

平成18年度 修士論文

複雑な背景における物体検出法の研究
A Study on Object Detection in Complex
Background

指導教授 渡辺 裕 教授

2007年3月

早稲田大学大学院 国際情報通信研究科
国際情報通信学専攻

4205A107-8

伊 谷 裕 介

Yusuke ITANI

目次

| | | |
|------------|------------------------------|-----------|
| 第1章 | 序論 | 1 |
| 1.1 | 研究の背景 | 1 |
| 1.2 | 本研究の目的 | 2 |
| 1.3 | 本論文の構成 | 4 |
| 第2章 | 物体検出の基本技術と利用用途 | 6 |
| 2.1 | まえがき | 6 |
| 2.2 | フレームワーク | 6 |
| 2.3 | 基本技術 | 10 |
| 2.3.1 | 背景差分法 | 10 |
| 2.3.2 | フレーム間差分法 | 11 |
| 2.3.3 | オプティカルフローを用いた方法 | 12 |
| 2.4 | 利用用途 | 13 |
| 2.4.1 | 工業用途 | 13 |
| 2.4.2 | 文書画像 | 13 |
| 2.4.3 | 医療画像 | 13 |
| 2.4.4 | 動画画像圧縮 | 14 |
| 2.4.5 | 顔認識 | 15 |
| 2.4.6 | 特徴点抽出 | 16 |
| 2.4.7 | 監視 | 16 |
| 2.4.8 | 動画画像の画質改善 | 17 |
| 2.5 | むすび | 22 |
| 第3章 | 複雑な背景における物体検出のための従来手法 | 23 |
| 3.1 | まえがき | 23 |
| 3.2 | 局所性に注目した手法 | 23 |
| 3.2.1 | 空間的な動きの強さ | 24 |
| 3.2.2 | 時間平均した動きの強さ | 25 |
| 3.2.3 | 時間的な動きの一様性 | 26 |
| 3.2.4 | 問題点 | 26 |
| 3.3 | 空間的，時間的な相関を用いた手法 | 30 |
| 3.3.1 | 手法 | 30 |

| | | |
|------------|-------------|-----------|
| 3.3.2 | 影及び高輝度領域の除去 | 30 |
| 3.3.3 | 背景の更新 | 31 |
| 3.3.4 | 問題点 | 31 |
| 3.4 | むすび | 32 |
| 第4章 | 提案手法 | 36 |
| 4.1 | まえがき | 36 |
| 4.2 | 予備実験 | 36 |
| 4.2.1 | 特徴量の定義 | 36 |
| 4.2.2 | 予備実験結果 | 37 |
| 4.3 | 提案手法 | 45 |
| 4.4 | むすび | 46 |
| 第5章 | 評価実験 | 48 |
| 5.1 | まえがき | 48 |
| 5.2 | 実験条件 | 48 |
| 5.3 | 評価尺度 | 48 |
| 5.3.1 | 正解マスクの作成 | 48 |
| 5.3.2 | 再現率, 適合率の定義 | 50 |
| 5.4 | 実験結果 | 51 |
| 5.4.1 | 結果画像 | 51 |
| 5.4.2 | 再現率, 適合率 | 52 |
| 5.5 | むすび | 52 |
| 第6章 | 結論 | 58 |
| 6.1 | 総括 | 58 |
| 6.2 | 今後の課題 | 59 |
| | 謝辞 | 60 |
| | 参考文献 | 61 |
| | 図一覧 | 64 |
| | 表一覧 | 65 |
| | 研究業績 | 66 |

第1章

序論

1.1 研究の背景

人物や自動車などの移動物体の位置を認識・追跡しその運動軌跡を復元する技術は、各種監視・防犯システム^(1, 2, 3)や店舗における顧客の振る舞いを把握するなどのマーケティングへの応用、あるいは交通量調査や駐車場の空き検出などITSへの応用⁽⁵⁾といった広範囲な応用分野が期待でき、注目される技術である。また、この他にもバーチャルリアリティやヒューマンインタフェースといった分野、あるいはロボットのナビゲーションシステムに関する分野への応用も考えられる。

人間を検出、認識し移動軌跡を追跡する手法は多く存在している。代表的なものとして、アクティブバッチセンサ⁽⁹⁾やRF-ID⁽¹⁰⁾など、対象物に対してセンサを装着させる手法がある。しかしながらこの手法は、検出する物体に対しセンサを取り付ける必要があるため、あらかじめ検出させるための物体を決めておく必要がある。会社の社員専用の位置情報システムなどには利用が可能であるが、防犯、マーケティングの目的に応用することは難しい。非接触型センサを用いた手法としては床に対して圧力感知や張力感知ワイヤを設置することなどが考えられるが、得られる情報は床に対して設置した際の足あとという離散的な情報であり、人物数が少なければ検出できるが、同時に検出する人間の数が増えた場合には検出することが難しい。さらに設置のコストが高くなりがちであるため、広範囲での利用は難しい。また、一般的に認識対象とされないもの、例えば小鳥や小動物、落ち葉などによる影響に敏感であるという問題もある。広範囲での追跡を可能とする手法として、携帯電話などに搭載されたGPSを用いた位置検出システムなどが提案されているが、逆に屋内環境など限定された空間における観測には精度の面で問題があるとされている。このように、特殊な装置やセンサ類を用いる方法ではその目的に特化したものを使用する性質から認識限界範囲内において比較的精度の高い安定した位置認識・追跡を可能とするが、その反面特殊な用途に限定されることや敷設と保守・点検作業が煩雑となるなどの問題が考えられる。

これに対し、こうした認識・追跡技術を視覚システム、すなわち画像・映像処理を用いて行う研究が各方面で活発に行われており、様々な手法が提案されている。

米国では、1997年から2000年の間に、CMU (Carnegie Mellon University) と David Sarnoff Reserch Center との共同研究として主に軍事目的での使用を対象とした大規模な認

識システムである，VSAM(Video Surveillance and Monitoring)⁽⁶⁾⁽⁸⁾ システムに関するプロジェクトが行われている⁽⁷⁾．CMU キャンパスに12台のカメラを配置しテストシステムを構築した．システムは動画像理解技術により検出した進入物体を複数のカメラが協調してトラッキングし，その状況をリアルタイムで監視員に提示する．これにより監視員の負担軽減と作業効率化に大きく貢献し，新しいビデオ監視システムとして期待されている．

このように映像処理による監視システムが注目されている背景としては近年コンピュータの演算精度の向上や記憶媒体の容量増加などが急速に進み，画像処理ハードウェアの性能が飛躍的に高まったことに伴い，動画・静止画を含めた画像処理技術が目覚ましい進歩を遂げていることが挙げられる．^(1, 2, 3) また，画像を用いる方法には以下の利点があることも理由の一つである．

- 対象物体に対して非接触である
- 小規模な空間から比較的大規模な空間までの観測に対応できる
- カメラを使用することから一般的に特殊なセンサ類に比べ低コストで実現可能
- 複数ターゲットの同時観測，及び検出が可能である．
- 人間が理解可能な可視化された情報が取得できる
- 様々な認識に容易に拡張できる

現状では，監視カメラにより得られた画像は集中監視センターなどに伝送されて監視員が表示映像を確認する例が多い．しかし，光ファイバの敷設などにより多くの画像が伝送可能となり，一人が担当するモニタ数が多くなったため，監視作業の負担が大きい⁽⁴⁾．人間に代わり自動的に異常を発見するための様々な画像認識技術が開発されており，機械による自動監視に対するニーズは極めて高い．近年各分野で物体認識，追跡技術が注目をあびている背景には，こうした自動化ならびに高精度化を実現することが産業界などにおける省力化，省コスト化につながる技術としてその要求が高まっていることがあげられる．

自動監視システムのためには，画像処理を用いて(1)移動物体検出(2)位置認識(3)位置追跡の3つの課題を解決する必要がある(1)では画像中の移動物体領域抽出(2)ではおもに画像情報からの3次元情報復元が課題となる(3)では物体の時系列上の対応付け，すなわち物体の同定が課題となる．また，この他にも照明変動の変化，背景変動，物体同士による相互オクルージョン(遮蔽)，障害物によるオクルージョンなどの問題があり，これらの解決も重要である．

1.2 本研究の目的

監視カメラによる映像監視の流れを図1-1に示す．映像監視ではまず，映像から動物体を検出し，追跡する．そして，その物体が何であるか分類し，物体の構造を把握する．実世界での物体検出では，机上や実験室で想像している以上に，物体検出の識別器を惑わせる様々

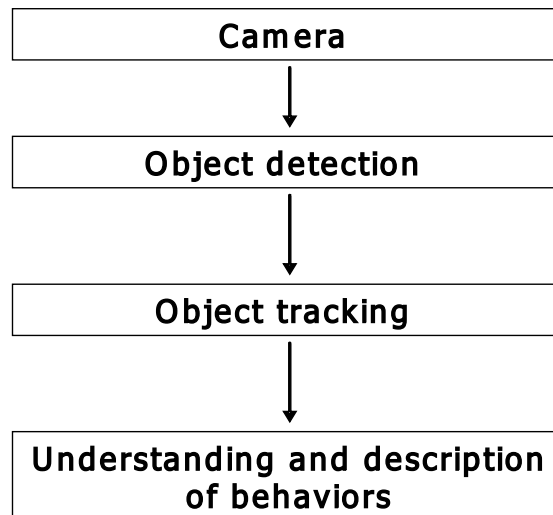


図 1-1: 監視カメラ映像の処理の流れ

な現象が発生する．そのため，実際の観察現場における様々な状況変化の下で安定して観察対象を発見・追尾・確認できるような手法が求められる．物体検出を困難にする様々な現象や制約条件を表 1-1 にまとめる．

本研究では屋外環境を想定している．屋外において，特に重要施設の映像監視セキュリティにおいて，見逃し（検出漏れ）が許されないと同時に日照条件や木の揺れ，光の乱反射など環境の変動に起因する誤検出を低減することが求められる．

光の変化に対する手法として，複数背景手法⁽²⁸⁾の手法では，様々な照明条件下であらかじめ取得・蓄積された背景画像を用いる．状態推定を逐次モンテカルロ法を用いて行い，この蓄積画像の重み付き線形和として背景画像を複数作成する．その中から最も確からしい画像を選択する．確からしさは入力画像との差分をとり，その差分値が閾値以下の画素数により評価する．

木の揺れの問題に対処するためには，動物体と背景の特徴を踏まえた特徴量を定義することが重要となる．従来手法として，画像中に現れる局所的な動きに着目し，動物体の動きとその他背景変動を分離する手法が提案されている⁽¹⁹⁾．この手法では代表速度の検出精度が悪い場合，違った箇所を時間軸をさかのぼりながら追跡することになり誤検出の原因となる．また，フレーム間の追跡で得られた代表速度が一定でなくとも，時間平均の特徴量が小さくなり誤検出の原因となる．そこで，本研究では木の揺れなどの背景変動に着目し，背景変動に対してロバストに物体を検出する手法について検討する

1.3 本論文の構成

以下に本章以降の構成を示す。

第2章 物体検出で使われる基本技術の説明と、物体検出を用いた様々な利用用途についてまとめる。

第3章 今回想定している複雑背景における物体検出の従来手法について述べる。従来手法による物体検出の結果を示し、従来手法の問題点について述べる。

第4章 提案手法のための予備実験について述べる。予備実験により、領域ごとに輝度を平均した画像を用いることで、背景変動によらず物体を検出できることを示す。さらに、予備実験を元に複雑背景における物体検出手法を提案し、その概要を述べる。提案手法では、まず領域に分割する。さらに領域ごとに輝度を平均し差分を取る。

第5章 前章で提案した手法をシミュレーションにより評価、検討を行う。まず、画像による比較を行い、提案手法により精度良く物体を検出できることを示す。次に、再現率と適合率について述べ、これらを用いた評価を示す。

第6章 本研究の総括と今後の課題について述べる。

表 1-1: 物体検出に影響する要因

| 項目 | 現象の説明 | 識別への影響 |
|---------------------|---|--|
| 緩やかな照明変化 | 時間と共に変化する日照変化など | 単純に記憶した背景との差分をとると大きな残差が残り誤検出する。 |
| 急激な照明変化 | 雲や遮蔽物の移動・照明の点灯/消灯による瞬時的照明変化 | 画面全体に物体が現れたかのような過検出が生じる |
| 背景の構造変化 | 移動体が通り過ぎたあとに痕跡が残り、背景が変化する。背景物体が急に移動する | 新たな物体が出現したように誤検出される |
| 揺れる・点滅する | 木や旗がはためくことによる画素の繰り返し変動、信号や広告照明など短周期で点滅する照明とその反射、水面のさざ波や反射など | 周期的である場合が多いが、変化が画像全体で一様ではないことが多く、過検出を発生しやすい。 |
| 検出対象ではない前景 | カメラの前に現れた虫や鳥、背景に投射されたスポットライト | 過検出 |
| 検出対象の Appearance 変化 | 検出対象の形状・色・模様が時間とともに変化する。立体的な対象の視点移動や対象の個体差など | 未検出 |
| 対象物体や他の物体による影 | 検出対象や背景上に他の物体や自分自身の影が投射される | 影の部分を誤検出する |
| 検出対象の長時間静止と再移動 | 検出対象が静止して画像上の変化が止まる | 静止後に見失う。移動後に背景を過検出 |
| 検出対象の隠蔽 | 検出対象の手前にそれを隠蔽する物体が存在 | 部分的にしか見えない。検出不能・誤検出 |
| 背景と類似した対象 | 輝度・色・テクスチャなどの特徴をとっても背景と前景が良く似ている。 | 検出不能 |
| 実例のない前景や背景 | めったに発生しない背景内の現象。システム稼働後に新たに生じる検出対象 | 検出不能・誤検出 |
| 背景の部分的隠蔽 | 歩行者や交通などが途切れることがなく、完全な背景が観測できない。 | 単純に背景を求めると平均化された検出対象が背景に混入し、対象の検出感度が低下する。 |
| 移動するプラットフォームからの撮影 | 自律移動ロボットや自動車などカメラを搭載したプラットフォームが進むため、背景も検出対象も大きく動く | 通常の背景差分では検出不能 |
| カメラの非線形性 | 黒つぶれ・白とびなどカメラの非線形現象 | 照明が変化しただけなのに新たな物体が出現したような誤認識を生じる |
| カメラの揺れ | 風や振動によってカメラが揺れることにより、背景を含めすべての情景が動く | 誤検出または検出感度の低下 |
| 視界不良 | 極度のノイズや照明の直射による | 検出不能 |

第2章

物体検出の基本技術と利用用途

2.1 まえがき

本章では、物体検出の手法として基本となる技術と、物体検出の利用用途について述べる。まず始めに、物体検出の基本的なフレームワークについて述べる。物体検出は、前景と背景の2クラス識別問題であり、いくつかの識別器を組み合わせるフレームワークについて述べる。また、物体検出でよく使われる、基本技術について述べる。基本技術の概要と、それを利用する際の問題について述べる。最後に、物体検出技術を用いた利用用途について述べる。それぞれの利用用途に対し、どのように物体検出を利用しているかについて述べる。

2.2 フレームワーク

物体検出とは入力画像を物体の映っている領域とそれ以外に分類する問題であると考えられる。この場合、問題は前景と背景との2クラス識別問題となり、その枠組みは図2-1で表される。このような枠組みで表現できる物体検出とは、物体と背景とが単一の特徴空間において分離可能な分布を持つ場合であり、たとえば明るさ・カラー・時間変化など画素ごとの特徴量やフィルタバンク・周波数解析により得られる空間的パターン特徴などが評価基準として用いられる。

ところが、現実の応用事例において物体検出を単純なクラス識別問題として定義することは難しい。その理由は、基準データまたは学習サンプルを集める難しさに原因がある。たとえば図2-2において、物体と背景とを識別する識別器に与えられるモデルは、入力画像から得られる特徴ベクトルとそれに対する正解との組み合わせとして学習アルゴリズムに与えられ、識別結果の誤りを最小化するようなモデルが生成される。識別の精度を良くするためには、適切なメトリックと最適な識別境界を決めるための基準データまたは統計的に十分有意な多くの事例を必要とする。

ところが、現実には以下に述べるような理由で、これらを獲得することが困難なことが多い。まず、第一にモデル学習時にすべてのバリエーションを持つサンプルが集められない。たとえば、人物を検出する場合、人物には体格・姿勢・着衣など多くのバリエーションがあり、ある視点から見た全変化範囲を網羅する画像例を集めることは困難である。人とは異なり姿勢や形状の変化の少ない人工物である自動車であっても、車種・姿勢・彩色・積載物な

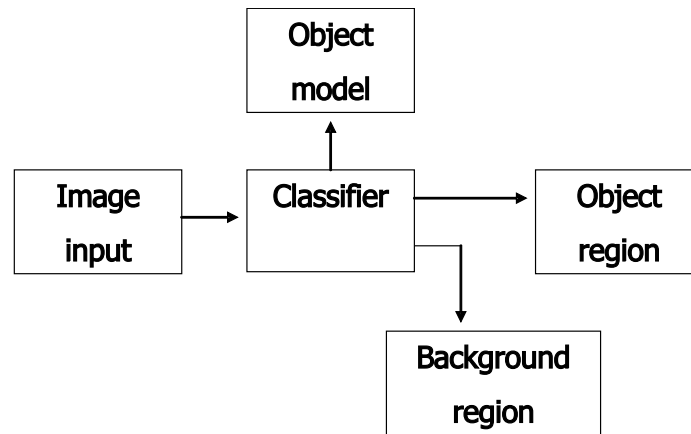


図 2-1: 物体検出の基本的フレームワーク

どで予想できないアピランスの変化がある．たとえば学習後に発売された新製品は実例としてモデリングできない．次に，モデルの精度を高めようとして情景に関する制約条件をうまく使おうとしても，物体と背景とでは利用できる制約条件が異なっている．たとえば固定視点から見た静止背景は画素ごとの重なりがよい．したがって同じ画素は背景上の同一の点を表していると考えると都合がいいが，前景物体は移動するのでこの制約条件は当てはまらない．

