

## 逐次背景推定による移動物体検出

三好 誠 タン ジュークイ 金 亨燮 石川 聖二

九州工業大学機械知能工学科 〒804-8550 福岡県北九州市戸畑区仙水町 1-1

E-mail: {miyoshi, etheltan, ishihawa}@ss10.cntl.kyutech.ac.jp, kim@cntl.kyutech.ac.jp

あらまし 本論文では、ゆっくりと動くカメラによって撮影された映像に対して、背景差分を用いて移動物体を検出する手法について述べる。提案手法は背景画像を逐次的に推定するために Temporal Median Filter を用いる。カメラが動くため、連続フレーム間のカメラ運動を推定し、カメラ運動でフレーム間補償を行う。そして、補償後に得られる共通画像部分に対して Temporal Median Filter を適用する。さらに、カメラ運動推定をより正確に行うために、局所相関マップを計算し、推定される動きベクトル内のアウトライヤを除去する。提案手法を移動する車両の前席に取り付けたカメラで撮影された映像に適用し、良好な結果を得た。

キーワード 物体検出, 移動カメラ, 背景差分, Temporal Median Filter, 局所相関マップ

## Extracting Moving Objects from a Video by Sequential Background Detection

Makoto MIYOSHI Joo Kooi TAN Hyungseop KIM and Seiji ISHIKAWA

Kyushu Institute of Technology Department of Mechanical & Control Engineering

1-1 Sensuicho, Tobata-ku, Kitakyushu-shi, Fukuoka, 804-8550 Japan

E-mail: {miyoshi, etheltan, ishihawa}@ss10.cntl.kyutech.ac.jp, kim@cntl.kyutech.ac.jp

**Abstract** This paper describes a technique for extracting moving objects from a video image sequence taken by a slowly moving camera by background subtraction. A temporal median filter is applied which detects the latest background images sequentially. The camera motion is estimated using a registration algorithm and the temporal median filter is applied to the common image area among a set of successive image frames to extract the background. In order that the camera motion estimation may become more exact and effective, local correlation maps are calculated which can exclude outliers in estimated motion vectors. The technique was applied to the video images taken from a camera set at the front seat of a car, and satisfactory results were obtained.

**Keyword** Object detection, Mobile camera, Background subtraction, Temporal median filter, Local correlation map

### 1. まえがき

近年の自動車技術の急速な発展により、事故を未然に防ぐための自動技術の開発も活発化している。車両周辺に存在する危険を検知するために、車へのカメラの搭載が常識化しつつある。特に、車両近くに立っている人物、または移動する人物の検出は非常に重要である。車載カメラ映像からの歩行者検出に関しては、[1]で述べられるように様々な手法が提案されている。Gavrila 他[2]は、歩行者の輪郭形状の階層型テンプレートマッチングを提案した。Viola & Jones は、顔検出のための効果的な手法である AdaBoost[3]を提案し、それを歩行者検出に応用した[4]。しかし、これを移動する車に適用するのは困難である。Curio 他[5]は人物の歩行モデルを用いた手法を提案した。この手法は車や自転車と歩行者を区別することができるが、計算コス

トが高いという問題点がある。ステレオカメラシステムを用いた手法は Zhao & Thorpe[6]によって提案されたが、ノイズの原因となる路側物体が多いため、市街地の道路への適用は困難である。Zhang 他[7]は、オプティカルフローを用いて、拡張焦点 (FOE) と呼ばれる消失点を計算し、カメラ (自動車) と異なる動きをする移動物体を検出する手法を提案した。しかし、この手法は車両が直進する場合に限られ、車両と平行して移動する歩行者を検出することができない。

上に述べた手法は、車両が通常速度で移動する場合を仮定している。しかし、車両が停止、または徐行するような横断歩道や交差点付近においても、運転手は信号などに気を取られ、歩行者などの移動物体を見落とすことがある。このような状況下での移動物体検出も重要な問題である。そこで本論文では、横断歩道や

