

NAIST-IS-DD0661025

博士論文

自動車運転支援のための車載単眼カメラを用いた 走行環境の認識手法に関する研究

山口 晃一郎

2010年9月24日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士(工学)授与の要件として提出した博士論文である。

山口 晃一郎

審査委員：

横矢 直和 教授 (主指導教員)

加藤 博一 教授 (副指導教員)

山澤 一誠 准教授 (副指導教員)

自動車運転支援のための車載単眼カメラを用いた 走行環境の認識手法に関する研究*

山口 晃一郎

内容梗概

自動車の安全性，利便性を向上させる運転支援システムの実現のためには，自車両の周囲の走行環境を認識することが要求される．走行環境の認識のために用いられる様々なセンサの中で，画像センサは歩行者や白線など自動車の走行時に重要となる様々な対象を検出することが可能であり，人の視覚のように走行シーン全体を認識できる可能性を持っていることから，重要な役割を果たすことが期待されている．本研究では，車載単眼カメラを用いた走行環境の認識技術として，自車両の運動を推定する手法，道路面領域を推定する手法，歩行者を検出するための特徴量を抽出する手法を提案する．1章では，画像を用いた走行環境認識技術に関する従来研究を概観し，本研究の目的と研究方針を明確にする．2章では，車載単眼カメラで撮影された時系列画像を用いて自車両の運動を推定する手法について述べる．また，道路面にパターンがないシーンや他の移動物体が存在するシーンのような一般走行環境で撮影した実画像を用いた評価実験により提案手法の有効性を確認する．3章では，2章で述べた自車両の運動推定手法により得られる自車両運動，特徴点の3次元位置，道路平面パラメータを用いて，道路面領域を推定する手法を提案する．また，実走行画像を用いた道路面領域の推定結果を示す．4章では，画像から歩行者を検出するための特徴量として，濃度勾配方向に関する2次元ヒストグラムを用いる特徴量を提案する．歩行者画像データベースによる識別性能の評価実験を行い，従来手法より高い識別性能が実現できるこ

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 博士論文，NAIST-IS-DD0661025, 2010年9月24日.

とを示す。最後に 5 章では，本論文を総括し，今後の展望と課題について述べる。

キーワード

カメラパラメータ推定, 三次元復元, パターン認識, 物体検出, 運転支援システム

Recognition of Road Environments Using a Vehicle Mounted Monocular Camera for Automotive Driving Assistance*

Koichiro Yamaguchi

Abstract

In advanced driving assistance systems for safety and convenience, it is required to recognize the road environment surrounding the vehicle. By using vision sensors, important objects such as pedestrians and lanes in road environments can be detected. Therefore, it is expected that vision technologies play a vital role in advanced driving assistance systems. This thesis proposes three methods for recognizing road environments using a vehicle mounted monocular camera: (i) ego-motion estimation, (ii) road region estimation, (iii) novel feature construction for pedestrian detection. Chapter 1 gives a perspective of the study in the area of road environment recognition. Chapter 2 describes a method for estimating the ego-motion of the vehicle from a sequence of monocular images. Chapter 3 describes a method for estimating road regions in images by utilizing the method for estimating the ego-motion given in Chapter 2. Chapter 4 presents a novel feature descriptor based on two dimensional histograms of oriented gradients for pedestrian detection. Finally, Chapter 5 summarizes the present study.

Keywords:

camera parameter estimation, 3D reconstruction, pattern recognition, object detection, driving assistance system

*Doctoral Dissertation, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0661025, September 24, 2010.

目次

1. 序論	1
1.1 走行環境認識技術の概観	1
1.1.1 運動推定技術および構造推定技術の従来研究とその技術課題	4
1.1.2 歩行者検出技術の従来研究とその技術課題	10
1.2 本研究の目的	16
1.3 本論文の構成	18
2. 車載単眼カメラを用いた自車両の運動推定	19
2.1 序言	19
2.2 自車両の運動推定の概要	19
2.3 運動パラメータの推定	23
2.3.1 運動推定のための特徴点の選択	23
2.3.2 運動パラメータの計算	26
2.4 3次元構造の推定	30
2.5 移動物体の検出	32
2.5.1 アウトライアの判定	32
2.5.2 移動物体領域の検出	33
2.6 道路平面の推定とスケールの決定	35
2.6.1 道路平面パラメータの推定	35
2.6.2 スケールの推定	37
2.7 実験	37
2.7.1 移動物体のないシーン	38
2.7.2 移動物体の存在するシーン	40
2.8 結言	54
3. 車載単眼カメラを用いた道路面領域の推定	55
3.1 序言	55
3.2 道路面領域の推定法の概要	55

3.3	道路平面に関する射影変換に基づく平面領域の検出	57
3.4	3次元構造からの高さ画像の生成	61
3.5	道路面領域の推定	63
3.6	実験	66
3.7	結言	68
4.	勾配方向の2次元ヒストグラム特徴を用いた歩行者検出	72
4.1	序言	72
4.2	特徴量計算手法の処理の概要	73
4.3	勾配方向の2次元ヒストグラム特徴	75
4.3.1	勾配の計算	75
4.3.2	2次元ヒストグラムの生成	78
4.3.3	正規化と特徴ベクトルの生成	83
4.4	従来研究との関係	84
4.5	歩行者認識性能の評価実験	86
4.6	結言	95
5.	結論	96
5.1	本論文のまとめ	96
5.2	今後の検討課題と展望	97
	謝辞	100
	参考文献	101
	付録	120
	A. 8点アルゴリズム	120
	B. サポートベクターマシン	121

目 次

1	2枚の画像と道路平面の位置関係	6
2	Sliding Window による検出の枠組み	12
3	勾配方向の1次元ヒストグラムと2次元ヒストグラム	15
4	自車両の運動推定手法の処理の流れ	20
5	特徴点選択の例	25
6	2枚の画像の関係	27
7	対応点とエピポーラ線の関係	28
8	特徴点の再投影誤差	30
9	特徴点の画像座標の補正	31
10	移動物体領域の検出	34
11	道路候補点とその3次元位置	36
12	移動物体のないシーンの評価実験画像 (シーン1)	41
13	移動物体のないシーンの評価実験画像 (シーン2)	42
14	移動物体のないシーンの評価実験画像 (シーン3)	43
15	自車両の推定走行軌跡 (シーン1)	44
16	自車両の推定位置の誤差 (シーン1)	45
17	自車両の推定走行軌跡 (シーン2)	46
18	自車両の推定位置の誤差 (シーン2)	47
19	自車両の推定走行軌跡 (シーン3)	48
20	自車両の推定位置の誤差 (シーン3)	49
21	運動推定に用いた特徴点とインライア/アウトライア判定結果	50
22	移動物体の存在するシーンの評価実験画像と移動物体の検出結果	51
23	交差点での提案手法 (一般的な SfM) による推定ヨーレート	52
24	交差点での提案手法による推定ヨーレート	52
25	推定ヨー方向の誤差	53
26	道路面領域の推定の処理の流れ	56
27	道路平面上の点の射影変換	57
28	射影変換による重ね合わせ処理	59

29	三角形と画素の関係	62
30	特徴点から生成された三角網と高さ画像	64
31	道路面領域の境界の条件	65
32	道路面領域の推定結果(昼間): (a) 正解画像, (b) 射影変換による重ね合わせのみの結果, (c) 提案手法の結果	69
33	道路面領域の推定結果(夜間): (a) 正解画像, (b) 射影変換による重ね合わせのみの結果, (c) 提案手法の結果	70
34	道路面領域推定の再現率と適合率の関係	71
35	Sliding Window による歩行者検出処理の流れ	74
36	勾配方向の2次元ヒストグラム	75
37	2D HOG 特徴による歩行者検出処理の流れ	76
38	勾配と G-O ベクトル	78
39	2次元ヒストグラム生成のための周辺領域	80
40	オフセット領域の設定例	81
41	オフセット領域の重み	82
42	画素シフト時の注目画素とオフセット画素の関係	83
43	GLAC 特徴のオフセットパターン	84
44	CoHOG 特徴のオフセットパターン	85
45	Daimler データベースの歩行者, 非歩行者画像例	87
46	夜間近赤外画像データベースの歩行者, 非歩行者画像例	87
47	Daimler データベースの評価結果: 従来手法との性能比較	90
48	Daimler データベースの評価結果: オフセット領域	90
49	Daimler データベースの評価結果: 周辺領域の形状	91
50	Daimler データベースの評価結果: 半径の分割数	91
51	Daimler データベースの評価結果: 周辺領域の半径	92
52	Daimler データベースの評価結果: 周辺領域の半径とオフセット領域の方向数	92
53	Daimler データベースの評価結果: 勾配の量子化方法	93

54	Daimler データベースの評価結果: 1次元ヒストグラムと2次元ヒストグラム	93
55	近赤外画像データベースでの評価結果	94
56	SVM の識別平面とマージン	122

表 目 次

1	特徴点追跡に基づく方法の従来手法	10
2	自車両の運動推定の評価実験に用いたカメラの仕様	38
3	運動推定実験の評価シーン	39
4	道路面領域の推定の評価実験に用いたカメラの仕様	66
5	微分オペレータ	77
6	歩行者データベースの仕様	86

1. 序論

現在，自動車の安全性，利便性を向上させるために，様々な運転支援システムの開発が行われている [国土, GMM83, Jon02]. 運転支援システムには，車両，歩行者やその他の立体物との衝突を防止，軽減するための衝突防止支援システム，自車線からの逸脱に対して警報，自動操舵を行う車線逸脱防止システム，一時停止の見落としに対して警報する一時停止支援システムなどがある．自動車の運転支援のための走行環境認識技術は，移動ロボットの自律移動のための技術と非常に関連しているが，走行環境には，他の走行車両や二輪車，歩行者などの移動物体を含む様々な種類の物体が存在する上，天候や昼夜の違いによる変化などがあるため，環境が非常に多様となる点が大きな課題であり，移動ロボットの自律移動技術との大きな違いとなる．走行環境を認識する技術では，画像センサ (カメラ) やミリ波レーダ，レーザレーダがセンサとして用いられている．それらのセンサの中で，画像センサは天候や照明の変動に対する頑健性が低いことや，距離計測の点でミリ波レーダ，レーザレーダより精度が劣るといった課題があるものの，空間分解能は高く，歩行者や白線，標識，信号など自動車の走行時に重要となる様々な対象を検出することが可能であり，人の視覚のように走行シーン全体を認識できる可能性を持っていることから，運転支援システムにおいて重要な役割を果たすことが期待されている．

本章では，まず，画像センサを用いた走行環境認識の従来手法について概観し，現状の課題について取り上げる．続いて，本研究の位置づけと目的を明らかにする．

1.1 走行環境認識技術の概観

画像センサを用いた走行環境認識技術は，(i) 自車両の運動を推定する技術，(ii) 走行環境の構造を推定する技術，(iii) 走行環境における特定物体を検出する技術の3つに大別することができる．以下，分類した技術それぞれについてその概要を述べる．

自車両運動の推定

運転支援システムが認識すべき走行環境において、自車両はその中心に存在するものであり、走行中の自車両の運動を高精度に推定することができると、周囲の物体との衝突の可能性の判定や一時停止支援のための自車両の位置の推定にとって非常に有用な情報となる。また、車載カメラから得られる画像を用いた画像認識において、他の車両や歩行者などの移動物体が存在する動環境の中で、自車両の動きによりカメラ自体が大きく運動することは、物体の検出、追跡などの処理を困難にする、大きな問題の1つである。そのため、自車両運動の推定はその課題を解決するために利用できる重要な技術となる。車両には、ヨーレートセンサや車速センサ、カーナビ用GPSなど自車両運動を計測するためのセンサが一般的に搭載されているが、一時停止支援のための自車両の位置推定や画像認識のための運動推定として用いるには精度が十分でない。さらにそれらのセンサでは、車両の走行時に加減速や道路の凹凸などにより発生する上下方向の振動(ピッチ変動)を正確に計測することは難しい。

そこで、車載カメラで撮影した時系列画像を用いて自車両の運動を推定する方法について研究が行われている [山口08]。これは、連続する画像間での画像上の動きを検出することにより、カメラの運動、つまり自車両の運動を推定する方法であり、精度やロバスト性の向上を目指して様々な手法が提案されている。画像を用いた運動推定技術の利点として、物体検出などに用いられる外界監視カメラをそのまま利用できること、ピッチ変動も含めた3次元運動を推定可能であること、カメラ自体の運動を推定するため、物体追跡などの他の画像認識技術への運動情報の利用が容易であることが挙げられる。一方、画像による運動推定の課題として、前述したような周囲の移動物体の存在があり、それらの影響を受けて推定精度が劣化する可能性がある。

走行環境の構造推定

自車両の周囲の走行環境の3次元構造を推定し、すべての立体物の存在を検出できれば、他の物体との衝突を防ぐことができる。そのような任意の

立体物を検出するための方法として、ステレオカメラを用いる方法が提案されており [実吉 97, 木村 07], 実用化の例もある [Lex06, 富士 08]. 一方, 単眼カメラを用いる場合には, 3次元構造を推定する方法として特徴点の追跡による *Structure from Motion* (SfM) [HZ04] に基づく手法が多く研究されている. また, 走行中の衝突を防ぐために最低限必要な情報は, 走行可能な領域を検出することであると考える, 走行環境のすべての構造を推定するのではなく, 道路面などの平坦領域を検出する方法についても研究が行われている.