このように考えると，一つの識別器だけで物体と背景とを誤りなく識別するよりも，背景と前景に関する制約条件をうまくモデルに取り入れて背景と物体とを別々のメトリックでモデル化して個別の識別器を構成し，順次候補を絞りこんでいくカスケード検出のフレームワーク (図 2-3) や，並列検出器の統合 (図 2-4) が解決策として考えられる．カスケードかパラレルかのアーキテクチャの選択は，個別の検出器の性能と処理コストのトレードオフによって決められる．カスケード構成では個別の識別器で誤って検出対象物体がふるい落とされると誤りを回復不可能なのに対し，並列アーキテクチャではそれぞれの識別器の出力信頼度の強さによって回復可能であるというメリットがあるが，共通する学習サンプルを集めなければならないうえに，最適化すべきパラメータが多いので学習のコストも高いことが欠点である．そのため多くの実用化例ではカスケード構成をとることが多い．なお，顔検出をは

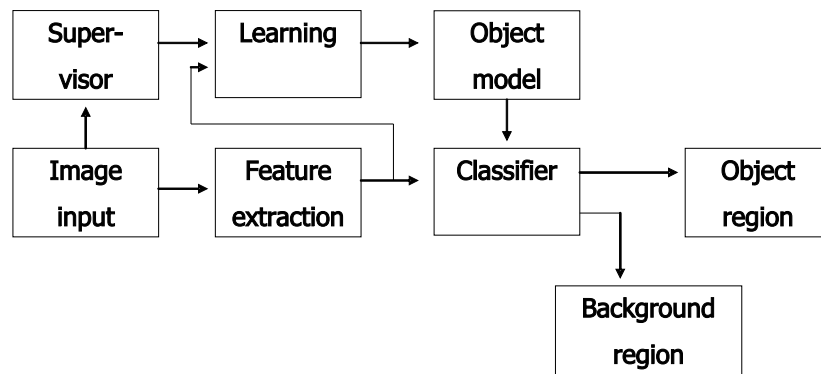


図 2-2: モデル学習を考慮したアーキテクチャ

じめとして最近多用される多数の弱識別器にブースティング学習を組み合わせたアーキテクチャの改良では、カスケードとの併用も見られる⁽²³⁾。

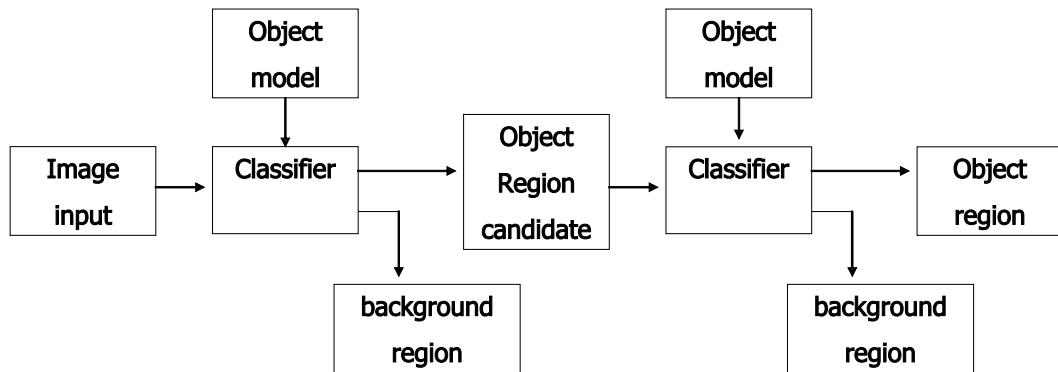


図 2-3: 実用的な物体検出のためのカスケード構成

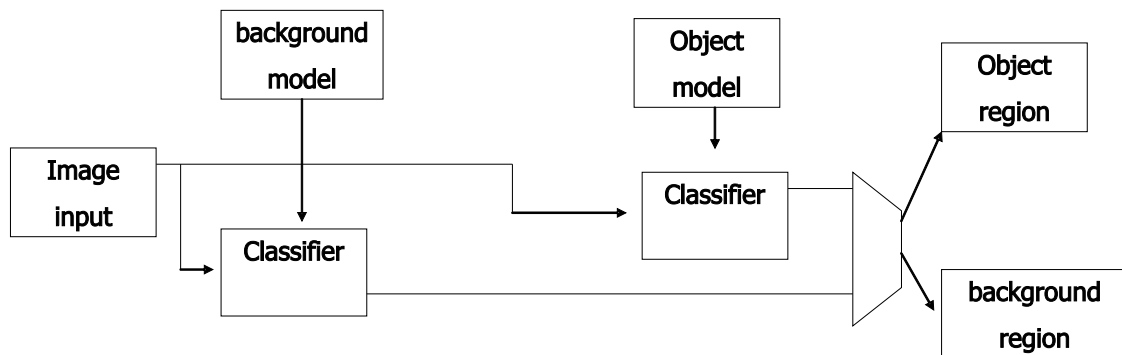


図 2-4: 個別の識別器の出力を並列統合する枠組み

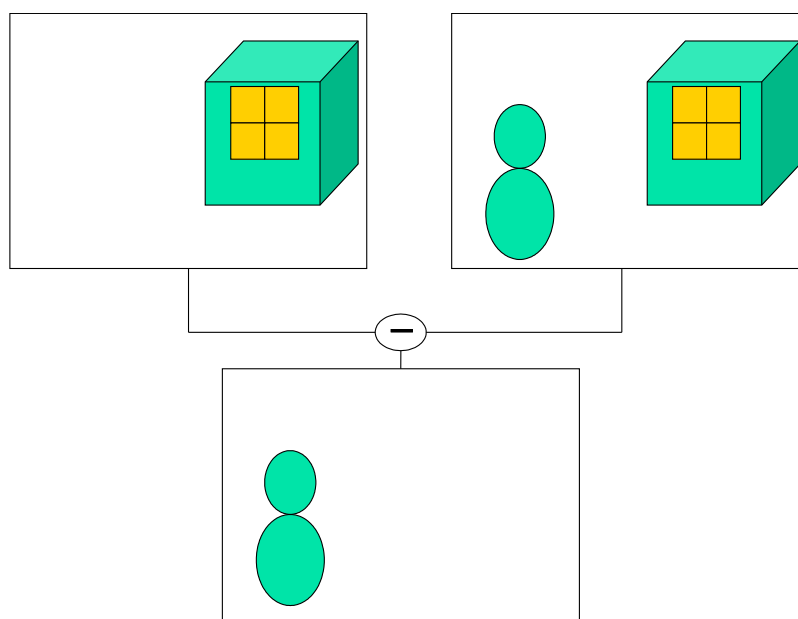


図 2-5: 背景差分法のしくみ

2.3 基本技術

2.3.1 背景差分法

背景差分法とは、あらかじめ移動物体の存在しない背景画像を取得しておき、入力画像と背景画像との差分処理に応じた2値化を行うことによって移動物体領域を抽出する手法である。この背景差分法の基本アルゴリズムを次に示す。ここで、対象とする画像はRGBの各成分を有するRGB画像とする。

$$d_i(x, y, t) = f_i(x, y, t) - b_i(x, y, t) \quad (2.1)$$

$$g_i(x, y, t) = \quad (2.2)$$

ここで、 x, y は画像座標位置、 t は時刻、 $f_i(x, y, t), b_i(x, y, t), g_i(x, y, t)$ はそれぞれ入力画像、背景画像、背景差分画像である。 D は背景差分閾値である。また、 $i = r, g, b$ は各RGB成分を表し、 x, y とともに以後は省略する。背景差分法では、個々の物体領域全体を塊として抽出することができるので、物体抽出性は高い。しかしその反面、物体の中で背景濃度に近い部分が存在する場合、物体を欠落して認識してしまう問題や、時間経過と共に照明の変動など撮影環境が変化する状況においては背景の一部が除去されずに残存してしまい、抽出精度が落ちるなどといった問題がある。また、そもそも移動物体の存在しない背景画像を取得するのが困難な状況も考えられる。これらの問題が生じる環境でも正しく処理を行うためには、背景画像を適応的に生成する必要がある。

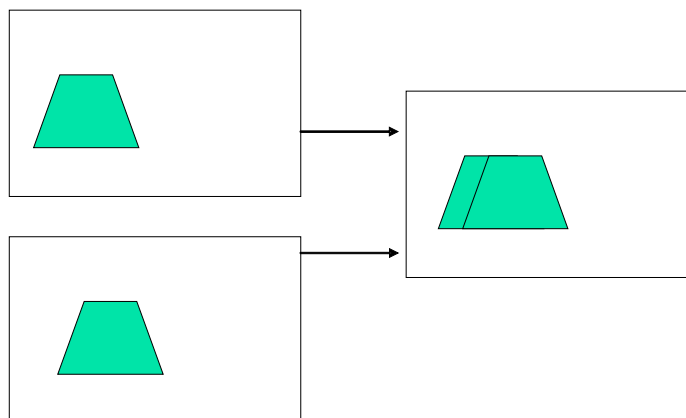


図 2-6: 2 枚のフレーム間差分画像を用いた物体の抽出例

2.3.2 フレーム間差分法

これは、例えば道路上を通過する自動車を撮影するために屋外に設置されたカメラを用いる場合など、天候や日照その他が時間的に変化するため、背景にあたる画像を容易に決定することができない場合がある。このようなときにフレーム間差分を用いる場合が多い。

フレーム間差分法は、時間微分法とも呼ばれ、時間的に連続した隣接フレーム間で差分をとることにより濃淡値が時間的に変化した部分、すなわち移動物体と考えられる領域の一部を抽出する方法である。図 2-6 にその例を示す。

しかしながら、隣接した 2 枚のフレームだけを用いる方法では、フレーム間隔が小さいと物体の移動量も小さくなり、物体のごく一部しか抽出できない。逆にフレーム間隔が大きいと物体が一つである場合は全領域を抽出できるが、複数物体の場合に他の物体と頻りに重なるとその形状を大きく損なう。この問題を防ぐために、連続する 3 枚を用いた方法がある。図 2-7 にその例を示す。

この方法では、図 2-6 のようにそれぞれ連続する 2 枚のフレーム間差分から求めた二枚のマスク画像の論理積を求めることによって、中間のフレーム中の移動物体に対する正確なマスクを生成している。原理的には単純であるが、その効果は大きい。この方式では、移動物体同士が重なった領域に穴が生じることがある。このため、差分画像を左端から水平右方向、右端から水平左方向にそれぞれ走査し、差分値の大きな画素を検出後、それらの中間を移動領域であると強制的にみなして移動物体を塊として抽出することによりこの問題をさけることが考えられる。⁽²²⁾

フレーム間差分法では、照明変動や影、反射像の影響などに頑健であるという利点があるが、移動物体の濃度値が一定であるような場合は移動した部分であっても差分値が 0 となっ

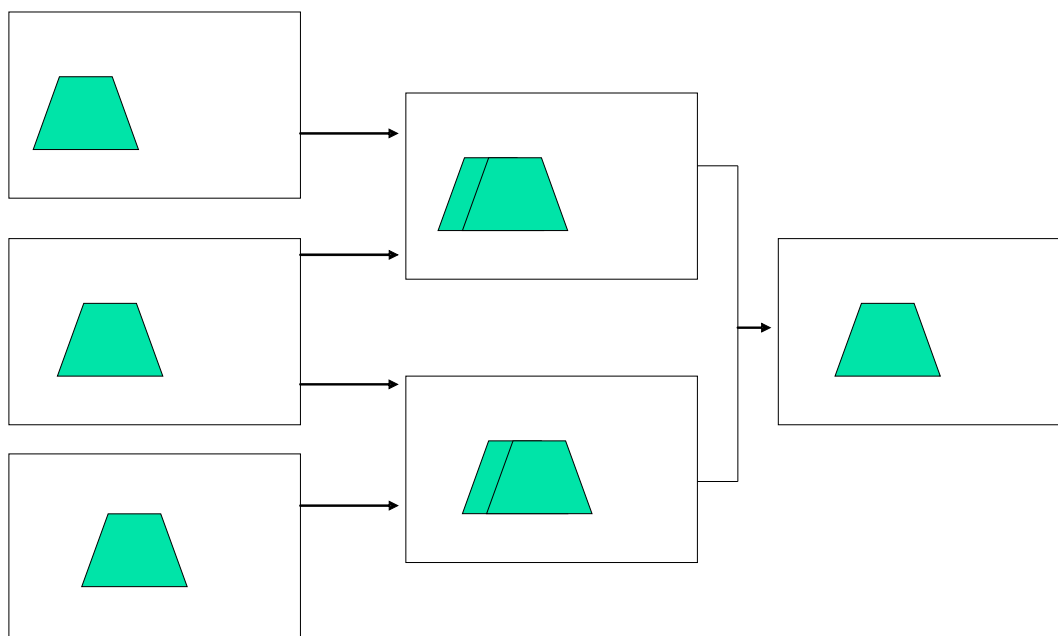


図 2-7: 3 枚のフレーム間差分画像を用いた物体の抽出例

て検出できないという問題や、移動物体が一時的に静止している場合には対応できないといった問題がある。

2.3.3 オプティカルフローを用いた方法

オプティカルフローとは、動画像中の移動物体の動きを知るために良く用いられるもので、その速度ベクトルの分布のことをさす。MPEGなどの動画像圧縮において動き補償に用いられる動きベクトルとの違いは、オプティカルフローのほうがより短いフレーム間での時間間隔を想定しており、速度場での解析が中心となる点である。このオプティカルフローは時空間勾配法などを用いて求められ、特に剛体の動きを求めるのに有用である。

このオプティカルフローを移動物体検出に応用した手法としては、オプティカルフローからの物体形状復元を応用した手法や、ステレオカメラ法による距離情報とオプティカルフローを統合した手法、あるいはオプティカルフローを用いた一様フロー領域分割と物体運動予測に基づく投票による手法などが提案されている。オプティカルフローではシーンの全体的な輝度値の変化に敏感でないことから、照明条件の変化に対して比較的頑健であるといった利点があるが、オプティカルフローには求めるべき変数に対して非線形で問題が複雑になること、実用的な精度が得られないこと、計算量が多くリアルタイム処理に適さないなどの

問題があり、盛んに応用されるまでにはいたっていない。これらの問題に対し、オプティカルフロー自体の精度、信頼性を考慮した手法が多く提案されており、またリアルタイムでの処理を目指した専用ハードウェアを用いた手法なども提案されているが、物体が一時的に静止している場合において抽出結果が得られないことや、背景領域に局所的かつ微小変動する部分がある場合ランダムなフローが抽出されて解析が困難になるなどの問題が残る。

2.4 利用用途

2.4.1 工業用途

工場における組み立て作業の自動化などに用いられる産業用ロボットビジョンでは、既知の二次元、三次元形状と外環とを持つ工業用部品を、位置・姿勢が自由な条件で検出し、その位置と姿勢を求める⁽¹⁷⁾。

対象は工業部品のようにほぼ同じ形状と模様を持つので、認識対象の見え方をモデルとして持つことが一般的である。見え方のモデルとしては、人工的な照明を当ててその反射パターンを二値画像として粒子解析するものが多いが、見え方を二次元パターンとしてパターン間の距離計算を行うテンプレートマッチングなどが実用化されている。

また、工業用途として、外観検査がある。外観検査においては欠陥や異常が検出対象となるが、欠陥や異常は発生頻度が少なく実例を用いてその外観をモデル化することができない。そのため、正常な外観を持つサンプルがモデルとして用いられ、モデルから離れた外観を持つ領域を欠陥として抽出する枠組みがとられる。

工業用品では、認識対象の周りを取り囲む安定した環境が作りやすく、対象の形や模様が均一であることが多い上に、高いスループットとコストパフォーマンスが求められる。

2.4.2 文書画像

文書画像⁽¹⁷⁾において、まず無地の背景の上に記された文字や記号を正確に抽出する技術、特に照明の輝度の不均一を補いつつ、文字の掠れを許容するための二値化アルゴリズムとして技術が発展した。この場合、文字・図・記号が物体であり、それ以外の部分が背景であると解釈できる。

1980年代以降、従来の無地背景ではなく写真やイラストの上に書かれた複雑な背景の文書画像からの文字抽出が行われた。また画像と文字の混在する複合文書が扱われ、文書画像を文字領域とその他の領域とに分割する手法が研究された。この場合、個別の文字ではなく、その並びや領域がそれぞれ物体として扱われていることになる。

2.4.3 医療画像

医療画像処理において、まず、レントゲン画像、CT/MRIなどの再構成画像、超音波エコー反射画像などの画像を対象に臓器や器官の抽出が行われた。ただし、医療画像の場合

には、診断対象の臓器が存在する画像上の位置が撮影時の設定により決まっていることが多く、臓器の存在と概略の位置を前提として、その輪郭や境界を正確に求めるという事例が多い。

X線画像における骨組織のように、高コントラストに撮影される場合は少なく、多くの認識対象の境界は他の器官と重なって映ったり、ノイズと見分けるのが困難だったりする。そのため、低コントラストの画像への既知の形状の輪郭の当てはめ問題として取り込まれている。

また、腫瘍などの病変部位の抽出も医療画像の物体検出技術として重要である^(?)。この場合、大きさや形状がまちまちであることが多く、テクスチャなど内部のパターンの持つ特徴や輪郭の特徴などを用いて正常組織との識別問題として取り込まれることが多い⁽¹⁷⁾。

2.4.4 動画画像圧縮

高レベルのソースモデルを実際の画像符号化に適用するために、動きに対して規則性のある領域を得ることが必要である⁽¹⁸⁾。このアプローチとして

- 輝度の一様性を仮定してクラスタリングや領域分割・成長法により領域分割を行った後、その領域について動きを求める。
- 動き情報を直接用いて、共通の動きパラメータで記述される画素集合として領域を分割する。

