

動的背景差分を用いる歩行者飛び出し検出システム

KASAHARA, Masaki / 笠原, 正樹

(出版者 / Publisher)

法政大学大学院情報科学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 情報科学研究科編 / 法政大学大学院紀要. 情報科学研究科編

(巻 / Volume)

9

(開始ページ / Start Page)

95

(終了ページ / End Page)

100

(発行年 / Year)

2014-03

(URL)

<https://doi.org/10.15002/00010528>

動的背景差分を用いる歩行者飛び出し検出システム

A System of Detecting Walkers Using Dynamic Background Subtraction

笠原 正樹

Masaki Kasahara

法政大学情報科学研究科情報科学専攻

E-mail: masaki.kasahara.4h@stu.hosei.ac.jp

Abstract

We proposed "Dynamic Background Subtraction" system for detecting walkers from mobile camera scenes, in order to prevent traffic accidents among automobiles and the walkers. This system separated the walkers from dynamic background of the scene. The system was based on the fact that front street scene of the mobile camera extended from one point: Infinite point. Analyzing the scene extensions, current scene was precisely predicted from the previous one. Stereo-camera provided us depth information for the scene prediction. The difference between the predicted scene and the real one removed background objects in Region of Interested, where the walkers run out into street in the scenes. In order to reduce the background noises made by vibration and/or curve movement of the automobile, the predicted scene was shifted and subtracted from the real one. Morphological operation indicated the walkers' residual. The proposed system was characterized by its simplicity in principle and high potentiality in easily realizing the system with low cost. The method was able to prevent the miss-extraction of background objects and un-extraction of walkers. In this paper, the principle and the procedures of the method were described, and the experimentation which detected walkers from mobile scenes by using the method showed its utility.

1. まえがき

近年、歩行者と自動車による交通事故は、運転手の脇見運転や安全不確認が主な原因となっている[1]。これを防ぐため、自動車にカメラやレーダーを設置し、それによって道路へ飛び出す歩行者を検出し、運転手に警告または自動的にブレーキをかけるシステムが開発されてきた。レーダーによる検出は、トヨタの PCS[2]等で実用化されている。一方でレーダーよりもコストが低いため、ホンダのインテリジェント・ナイトビジョンシステム[3]やスバルのアイサイト[4]など、カメラの画像による検出システムも実用化されている。

車載カメラの画像から歩行者を検出する方法は、歩行者の外観が多様多岐なため、色情報に依存しない輪郭エッジなどの形状マッチングが主流となっている。具体的には HOG[5]やウェーブレット[6]の利用、テンプレートマッチング[7]等の手法が挙げられる。しかし、いずれの手

法でも、街路樹や標識など歩行者以外の画像の背景物による誤検出や、帽子の着用など歩行者の特殊な形状による未検出が大きな課題となっている。

提案するシステムでは動的背景差分法を用いて、車載カメラ画像の背景を除去し歩行者の飛び出しを検出する。本来の背景差分法は、監視カメラのような定点カメラで、時間的に不変な背景を 1 フレーム前の画像と現在のフレームで差分して除去し、移動物体のみを抽出する手法である[8]。しかし、車載カメラのようにカメラ自体が動く場合、背景も見かけ上の動きを持つため、従来の背景差分では背景除去が困難となる。提案した動的背景差分は、車載カメラの背景がある 1 点から放射状に広がることを利用し、1 フレーム前の画像から現在のフレームの背景を予測生成し、生成された背景と実際の現在フレーム画像との差分をとることによって背景を除去する。これにより、背景の動きと異なる道路へ飛び出す歩行者を検出できる。この手法は非常に単純な原理と低コストでカメラ画像から歩行者の検出を実現でき、かつ形状マッチング法の課題である背景物の誤検出や歩行者の未検出を防ぐことが可能である。本論文ではこのステレオカメラによる車載映像で動的背景差分を行い、実験を通して本手法による歩行者検出の有用性を示す。

2. 原理と処理手順

提案した動的背景差分法は、最初に無限遠点をあらかじめ導出し、無限遠点から放射状に動く背景を分析することで、1 フレーム前の画像から現在の背景を予測生成する。背景の予測生成には背景の物体とカメラの距離が必要であり、この距離情報をステレオカメラによって得る。画像の歩行者の飛び出しが想定される領域 ROI(Region of Interested)内で、生成された背景と実際の現在フレーム画像を差し引きし、背景を除去する。この時、車のカーブや振動によって発生する背景の雑音を軽減するため、予測画像を平行移動させながら差分を行う。最後に残った微小な雑音を除去し、差分結果から歩行者の残差を抽出する。以下に動的背景差分法の詳細を述べる。

