

時空間特徴を利用した移動物体追跡に関する一考察

A Consideration on Object Tracking based on Spatial-Temporal Features

武田 吉史†
Yoshifumi Takeda

田中 章†
Akira Tanaka

河口 万由香†
Mayuka F. Kawaguchi

宮腰 政明†
Masaaki Miyakoshi

1 はじめに

動画画像圧縮における動き補償や監視システムによる不審人物の追跡等に代表されるように、動画画像から移動物体を検出・追跡する処理は、動画画像処理における重要な要素技術の一つであり、これまでに多くの手法が提案されている [1, 2, 3]。文献 [1] の手法では、画像として変化のあった領域に対し、ある特徴量を算出し用いることで、当該領域の変化が単なる背景変動であるか歩行者などの移動物体であるかを区別している。文献 [2, 3] では、輝度の類似性に基づいて変化領域を複数の小領域に分割し、その小領域の画像としてのパターンが不変であるとの仮定の下で、各フレームで画像マッチングを行うことで当該小領域を追跡し、最終的にそれらを統合することで移動物体を検出するという手法がとられている。これら手法に共通することは、移動物体に何らかの条件を仮定し、それに基づいて判別を行っていることであると言える。このことは当該条件で記述できない移動物体の検出、追跡が困難であることを意味する。さて、動画画像における移動物体が有する最も基本的かつ重要な性質は、それが時空間的に連続であることである。

本稿では、この考え方のみに基づき、任意の移動物体を検出・追跡可能な手法の構成を試みる。

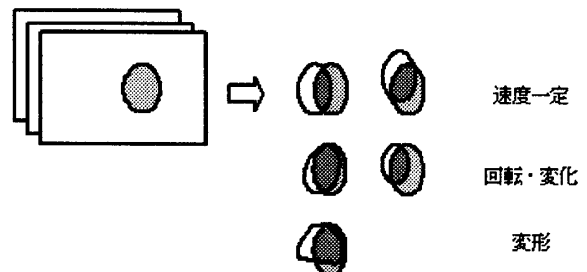
2 時空間連続性の解釈

現実世界においては、人などのオブジェクトは時間に関して連続的に移動する。しかし、動画画像情報をコンピュータで扱う場合には、ある一定のフレームレートでサンプリングを行う必要があり、それによって生じるフレーム間の連続性のとらえ方によって追跡・検出手法が異なってくる。文献 [1, 2, 3] においてはフレーム間での移動領域の存在を「輝度値の類似したものが隣接フレーム近辺に存在する」と言い換えて利用して、また、「同一変化領域はフレーム間で重なりをもつ」ととらえることで別の手法による追跡が可能になる。しかし、前者は「はじめに」でも述べたように類似性の条件が重要であり、後者では変化領域の重なり評価に速度や大きさ、拡大・縮小の様子などの領域固有の変数が必要となり対応がとりにくくなる (図 1 参照)。

そこで連続性について「各時刻の隣接フレーム間差分も連続している」とする考えを軸とすることで輝度類似性条件によらず、領域の固有要素にも依存しないオブジェクト追跡手法の構成を試みる (図 2 参照)。

