

# 時空間に拡張した背景差分による映像中の移動物体検出

遠藤 広人<sup>†1</sup> 阿部 亨<sup>†2,†1</sup> 菅沼 拓夫<sup>†2,†1</sup>

<sup>†1</sup> 東北大学大学院情報科学研究科 <sup>†2</sup> 東北大学サイバーサイエンスセンター

## 1 はじめに

ビデオ監視システムは、映像中の種々の物体を対象とした監視が可能であるため、現在、幅広い用途で使用されている。ビデオ監視では、一般に、監視対象を映像から抽出する前処理が必要となり、そのためには、背景差分と呼ばれる手法が広く用いられている。背景差分は、入力画像中で、背景モデル（背景画像）と見え方（色など）が異なる箇所を移動物体として検出する手法であり、対象を限定することなく様々な移動物体を検出することができる。しかし、実際のビデオ監視では、映像から抽出すべき対象が用途により異なるため、従来は、背景差分が検出した様々な移動物体の中から、目的の対象を選択する処理が別途必要となっていた。

本稿では、背景差分における入力画像と背景モデルの比較単位を時空間に拡張することで、背景モデルとの見え方の違いだけでなく、映像中での動きの状態についても評価できるようにし、その評価結果に基づき移動物体を選択的に検出可能な手法を提案する。映像中での物体の動きの状態は、物体の種類を区別する際の有効な特徴と考えられるため、提案手法を用いることで、ビデオ監視における前処理の効果的な実現が可能になると期待できる。

## 2 関連研究

背景差分は、背景との見え方の違いにより、画像中の前景（移動物体）の検出を行う手法である。具体的には、背景モデルを用意し、入力画像と比較することで、画素値など見え方が異なる箇所を前景として抽出する。従来、背景差分を用いて画像から車両を検出する手法 [1] などが提案されている。しかし、背景との見え方の違いは、物体の個々の状況により大きく異なるため、見え方の違いのみで特定の種類の物体を選択的に検出することは困難である。

また、映像中での速度を推定したオプティカルフローを併用することで、背景差分による移動物体の検出精度の向上や、速度に基づく物体の選択的な検出を図る手法が提案されている [2, 3]。しかし、ビデオ監視のための前処理として監視対象を映像から抽出するような目的では、映像中での速度に関する詳細な情報が必ずしも必要ではなく、オプティカルフローの推定を別途行うこれらの手法は、処理コストが必要以上に高くなるという問題がある。

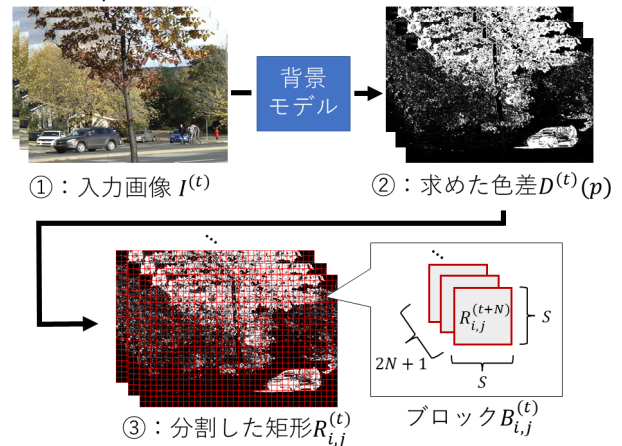


図 1: 比較単位を時空間に拡張した背景差分

## 3 提案手法

### 3.1 提案手法の概要

提案手法での処理の概要を図 1 に示す。提案手法では、背景差分における入力画像と背景モデルの比較単位を時空間に拡張する。具体的には、各入力画像（ビデオの各フレーム） $I^{(t)}$  を、大きさが  $S \times S$  画素の矩形  $R_{i,j}^{(t)}$  に分割し、連続した  $2N + 1$  枚の  $I^{(n)}$  中の  $R_{i,j}^{(n)}$  ( $n = t - N, \dots, t + N$ ) からなるブロック  $B_{i,j}^{(t)}$  を比較単位に用いる。各  $B_{i,j}^{(t)}$  において、背景モデルとの見え方の違いを評価することで、 $B_{i,j}^{(t)}$  が背景と移動物体のどちらに対応するかを判定し、さらに境界部分に対応する  $B_{i,j}^{(t)}$  では動きの状態（映像中での移動の程度）も推定する。