走行環境の構造推定には, ステレオ法や SfM など三角測量の原理を用いる方法や平面の射影変換を利用する方法など幾何学的なアプローチが多く提案されてきたが, 最近, 学習による統計的アプローチの研究も行われている. これらの手法では, あらかじめ多数の学習画像を用いて学習を行うことで, 1枚の画像のみを用いて, その画像を領域分割し, 各領域の面の向きの推定 [HEH07, BKY⁺08] や各領域が属する物体の種類の識別 [KYNN] を行う. さらに, 画像とともにレーザレンジファインダのデータも学習に利用することで, 1枚の画像のみから各画素の距離を推定する方法も提案されている [SSN09]. これら学習型の手法を用いることで, 様々な構造情報を推定することが可能となるが, 実環境へ適用するときには解くべき課題として, 多様なシーンへの対応と計算量の削減がある.

特定物体の検出

走行環境において意味のある, 特に重要な対象として, 車両, 歩行者などの走行環境を移動する物体と白線や道路標示, 信号, 標識などの標示物があり, 各対象を検出する手法の研究が行われている.

標示物の中で白線を検出する技術は, 前方向きの単眼カメラによる方法 [YJ97, 二宮 03, WN05] やステレオカメラを用いる方法 [BB98] など数多くの研究がこれまでに行われ, 既に車線逸脱防止システムや車線維持支援システムにおいて実用化されている [トヨ 04, 本田 02]. また, 信号 [LKK04, 木村 06], 標識 [内村 00, 高木 09], 道路標示 [LHJ07, 野田 09] の認識手法の研究も多数行われている. 標識や道路標示に関しては種類が多様であることが1つ

の課題であるが、認識対象とする標識の種類を制限速度標示のみなどに限定した場合にも支援システムとしては成立することから、既に実用化されたものも存在する [Dai08].

走行環境において周囲に存在する他の車両は衝突する可能性の高い対象であるため、画像から車両を検出する手法の研究は長く行われてきた [SBM06]. 画像から車両を検出する方法としては、左右の対称性を利用する方法 [Kue91, BBFN00] や垂直、水平エッジの特徴を利用する方法 [Sri02, OTFO03] など車両の見え方に対する知識を利用した知識ベースの方法と、主成分分析 (Principal Component Analysis: PCA) により抽出した特徴とニューラルネットを用いる方法 [MACH96] や Gabor フィルタとサポートベクターマシン (SVM) による方法 [SBM02] などの学習ベースの方法がある. 車両、特に先行車は、画像での見え方の多様性は小さいことから、これまでに開発された手法により実用化が可能な検出性能が実現されている.

一方、歩行者との衝突は重大事故につながる可能性があることから、近年、歩行者を検出する手法の研究が盛んに行われている [DWSP09, EG09]. 歩行者は、姿勢や服装、荷物の有無など見え方の多様性が大きく、部分的な隠れも多いことから、車両と比べて検出の難易度が高い対象である. 現在は学習型統計的手法に基づく方法が主流であり、様々な特徴量と識別器が提案され、検出手法は発展し続けているものの、まだ実用化に向けて十分な性能が達成できているとは言い難く、今後さらなる性能向上が要求されている技術である.

以上で述べた画像を用いた走行環境認識技術のうち、運動推定技術と幾何学的な手法に基づく構造推定技術の従来研究、および歩行者検出技術の従来研究について詳しく説明し、それぞれの技術課題を明確にする.

1.1.1 運動推定技術および構造推定技術の従来研究とその技術課題

車載カメラで撮影した画像から自車両の運動を推定する技術および構造を推定する技術は、道路面を基準とする方法と特徴点追跡に基づく方法の2つに大きく

分けることができる。道路面を基準とする方法は、道路面領域の画像上の動きから自車両の運動を推定、または画像上の動きから道路面領域を推定する方法であり、特徴点追跡に基づく方法は、画像間での特徴点の対応関係から自車両の運動の推定と3次元構造の復元を行う方法である。なお、車載カメラの画像から自車両の運動を推定するときには、カメラの運動を自車両の運動と考えるものとする。

道路面基準の方法 道路面を基準とする方法では、道路面を平面と仮定し、時系列画像間で発生する画像上での動きと、カメラ運動およびカメラと道路平面の位置関係を定式化することで、カメラ運動と道路平面の推定を行う。図1に、異なる位置で撮影された2枚の画像と道路平面の関係を示す。3次元空間において道路平面上にある点 P が2枚の画像へ投影されたとき、2枚の画像それぞれへ投影された2つの点 p_1, p_2 の画像座標は、カメラ運動と道路平面のパラメータを用いて変換することが可能である。カメラ運動のパラメータは2枚の画像間での回転成分と平行移動成分から構成され、道路平面のパラメータは基準とするカメラからの距離と平面の向き(法線方向)で構成される。したがってこの関係を用いると、道路平面上にある点に関して2枚の画像間での投影点の対応関係が分かれば、カメラ運動や道路平面のパラメータを推定することが可能である。また逆に、カメラ運動と道路平面パラメータが既知であれば、画像間での対応関係から、画像上の点が道路平面上の点の投影点であるか否かを判定することができ、すべての道路平面上の投影点を検出することで道路領域を推定することが可能となる。画像間での点の対応関係とカメラ運動、道路平面パラメータの関係は、オプティカルフローの近似式 [LHP80]、または2次元射影変換行列 (Homography 行列) [HZ04] により表すことができる。オプティカルフロー近似式は、射影変換行列を用いる場合と比べて、計算式が単純であるが、画像間でのカメラ運動が大きい場合には近似誤差が増大するという問題がある。

道路面基準の方法で自車両の運動を推定する手法として、画像上での道路領域を仮定し、その領域のオプティカルフローを用いる方法が提案されている [SK99]。しかし、一般的に道路領域にはほとんどテクスチャが存在せず、実線の白線上の点では開口問題 (Aperture Problem) があるため、道路領域では正しくオプティカルフローを求めることが難しい。そのため、道路面基準の方法では直接法 (Direct

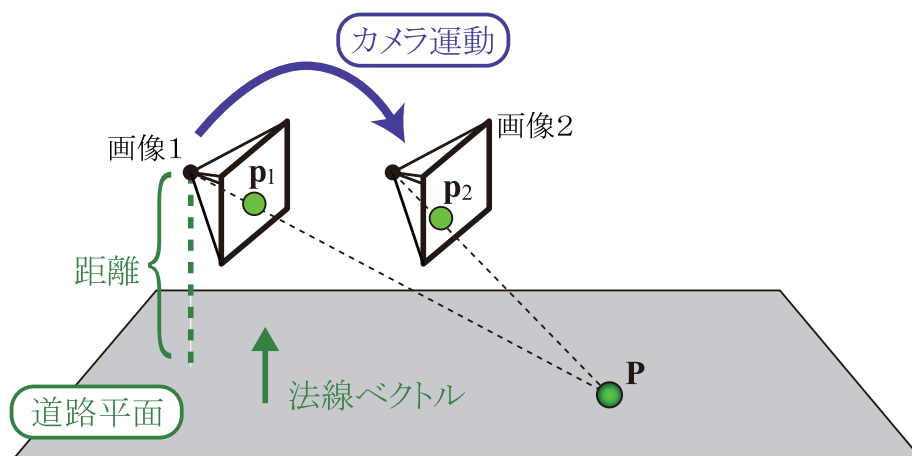


図1 2枚の画像と道路平面の位置関係

Method) [HJ88] が用いられることが多い。直接法では、推定するパラメータに関する誤差を画像の輝度から直接計算する関数を定義し、その誤差関数を最小化することで最適なパラメータを推定する。オプティカルフローなど、画像間での対応関係を直接計算する必要がないため、テクスチャのほとんどない道路シーンにおいても適用することが可能となる。単眼カメラを用いた直接法による手法 [SMS00, KK03] では、時系列単眼画像から自車両の運動の推定と道路領域の検出を行う。車載単眼カメラでは、時刻間のカメラ運動とカメラと道路平面の位置関係の両方が変化するため、直接法で最小化すべき誤差関数では、カメラ運動と道路平面の両方のパラメータを推定パラメータとして定義する必要がある。しかし、直接法では推定するパラメータの数が増えたときには安定に推定することが難しいという課題があるため、これらの手法では車両の運動パラメータの数を減らした近似運動モデルを用いることで対応している。さらに、この直接法での誤差関数の計算は画像上での道路領域を用いて行う必要があるが、走行中の画像において一般的に道路領域は既知ではない。そこで、画像を複数の局所領域に分割し、局所領域ごとに直接法で運動パラメータの推定を行う。そして各局所領域で推定された運動パラメータを、前の時刻での推定運動パラメータや他のセンサ情報と比較することにより、局所領域ごとに道路面らしさを算出することで、運動

パラメータの推定と道路面領域の検出を同時に行うことを実現している [SMS00]. また, ステレオカメラを用いる方法 [奥富 02] では, カメラ間の関係は較正済みとし, 道路平面とカメラの位置関係を表すパラメータを直接法により推定する. 較正済みのステレオカメラを用いることにより, カメラ間の位置関係は既知となるため, 直接法を使って道路平面パラメータを安定に推定することができる. そして, 得られた道路平面パラメータを用いて各時刻ごとに適切な 2 次元射影変換を行い, 道路面の領域を検出することが可能となる. さらに, 推定した道路平面パラメータを用いて道路を仮想的に真上から見た画像に変換し, 時刻間で変換した画像の道路領域をマッチングすることで自車両の運動推定を行うことができる [SO06].

これらの直接法を用いる道路面基準の方法により自車両の運動を推定することが可能であるが, 白線のような直線しかパターンがない場合や前方の車両などにより道路面領域が隠れている場合には道路面基準の方法は適用は難しい. また, 単眼カメラを用いる場合には, 直接法ではパラメータ数が多いときに安定に推定が行えない課題があるため, 運動パラメータ数を減らした近似運動モデルを使用する必要がある. しかし, 車載カメラでは削減された運動パラメータが無視できない場合も発生し, 近似運動モデルでは誤差が大きくなる可能性がある.

特徴点追跡に基づく方法 特徴点追跡に基づく方法は, 画像から検出された特徴点を時系列画像間で追跡し, 得られた画像間での特徴点の対応関係から, カメラ運動と 3 次元構造を推定する方法であり, コンピュータビジョンの分野で研究されてきた SfM や移動ロボットの分野で研究されている Visual SLAM (Simultaneous Localization and Mapping), Visual Odometry と呼ばれている技術である. 最近では移動ロボットのための技術として盛んに研究が行われており [HHH07], 屋外環境への適用も行われている. 特徴点追跡に基づく手法では, (a) 特徴点の検出, 画像間での特徴点の対応付け, (b) カメラ位置姿勢と 3 次元構造の初期化, (c) 新規画像の追加, 更新という処理を基本的に行う. 以下, 各処理について簡単に述べる.

