

コンピュータ工学特別研究報告書

題目 背景差分による移動物体の検出に関する
基礎的な検討

学生証番号 946380

氏名 岡部 良亮

提出日 平成25年2月5日

指導教員 蚊野 浩

京都産業大学
コンピュータ工学部

要約

本論文は、バスケットボールの試合など、チームスポーツにおける選手の動きを画像処理で検出する技術について基礎的な検討を行った。固定されたカメラで撮影した映像から移動物体を抽出するために背景差分が使われる。本研究では、背景差分の手法として、最も基本的な方法であるフレーム差分法、一つの平均背景画像とその統計量を背景モデルとして、背景モデルと入力画像の差分から前景領域を表示する平均背景法および、複数の背景モデルを学習する、より高度なコードブック法を実装し、それらの性能を比較した。その結果、フレーム差分法および平均背景法に比べ、コードブック法を用いることで明確な前景領域が抽出できた。また、背景の変化が激しい画像に適用した場合にも正しく背景モデルを学習しており、抽出すべき対象を前景領域として抽出することができた。

目次

第1章 序論	・・・	4
第2章 固定カメラの映像を用いた移動物体の検出		
2.1 映像からの移動物体の抽出	・・・	5
2.2 背景差分による移動物体の抽出	・・・	6
2.2.1 フレーム差分法	・・・	7
2.2.2 2つの平均背景法	・・・	11
2.2.3 フレーム差分法と平均背景法の課題	・・・	16
2.3 コードブック法の概要	・・・	17
第3章 コードブック法の詳細な説明		
3.1 背景モデルデータ構造	・・・	18
3.2 アルゴリズム	・・・	20
第4章 実験と考察		
4.1 実験に用いた動画像	・・・	23
4.2 フレーム差分法による移動物体の検出結果	・・・	30
4.3 平均背景法による移動物体の検出結果	・・・	31
4.4 コードブック法による移動物体の検出結果	・・・	32
第5章 結論		
5.1 成果	・・・	33
5.2 課題	・・・	33

参考文献

謝辞

付録

1. 序論

画像センサとコンピュータ技術の進歩によりカメラの機能・性能が格段に上がった。画質が向上し、顔認識やシーン判別等に利用する映像解析技術の進歩がめざましい。例えば、監視カメラは遠くの景色を鮮明な解像度で撮影することができるようになった。その映像を解析することで、不審な人物を容易に発見する事が可能になり、安心・安全な社会に寄与している。

カメラで撮影した画像・映像を解析する技術はスポーツ映像にも応用されている。最近のビデオ映像レコーダーには野球やサッカーの試合における重要なシーンを抽出する機能が実装されている。また、スポーツの試合で実際に使用されている技術もある。ウインブルドンでは、テニスボールのイン／アウトに微妙な判断が必要な場面で映像処理技術が利用される。「The Hawk-Eye Officiating System」はボールの位置や軌道を分析し、3次元グラフィックスを用いて軌道を表示することで、ライン判定の補助を行う。

コンピュータを使った映像解析技術で、チームスポーツに参加する選手の動きを追跡することも実用化されている。TRACAB 社のトラッキング・システム「TRACAB image tracking system」は、サッカーの試合中の選手やボールの動きを定量化するシステムである。試合中の各選手の総移動距離やプレー・エリア等をデータ化することができる。

TRACAB 社のトラッキング・システムが実現しているサッカー選手の動きを抽出するような技術を、一般に、映像からの移動物体の追跡技術とよぶ。本研究では、固定カメラから撮影したサッカーやラグビー等のスポーツシーンの映像から、移動物体である選手を検出することを目的とした。これが実現できれば、選手個人の行動履歴や運動量をデータとして抽出して詳細な選手のデータやプレー中の行動を見返すことができる。今回はスポーツ映像を解析する前処理として、動画像から背景差分を用いて移動物体である動き領域(前景領域)を抽出する3種類の手法を開発し、比較検討した。

2. 固定カメラの映像を用いた移動物体の検出

2. 1 映像からの移動物体の検出

人間は目に映る画像から、その中に存在する個々の物体を、簡単に区別することができる。例えば、サッカーの試合を観戦したとき、メッシの動きがいくらすばやくとも、その動きを見失うことはない。一方、その機能をカメラとそれに接続したコンピュータによって完全に実現することはいまだに達成されていない。

カメラで、サッカーの試合のようなチームスポーツの映像を撮影し、その映像から個々の選手の動きを追跡することは困難な問題である。しかし、序論で述べたように、これにつながる技術は部分的に実用化されている。そこで、本論文では、カメラで撮影したチームスポーツの映像から、個々の選手の動きを追跡するという機能を、比較的、単純な状況において実現する技術を研究することとする。

チームスポーツを撮影した映像から、選手の動きを検出・追跡する手法の概要を説明する。最も一般的な状況は、手持ち撮影したビデオカメラの映像から、選手を検出し追跡するケースである。この場合、全体が変化する映像から、背景とは異なる動きをする選手を抽出する必要がある。これを実現するには、1枚の画像から任意姿勢の人物を検出するような、高度な画像認識技術が必要になる。今回の研究では、このような一般的な状況は考慮しない。

固定されたビデオカメラで撮影した映像から、移動物体を検出・追跡する手法には、背景差分を用いる方法と、動きベクトル（オプティカルフロー）を用いる方法がある。背景差分を用いる方法は、比較的単純なアルゴリズムであっても安定して移動物体を抽出することができる。一方、動きベクトルを用いる方法はアルゴリズムが複雑である。まず、連続したフレーム間で微小な小領域ごとに動きベクトルを抽出し、多数の動きベクトルをクラスタリングすることで、移動物体を検出する。しかし、動きベクトルに誤差が発生しやすいので、選手に正確に対応するようにクラスタリングすることは難しい。

2. 2 背景差分による移動物体の検出

本研究では移動物体の検出に背景差分を使う。この手法は、監視カメラのような固定されたカメラで移動物体を検出する場合に有効な方法である。その理由は、固定カメラで撮影された映像は、移動物体以外の背景領域に変化が少なく、移動物体領域を前景として抽出することが容易だからである。

背景差分法では背景画像を用いる。背景画像を生成するための手法がさまざまに提案されている。それらの手法で生成されるものは、1枚の画像である場合もあるが、背景の様子をモデル化したデータ構造である場合もある。これらを総称して背景モデルとよぶ。

背景モデルを生成する代表的な手法は、移動物体がない状態の画像を固定カメラで撮影し、その画像を背景画像として採用することである。高度な方法では、背景領域に生じる信号変化を確率モデルで表現するものがある。背景モデルを生成する手法にはさまざまなものがあるが、このプロセスを背景モデルの学習とよぶ。背景モデルが構築されると、背景モデルを現在の入力画像と比較し、既知の背景部分を取り除く。除去の後に残った領域を、前景領域とよぶ。

背景差分を用いる場合、入力映像において、検出したい移動物体が存在する領域以外は完全に静止していることが理想である。しかし、通常、屋外であれば木々の揺れ、屋内であれば照明の変化による影響などで、背景部分がある程度変化することは避けられない。背景モデルの生成において、次のような問題を考慮する必要がある。

- 移動物体の背景化、逆に背景物体の移動：これらによって背景モデルが時間的に変化する場合がある。
- ゆるやかな明るさ変化：ゆるやかな照明の変化によって背景モデルが変化する場合がある。
- 急激な明るさ変化：急な照明の変化によって背景モデルが変化する場合がある。
- 木々の揺れのような定常的変化：固定的な背景部分と定常的な変化が発

生している部分では、背景モデルのパラメータが異なる。

- 背景と前景の画素値の類似：前景物体の画素値が背景モデルの画素値に類似していると、画素値レベルの処理では、区別ができない。
- 背景モデルの作成：前景物体が存在しない映像があれば、背景モデルを作成することが容易である。
- 前景に生じる穴：均一色の物体が移動している時、フレーム差分法では物体内側の全画素を検出することが難しい。フレーム差分に限らず、移動物体の全体を抜けなく前景領域として抽出することは難しい。
- 影の影響：前景物体の影が生じる領域は、物体が存在しない領域であるが、前景領域として抽出されやすい。

これらの問題は背景差分において、常に課題となっている。そのため、これらのシーンに適応した動作を行う背景差分法が必要である。

次からの節で、最も基本的な背景差分法であるフレーム差分法と平均背景法について説明する。そして、今回の研究素材であるコードブック法を説明する。なお、本研究で用いる動画像はカラー画像であるが、背景差分の処理は濃淡画像に対して行った。従って、特に明示しない限り、カラー画像から濃淡画像への変換を伴う。本論文では、説明を簡単にするために、この色変換の処理の記述を省略している。