の2つが考えられる。画像符号化には頑健性と処理速度を考慮して領域分割・成長法やクラスタリングが使われる。前者のアプローチの代表的な手法として、“watershed”と呼ばれる領域成長法が知られている。watershed法では画像中の均一領域にマーカとの境を輪郭として抽出する。高速な演算アルゴリズムが存在し、かつ必ず閉輪郭が得られることに特徴がある。

後者のアプローチとして、画像 R が領域内で動きが一様な c 個の領域断片に分解されると考える。位置 (x, y) の観測された動きベクトルを $v(x, y)$ 、パラメータで推定される動きベクトルを $v(x, y) = (x' - x, y' - y)^t$ として、 $\sum_w \sum_{(x, y)} |v(x, y) - v(x, y)|$ を最小化するように動きパラメータと領域 σ_i を決定することが問題となる。動きベクトルをもとに領域分割を行う方法は、はじめに得られた動きベクトルの精度に依存している。一般に領域分割に利用できるだけ十分に動きベクトルが安定して得られていない場合が多い。そこで動きパラメータにより予測される輝度と観測される輝度との誤差和を直接最小化しつつ輝度誤差を距離とした画素のクラスタリングにより領域分割を行う手法が提案されている。これらの手法では、直接輝度誤差を最小とする領域分割と動き推定を行うため、画像再構成を目的とする画像符号化に適している。

このように領域に分割し、画像シーケンスを複数のレイヤに分割してそれぞれ圧縮し、再構成時にレイヤの隠蔽関係を考慮して合成することにより圧縮率を高めたり精度を最適に制御することができる。

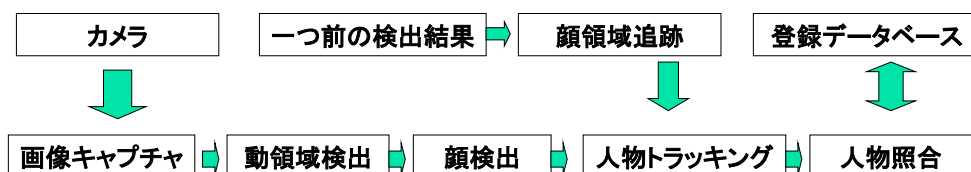


図 2-8: 顔認証システム

2.4.5 顔認識

人の顔は指紋や静脈と異なり，日常生活において常にさらけ出された数少ないバイオメトリクスの一つである．双子は区別できないなど問題はあがるが，入手しやすいという利点によりパスポートや犯人の指名手配などに利用され有効な人物認証の一つである．

顔認証は，入手した顔画像をあらかじめ登録された複数人の顔画像と照合する技術である．凹凸のある3次元をした顔を2次元に射影するため，顔の一方しかとらえられない，また照明の当たり方により見え方は様々である．自動監視顔認証システムの流れを図2-8に示す．⁽¹³⁾

顔画像検出

監視カメラから送られてきた画像をキャプチャする．次に背景差分を用いて動きがあった領域のみを抽出し，この領域に対して顔検出処理を行う．現在顔の検出には2つの方法がある．一つはニューラルネットワークを用いた複数人顔検出処理であり，もう一つは一つ前の画像から検出された顔領域をオプティカルフローにより領域検出を行い，顔位置を推定する方法である．そして検出された顔に対して顔認識処理を行うことにより検出された人物が登録されているかを判断する．

また，顔の位置検出は，Viola と Jones が Haar-like 特徴を用いた Boosting 識別器による方法⁽¹⁴⁾ を発表してから高い検出率が実現されている．特徴量として，照明変動に強い Haar-like 特徴を用いた点が大きく貢献しているが，膨大な数の特徴量を Boosting で自動選択し，さらに非線形でかつ汎化性能が高い Adaboost の利用も効果的である．Boosting は，少ない特徴で早期に非顔を判断するカスケード方式により高速化され，弱識別器の信頼性を用いた Real Adaboost⁽¹⁴⁾ で性能が向上している．これらの成果により，現在では任意の回転，向きを持った顔をリアルタイムで検出できている．

2.4.6 特徴点抽出

顔画像の正規化を行うために顔の部位の特徴点として、円形の瞳・鼻孔とコーナーエッジからなる目じり・目頭・口端を精度良く抽出する。(13) 瞳・鼻孔の抽出には、それらの特徴である暗い円形領域を取り囲む明るい領域を見つけるために、分離度フィルタと画像マッチングを併用し、高速に精度良く位置を同定する。目じり・目頭・口端の抽出には、それらの特徴であるコーナーエッジを見つける Harris フィルタと画像マッチングを併用する。さらに、特徴点の幾何学的位置の拘束条件を用いて、抽出された複数の候補から最もそれらしい組み合わせを判定することにより、個々の部位の位置を総合的に決定する。画像マッチングも局所的な当てはめの他に、幾何的な関係を利用して段階的により広範囲に行い、特徴点の位置決めをする。

顔向き補正

人が積極的に協力してくれない限り、認証に理想的な顔画像が取得できないという問題がある。そこで、顔の姿勢変化を正規化する処理が必要となる。あらかじめ準備した3次元の顔形状モデルを用いて、任意の2次元顔画像の部位の位置と姿勢を正規化する。個々の人の3次元顔形状モデルを準備できないので、平均的な顔形状モデルを準備し、位置あわせに必要な特徴点と2次元顔画像から得られる特徴点を利用して変換のための写像行列を推定する。

顔識別

顔向きや照明変化に左右されない画像が得られた後、個々人の顔画像の特徴を辞書として登録する。目的は他の人との識別であるため、個々人に固有な特徴量を抽出し辞書とする。これまでに個人から得られる複数の顔画像が成す部分空間を求め、さらにその部分空間が個々に大きな正準角を成すような空間に写像し部分空間どうしの距離を離すことにより識別能力を向上させてきた。個々人の特徴量が他者と異なることから、個々人の部分空間が互いに直交するように変換することによってこれまでよりも高い識別精度が得られるようになった。

2.4.7 監視

基本処理

監視の対象は移動物体または異常事象である。物理量を計測して直接検出できるものではなく、まず背景差分やフレーム間差分により変化領域を抽出する。

監視システムが屋外で適用される場合には照明や陰影が変化する。背景物体が揺れる。対象物体が停止するなど変化領域が対象物体かどうかの判断が困難なことが多い。屋内であっても夜間に適用されることが多く、照明が不十分、室外の照明変化が影響して誤認識するなどがある。

検出された変化領域が対象物体かそうでないかを判断するために物体のモデリングが必要である。道路監視や車両についてはITS関連分野で、人については動作解析を含めた詳細なものが表情・ジェスチャ認識の分野で多くの提案があり、テレビ電話または個人認証のための顔のモデリング、移動軌跡を求めるためのモデリングなどもある。

物体計測処理

フレームごとに監視対象の検出処理を行い、時系列処理で誤りを修正する手法、フレームごとには変化領域抽出処理にとどめ、時空間画像に変換してから監視対象の検出処理を行う手法などがある。精度についても、駅の乗り換え通路の人の流れのような群の係数の場合、店内の顧客動線のように同一人物を追跡する必要がある場合、さらには顔部分を検出して個人認証をする必要がある場合などがある。

状況の理解

交通事故の検出、駐車違反の検出などを目的とする場合には単に車両を検出するだけでは不十分である。事故の検出のためには壁に衝突したなどの判定のため、車両の2次元位置の時系列変化を計測する必要がある。また、駐車違反を検出するためには車両が停止して、ドアを開けてドライバーが車外に出たなど動作の時系列を検出し、さらに顔画像を獲得して個人を特定することが必要となる。移動物体間の相互関係に基づいた事象の推定が行われる。

対象物体の3次元位置の推定

対象物体の誤検出の原因が陰であったり、路面に書かれた標識であったり、背景の無意味な変化であったりすることが多く、対処に多くの工夫がされている。一方では複数のカメラを用いることが可能になってきたので、ステレオのカメラが用いられることも多くなった。厳密な3次元計測を目的とするのではなく、変化領域の判定に距離情報を用いるなど顕著に精度を高めることも可能である。

2.4.8 動画像の画質改善

JPEG 2000 のビットストリームを復号後、特定の輝度値の変化パターンを検出し、輝度値の補正を行う画質改善手法がある⁽²¹⁾。これは、前後フレームを用いたフィルタを用いている。しかし、このフィルタだけでは動領域の雑音が残ってしまうため、動物体と背景を分ける必要がある。

この手法には、輝度値の変動を抑えるフィルタと、動領域と静止領域を判定する部分とからなる。以下それぞれについて説明をする。

フィルタ

フィルタは、フリッカーの検出，補正からなる．

- フリッカーの検出

前後フレームの画素値の波形に注目し，波形が振動していたらその部分をフリッカー雑音が発生している場所とし検出する．前フレームや後フレームの値と異なる値であったとしても，振動している波形でない場合はフリッカー雑音とみなさない．

以上のことを数式で表すと以下のようになる．

$$V_i - V_{i-1} > 0, V_i - V_{i+1} > 0 \quad (2.3)$$

$$V_i - V_{i-1} < 0, V_i - V_{i+1} < 0 \quad (2.4)$$

式 (1),(2) の場合にはフリッカー雑音として認識する．

$$V_i - V_{i-1} > 0, V_i - V_{i+1} < 0 \quad (2.5)$$

$$V_i - V_{i-1} < 0, V_i - V_{i+1} > 0 \quad (2.6)$$

式 (3) , (4) の場合には，前フレームとの値は異なっているが，振動していないため，フリッカー雑音として認識しない．

- 補正

フリッカー雑音の検出後，検出された箇所に対して補正を行う．補正は前フレームの輝度値にそろえることによって行う．

図 2-9 にフリッカー雑音の検出，補正の様子を示す．

静動領域判定

図 2-10 にある一点による輝度変化を示す．このように，静止領域から動領域に変わるとき，輝度が大きく変化することが確認される．そのため，閾値を設定することにより，この変化を検出することが可能であると考えられる．

このため，静動領域判定は以下の方法により行う．

- 8 × 8 のブロックに分ける
- ブロックごとに，前フレームの対応するブロックとの差分値 D を計算する

$$D = V_i - V_{i-1} \quad (2.7)$$

- ブロック内で差分値の 2 乗の和がある閾値を超えた場合，その領域を動領域とする．

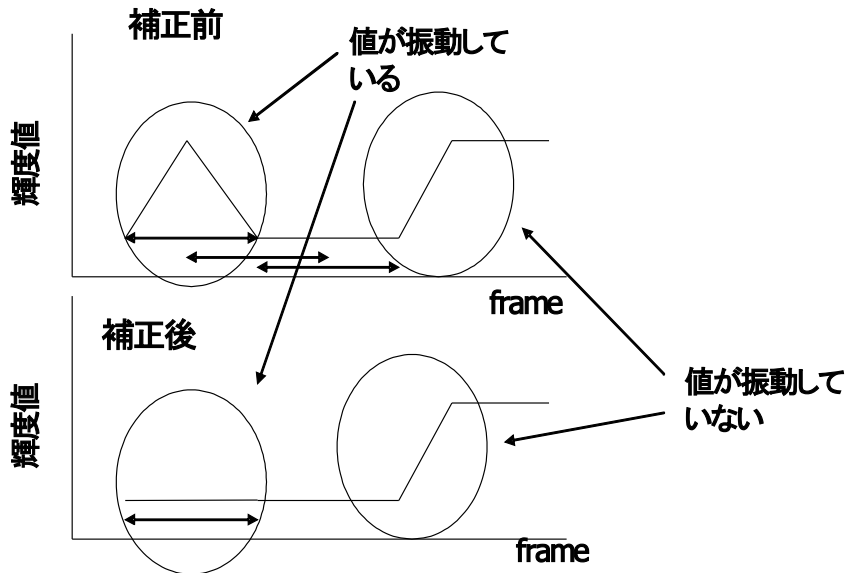


図 2-9: フリッカー雑音の検出, 補正の様子

動領域

$$\sum D^2 > threshold \quad (2.8)$$

静止領域

$$\sum D^2 < threshold \quad (2.9)$$

実験結果

図 2-11 に示すように, 動領域にこのフィルタをかけると, 残像が残ることがある. また, この手法による, 領域判定の結果を図 2-12 に示す.

図 2-13 に示すように, 領域判定をすることで提案フィルタのみのときに発生した残像が残らずに処理ができた.

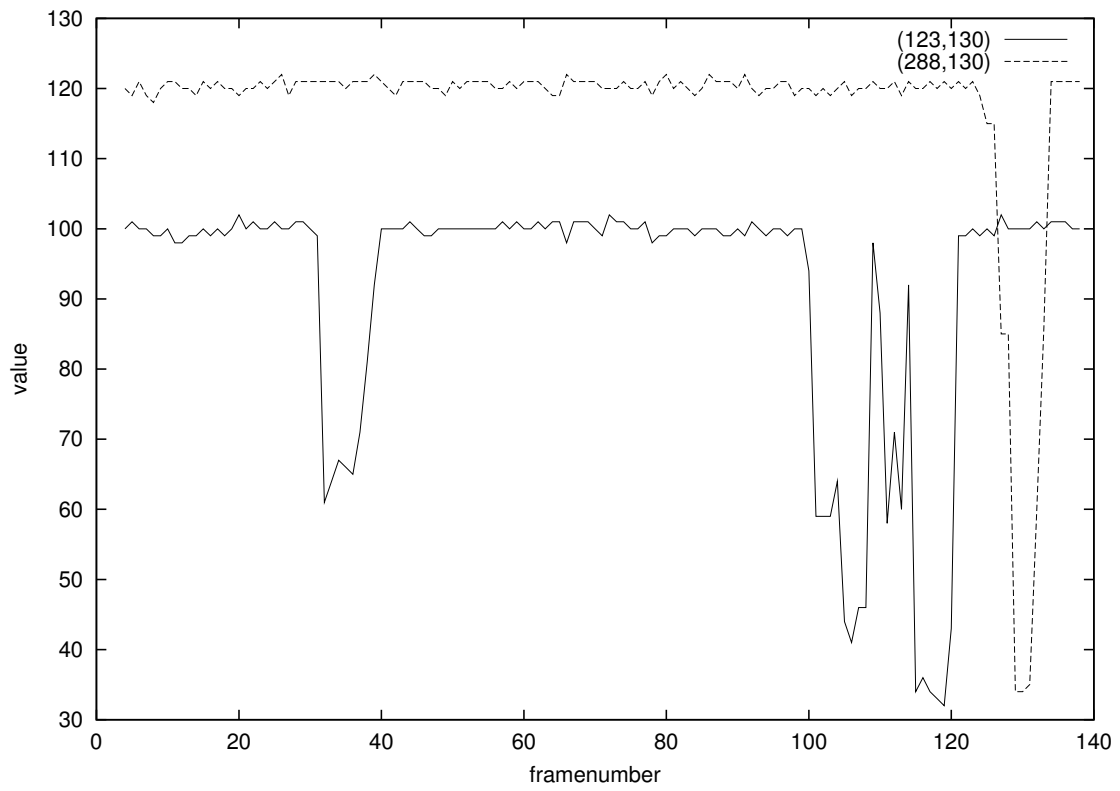


図 2-10: 輝度変化の様子



図 2-11: 残像



図 2-12: 動領域判定



図 2-13: フィルタ処理後

2.5 むすび

本章では、物体検出の基本技術、および利用用途について述べた。

まず、物体検出の基本技術について述べた。基本技術として、背景差分法、フレーム間差分法、オプティカルフローを用いた手法について説明し、その問題点について述べた。

また、物体検出の利用用途について述べた。利用用途として、工業用途、医療画像、顔検出、監視について述べ、それぞれどのように利用しているかについて述べた。

第3章

複雑な背景における物体検出のための従来手法

3.1 まえがき

本章では、複雑な背景における物体検出の従来手法について述べる。従来手法は、物体の動きに着目した手法と空間的な相関に着目した手法に分けられる。前者の手法として、画像の局所的な動きに着目した手法について述べる。これは、空間的な変化と時間的な動きの変化に注目する。後者として、空間的、時間的相関に着目した手法について述べる。それぞれについて、概要を述べ、実験結果を示す。最後に、問題点を述べる。

3.2 局所性に注目した手法

複雑背景に対する手法として、画像中に現れる局所的な動きに着目し、動物体の動きとその他背景変動を分離する手法が提案されている⁽¹⁹⁾。

これは、水面の光の乱反射の動き、及び木の揺れの動きは、

- 空間的累積結果に動きがゼロの瞬間があること
- 時間的累積結果におけるピークの相関値と原点における相関値との差は不明瞭である。

という特徴があること、これに対して人の動きは、

- 時間的累積結果に動きがゼロの瞬間がない
- 時空間的累積結果に単峰性ピークが見られ原点における相関値との差が明瞭である。

という特徴による。また、空間平均した動きの強さや時間的な動きの一様性には人の動きと背景の変動の間に違いがみられる。そのため

- 空間平均した動きの強さは高い値
- 時間平均した動きの強さは高い値
- 各時刻の空間平均した動きの強さはいずれも高く、途中で動きがゼロにならない

の3つの判定基準を用いて多元的に判定する。

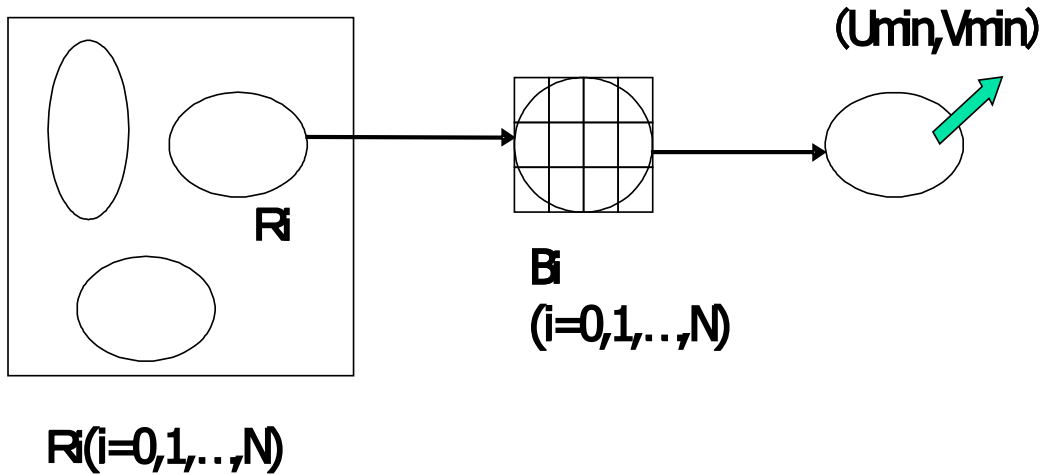


図 3-1: 従来手法

3.2.1 空間的な動きの強さ

空間平均した動きの強さとは、変化領域内で求めた局所的な相関値マップのうち、ゼロでないもののピーク位置がどれだけまとまりがよいかを表す尺度である。

図 3-1 に示すように、まず背景差分により検出した領域 R_j の中に割り当てた各ブロックごとに、時刻 t の輝度パターンをテンプレートとして切り出し、時刻 $t - 1$ の画像を探索画像として以下の式で示すように輝度差分の絶対値総和 (SAD) を計算する。

$$S_{B_i}(u, v) = \sum_{(x,y) \in B=i} |I_{t-1}(x + u, y + v) - I_t(x, y)| \quad (3.1)$$

ここで、領域 R_j 内の各ブロック B_i の相関値マップ $S_{B_i}(u, v)$ のうち、動きがゼロでないものを累積し、正規化した相関値マップのピークを判定する。各局所相関値マップのピークが空間的に一様な場合、各ピークが強めあい累積した相関値マップに明瞭なピークが現れるが、空間的に不均一な場合、各ピークが弱めあい累積した相関値マップのピークは平坦となる。

$$S_{R_j}^* = \frac{1}{N_{R_j}} \sum_{B_i \in R_j} \{W_{B_i} \cdot S_{B_i}(u, v)\} \quad (3.2)$$

ただし, W_{B_i}, N_{R_j} を次のように定める.

$$W_{B_i} = 1(\operatorname{argmin}_{(u,v)} S_{B_i} \neq (0,0)) \quad (3.3)$$

$$N_{R_i} = \sum_{B_i \in R_i} W_{B_i} \quad (3.4)$$

この $S_{R_j}^*$ を用いて, 空間的な動きの強さを次のように定式化する.

$$F_1 = S_{R_j}^*(0,0) - S_{R_j}^*(u_{\min}, v_{\min}) \quad (3.5)$$

ただし, $(u_{\min}, v_{\min}) = \operatorname{argmin}_{(u,v)} S_{R_j}^*(u, v)$ である.

