法について提案する。

使用する画像センサに関しては、広範囲な環境を常時、監視できることが望ましい。そのとき、複数カメラ間での切替えやカメラの移動・回転といった監視者への負担を軽減する必要がある。そのため、1台のカメラで環境全体の情景を取得できる全方位画像センサのみを使用する。全方位画像センサには、球面ミラーを用いたもの [21]、円錐ミラーを用いたもの [22]、双曲面ミラーを用いたもの [23] などが考案されているが、本研究では、監視に必要な画像センサの側方領域と下方領域の情報を効率的に取得でき、透視投影という光学的特性を持つことで提示用の平面透視投影画像の生成が容易であるという2点から双曲面ミラーを用いた全方位画像センサを使用する。しかし、従来の CCD カメラを用いた全方位画像センサでは行動分析や画像提示に十分な解像度が得られないため、新たに、高解像度で撮

像可能な双曲面ミラーを用いた全方位画像センサを製作し、それを用いる。

イベント検出に関しては、広範囲な監視環境内の任意地点で起こるイベントの検出が必要であり、リアルタイムでの処理が要求される。また、システム構築の容易さを考慮すると、テンプレートの設定やマップ情報の事前取得が必要のない方法が適当と考え、[17, 18]らの動き解析に基づいた方法を拡張した広範囲な環境下でのイベント検出を行なう。

提案手法における処理の流れは、まず高解像度全方位画像センサで得られた入力画像から、注視すべき対象物を検出する。検出された個々の対象物について、重心距離、カラーヒストグラムを用いて時系列間に対応付けを行ない対象物を追跡する。さらに、時系列差分、対応する領域の有無、重心距離による位置変化を調べること対象物の動きを解析し、この結果をあらかじめ設定されたイベント条件と照らし合わせることにより、イベントの検出を行なう。同時に、検出された対象物の存在領域を入力画像から切り出し、透視投影画像に変換して追跡画像として提示する。以上の処理により、提案手法では常時広範囲な環境が監視でき、環境内に存在する対象物の検出と追跡・動き解析によるイベント検出、そして、対象物の画像とそれに対応するイベント検出結果の複数提示が可能となる。

以下、2章では本研究で使用した高解像度全方位画像センサについて述べる。3章では全方位画像センサで得られる全方位画像から注視すべき対象物を検出し、追跡する手法について述べ、4章では注視対象物の時系列における動きの解析と動き情報を基にしたイベント検出手法について述べる。5章では提案手法によるイベント検出の実験結果について述べる。最後に6章では本論文のまとめと今後の課題について述べる。

2. 高解像度全方位画像センサ

本章では、本研究で使用した高解像度全方位画像センサについて述べる。今回使用した画像センサは従来より我々の研究室で用いられている全方位画像センサ HyperOmni Vision [23] を高解像度化したものであり、以下の仕様を満たすものである。

- ・ 高解像度カメラとしてハイビジョンカメラを使用する
- ・ 画角 (視野) は水平より 30 度上まで確保確保する
- ・ ズーム機能のついたレンズを使用する
- ・ 屋内でも使用可能にする

これらの要求を基に設計された高解像度全方位画像センサ [24] について、2.1 節ではセンサの構成を示し、2.2 節で本センサの特徴について簡単に述べる。

2.1 高解像度全方位画像センサの構成

まず、高解像度全方位画像センサの外観と構成を図 2 に示し、設計に使用した機器と各種パラメータを表 1 に示す。

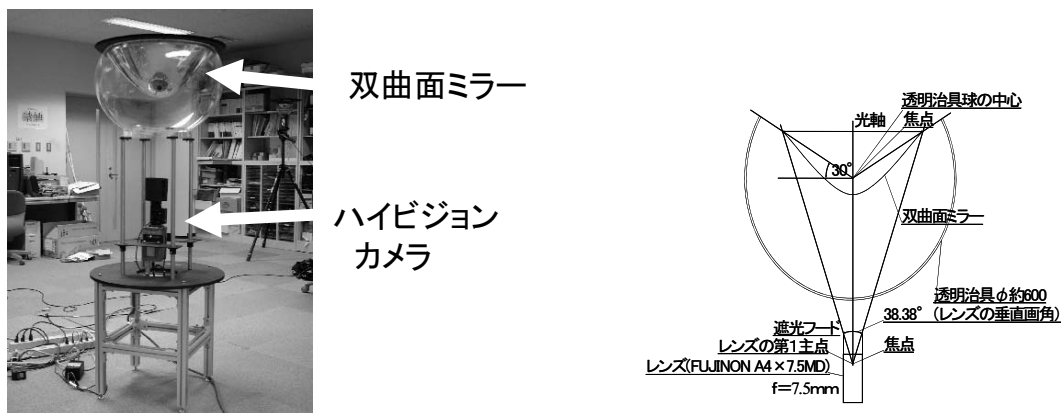
本センサは図 2(a) に示すように、鉛直下向きに取り付けられた双曲面ミラーに写り込む像をその下部に取り付けたハイビジョンカメラで撮像するものである。ミラー部の双曲面には回転 2 葉双曲面を用い、図 3 に示すように 2 焦点 O_M , O_C が存在し、ミラー焦点とカメラのレンズ中心がそれぞれ O_M , O_C に位置するように設置される。さらに、図 3 に示すような Z 軸を鉛直軸とする 3 次元ワールド座標系 (X, Y, Z) を考えると、双曲面ミラーの面、ミラー焦点 O_M 、及びカメラのレンズ中心 O_C は式 (1) で表せる (a, b, c の値は表 1 参照)。

$$\begin{aligned} \text{ミラー曲面} & : \frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \quad (Z > 0) \\ \text{ミラーの焦点 } O_M & : (0, 0, +c) \quad (c = \sqrt{a^2 + b^2}) \\ \text{カメラのレンズ中心 } O_C & : (0, 0, -c) \end{aligned} \tag{1}$$

(ただし a, b, c : 双曲面パラメータ)

表 1 高解像度全方位画像センサの仕様

ハイビジョンカメラ DXC-H10 (SONY)	
撮像素子	2/3in.3CCD
有効画素数	1920 × 1035pixel
撮像面	9.59 × 5.39mm
レンズ A4 × 7.5MD (FUJINON)	
焦点距離 (ワイド端)	7.5 mm
ズーム比	4 倍
最大口径比	F2.8
最至近撮影距離 (M.O.D.)	500mm
双曲面ミラー	
双曲面パラメータ a	163.7 mm
双曲面パラメータ b	236.8 mm
双曲面パラメータ c	287.9 mm
ミラーの直径	500 mm



(a) センサの外観

(b) センサの構成

図 2 高解像度全方位画像センサ

また、双曲面ミラーの焦点 O_M とカメラのレンズ中心 O_C との位置関係ならびに双曲面の特性を考慮すると、3次元環境中の点 $P(X, Y, Z)$ と全方位画像上の写像点 $p(x, y)$ の間に次の関係式 (2) が導かれる。この関係式により、全方位画像から平面透視投影画像作成のための対応点計算が可能となる。

$$\begin{aligned}
 x &= X \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}} \\
 y &= Y \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}}
 \end{aligned} \tag{2}$$

2.2 高解像度全方位画像センサの特徴

今回使用した高解像度全方位画像センサはカメラ部にハイビジョンカメラを用いることで、従来の全方位画像センサに比べて比較的解像度の高い全方位画像を得ることができる。そして、双曲面ミラーを用いた全方位画像センサの特徴として、得られる全方位画像は側方領域の情報を多く含むだけでなく、センサ自身の写り込みを除く下方領域の情報も含んでいることである。このことは人間の目線の

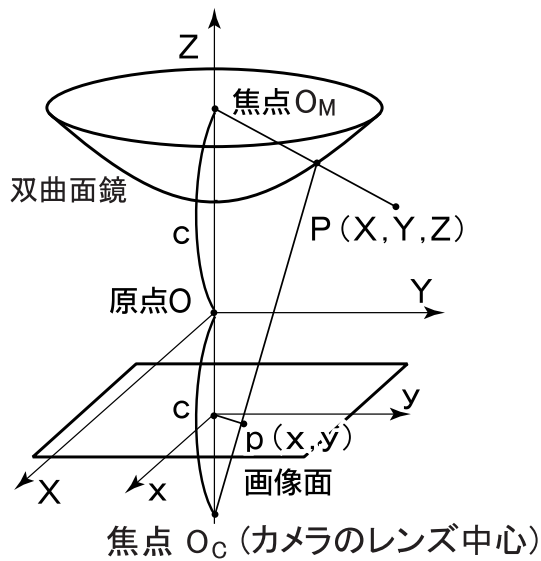
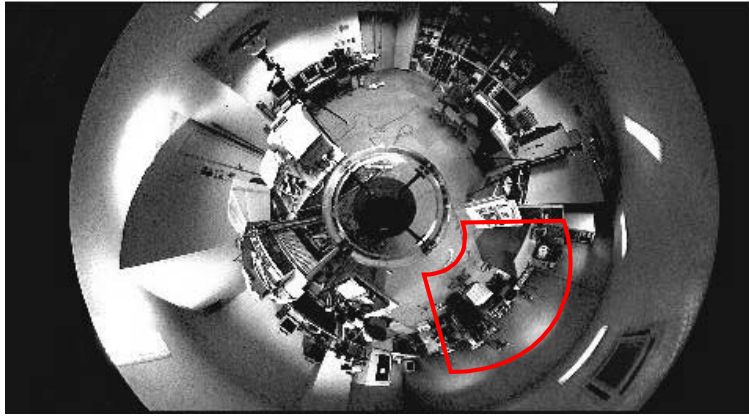