2. 2. 1 フレーム差分法

移動物体がない理想的な背景画像を得ることは難しい。フレーム差分は、特別な背景モデルを生成しない背景差分法である。ここでは、時間的に連続する画像（フレーム）の差分画像から、移動物体を検出する。この方法で安定して前景領域を抽出できるのは、背景に変化が無く、移動物体の濃淡や色彩が背景領域と明確に異なり、しかも、移動物体が高速に動くような場合である。

図 1 に処理の流れを示す。最初に各変数を初期化し、第一フレームを `frameTime1` に入力する。次いで、動画像の各フレームを順次処理するループに入る。まず、`frameTime2` に `frameTime1` の画像をコピーする。次のフレームを `frameTime1` に入力する。`Foreground` に `frameTime1` と `frameTime2` の差分画像を代入する。`Foreground` をしきい値 `th` で 2 値化する

ることで、前景領域を抽出する。

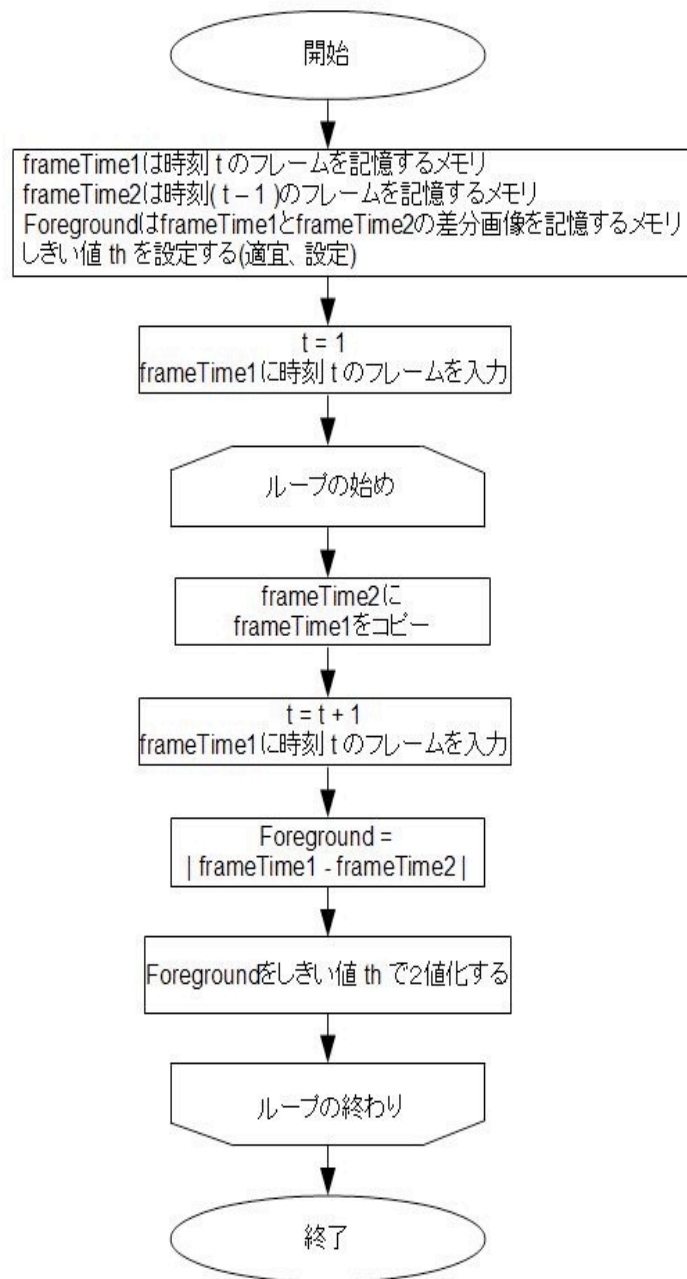


図 1 : フレーム差分法のフローチャート

入力画素に生じたノイズの影響で、2値化した差分画像に小さな粒状のノイズが発生する。これを除去するために、2値画像処理の一つである膨張収縮処理を適用する。膨張収縮処理は、膨張処理とよばれる処理と収縮処理とよばれる処理を繰り返すことで、2値画像に含まれる粒状のノイズを除去する処理である。

収縮処理は、注目する前景領域の画素に背景領域の画素が接しているとき、注目画素を背景に変換して1画素分縮める処理である。図2に収縮処理の例を示す。図2を見ると、背景領域と8近傍で接している前景領域が、背景領域に変換されている。木の葉の揺れや、細かなノイズを前景として抽出したとき、この方法でノイズを除去することが可能である。

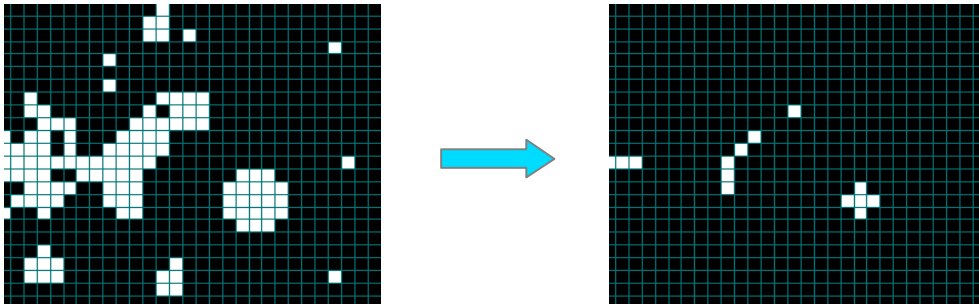


図2. 収縮処理による前景領域の除去
(黒い部分が背景領域、白い部分が前景領域)

膨張処理は、注目する前景領域の画素に背景領域の画素が接しているとき、接している背景領域を注目画素の値に変換して前景領域を1画素分膨らませる処理である。図3に膨張処理の一例を示す。図3を見ると、前景領域に接している背景領域の8近傍を前景領域に変換して前景を補足している。抽出したい移動物体が、小さな前景領域として抽出された場合、視認しやすくなる。

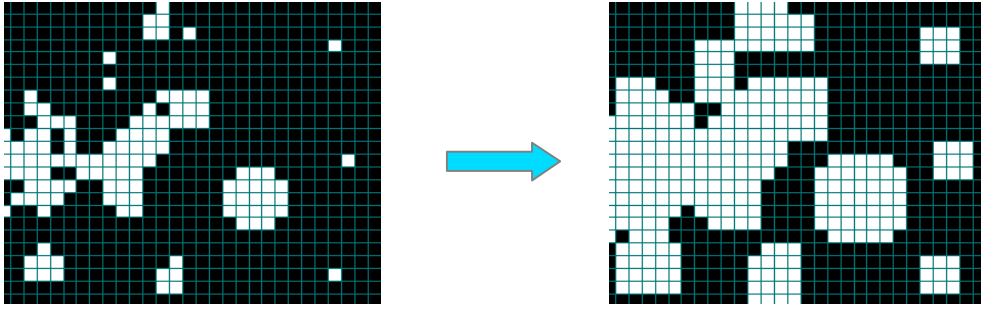


図3. 膨張処理による前景領域の補足

膨張収縮処理は収縮処理と膨張処理を繰り返すことである。これによって、元画像にあったノイズの除去と、抽出したい前景領域の画素の補足を行う。元画像のノイズが除去され、抽出する前景領域が視認しやすくなる。前景領域に属する画素が集合している部分以外は、収縮によって除去されるため、元画像より多少荒い前景領域になる。