3 提案手法

時刻 t における位置 x の画像輝度値を $I_t(x)$ 、推定背景輝度値を $B_t(x)$ とすると、このとき得られる変動領域 R_t (オブジェ



すべての領域で追跡可能であるが領域同士の対応がとりにくい

図1 重なり評価の困難性

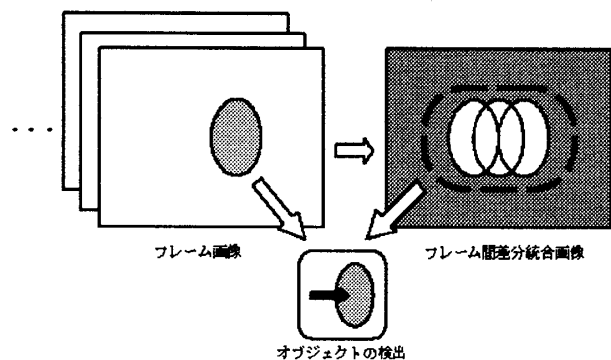


図2 フレーム間差分からの追跡

クト候補領域) は

$$R_t = \{x \mid |I_t(x) - B_t(x)| > \theta\} \quad (1)$$

と示される。ただし θ は閾値である。続いて、変動領域 R_t が隣接フレームで同一のものであるという評価方法をフレーム間差分画像から導き出す。

フレーム間差分画像 $\text{Diff}(x, 0, N-1)$ を

$$\text{Diff}(x, 0, N-1) = \begin{cases} 1, & \text{if } \exists j \in \{0, \dots, N-1\} \\ & \text{s.t. } |I_{t-j}(x) - I_{t-(j+1)}(x)| > \theta_{\text{diff}} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

と定義する。ただし θ_{diff} は閾値である。 $\text{Diff}(x, 0, N-1)$ はある対象が時刻 $t_0 - (N-1)$ から t_0 まで移動する差分画像である。

フレーム間差分は位置 x に対して背景画像に前景画像が出現したときと前景画像が位置 x から抜けて背景画像に戻ったと

†北海道大学大学院情報科学研究科 CS 専攻

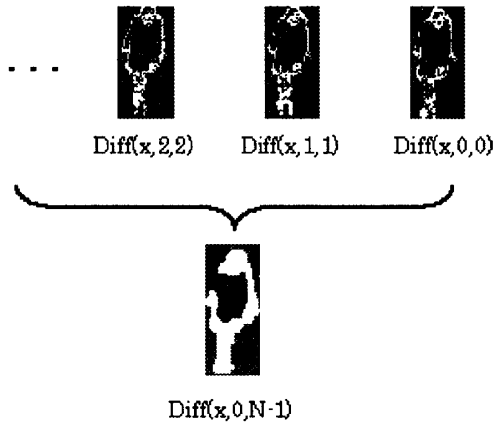


図3 Diff(x, 0, N - 1)の様子

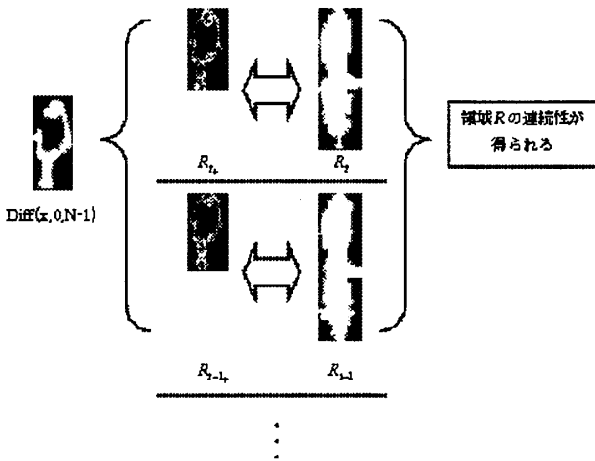


図4 フレーム間差分からの領域追跡

き、および前景画像が複数重なりあう場合に値を持つ。前者二つの差分の関係はあるオブジェクトの位置 x に対する出現差分と消失差分の関係であると考えられるためオブジェクトの移動について考える際には同じ意味をもつので、出現差分画像だけを考慮することにする。画像 $\text{Diff}(x, 0, N - 1)$ に対して、出現差分画像を

$$R_{t-i_+} = \{x | \text{Diff}(x, i, i) = 1 \cap x \in R_{t-i}\} \quad (3)$$

によって定義する。この範囲 R_+ は導出した $\text{Diff}(x, 0, N - 1)$ と同様に扱えるために時刻ごとに分割することが可能であり、さらにその分割画像は連なりをもつ。

この範囲 R_+ を用いるとオブジェクトの連続性の表現は「連続するフレーム間において、出現差分画像 R_{t+} は隣接フレームでつながっている」とすることができる。つまり出現差分画像 R_{t+} が連続していることがわかれば、その出現に対応している変化領域（オブジェクト領域） R_t も連続しているということになり、重なりが変わる追跡指標の一つとして、出現差分画像 R_{t+} の連なりが定義できる。（図4参照）