図 3 高解像度全方位画像センサの視覚系の構成

高さに設置されたセンサによって、監視環境に存在するセンサ側方の人物からセンサの下方に置かれた物まで、およそ監視に必要な領域を効率良く含んだ画像が取得できることを示している。また、双曲面の特性によりその画像は透視投影の性質をもち、図 4(a) に示すような全方位画像を図 4(b) に示すようなミラー焦点 O_M から見た通常の平面透視投影画像に変換することができる。本研究ではこの透視投影画像を監視者への提示画像として利用する。

さらに、解像度の違いを確認するために従来の全方位画像センサ HyperOmni Vision で撮像された全方位画像と透視投影画像を図 5 に示す。図 4 と図 5 の違いから高解像度全方位画像センサから得られる提示用画像は非常に有用なものであることがわかる。その上、ズームを変更することで図 6 に示すような撮像範囲の異なる画像を得ることができ、監視環境や目的に応じた撮像が可能であることがわかる。



(a) 全方位画像

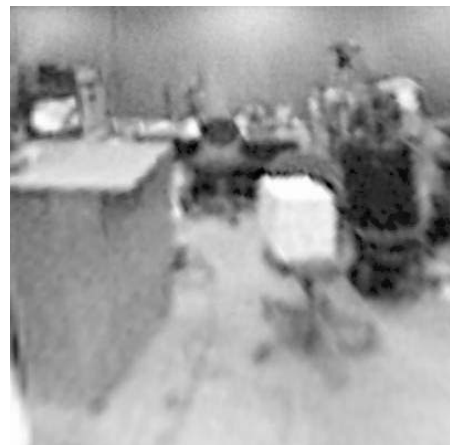


(b) 透視投影画像

図 4 高解像度全方位画像センサによる全方位画像と透視投影画像



(a) 全方位画像



(b) 透視投影画像

図 5 HyperOmni Vision による全方位画像と透視投影画像

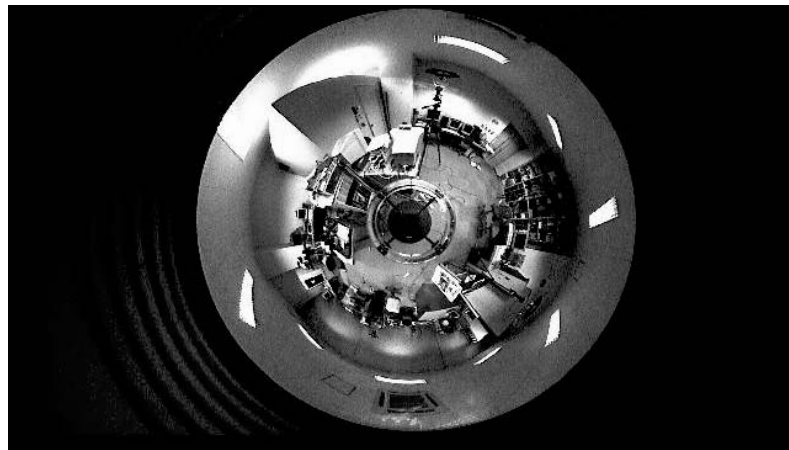
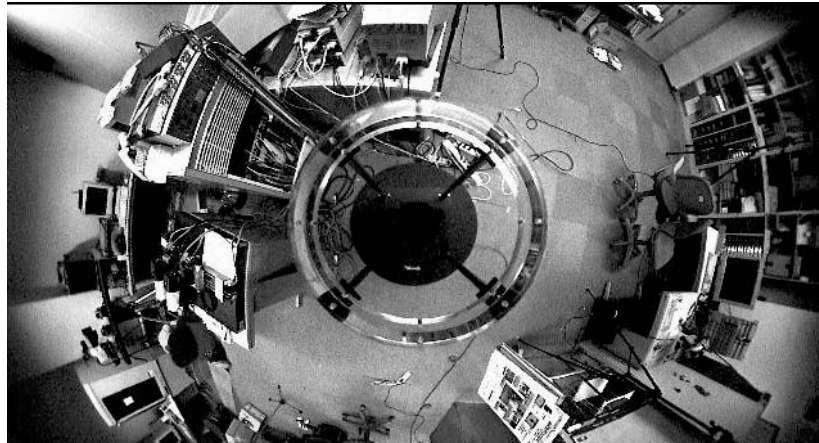


図 6 異なるズームにおける全方位画像の違い

3. 対象物の検出と追跡

遠隔監視システムにおいて、環境内に存在する注視すべき対象物や環境内で起こる異常事態を自動で検出し、監視者に知らせることは非常に重要である。また、注視対象物の動きや環境内の変化を自動で追跡し、画像提示することによって監視者への負担を大きく軽減することができる。さらに、追跡された対象物の動きを解析し、特定のイベントを検出することで監視者にとって必要な情報だけを抽出することができる。こうした重要な機能のうち、ここでは、画像センサで撮像された監視環境内から注視すべき対象物を自動検出し、その対象物を追跡する方法について述べる。

一般に、監視環境において監視者が注視すべき対象というのは、環境内を移動する物体や環境内で起こるその他の変化である。ここでは高解像度全方位画像センサにより得られる全方位画像から移動物体やその他の変化を検出し、追跡する手法について述べる。

以下、3.1節で注視対象物の検出手法について述べ、3.2節で注視対象物の追跡手法について述べる。

3.1 対象物の検出

画像内から対象物を検出・追跡するための方法はいくつか挙げられるが、代表的なものとして背景差分やオプティカルフローを用いた方法 [25, 26] などがある。背景差分を用いる方法の特徴は処理が簡単なため高速に行なえ、複数の対象物も同時に検出することができる。さらに、他の方法に比べて検出の信頼性が高いため監視システムにおいて多く利用されている [6, 7, 8, 9, 10]。しかし、複数の対象物が重なった場合には 1 つの領域として検出されるため、追跡する場合には何らかの処理によって個々の対象物を識別する必要がある。オプティカルフローによる方法では、移動ベクトルを計算するため、速度の異なる物体が重なった場合でも個別に検出することができ、その移動ベクトルによって対象物の追跡が容易になる。しかし、信頼性の点で問題があり、精度を向上させるには膨大な計算量が必要となる。

その他の方法として特定の物体を環境内から検出するテンプレートマッチングによる方法 [27, 28] も考えられる。この方法では、フレーム間での対象物の同定も同時に行なえるため対象物の追跡には有利であるが、本研究では入力が全方位画像という特殊な画像のためマッチングの処理が複雑であり、テンプレートに関しても対象物の位置や向きに影響を受けやすく信頼性と計算量の面で問題がある。

本研究においては、遠隔監視という目的上、処理の実時間性と信頼性が重要であると考え、背景差分による方法を用いる。背景差分の処理は入力画像の各フレーム毎に行ない、各フレームで得られる変化領域を対象物領域として実世界における対象物として抽出する。さらに、抽出された個々の対象物領域から、時系列での追跡や、次章で述べる動き解析のために必要な情報を取得する。まず、3.1.1節で背景差分を用いた対象物領域の抽出について述べ、3.1.2節で抽出された領域から対象物の追跡や行動解析に必要な情報の取得について述べる。

3.1.1 背景差分による対象物領域の抽出

まず、最初に環境内に対象物が存在しない時点での画像を背景画像として取得する。背景画像の一例を図 7(a) に示す。以降の入力画像と背景画像の間で差分を求め、得られる差分画像を固定閾値で 2 値化することによって 2 値画像を得る。入力画像の一例を図 7(b) に示し、差分画像と 2 値画像をそれぞれ図 7(c)、図 7(d) に示す。このとき画像内のカメラの写り込み部分は 2 値化処理時にカットする。こうして得られた 2 値画像は環境内で変化のあった変化領域と変化のない背景領域で構成されるが、このうちの変化領域を実環境内の監視対象物を表す対象物領域として検出する。それには、2 値画像上の変化領域をまとめた領域毎に分類し、それぞれの変化領域を実環境の個々の対象物に対応させる必要がある。また、図 7(d) に示すように実際の 2 値画像には実環境内の対象物を表す領域以外にもノイズによる領域が多く存在するため、これらの除去も行なわなければならない。