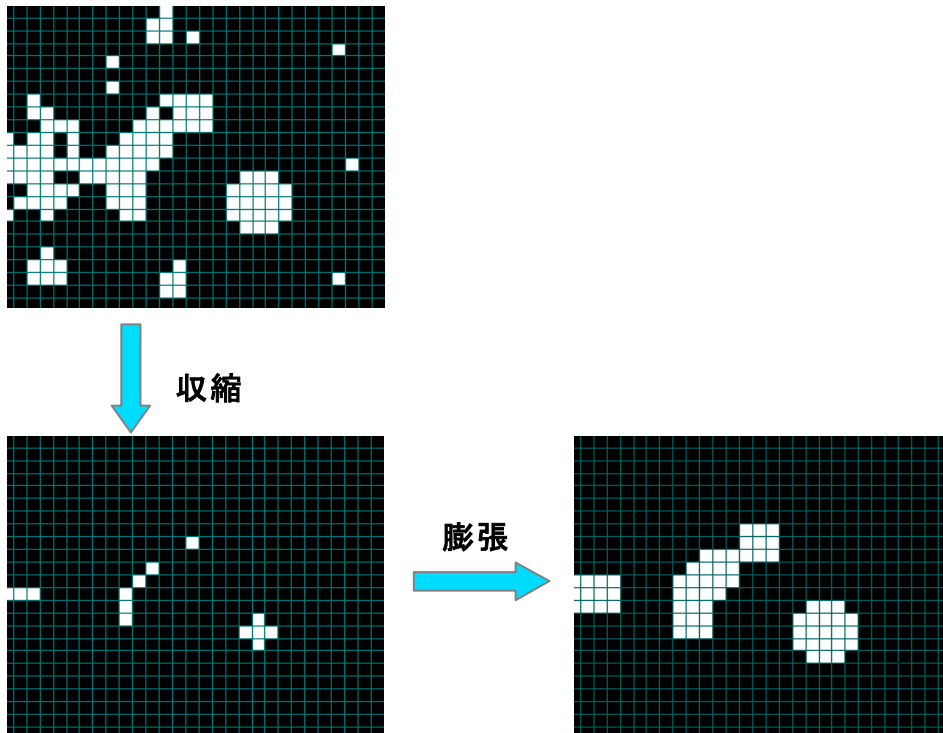


図4. 膨張収縮処理による前景領域の抽出

2. 2. 2 2つの平均背景法

観測される対象シーンの中に樹木があり、それが風で揺れていることがある。この場合、背景画像の画素値に時間的な変動がある。そのため、ある瞬間に撮影したシーンを背景画像として取り込み、それを用いた背景差分法により移動物体検出を行うと、樹木の小枝が移動物体として検出されてしまう。このような、画素値の定常的な変動を考慮して移動物体を検出する必要がある。

平均背景法は、移動物体が存在しない複数枚の背景画像の平均を背景モデルとし、これを用いて前景画像を抽出する方法である。このとき、前景領域を抽出するために、しきい値が必要になる。このしきい値の設定方法にいくつかの考え方がある。

最も簡単なものは、画像全体に一つのしきい値を設定する方法である。しかし、この方法では、樹木の揺らぎを検出しないように大きなしきい値を設定すると、本来抽出すべき前景領域の抽出に失敗する。逆に、小さなしきい値を設定すると、樹木の揺らぎを前景領域として抽出する。これを解決するために、画素ごとに適したしきい値を設定する方法がある。樹木の揺らぎが観察される領域では大きなしきい値を設定し、そうでない領域には小さなしきい値を設定する。画素値の揺らぎが正規分布に従うならば、しきい値はその分散に応じて決めることが望ましい。

画素ごとに分散を計算するには、背景のみの映像を2回読み込んで処理しなければならない。すなわち、1回目は平均を計算するために、2回目に分散を計算するために。このように背景モデルを生成するために、映像を2回読み込むことは、場合によっては実行が難しい。映像を1回読み込むだけで、背景画像の画素値のばらつきを推定する方法として、背景画像だけでフレーム差分を行い、差分画像の画素値のばらつきを用いる方法がある。以下に、この2つの平均背景法を説明する。

まず、第一の平均背景法（差分しきい値平均背景法とよぶことにする）について説明する。

背景モデルを学習するために、あらかじめ、移動物体が存在しない映像を準備する。この映像に含まれる全フレームの平均を計算し、平均画像とする。それとともに、フレーム差分画像の平均を計算し、差分平均画像とする。差分平均画像は、定常的な変化が生じている画素の値が大きく、変化が生じていない画素の値が小さい。

平均画像と差分平均画像を用いて、次の式によって高しきい値画像、低しきい値画像を計算する。

$$\text{高しきい値画像} = \text{平均画像} + k \cdot \text{差分平均画像}$$

$$\text{低しきい値画像} = \text{平均画像} - k \cdot \text{差分平均画像}$$

ここでkは正の定数である。入力画像を高しきい値画像、低しきい値画像と比較し、各画素の値が、2つのしきい値の間であれば背景、なければ前景と判断する。

図5のフローチャートは、次のようになる。

- 背景モデルを学習するため、移動物体が存在しない映像を準備する
- その映像から平均画像と差分平均画像を計算する
- 高しきい値画像、低しきい値画像を計算する
- 移動物体を抽出するループの開始、 $t=1$
- `frameTime` に時刻 t のフレームを入力
- `frameTime` と高しきい値画像と低しきい値画像を比較することで、前景領域を抽出
- $t=t+1$ 、□に戻る

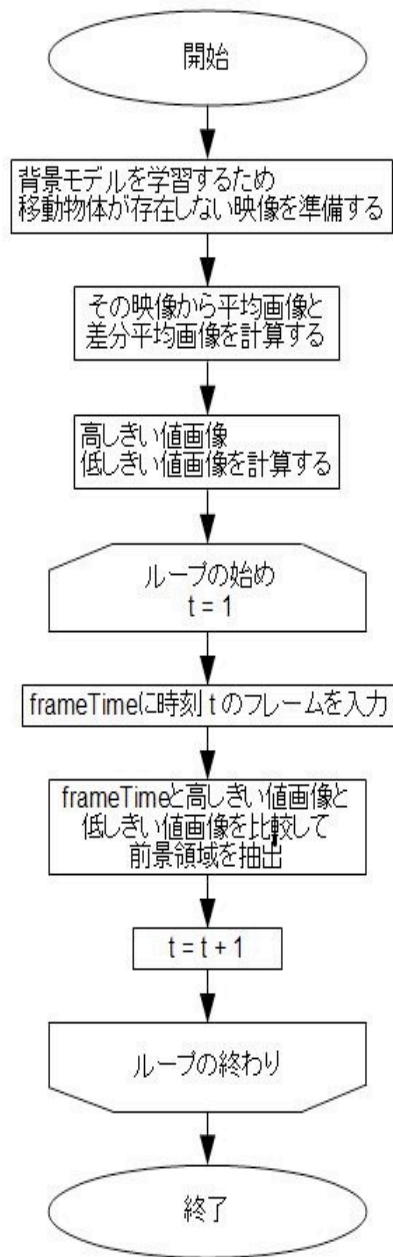


図 5. 差分しきい値平均背景法のフローチャート

次に、第二の平均背景法（分散しきい値平均背景法とよぶことにする）について説明する。この方法では、しきい値は分散の平方根（ σ 、標準偏差）の定数倍に設定される。画素値の分布が正規分布にしたがうならば、図6に示すように、しきい値は 3σ 程度に設定するのが妥当である。

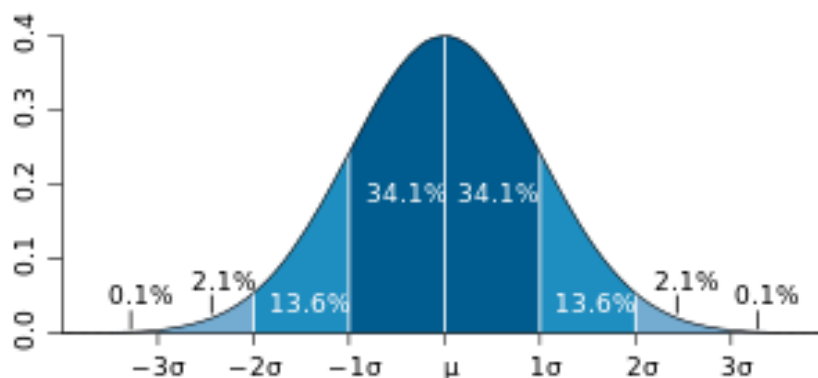


図6. 正規分布図

図7のフローチャートは、図5とほぼ同じである。2つの方法の違いは、高しきい値と低しきい値に設定する値の計算方法である。

$$\text{高しきい値画像} = \text{平均画像} + k \cdot \text{標準偏差画像}$$

$$\text{低しきい値画像} = \text{平均画像} - k \cdot \text{標準偏差画像}$$

- 移動物体が存在しない映像を準備する
- その映像から平均画像と分散画像を計算する
- 高しきい値画像、低しきい値画像を計算する
- 移動物体を抽出するループの開始、 $t=1$
- `frameTime` に時刻 t のフレームを入力
- `frameTime` と高しきい値画像と低しきい値画像を比較することで、前景領域を抽出
- $t=t+1$ 、□に戻る