3.2.2 時間平均した動きの強さ

時間平均した動きの強さは, 変化領域を追跡する過程で得られる空間的に累積した相関値マップのうち, 原点でない相関値マップのピーク位置がどれだけまとまりがいいかを表す尺度である. 前節で求められた変化領域の代表速度 (u_{\min}, v_{\min}) を用いて, 領域 R_i に属する全ブロック B_i を (u_{\min}, v_{\min}) だけ移動させた位置において時刻 $t-1$ の画像から輝度パターンをテンプレートとして取り出し, 時刻 $t-2$ の画像を探索画像として相関値マップを計算する. 以後同様に各時刻の代表速度を用いて順次時間軸をさかのぼる方向に追跡処理を全 L 回繰り返す.

空間的な動きの強さと同様に, 時間的な動きの強さに関しても累積した相関値マップ $S_{R_j}^{*(k)}(u, v)$ のうち, 動きがゼロでないものを累積投票し, 正規化した相関値マップ $S_{R_j}^{**}(u, v)$ のピークを判定することにより調べる.

$$S_{R_j}^{**}(u, v) = \frac{1}{N} \sum_{k=0}^{L-1} \{W_{R_j}^{(k)} \cdot N_{R_j}^{(k)} \cdot S_{R_j}^{*(k)}\} \quad (3.6)$$

ただし, $W_{R_j}^{(k)}, N$ は以下の式で求める.

$$W_{R_j}^{(k)} = 1(\operatorname{argmin}_{u,v} W_{R_j}^{*(k)}(u, v) \neq (0,0)) \quad (3.7)$$

$$N = \sum_{k=0}^{L-1} (W_{R_j}^{(k)} \cdot N_{R_j}^{(k)}) \quad (3.8)$$

空間的な動きの強さは以下の式で定義される.

$$F_2 = S_{R_j}^{**}(0,0) - S_{R_j}^{**}(u_{\min 2}, v_{\min 2}) \quad (3.9)$$

3.2.3 時間的な動きの一様性

時間的な動きの一様性とは、追跡過程で得られる変化領域の代表速度全体のうち、ゼロでない動きベクトルが占める割合を表す尺度であり、追跡過程における前述の空間平均した動きの強さがしきい値を超えたものの数で表せる。以下の式で定義する。

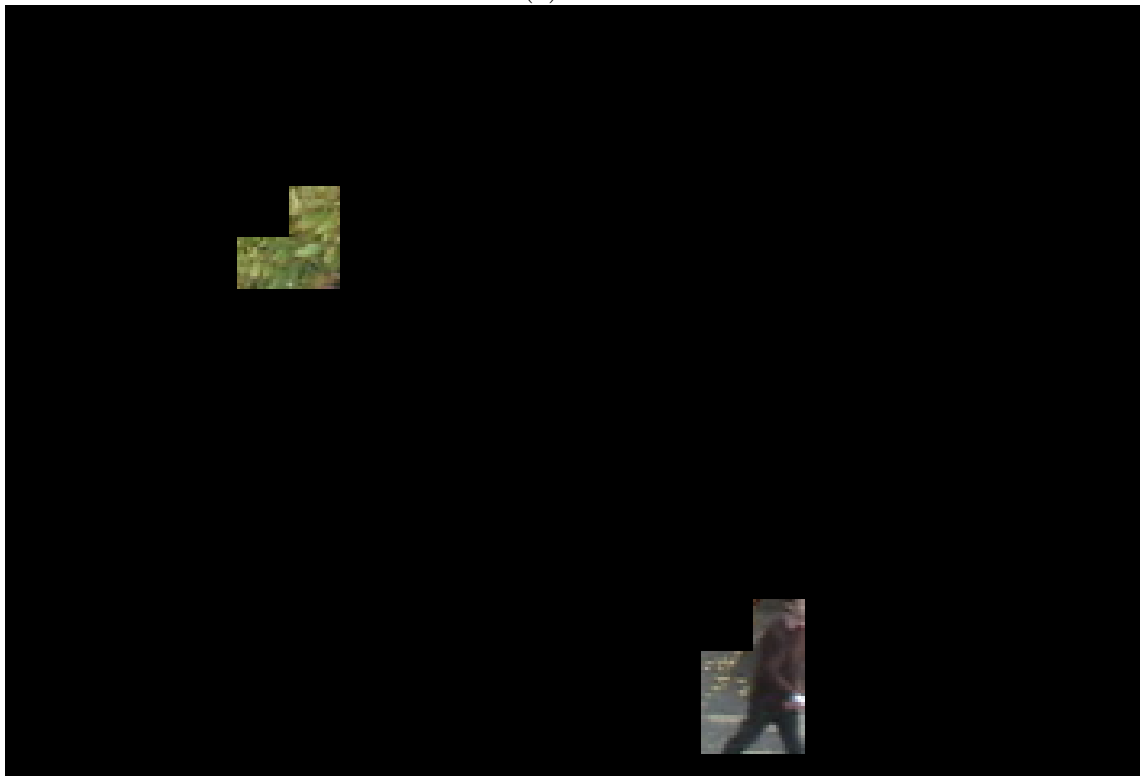
$$F_3 = \sum_{k=0}^{L-1} W_{R_j}^{(k)} \quad (3.10)$$

3.2.4 問題点

実験結果を図 3-23-33-4 に示す。この手法では代表速度の検出精度が悪い場合、違った箇所を時間軸をさかのぼりながら追跡することになり誤検出の原因となる。また、フレーム間の追跡で得られた代表速度が一定でなくとも、時間平均の特徴量が小さくなり誤検出の原因となる。



(a) 原画

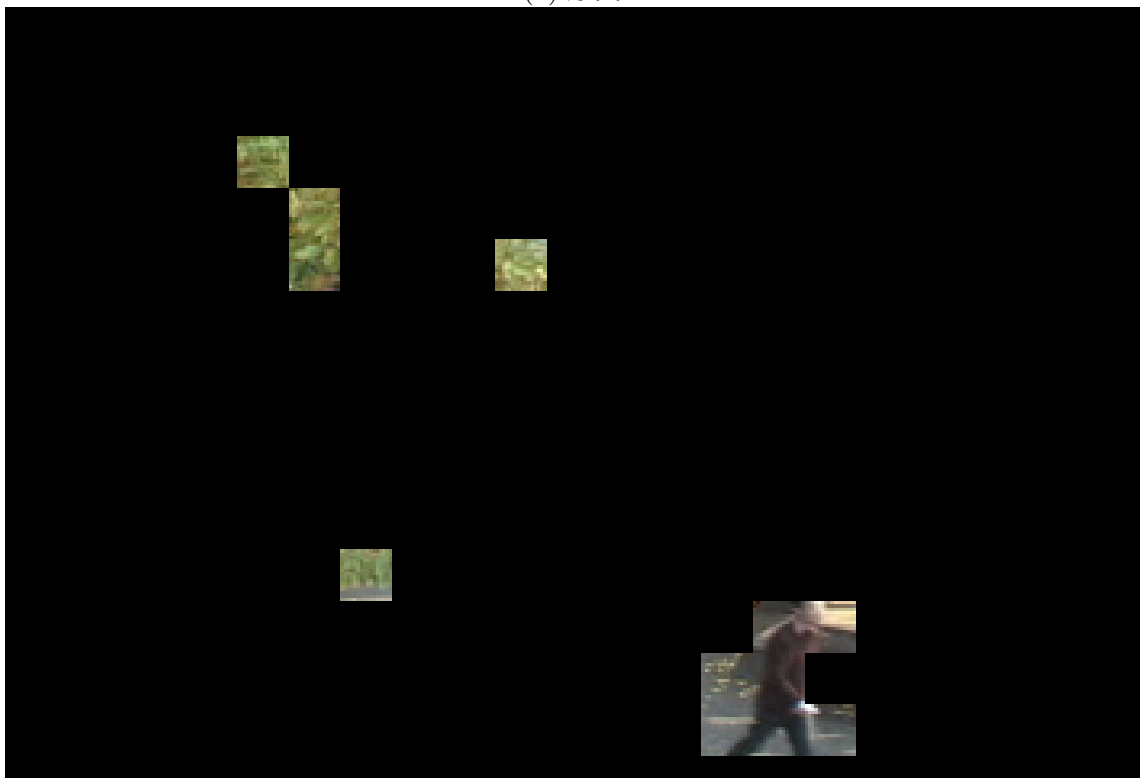


(b) 結果

図 3-2: 従来手法の実験結果 (1 フレーム目)



(a) 原画



(b) 結果

図 3-3: 従来手法の実験結果 (2 フレーム目)



(a) 原画



(b) 結果

図 3-4: 従来手法の実験結果 (3 フレーム目)

3.3 空間的，時間的な相関を用いた手法

3.3.1 手法

背景に影や高輝度の領域をふくんだ動画像における物体検出法が提案されている。(24)

従来では，画素ごとに一定フレームの中央値及び標準偏差を用いて，背景のノイズを想定し背景差分法を用いていた(25)．多数のオブジェクトが存在する場合，この方法では背景のノイズをきちんと測定することができないという問題がある．この手法では，この問題に対応するため，フレーム一枚一枚の輝度のヒストグラムを考えたとき，最初に来るピークが背景の平均ノイズであるとしている．

また，以下の2つの仮定を用いる．

- 空間的な一様性

一般的に，前景の画素は孤立して現れることない．つまり，周囲の画素と結びついて現れる傾向にある．そのため，以下の式を用いる．

$$D_t = |I_t(x, y) - \lambda(x, y)| * A \quad (3-11)$$

$$A = w = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (3-12)$$

これは， (x, y) を中心とする3x3の隣接する画素を重みづけしたものである．

ここで， $I_t(x, y)$ はフレーム内の位置 (x, y) における輝度値， $\lambda(x, y)$ は (x, y) を中心として $I_t(x, y)$ の平均分散である．

- 時間的な一様性

前景オブジェクトは連続的に動いているので，複数のフレームの同じ位置の画素は前景である傾向にある．つまり，もし (x, y) の画素がフレーム t で前景であるならば， $t+1$ フレーム目の (x, y) の画素は前景である可能性が高い．そのため，上の式を用いて，以下のように定式化する．

$$ND_t(x, y) = \frac{1}{2}(D_t(x, y) + D_{t-1}(x, y)) \quad (3-13)$$

以上の式を用いて，以下の式でしきい値処理を行う．

$$I_t(x, y) \in \text{foreground} \text{ if } ND_t(x, y) > k\sigma_m \quad (3-14)$$

ここで， σ_m はフレーム内の輝度のヒストグラムにおける，最初のピーク値を表す．

3.3.2 影及び高輝度領域の除去

影の領域として観測される輝度は入射光に比例している，つまり周りの背景の輝度と比べて同じスケールで暗くなると仮定する．

このとき、 $R(x, y)$ を (x, y) を中心とした 3×3 の領域とし、影の画素を以下のように定義する。

$$std_R\left(\frac{I_t(x, y)}{\lambda(x, y)}\right) < L_{std} \text{ and } L_{low} \leq \left(\frac{I_t(x, y)}{\lambda(x, y)}\right) < 1, \quad (3.15)$$

ここで、 $std_R\left(\frac{I_t(x, y)}{\lambda(x, y)}\right)$ は領域 R における $\frac{I_t(x, y)}{\lambda(x, y)}$ の標準偏差である。 L_{std} は隣接領域の最大分散を調節するしきい値であり、 L_{low} は非常に暗いオブジェクトを影として誤分類しないためのしきい値である。

同様に、高輝度領域を以下の式で定める。

$$std_R\left(\frac{I_t(x, y)}{\lambda(x, y)}\right) < L_{std} \text{ and } 1 \leq \left(\frac{I_t(x, y)}{\lambda(x, y)}\right) < L_{high}, \quad (3.16)$$

ここで L_{high} は非常に明るいオブジェクトを誤分類しないためのしきい値である。 L_{std} , L_{low} , L_{high} はいくつかの動画像による実験により、以下のように定める。

$$L_{std} = 0.05, L_{low} = 0.5, L_{high} = 1.3.$$

3.3.3 背景の更新

背景の更新は長い間影や高輝度として分類されたそれぞれの画素に対して適用する。 $C_t(x, y)$ を位置 (x, y) の画素が影や高輝度領域として分類された場合 1 を返す二値のマスクとする。そして M を更新時間（背景を更新するのに使われるフレーム数）とする。このとき、時間 t のフレームに対して、以下の式を計算する。

$$A_t(x, y) = \sum_{t-M+1}^t C_i(x, y), \quad (3.17)$$

これは、画素 (x, y) が前 M フレームに対し影もしくは高輝度領域として表された回数を示している。背景モデルは、以下の式を満たす場合に再計算される。

$$A_t(x, y) \geq pM, \quad (3.18)$$

ここで、 p は $0 \leq p \leq 1$ であり、背景更新を必要とする更新時間の最小を決めるパラメータである。ここで、実験的に、 p , M を定める。

$$p=0.8, M=50$$

3.3.4 問題点

この手法を用いたときの結果を図 3-53-63-7 に示す。

実験結果を見るとわかるように、影や日照変動など光の影響に関しては誤検出がないが、木の揺れなどの背景の変動には弱いことがわかる。そのため、この手法では十分ではない。

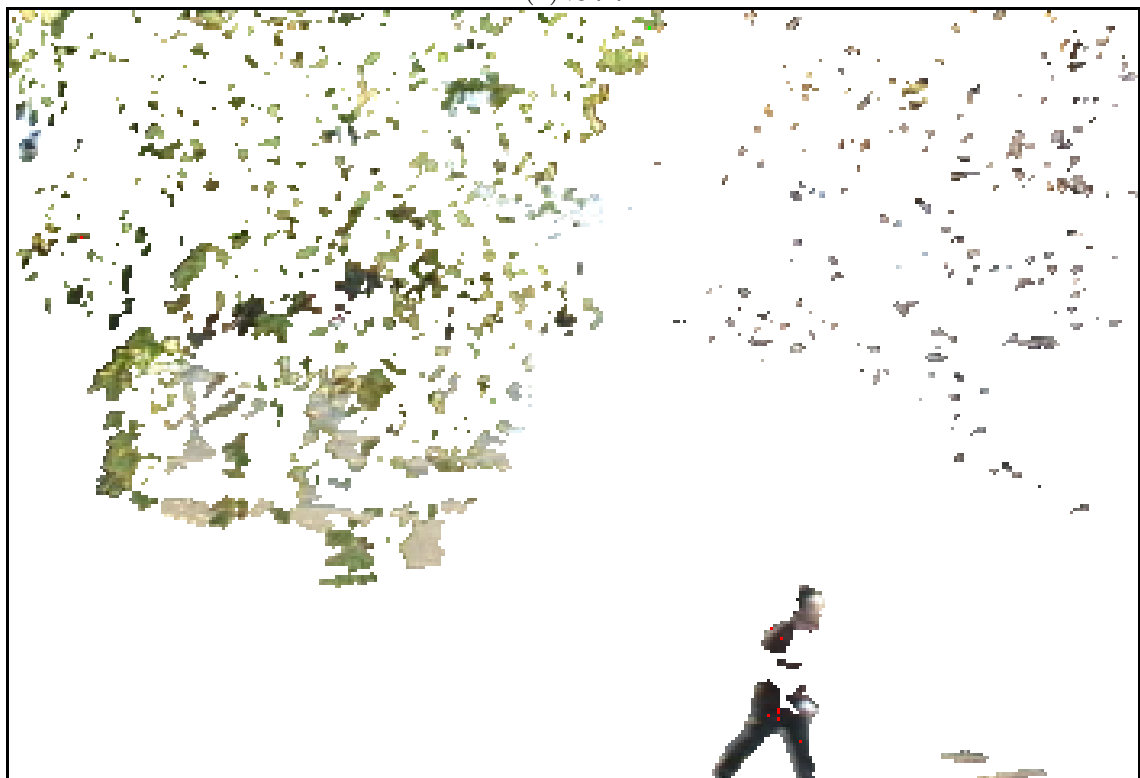
3.4 むすび

本章では、本研究で想定している、複雑な背景に対する物体検出法の従来手法について述べた。局所性に注目した手法として、空間平均した動き、時間平均した動きの強さ、各時刻の空間平均した動きの強さに注目した手法について述べた。この手法は代表速度の検出精度が悪い場合、違った箇所を時間軸をさかのぼりながら追跡することになり誤検出の原因となる。また、フレーム間の追跡で得られた代表速度が一定でなくとも、時間平均の特徴量が小さくなり誤検出の原因となることを述べた。