交差点付近での移動物体の検出を目的とする。

我々は移動物体を含むビデオ映像から背景画像を推定するために、Temporal Median Filter という手法を適用する。背景画像を取得できれば、歩行者のような移動物体は現フレーム画像と背景画像との差分によって抽出することができる。正確な背景画像を得るために、局所相関マップを用いてカメラ運動を推定する。提案手法では、カメラは固定、またはゆっくりと移動するものと仮定する。つまり本研究は、車両が徐行、停止、再始動するような横断歩道や交差点付近の交通シーンを想定している。

## 2. 提案手法

### 2.1. Temporal Median Filter

画像処理におけるメディアンフィルタは、注目画素とその8近傍画素値内のメディアン値を求め、そのメディアン値を注目画素値と置き換えるノイズ除去のための平滑化フィルタである。このメディアンフィルタを空間方向ではなく、時間方向に用いる。ゆえに、本論文ではこの手法を Temporal Median Filter と呼ぶ。

固定カメラを仮定し、カメラから得られるビデオ映像を  $f(m,n,t)$  ( $m=1,2,\dots,M; n=1,2,\dots,N; t=1,2,\dots,T$ ) とする。このとき、画像  $f(m,n,t)$  に対して Temporal Median Filter を適用して得られる画像  $g(m,n,t)$  ( $m=1,2,\dots,M; n=1,2,\dots,N; t=1,2,\dots,T$ ) は以下の式で定義される。

$$g(m,n,t) = \text{med}\{f(m,n,t') \mid t' = t - F + 1, t - F + 2, \dots, t - 1, t\} \quad (1)$$

ここで、 $\text{med}\{\}$  は  $\{\}$  内のメディアン値を求める関数である。

パラメータ  $F$  は3以上の奇数である。これは  $F$  枚の連続フレーム画像に対して、Temporal Median Filter を適用することを表す。 $F$  はカメラの前を通る物体の移動速度に対応して定義される必要がある。画像上で  $M[\text{pixel}]$  の幅を持つ物体が連続フレーム間で水平方向に  $V[\text{pixel}]$  移動するとき、物体が重ならないために必要なフレーム数  $\alpha$  は以下の式で定義される；

$$\alpha = \left\lceil \frac{M}{V} + 1 \left( \frac{M}{V} - \left\lfloor \frac{M}{V} \right\rfloor \right) \right\rceil \quad (2)$$

ここで、 $\lfloor x \rfloor$  は  $x$  を超えない最大の整数であり、 $1(x)$  は以下のようなステップ関数である。

$$1(x) = \begin{cases} 1 & \dots x > 0 \\ 0 & \dots x \leq 0 \end{cases} \quad (3)$$

このように、 $\alpha$  は整数であり、パラメータ  $F$  は以下の不等式を満たすように設定される。

$$F \geq 2\alpha + 1 \quad (4)$$

この不等式は、画像上に最も長く存在する物体上の画素値においても、正確に背景として復元できるフレーム数を表している。

### 2.2. 局所相関マップ

カメラが固定されている場合は、式(1)をビデオ映像にそのまま適用すれば良好な結果が得られる。しかし、カメラが動く場合は、カメラ運動を補償する必要がある。我々はカメラ運動が比較的小さいものと仮定し、以下に述べる手法によって、フレーム間補償を行う。

まず、カメラ運動を検出するために、画像内に複数の矩形領域を設定する。各領域は式(5)で定義される SAD(Sum of Absolute Difference)を用いて、連続フレーム画像上で探索され、2枚の画像間の領域の移動量が動きベクトルとなる。領域内に移動物体が含まれる場合や、誤対応等が原因でアウトライヤが生じる可能性があるため、カメラ運動を推定する上で信頼性の高い動きベクトルを見つける必要がある。これは RANSAC[8]を用いたアウトライヤの除去を行うことによって実現できる。連続フレーム間のカメラ運動を推定できれば、補償を行うことにより連続  $F$  フレーム間の共通画像部分が得られ、その領域に対して式(1)を適用すれば背景画像を推定できる。