2.1. 無限遠点

車が直進する場合、車載カメラ画像中の静止している背景はある 1 点から放射状に広がっていく。この点を無限遠点と呼ぶ。無限遠点はカメラの配置や向きによって変わるが、その配置と向きを保てればいかなるシーンでも導出した無限遠点が使え、したがって、車が直進した時の数枚のフレーム画像から、あらかじめ無限遠点を導出する。

無限遠点の導出として、まず複数枚の車載カメラ画像から、複数の背景の対応点のグループを抽出する。対応点の抽出にはオプティカルフロー[9][10]を使用した。1フレーム(1/30 秒)間における2枚の画像から、オプティカルフローによって背景点のフローベクトルを算出する。ベクトルの中には、本来の背景点の動きでは有り得ない方向や大きさを持つ誤りベクトルも算出される。誤りベクトルを除くため、算出された全てのベクトルの大きさの平均値を求め、この平均値を下回る大きさを持つベクトルを背景点のフローと定め、その始点と終点の位置を取得する。これを n フレーム繰り返し、 m 個の対応点グループ、すなわち合計で $n \times m$ 個の対応点が抽出される。

図1に無限遠点導出原理の概略を示す。白点で示した座標 $(\bar{x}_j, \bar{y}_j)$ は対応点グループ j における重心である。ここで無限遠点の候補である点 I' を仮定する。これらの点は無限遠点 I を始点とする半直線上に位置するはずなので、候補点 I' を始点とし各対応点グループの重心を通る半直線を考え、グループ内の全ての点との距離の二乗和 S が最小となるように I' を走査し、無限遠点 I を定める。最小二乗和 S は次の式(1)のように示せる。

$$S = \min_{x,y} \sum_{j=1}^m \sum_{i=1}^n \left(\frac{\alpha_j x_{ij} - y_{ij} + \beta_j}{\sqrt{\alpha_j^2 + \beta_j^2}} \right)^2, \quad (1)$$

$$\alpha_j = \frac{\bar{y}_j - y}{\bar{x}_j - x}, \beta_j = \frac{\bar{x}_j y - x \bar{y}_j}{\bar{x}_j - x},$$

導出の際に、カメラのレンズによる画像の歪みに基づく誤差を避けるため、カメラキャリブレーションによって歪みを事前に補正する[11][12]。

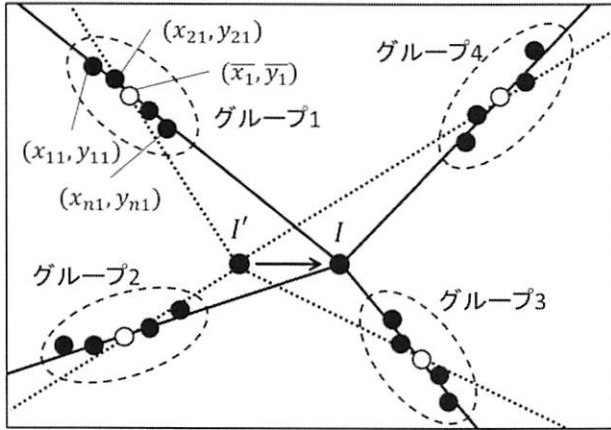


図1 無限遠点の導出

図2にオプティカルフローと無限遠点を描画したものを示す。この画像でのオプティカルフローは、無限遠点との関係を明確に図示するため、1秒間におけるフローを描画した。画像中の矢印で示した点が無限遠点である。ほとんどの背景点のフローが無限遠点に向かって伸びているのがわかる。



図2 オプティカルフローと無限遠点

2.2. 背景の予測

背景の点は無限遠点 I を始点とした半直線上を移動するので、この移動を分析することにより1フレーム間における背景の変化を予測することができる。図3に背景と画像の点の関係を示す。あるフレームにおける背景の位置 W_0 が、次のフレームでカメラの直進によって W_1 に移動するとき、画像の点の移動 $P_0 \rightarrow P_1$ を考える。カメラの位置を O 、焦点距離を f 、 W_0 とカメラの距離を Z 、進行方向との距離を L 、 $W_0 \rightarrow W_1$ の移動量を ΔZ とし、カメラの進行方向は線分 IO の延長上とする。図2より、 $\triangle OIP_0$ と $\triangle OZ_0W_0$ 、及び $\triangle OIP_1$ と $\triangle OZ_1W_1$ は相似であり、ゆえに無限遠点 I と点 P_0 及び P_1 の関係は次の式(2)のように表せる。

$$\left. \begin{aligned} \frac{\overline{IP_0}}{f} &= \frac{L}{Z}, \\ \frac{\overline{IP_1}}{f} &= \frac{L}{Z - \Delta Z} \end{aligned} \right\} \quad (2)$$