また、空間的、時間的な相関を用いた手法として、空間的一様性、時間的一様性に着目した手法について述べた。この手法は影や日照変動など光の影響に関しては誤検出がないが、木の揺れなどの背景の変動には弱く、想定している環境では十分な精度が得られないことを述べた。



(a) 原画



(b) 実験結果

図 3-5: 従来手法 2 の実験結果 (1 枚目)



(a) 原画



(b) 実験結果

図 3-6: 従来手法 2 の実験結果 (2 枚目)



(a) 原画



(b) 実験結果

図 3-7: 従来手法 2 の実験結果 (3 枚目)

第4章

提案手法

4.1 まえがき

本章では、まず背景変動のある領域と検出する動物体の特徴の違いをみる予備実験とその結果について述べる。そしてその特徴量を用いた物体検出法の提案手法について述べる。

予備実験では、まず今回使う特徴について述べる。特徴量として、画像を領域ごとに輝度を平均し、そしてそれをフレーム間で差分を取ることを述べる。次に、その特徴量により、背景変動と動物体を分離できることを実験結果で示す。そしてこの実験結果を基に、提案手法について述べる。提案手法は、まず領域分割を行う。分割の仕方として、色情報と動きの向きを用いることを述べる。そして、予備実験で示した特徴を用いて背景変動と物体を分離する手法について述べる。

4.2 予備実験

4.2.1 特徴量の定義

本節では、背景変動と動物体を分類するために、必要となる特徴について述べる。図 4-1 のように、木が左から右側に動いた場合を考える。この場合、左上のブロックは木のない領域に対し木が入る変化が起きる。そのため、輝度値が下がる変化である。一方で、右下のブロックは木がなくなる変化が起きる。そのため、輝度値が増える変化である。つまり、左上のブロックと右上のブロックとでは全く逆の変化が起きていることがわかる。これより、時間 t のフレームの (x, y) 座標における輝度を $f(x, y, t)$ とすると、木などの揺れの場合、注目画素の輝度値 $f(x, y, t)$ の時間方向の差分 $f(x + i, y + j, t) - f(x + i, y + j, t - 1)$ をその周囲で和をとると 0 に近づくことが予想される。

$$\sum_{i,j} f(x + i, y + j, t) - f(x + i, y + j, t - 1) > TH \quad (4.1)$$

$$\sum_{i,j} f(x + i, y + j, t) - f(x + i, y + j, t - 1) < TH \quad (4.2)$$

つまり、式 (4.1) のときは動物体、式 (4.2) のときは背景であると考えられる。また、木の揺れなどの周期性を持つ変化の場合、時間方向の変化を考えたとき、その領域の輝度の差分

和の変化はほとんどないと考えられる。

ここで、式(4.1)(4.2)はある領域を規定したときに、その領域の差分和を考えている。そのため、領域ごとに輝度を平均した画像を考え、その差分を考えることと同じである。

$$\frac{1}{N} \sum_{i,j} f(x+i, y+j, t) - f(x+i, y+j, t-1) \quad (4.3)$$

$$= \sum_{i,j} \frac{f(x+i, y+j, t) - f(x+i, y+j, t-1)}{N} \quad (4.4)$$

$$= \sum_{i,j} \frac{f(x+i, y+j, t)}{N} - \sum_{i,j} \frac{f(x+i, y+j, t-1)}{N} \quad (4.5)$$

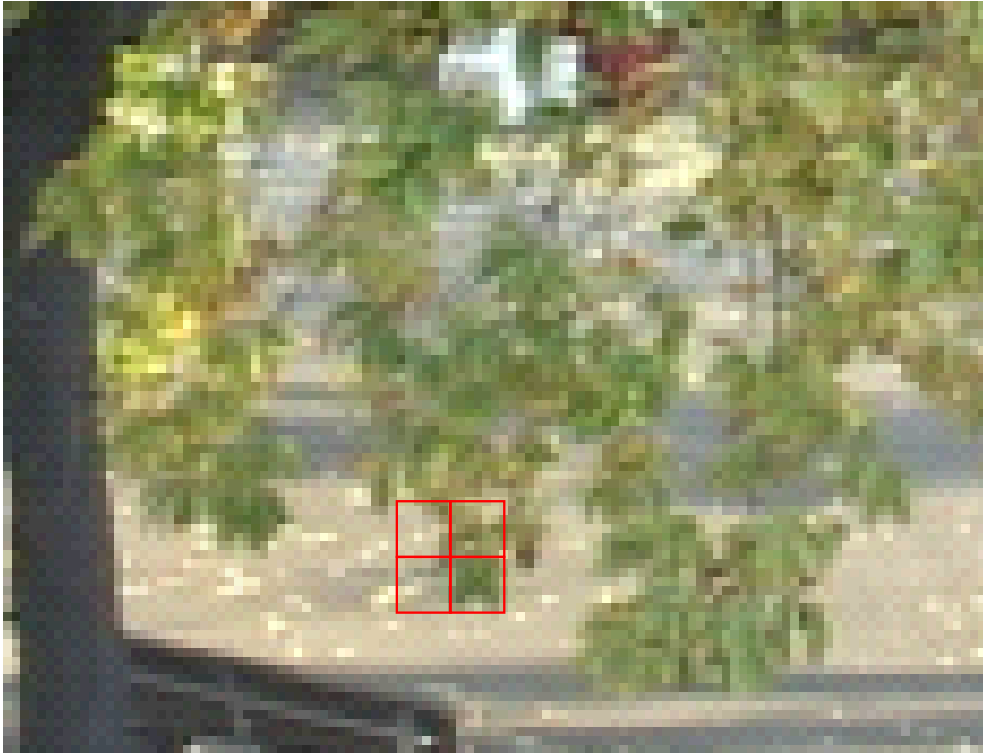
フレーム間差分と、領域ごとに平均化したフレーム間差分の性質の違いを見るために予備実験を行った。実験では、フレームを8×8のブロックに分け、ブロックごとに輝度値を平均化した画像を用いた。この画像を図4-4に示す。実験に用いた原画像を図4-2に示す。

4.2.2 予備実験結果

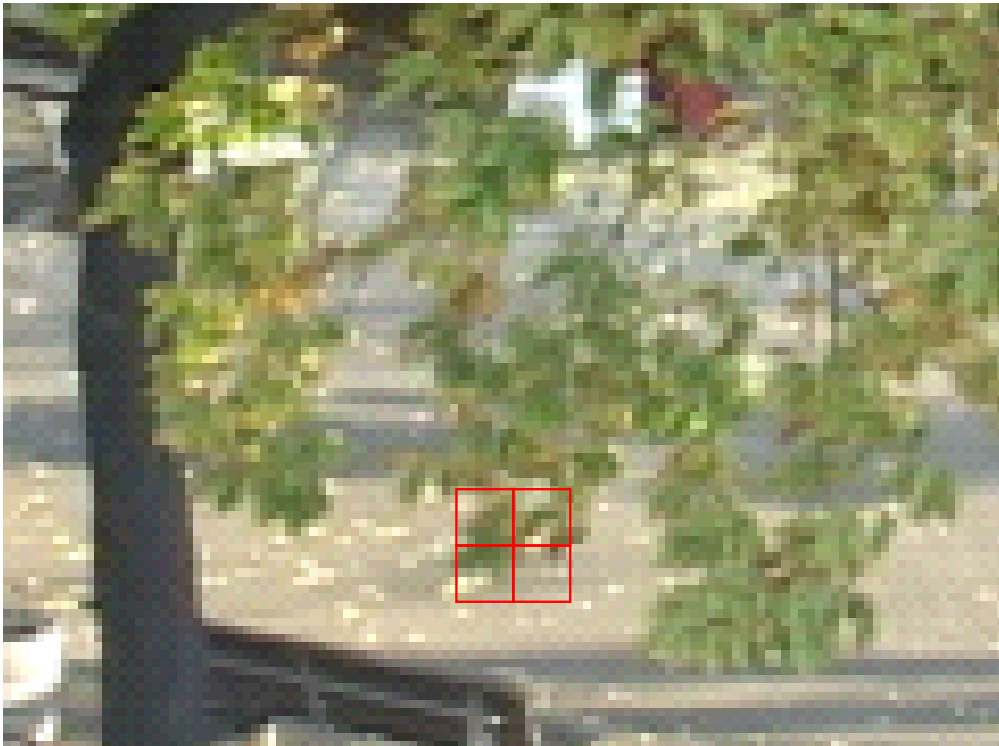
図4-3, 4-5に単純なフレーム間差分の結果と平均画像を用いたフレーム間差分の画像の結果を示す。図4-3, 4-5より、フレーム間差分の画像は木の揺れによる変化が目立つのに対し、平均画像を用いた場合のフレーム間差分はそれが目立たない。そして人の領域は平均画像の差分であっても目立つことが確認できる。

また、その際の木領域、人が動いた領域、木などが背景におけるフレーム間の輝度の差分値のグラフを図4-6, 4-7に示す。図4-6, 4-7より、人の動きの変化に比べ、木の揺れや背景のほうが輝度の変化が小さいことが確認される。

また、平均画像の輝度差分値を時間方向に3フレーム分足したもののグラフを図4-8に示す。これより、時間方向の和も考慮することで、より精度良く分類が可能になることがわかる。