### 3.2 入力画像と背景モデルの比較

各入力画像  $I^{(t)}$  と背景モデル  $M$  の画素  $p$  における画素値 (Lab 表色系) を各々  $I^{(t)}(p), M(p)$  とすると、両者の色差  $D^{(t)}(p)$  は式 (1) で求められる。

$$D^{(t)}(p) = \|I^{(t)}(p) - M(p)\|^2 \quad (1)$$

また、 $I^{(t)}$  中の矩形  $R_{i,j}^{(t)}$  における平均色差  $DR_{i,j}^{(t)}$ 、ブロック  $B_{i,j}^{(t)}$  における平均色差  $DB_{i,j}^{(t)}$  は、式 (2), (3) で各々求められる。

$$DR_{i,j}^{(t)} = \frac{1}{S^2} \sum_{p \in R_{i,j}^{(t)}} D^{(t)}(p) \quad (2)$$

$$DB_{i,j}^{(t)} = \frac{1}{2N + 1} \sum_{n=t-N}^{t+N} DR_{i,j}^{(n)} \quad (3)$$

Moving object detection using spatiotemporal background subtraction

Hiroto ENDO<sup>†1</sup>, Toru ABE<sup>†2,†1</sup>, and Takuo SUGANUMA<sup>†2,†1</sup>

<sup>†1</sup> Graduate School of Information Sciences, Tohoku University

<sup>†2</sup> Cyberscience Center, Tohoku University

### 3.3 移動物体の検出

ブロック  $B_{i,j}^{(t)}$  での平均色差  $DB_{i,j}^{(t)}$  が大きければ、 $B_{i,j}^{(t)}$  は、移動物体の領域を含む可能性が高いと考えられる。そこで、 $DB_{i,j}^{(t)}$  が閾値以上なら、入力画像  $I^{(t)}$  の矩形  $R_{i,j}^{(t)}$  を移動物体に対応するものとして検出する。

### 3.4 動きの状態の推定

ブロック  $B_{i,j}^{(t)}$  を構成する矩形  $R_{i,j}^{(n)}$  ( $n = t - N, \dots, t + N$ ) での平均色差  $DR_{i,j}^{(n)}$  が入力画像  $I^{(n)}$  で変化している場合、 $B_{i,j}^{(t)}$  は、移動物体の境界部分を含む可能性が高いと考えられる。

ここで、 $DR_{i,j}^{(n)}$  の変化量は、画像中での物体の動きの状態だけでなく、物体と背景の色差にも影響される。そのため、 $B_{i,j}^{(t)}$  の近傍には全体が物体領域に対応するブロックが存在すると考え、隣接するブロックで最大の平均色差  $DB_{i,j,max}^{(n)}$  を物体と背景の平均色差とみなし、これにより  $DR_{i,j}^{(n)}$  を正規化した  $\text{norm}.DR_{i,j}^{(n)}$  を式 (4) で求める。

$$\text{norm}.DR_{i,j}^{(n)} = DR_{i,j}^{(n)} / DB_{i,j,max}^{(n)} \quad (4)$$

理想的な状況では、この値は、物体と背景の色差によらず、 $R_{i,j}^{(n)}$  が含む物体領域の面積（画素数）のみに対応する。従って、各  $B_{i,j}^{(t)}$  内で  $I^{(n)}$  に対する  $\text{norm}.DR_{i,j}^{(n)}$  の平均変化量  $a_{i,j}^{(t)}$  を求めることで、画像中での物体の動きの状態（移動の程度）を推定することが可能になる。平均変化量  $a_{i,j}^{(t)}$  は、物体の移動方向などの詳細を表すものではないが、移動の速さは概ね反映しているため、移動物体の種類を区別する際の有効な特徴になると考えられる。

また、 $B_{i,j}^{(t)}$  内で  $\text{norm}.DR_{i,j}^{(n)}$  を線形近似した際の誤差  $e_{i,j}^{(t)}$  が大きければ、ブロックに収まらない高速な移動が生じたなど、動きの状態のより詳しい推定も可能になると考えられる。