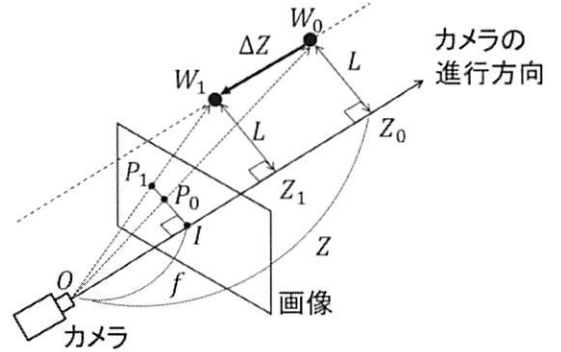


図3 背景と画像の点の関係

ここで現在の背景を予測生成するための比率 $\overline{IP_1}/\overline{IP_0}$ を拡大率 k とする。カメラ(自動車)の速度を v 、1フレーム当たりの時間間隔を $\Delta t (= 1/30 \text{ 秒})$ とすると、式(2)より k は式(3)で表すことができる。

$$k = \frac{\overline{IP_1}}{\overline{IP_0}} = \frac{Z}{Z - \Delta Z} = \frac{Z}{Z - v\Delta t} \quad (3)$$

式(3)より、拡大率 k は速度 v とカメラとの距離 Z の2変数によって決まる。図4に速度と距離による拡大率の変化を示す。距離 Z は、ステレオカメラによるデプス画像の距離情報を用いて求める。

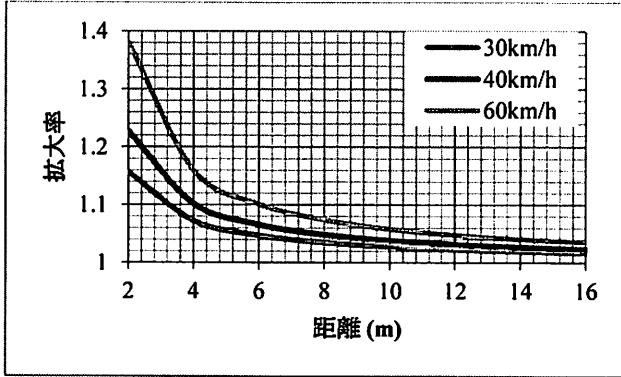


図4 速さと距離による拡大率の変化

2.3. 距離画像の取得

2つのカメラからなるステレオカメラにより、カメラと対象物の距離によってグレー画像の階調値（デプス値）が決まるデプス画像が取得できる。図5にステレオカメラの概略図を示す。物体 W と両カメラとの距離を Z 、左カメラの光軸との距離を X 、焦点距離を f 、ピクセルとメートルの変換係数を c 、右カメラと左カメラ間の距離（基線長）を b とし、左画像上の物体の点 x_l と右画像上の点 x_r を考える。図5より、 $\triangle O_l f_l x_l$ と $\triangle O_l W_l W$ 、及び $\triangle O_r f_r x_r$ と $\triangle O_r W_r W$ は相似であり、ゆえに x_l と x_r は次の(4)式で表せる。

$$\left. \begin{aligned} \frac{x_l}{f c} &= \frac{X}{Z}, \\ \frac{x_r}{f c} &= \frac{X - b}{Z}, \end{aligned} \right\} \quad (4)$$