(a)1 枚目



(b)2 枚目

図 4-1: アプローチ



(a)1 枚目



(b)2 枚目

图 4-2: 原画像

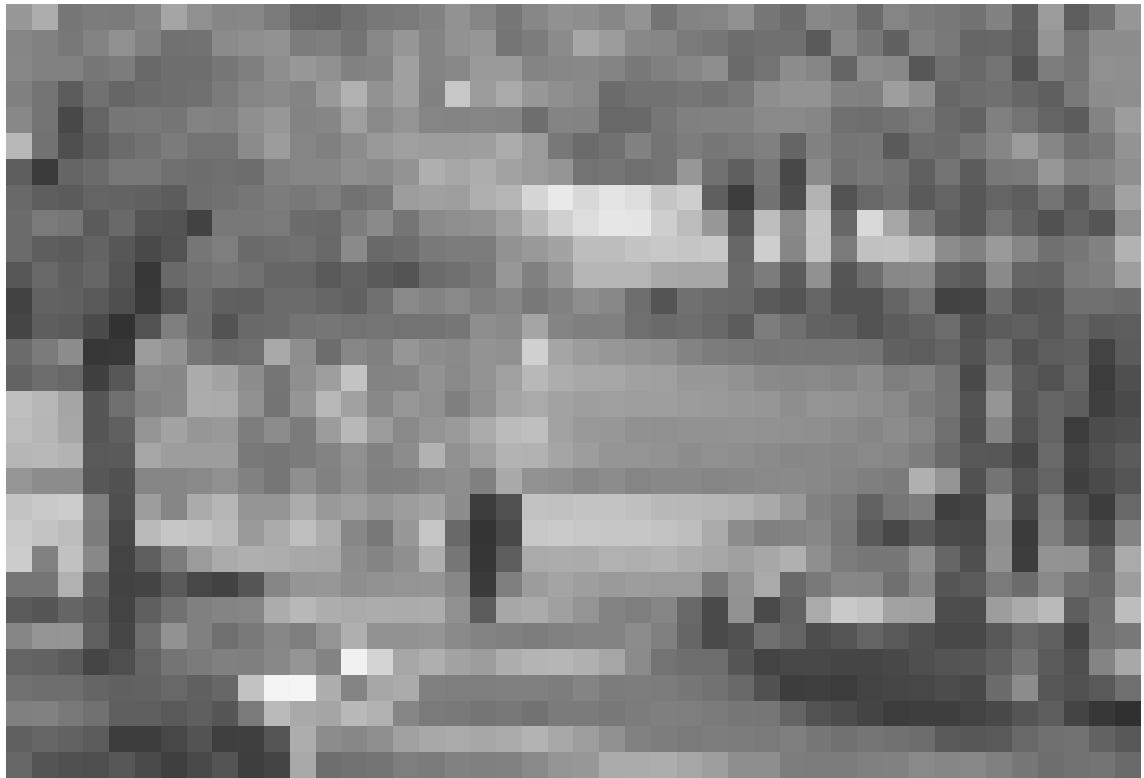


(a)1 枚目

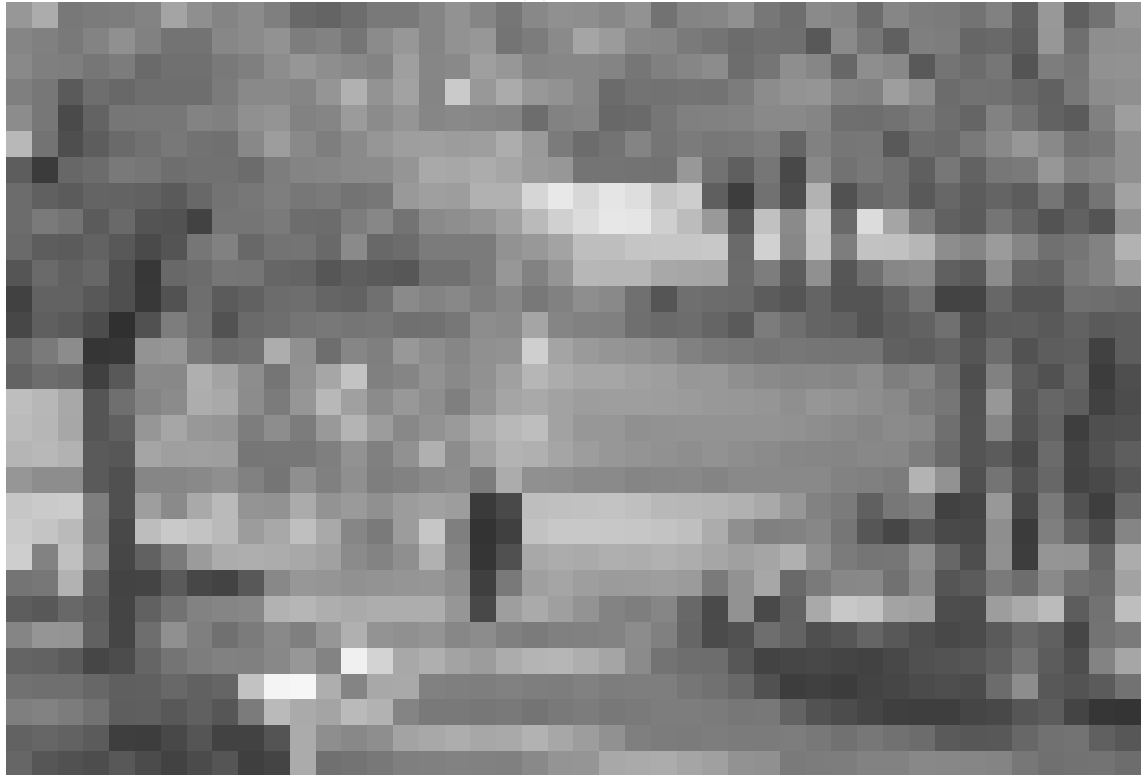


(b)2 枚目

図 4-3: フレーム間差分



(a)1 枚目



(b)2 枚目

图 4-4: 平均画像



(a)1 枚目



(b)2 枚目

図 4-5: 平均画像のフレーム間差分

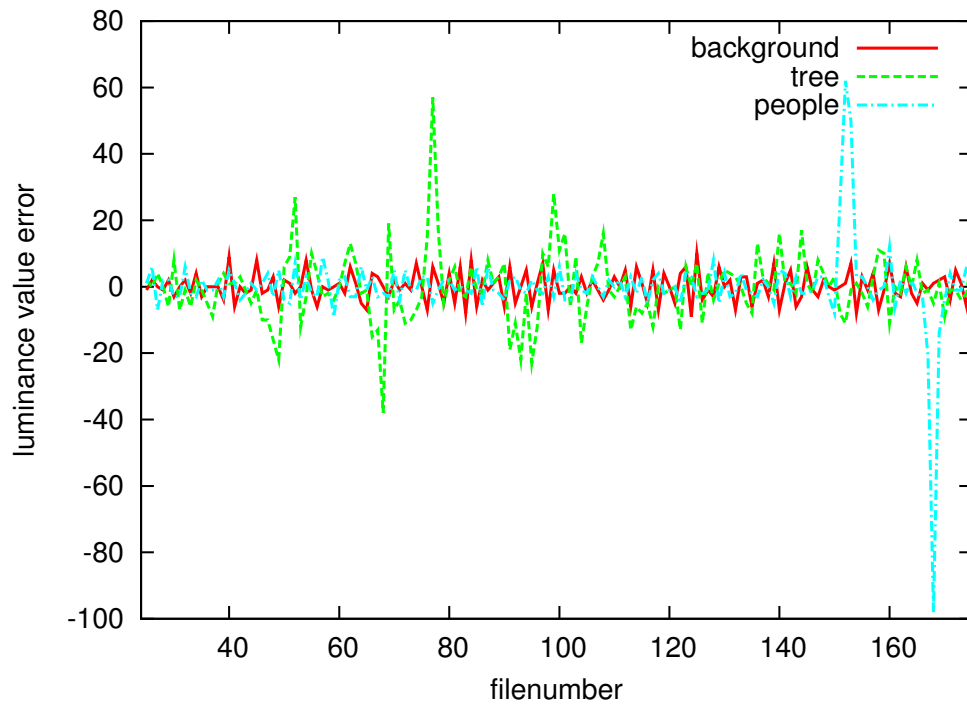


図 4-6: フレーム間差分の輝度変化

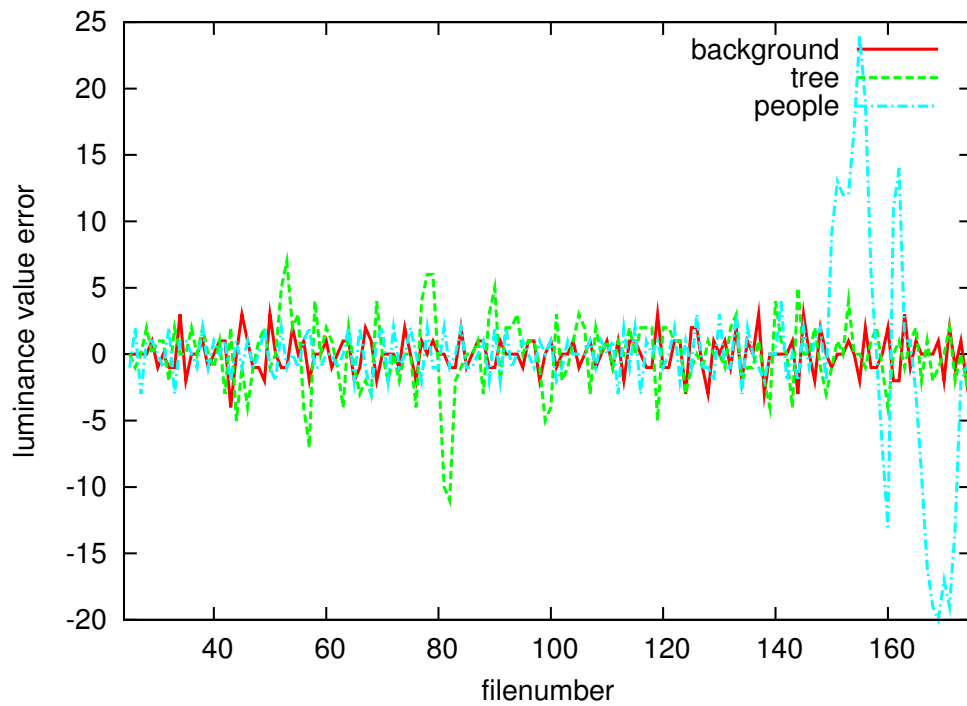


図 4-7: 平均画像の輝度変化

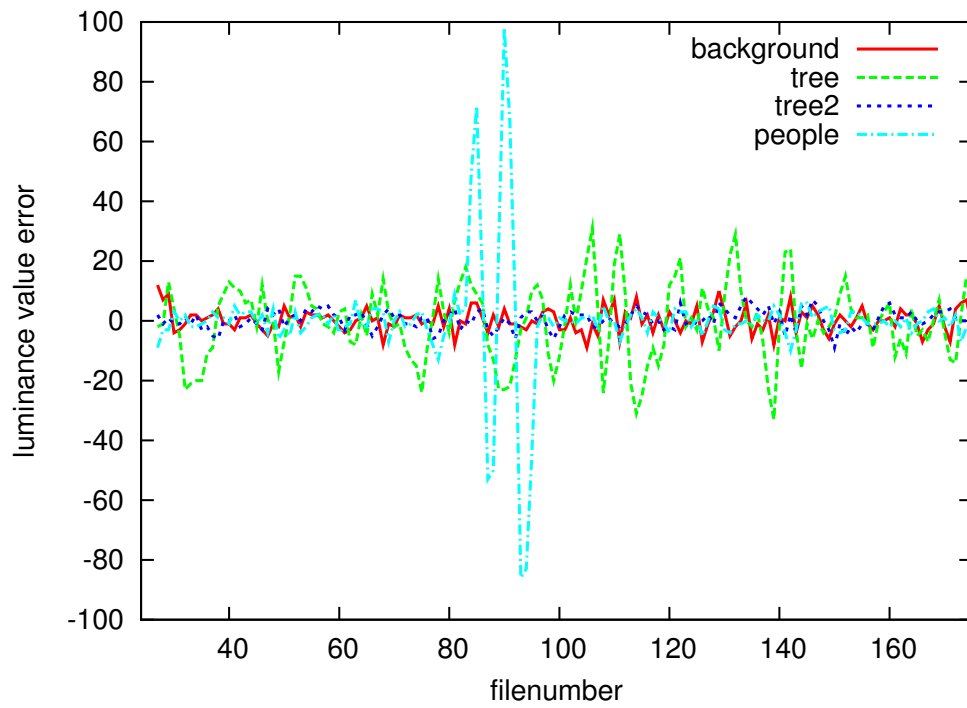


図 4-8: 平均画像の輝度変化 (3 フレーム蓄積)

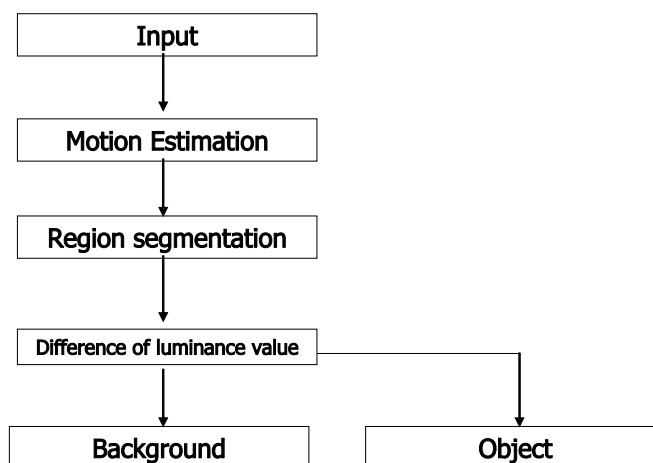


図 4-9: 提案手法の流れ

4.3 提案手法

本節では、前節の予備実験の結果を踏まえ、提案手法について述べる。提案手法の流れを図 4-9 に示す。

提案手法では、まず領域分割を行う。領域分割は以下の手順で行う。

- 8x8 のブロックにわけると。
- 隣接ブロックを比較し、濃度平均がしきい値 $THnode$ 以下の場合は統合する。

$$f(x, y, t) = \frac{1}{3}(\sum R(x, y, t) + G(x, y, t) + B(x, y, t)) \quad (4.6)$$

$$\frac{1}{N} \sum_{i,j} f(x, y, t) - f(x + i, y + j, t) > THnode \quad (4.7)$$

- 動きベクトルの向きが類似している場合は、同じ領域として統合する。

$$argMV < THarg \quad (4.8)$$

今回、しきい値 $THarg$ は 60 度とした。

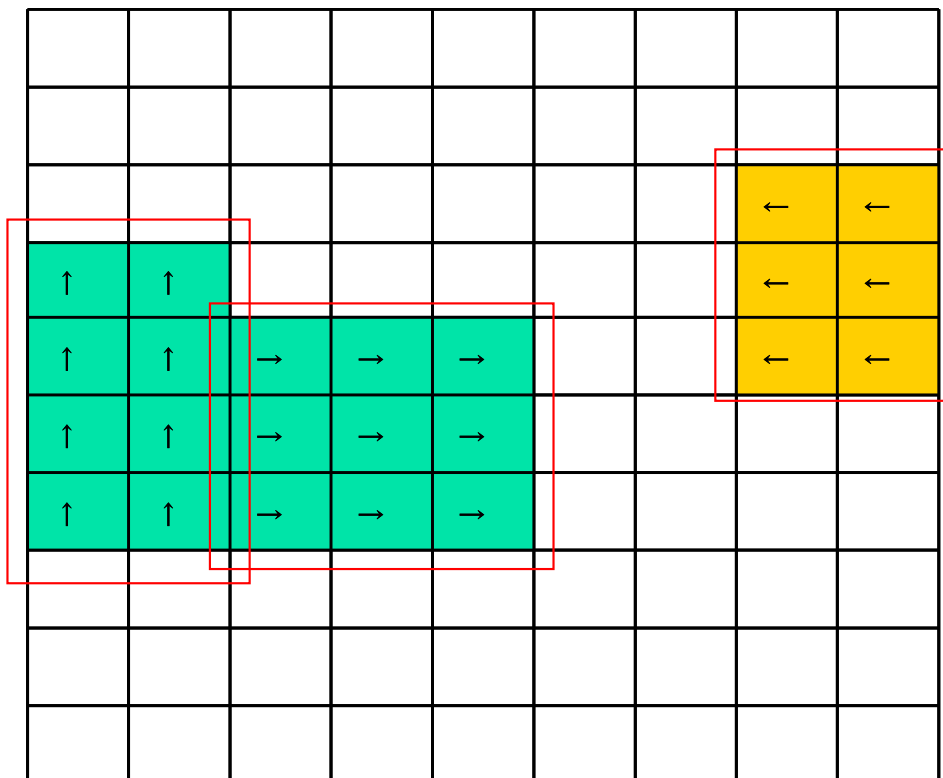


図 4-10: 領域分割

領域分割の例を図 4-10 に示す . 図 4-10 のように , 色と動きの向きが似ている領域を統合する . 領域分割の後 , 領域ごとに輝度の差分値の和を計算する .

$$\sum_{i,j} f(x+i, y+j, t) - f(x+i, y+j, t-1) \tag{4.9}$$

その差分値を 3 フレーム分の和を求める . そして領域内のブロック数 N で割り平均をとり , しきい値 TH により , 動物体と背景に分類する .

$$\frac{1}{N} \sum_{i,j,k} f(x+i, y+j, t) - f(x+i, y+j, t-k) > TH \tag{4.10}$$

$$\frac{1}{N} \sum_{i,j,k} f(x+i, y+j, t) - f(x+i, y+j, t-k) < TH \tag{4.11}$$

つまり式 (9) を動物体 (10) を背景とする .

4.4 むすび

予備実験では , 画像を 8×8 ブロックに分割し , ブロックごとに輝度を平均化する . 輝度を平均した画像を連続 2 フレームでフレーム間差分を取った画像と , 平均化することなく差

分を取った画像を比較した。これにより、領域ごとに輝度を平均し差分を取ることで、背景変動が起きている領域と検出した動物体を分離できることを示した。また、その値を3フレーム分蓄積することで、より分離の精度が向上することを述べた。

この予備実験を基に、提案手法について述べた。提案手法では、まず、領域分割を行う。領域分割は、色が似ている、また動きの向きが似ている領域を同じ領域として分割する。その後、領域ごとに輝度を平均し、隣接フレーム間で差分をとる。そして、その差分値を3フレーム分足すことにより、動物体と背景変動を分離する。

第5章

評価実験

5.1 まえがき

本章では、提案手法の有効性を示すための実験とその結果について述べる。まず、実験条件について述べる。また、今回用いた評価尺度について述べる。評価尺度として、再現率と適合率の定義について述べ、それぞれがどのような意味を持つかについて述べる。そして、前章で述べた提案手法を用いた結果を示す。結果は検出画像、再現率、適合率を示す。従来手法に比べよい精度で検出できることを示し、提案手法が有効であることを示す。

5.2 実験条件

実験条件を表 1 に示す。

表 5-1: 実験条件

| | |
|---------|-------|
| フレームレート | 15fps |
| 解像度 | CIF |
| カメラ | 固定 |

5.3 評価尺度

提案手法の精度は、処理結果とあらかじめ手動で作成した正解マスクを比較することで評価する。評価には再現率と適合率を用いる。本節では、正解マスクの作成条件、また再現率と適合率の定義を行う。

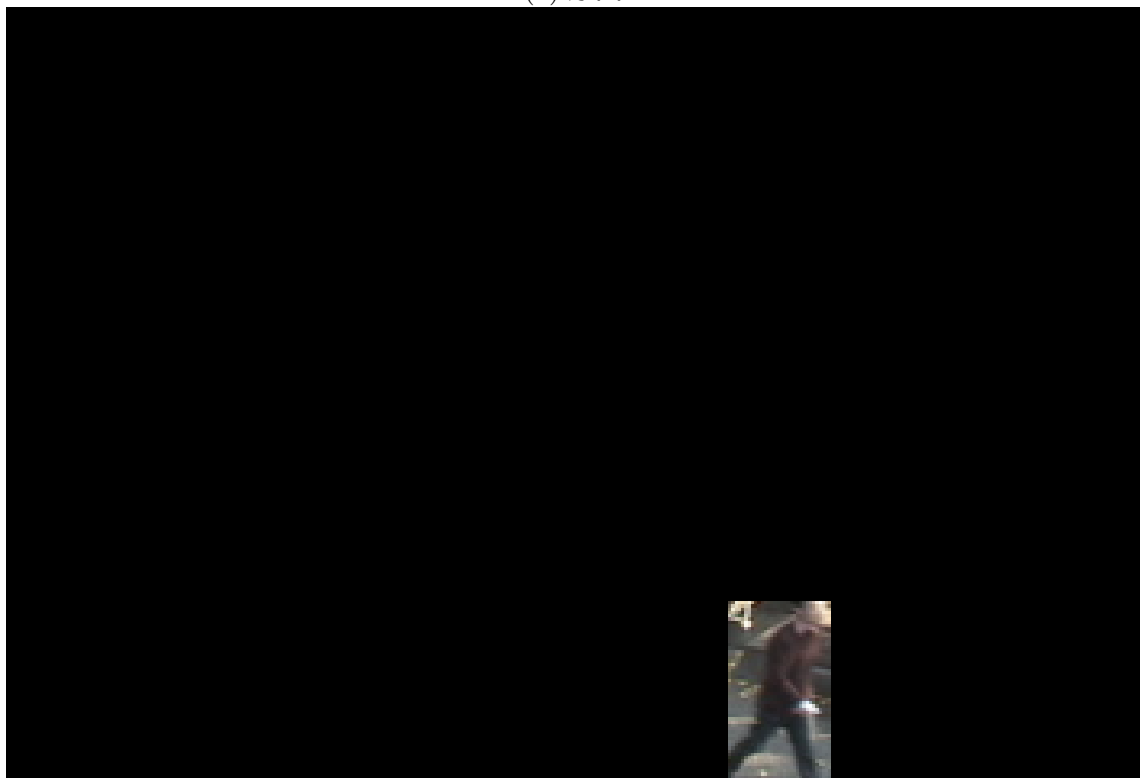
5.3.1 正解マスクの作成

提案手法の検出精度を求めるために、正解マスクを作成する。正解マスクとは、動物体を正確に検出できた結果画像である。正解マスクをあらかじめマニュアルで作成し、処理結果とこの正解マスクを比較することで提案手法の精度を測定する。

正解マスク例を図 5-1 に示す。



(a) 原画



(b) 正解マスク

図 5-1: 正解マスク

5.3.2 再現率，適合率の定義

再現率 (Recall)，適合率 (Precision) は次の式で定義される。

$$Recall = \frac{\text{正しく検出された } 8x8 \text{ ブロックの個数}}{\text{正解マスク中の } 8x8 \text{ ブロックの個数}} \quad (5.1)$$

$$Precision = \frac{\text{正しく検出された } 8x8 \text{ ブロックの個数}}{\text{検出された } 8x8 \text{ ブロックの個数}} \quad (5.2)$$

この再現率，適合率がどのような値を示すかを次にまとめる。

- 検出漏れ，誤検出となるマクロブロックが共に少ない場合

まず，再現率に関しては，分母は「正解マスク中の個数」であるため，一定値をとる．これに対して，分母は「正しく検出された個数」であるが，検出漏れとなるマクロブロックが少ないため，分母の値と分子の値はほぼ等しくなる．従って，再現率は1に近い値をとる．

次に適合率は「検出された個数」の中に「正しく検出された個数」がどれだけ含まれるかを示している．誤検出されるマクロブロックが少ないため，分母は正解マスク中の個数とほぼ等しい値をとる．また，検出漏れとなる個数が少ないため，分子も正解マスク中の個数とほぼ等しい値をとる．従って，適合率も1に近い値をとる．

- 検出もれ，誤検出となるマクロブロックが共に多い場合

まず，再現率に関しては，分母が一定値であるのに対し，分子は検出漏れが多いために小さい値をとる．したがって再現率は低い値をとる．

一方，適合率に関しては，誤検出となる個数が多いため，「検出される個数」が多く，また検出漏れとなる個数がおおいため，「正しく検出された個数」は小さい値をとる．従って適合率も低い値をとる．