(a) 特徴点の検出, 対応付け

画像から特徴点を検出し, 画像間で特徴点を対応付ける. この特徴点の対応関係から, カメラ運動の推定と 3次元構造の復元を行う. 特徴点検出手法としては, 安定にコーナー点を検出する手法として知られる Harris オペレータ [HS88] や Shi-Tomasi のコーナー検出法 [ST94] がよく用いられる. 高速に特徴点を検出する手法として, FAST (Features from Accelerated Segment Test) [RD06] や CenSurE (Center Surround Extrema) [AKB08] などの手法も提案されている. これらの手法により各画像において検出された特徴点を, 正規化相関などを用いて画像間で対応付けることで特徴点の対応関係を得ることができる. また, SIFT (Scale Invariant Feature Transform) [Low04] や SURF (Speeded-up Robust Features) [BETG08] といった手法では, 特徴点検出においてスケール不変性を, 特徴点の特徴量記述に回転不変性を持っており, 画像間で位置, 姿勢が大きく変化した場合においてもよりロバストに特徴点の対応付けが可能である.

(b) 初期化

単眼カメラを用いる場合には, 特徴点の対応関係からカメラの位置関係と特徴点の 3次元位置をまず求める必要がある. 2枚の画像間の特徴点对応から, カメラの位置関係を求める方法として, 8点アルゴリズム [Har97] がある. 較正済みカメラを用いるときには, 5点の対応関係から 5点アルゴリズムにより 2つのカメラの位置関係を計算可能である [Nis04]. 実画像から検出された特徴点对応を用いる場合には, 誤対応が通常含まれるため, 5点アルゴリズムまたは 8点アルゴリズムと RANSAC (Random Sample Consensus) [FB81] などのロバスト推定手法と組み合わせて, 誤対応の検出とカメラ位置関係の推定を行う方法が一般的に用いられる. また, 2つのカメラの位置関係が得られると, カメラの位置関係と特徴点の対応関係から三角測量の原理により特徴点の 3次元位置を計算できる [HS97, KSN08].

(c) 新規画像の追加, 更新

初期化により, 特徴点の 3次元位置が計算されると, 次の時刻の画像など, 新たに画像が追加されたときに, 既知の特徴点の 3次元位置と新規画像で

の特徴点位置から，新規画像が撮影されたカメラの位置，姿勢を計算できる．これは PnP (Perspective-n-Point) 問題と呼ばれ，カメラが較正済みであれば 3 点の特徴点の 3 次元位置と画像上の位置から P3P [GHTC03] によりカメラの位置，姿勢が計算可能である．この場合にも，特徴点の誤検出，誤対応が通常発生するため，RANSAC と組み合わせて用いられる．また，ステレオカメラを用いる場合には，ステレオペアから特徴点の 3 次元位置を計算できるため，2つの時刻間のステレオペアの位置関係を求める初期化処理として P3P が用いられる．

さらに，カメラの位置関係と特徴点の 3 次元位置の推定精度を向上させるための処理が加えられることもある．オフラインの SfM では，すべての画像と特徴点を用いて，再投影誤差を最小化することによりカメラの位置姿勢と特徴点の 3 次元位置を最適化するバンドル調整 (Bundle Adjustment) [TMHF99] が一般的に適用されるが，計算コストが大きいため，車載カメラで走行中に運動を逐次推定する目的のためには適用が難しい．そこで，最新の少数の画像のみを用いて局所的なバンドル調整を行う方法や，拡張カルマンフィルタ，パーティクルフィルタ [IB98]，ラオ・ブラックウェル化パーティクルフィルタ (Rao-Blackwellised Particle Filter) [DdFMR00] などの状態推定フィルタを用いて各時刻のカメラ位置姿勢と特徴点の 3 次元位置を推定する方法が用いられる．

以上のような処理を利用してカメラ運動と 3 次元構造を逐次的に推定する従来手法の例を表 1 にまとめる．それぞれ，特徴点検出，初期化，更新の各処理に前述した手法を組み合わせて運動推定と 3 次元構造の復元を行っている．これらの従来手法のうち，[KM07, Dav03] は拡張現実感 (AR) 向けの小さい環境を想定しており，[CLCG06, KAS07, MLD+06] は走行環境に近い屋外環境で適用されているものの，周囲に車両などの移動物体が存在する通常の走行環境を想定していない．

特徴点追跡に基づく手法では，道路面基準の方法のように道路領域の情報のみを利用するのではなく，画像中に存在する様々な物体上の特徴点の時刻間での対応関係を利用するため，道路にパターンがないシーンなどでもカメラ運動を推定することが可能となり，また精度を向上させるための様々な手法が提案されてい

表 1 特徴点追跡に基づく方法の従来手法

手法	カメラ	特徴点検出	初期化	更新
[NNB04]	単眼・ステレオ	Harris	5点法・P3P	—
[CLCG06]	ステレオ	簡易コーナー検出	P3P	局所的 BA
[KAS07]	ステレオ	CenSurE	P3P	局所的 BA
[MLD ⁺ 06]	単眼	Harris	5点法	局所的 BA
[KM07]	単眼	FAST	5点法	BA
[Dav03]	単眼	Shi-Tomasi	PF で推定	EKF
[SEL07]	単眼・ステレオ	SIFT	確率的表現	RBPF

局所的 BA: 局所的バンドル調整

EKF: 拡張カルマンフィルタ

PF: パーティクルフィルタ

RBPF: ラオ・ブラックウェル化パーティクルフィルタ

る。しかし、実際の走行環境へ適用する際にはまだ課題が存在する。まず、道路環境には他の車両などの移動物体が存在するため、移動物体上の特徴点对応を用いると運動を誤って推定する。また、一般的に道路領域にはテクスチャがほとんどないため、道路領域に検出される特徴点は少なく、遠方の建物などの背景に特徴点が多く検出される。そのような特徴点の分布の偏りがあると、運動の推定精度が劣化する。

以上で述べた画像を用いた自車両の運動推定と構造推定の従来技術では、道路面にほとんどテクスチャがない場合や近くに移動物体が存在するような場合という走行環境において一般的に発生する状況に対応することが難しいという大きな課題がまだ存在する。

1.1.2 歩行者検出技術の従来研究とその技術課題

国内の交通事故による死者数を状態別にみると、歩行者が全体の 3 分の 1 以上を占め、最も高い割合となっており [警察 10]、交通事故による死傷者を低減する

という運転支援システムの一つの目的において、歩行者を検出する機能は非常に重要なものである。歩行者を立体物の中の1つとして検出することも有用ではあるが、レーザレーダを用いた場合においても遠方の歩行者を検出することは難しい。また、高度な運転支援システムを実現するためには、電柱やその他の立体物と区別して歩行者を歩行者として認識できることは重要である。そのため、画像から歩行者を検出する技術の確立が望まれている。画像からの歩行者検出を困難にしている要因として次のことが挙げられる。

- 歩行者は姿勢、服装が多様であり、また荷物を持っている歩行者も存在し、その荷物の大きさ、種類も様々である。
- 近くから遠方の歩行者まで検出することが要求されるため、画像上での大きさが広い範囲にわたる。そして、遠方の歩行者では画像の大きさが非常に小さい(例えば、カメラの水平画角が40度、解像度が640×480画素の場合、60m先にいる身長1.6mの歩行者の画像上の高さは約24画素)。
- 走行環境では、都心部や郊外などのシーン、時間、天気による日照条件の違いがあり、背景も様々に変化する。また、車両や電柱、木などにより歩行者の体の一部が隠れることもある。

これらの課題に対応する方法として学習型の統計的アプローチを用いた歩行者検出手法の研究が活発に行われ、この数年で大きな進歩を遂げてきた。以下では、歩行者検出技術の従来研究について述べる。なお、走行環境に限定せず、立っている人を検出対象とする人検出手法や人を含む様々な物体を検出する一般物体検出手法についても、ここでは歩行者検出手法と同等のものとして扱う。

入力画像中の様々な大きさの歩行者を検出するための方法として、Sliding Windowの枠組みが一般的に用いられている。これは、図2に示すように、大きさを変化させたウィンドウを順に用いて入力画像を走査することで、すべての位置、大きさのウィンドウ画像を切り出し、切り出されたウィンドウごとに歩行者か否かを識別する枠組みである。Sliding Windowを用いない方法としては、特徴点を利用する方法があり [LLS08, CZ07]、Implicit Shape Model [LLS08] では、特徴点の特徴量を量子化して生成する Visual Word を用いた一般化ハフ変換から歩

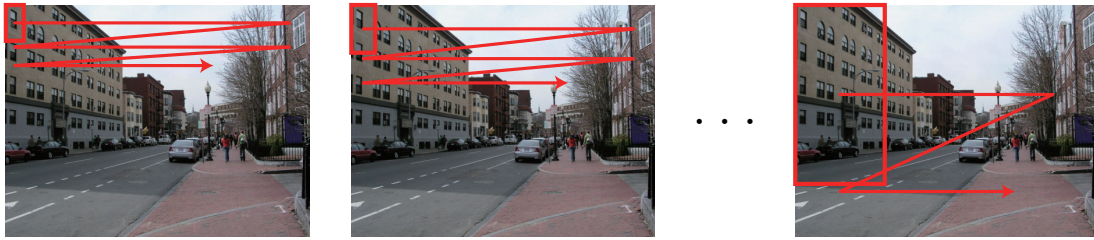


図 2 Sliding Window による検出の枠組み

行者の位置，大きさの検出を行っている．しかし，最新手法では Sliding Window の枠組みが用いられることが多い．Sliding Window の大きな問題として，すべての位置，大きさのウィンドウに対して識別処理を行うため，処理量が増大するということが挙げられる．そのため，計算量を低減するための方法として，顔検出手法として提案された Viola-Jones の方法 [VJ04] のようなカスケード構造を適用する方法 [ZYCA06, Lap09, VGVZ09, FGM10] や，Branch and Bound Search により識別処理を行う回数を減らす Efficient Subwindow Search [LBH09] が提案されている．Sliding Window の枠組みを用いる場合，切り出されたウィンドウから抽出する特徴量と特徴量から識別を行う識別器に性能は依存することになるため，特徴量抽出と識別器に関して様々な手法が提案されている．以下，それぞれ代表的な特徴量，識別器に関して述べる．

特徴量

- Haar ウェーブレット [PP00, VJ04]
ウィンドウ内の様々な位置，大きさの矩形を上下，左右，斜めに 2 分割したときの輝度の差を特徴とする．Viola-Jones の方法 [VJ04] で提案された Integral Image を使用することにより矩形の大きさに依存しない高速な計算が可能である．
- HOG (Histograms of Oriented Gradients) [DT05]
ウィンドウをセルと呼ぶ局所領域に分割してセル内の濃度勾配情報を勾配方向のヒストグラムで表し，隣接する複数のセルを合わせた領域 (ブロック)

でヒストグラムを正規化して生成する特徴量であり，局所領域でのヒストグラム化とブロックでの正規化処理により，姿勢の変化による位置ずれや照明変動による明るさの変動に頑健な特徴表現となっている．HOG 特徴は，歩行者識別性能が高い特徴として知られており，その改良手法として，複数種類のセルの分割を行って生成した勾配方向ヒストグラムを連結した特徴量を用いる方法 [MBM08] や，ベースとなる特徴量として HOG 特徴を用いて，全体モデルと部分モデルを用いる方法 [FGMR10] が提案されている．

- Bag of Features (BoF) [CZ07, VGVZ09]
SIFT 特徴量を量子化して Visual Word で表し，その Visual Word のヒストグラムを生成して特徴量とする．入力画像中に対象物体が含まれているかどうかのみを判定する Classification と呼ばれる認識技術において，広く利用されている特徴量である [LSP06]．
- Local Binary Pattern (LBP) [WHY09]
ウィンドウ内の各画素に対して，周囲の画素との大小関係から決定されるコードを割り当て，そのコードのヒストグラムを生成して特徴量とする．HOG 特徴と組み合わせて用いた場合に，エッジ情報の少ない領域を含む画像に対して認識性能を向上できるという報告がある [WHY09]．
- CSS [WMSS10]
ウィンドウを分割して生成した局所領域の色ヒストグラムの自己相似性 (Self Similarity) を特徴とする．一般的に歩行者の服装は多様であるため，色情報をそのまま特徴量とすることは難しいが，一人の歩行者では上半身，下半身の服の色，パターンはそれぞれ一様であることが多いため，肩や腕，足の色分布は左右で類似している．そこで，色ヒストグラムの自己相似性を用いて，この局所領域間の色分布の類似性を表す特徴量を生成する．
- HOF [DTS06]
1 枚の画像のAppearance情報ではなく，2 枚の時間的に連続する画像からの動き情報を利用した特徴量を生成する．車載カメラのようにカメラ自体が動いている場合には，画像間で発生する動きは歩行者自体の動きだけで