そこで、これらの問題を解決するためにラベリングによる領域の分類と面積による領域の切捨てを行なう。注目画素の近傍 8 画素の連結性を調べ、連結する画素には同一のラベルを貼ることで変化領域の分類を行ない、同一ラベルが付けられた領域を 1 つの対象物領域として検出する。そのとき、分類された同一ラベル領

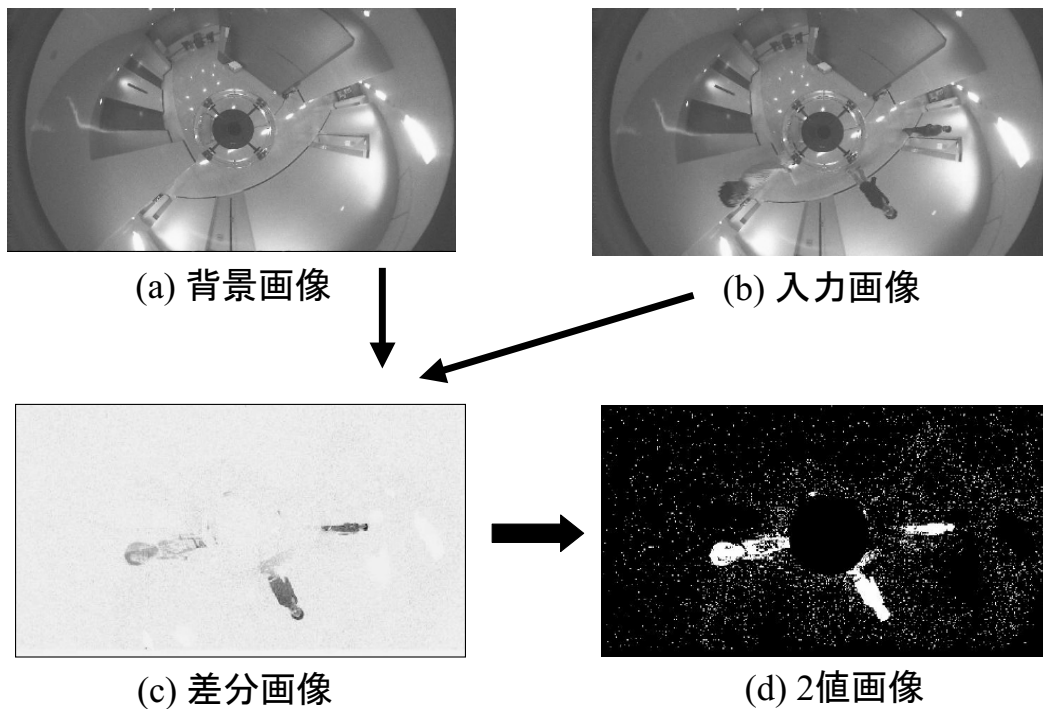


図 7 背景差分による領域抽出

域に含まれる画素数を調べ、その値が一定数に満たない小さな領域はノイズ領域として切り捨てる。実際にラベリング処理と面積による切捨て処理を行ない、対象物領域を検出した結果の一例を図 8 に示す。

ここまでの背景差分、ラベリング、面積による切捨て処理によって、入力の全方位画像から注視対象物となる領域を抽出することができる。しかし、背景差分とラベリングの処理には、入力画像全体の走査が必要であり、本研究のように解像度の高い画像センサを用いた場合、その処理時間が問題となる。そこで、提案手法では、入力画像に対してこれらの処理を行なう際、縦横ともに画素を間引いて一定画素おきに行なう。注視対象物となる領域はある一定値以上の大きさを持っており、背景差分においてもまとまった領域として抽出される。さらに、ここで欲しい情報は対象物の細かな形状や色ではなく、存在する領域のみなので画素を間引いた場合でも不都合は生じない。しかも、走査対象にならないような微小領域もあ

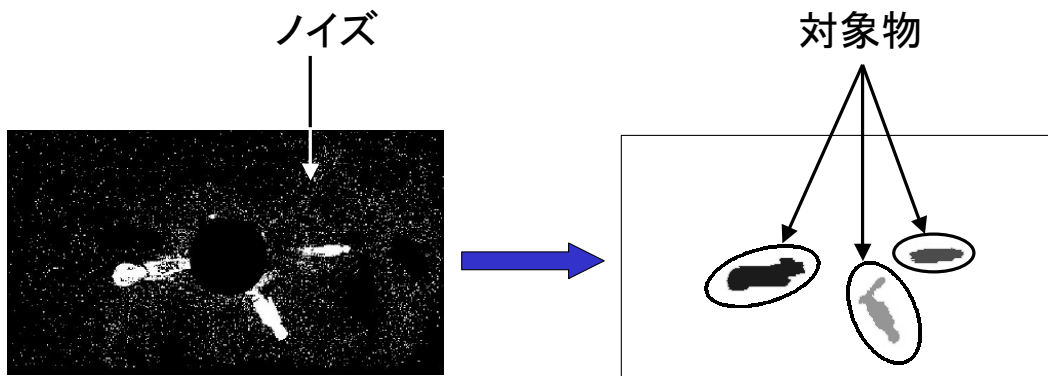


図 8 ラベリング処理と面積による切捨て処理

らかじめ削除することができ、全体の処理速度を大幅に向上できる。しかし、対象物と背景が同系色である場合、差分が少なく 2 値化によって同一の対象物を表す変化領域が分割され、複数の対象物領域として検出されることがある (図 9 参照)。このとき、分割された領域は、一旦別々の領域として検出するが対象物の追跡時に結合を行なう。領域の結合については 3.2.3 節で詳しく述べる。

3.1.2 領域情報の取得

検出された個々の対象物領域について、対象物の追跡や動き解析のために表 2 に示すような領域情報を取得する。このうち、領域番号、面積、重心座標は対象物検出のラベリング処理時に取得し、対象物の検出・追跡を行なうための各処理に利用する。種類情報については、3.2 節の対象物領域の対応付け、及び 4 章で述べる対象物の動き解析によって取得し、動き解析やイベント検出に利用する。

3.2 対象物の追跡

注視すべき対象物を追跡するためには、連続する時系列フレーム間で同一の対象物を表す対象物領域を対応付ける必要がある。連続するフレーム間で対象物領域の対応付けを行なうために、まず連続する前後 2 フレーム間で対象物領域を対応付ける。前後 2 フレーム間での対応付けは現在の入力画像と直前の入力画像の

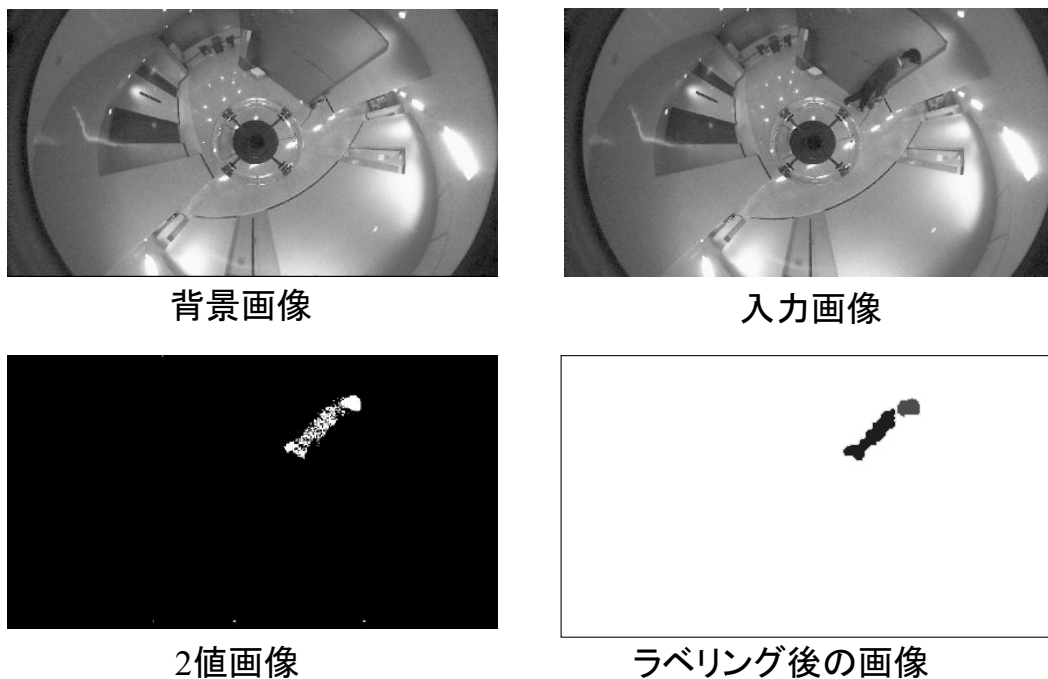


図 9 同系色による対象物の分割

2 フレームを用いて行ない、これを時系列で順次繰り返すことで連続フレーム間での対応付けとする。そのとき、監視環境内に存在する複数の対象物が重なることで、検出される対象物領域の数が減少したり、3.1.1節で述べたように本来単一の対象物領域の分割によって、フレーム間で対象物領域の数が変化する場合があります。こうした状況を考慮して、2 フレーム間での対応付けを前後 2 フレーム間で双方向に行ない、複数の対象物領域どうしの対応付けを実現する。フレーム間での対象物領域の対応付けの基準となる尺度には対象物領域の重心間距離、後ほど述べる領域間距離、カラーヒストグラムを利用する。まず、対象物領域の重心間距離を用いて前後 2 フレーム間における対象物領域の結びつきを調べる。その結びつきが単一対象物と複数対象物の場合には領域間距離を用いて対象物領域の分離や結合を行なう。また、カラーヒストグラムは、一度、他の対象物領域と重なって集合領域となった対象物領域が再び分離したときに、集合前のどの対象物領域と対応するのかを調べるために使われる。