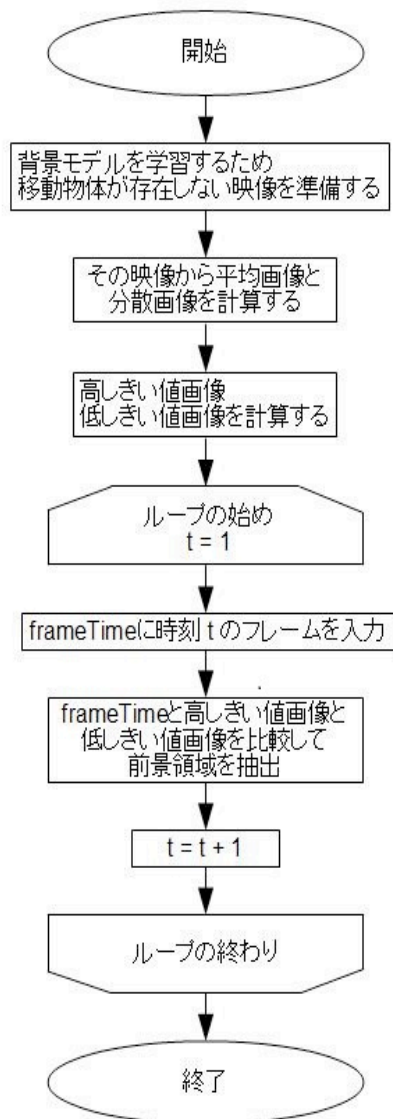


図7. 分散しきい値平均背景法のフローチャート

2. 2. 3 フレーム差分法と平均背景法の課題

フレーム差分法には次のような課題がある。

- 連続するフレーム間で移動物体の動きが小さいとき、物体の輪郭だけが抽出され、内部が背景になる傾向がある。
- 木の葉のゆれのような定常的な動きを前景として抽出してしまう。

後で述べるように、手がゆっくりと通り過ぎるサンプル動画がある。本来、手の全体を前景領域として抽出すべきだが、手の中が背景として抽出されてしまう。このように均一の色領域が、連続するフレーム間で続いた場合、色領域の全体を抽出することが難しい。また、背景に変化のある画像に適用した場合、背景領域の動きも、前景領域として抽出する。

2つの平均背景法には次のような課題がある。

- 日光によってできる影のように時間とともに背景が少しずつ変化すると、背景モデルを更新しない限り、そこを抽出してしまう。
- 背景領域に定常的な信号変化があるとき、そこに移動物体が存在すると検出がむずかしい。

日光によってできる建物の影の境界部分では、常に画素値が変動しているので、抽出された前景画像にノイズが生じる。背景の変化が激しい動画に適用した場合、定常的に動く背景が常に前景として抽出されるので、前景領域抽出の際に誤検出が生じる。

2. 3 コードブック法の概要

フレーム差分法では、移動物体の輪郭だけを抽出してしまい、前景内部が背景になってしまう傾向がある。また、背景の定常的な動きを前景として抽出してしまう。平均背景法は、背景モデルとして画素ごとにしきい値を設定し、画素ごとに前景を判定する。しかし、背景が常に変化する画素に注目すると常に画素値が変動しており、閾値の範囲外である画素値が多く観測される。(ここで説明した平均背景法を使っても、背景の範囲外になる画素値が多く現れる。ある程度の画素値の変動には使えるが、あまりに大きく変動する場合には、検出を間違ふ。)

この問題を解決する方法として、各画素に背景モデルを複数作成して、常に変動する画素値に対応する手法が必要である。そのような手法として、コードブック法を実装して、フレーム差分法と平均背景法の問題が解決できるか検証する。

3. コードブック法の詳細な説明

3. 1 背景モデルデータ構造

コードブック法は K.Kim らによって研究された方法である[1]。OpenCV に関する書籍[3]において、この方法がフレーム差分法と平均背景法よりも優れた方法として紹介されている。

コードブック法は、一つの画素に、複数の背景モデルを保持する手法である。複数の背景モデルは個々をエン트리とよび、エントリ内に閾値を持ったボックスが存在すると考える。

このボックスは、ボックス自体の大きさの値と、学習用閾値を持つ。ボックス自体の大きさは、前景領域の判別をする際の閾値となる。学習用閾値は、観測された画素値がその範囲内のとき、ボックス自体の大きさを成長させるための閾値である。観測された画素値がボックスの学習用閾値内のとき、ボックス自体の大きさの値と、学習用閾値の両方を成長させる。背景モデルが1つの場合、ほんの少し画素値が変動しても、閾値を超えた時点で前景として抽出する。しかし、コードブック法の場合、画素値が少し変化する程度なら、ボックス自身が閾値を成長させることで、抽出すべきでない前景抽出を抑えることができる。

また学習用閾値外の画素値が観測された場合、新しいエントリを作成する。多くのエントリを作成することで、定常的な動きをする画素値に対応でき、木の葉の揺れのような画素も背景として見なすことができる。

エントリを多く作成しすぎると処理動作が重くなり、前景を判別する閾値の判定が多く、抽出すべき前景までも背景と見なしてしまう。これを回避するため、エントリが作成されて以降、更新がほとんど無い古くなったエントリを除去する。一度大きな画素値の変化が観測されて、エントリが作成されても、ほとんど観測されない画素値の閾値を持つエントリならば不要である。

このような、画素値の範囲を指定したボックスの組をコードブックとよぶ。図8にボックス形成のイメージを紹介する。

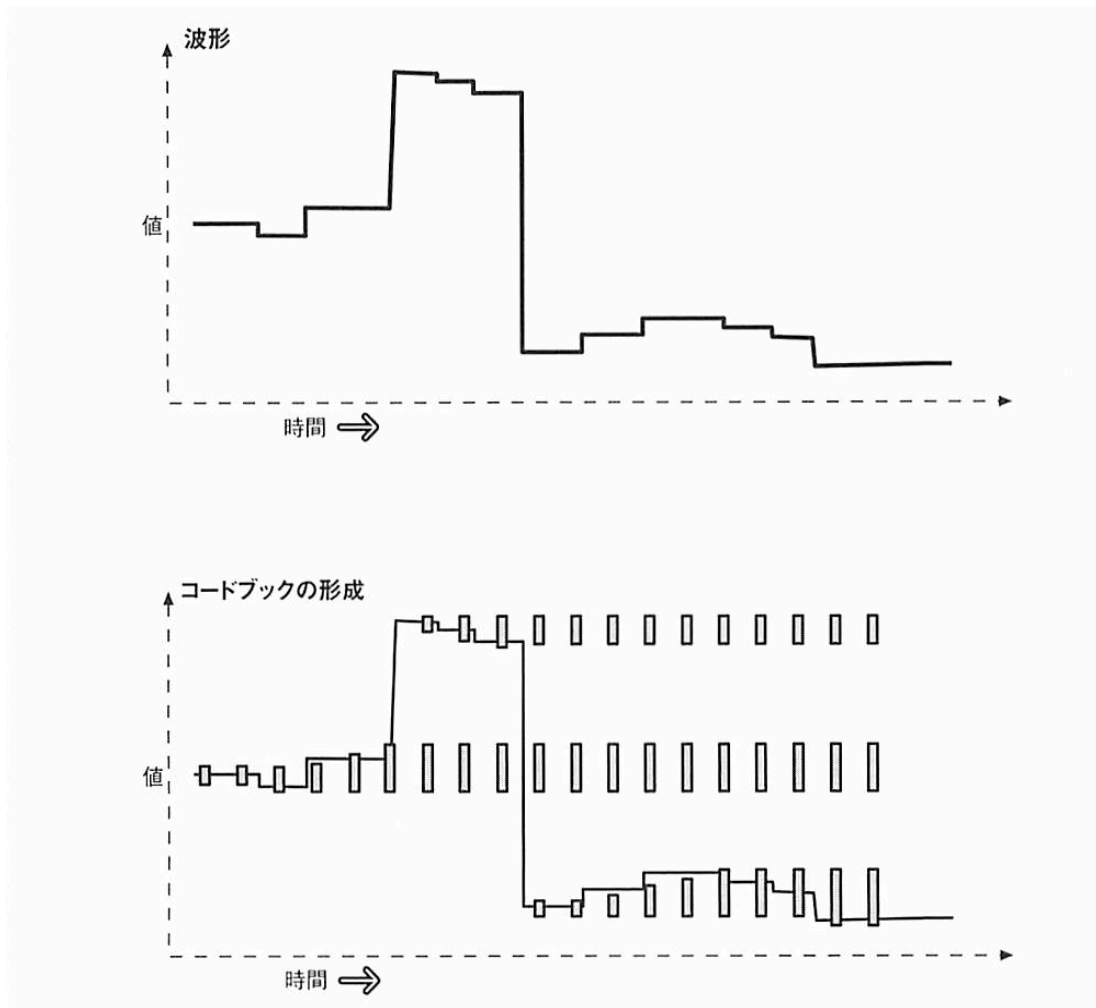


図 8.