なく、カメラの動きの影響があるため、HOF 特徴では 2 つの異なる位置の局所領域内のオプティカルフローの差のベクトルを用いて、方向ごとのヒストグラムを生成する。ウィンドウ内の異なる位置のフローの差を用いることにより、部分ごとに異なる動きを示す歩行者の歩行時の手足の動きを特徴として表現することが可能である。

動き情報を用いているため、静止している歩行者の検出には向いていないが、HOG とともに用いることにより、横断する横向きの歩行者の検出性能が向上することが報告されている [WWS09]。

識別器

- SVM [Vap98]

Sliding Window の枠組みでは、大量のウィンドウ画像に対して識別処理を行う必要があるため、計算が高速で識別性能の高い線形 SVM が用いられることが多い。線形 SVM は、高次元の特徴空間で、歩行者と非歩行者を分離する超平面を学習するものであり、計算量は特徴次元数のオーダーである。HOG のようなヒストグラム特徴に対しては、 χ^2 -RBF カーネルの非線形 SVM [ZMLS07] が最も高い性能を示すことが知られているが、先に述べたように計算量の問題から実際に非線形 SVM を適用する例は少ない。ただし、例外として Histogram Intersection カーネルの非線形 SVM があり、識別時に高速計算を行う手法が提案されている [MBM08]。この高速計算法は加法性 (Additive) を持つ非線形カーネルに適用でき、これにより線形 SVM より高い識別性能の識別器を計算量を大きく増大させることなく利用できる。

- AdaBoost [FS97]

AdaBoost は、複数の弱識別器を組み合わせて識別性能の高い強識別器を構築する方法であり、最適な弱識別器を繰り返し選択、追加することで学習が行われる。通常、弱識別器は特徴と関連付けられており、AdaBoost が最適な特徴を自動的に選択する特性を持つことから、Viola-Jones の手法 [VJ04] のように、膨大な特徴空間から効果的な特徴を選択した識別器を構築すること

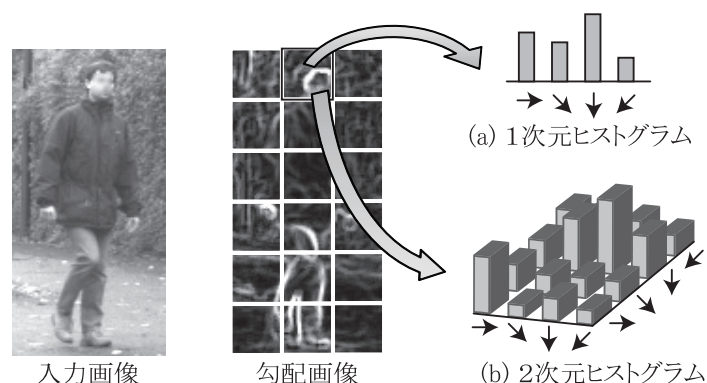


図 3 勾配方向の 1 次元ヒストグラムと 2 次元ヒストグラム

が可能である．同じ特徴量を用いた場合，AdaBoost は線形 SVM より識別性能が低いことが報告されている [WWS09] が，より識別性能の高いブースティングによる識別器として MCBost (Multiple Classifier Boosting) [KC08]，MPLBoost (Multiple Pose Learning Boosting) [BPDB08] が提案されている．これらは，複数の強識別器から成る識別器で，歩行者サンプルのクラスタリングと識別器の学習を同時に行う．計算量も各強識別器の弱識別器数の合計のオーダーである．

以上の特徴量，識別器を組み合わせる Sliding Window により歩行者検出を行う方法が多く提案されているが，単独の特徴量として最も高い性能を示しているのは濃度勾配方向のヒストグラムを用いる HOG 特徴である．これは，一般物体認識アルゴリズムの認識性能を競う PASCAL Challenge [EGW⁺] において，最もよい性能を実現している方法 [FGMR10, VGVZ09] において用いられており，また他の特徴量と組み合わせる方法 [WS08, WWS09, WHY09, WMSS10] においても基本特徴量として使われていることから明らかである．そこで，濃度勾配方向のヒストグラムを用いる新たな特徴量の提案も行われている．HOG 特徴は，図 3(a) に示すような局所領域ごとの濃度勾配方向ヒストグラムを用いているが，そのヒストグラムの要素間の共起性を計算することによって歩行者の形状情報をより高度に表現する方法がある．山内ら [山内 09] は，Real AdaBoost の

弱識別器の出力を演算子によって結合することにより特徴量間の共起性を表現した。Caoら [CYNN09] は、HOG 特徴で計算されるヒストグラムの要素間の相関値を、局所領域内および隣接する局所領域間で計算することにより、勾配方向の共起性を表す特徴量を生成した。これらは、局所領域で生成した勾配方向ヒストグラムの要素間の共起性を表す特徴量であるが、また一方、勾配の共起性をより低いレベルで表現する方法として、図 3(b) に示すように、2つの画素の勾配方向の共起性を 2次元の勾配ヒストグラムで表す方法が提案されている。Kobayashiら [KO08] は、勾配に関する位置、方向の自己相関を計算する GLAC (Gradient Local Auto-Correlation) 特徴を、Watanabeら [WIY09] は、局所領域における勾配方向の組み合わせを表す 2次元ヒストグラムを用いて生成する CoHOG (Co-occurrence Histograms of Oriented Gradients) 特徴を提案した。これらはともに HOG 特徴より高い識別性能を実現し、2つの画素の勾配方向の組み合わせを用いた 2次元ヒストグラムが歩行者識別に有効な特徴量であることが示されている。

以上で述べた歩行者検出手法の従来技術では、まだ検出性能の大きな向上が必要であるという課題がある。そのため、マルチカーネル学習 [VGVZ09] など識別器の改良やカスケード構造などの新たな識別器の構成方法の提案、そして特徴量においても複数の特徴の組み合わせ、特に動きなど新たな手掛かりとなる特徴の組み合わせが必要となる。さらに、アピランス特徴としても HOG 特徴を超える濃度勾配情報を用いた特徴や、他のパターンを用いる特徴を開発していくことが重要である。

1.2 本研究の目的

本研究では、画像を用いた走行環境認識技術として、1.1節で述べた3つの分類それぞれから、(i) 車載単眼カメラを用いた自車両の運動の推定、(ii) 道路面領域の推定、(iii) 歩行者の検出を行う技術を構築することを目的とする。使用するセンサは、車両前方に向けられた車載単眼カメラとする。これは、実用化の観点から考えたとき、単眼カメラのみで要求される機能が実現できれば、コストや搭載性の点で有利となり、様々な車種へ展開可能となるためであり、また単眼画像を用いて実現された機能は、ステレオカメラを用いた場合や他のセンサと組み合わせ

せた場合にも利用できることが期待できるためである。なお本研究では、車載単眼カメラを用いた新たな走行環境認識技術の開発を主目的とし、昼間、晴れの基本的な走行環境への適用を想定する。以下では、本研究での各技術における方針を述べる。

(i) 自車両の運動の推定

1.1.1節で述べた従来手法の課題を解決し、一般走行環境で自車両の運動を推定可能とすることを目的とする。まず、道路面にパターンがない場合や道路面に隠れがある場合にも対応するため、特徴点追跡に基づく方法を取り、基本的な計算方法に関しては従来手法に従う。そして、一般走行環境で運動推定の精度が劣化する、または正しく推定できなくなる原因であった特徴点分布の問題を解決するため、自車両の運動を推定するために用いる特徴点を選択する方法を導入する。

(ii) 道路面領域の推定

走行環境の構造の推定において、車両の走行時に最も重要な構造情報である道路面の領域を検出することを目的とする。1.1.1節で述べたように従来の道路面領域の検出手法では、自車両の運動推定を道路面基準で行っていたため、近似運動モデルとなっていた問題があったが、本手法では、(i)の自車両運動の推定法で推定された自車両運動と3次元構造情報を利用して道路面領域を検出することで、高精度に道路面領域を検出することを可能にする。

(iii) 歩行者の検出

歩行者検出技術の性能向上のため、基本となる特徴量である濃度勾配方向のヒストグラム特徴を、より識別性能の高い特徴量にすることを目的とする。そのため、勾配方向の2次元ヒストグラムを用いる特徴量を生成する。2次元ヒストグラムを計算するための画素の組の組み合わせ方を効果的に設定することで、従来手法より識別性能を向上させることができ、また特徴次元数を大きく増大させることもないため、他の特徴量との組み合わせも容易な特徴量を生成することができる。

1.3 本論文の構成

本論文の構成は次の通りである。第2章では、車載単眼カメラを用いた自車両の運動推定手法について述べる。また、実走行環境で撮影した画像を用いた運動推定の評価実験を行い、本手法の有効性を確認する。第3章では、単眼カメラから自車両の運動推定と3次元構造推定を行った後に、画像中の道路面領域を検出する手法について述べる。評価実験では、昼夜の実画像を用いて提案手法の推定結果を示す。第4章では、濃度勾配方向の2次元ヒストグラム特徴による歩行者検出について説明し、従来の2次元ヒストグラムとの違いを述べる。実験では、公開データベースを用いて従来手法との性能比較を行うとともに提案手法のパラメータ設定による性能の違いについて考察する。第5章では、本論文のまとめと今後の課題を述べる。

2. 車載単眼カメラを用いた自車両の運動推定

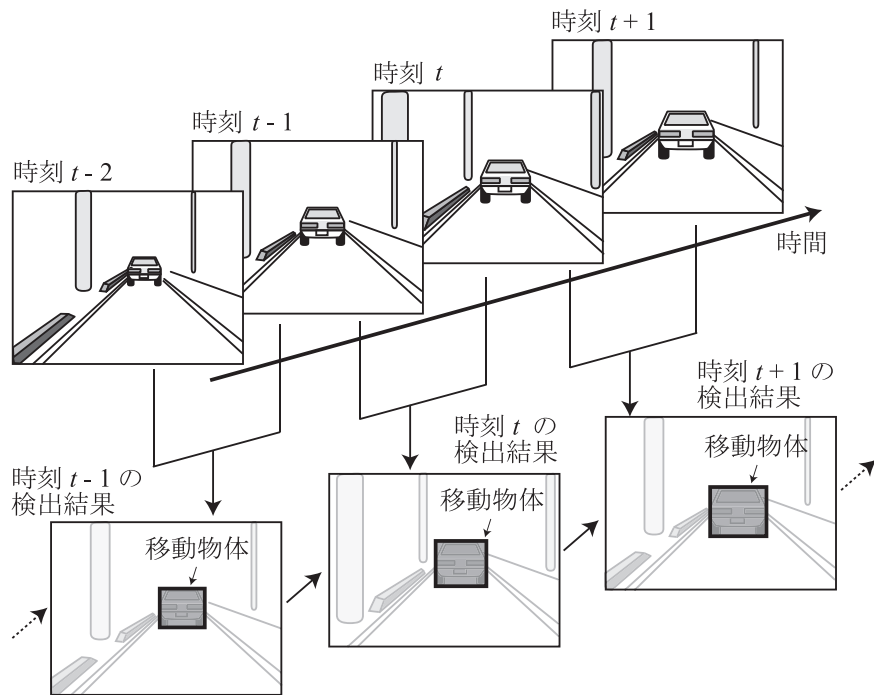
2.1 序言

画像からの車両，移動体の運動の推定は，Visual SLAM や Visual Odometry として研究が行われており [HHH07]，Structure from Motion (SfM) アルゴリズムに基づいて逐次的に処理を行う方法 [NNB04, CLCG06, KAS07, MLD⁺06] や，様々な状態推定フィルタを用いる方法 [Dav03, SEL07] が提案されている．しかし，それらの研究で想定されている環境は，非舗装路 [NNB04, KAS07] や，道路環境であっても他の車両などの存在がほとんどない走路 [CLCG06, MLD⁺06] であり，一般道路環境を対象としているものはほとんどない．