表 2 検出された対象物の領域情報

情報	意味		用途
領域番号	対象物領域が固有に持つ識別番号		対象物の同定
面積	対象物領域に含まれる画素数		領域の結合
種類	対象物領域の状態を次の 5 つに分けて保持する		動き解析 イベント検出
	新規領域	新しく出現した対象物領域	
	静止物体	自ら移動しない物体	
	動物体	自ら移動する物体	
	集合物体	対象物の重なりによる見かけ上 1 つの領域	
	分離領域	1 つの対象物から分離してできた対象物領域	
重心座標	対象物領域に含まれる x, y それぞれの平均座標		時系列での対応付け

これらの対応付けの結果に基づき、連続するフレーム間で各対象物領域の同定を行ない、対象物を追跡する。追跡の結果は提示画像として監視者に示される。

以下、3.2.1節では対象物領域の重心間距離を用いた対応付けについて述べ、3.2.2節では対象物領域を同定する時に用いるカラーヒストグラムを用いた対応付けについて述べる。そして、3.2.3節で連続フレーム間における対象物領域の同定について述べ、3.2.4節では、追跡結果を監視者へ提示するための透視投影画像生成について述べる。

3.2.1 重心間距離による対応付け

まず、対応付けを行なう前後 2 フレームをそれぞれ、前フレーム、現フレームとし、各フレームに存在する対象物領域をそれぞれ、 $P = \{p_i | i = 1, \dots, n\}$ 、 $N = \{n_j | j = 1, \dots, m\}$ とする。また、 p_i と n_j の重心間距離を $D(p_i, n_j)$ として以下の手順で対応付けを行なう。

1. N 中の各 $n_j (j = 1, \dots, m)$ に対して、

- (a) P 中の $p_i (i = 1, \dots, n)$ と重心間距離を用いて順に比較を行ない, $D(p_k, n_j)$ が最小となる $p_k (1 \leq k \leq n)$ を見つける
 - (b) $D(p_k, n_j)$ が閾値内であれば, p_k と n_j を対応付ける
2. P 中の各 $p_i (i = 1, \dots, n)$ に対して,
- (a) N 中の $n_j (j = 1, \dots, m)$ と重心距離を用いて順に比較を行ない, $D(p_i, n_l)$ が最小となる $n_l (l \leq l \leq m)$ を見つける
 - (b) $D(p_i, n_l)$ が閾値内のとき,
 - i. p_i と n_l が対応済みの場合はそのまま
 - ii. n_l が他の $p_k (1 \leq k \leq n)$ とすでに対応付けされている場合, 前フレームにおける複数の対象物領域と現フレームにおける単一対象物領域の対応付けとして p_i と n_l を対応付ける
 - iii. p_i が他の $n_k (1 \leq k \leq m)$ とすでに対応付けされている場合, 前フレームにおける単一対象物領域と現フレームにおける複数の対象物領域の対応付けとして p_i と n_l を対応付ける

3.2.2 カラーヒストグラムを用いた対応付け

カラーヒストグラムは画像特徴量の 1 つであり [29], 類似画像の検索などに利用されている [30, 31]. 本研究では, RGB 空間を用いたカラーヒストグラムの生成をおこない, その類似度評価を対象物領域の同定のために利用する. 対象物領域 I と対象物領域 J を比較する場合を例に以下にその手順を示す.

1. RGB 空間を各軸それぞれ 8 等分し, 合計 512 のセルに分割する (色空間 $C(c_1, c_2, \dots, c_{512})$ とする)
2. 対象物領域 I に含まれる各画素を調べ, 色空間 C 中の $c_i (i = 1, 2, \dots, 512)$ に落ちる画素数を I_i とする
3. 対象物領域 J に含まれる各画素を調べ, 色空間 C 中の $c_i (i = 1, 2, \dots, 512)$ に落ちる画素数を J_i とする

4. 生成されたカラーヒストグラムの比較のために、以下の式を用いて評価を行なう

$$\text{類似度評価関数 } S = \sum_{k=1}^{512} \left| \frac{I_k}{S_I} - \frac{J_k}{S_J} \right|$$

(ただし、 S_I, S_J はそれぞれ対象物領域 I, J の面積である)

3.2.3 時系列間での対象物領域の同定

3.1節で述べたように、検出される単一の対象物が必ずしも実世界での単一の対象物と結び付くとは限らない。そのため、対応付けされた対象物領域について、同一フレーム内の隣合う対象物領域の領域間距離を用いて対象物領域の分離や統合を行ない、時系列間において各対象物領域を同定する。ここで、領域間距離とは隣接する対象物領域に含まれる画素のうち八近傍距離が最も小さくなる画素間の距離と定義する。さらに、前節で対応付けされた2フレーム間での対象物領域の関係を以下の5つの場合に分けて説明する。

(1) なし 単一対象物領域

図10のように、現フレームの対象物領域 n_i に対して前フレームで対応する対象物領域が存在しない場合、 n_i は現フレームで新たに出現したのものとしてその対象物領域の領域情報の種類欄に新規領域のラベルを貼る(表2参照)。また、現フレームからの追跡を開始し、イベント検出のための動きを解析する。なお、動き解析については、後の4章で詳しく述べる。

(2) 単一対象物領域 単一対象物領域

図11のように、現フレームにおける対象物領域 n_i に対応する前フレームでの対象物領域がただ1つ p_k しか存在しない場合、現フレームの n_i を前フレームでの対象物領域 p_k と同定し、 p_k の面積、重心座標以外の領域情報を継承する。

(3) 単一対象物領域 なし

図12のように、前フレームの対象物領域 p_i に対応する現フレームでの対象物領域が存在しない場合、対象物領域 p_i は現フレームにおいて対象物領域

が消失したとして、対象物領域 p_i の領域情報の継承を行わず、対象物領域 p_i の追跡を終了する。

(4) 複数の対象物領域 単一对象物領域

図 13のように、直前のフレームに存在する複数の対象物領域 $p_i, p_j (i < j)$ が現フレームにおいて、1つの対象物領域 n_k に対応付けされる場合は、現フレームで対象物領域 p_i, p_j が重なり、その結果、単一对象物領域として検出されたと考える。そこで、 n_k の領域情報の種類欄に集合領域のラベルを貼り（表 2参照）、 p_i の領域情報を継承する。さらに、集合される対象物領域 p_j の領域情報とカラーヒストグラムを集合情報として保持する。

(5) 単一对象物領域 複数の対象物領域

図 14のように、前フレーム中の1つの対象物領域 p_i が現フレーム中の複数の対象物領域 n_k, n_l に対応付けされる場合、対象物領域 p_i が現フレームにおいて分離して新たに対象物領域が出現したか、もしくは、背景差分時のノイズによって本来1つの対象物領域が分割されて複数の対象物領域として検出されたかのいずれかであると考えられる。対象物領域が分離した場合は分離後の各対象物領域の距離が一定以上離れるが、ノイズによる対象物領域の分割ではその距離が少なくなると考えられる。そこで、これらを区別するために、図 15に示すように隣接する対象物領域領域の領域間距離を求め、その距離が固定閾値内であれば対象物領域の分割と判断しそれらの領域を結合する、それ以外を対象物領域の分離とする。

対象物領域を結合する場合 領域間の八近傍距離が固定閾値内であるとき、現フレーム内の対象物領域 n_k と n_l を同一の対象物領域として結合する。このとき、実際に2つの領域間のすき間の画素を埋めることはせず、2つの領域を1つの領域として対象物領域 p_i の領域情報を継承する。このとき新たな面積として各領域の合計を用い、重心座標に関しては、新しい重心座標 $G(g_x, g_y)$ を対象物領域 n_k, n_l の重心座標 $G_k(x_k, y_k)$, $G_l(x_l, y_l)q$ とそれぞれの面積 S_k, S_l より以下の式によって算出する。

$$g_x = \frac{S_k x_k + S_l x_l}{S_k + S_l}, \quad g_y = \frac{S_k y_k + S_l y_l}{S_k + S_l}$$

対象物領域を分離する場合 領域の分離と判定された場合, 対象物領域 p_i が集合領域であれば, p_i が保持する集合情報中のカラーヒストグラムを n_k, n_l のカラーヒストグラムと順に比較し, 対応付けられるものがあれば対応する領域の領域情報を継承する. 対応の取れなかったものは p_i の領域情報を継承する.