コードブック形成のイメージ。波形は画素値の変動を表す。下図の四角はエントリであるボックスを表しており、画素値がボックスの範囲内にある場合、ボックス自身が閾値を広げて成長する。画素値が大きく変動してボックスの閾値外になったとき、新しいエントリを作成する。

3. 2 アルゴリズム

各変数の説明

- **LearnHigh** : ボックスの学習用閾値の上辺値
- **LearnLow** : ボックスの学習用閾値の下辺値
- **max** : ボックスの上辺値
- **min** : ボックスの下辺値
- **minMod, maxMod** : 前景判別のためのしきい値 (適宜、設定する)
- **last_update_time** : エントリが作成された時間
- **stale** : エントリが作成された時間から、次に更新されるまでの時間
- **cbBounds** : ボックスの学習境界。境界値は 10 程度が最適、あまり大きいとエントリの作成ができない。逆に小さいと、定常的な画素値に対応できない。

図 9 に処理の流れを示す。はじめに、コードブック構造体を定義する。コードブック構造体の配列は、画像の縦横サイズとする。動画からフレームを取得する。ここから処理毎に分けて説明する。

1. 背景の学習

次のフレームを取得する。

ボックスの学習用閾値である **LearnHigh**、**LearnLow** の値を設定する。

ボックスの大きさを更新する。**RGB** の各チャンネル毎の画素値について、**LearnLow** < 画素 < **LearnHigh** の範囲内のとき、範囲内である画素値がボックスの **RGB** の **min**、**max** を更新する。エントリの除去のために、更新までの時間をカウントする値 **stale** を更新する。

観測された画素値が、全てのエントリで学習閾値範囲外だったとき、新しいコードブックエントリを作成する。作成されたエントリのボックスの値は、観測された画素値を元に設定する。ボックスの学習閾値である **LearnHigh** は、画素値から学習境界値 **cbBounds** を加算した値、**LearnLow** は **cbBounds** を減算した値を代入する。ボックス自体の大きさとなる上辺値 **max**、下辺値 **min** は、観測された画素値をそのまま代入する。**last_update_time** は背景学習処理の総実行時間の値を代入、**stale** の値は 0 に代入する。

学習用閾値 **LearnHigh**、**LearnLow** を更新する。RGB の各チャンネル毎に、観測された画素値が **LearnHigh** より高い値のとき、**LearnHigh** を +1 して学習閾値を広げる。**LearnLow** より低い値のときは、**LearnLow** を -1 して学習閾値を広げる。

背景モデルの学習期間の間、1.の処理をループする。**Time** が学習期間である画像の枚数を超えたとき、ループを抜ける。

2. 古くなったエントリの除去

エントリが作成されてから、一定期間更新されなかった古いエントリを除去する。各エントリについて、**stale** が総実行時間の半分の値のとき、そのエントリを除去する。

次のフレームを取得する。

3. 前景領域判別

画素値が前景領域か判別を行う。全てのエントリの RGB の各チャンネルそれぞれに、条件式 $\text{min} - \text{minMod} < \text{画素値} < \text{max} + \text{maxMod}$ を適用する。画素値が条件式を満たしている場合、戻り値 255 (前景) を返し、満たさない場合は戻り値 0 を返す。閾値 **minMod**、**maxMod** の値は適用する動画によって、最適な閾値範囲になる数値を設定する。

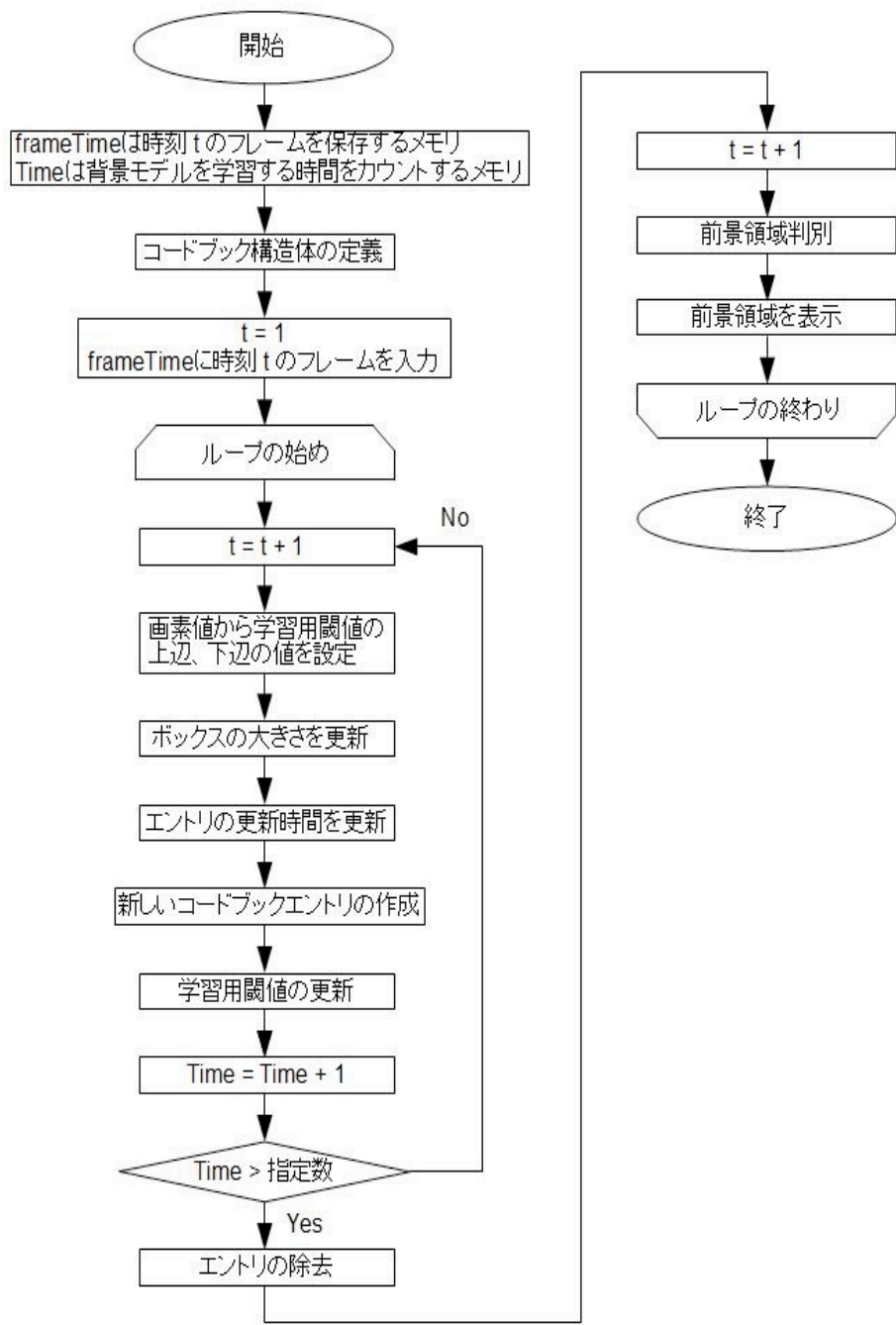


図 9. コードブック法のフローチャート

4. 実験と考察

4. 1 実験に用いた動画像

hand.avi : 320×240 画素の 80 秒の動画像。2416 フレームの画像で構成される。研究室に設置したカメラの前を手が通り過ぎる映像である。背景の変化がない場合の、前景抽出性能を確認するためのテスト動画。

treehand.avi : 320×240 画素の 2 分 28 秒の動画像。4458 フレームの画像で構成される。14 号館 2 階にカメラを設置して、外の風景を撮影した映像である。背景で木の葉が揺れている状態で、その前を手が通り過ぎる。背景が定常的に変化している場合の、前景抽出性能を確認するためのテスト動画。

basketball.avi : 320×240 画素の 2 分 10 秒の動画像。3890 フレームの画像で構成される。体育館の 2 階にカメラを設置して、バスケットボールの試合を撮影した映像である。総合的な性能を見るためのテスト動画である。

実験で使用する動画は、全て固定したカメラで撮影した。全ての動画において、背景のみが撮影された期間が存在し、それらを用いて背景モデルを作成する。検出結果の画像は 30 フレームで区切った画像を添付している。図 10～図 27 が 3 種類のテスト映像と、それらに各手法を適用した処理後の画像である。これらの処理画像をもとに、4. 2 節から考察を行う。