- 検出漏れとなる個数が多く，誤検出となる個数が少ない場合

まず，再現率に関しては，分母が一定値であるのに対し，分子は検出漏れが多いために小さい値をとる．従って再現率は低い値となる．

一方，適合率に関しては，誤検出となる個数が少ないため，検出されたものの大半が正しく検出されたものとなる．従って適合率は1に近い値となる．

- 検出漏れとなる個数が少なく，誤検出なる個数が多い場合

まず，再現率に関しては，検出漏れが少ないために分母と分子がほぼ等しい値となるため，再現率は1に近い値となる．

一方，適合率に関しては，誤検出となる個数が多いため，検出された個数の大半が誤って検出されたものとなる．従って，適合率は低い値となる．

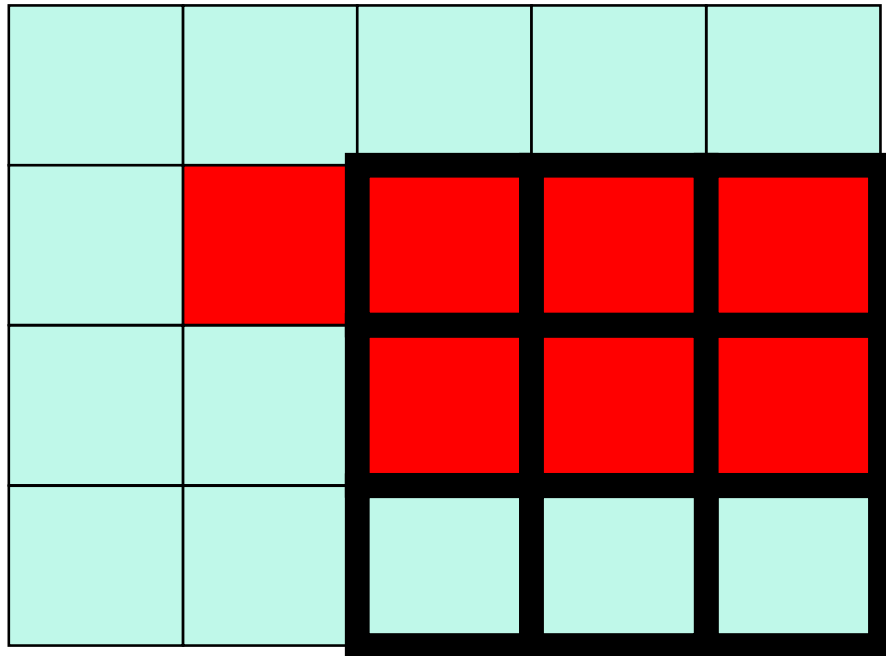


図 5-2: 再現率と適合率

以上より、再現率は検出漏れに関する精度、適合率は誤検出に関する精度を表す指標として用いることができる。

図 5-2 に再現率と適合率の図を示す。図の網がけブロックが正解マスク、太字のブロックが検出されたブロックである。図 5-2 の例だと、

- 再現率 = $\frac{6}{7} = 86\%$
- 適合率 = $\frac{6}{9} = 67\%$

と計算される。

5.4 実験結果

5.4.1 結果画像

実験結果の画像を図 5-3, 5-4, 5-5 に示す。木が揺れている領域に注目すると、その領域での検出はほとんど見られず、木の揺れによる誤検出を減らすことができている。また、検出すべき動物体の領域に注目すると、人がいる領域をきちんと検出している。

5.4.2 再現率，適合率

フレームごとに再現率と適合率を求めたグラフを図5-6に示す．提案手法による過検出，未検出の例を図5-7,5-8に示す．再現率と適合率の式を示す．

$$\text{再現率} = \frac{\text{正しく検出された } 8 \times 8 \text{ の動物体ブロック数}}{\text{正解マスク中の } 8 \times 8 \text{ ブロック数}} \quad (5.3)$$

$$\text{適合率} = \frac{\text{正しく検出された } 8 \times 8 \text{ 動物体ブロック数}}{\text{検出された } 8 \times 8 \text{ ブロック数}} \quad (5.4)$$

再現率についてはフレーム平均で93%であり，検出すべき領域が検出できていることがわかる．一部再現率が低下しているフレームがあるが，これは図5-8のように，頭や手など人の一部が欠けた場合である．そのため，動物体が存在するという情報の漏れは無い．一部が欠けてしまった理由は，領域分割の際にブロックに分割するが，そのブロック内で検出対象となる動物体の割合が少ない場合にそのブロックを背景であると誤検出したためであると考えられる．

適合率はフレーム平均で85%である．再現率に比べ低い，これは図5-7のように動物体の周りの背景領域を検出したためである．動物体の周りの背景領域を検出した理由は，動物体が通ってきた領域が，動物体の領域から背景に変わるときの輝度変化と，新しく背景から動物体が変わるときの輝度変化が同じような変化をするためであると考えられる．

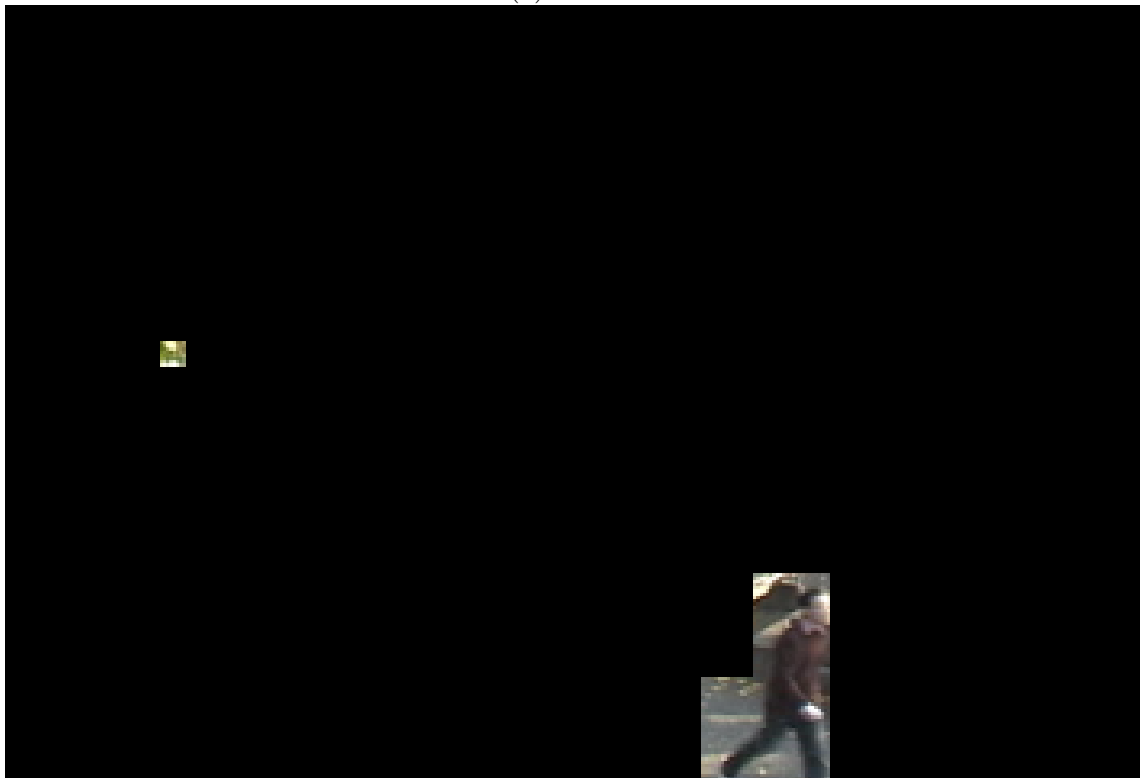
今回の目的は動物体の検出であるため，背景領域を含んでいたり，多少欠けていても問題はないと考えられる．そのため，提案手法は有効であると考えられる．

5.5 むすび

本章では，提案手法の有効性を示すための実験とその結果，考察について述べた．まず，提案手法の評価手法について述べた．評価の際に必要な，検出の正解マスクについて述べ，例を示した．また再現率と適合率について説明し，その意味について述べた．実験結果により，結果の画像により従来手法に比べて提案手法のほうがよい精度で検出できることを示した．そして，再現率と適合率を用いた評価を行った．これにより，提案手法は再現率，適合率ともに高い値を示すことを示した．また，過検出，未検出の例を示した．そして，過検出は動物体に通ってきた領域が，動物体の領域から背景に変わるときの輝度変化と，新しく背景から動物体が変わるときの輝度変化が同じような変化をするためであることを述べた．また，未検出は，領域分割の際にブロックに分割するが，そのブロック内で検出対象となる動物体の割合が少ない場合にそのブロックを背景であると誤検出したためであることを述べた．



(a) 原画像

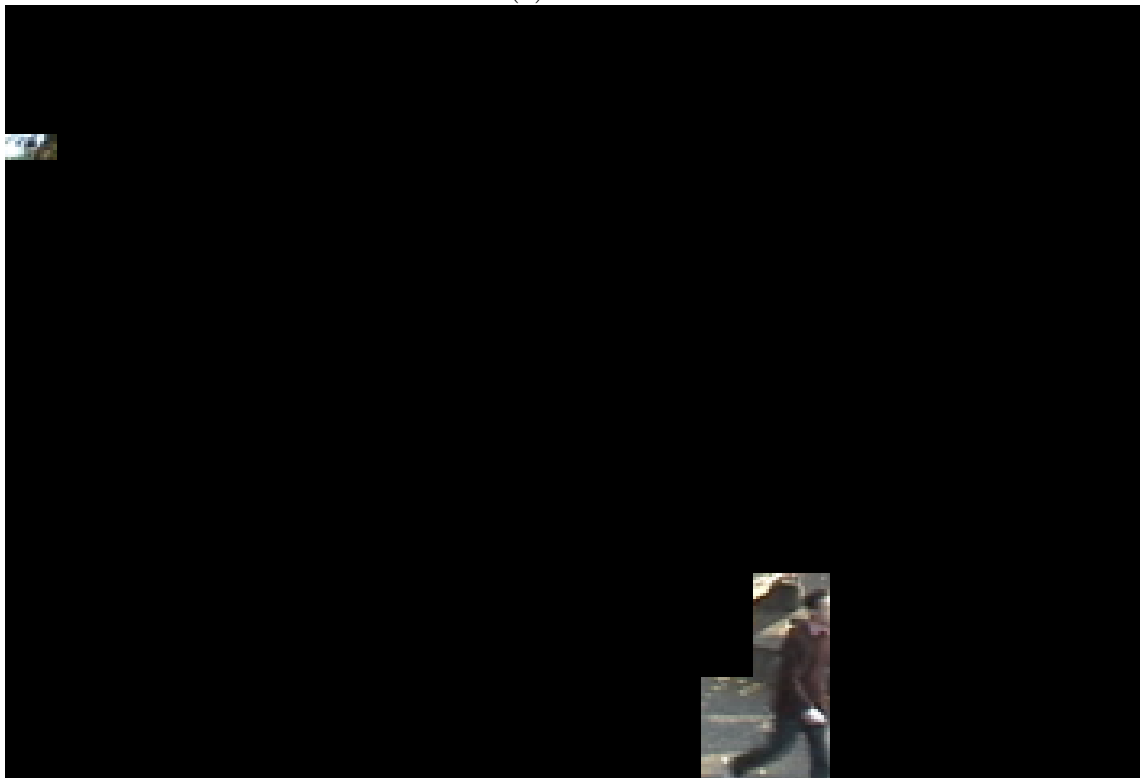


(b) 検出結果

図 5-3: 実験結果 (1 枚目)



(a) 原画像

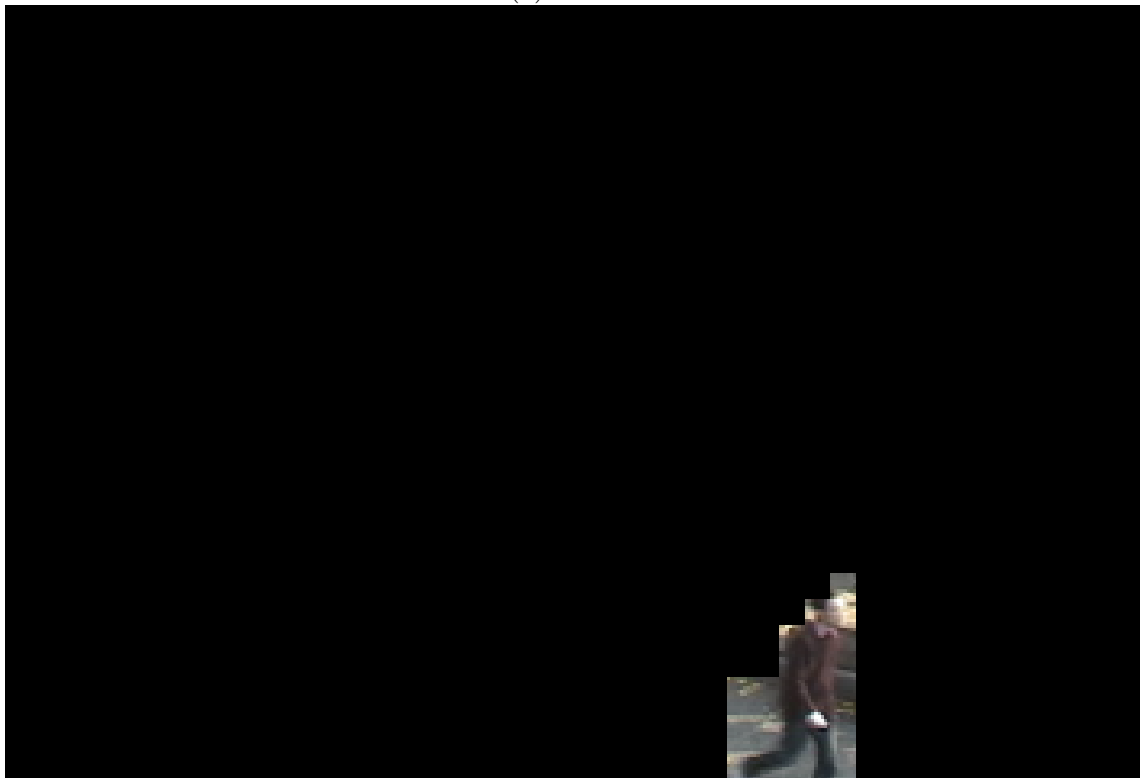


(b) 検出結果

図 5-4: 実験結果 (2 枚目)



(a) 原画像



(b) 検出結果

図 5-5: 実験結果 (3 枚目)