対象物領域 p_i が集合領域でない場合, 新規の分離領域として n_k, n_l ともに p_i の領域情報を継承し, 領域情報中の種類欄には分離領域のラベルを貼る (表 2 参照).

3.2.4 監視者への提示

前節までの処理により監視環境内に存在する注視すべき対象物を対象物領域として検出し, 追跡することが可能となった. ここでは, 対象物の行動を監視するために監視者へ画像提示を行なう. 提示用の画像は, 対象物領域の重心座標を中心とする一定の領域を入力的全方位画像から切り出し, 2.1節の式 (2) の変換を用いて生成される透視投影画像である. 透視投影画像は環境内の注視対象の数に応じて生成され, 領域番号 (表 2 参照) が同一の対象物領域が存在する限り, 一つの提示画像が出力される.

透視投影画像生成に関して, 切り出される領域内の全点に関して対応点の計算を行なうとその処理時間が問題となり, リアルタイムでの処理は困難となる. そこで, 全方位画像からの透視投影画像生成には尾上らの方法 [32] を利用する. 具体的には, 図 16 に示すように生成する透視投影画像を格子に分割し, 各格子点に対応する全方位画像上の点はソフトウェアによる正確な計算, 格子内部の各点はハードウェアのテクスチャマッピング機能による変換を行なう. これにより高速に透視投影画像を生成することが可能となる. 提示画像の例を図 17 に示す.

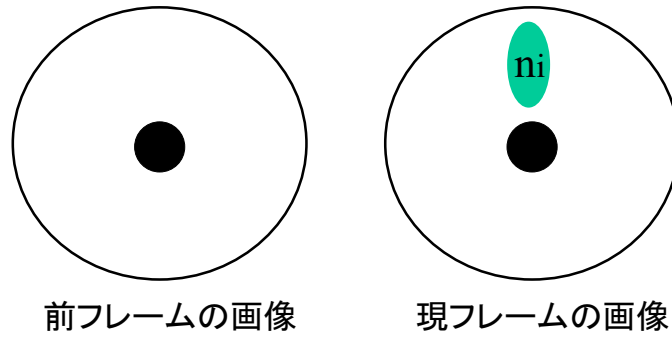


図 10 フレーム間での対象物の同定 (なし 単一対象物領域)

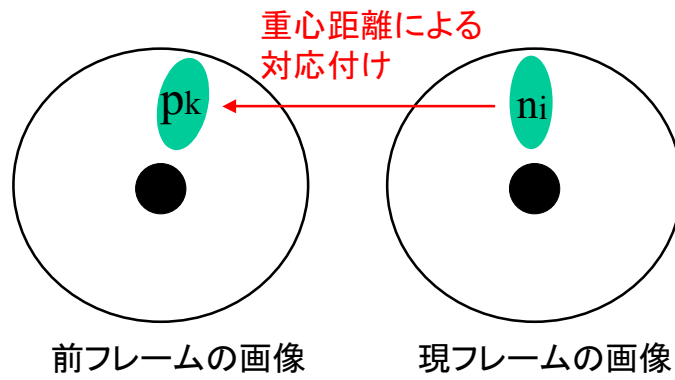


図 11 フレーム間での対象物領域の同定 (単一対象物領域 単一対象物領域)

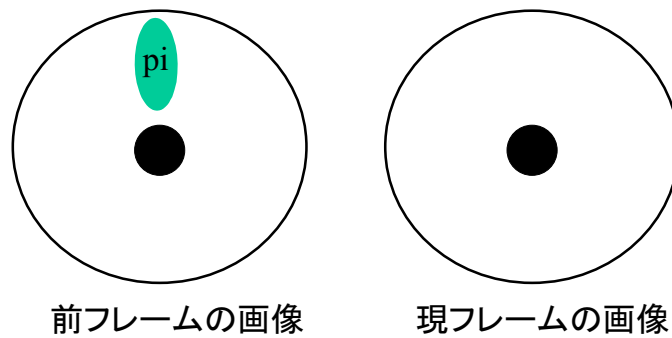


図 12 フレーム間での対象物領域の同定 (単一対象物領域 なし)

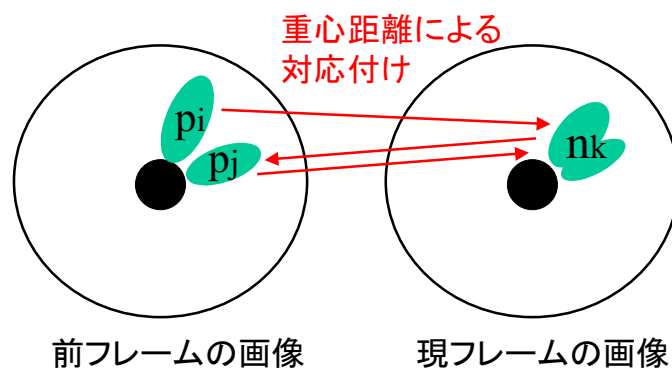


図 13 フレーム間での対象物領域の同定 (複数の対象物領域 単一対象物領域)

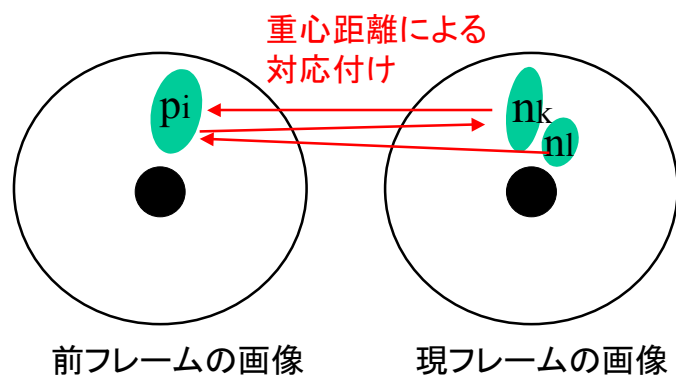


図 14 フレーム間での対象物領域の同定 (単一対象物領域 複数対象物領域)

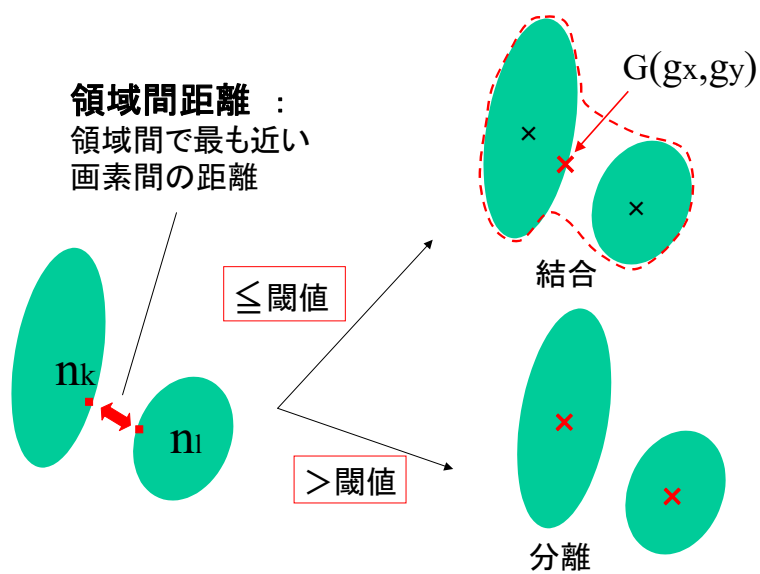


図 15 領域間の八近傍距離による分離・分割の判定

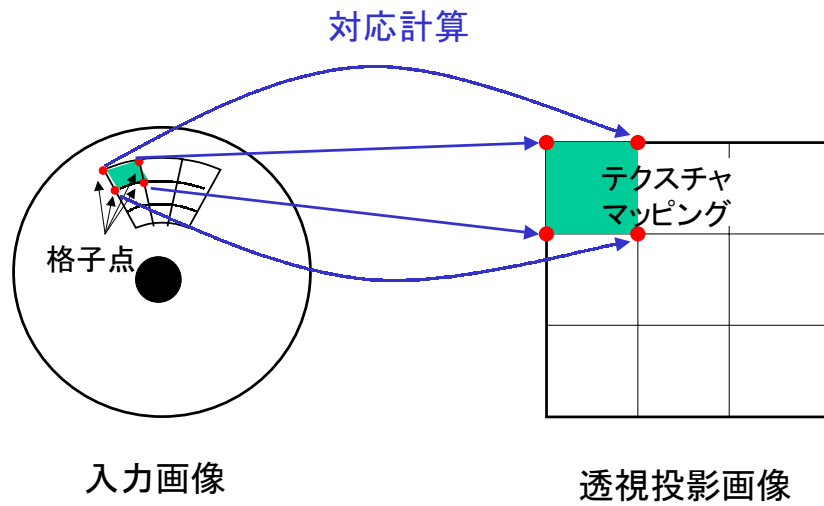


図 16 全方位画像と透視投影画像の対応点計算



図 17 監視者への提示画像

4. 対象物の動き解析とイベント検出

監視環境内から注視すべき対象物を自動で検出し、追跡・提示を行なうことで監視者へ注意を促すことができる。さらに、その対象物の動きを解析し、環境内で起こる変化をイベントとして検出することにより、目的に応じた必要情報だけを効率良く得ることが可能となり、監視者の負担を大きく軽減することができる。また、こうしたイベント情報に基づいて記録されたストリームの要約を作成することで、膨大なストリームデータの管理が容易になる。