・ hand.avi

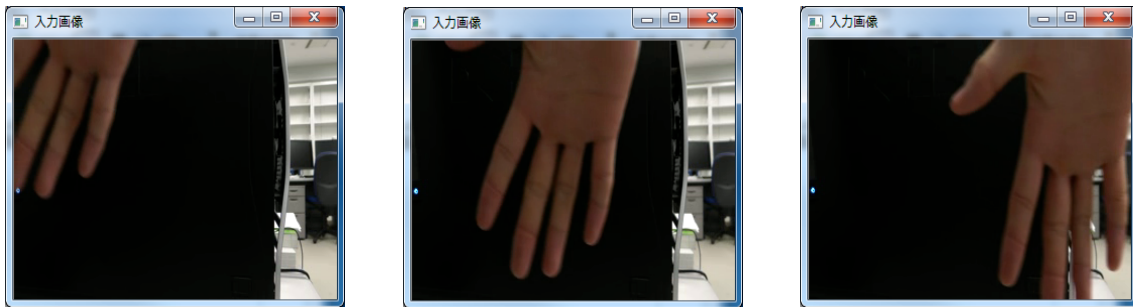


図 1 0. 入力動画(左から 1000 フレーム,1030 フレーム,1060 フレーム)

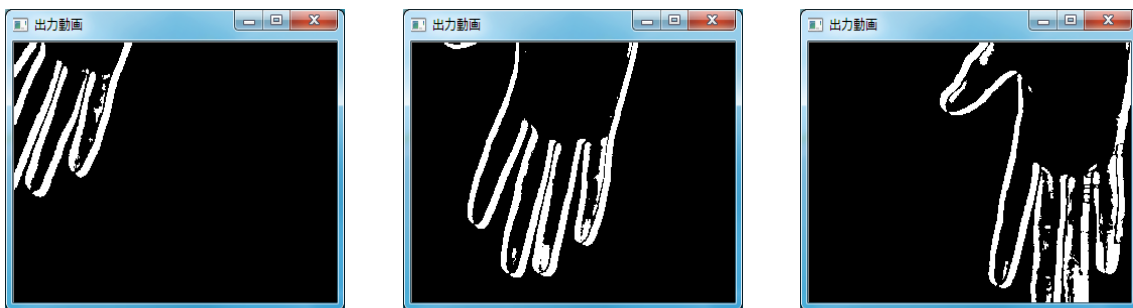


図 1 1. フレーム間差分法を適用した検出結果

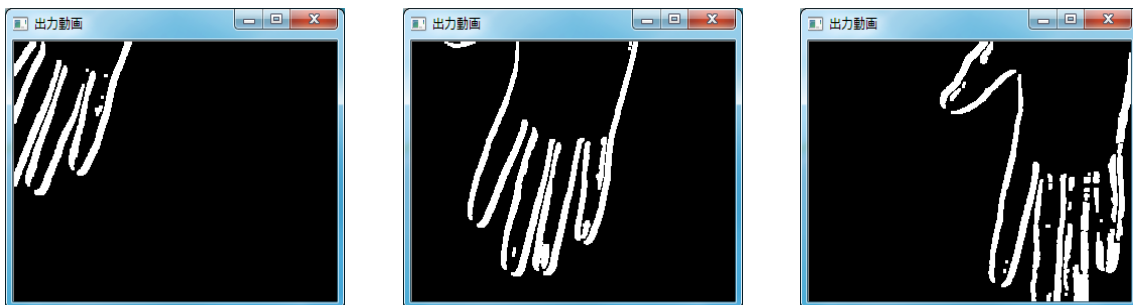


図 1 2. 図 1 1 に膨張収縮処理を適用した検出結果

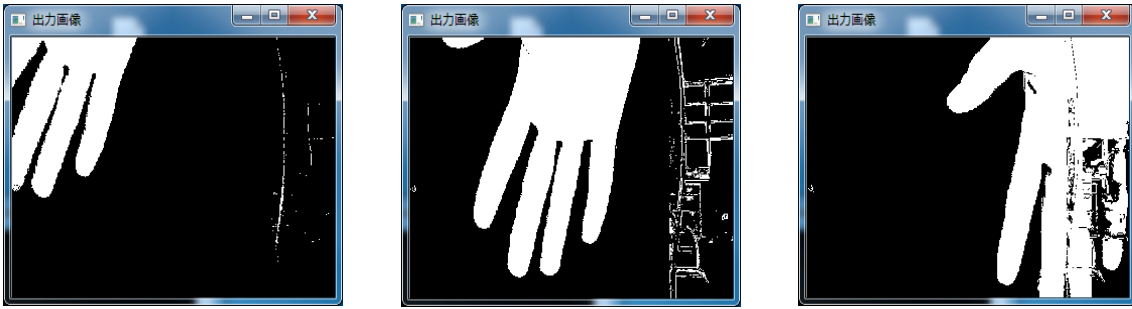


図 1 3. 差分しきい値平均背景法(しきい値 High=7.0 Low=6.0)を適用した検出結果

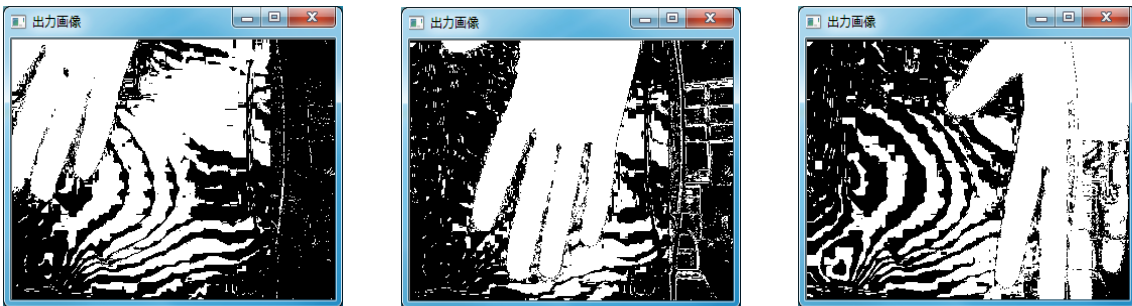


図 1 4. 分散しきい値平均背景法(しきい値 $\pm 3\sigma$)を適用した検出結果

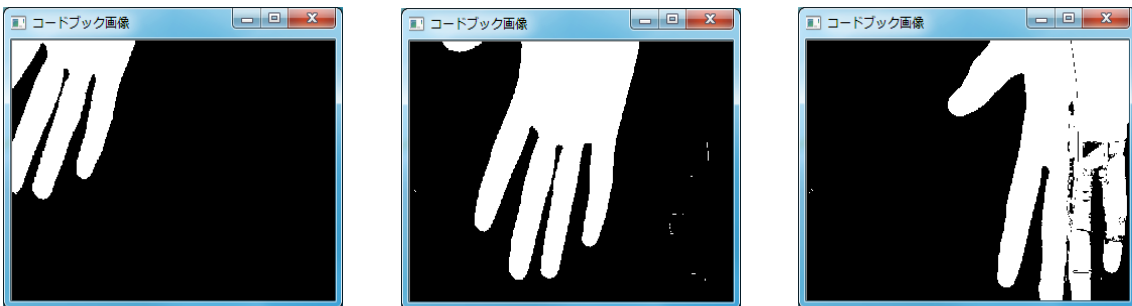


図 1 5. コードブック法(しきい値 maxMod={30,30,30} minMod={30,30,30})を適用した検出結果

・ treehand.avi



図 1 6. 入力動画(1940 フレーム,1970 フレーム,2000 フレーム)