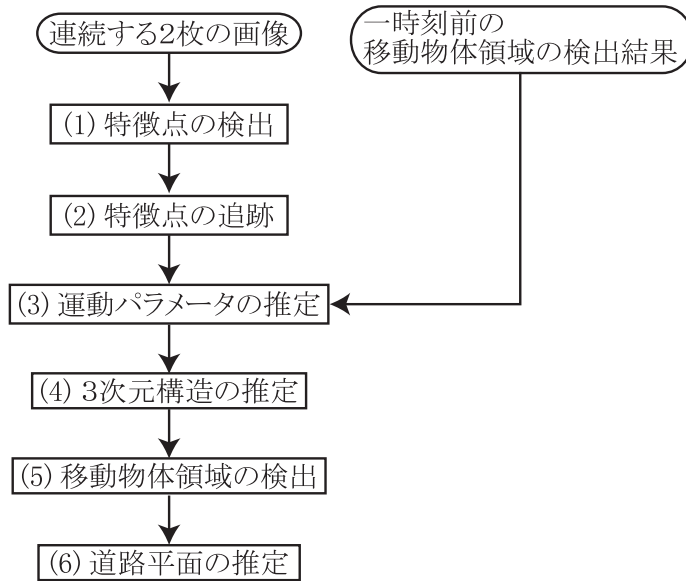
本章では，一般道路環境において単眼時系列画像を用いて自車両の運動推定を行う手法を提案する．提案手法は，特徴点追跡による SfM に基づく方法であるが，一般道路環境での道路面にパターンがほとんどないこと，および画像中に他の移動物体が存在するという問題を解決するため，運動推定に利用する特徴点を選択する処理を新たに導入する．そのため，各時刻において移動物体の検出を行い，それを次の時刻での特徴点選択に利用する．また，単眼カメラを用いる場合，平行移動量の絶対的な大きさ (スケール) は原理的に計算することができないが，提案手法では道路面の位置を推定し，車両に設置したカメラから道路面までの距離を固定と考えることにより，スケールの推定を行う．これにより，車載単眼カメラのみを用いて自車両の 3 次元運動を決定する 6 自由度すべてのパラメータを推定することが可能となる．

2.2 自車両の運動推定の概要

提案手法では時系列画像間で特徴点を連続的に追跡して，各時刻間でのカメラ運動を自車両の運動として推定する．処理の流れを図 4 に示す．なお，提案手法では車両前方前向きに設置されたカメラを使用し，カメラは較正済み，つまり内部パラメータは既知であるとする．図 4(a) に示すように，各時刻で現時刻の画像と一時刻前の画像，つまり時刻 t においては時刻 t の画像と時刻 $t-1$ の画像を用いる．さらに，時刻 $t-1$ での移動物体の検出結果も利用して，自車両の運



(a) 処理の流れの概要



(b) 各時刻での処理フロー

図 4 自車両の運動推定手法の処理の流れ

動推定を行う。なお初期時刻では、一時刻前において移動物体は検出されていないものとして処理を行う。また、特徴点は時刻間で連続的に追跡し、各時刻においてその種類を判定する。特徴点の種類は、(i) アウトライア、(ii) 道路面上の点、(iii) それ以外の点の3種類である。

各時刻における処理の流れを図4(b)に示す。入力は一時刻前の画像と現時刻の画像の2枚の画像、および一時刻前の処理において検出された移動物体領域である。また、一時刻前までの処理で検出、追跡された特徴点の情報も保持しているものとする。

(1) 特徴点の検出

入力された一時刻前の画像から Harris オペレータ [HS88] により新規特徴点を検出し、一時刻前までに追跡された特徴点群に追加する。新規特徴点の追加は次のようにして行う。

- (i) 一時刻前の画像のすべての画素について以下で定義される Harris オペレータのコーナー値 $M_c(x, y)$ を計算する。

$$M_c(x, y) = \det(\mathbf{A}(x, y)) - k \cdot \text{trace}(\mathbf{A}(x, y)), \quad (1)$$

$$\mathbf{A}(x, y) = \sum_{\mathbf{p} \in W(x, y)} \begin{bmatrix} I_x(\mathbf{p})^2 & I_x(\mathbf{p})I_y(\mathbf{p}) \\ I_x(\mathbf{p})I_y(\mathbf{p}) & I_y(\mathbf{p})^2 \end{bmatrix}. \quad (2)$$

なお、 I_x, I_y は画像の輝度 I の x 方向、 y 方向の勾配、 $W(x, y)$ は (x, y) を中心とする所定サイズのウィンドウである。また、 k はパラメータであり、通常 0.04~0.06 の値が用いられる。

- (ii) Harris のコーナー値が極大かつしきい値以上となる画素を特徴点として検出する。極大値の判定は、その画素を中心とする 5×5 のウィンドウ内でコーナー値が最大であるか否かにより行う。
- (iii) 検出された特徴点のうち、一時刻前までに追跡されている特徴点との画像上の距離が所定の距離以上である点のみを抽出し、新規特徴点として追加する。

以上のようにして検出された新規特徴点を一時刻前までに追跡された特徴点群に加えた特徴点集合を以降の処理に用いる。

(2) 特徴点の追跡

一時刻前の画像での特徴点に関して、一時刻前の画像と現時刻の画像の間のオプティカルフローを Lucas-Kanade 法 [LK81] により計算することにより、特徴点の時刻間での対応関係を求める。Lucas-Kanade 法は 2 枚の画像間で輝度パターンの差が最も小さくなる点を Gauss-Newton 法により計算する手法であり、誤差関数には SSD(Sum of Squared Differences) が用いられる。

(3) 運動パラメータの推定

特徴点の対応関係から画像間の運動パラメータを推定する。このとき、一時刻前での移動物体の検出結果を利用して運動推定に用いる特徴点を選択することで、道路環境での高精度な推定を実現する。3次元空間での運動は、3自由度の回転と3自由度の平行移動から成る6つのパラメータで表されるが、単眼時系列画像を用いる場合、平行移動のスケールを推定することは原理的にできないため、この処理の段階では平行移動のスケールを除いた5つのパラメータを推定し、平行移動ベクトルの大きさは1に正規化する。なお平行移動のスケールは、後の道路平面推定の処理で求める。

(4) 3次元構造の推定

検出された特徴点の対応関係と推定された運動パラメータを用いて、三角測量によりすべての特徴点の3次元位置を計算する。

(5) 移動物体領域の検出

まず、推定された運動パラメータから決定されるエピポーラ拘束に従わない特徴点および計算された3次元位置で負の距離を持つ特徴点をアウトライアとして判定する。そして、それらのアウトライア点から移動物体領域を検出する。

(6) 道路平面の推定

特徴点の3次元位置の分布から道路面平面を推定し、カメラと道路平面の位置関係に関する知識を利用することで平行移動のスケールを決定する。こ

のとき、推定した道路平面を利用して各特徴点が道路平面上の点かどうか判定する。

以上の処理により、連続する画像間での自車両の3次元運動パラメータをすべて推定することができる。なお、各時刻での移動物体の検出結果および特徴点の種類判定結果は次の時刻での処理に利用する。以下、運動パラメータの推定、3次元構造の推定、移動物体領域の検出、道路平面の推定の各処理について詳しく説明する。

2.3 運動パラメータの推定

運動パラメータの推定では、まず抽出された特徴点から、運動推定に利用する特徴点を選択する。次に、選択された特徴点の対応関係に基づき画像間の運動パラメータを推定する。

2.3.1 運動推定のための特徴点の選択

図 5(a) に一般道路環境を走行して撮影した画像から検出される特徴点の例を示す。このような時系列画像に対して通常の SfM アルゴリズムを適用する場合、以下の2つの大きな問題がある。

1. 道路領域にはテクスチャパターンがほとんどないため、道路に検出される特徴点は少なく、他の車両や遠方の建物などの背景には多くの特徴点が検出される。

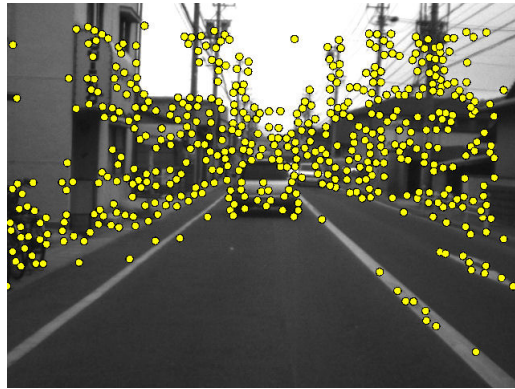
車両前方に向けられた車載カメラでは、様々な距離の物体が画像に写り、それらの距離の分布範囲は非常に広いが、近距離のものは道路領域が多く、特徴点の多い建物などは遠距離にあることが多い。回転成分と平行移動成分をそれぞれ高精度に推定するためには、様々な距離にある特徴点を用いることが望ましいが、抽出された特徴点の分布に偏りがあると運動の推定精度が劣化する。例えば、遠方にある特徴点の数が近距離の特徴点の数と比べて非常に多い場合、多数の遠方の特徴点にのみ従い、少数の近距離の特徴点に従わない運動パラメータが推定されることがある。

2. 画像中には車両などの移動する物体が含まれる。

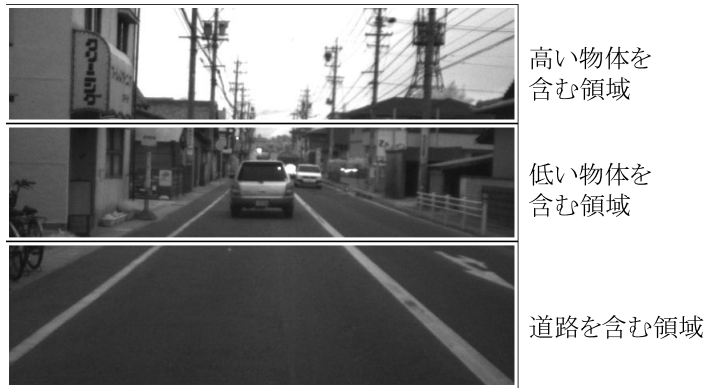
通常の SfM アルゴリズムでは、静止環境をカメラが移動することを仮定しているため、移動物体上の特徴点を含んだ特徴点对応を用いると正しい運動が推定されない。移動物体上の特徴点数が非常に少数であれば、RANSAC などのロバスト推定手法を利用して移動物体上の特徴点の影響を除去することは可能であるが、移動物体上の特徴点数が多いときや前述したもう一つの問題である特徴点の分布の偏りがあるときには正しく移動物体上の特徴点を除去できない場合が発生する。

これらの2つの課題に対応するため、提案手法では以下の方法により特徴点を選択する。

1. 特徴点の分布の偏りを低減するため、図 5(b) に示すように画像を3つの領域に分割する。車載カメラで撮影された画像をこのように分割したとき、下の領域は道路面、真ん中の領域は低い物体、上の領域は高い物体をそれぞれ多く含むと考えられ、下の領域では低い位置にある近距離の特徴点、真ん中の領域では様々な距離にある高さのある特徴点、上の領域には高い位置にある遠距離の特徴点が多く抽出される。それぞれの領域から特徴点を選択することで、3次元空間で様々な位置にある特徴点を偏りなく抽出できるため、インライアとなる特徴点の数を評価基準とする RANSAC を適用しても問題が起こりにくい。なお、領域の分割位置はカメラの設置高さおよび俯角により、あらかじめ決定するものとし、実験では下の領域は 30m 先までの道路面が含まれる領域、上と真ん中の領域は残りの領域を同じ大きさに2分割した領域とした。これは、通常走行時において、道路面が多く含まれる、できるだけ大きい領域を下の領域としたものであるが、分割位置のわずかな変化は運動推定精度には大きく影響しない。また、それぞれの領域から抽出する特徴点の数の比は、テクスチャの少ない道路面を含む下の領域では正しく追跡できる特徴点数が少ないことを考慮して、(上の領域):(真ん中の領域):(下の領域) = 2 : 2 : 1 とする。



(a) 画像から検出された全特徴点



(b) 画像の3領域の分割



(c) 運動推定用に選択された特徴点

図 5 特徴点選択の例

2. 一時刻前の移動物体の検出結果を用いて，移動物体上の特徴点を除去する．
なお，移動物体の検出法については後の節で説明する．

図 5(a) は，画像全体から Harris オペレータのコーナー関数値が大きい順に特徴点を検出した結果であり，建物などの背景部分に多数の特徴点が検出されて，分布に偏りが大きいことが分かる．さらに，移動物体である先行車両上にも多くの特徴点が検出されている．一方，図 5(c) に示す提案手法で選択された特徴点は画像中で分散しており，3次元空間においてもさまざまな位置にある点を選択されていることが確認できる．また，先行車両上の特徴点も除去されている．これは一時刻前に先行車両が移動物体領域として検出されたためである．前の時刻で移動物体が正しく検出されなかったときには，移動物体上の特徴点を除去することができないが，画像分割により特徴点を分散して選択することにより，移動物体に特徴点が集中することを防ぎ，その影響を軽減することができる．以上のようにして選択された特徴点の画像間での対応関係から運動を推定することにより，一般道路環境においても精度良く推定を行うことができる．