本章では、まず、4.1節において検出された対象物の動きを解析する方法について述べ、4.2節で、検出するイベントの設定とその検出手法について述べる。

4.1 対象物の動き解析

検出された対象物の動きを解析するために、時系列における対象物領域の状態変化を調べる。そのため、時系列差分を取ることで対象物領域の変化を調べるが、ここでは対象物の検出とは異なり解像度の高い画像センサの利点を活かすため、対象物領域内の全画素について時系列差分を取る。

提案手法では、本節で述べる動き解析の結果を用いて次節で述べるイベント検出を行なう。対象物の動き解析には、現在の入力フレームと直前のフレームを用い、直前のフレーム（以後、処理フレームとする）から現在の入力フレーム（同様に、次フレームとする）を見た時に、フレーム内の対象物領域がどのように変化しているかを調べる。処理フレームに存在する対象物領域に対して、次フレームと時系列差分を取り、その領域内の画素変化が一定割合を越えるかどうかを調べることで対象物の動きの有無を判断する。また、処理フレーム内の各対象物領域に対して3.2.3節で述べた対応付けの結果により、次フレームにおいて対応する領域の存在を調べる。さらに、対応する領域が存在する場合それらの重心間距離によって対象物の位置変化を調べる。これら前後2フレーム間での動き解析を基本として、実際の処理では一定フレームの間連続して解析を行ない得られた結果を動き解析の結果とする。数フレームの間連続して動きを解析するのは、影や光の反射といった一時的に出現するノイズ領域を除去するためである。また、時系列差分

と同時に対応する領域の有無を調べるのは、時系列差分で現れる大きな変化が対象物の動きだけによるものではなく、対象物領域の消失によって生じた可能性も考えられるからである。

以上の処理の結果、設定するフレーム数の間連続して領域の対応が取れ、かつ時系列の変化が大きい対象物は「動きあり」、領域の対応が取れ、時系列の変化が小さい対象物は「動きなし」、領域の対応が取れない対象物は「対応なし」と分類される。さらに、「動きあり」と分類される対象物については対応する領域間の重心距離による対象物の位置変化を調べておく。これにより、対象物が一定の位置で動きを持つのか、対象物の移動によって動きが生じるのかが判断でき、「動きあり」を「移動あり・動きあり」と「移動なし・動きあり」に再分類して判定する。図 18に動き解析処理のフローチャートを示す。

4.2 イベント検出

ここでは、提案手法において検出の対象となるイベントの設定方針とその検出手法について述べる。本研究におけるイベント検出では監視ビデオの要約作成への応用を考慮している。そのため、検出すべきイベントとしては、ある特定の対象物が起こす特殊な行動ではなく、監視環境下で一般的に行なわれると予測される各種行動が検出の対象となる。また、特定の環境や状況に依存したイベント検出ではなく、汎用性の高いシステムにするため、環境内のマップ情報や環境情報を用いない手法を提案する。

そこで、提案手法では環境内に存在する対象物の時系列における動き解析の結果のみを用いたイベント検出を行なう。以下、表 3に提案手法で設定する各種イベントを示した後、各イベントの検出方法について述べる。

出現イベント 処理フレーム内に存在する新規領域(表 2参照)について、前節 4.1で述べた動き解析を行なう。その結果、「対応なし」以外に判定されればこの新規領域の種類を動物体書き換え、環境内に新たに対象物が現れたものとして出現イベントとする。それと同時に提示画面への出力を行ない監視者に通知する。

表 3 イベントの設定

イベント名	イベントの説明
出現 (APPEAR)	環境内に新しく対象物が侵入したとき
退出 (EXIT)	対象物が環境内から退出したとき
静止 (STOP)	対象物の位置変化がなく、動作もないとき
作業 (WORK)	対象物が一定位置で動作を行なっているとき
オブジェクト (OBJECT)	対象物が環境内に静止物体を置いたとき 対象物が環境内の静止物体を取ったとき

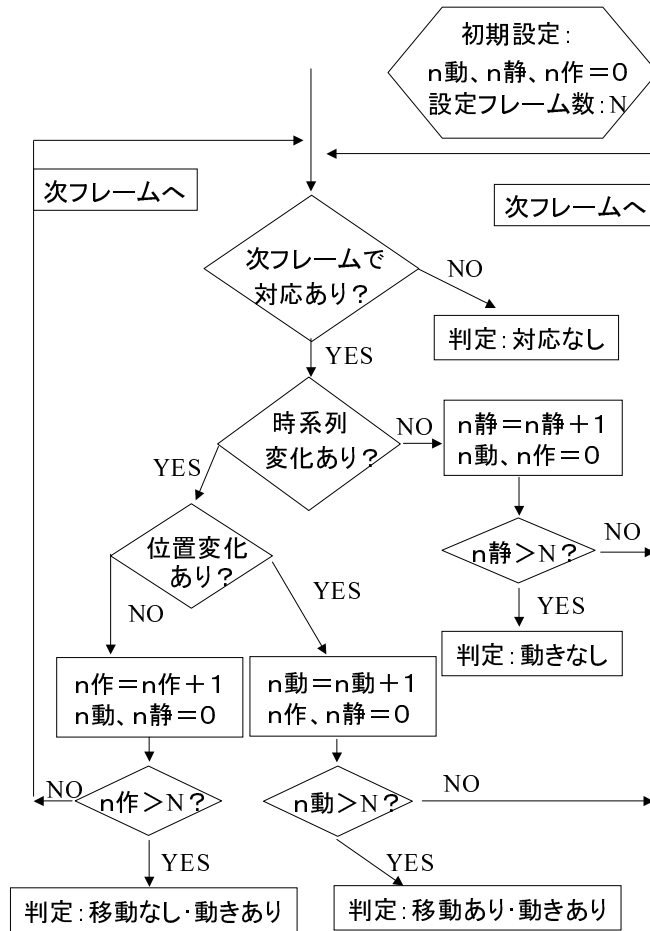


図 18 動き解析の処理の流れ

退出イベント 処理フレームに存在する動物体もしくは集合領域(表2参照)について、動き解析の結果、「対応なし」と判定された場合、動物体が環境内から退出したとして退出イベントとする。

静止イベント 処理フレーム内に存在する動物体について動き解析を行なう。その結果、「動きなし」と判定された場合、その動物体は定位置で動きがないものと判断し、静止イベントとする。同時に提示画面への出力を行ない監視者に通知する。

作業イベント 処理フレーム内に存在する動物体について動き解析を行なう。その結果、「移動なし・動きあり」と判定された場合、その動物体は一定の場所で何らかの動作を行なっていると判断し、作業イベントとする。また同時に提示画面への出力を行ない監視者に通知する。

オブジェクトイベント 処理フレームで分離領域として出現した領域のうち、動き解析の結果、「動きなし」と判定された場合、その領域は動物体が環境へ作用することで出現した静止物体と判断する。このとき環境への作用として考えられるのは、動物体が環境内に新しく静止物体を追加したか、環境内の静止物体を取り除いたかなので、これをオブジェクトイベントとする。また別の場合として、追加された物体を取り除いたときや取り除いた物体を元に戻したときが考えられる。これらの場合は、一度オブジェクトイベントが発生して静止物体が出現した後、その領域が背景差分で検出されなくなったときなので、静止物体の出現時からその領域周辺を調べ続け、領域がなくなった時点でイベント発生とする。それと同時に、提示画面への出力による監視者への通知を行なう。

5. イベント検出実験

本章では、提案手法によるイベント検出が実際の監視システムにおいて有効であるかを検証するために行なった実験について述べる。実験では、2章で述べた高解像度全方位画像センサを用いて環境を監視し、3, 4章で述べたアルゴリズムによって注視すべき対象物の検出・追跡と対象物の動き解析によるイベント検出を行なうシステムを構築した。