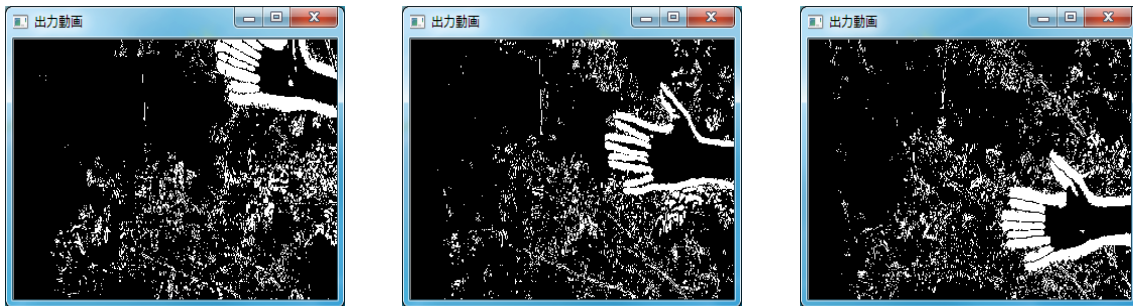


図 1 7. フレーム間差分法を適用した検出結果

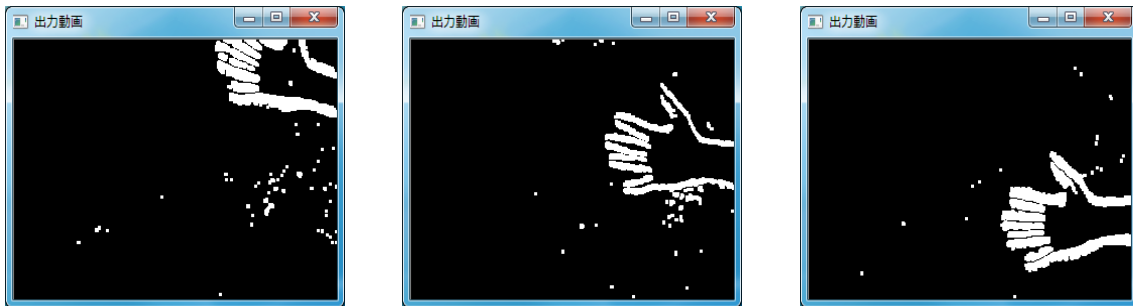


図 1 8. 図 1 7 に膨張収縮処理を適用した検出結果

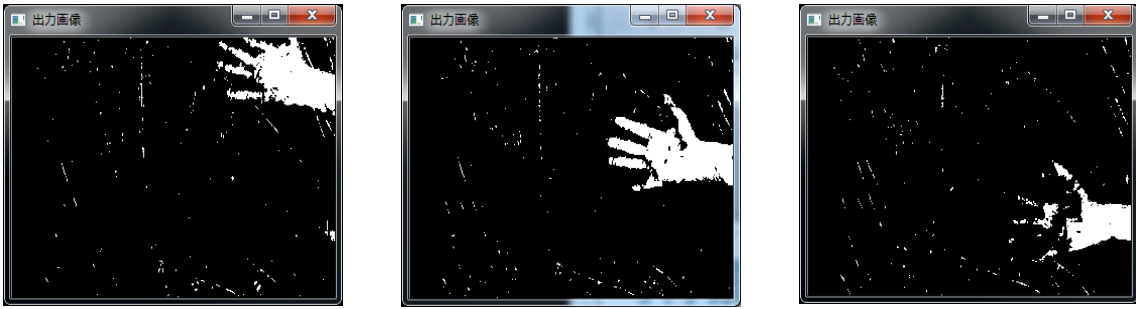


図 19. 差分しきい値平均背景法(しきい値 High=65.0 Low=65.0)を適用した検出結果

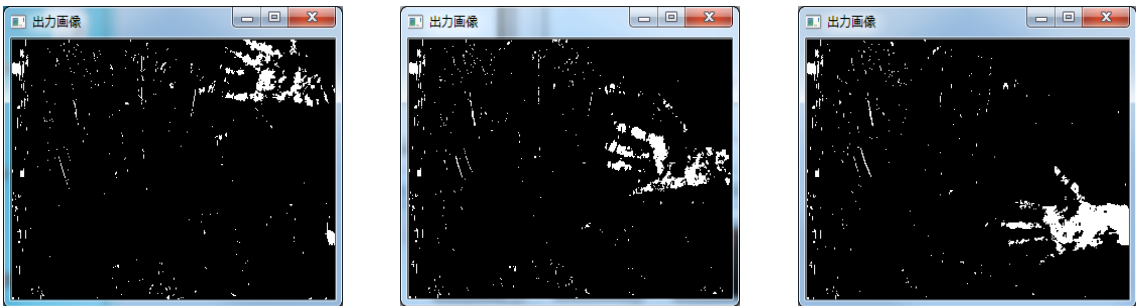


図 20. 分散しきい値平均背景法(しきい値 $\pm 3\sigma$)を適用した検出結果

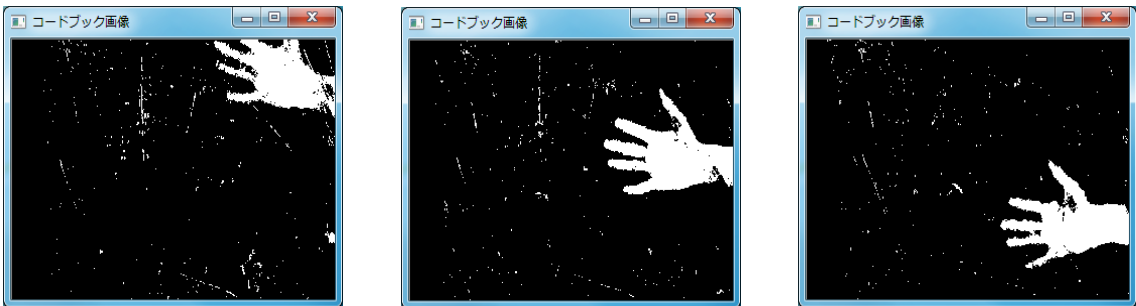


図 21. コードブック法(しきい値 $\max\text{Mod}=\{30,30,30\}$ $\min\text{Mod}=\{30,30,30\}$)を適用した検出結果

・ basketball.avi



図 2 2. 入力動画(1970 フレーム,2000 フレーム,2030 フレーム)