式(4)を Z について整理すると、式(5)が得られる。 x_l と x_r の差 d によってデプス画像は生成される。

$$Z = f c \frac{b}{x_l - x_r} = f c \frac{b}{d}, \quad (5)$$

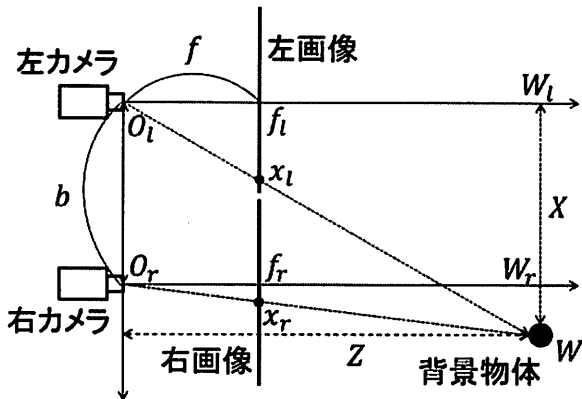


図5 ステレオカメラシステムと距離情報

デプス画像の例は図7(b)に示す。図6のグラフは、距離による画像のデプス値の変化を示す。デプス値の変化は、カメラの光学ズーム倍率によって変わるので、グラフには本研究の実験で使用したカメラの最大倍率3倍の曲線も示す。このデプス値を距離情報として利用する。

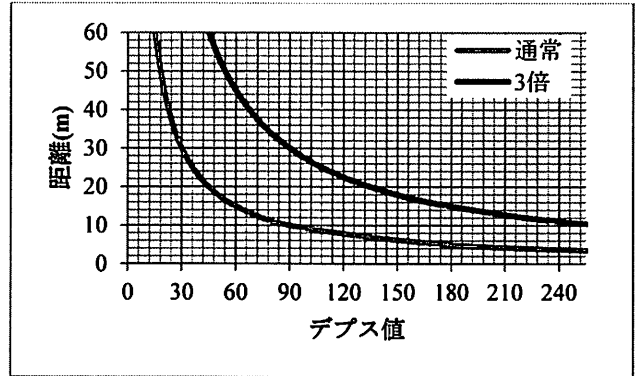


図6 距離とデプス値のグラフ

2.4. 予測背景の生成

背景の予測生成に必要な2変数のうち、カメラとの距離 Z はデプス画像の距離情報により取得可能となった。車速 v については、正確に取得するには速度センサー等の外部デバイスが必要となるが、そのようなデバイスがない場合でも Z が既知ならば式(3)を変形して、次の式(6)で v が求まる。また式(6)以外に、1秒間における画像の背景物のデプス値の変化でも速度は取得可能である。

$$v \Delta t = Z \frac{\overrightarrow{IP_1} - \overrightarrow{IP_0}}{\overrightarrow{IP_1}}, \quad (6)$$

予測画像は、左カメラの画像に重なる形で生成する。まず、1フレーム前の左右画像からデプス画像を生成する。次に、前フレーム画像をデプス値毎に領域分けし、領域画像を取得する。デプス値は0-255の階調値が存在するが、デプス値0は空やオクリュージョン等を含むため除外し、最大で255枚の領域画像が生成される。同時にデプス値毎の拡大率を求め、領域画像を対応する拡大率だけ無限遠点を中心に拡大する。最後に拡大した領域画像を合成し、前フレーム画像の解像度に合わせてトリミングして予測背景を得る。

図7に予測背景の生成の過程を示す。7(a)が1フレーム前の左画像で、右画像とともに7(b)のデプス画像が得られる。その後、デプス値に従い、7(c)の領域画像が取得される。最後に領域画像はデプス値に対応して拡大され、7(d)の予測画像が生成される。虫食いになっている部分はデプス値0、または拡大合成時のすき間である。

2.5. ROIの設定

処理の効率を上げるため、歩行者を検出する領域ROIを設定する。このROIは空や地面といった歩行者が飛び出しえない場所や、対向車による誤検出を防ぐため対向車線を除いた、自転車線と隣接する歩道に絞った領域であり、この中で差分処理を行う。