以上のことを各フレームで逐次的に実行することでオブジェクトの追跡が実現できる。提案手法の処理を以下にまとめる。



図5 実験対象動画

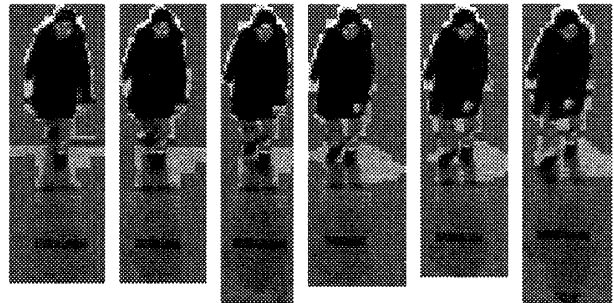


図6 提案手法による追跡結果

1. 時刻 t で差分画像 $\text{Diff}(x, 0, N - 1)$ を作成し、画像内でラベリングする。そのときのラベル l に対する差分画像を $\text{Diff}^{(l)}$ とする。
2. $\text{Diff}^{(l)}$ を構成する各時刻のフレーム間差分画像に対して、出現差分画像 R_{t+} を作る。
3. 導出した出現差分画像 R_{t+} が同一のラベルをもつ変化領域 R は時刻間で連続していると考え、これをもって R の追跡とする。
4. 同様に変化領域の追跡をすべてのラベル l に対して行う。

4 実験

4.1 実験結果

固定カメラから得られた動画像に対して、提案したフレーム間差分追跡法を適用した。動画像のフレームレートは 30fps でサイズは 360×240 とした。背景画像 $B_t(x)$ の推定手法には文献 [4] の混合正規分布を基にする手法を用い、また変化領域検出閾値 θ 、フレーム間差分閾値 θ_{diff} については適切になるように試行錯誤により設定した。カメラから得られた画像の例を図5に示す。また、提案手法による同一領域の追跡結果を図6に示す（灰色部分はオブジェクトではない領域）。結果の画像からフレーム間差分による追跡がうまく行われていることがわかる。

4.2 考察

本手法では、領域内の輝度の様子の類似度を用いていないために、比較領域の大きさ変化や縮小・拡大についての考慮をすることなく追跡が可能になっている。また、移動速度が大きい場合にもフレーム間差分の画像は同一オブジェクトで連続して

いると考えられるので、輝度差分考慮時のマッチング範囲の条件も考えることなく処理できる。

5 おわりに

動画像からオブジェクトを検出するための追跡処理に対してフレーム間差分の連続性を用いた手法を提案し、その有効性を検証した。考察で述べたように、各時刻でのオブジェクト同士の輝度対応をとることがないために、変形するオブジェクトや、領域内の輝度変化による追跡誤対応といった問題に対する解決策の一つになるのではないかと考えられる。

今回の提案手法ではオブジェクトがフレーム間で移動がある場合のみに検出が行われている。これは対象が移動しない場合にはフレーム間差分が存在しなくなるためである。また、移動が非常に遅い場合にもフレーム間差分の値は時間上でまばらに存在してしまうために連続性が導き出せずに追跡できないという問題が残る。さらに、変化領域抽出手法に依存して結果の画像のオブジェクトの正確性が変わる。これらの問題の解決が今後の課題となる。

参考文献

- [1] 羽下 哲司, 鷺見 和彦, 八木 康史, “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-DII, No.5, pp.1104-1111, 2004
- [2] Andrca Cavallaro, Olivier Steiger, Touradj Ebrahimi, “Tracking Video Objects in Cluttered Background,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.15, No.4, pp.575-584, 2005
- [3] Yaakov Tsaig, Amir Averbuch, “Automatic Segmentation of Moving Objects in Video Sequences: A Region Labeling Approach,” *ibid.*, Vol.12, No.7, pp.597-612, 2002
- [4] Chris Stauffer, W.E.L Grimson, “Adaptive Background Mixture Models for Real-Time Tracking,” *In Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp.246-252, 1999