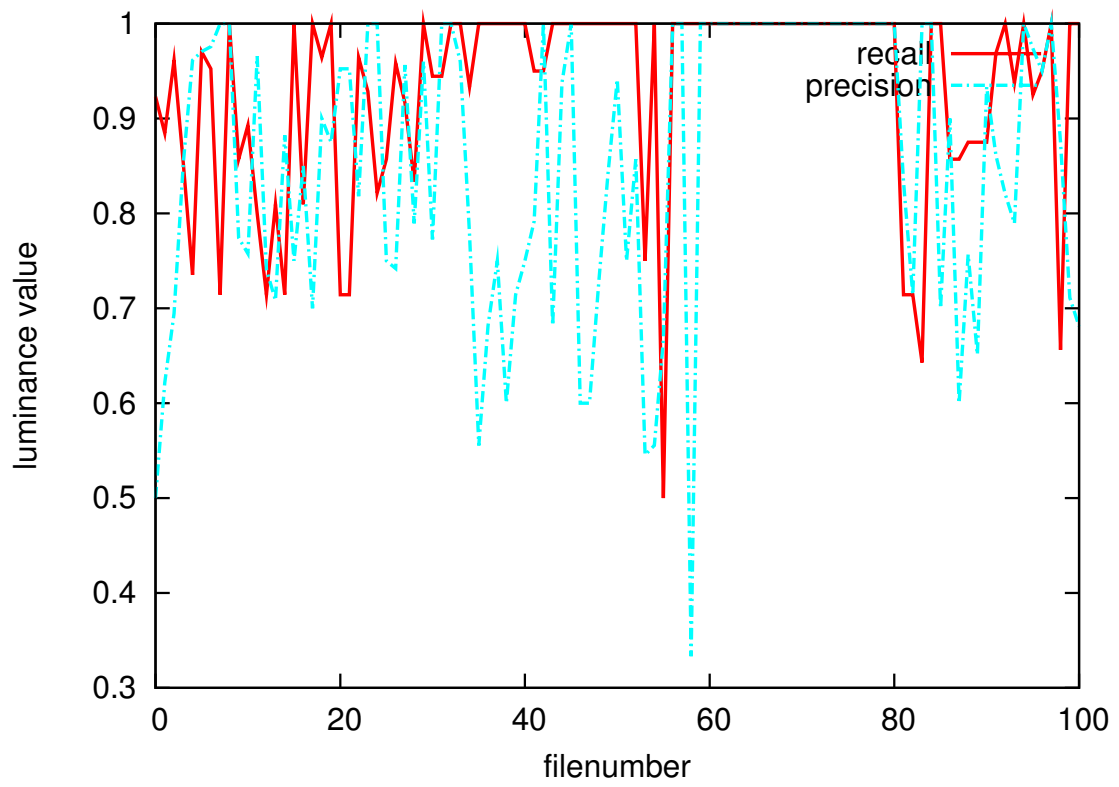


図 5-6: 再現率と適合率

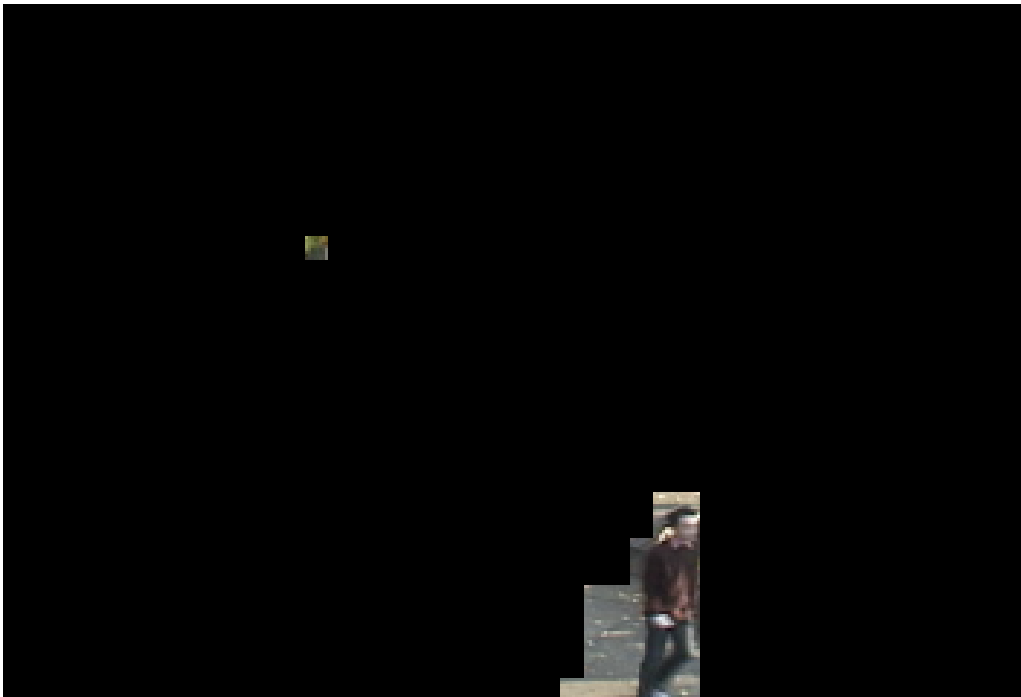


図 5-7: 過検出

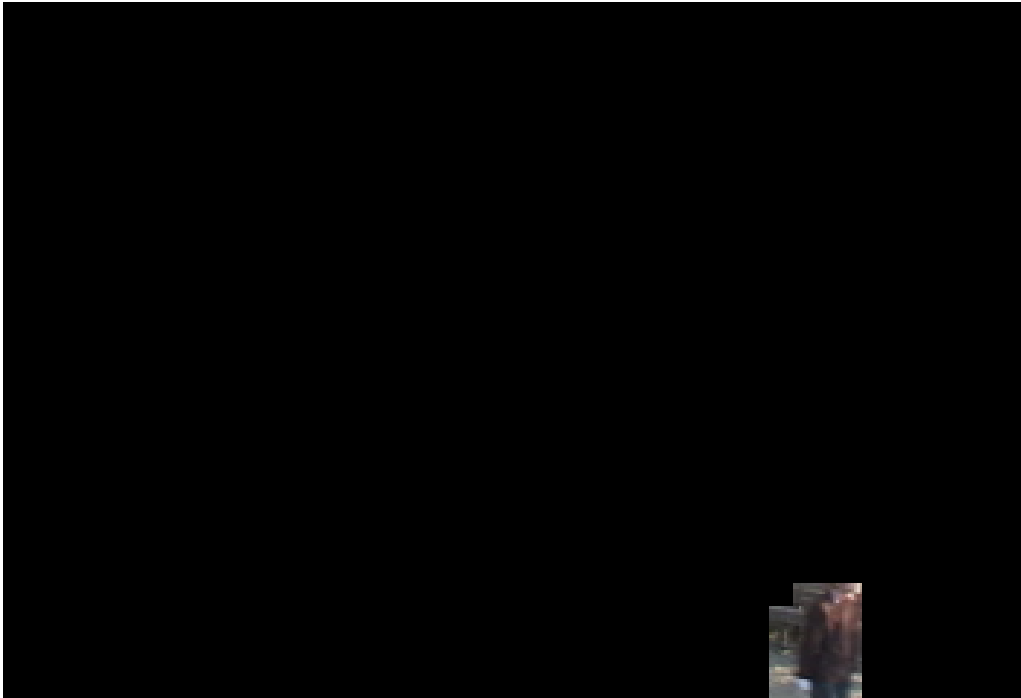


図 5-8: 未検出

第6章

結論

本章では本論文を総括し、今後の検討課題を述べる。

6.1 総括

本研究は、監視カメラシステムを対象とし、その処理の第一段階である物体検出に着目した。その中で、屋外環境で問題となる、木の揺れなど背景変動に注目し、このような複雑な背景における物体検出法について提案した。提案手法は、ブロック内の輝度変化を平均し、フレーム間差分を行う。実験により本手法の有効性を示した。これにより、背景変動を含む映像から、動物体を検出することが可能となる。

第1章では本研究の背景と目的、および本論文の構成について述べた。

第2章では物体検出の基本技術について述べた。まず、物体検出の基本構成を述べた。背景と前景に関する制約条件をうまくモデルに取り入れて背景と物体とを別々のメトリックでモデル化して個別の識別器を構成し、順次候補を絞っていくカスケード検出のフレームワークについて述べた。基本技術として、背景差分法、フレーム間差分法、オプティカルフローを用いた手法について説明し、その問題点について述べた。また、物体検出の利用用途について述べた。利用用途として、工業用途、医療画像、顔検出、監視について述べ、それぞれどのように利用しているかについて述べた。

第3章では本研究で想定している、複雑な背景に対する物体検出法の従来手法について述べた。局所性に注目した手法として、空間平均した動き、時間平均した動きの強さ、各時刻の空間平均した動きの強さに注目した手法について述べた。この手法は代表速度の検出精度が悪い場合、違った箇所を時間軸をさかのぼりながら追跡することになり誤検出の原因となる。また、フレーム間の追跡で得られた代表速度が一定でなくとも、時間平均の特徴量が小さくなり誤検出の原因となることを述べた。また、空間的、時間的な相関を用いた手法として、空間的一様性、時間的一様性に着目した手法について述べた。この手法は影や日照変動など光の影響に関しては誤検出がないが、木の揺れなどの背景の変動には弱く、想定している環境では十分な精度が得られないことを述べた。

第4章では、予備実験と、それを基にして今回提案する手法について述べた。予備実験では、画像を 8×8 ブロックに分割し、ブロックごとに輝度を平均化する。輝度を平均した画

像を連続2フレームでフレーム間差分を取った画像と、平均化することなく差分を取った画像を比較した。これにより、領域ごとに輝度を平均し差分を取ることで、背景変動が起きている領域と検出した動物体を分離できることを示した。また、その値を3フレーム分蓄積することで、より分離の精度が向上することを述べた。この予備実験を基に、提案手法について述べた。提案手法では、まず、領域分割を行う。次に、領域ごとに輝度を平均し、その差分をとる。その差分値を3フレーム分足すことにより、動物体と背景変動を分離する。この処理の流れについて述べた。

第5章では、提案手法の有効性を示すための実験とその結果、考察について述べた。まず、提案手法の評価手法について述べた。評価の際に必要な、検出の正解マスクについて述べ、例を示した。また再現率と適合率について説明し、その意味について述べた。実験結果により、結果の画像により従来手法に比べて提案手法のほうがよい精度で検出できることを示した。次に、再現率と適合率を用いた評価を行った。これにより、再現率、適合率ともに高い値を示すことを示した。また、過検出、未検出の例を示した。過検出は動物体が出てきた領域が、動物体の領域から背景に変わるときの輝度変化と、新しく背景から動物体になるときの輝度変化が同じような変化をするためであることを述べた。また、未検出は、領域分割の際にブロックに分割するが、そのブロック内で検出対象となる動物体の割合が少ない場合にそのブロックを背景であると誤検出したためであることを述べた。

第6章は結論であり、本論文の総括および今後の課題について述べている。

6.2 今後の課題

今後検討すべき課題を以下に示す。

- 解像度と検出結果の関係

本研究の提案手法では、ブロックに分けて輝度の平均を取った。そのため、実際には解像度を落とした画像を用いて比較を行っていることと同じである。そのため、解像度がどの程度のときに背景変動と動物体を分離できるのかという関係性を考える必要があると思われる。

- 多重解像度解析

本研究は、低解像度の画像を解析し、その結果を基に高解像度の映像を解析する多重解像度解析との関係が深いと考えられる。そのため、多重解像度解析との関係を考え、今後発展させていくことが課題であると考えられる。

- トラッキング

本研究は、監視カメラの第一段階の処理として位置づけられている。本手法を基に、追跡技術について検討する必要がある。

謝辞

本研究の機会及び素晴らしい実験環境を与えて下さり、貴重な時間を割いて研究の方向性を御指導頂きました渡辺 裕 教授に心から感謝致します。

本研究を進めるにあたり、日頃から惜しみなく御指導して頂きました早稲田大学国際情報通信研究科富永 英義 教授に深く感謝致します。

貴重な御意見、様々な御提案を頂いたA V Sゼミの皆様に御礼申し上げます。特に、29-7号館106教室で常に研究の進捗を気にして下さった博士課程2年河村 圭 氏に心から感謝致します。渡辺研合宿幹事として川奈合宿の企画や旅行を企画して下さった修士課程1年守屋 宏美女史に感謝いたします。ゼミの幹事として、時間の管理や進行をしてくださった修士課程1年石井 祐一氏に感謝いたします。

渡辺研究室の皆様に、御礼申し上げます。特に、同期として一緒に研究や行事を共にした加藤 幸一君、後藤 崇行君、山本 勇樹君に感謝いたします。学部のところから貴重なご意見をいただいた博士課程2年石川 孝明氏に感謝いたします。

富永研究室の皆様に、御礼申し上げます。M2合宿を計画するなど、幹事としてがんばってくれた山本 高大君、田代 裕和君に感謝いたします。分室システム管理として、分室のネットワークの管理をしてくれた秋間 佑輔君に感謝いたします。その他同期の方々に感謝いたします。

最後に、私をここまで育てて下さった家族に深く感謝します。

平成 19 年 2 月 06 日

参考文献

- (1) Arun Hampapur, Lisa Brown, Jonathan Connell, Ahmet Ekin, Norman Haas, Max Lu, Hans Merkl, Sharath Pankanti, Andrew Senior, Chiao-Fe Shu, and Ying Li Tian, "Smart video Surveillance," IEEE signal processing magazine, pp38-51, Mar. 2005.
- (2) 橋本, "映像監視における CV 実用技術の現状と課題," 情報処理学会 CVIM 研究会, Nov.2005.
- (3) Weiming Hu, Tienju Tan, Liang Wang, and Steve Maybank, "A Survey on Visual Surveillance of Object Motion and Behavior," IEEE transactions, Man and Netics-Partic:Apprications and Reviews, Vol.34, No.3, Aug. 2004.
- (4) 小沢, "サーベイランスの現状と今後," 映像情報メディア学会誌, vol.57, NO.9, pp.1056-1057, Sep.2003.
- (5) 国土交通省, <http://www.mlit.go.jp/road/ITS/j-html/index.html>
- (6) T.Kanade, R.T.Collins, A.J.Lipton, P.Anandan, P.Burt and L.Wixson, "Cooperative Multi-Sensor Video Surveillance," Proceeding of DARPA Image Understanding Workshop 1997, Vol.1, pp3-10, 1997.
- (7) R.Collins, et al, "A system for video surveillance and monitoring: VSAM final report," Technical report CMU-RITR-00-12, Robotics Institute, CMU, May 2000.
- (8) 藤吉弘亘, 榎本暢芳, 長谷川修, 金出武雄, "アクティビティモニタリング-屋外監視映像の要約と WWW 上表示・検索システム-, " 第7回画像センシングシンポジウム June, 2001.
- (9) R.Want, A.Hopper, V.Falcao and J.Gibbons, "The Active Badge Location System," ACM Transaction on Information System, Vol.10, No.1, pp.91-102, 1992.
- (10) 佐藤典幸, 長尾確, "個人用知的移動体における実世界と連動した情報コンテンツの利用," 情報処理学会第67回全国大会論文集, Mar.2005.
- (11) 川村邦明, 清水幸一, 山本克之, "痴呆性徘徊対処システムの開発-移動体通信とGPSによる徘徊者定位の試み," 信学技法MBF, pp94-96, Jul.1994.
- (12) 堀 修, "人物画像処理技術とその応用," 信学技報PRMU 2006-19, pp.107-112, May 2006.
- (13) 岸場秀行, "サーベイランス用顔画像検出技術," 映像情報メディア学会誌, vol.57, NO.9, pp.1056-1057, Sep.2003.

- (14) P.Viola and M.Jones, "Rapid object detection using a boosted cascade of simple features," Proc.IEEE Conf.Computer Vision and Pattern Recognition(CVPR),vol.1,pp.511-518,2001
- (15) C.Huang,H.Ai,Y.Li,S.Lao, "Vector Boosting for Rotation Invariant Multi-view Face Detection," Proc.Intl.Conf.on Computer Vision(ICCV),pp.446-453,2005
- (16) Tao Zhao, and Ram Nevatia, "Tracking Multiple Humans in Complex Situations," IEEE transactions on pattern analysis and machine intelligence, vol.26, No.9, Sep. 2004.
- (17) 鷺見, 関, 波部, "物体検出-背景と検出のモデリング," 情報処理学会 CVIM 研究会, pp.79-98, Sep.2005.
- (18) 栄藤稔, "動画像符号化におけるコンピュータビジョン," 信学論 D-2, Vol.80-D-2, No.2, Feb.1997
- (19) 羽下, 鷺見, 八木, "変化領域内の動きの時空間特徴に着目した屋外情景における歩行者の検出," 信学論 D-II, Vol.J87-D-II, No.5, pp.1104-1111, May. 2004.
- (20) 吉田, "参照画素の状態に基づく背景差分法," 信学論 A, Vol.J88-A, No.11, pp1226-1234, Nov.2005.
- (21) 伊谷, 渡辺, "Motion JPEG 2000 における前後フレームを用いたフリッカー低減手法の検討", FIT2005, J-061, Sep. 2005
- (22) 安居院 猛, 長尾智晴, "C 言語による画像処理入門," 昭晃堂, Apr.2005.
- (23) P.Viola,MJ.Jones,D.Snow, "Detection Pedestrians Using Patterns of Motion and Appearance," Proceedings of the Ninth IEEE International Conference on Computer Vision(ICCV03),2003.
- (24) Julio Cezar et al, "A Background Subtraction Model Adapted to Illumination Changes," IEEE International Conference on Image Processing 2006 (ICIP2006), TP-P1.2, Oct.2006.
- (25) I.Haritaoglu,D.Harwood,and L.S.Davis, "W4:Real-time surveillance of people and their activities," IEEE Transactions on Pattern Analysis and Machine Intelligence,vol.22,no.8,pp809-830,Aug.2000.
- (26) Giger,M.et al, "Guest editorial computer-aided diagnosis in medical imaging," IEEE Transactions Medical Imaging, Vol.20, No,12, pp1205-1208, Dec.2001.
- (27) 和島直哉, 高橋悟, 伊藤誠也, 佐藤雄隆, 金子俊一 "Radial Reach Filter を用いたブロック識別法による移動物体追跡," 画像電子学会誌, vol. 35, no. 4, pp.306-313, Sep.2006 .

- (28) 土田, 川西, 村瀬, 高木, “背景差分法による物体検出を目的とした逐次モンテカルロ法による背景推定,” 信学論 D-II, Vol.J87-D-II, No.5, pp.1062-1070 May. 2004.

図一覧

| | | |
|------|------------------------|----|
| 1-1 | 監視カメラ映像の処理の流れ | 3 |
| 2-1 | 物体検出の基本的フレームワーク | 7 |
| 2-2 | モデル学習を考慮したアーキテクチャ | 8 |
| 2-3 | 実用的な物体検出のためのカスケード構成 | 9 |
| 2-4 | 個別の識別器の出力を並列統合する枠組み | 9 |
| 2-5 | 背景差分法のしくみ | 10 |
| 2-6 | 2枚のフレーム間差分画像を用いた物体の抽出例 | 11 |
| 2-7 | 3枚のフレーム間差分画像を用いた物体の抽出例 | 12 |
| 2-8 | 顔認証システム | 15 |
| 2-9 | フリッカー雑音の検出，補正の様子 | 19 |
| 2-10 | 輝度変化の様子 | 20 |
| 2-11 | 残像 | 20 |
| 2-12 | 動領域判定 | 21 |
| 2-13 | フィルタ処理後 | 22 |
| 3-1 | 従来手法 | 24 |
| 3-2 | 従来手法の実験結果 (1 フレーム目) | 27 |
| 3-3 | 従来手法の実験結果 (2 フレーム目) | 28 |
| 3-4 | 従来手法の実験結果 (3 フレーム目) | 29 |
| 3-5 | 従来手法 2 の実験結果 (1 枚目) | 33 |
| 3-6 | 従来手法 2 の実験結果 (2 枚目) | 34 |
| 3-7 | 従来手法 2 の実験結果 (3 枚目) | 35 |
| 4-1 | アプローチ | 38 |
| 4-2 | 原画像 | 39 |
| 4-3 | フレーム間差分 | 40 |
| 4-4 | 平均画像 | 41 |
| 4-5 | 平均画像のフレーム間差分 | 42 |
| 4-6 | フレーム間差分の輝度変化 | 43 |
| 4-7 | 平均画像の輝度変化 | 43 |
| 4-8 | 平均画像の輝度変化 (3 フレーム蓄積) | 44 |
| 4-9 | 提案手法の流れ | 45 |

| | |
|---------------------------|----|
| 4-10 領域分割 | 46 |
| 5-1 正解マスク | 49 |
| 5-2 再現率と適合率 | 51 |
| 5-3 実験結果 (1 枚目) | 53 |
| 5-4 実験結果 (2 枚目) | 54 |
| 5-5 実験結果 (3 枚目) | 55 |
| 5-6 再現率と適合率 | 56 |
| 5-7 過検出 | 56 |
| 5-8 未検出 | 57 |

表一覧

| | |
|---------------------------|----|
| 1-1 物体検出に影響する要因 | 5 |
| 5-1 実験条件 | 48 |

研究業績

| | 題 名 | 発表年月 | 発表掲載誌 | 連名者 |
|-----|---|-------------|------------------|---------------------------------|
| 国内 | | | | |
| (1) | MotionJPEG2000 におけるフリッカー低減手法の検討 | 2005 年 3 月 | 信学春季全大 | 伊谷 裕介 石川 孝明 渡辺 裕 富永 英義 |
| (2) | Motion JPEG 2000 における静動領域判定を用いたフリッカー低減手法の検討 | 2005 年 7 月 | 情処研報 2005-AVM-49 | 伊谷 裕介 渡辺 裕 |
| (3) | Motion JPEG 2000 における前後フレームを用いたフリッカー低減手法の検討 | 2005 年 9 月 | FIT2005 | 伊谷 裕介 渡辺 裕 |
| (4) | 複雑な環境における物体検出法の一検討 | 2005 年 11 月 | IMPS2005 | 伊谷 裕介 渡辺 裕 |
| (5) | 複雑な背景における物体検出法の検討 | 2006 年 12 月 | 信学会 P R M U 研究会 | 伊谷 裕介 渡辺 裕 |
| (6) | 複雑背景における物体検出法 | 2007 年 3 月 | 信学春季全大 | 伊谷 裕介 渡辺 裕 |
| (7) | 照明変化のある環境下での移動物体検出の検討 | 2007 年 3 月 | 信学春季全大 | 谷 誠一 伊谷 裕介 渡辺 裕 富永英義 |