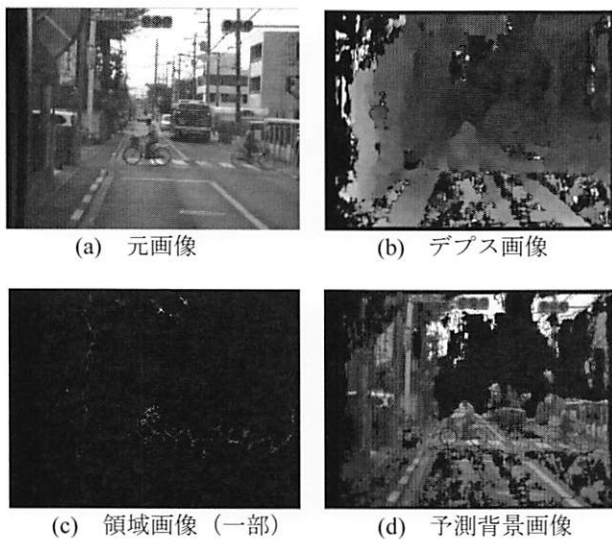


図7 予測背景生成の流れ

ROIを設定するため、距離とピクセルの関係を分析する。縦横 30x20cm の看板を 5m 毎に撮影し、640x480 の解像度における各距離の看板のピクセルの大きさを計測した。図8のグラフは、カメラとの距離によって長さ 1m の物が画像では何ピクセルになるかを示す。このグラフを基に、ROI を距離 5m (光学 3 倍の時 15m) における、横が 1 車線分 2.5m に相当する 375 ピクセル、縦が人の平均身長 1.7m に相当する 255 ピクセル、無限遠点を中心に設定した。設定した ROI を図9に示す。

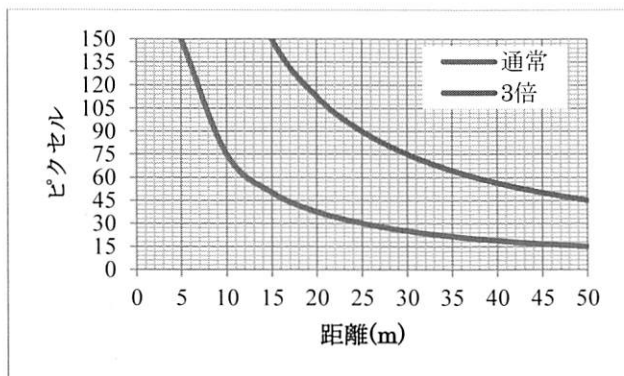


図8 距離とピクセルの関係



図9 設定した ROI

2.6. 差分と検出

単純な背景差分では、前フレーム画像と現フレーム画像での同じ位置のピクセルの RGB 値の差をとり、その差が設定した閾値を下回れば背景とみなし除去する。ゆえに移動物体が画像中に存在する場合、差が閾値を大きく上回り、残差として残る。

動的背景差分でも、予測生成した背景と現在のフレーム画像を差し引き、歩行者の残差を抽出する。2.5 節で述べた ROI 内で差分を行う。予測生成した背景には、デプス値 0 部分の除外や拡大合成時のすき間等で虫食い部分があるので、その部分以外で差分する。また、車の振動やカーブによる背景の残差を減らすため、2.7 節で述べる予測画像の平行移動をしながら差分を行う。

差分結果では、歩行者の残差以外に、無限遠点やデプス値の誤差等で発生する背景部分の残差が細かな点や線状に雑音として残る。雑音を除去し歩行者の残差のみを抽出するため、背景部分を黒、残差部分を白とした 2 値の残差画像を一旦出力する。

2 値画像での細かな雑音を除去する方法として、モルフォロジー変換がある。この変換は、白画素領域を増やす膨張と、黒画素領域を増やす収縮処理を交互に行う。まずオープニングという収縮の次に膨張を行う処理で細かい雑音を除去し、その後クロージングという膨張の次に縮小を行う処理で歩行者の残差を明確に抽出する。オープニングでの膨張収縮回数は 1 回、クロージングでは 16 回行った。最後に雑音処理を行った残差画像を基に、現在のフレーム画像に残差の位置を矩形で示す。

図 10 に雑音除去と歩行者検出を示す。10(a)は、現在のフレーム画像と図 7(d)の予測背景画像との差分で生成された、残差画像である。10(b)がオープニングによる雑音除去後であり、背景の細かな雑音がほとんど消え、一方歩行者の残差が残った。10(c)がクロージングによる残差の抽出であり、10(b)で小さく散らばっていた歩行者の残差が 1 つの大きな残差に纏まった。最後に 10(c)での残差の位置を、10(d)のように現フレーム画像に矩形で示す。