2.3.2 運動パラメータの計算

前節で述べた方法により選択された特徴点から，連続する画像間の運動パラメータを推定する．求める運動は，図 6 に示すような画像間の回転と平行移動から構成される．図 6 に示すように，選択された特徴点の一時刻前の画像座標と現時刻での画像座標をそれぞれ $\mathbf{m}_i^p, \mathbf{m}_i^c$ ($i = 1, \dots, n_s$) で表す． n_s は選択された特徴点の数である．なおここでは，一時刻前の画像を撮影したカメラの座標系を基準座標系として考える．このような画像の位置関係のとき，各対応点 $\mathbf{m}_i^p, \mathbf{m}_i^c$ は以下で示すエピポーラ拘束式を満たす．

$$\mathbf{m}_i^{cT} \mathbf{F} \mathbf{m}_i^p = 0. \quad (3)$$

ただし， \mathbf{F} は基礎行列 (Fundamental Matrix) であり，カメラの内部パラメータと外部パラメータ (画像間の運動パラメータ) により決定する行列である．この基礎行列を特徴点の対応関係集合 $\{\mathbf{m}_i^p, \mathbf{m}_i^c\}, i = 1, \dots, n_s$ を用いて推定する．

基礎行列は 8 点以上の対応点が与えられれば，8 点アルゴリズム [Har97] により計算することが可能であるが，実画像を用いて対応点を求めた場合には，対応

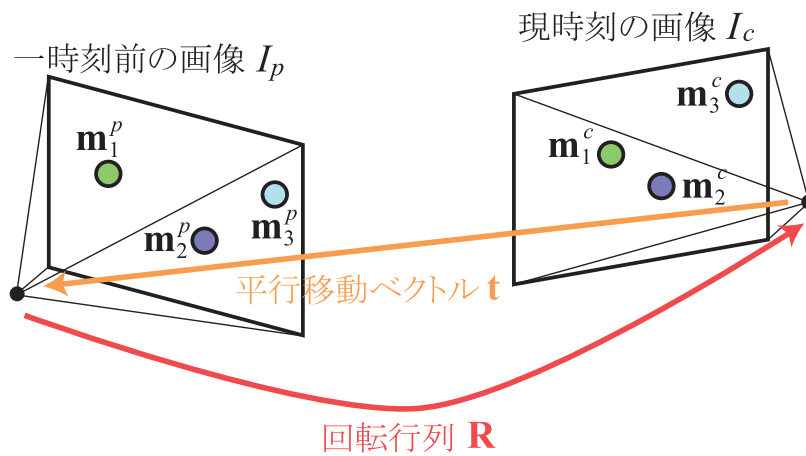


図 6 2 枚の画像の関係

点集合の中に誤対応している点が含まれることが避けられない。そこで、8 点アルゴリズムと RANSAC [FB81] により基礎行列を計算する。計算手順は以下の通りである。

(1) 特徴点の選択

特徴点集合からランダムに 8 点を選択する。この 8 点は前節で述べた分割領域の位置に関わらずランダムに選択する。

(2) 基礎行列候補の計算

選択された 8 点の対応点から 8 点アルゴリズム (付録 A 参照) により基礎行列を計算し、基礎行列候補 F_e とする。

(3) 基礎行列候補のスコア計算

計算された基礎行列候補に従う対応点数 k_e をスコアとして計算する。対応点が基礎行列候補に従うか否かはエピポーラ拘束に従うか否かであり、特徴点とエピポーラ線からの距離がしきい値以下であるかどうかにより判定する。

図 7 に示すように、対応点が m_i^p と m_i^c で表されるとき、基礎行列候補を F_e とすると、一時刻前の画像の特徴点 m_i^p に対応する現時刻の画像上での

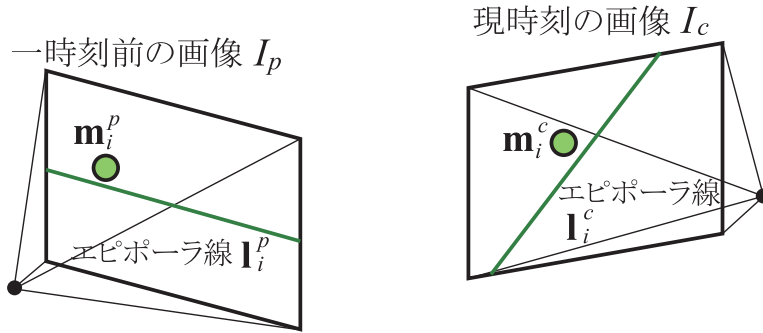


図 7 対応点とエピポーラ線の関係

エピポーラ線 l_i^c は、 $l_i^c = F_e m_i^p$ 、現時刻の画像の特徴点 m_i^c に対応する一時刻前の画像上でのエピポーラ線 l_i^p は、 $l_i^p = F_e^T m_i^c$ で計算される。したがって、一時刻前の画像での特徴点 m_i^p とエピポーラ線 l_i^p との距離、および現時刻の画像での特徴点 m_i^c とエピポーラ線 l_i^c との距離との距離がしきい値以下であるとき、対応点 m_i^p 、 m_i^c は基礎行列 F_e に従うと判定する。なお実験において、しきい値は 1 画素とした。

(4) 終了判定

基礎行列候補 F_e に従う特徴点数 k_e が、これまでに計算された基礎行列候補に従う特徴点数の最大値 k_{\max} より大きければ、基礎行列候補 F_e を推定基礎行列 F とし、 $k_{\max} = k_e$ に更新する。これまでの繰り返し数が、以下の式で計算される必要繰り返し数 r_{req} 以上であれば、推定基礎行列 F を出力として終了、そうでなければ (1) から繰り返す。

$$r_{\text{req}} = \frac{\log(1 - p_{\text{req}})}{\log(1 - (k_{\max}/n_s)^8)} \quad (4)$$

なお、 n_s は前節で述べた手法により選択された全特徴点数、 p_{req} は RANSAC による繰り返しにより、少なくとも 1 度はインライアのみ 8 点を選択された確率の要求値である。実験では、 $p_{\text{req}} = 0.99$ としている。

8 点アルゴリズムと RANSAC により基礎行列 F が推定されると、次に運動パ

ラメータを計算する。基礎行列 \mathbf{F} は、以下の式のように表される。

$$\begin{aligned}\mathbf{F} &= \mathbf{K}^{-\text{T}}[\mathbf{t}]_{\times}\mathbf{R}\mathbf{K}^{-1} \\ &= \mathbf{K}^{-\text{T}} \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \mathbf{R}\mathbf{K}^{-1},\end{aligned}\quad (5)$$

$$\mathbf{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}.\quad (6)$$

ここで、 $\mathbf{t} = (t_x, t_y, t_z)^{\text{T}}$ 、 \mathbf{R} は、図 6 に示すような画像間の平行移動ベクトル、回転行列であり、 \mathbf{K} はカメラのキャリブレーション行列を示す。キャリブレーション行列は、カメラの内部パラメータを要素として持つ行列で、 f_x, f_y は水平方向、垂直方向の焦点距離、 s は歪み係数、 (c_x, c_y) は画像中心の座標を表す。なお、本手法では $s = 0$ とする。カメラは較正済みであるため、キャリブレーション行列 \mathbf{K} は既知である。したがって、回転行列 \mathbf{R} および平行移動ベクトル \mathbf{t} は、基礎行列を分解することにより計算することができる。ただし、平行移動のスケールを決定することはできないため、ここでは平行移動ベクトル \mathbf{t} の大きさを 1 とし、後段の道路平面推定の処理でスケールを求めることとする。

運動パラメータの最適化 基礎行列 \mathbf{F} を分解して得られた運動パラメータ \mathbf{R}, \mathbf{t} を用いて、対応点の再投影誤差最小化を行うことにより推定精度を向上させる。この最適化処理は以下に述べる流れで行う。

- (1) 2.3.1 節で述べた方法により選択された特徴点 $\mathbf{m}_i^p, \mathbf{m}_i^c$ ($i = 1, \dots, n_s$) のうち、基礎行列 \mathbf{F} に従わない点をアウトライアとして除去する。なお、アウトライアを除いた特徴点群を $\{\mathbf{m}_i^p, \mathbf{m}_i^c\}, i = 1, \dots, n'_s$ で表す。 $n'_s \leq n_s$ は、アウトライアでない特徴点の数である。
- (2) 特徴点 $\{\mathbf{m}_i^p, \mathbf{m}_i^c\}, i = 1, \dots, n'_s$ の 3 次元座標 $\{\mathbf{Q}'_i\}$ を、基礎行列 \mathbf{F} を分解して得られた運動パラメータ \mathbf{R}, \mathbf{t} を用いて計算する。3 次元座標の計算方法は次節で説明する。

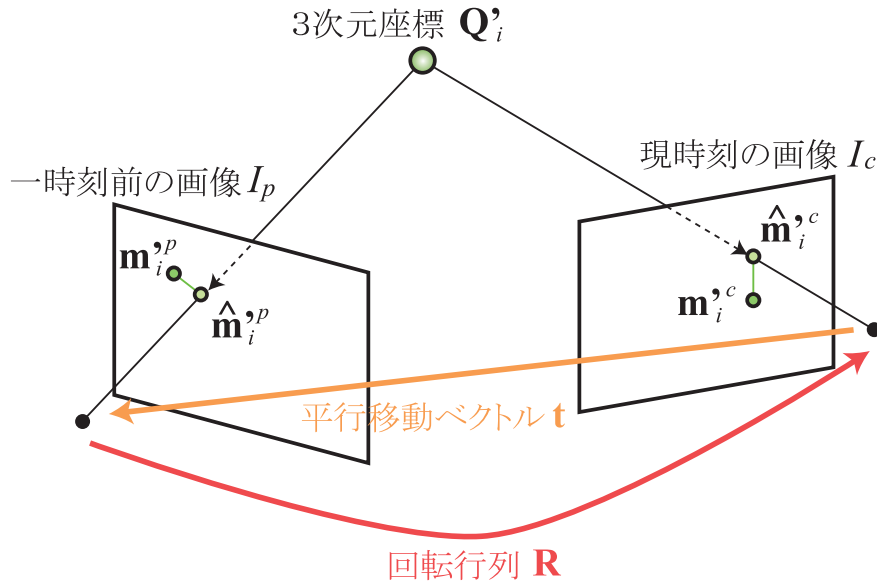


図 8 特徴点の再投影誤差

- (3) 運動パラメータ \mathbf{R}, \mathbf{t} と特徴点の 3 次元位置 $\{\mathbf{Q}'_i\}, i = 1, \dots, n'_s$ を非線形最小二乗法を用いて、再投影誤差を最小化するように最適化する．このとき運動パラメータは、3 次元の回転成分 (r_x, r_y, r_z) と 3 次元の平行移動成分 (t_x, t_y, t_z) の 6 次元パラメータとして表す．特徴点の再投影誤差は図 8 に示すように、一時刻前の画像での画像点 \mathbf{m}_i^p と 3 次元点 \mathbf{Q}'_i を一時刻前の画像へ投影した点 $\hat{\mathbf{m}}_i^p$ の距離、および現時刻の画像での画像点 \mathbf{m}_i^c と 3 次元点 \mathbf{Q}'_i を現時刻の画像へ投影した点 $\hat{\mathbf{m}}_i^c$ の距離により定義され、すべての特徴点の再投影誤差の二乗和を最小化する運動パラメータおよび特徴点の 3 次元位置を Levenberg-Marquardt 法 [Mar63] により求める．

2.4 3 次元構造の推定

前節で述べた方法により推定された画像間の運動パラメータを用いて、すべての特徴点の 3 次元位置を推定する．なおここでは、2.3.1 節で選択した特徴点群 $\{\mathbf{m}_i^p, \mathbf{m}_i^c\}, i = 1, \dots, n_s$ でなく、画像間で追跡されたすべての特徴点群 $\{\mathbf{x}_i^p, \mathbf{x}_i^c\}, i =$

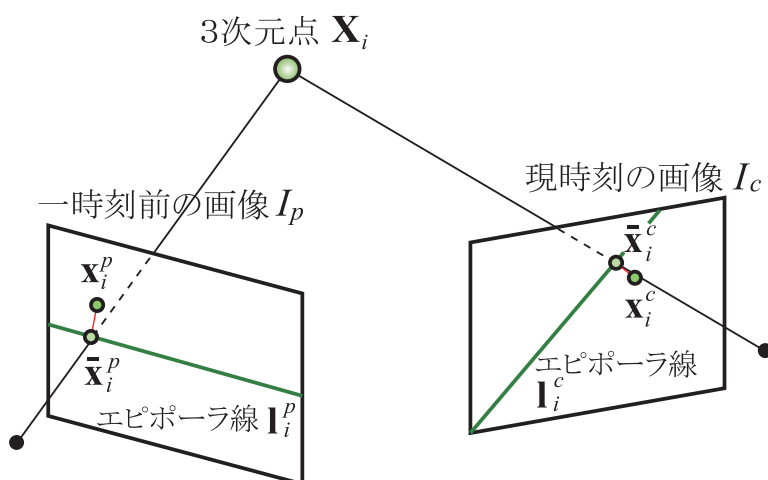


図 9 特徴点の画像座標の補正

$1, \dots, n$ を用いる. n は全特徴点数である. 各特徴点の対応関係 $\mathbf{x}_i^p, \mathbf{x}_i^c$ から, その 3次元点 \mathbf{X}_i への計算では, 特徴点の画像座標の補正を行ってから, 3次元座標を求める.