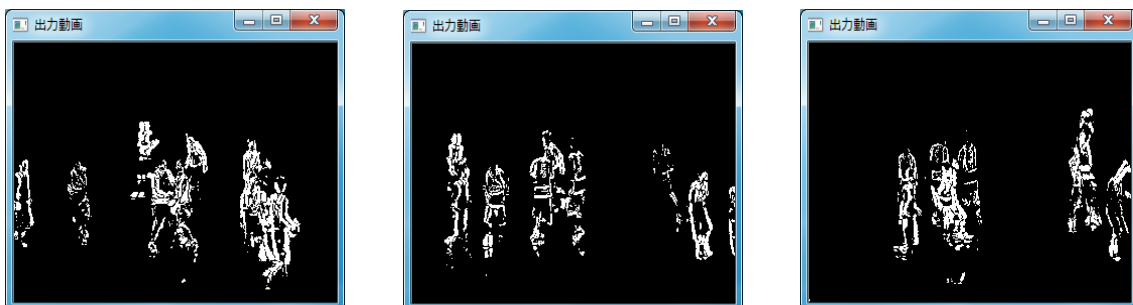


図 2 3. フレーム間差分法を適用した検出結果

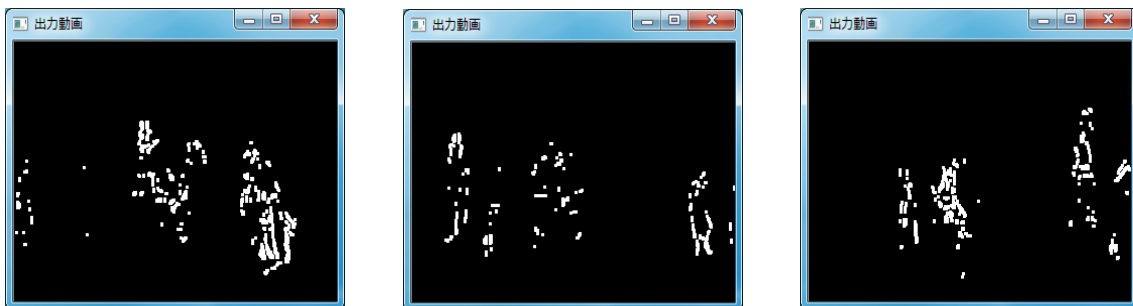


図 2 4. 図 2 3 に膨張収縮処理を適用した検出結果

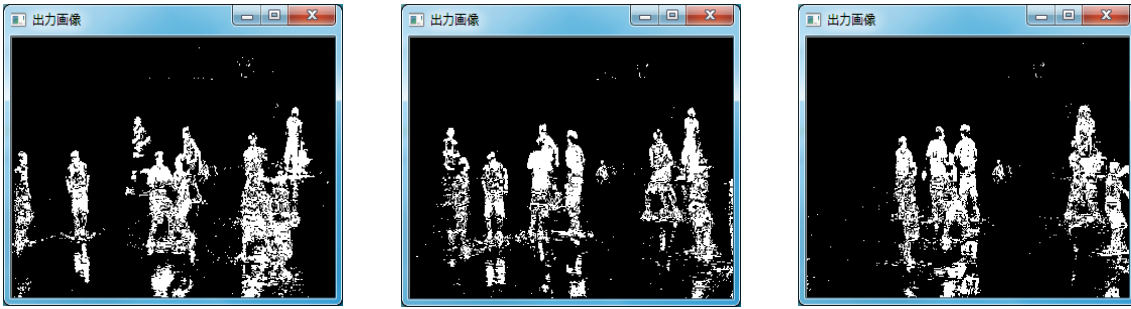


図 2 5. 差分しきい値平均背景法(しきい値 High=7.0 Low=6.0)を適用した検出結果

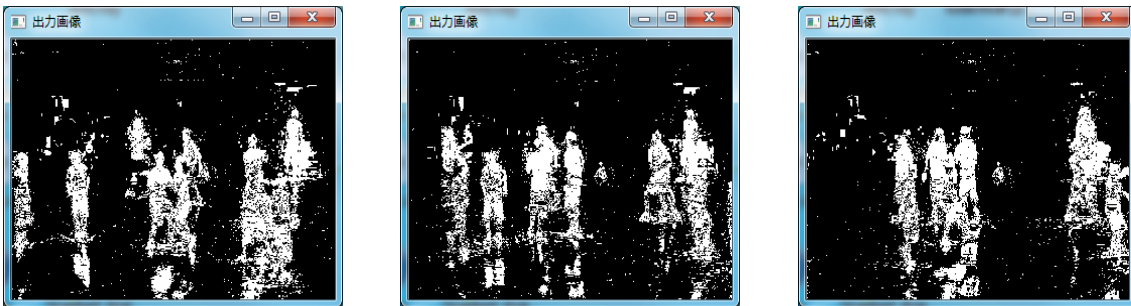


図 2 6. 分散しきい値平均背景法(しきい値 $\pm 3\sigma$)を適用した検出結果

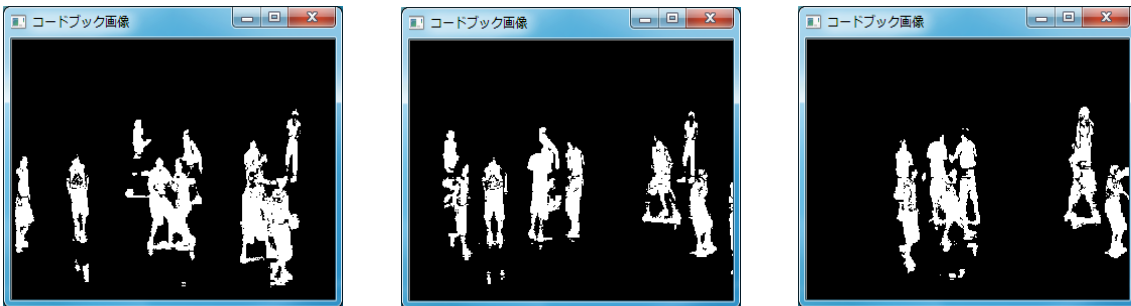


図 2 7. コードブック法(しきい値 maxMod={30,30,30} minMod={30,30,30})を適用した検出結果

4. 2 フレーム差分法による移動物体の検出結果

- hand.avi

図1 1では、移動物体である手の輪郭部分だけが前景領域として抽出されている。手の全体は均一の色領域が続いているため、フレーム差分による抽出結果は、手に穴が空いているように表示される。図1 2の膨張収縮処理は、処理前の前景画像に細かなノイズがほとんどないので、あまり変化は見られない。

- treehand.avi

図1 7では、背景の木の枝と葉の揺れが、前景領域として抽出されている。抽出すべき手は輪郭部分のみ表示されている。図1 8の膨張収縮処理では、背景として扱う木の葉の揺れをある程度除去できており、移動物体を含む前景を比較的良く抽出できている。

- basketball.avi

図2 3では、選手の服と床の色が似ている部分は、前景抽出されていない場所がある。図2 4の膨張収縮処理を適用した画像は、選手を含む前景領域が荒く表示されてしまい、効果的では無い。

4. 3 平均背景法による移動物体の検出結果

・ hand.avi

差分しきい値平均背景法を用いた場合、手の前景領域が内部まで表示されている。画像右側の、本棚や椅子の下の影で誤検出が生じており、前景領域として抽出されている。1060 フレームで、手の一部が背景領域になっている。手のひらの影と、後ろのディスプレイの画素値の差が近いため、背景として抽出されたと思われる。

分散しきい値平均背景法を用いた場合、黒い背景部分に大きく前景が出ている。おそらく、黒い部分の画素値が0に近く、標準偏差の値を引いた時、閾値を超えているのだと思われる。

・ treehand.avi

差分しきい値平均背景法を用いた場合、手の前景領域は内部まで表示されている。背景の木の葉の揺れの抽出も少なくなっているが、2000 フレームで手の前景領域が、背景として抽出されている部分が出ている。これはしきい値が高く設定されているためである。

分散しきい値平均背景法を用いた場合、木の葉の揺れが、比較的背景として抽出されている。抽出すべき手の前景領域は、背景抽出されている部分が大きくなり、手全体を前景として抽出していない。

・ basketball.avi

差分しきい値平均背景法を用いた場合、荷物や、ゴール横の布にできている影のノイズが抽出されている。コートの上に映る影と、光の反射で映る選手の姿が映っている部分も、前景抽出されている。人物は、比較的抽出できているが、床と服の画素値が近似している部分は前景抽出されにくい。髪の毛と、黒い壁が被っている部分は、画素値が似ているため背景抽出されている。

分散しきい値平均背景法を用いた場合、画像全体に、細かなノイズが増えている。分散を用いない場合と比べ、ゴール横の布にできた影のノイズを、多く抽出している。

4. 4 コードブック法による移動物体の検出結果

- **hand.avi**

手の前景領域は内部まで表示されている。ディスプレイや椅子の影、PCの境目部分で、背景領域が手の内部に多少できている。

- **treehand.avi**

木の葉の揺れの前景抽出も少なく、抽出すべき手の前景領域も、ある程度ははっきりと抽出されている。

- **basketball.avi**

選手が比較的、はっきりと前景抽出されている。人影は、はっきりと前景抽出されている部分があるが、床に映った選手の姿の前景抽出は平均背景法に比べると少ない。

5. 結論

5. 1 成果

フレーム差分法や平均背景法を用いた前景抽出に比べ、コードブック法を用いた場合、はっきりとした前景領域が表示された。定常的な動きをしている木の葉の揺れも、ほぼ背景として抽出されており、手の前景領域の抽出ができていた。

5. 2 課題

今回検討した手法では、定常的な動きをする背景は対応できたが、人物に追従して動く影を除去することはできなかった。影の画素値が黒に近いのを利用して、平均画像から影だけを認識するような処理を施すことが課題である。

現状では、抽出する画像の状況によって、最適なしきい値を設定する。背景モデルを作る場合は、更に学習の期間を長くする等が考えられる。今後は、様々な状況に対応した、更に高度な手法を考え実装することが課題となる。

参考文献

- [1]. Kyungnam Kim, Thanarat H.Chalidabhongse, David Harwood, Larry Davis,
“Real-time foreground-background segmentation using codebook model,”
AVSS '09. Sixth IEEE International Conference on Advanced Video and
Signal Based Surveillance, pp.454-459, 2009

- [2]. Kentaro Toyama, John Krumm, Barry Brumitt, Bran Meyers,
“Wallflower: Principles and Practice of Background Maintenance,”
Proceedings of the 7th IEEE International Conference on Computer
Vision, pp.255-261, 1999.

- [3]. Gary Bradski, Adrian kaehler 著, 松田 晃一 訳
“詳解 OpenCV コンピュータビジョンライブラリを使った画像処理・認識”,
O'Reilly Japan pp.267-298, 2009.

謝辞

本研究を進めるにあたり、ご指導を頂いた卒業論文指導教員の蚊野浩教授に感謝致します。また、実験の協力を快く引き受けてくださり、そして多くのご指摘を下さいました蚊野研究室の先輩・同期の皆様に感謝いたします。

付録

本研究で使用したプログラム

1.cvAbsDiff.sln

フレーム差分法を使用するプログラム

2.Ave_backg.sln

差分しきい値平均背景法を使用するプログラム

3.Ave_backg_var.sln

分散しきい値平均背景法を使用するプログラム

4.Codebook.sln

コードブック法を使用するプログラム