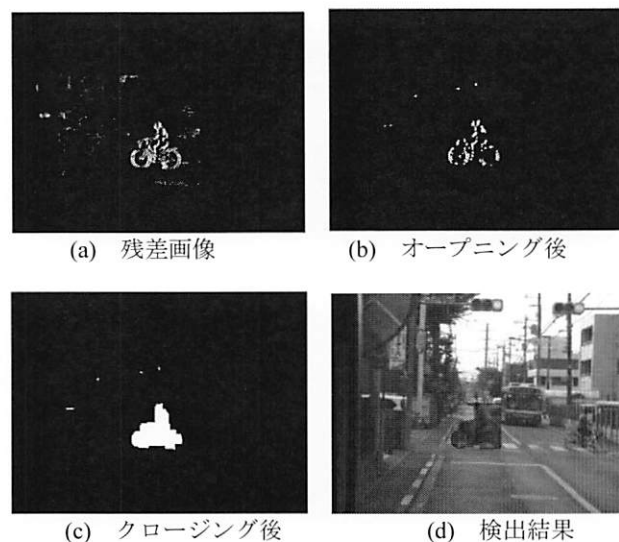


図10 差分と検出処理

2.7. 振動とカーブ

車は地面の状態によって上下左右に小刻みに振動し、それによってカメラ画像の背景の残差が大きく残り、背景部分の誤検出を引き起こす。振動による背景の残差を軽減するため、振動の揺れ幅を分析し、揺れ幅内で画像を平行移動させながら差分を行う。

図 11 に実際の背景点と理論上の背景点の軌跡との関係を示す。このグラフは、1 秒(30 フレーム)間における背景の点をトレースし、本来の無限遠点を始点とした背景点の軌跡と比較した。直線は背景点の軌跡、点は実際の背景点の座標を示す。トレースの結果、1 フレーム間における振動幅の最大は上下方向で ± 5 ピクセル、左右方向で ± 10 ピクセルとなった。この範囲で予測画像を上下左右に平行移動し、現フレーム画像との差分での残差率 r_R を取得する。残差率は、ROI 内の全ピクセル数を P 、残差のピクセル数を R とし、次の(7)式で示す。この残差率が一番低い時に、2.6 節で述べた雑音除去と検出工程に移る。

$$r_R = \frac{R}{P} \quad (7)$$

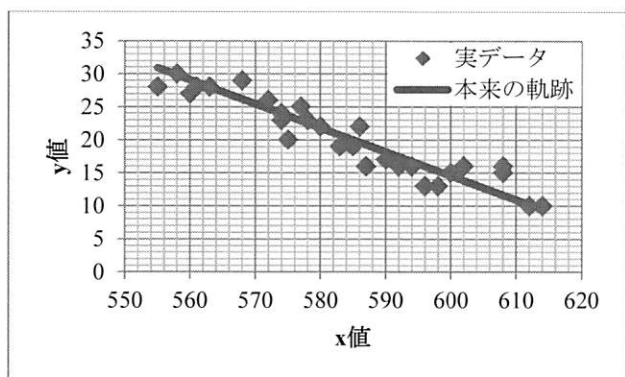
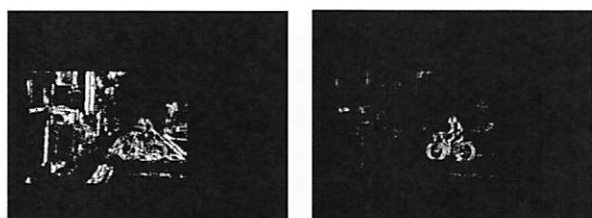


図 11 実際の背景点と理論上の点の軌跡

図 12 は、図 10(a)の残差画像での平行移動補正の有無による結果を示す。図 12(a)は平行移動無しで差分を行った結果であり、背景成分の残差も多く残ったが、12(b)では背景の残差が大きく減り、歩行者の残差が精度良く抽出できているのがわかる。残差率は(a)が 17.2%、(b)が 3.5%となった。