## 4 実映像を対象とした実験

実映像に提案手法を適用する実験を行った。実験には、公開されているデータセット [4] 中の、歩行する人物、走行する車両、揺れている木の葉など、異なる速度で移動する物体が同時に映っている映像 ( $720 \times 480\text{px}$ ) を用いた。入力画像の例を図 2 (a) に示す。また、提案手法中のパラメータは、 $S = 20\text{px}$ ,  $2N + 1 = 3$  枚とした。

図 2 (b) に、各ブロック  $B_{i,j}^{(t)}$  で得られた平均色差  $DB_{i,j}^{(t)}$  を示す。図中、人物（画像の下部右）、車両（画像の下部左）、木の葉（画像の上部）の箇所で値が大きくなっており、 $DB_{i,j}^{(t)}$  に対する閾値処理で移動物体の検出が可能であることが確認できる。

図 2 (c) に、各  $B_{i,j}^{(t)}$  での平均変化量  $a_{i,j}^{(t)}$  を示す。

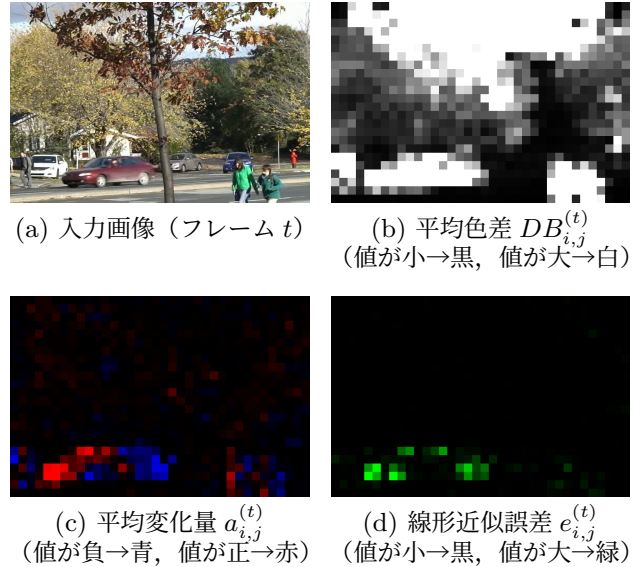


図 2: 提案手法を実映像に適用した例

図中、移動速度が比較的遅い木の葉の箇所では  $|a|$  が小さくなっている。一方、移動速度が比較的速い車両、人物の前後では  $|a|$  が大きくなっており、さらに、移動方向の前方では  $a > 0$  (赤色)、後方では  $a < 0$  (青色) になっている。このことから、 $a_{i,j}^{(t)}$  を用いることで、動きの状態に基づく物体の選択的な検出が可能であることが確認できる。

図 2 (d) に、各  $B_{i,j}^{(t)}$  での線形近似誤差  $e_{i,j}^{(t)}$  を示す。図中、車両の前後では値が大きくなっており、人物の前後を含む他の箇所では値が小さくなっている。これは、車両の動きが、今回設定したブロックには収まらない程高速であったためと考えられる。

## 5 おわりに

本稿では、背景差分を時空間に拡張することで、映像中での動きの状態に基づき、移動物体を選択的に検出可能な手法を提案した。今後、提案手法の詳細な設計を行い、種々の映像を対象とした実験により提案手法の有効性を検証する予定である。

## 参考文献

- [1] Guo, J. et al.: Vehicle detection counting algorithm based on background subtraction algorithm and SORT, *Int. Conf. Adv. Commun. Technol.*, pp.319–325 (2021).
- [2] Li, W. et al.: Foreground detection based on optical flow and background subtract, *Int. Conf. Commun. Circuits Syst.*, pp.359–362 (2010).
- [3] Parks, D.H. and Fels, S.S.: Evaluation of background subtraction algorithms with post-processing, *IEEE Fifth Int. Conf. Adv. Video Signal Based Surveillance*, pp.192–199 (2008).
- [4] Wang, Y. et al.: CDnet 2014: An expanded change detection benchmark dataset, *IEEE Conf. Comput. Vision Pattern Recognit. Workshops*, pp.393–400 (2014). <http://changedetection.net/>