特徴点の画像座標の補正 実画像を用いて検出した特徴点の画像座標および推定された運動パラメータには誤差があるため, 2枚の画像それぞれの投影中心から特徴点へ延ばした2本の光線は3次元空間で交わらない. そこで, 2本の光線が3次元空間で交わるように特徴点の画像座標を補正する. 2本の光線が3次元空間で交わるのは, 特徴点対応が式(3)のエピポラ拘束式を満たすときであるため, 特徴点 $\mathbf{x}_i^p, \mathbf{x}_i^c$ を図9に示すような点 $\bar{\mathbf{x}}_i^p, \bar{\mathbf{x}}_i^c$ に変換する. $\bar{\mathbf{x}}_i^p, \bar{\mathbf{x}}_i^c$ は, 以下の式を最小化することにより求めることができる.

$$\|\mathbf{x}_i^p - \bar{\mathbf{x}}_i^p\|^2 + \|\mathbf{x}_i^c - \bar{\mathbf{x}}_i^c\|^2 \quad (7)$$

ただし, $\bar{\mathbf{x}}_i^p, \bar{\mathbf{x}}_i^c$ はエピポラ拘束式を満たす対応点であるため, 以下が成立する.

$$\bar{\mathbf{x}}_i^{cT} \mathbf{F} \bar{\mathbf{x}}_i^p = 0. \quad (8)$$

つまり, $\bar{\mathbf{x}}_i^p, \bar{\mathbf{x}}_i^c$ は式 (8) を条件として, 式 (7) を最小化することにより計算することができる.

3次元座標の計算 一時刻前の画像を撮影したカメラ座標系を基準座標系としたとき, 一時刻前のカメラの投影行列 \mathbf{P}_p と現時刻のカメラの投影行列 \mathbf{P}_c は次のように表すことができる.

$$\mathbf{P}_p = \mathbf{K} [\mathbf{I} | \mathbf{0}] = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (9)$$

$$\mathbf{P}_c = \mathbf{K} [\mathbf{R} | \mathbf{t}]. \quad (10)$$

このとき, 以下の式を解くことにより, 2枚の画像で対応する点 $\bar{\mathbf{x}}_i^p = (\bar{x}_i^p, \bar{y}_i^p), \bar{\mathbf{x}}_i^c = (\bar{x}_i^c, \bar{y}_i^c)$ から, その特徴点の3次元点 \mathbf{X}_i を求めることができる.

$$\mathbf{A} \mathbf{X}_i = 0, \quad (11)$$

$$\mathbf{A} = \begin{bmatrix} \bar{x}_i^p \mathbf{p}_p^{3\top} - \mathbf{p}_p^{1\top} \\ \bar{y}_i^p \mathbf{p}_p^{3\top} - \mathbf{p}_p^{2\top} \\ \bar{x}_i^c \mathbf{p}_c^{3\top} - \mathbf{p}_c^{1\top} \\ \bar{y}_i^c \mathbf{p}_c^{3\top} - \mathbf{p}_c^{2\top} \end{bmatrix}. \quad (12)$$

なお, $\mathbf{p}_p^{i\top}, \mathbf{p}_c^{i\top}$ はそれぞれ投影行列 $\mathbf{P}_p, \mathbf{P}_c$ の第 i 行目の行ベクトルを表す.

以上の計算をすべての特徴点の対応関係 $\{\mathbf{x}_i^p, \mathbf{x}_i^c\}$ を用いて行うことにより, 特徴点の3次元点 $\{\mathbf{X}_i\}, i = 1, \dots, n$ を計算する.

2.5 移動物体の検出

2.5.1 アウトライアの判定

運動パラメータを推定し, その推定結果からすべての特徴点の3次元位置を計算した後, アウトライアの特徴点を検出する. アウトライアの特徴点は, 2.3.2節で推定した基礎行列 \mathbf{F} に対してエピポーラ拘束を満たさない点, およびその3次元点で奥行き方向の距離が負の値である点とする. エピポーラ拘束を満たすか否

かの判定は、前述したように対応点とエピポーラ線の距離がしきい値以下であるか否かにより判定する。このようにして検出されたアウトライア特徴点は、移動物体上の特徴点であるか画像間で誤対応した点であるかのどちらかであるが、誤対応点は運動パラメータの推定誤差や特徴点の対応誤差などによりある時刻でのみアウトライアとなり、移動物体上の特徴点は連続的にアウトライアとなる傾向が強い。そこで、時間的に連続して追跡している特徴点の各時刻でのアウトライア判定結果を保持し、3時刻以上アウトライアと判定された特徴点を抽出して移動物体の候補点とする。なお、判定基準は実験により経験的に3時刻以上に設定した。図10(a)に、抽出された移動物体の候補点とそのオプティカルフローの例を示す。

2.5.2 移動物体領域の検出

前節の処理により移動物体の候補点と判定された特徴点を、その画像上での位置およびオプティカルフローの大きさ、向きを基準としてグループ化し、図10(b)に示すように同じグループに属するすべての特徴点の外接矩形を移動物体領域として検出する。また、2点以上のグループとならなかった移動物体の候補特徴点は、誤対応として特徴点群から除去し、以降のフレームでは追跡しない。

この移動物体検出処理では、2点以上の特徴点がグループ化される必要があるため、画像中で小さい移動物体領域が検出されない場合や、グループ化された特徴点の外接矩形を移動物体領域とするため、移動物体の正確な領域とならない場合が発生する。しかし、運動推定に大きく影響する画像上で大きい領域を占める移動物体は検出されるため、2.3.1節で述べた特徴点選択手法により移動物体の影響を軽減し、運動推定の高精度化を行うことができる。

しかし、運動推定が誤っている場合には移動物体も正しく検出されず、それが次の時刻の推定にも影響する問題が発生する。そこで、多数の特徴点がアウトライアとなり、移動物体の領域が非常に大きくなった場合には、運動推定が誤っていると判断し、その移動物体領域を次の時刻の処理に利用しない。大きさの判定基準には、10m先の車両の大きさを用いる。



(a) 移動物体の候補点のオプティカルフロー



(b) 検出された移動物体領域

図 10 移動物体領域の検出

2.6 道路平面の推定とスケールの決定

計算された特徴点の3次元位置を用いて、3次元道路平面のパラメータを推定する。道路平面パラメータが推定されると、次に道路平面の位置から平行移動ベクトルのスケールを決定する。

2.6.1 道路平面パラメータの推定

まず、すべての特徴点から道路候補点を抽出する。道路候補点は、2.5.1節で検出されたアウトライア点以外の特徴点のうち、あらかじめ設定した道路候補領域内にある特徴点および前の時刻において道路面上にあると判定された特徴点である。道路候補領域はあらかじめ手動で設定するものとし、実験では2.3.1節で述べた3分割された下の領域の左右50画素を除いた領域とした(図11(a)の矩形領域)。次に抽出された道路候補点の3次元位置から道路平面パラメータを推定する。3次元道路平面は、以下の式で表される。

$$ax + by + cz = d \quad (13)$$

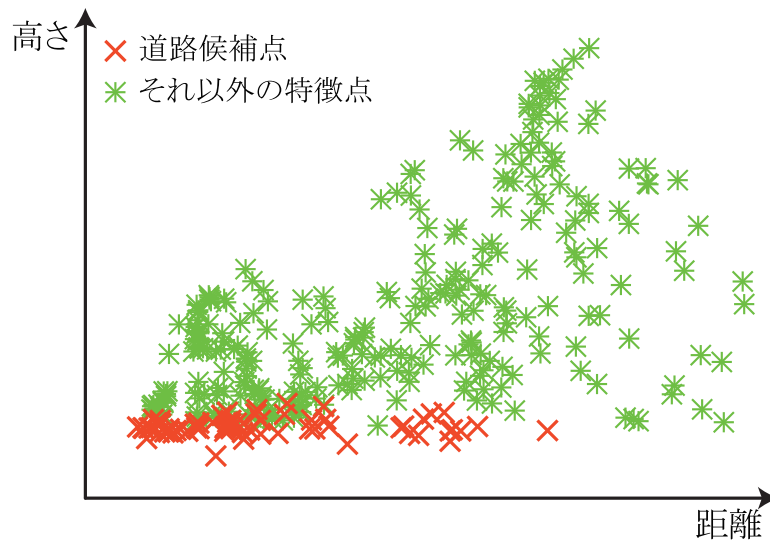
ここで、 $\mathbf{n} = (a, b, c)^T$ は道路平面の法線ベクトル、 d は道路平面とカメラ中心の距離であり、 $\|\mathbf{n}\| = 1$ であるとする。図11に抽出された道路候補点とその3次元位置の分布の例を示す。

次に、道路候補点に平面を当てはめることで道路平面パラメータを推定する。ただし、図11(a),(b)に示すように、道路候補点の中には、実際には道路上の点でない点や誤対応により誤った3次元位置を持つ点が含まれるため、LMedS (Least Median of Squares) を用いて、道路候補点に対して平面を当てはめ、道路平面パラメータを求める。LMedSによる道路平面パラメータの推定は以下のように行う。

- (1) 道路候補点からランダムに3点を選択する。
- (2) 選択された3点の3次元位置から道路平面パラメータ a, b, c, d を計算する。
このとき、車載カメラに対する道路平面の向きの範囲には制限を設けることができる。たとえばカメラに対してほぼ垂直な道路平面はあり得ない。そ



(a) 道路候補点



(b) 3次元構造を真横から見た図

図 11 道路候補点とその3次元位置

ここで、道路平面の法線ベクトル $\mathbf{n} = (a, b, c)$ が許容範囲外なら 1. へ戻り、3 点の選択からやり直す。

- (3) すべての道路候補点と求めたパラメータの道路平面との距離を計算し、その中央値を求める。
- (4) (1) から (3) を繰り返し、最も中央値が小さくなった道路平面パラメータ a, b, c, d を推定値とする。

2.6.2 スケールの推定

時系列画像から推定した画像間の平行移動成分にはスケールの不定性があるため、これまでの処理では平行移動ベクトルの大きさを 1 とし、特徴点の 3 次元位置を計算し、道路平面の推定を行った。本処理では、カメラの高さ、つまり道路平面からの距離を一定と仮定することにより推定した道路平面からスケール成分を求める。車両が静止した状態で計測したカメラの道路からの高さが d_{true} であるとき、道路平面の推定で得られた式 (13) の d を用いて、平行移動成分の大きさを d_{true}/d とする。本手法では、このようにスケールの推定を行うことで、画像のみから 3 次元運動のすべてのパラメータを推定することができる。