図1において、(a)はゆっくりと右に移動するカメラで撮影された連続フレーム画像、(b)はカメラ運動を推定するために設定された矩形領域と、連続フレーム間で求められた動きベクトル、(c)は推定されたカメラ運動で補償を行うことによって得られる共通画像部分を示している。

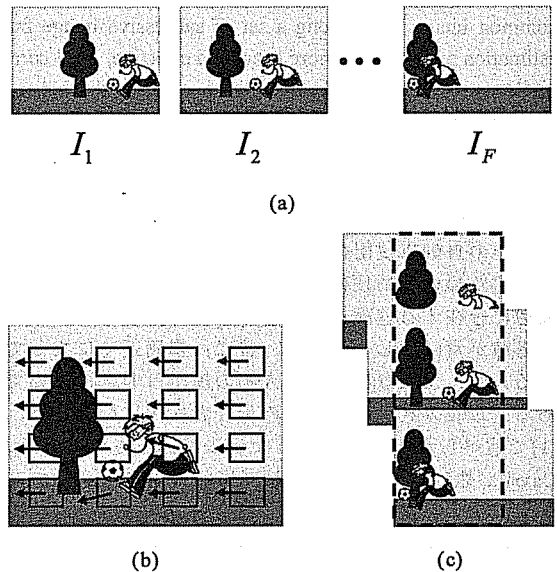


図1 移動カメラにおける手順：(a) 連続フレーム画像；(b) 局所領域と動きベクトル；(c) 共通画像部分。

しかし、領域内のテクスチャに依存して、動きベクトルを正確に求められない場合がある。この問題を解決するために、我々は局所相関マップを用いる。局所相関マップは監視カメラ映像における動き検出や解析に用いられている[9,10]。局所相関マップは、特定領域内の各画素で計算される残差を画像上でマップ化することによって得られる。点 $(u, v)$ における残差は以下の式で定義される。

$$E(u, v) = \sum_{j=1}^n \sum_{i=1}^m |f(i+u-1, j+v-1) - t(i, j)| \quad (5)$$

ここで、 $f$ は原画像であり、 $t$ はテンプレート画像、 $m$ 、 $n$ は $t$ の大きさである。 $E(u, v)$ はSADと呼ばれ、 $(u, v)$ は大きさ $UV$  ( $u=1, 2, \dots, U$ ;  $v=1, 2, \dots, V$ )の相関マップの要素である。図2に局所相関マップとその動きベクトルの例を示す。ここで、残差はグレイスケールで表されている(残差が小さいほど暗く、大きいほど明るい)。

図3は画像内の各固定矩形領域の局所相関マップを示している。ここで、円で囲まれたマップのように、動きベクトルを一意に求めることができない場合が発生する。このような状況は、道路のように矩形領域内のテクスチャが非常に少ない場合や、横断歩道のように強い直線成分を含む場合に頻繁に発生する。図4の(a)に信頼性の高いマップと、(b)に信頼性の低いマップの例を示す。

この問題を解決するために、我々は各局所相関マップのテクスチャ情報を用いてその信頼性を評価する。評価に用いる値として、テクスチャの強度を表す特徴量 $T_{mag}$ と、テクスチャの直線成分の強度を表す特徴量 $T_{line}$ の2つを定義する。具体的には、 $T_{mag}$ については加重マトリクスの空間フィルタリングを、 $T_{line}$ については周波数解析を用いることで求める。そして、以下の条件を満たす場合、その局所相関マップは信頼性が高いと判断し、その動きベクトルをカメラ運動推定に用いる。条件を満たさない場合は、その局所相関マップは信頼性が低いと判断し、用いない。