以下、5.1節では使用機器や撮影場所などの実験環境について述べる。そして、5.2節で実験結果について述べた後、5.3節において実験結果を考察する。

5.1 実験設定

まず、構築したシステムの構成を図 19 に示す。使用した画像センサは 2 章で述べた高解像度全方位画像センサであり、使用した計算機は SGI 社の Onyx2 IR2 (MIPS R10000 195MHz, 16CPU) である。また、高解像度全方位画像センサの各種パラメータは表 1 で示した通りである。ただし、ズームの設定は撮影環境や目的に合わせて、その都度調整を行なうものとする。また本研究では、監視を行なう環境を屋内環境に限定し、環境内を移動する物体の多くは人物であるという仮定に基づいて、高解像度全方位画像センサを人間の目線の高さに設定し、環境全体を効率良く撮像するために監視環境の中央にセンサを配置するものとする。図 19 に示すように、システムの入力は高解像度全方位画像センサで得られる全方位画像であり、その大きさは 1920×1035 画素である。出力は監視者へ提示される 2 つの透視投影画像であり、その大きさは 300×600 画素である。

構築したシステムでは、まず、監視環境内に設置された高解像度全方位画像センサによってその情景を全方位画像として取得する。得られた全方位画像データを逐次計算機に送り、提案手法で述べた処理を行なう。このとき、対象物検出のための背景差分、ラベリング処理は 3 画素おきに行ない、検出された対象物の動き解析のための時系列差分処理は対象物領域内の全画素を用いて行なう。また、各種閾値はヒューリスティックに設定している。背景差分、ラベリング処理の結果、監視者への提示用画像データを逐次生成し、監視者のいるモニターに表示する。提

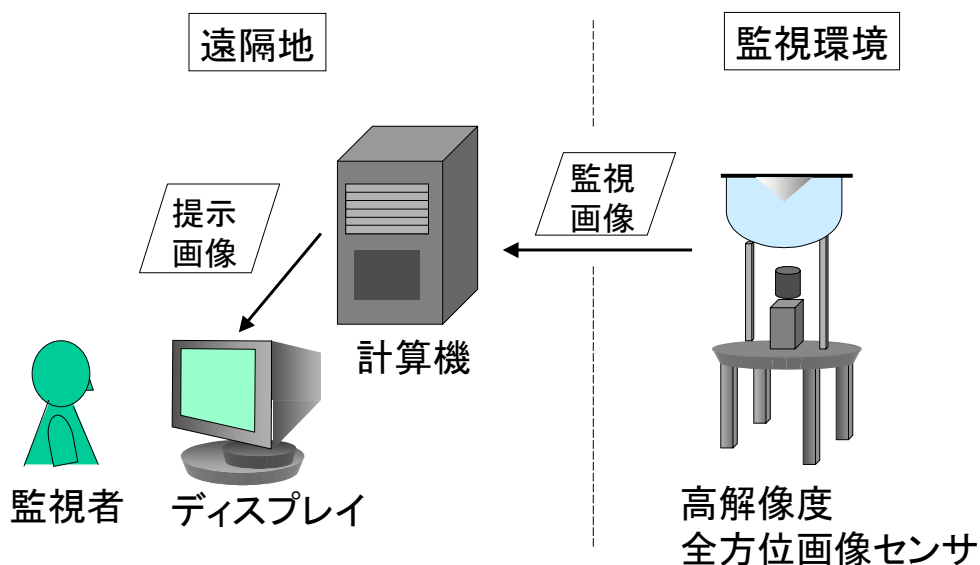


図 19 システム構成

示用の画像データとしては、環境内で検出された対象物を追跡した透視投影画像であり、追跡される対象物が起こすイベントの検出結果も同じ透視投影画像内に出力される。イベントの検出に関しては環境内に存在する全ての移動物体が対象物となるが、処理時間の問題から今回の実験では追跡により画像提示されるのは、同時に 2 人までとした。

次に、実験の設定について述べる。今回行なった実験は、屋内環境に高解像度全方位画像センサを設置し、環境内を出入りする人物の監視を行ない、発生するイベントを検出する。監視環境は図 20 に示すようなエレベーターホールを含む屋内の廊下であり、この中で複数の人物に設定したイベントに沿って行動してもらいその検出率を評価する。撮影は日照条件の変化による背景差分への影響を考慮して日没後に 20 分程度行なった。



図 20 監視環境

5.2 実験結果

前節で述べた実験設定に基づいて行なったイベント検出実験の結果を以下に示す。まず、図 21 に示すように対象物の存在しない実験環境を高解像度全方位画像センサで撮像し、背景画像とした。この環境内に対象物が侵入すると図 22(a) に示すような入力画像を得、3章で述べた対象物の検出、追跡処理によって図 22(b) のような提示画像を得た。

次に、設定した各イベントの検出の様子を図 24 ~ 図 5.2 に示す。

最後に、20 分間のイベント検出実験において、構築したシステムが検するイベントの数と、人が目視によって数えた実際のイベント数を表 4 において比較する。表 4 において誤検出数とはシステムが過剰に検出した数である。また、検出率は実際のイベント数に対するシステムの検出割合であり、正答率とはシステムが検出したイベント数に対する正しく検出されたイベント数の割合である。



図 21 監視環境の全方位画像

5.3 考察

実検の結果, 提案手法では $1/3s$ の更新間隔で監視環境内に存在する複数の対象物の検出・追跡を行なうことができ, 動き解析によって多地点で起こる複数のイベントを検出することができた. また, 対象物の追跡と検出したイベントの情報を設定した 2 つのウィンドウに表示することで監視者への提示も可能であった. イベントの検出に関しては, 「出現イベント」, 「退出イベント」については検出率や正答率からほぼ問題なく検出できていることがわかる. 「出現イベント」における誤検出の主な原因は, 背景差分によって対象物が分離してしまい, 新たに領域(偽領域)が出現してしまうことが挙げられる. そして, こうした偽領域が消失することで「退出イベント」の誤検出が生じることがわかった. また, 「静止イベント」, 「作業イベント」では未検出が多く, これらは実際に実際に静止や作業を行なっても動き解析で設定されたフレーム数に満たなかったためや, 時系列差分での動き判定の閾値によって生じることがわかった. 誤検出については, センサの放射方向に一直線に移動した場合に, 時系列での変化や重心座標の変化が少ないために起こることがわかった. オブジェクトイベントについては正答率からも, 誤って発生することのないことがわかった. 未検出の原因は主に, 処理速



(a) 入力画像



(b) 提示画像

図 22 対象物が侵入したときの入力画像と提示画像の一例

度が問題で時系列でのフレーム間の移動量が大きく、分離として生じるはずの領域が新規領域として出現することで起こることがわかった。その他の問題点としては、影による偽領域の出現や高速移動による追跡の失敗などが挙げられる。これらの問題点は今後の課題として残される。

提案手法によるイベント検出では、環境内で起こる複数のイベントを同時に検出することができ、使用する環境に合わせて提示画像を増やすことで監視者は必要な情報のみを得ることができる。また、画像提示されていない対象物についてもその動き解析とイベント検出は行なわれているため、監視ビデオの要約作成にもインデキシング情報として利用できる。

以上のように、広範囲な監視環境内で起こる複数のイベントを検出することで監視に必要な情報のみを得ることができ、監視ビデオの要約への応用も可能な本手法は、遠隔監視システムにおいて有効なイベント検出であると考えられる。



図 23 イベントの検出例 (出現イベント)



図 24 イベントの検出例 (退出イベント)



図 25 イベントの検出例 (静止イベント)



図 26 イベントの検出例 (作業イベント)



図 27 イベントの検出例 (オブジェクトイベント)

表 4 検出したイベント数の比較

イベント名	実際に 起こった数	システムが 検出した数	誤検出数	正答率 (%)	検出率 (%)
出現	34	39	5	87	100
退出	34	41	8	80	97
静止	15	23	15	34	53
作業	10	12	5	58	70
オブジェクト	20	13	0	100	65

6. おわりに

本論文では高解像度で撮像可能な全方位画像センサを用いて遠隔地の環境を広範囲に監視し、環境内の多地点で起こる複数のイベントを同時に検出する手法を提案した。提案手法では注視すべき対象物を背景差分により対象物領域として検出し、時系列差分によって各対象物の動きを解析し、その結果に基づいてイベント検出を行なう。そのとき、対象物領域の検出は画素を間引いて解像度の低い画像での処理、対象物の動き解析は全画素を用いた解像度の高い画像での処理によりセンサの特徴を損なわずに処理を高速に行なう。

また、実際にイベント検出システムを構築し実験を行なった。実験では環境内に存在する2人の注視物体の追跡・画像提示を1/3sの更新間隔で行なうことができ、それぞれの人物の動きを解析し、発生するイベントの検出、提示ができることを確認した。これにより、本研究で用いた手法は、(1)常時、広範囲の監視(2)対象物の自動検出といった監視システムにおける基本的な機能を満足するものであり、さらに、注視対象物の行動分析、イベント検出が可能なことで監視者の負担軽減を目的とした監視ビデオの要約作成に利用できる有用なものであると考えられる。