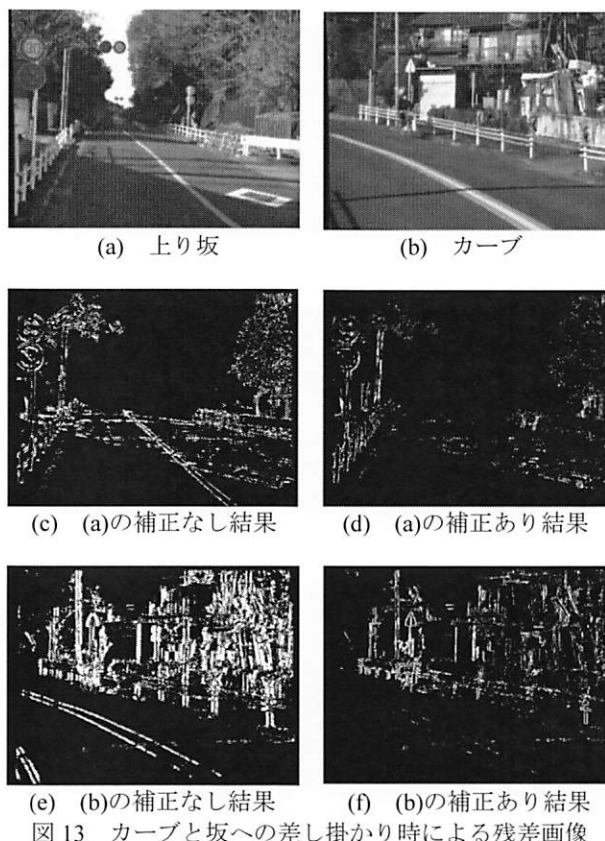


(a) 補正なし (b) 補正あり
図 12 平行移動補正の有無による残差画像

画像の平行移動による振動対策は、カーブや坂への差し掛かりにも適用可能である。これらの状況では車(カメラ)が曲線移動し、背景は直進が前提の背景予測と異なる動きをするが、1 フレーム間(1/30 秒)においての曲線

移動は非常に微小で局所的に直進すると近似でき、画像の位置ずれとして決できる。

図 13(a), (c), (d)は上り坂への差し掛かり、13(b), (e), (f)はカーブ時の、元画像と平行移動補正の有無の残差画像を示す。ここでは、全体的にどれだけ背景の残差が軽減できたかを示すため、ROI は設定せずに画像全体で評価を行った。表 1 に各シーンでの平行移動の有無による残差率を示す。図 13 と表 1 より、坂の差し掛かりとカーブの 2 シーンでも、平行移動ありの差分で背景の残差を減らすことができた。ROI を設定することによって、さらに背景の残差による誤検出を抑えることが可能となる。



(c) (a)の補正なし結果 (d) (a)の補正あり結果
(e) (b)の補正なし結果 (f) (b)の補正あり結果
図 13 カーブと坂への差し掛かり時による残差画像

表 1 各シーンの残差画像

シーン	補正なし残差率(%)	補正あり残差率(%)
上り坂	7.12%	2.58%
カーブ	18.27%	5.86%

3. 実験と考察

3.1. 実験

提案した動的背景差分法による歩行者検出の有用性を示すため、ステレオカメラを車に載せ、車が移動している状態での映像で、横断歩道を渡る歩行者を検出する実験を行った。実験ではより遠くの歩行者を検出するため、カメラレンズ光学 3 倍ズームをかけた映像を使用した。映像での車の時速はおよそ 10.8km/h、実験で使ったステレオカメラは FUJIFILM 社の FINE PIX REAL 3D、画像

フレーム数は 120 枚, 処理に使った CPU は Intel Core i5-2520M CPU(2.50GHz)である. 処理時間は全体で平均約 65 秒となった.

3.2. 結果と考察

図 14 に実験結果を示す. 14(a)はこの実験における無限遠点と ROI で, この ROI 内での歩行者を検出した. 14(b)は最も検出に成功した一例である. 横断歩道を渡る歩行者を検出できており, 背景物の誤検出もほとんどない. 14(c)は背景部分に大きく誤検出した例であり, 左の自転車と上部の背景部分の検出矩形が合体している. 14(d)は歩行者を検出できなかった例であり, 丸で示した歩行者に検出を示す矩形がかからなかった.

実験の結果, ほとんどの画像の ROI 内において, 全体的に背景部分の誤検出は少なくなかったが, 横断歩道を渡る歩行者の検出には成功した. 図 14(d)のように, 位置的に最もカメラから離れた場所を横断する歩行者の検出に失敗した例は 120 枚中 7 枚であり, 未検出率にすると 5.8%となった. また ROI 内で, 自転車と近くの 2 組の歩行者(図 14(a)右)は全て検出できた.

動的背景差分法による検出では, カメラと歩行者の距離が離れるほど画像の歩行者の移動量が少なくなり, 歩行者の残差が背景の雑音とみなされ検出されない. デプス値より, 図 14(d)での未検出の歩行者はカメラからおおよそ 65m 離れた場所を横断しており, カメラとの距離 65m 前後が検出の限度である. この距離は, 車が 80km/h で急ブレーキをかけた時の制動距離に相当し, 歩行者が飛び出する一般道での法定内速度 60km/h 以下での検出は充分可能である. このことから, 動的背景差分法による歩行者検出の有用性の高いことが示された.