$$(T_{mag} > \theta_{mag}) \wedge \neg (T_{line} > \theta_{line}) \quad (6)$$

ここで、 $\theta_{mag}$ と $\theta_{line}$ はそれぞれ実験的に決められる閾値である。

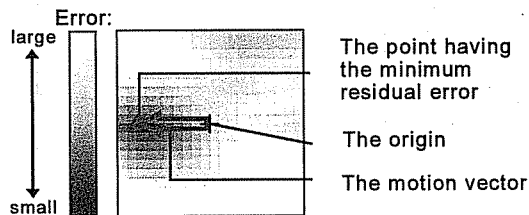


図2 局所相関マップと動きベクトル

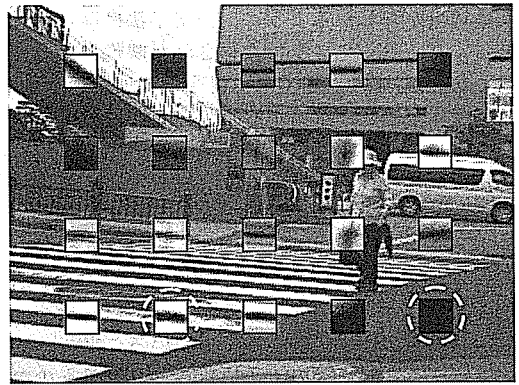


図3 画像内の局所相関マップ

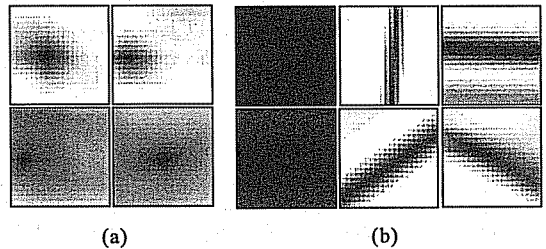


図4 局所相関マップの例：(a) 信頼性の高いマップ；(b) 信頼性の低いマップ。

### 3. 実験結果

画像内に移動する人物を含む車載カメラ映像に対して、提案手法の実験を行った。カメラは車内前席に固定され、車が市街地を走る間、車両前方の交通シーンを撮影した。この実験において、我々は局所相関マップの大きさを $20 \times 20$ [pixel]、数を20に設定した。処理時間は、2.4 GHzのPC上で約12[frame/sec]であった。

図5(a)は停止中の車から撮影された横断歩道を渡る歩行者を含む画像である。提案手法は、(b)が示すように移動物体を含んだ連続フレームから背景のみの画像を推定し、(c)が示すように歩行者を抽出している。

図6は車がゆっくりと左折する時に、横断歩道を渡る歩行者を検出した結果である。(c)が示すように歩行者はほぼ正確に抽出されている。

### 4. まとめ

本論文で提案する手法では、カメラの前を多くの移動物体が通るような困難な状況下で、逐次的に背景画像を推定するためにTemporal Median Filterを用いた。取得された背景画像は移動物体検出に用いられ、特に、車載カメラ映像からの歩行者検出に適用された。さらに、カメラ運動推定に用いる動きベクトルの検出精度を向上させるため、我々は局所相関マップの信頼性評価を用いた。

実験結果からわかるように、提案手法は固定カメラやゆっくりと動くカメラで撮影された映像に対しては有効であるが、車両が速く動く場合は精度が低下する。従って提案手法は、車両が徐行、停止、再始動するような横断歩道や交差点付近での適用を想定している。人対車の事故の多くがこのような場所で発生することを考えると、この技術は重要であると考えられる。

提案手法は逐次的に推定された背景画像を用いて移動物体を検出するため、屋外環境では不可避な照明変化の問題に対して頑健であるという利点を持つ。さらに、移動物体抽出に背景差分を用いることで、オプティカルフローベースの手法に比べて計算量も少なく、移動物体の形状を捉えやすい。また、単眼カメラシステムであるため、ステレオカメラシステムに比べて低コストであり、計算量も少ない。以上の点から、提案手法は安全運転支援システムにおいて有用な技術であると考えられる。