最後に今後の課題として、以下の点が挙げられる。

1. 複数台の全方位画像センサの使用

本研究では1台の画像センサにより監視を行なっているが、画像センサの台数を増やすことにより注視対象物の時系列間における対応付けの精度が向上すると考えられる。また、1台の画像センサでは他の領域と結合して検出されない新規の対象物が検出でき、誤対応を減らすことができる。さらに、異なる視点からの三角測量によって対象物の位置を推定することができるため、新たに位置情報を考慮したイベントの設定や対象物の移動経路の記録が可能となり、監視システムとしてより有用なものになると考えられる。

2. 時系列における領域の対応付け精度の向上

時系列間で対象物領域を対応付けするために、領域の重心間距離のみを用いて行なったが、重心間距離は領域の形状や大きさの影響を受けやすいため、

誤対応が発生した。そこで、対応付けの精度を向上させるために、重心間距離位以外に領域の面積や速度情報などを利用することが考えられる。

3. 実行速度の向上

処理全体をより高速化することで、時系列における対象物の誤対応を減らすことができる。また、対象物の高速移動による追跡の失敗も減らすことができる上、より細かな対象物の変化も捉えることができると考えられる。

4. 背景画像の更新

本研究では、背景画像は最初に入力された画像を使い続けているため僅かな照明条件の変化に対しても背景差分に影響が出てしまい、領域の誤検出となってしまう。そこで、背景画像を更新することで照明条件の変化に対してロバストな対象物の検出が行なえると考える。

5. 監視ビデオの要約作成

今回の実験では環境内の監視とイベント検出についてのみ行なったが、実際に長時間の監視ビデオのストリームに対してイベント情報を行ない、その情報を基にした要約ビデオの作成を行なう要求もある。その際は、要約されたビデオと元のビデオを比較し、ストリームの長さや情報量の違いを算出するとともに監視者への影響についての評価を行なう必要がある。

謝辞

本研究を進めるにあたり、終始暖かく御指導、御鞭撻を頂いたソフトウェア基礎講座 横矢 直和教授に心より感謝申し上げます。また、本学での研究におけるあらゆる面で、多大な御助言、御協力を賜りましたことを深く感謝致します。

副指導教官として御指導、御鞭撻を頂いたロボティクス講座 小笠原 司教授、ならびに、ソフトウェア基礎講座 竹村 治雄助教授に心より感謝致します。

さらに、本研究を進めるにあたり、御助言や研究環境の整備等で多大な御協力を頂いたソフトウェア基礎講座 山澤 一誠助手、ならびにソフトウェア基礎講座 岩佐 英彦元助手に心より感謝致します。

そして、ソフトウェア基礎講座の 佐藤 哲氏、神原 誠之氏、松宮 雅俊氏、町田 貴史氏の先輩方、二年間苦楽を共に過ごしてきた同輩方、学部の時代より共に学んできたマルチメディア統合システム講座 森中 雄氏、そして、一年間ではありますが様々な面でお世話になった後輩方に感謝の意を表します。

最後に、日々の研究室での生活を支えて下さったソフトウェア基礎講座事務補佐員 北川 知代女史、ならびにソフトウェア基礎講座元事務補佐員 福永 博美女史に心より感謝致します。

参考文献

- [1] D. Coombs and C. Brown: “Real-time smooth pursuit tracking for a moving binocular robot”, *Proc. of CVPR'92*, pp.23–28 (1992).
- [2] T. Uhlin, P. Nordlund, A. Maki and J. O. Eklundh: “Towards an active visual observer”, *Proc. of ICCV'95*, pp.679–686 (1995).
- [3] S. Rougeaux and Y. Kuniyoshi: “Velocity and disparity cues for robust real-time binocular tracking”, *Proc. of CVPR'97*, pp.1–6 (1997).
- [4] Q. Cai and J. K. Aggarwal: “Tracking human motion using multiple cameras.”, *Proc. of ICPR'96*, pp.68–72 (1996).
- [5] H. Mori, A. Utsumi, J. Ohya and M. Yachida: “Human tracking system using adaptive camera selection”, *Proc. of RO-MAN'98*, pp.494–499 (1998).
- [6] J. Gluckman and S.K. Nayar: “Egomotion and Omnidirectional Sensors.”, *Proc. of ICCV'98*, pp.999–1005 (1998).
- [7] D. Gutchess, A. K. Jain and S. W. Chen: “Automatic Surveillance Using Omnidirectional and Active Cameras”, *Proc. of ACCV'00*, pp. 228–233, (2000).
- [8] Q. Huang, Y. Cui, S. Samarasekera, and M. Greiffenhagen: “Auto Cameraman Via Collaborative Sensing Agents”, *Proc. of ACCV'98*, pp.418–425 (1998).
- [9] 三木, 山澤, 竹村, 横矢: “複数の全方位画像センサを用いた遠隔監視システム”, 信学技報, PRMU98-177 (1999).
- [10] 寺沢, 山澤, 竹村, 横矢: “3台の全方位画像センサを使った重なり判定による人物追跡”, 信学ソ大 D-12-58, pp.245 (2000).
- [11] 市川, 山澤, 竹村, 横矢: “高解像度全方位画像センサを用いたイベント検出”, 信学ソ大 D-12-52, pp.239 (2000).

- [12] S.K. Nayar: “Catadioptric Omnidirectional camera.”, *Proc. of CVPR’97*, pp.482–488 (1997).
- [13] J. Yamamoto, J. Ohya and K. Ishii: “Recognizing human action in time-sequential images using hidden markov models”, *Proc. of CVPR’92*, pp.379–387 (1992).
- [14] Y. Yacoob and M. Black: “Parameterized modeling and recognition of activities”, *Proc. of CVIU’99*, pp.232–247 (1999).
- [15] J. Davis and A. Bobick: “The representation and recognition of action using temporal templates”, *Proc. of CVPR’97*, pp.928–934 (1997).
- [16] R. Sagawa, H. Ishiguro and T. Ishida: “Real Time Tracking of Human Behavior with Multiple Omni-directional Vision Sensors”, *Proc. of SII’98*, pp.179–184 (1998).
- [17] F. Z. Brill, T. J. Olson and C. Tserng: “Event Recognition and Reliability Improvements for the Autonomous Video Surveillance System”, *Proc. of Image Understanding Workshop’98*, pp. 267–283 (1998).
- [18] T. J. Olson and F. Z. Brill: “Moving Object Detection and Event Recognition Algorithms for Smart Cameras”, *Proc. of Image Understanding Workshop’97*, pp.159–175 (1997).
- [19] U. Gargi, R. Ksturi and S. Antani: “Performance characterization and comparison of video indexing algorithms” *Proc. of CVPR’98*, pp.559–565 (1998).
- [20] Y. Rui and P. Anandan: “Segmenting Visual Actions Based on Spatio-Temporal Motion Patterns” *Proc. of ICCV’00*, pp.111–118 (2000).
- [21] J. Hong, X. Tan, B.Pinette, R.Weiss and E. M. Riseman: “Image-based Homing”, *Proc. of ICRA’91*, pp.620–625 (1991).

- [22] 八木, 川戸: “円すいミラーを用いた全方位視覚センサによる位置情報の獲得”, 信学論 (D-II), **J74-D-II**, No.1, pp.19–26 (1991).
- [23] 山澤, 八木, 谷内田: “移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案”, 信学論 (D-II), **J79-D-II**, No.5, pp.698–707 (1996).
- [24] 山澤, 市川, 寺沢, 竹村, 横矢: “全方位ビデオカメラを用いた行動記録システム”, 信学ソ大 PD-1-6, pp.333–334 (2000).
- [25] 長井, 久野, 白井: “時空間情報に基づく侵入者監視システム”, 信学技報 PRU95-204 (1996).
- [26] 二宮, 太田: “オプティカルフローによる移動物体の検知”, 電子情報通信学会, パターン認識・メディア理解研究会, 信学技報 PRMU97-28 (1997).
- [27] D. M. Gavrila and V. Philomin: “REAL-TIME OBJECT DETECTION FOR ”SMART” VEHICLES”, *Proc. of ICCV'99*, pp.87–93 (1999).
- [28] 村上, 菊地: “画像処理を用いた侵入者検知装置 — MP30”, 電子情報通信学会, 画像工学研究会, 信学技報 IE95-51 (1995).
- [29] Y. Gong: *Intelligent Image Databases Towards Advanced Image Retrieval*, Kluwer Academic Publishers (1998).
- [30] 山本, 岩佐, 竹村, 横矢: “可変構図テンプレートを用いた色情報に基づく類似画像検索手法の実験的評価”, 画像電子学会第 167 回研究会講演予稿, pp.13-18 (1998).
- [31] 山本, 岩佐, 竹村, 横矢: “色情報の空間分布を考慮した類似画像検索”, 電子情報通信学会, 画像工学研究会, 信学技報 EID98-171 (1999).
- [32] 山澤, 尾上, 横矢, 竹村: “全方位画像からの視線追従型実時間画像生成によるテレプレゼンス”, 信学論 (D-II), **J81-D-II**, 5, pp.880–887 (1998).