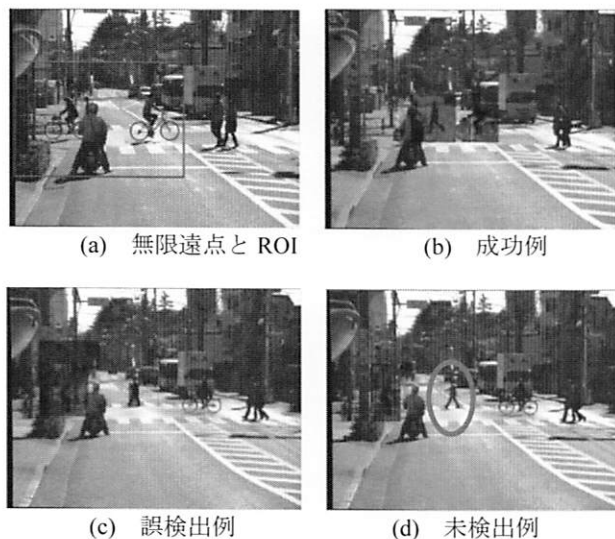


図 14 実験結果

4. むすび

本研究では, 車載カメラの画像から歩行者を検出するシステムとして, 背景部分を除去することによって歩行者を検出する動的背景差分法を提案し, 実験で有用性を示した. このシステムでは, 車載カメラの背景がある 1

点から放射状に動くことを利用して, 1 フレーム前の画像から, ステレオカメラの距離情報を用いて現在の背景を予測生成した. そして, 生成した背景と実際の現在フレーム画像を差分し, 画像の平行移動やモルフォロジー変換で振動やカーブによる背景の残差による雑音を減らし, 歩行者の残差のみを抽出した. 実験ではほとんどの画像において歩行者の検出に成功し, 提案した動的背景差分法による歩行者検出の有用性が高いことが示された.

今後の課題として, 前方を走る車の誤検出の対策が第一に挙げられる. 前方車も背景の動きに反する動きを持つため, 差分結果での残差が大きく残り誤検出を引き起こす. また, 夜間や雨天時の差分手法の確立も今後の課題である. さらに, 処理時間や背景の誤検出を軽減するなど動的背景差分法の精度を上げ, 本手法による歩行者検出システムを実用化に向けて開発していきたい.

文 献

- [1] 内閣府, “交通安全白書<平成 24 年度版>”, 第 1 編 第 1 部第 2 節, 日経印刷, pp.18, 2013.
- [2] S. Tokoro, K. Moriizumi, T. Kawasaki, T. Nagao, K. Abe and K. Fujita, “Sensor Fusion System for Pre-crash Safety System”, *IEEE Intelligent Vehicles Symposium*, pp.945-950, 2004.
- [3] T. Tsuji, H. Hattori, M. Watanabe and N. Nagaoka, “Development of Night-Vision System”, *IEEE Transactions on Intelligent Transportation Systems*, vol.3, No.3, pp.203-209, 2002.
- [4] 竹志夫, “スバル・レガシィ&新型アイサイト”, 自動車工学, pp.12-15, 2010.
- [5] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.886-893, 2005.
- [6] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, “Pedestrian detection using wavelet templates”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.193-199, 1997.
- [7] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, “Shape-based pedestrian detection”, *Proc. IEEE Intelligent Vehicles Symposium*, pp.215-220, Dearbon, USA, 2000.
- [8] K. Toyama, J. Krumm, B. Brumitt and B. Meyers, “Wallflower: Principles and practice of background maintenance”, *Proceedings of the 7th IEEE International Conference on Computer Vision*, pp.255-261, 1999.
- [9] C. Harris, M. Stephenes, “A combined corner and edge detector”, *Proceedings of the 4th Alvey Vision Conference*, pp.147-151, 1988.
- [10] B. D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision”, *Proceedings of the 1981 DARPA Imaging Understanding Workshop*, pp.121-130, 1981.
- [11] J.g. Frayer and D.C. Brown, “Lens distortion for close-range photogrammetry”, *Photogrammetric Engineering and Remote Sensing*, vol.52, pp.51-58, 1986
- [12] Z.Zhang, “A flexible new technique for camera calibration”, *IEEE transactions on Pattern Analysis and machine Intelligence*, vol.22, pp.1330-1334, 2000