### 文 献

- [1] D. M. Gavrila, S. Munder: "Multi-Cue Pedestrian Detection and Tracking from a Moving Vehicle," International Journal of Computer Vision, vol.73, no.1, pp.41-59, 2007.
- [2] D. M. Gavrila: "Detection from a moving vehicle," Proc. of the European Conference on Computer Vision, vol.2, pp.37-49, 2000.
- [3] P. Viola, M. Jones: "Robust real-time object detection", Proc. 2nd Int. W. Statistical & Computational Theories of Vision - Modeling, Learning, Computing, and Sampling, pp.1-25, 2001.
- [4] P. Viola, M. Jones and D. Snow: "Detecting Pedestrians Using Patterns of Motion and Appearance", International Journal of Computer Vision, vol.63, no.2, pp.153-161, 2005.
- [5] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas, C. Bruckhoff, T. Bergener, and W. von Seelen. "Walking pedestrian detection and classification," Proc. of the DAGM-Symposium Mustererkennung, pp. 78-85, 1999.
- [6] L. Zhao and C. Thorpe: "Stereo- and neural network-based pedestrian detection," IEEE Trans. on Intelligent Transportation Systems, vol. 1, no. 3, 2000.
- [7] Y. Zhang, S. J. Kiselewich, W. A. Bauson, "Robust moving object detection at distance in the visible spectrum and beyond using a moving camera," Proc. of CVPR Workshop, 2006.
- [8] M. A. Fischler, R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Communication of ACM, vol.6, no.24, pp. 381-395, 1981.
- [9] 森田俊彦, "局所相関演算による動きの検出と追跡," 信学論(D-II), vol.J84-DII, no.2, pp.299-309, Feb. 2001.
- [10] 羽下哲司, 鷲見和彦, 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環境における歩行者の検出," 信学論(D-II), vol.J87-DII, no.5, pp. 1104-1111, May. 2004.

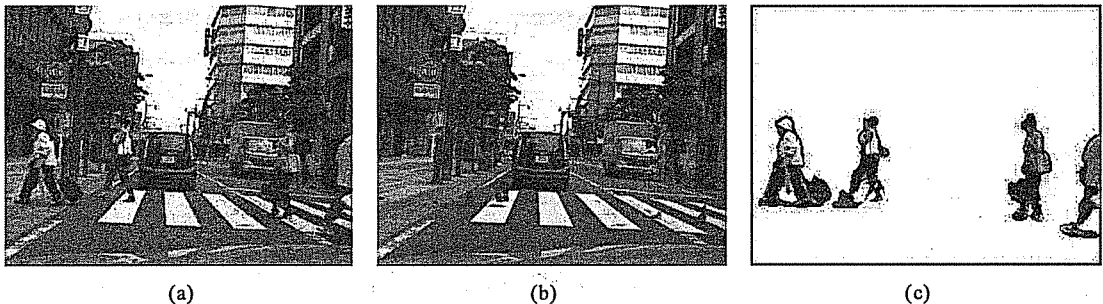


図5 実験結果1 : (a) 原画像 ; (b) 推定背景画像 ; (c) 移動物体抽出画像。

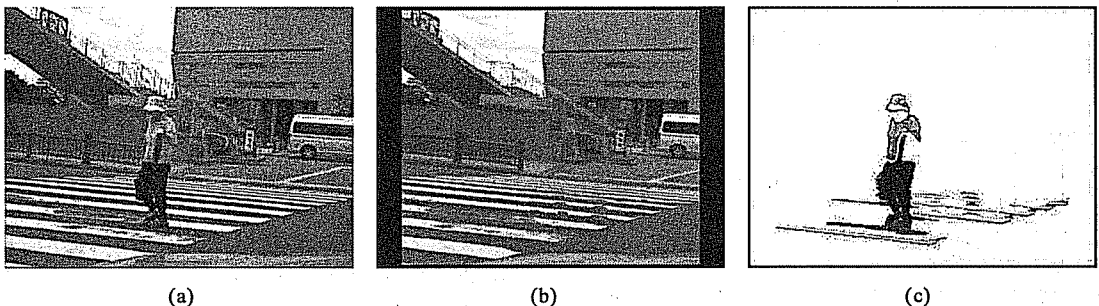


図6 実験結果2 : (a) 原画像 ; (b) 推定背景画像 ; (c) 移動物体抽出画像。