スケールを決定した後、アウトライアでないすべての特徴点に対して道路平面との距離を計算し、その距離がしきい値以下である特徴点を道路平面上の点として判定する。実験では、道路平面との距離のしきい値は 0.3m とした。このしきい値は、提案手法を適用したときの推定道路平面と道路面上にある特徴点の推定 3 次元位置の距離の分布から経験的に決定した。これらの道路平面上の点は、次の時刻において道路候補点として抽出され、道路平面の推定処理に用いられる。

2.7 実験

車載したカメラで撮影した道路環境の実走行画像に提案手法を適用し、車両の運動を推定する実験を行った。実験で用いた走行画像は (株) 豊田中央研究所が撮影したものであり、カメラの仕様は表 2 に示す通りである。なお、カメラはルームミラー横に前向きに設置した。

表 2 自車両の運動推定の評価実験に用いたカメラの仕様

カメラ	Flea (Point Grey Research Inc.)
画角	32.6×24.8 度
解像度	640×480 画素
フレームレート	30 フレーム / 秒
カメラ設置高さ (d_{true})	1.28m
俯角	下向き 1.8 度

実験では、画像から抽出された特徴点をすべて用いて運動の推定を行う一般的な SfM アルゴリズムを比較手法とした。一方、提案手法では特徴点選択法により 100 点の特徴点を選択し、運動推定に利用した。選択する特徴点の数は、様々な走行実験画像の道路面領域において安定に追跡できる特徴点の平均的な数から決定した。これは、選択する特徴点数を多くした場合には、道路面領域で誤対応する特徴点が増える、または道路面領域から抽出できる特徴点が他の領域の特徴点より少なくなるためである。特徴点選択と移動物体検出以外の処理は、比較手法と提案手法で同じである。使用した計算機は、CPU Core2 Duo 3.0GHz、メモリ 4GB であり、計算時間は 1 フレームあたり 0.15 秒であった。

2.7.1 移動物体のないシーン

最初の実験として、移動物体がほとんど存在しない 3 つのシーンで運動推定の評価を行った。評価シーンの時間、走行距離を表 3 に、画像を図 12, 13, 14 に示す。評価シーンはすべて昼間、晴であり、テクスチャのない道路領域を多く含むシーンである。

比較手法である一般的な SfM アルゴリズムおよび提案手法により得られた自車両の走行軌跡を図 15, 17, 19 に示す。自車両の走行軌跡は、各時刻で推定された運動を累積することにより計算したものである。また、走行軌跡の正解値と

表 3 運動推定実験の評価シーン

	時間	走行距離	車速
シーン 1	20 秒	200m	約 40km/h
シーン 2	24 秒	230m	約 40km/h
シーン 3	18 秒	130m	約 30km/h

して RTK-GPS (Real Time Kinematic GPS) の測位値を利用し、推定値と比較した。評価シーンでは、RTK-GPS の測位は常に FIX 解が得られた。使用した RTK-GPS は Trimble 社 Trimble5700 で、FIX 解の測位精度は 10mm である。なお、図 15, 17, 19 では、車両の初期位置および車両の初期移動方向は RTK-GPS と一致するように手動で合わせている。図 15, 17, 19 に示すように、提案手法で推定された走行軌跡は正解値に非常に近く、一般的な SfM による推定軌跡はややずれが生じている。図 16, 18, 20 に推定位置の誤差を示す。いずれの場合も提案手法は、一般的な SfM を適用した比較手法より誤差が小さいことが確認できる。図 16, 18 に示すように、シーン 1, シーン 2 では走行距離が増えても、提案手法による推定位置の誤差は大きく増大しないが、図 20 に示すように、シーン 3 では走行距離が増えるにつれ、提案手法の誤差は比較手法と同じように増大する結果となった。これは、図 14 に示すように、シーン 3 では 1 車線のみ狭い道路である上に駐車車両があるため、道路上に十分な数の特徴点が抽出できず、道路面の推定に誤差が生じ、自車両運動のスケール成分が誤ったことが原因と考えられる。図 19 で示す提案手法による推定走行軌跡は正解値と比べて方向が大きくずれていないことから、誤差の原因がスケールの推定にあることが分かる。なお、3 シーンでの 100m 走行時の位置誤差の平均は、比較手法が 2.6m、提案手法が 1.2m であった。

図 21 に、シーン 1 において比較手法と提案手法それぞれで運動推定に利用した特徴点を示す。図 21 で、白い点は RANSAC でインライアと判定された特徴点、黒い点はアウトライアと判定された特徴点である。図 21(a) に示す比較手法で用いられた特徴点は、建物部分に多く集中しており、画像下部の道路面上の特徴点

はアウトライアと判定されている。これは、RANSACがインライアとなる特徴点数のみを評価基準としているため、少数の道路領域の特徴点对応には従わなくても建物部分の多数の特徴点对応に従う運動パラメータが採用されてしまった結果である。一方、図 21(b) ではほとんどの特徴点がインライアと判定されており、特定の領域にアウトライアが集中していないことから、運動を正しく推定できると考えられる。

以上の移動物体のないシーンでの結果から、特徴点選択法のうち、画像から分散して特徴点を選択することが道路環境での運動推定に有効であることが確認できる。また走行軌跡の推定結果より、提案手法では画像のみから平行移動のスケールを含むすべての運動パラメータの推定精度を向上できることが分かる。

2.7.2 移動物体の存在するシーン

次に、交差点でのシーンを用いて評価を行った。図 22に評価に用いたシーンの画像を示す。これは、自車両は交差点を右折し、画像に写る先行車両は左に動いた後、直進するシーンである。自車両の速度は約 20km/h であった。

提案手法の移動物体検出の結果を、図 22の矩形で示している。先行車両が正しく移動物体として検出されていることが分かる。図 23, 24に、比較手法と提案手法それぞれで推定した画像間のヨーレートを示す。縦軸はヨーレート (度/秒)、横軸はフレーム番号である。推定したヨーレートはジャイロセンサの値と比較した。評価シーンにおいて、先行車両は 70 フレームから 170 フレームの間、左へ移動し、自車両は 170 フレームで右折し始める。図 23に示す比較手法では、移動する先行車両の影響で 70 フレームから 170 フレームの間、ジャイロと大きく異なる誤推定が頻繁に発生しているが、図 23に示す提案手法では、その間でも誤りの発生頻度が低い。図 25に 70 フレームからのヨー方向の累積誤差を示す。比較手法は、移動車両の影響で角度誤差が大きく増大しているが、提案手法ではその影響が軽減されていることが確認できる。この結果から、交差点のような前方に移動する車両が存在するシーンにおいて、移動物体を検出し、移動物体領域内の特徴点を運動推定に利用しない選択方法が有効であると言える。



frame 70



frame 170



frame 270



frame 370



frame 470



frame 570

図 12 移動物体のないシーンの評価実験画像 (シーン 1)



frame 0



frame 145



frame 290



frame 435



frame 580



frame 720

図 13 移動物体のないシーンの評価実験画像 (シーン 2)



frame 90



frame 180



frame 270



frame 360



frame 450



frame 540

図 14 移動物体のないシーンの評価実験画像 (シーン 3)

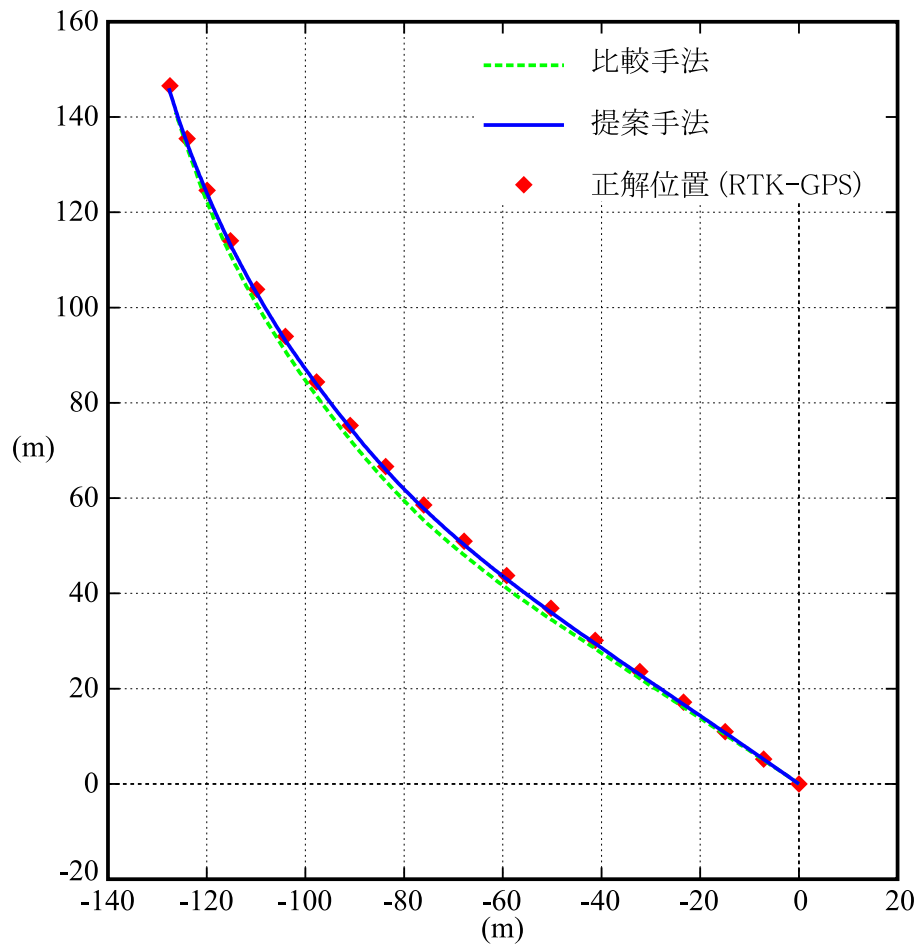


図 15 自車両の推定走行軌跡 (シーン 1)

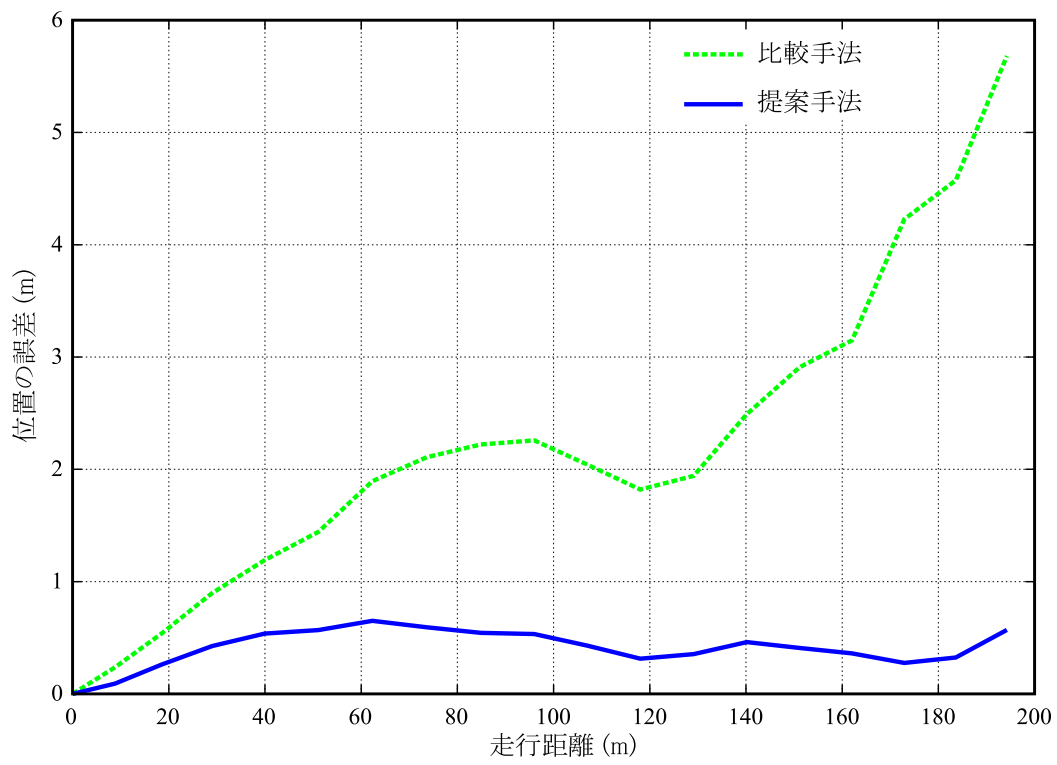


図 16 自車両の推定位置の誤差 (シーン 1)

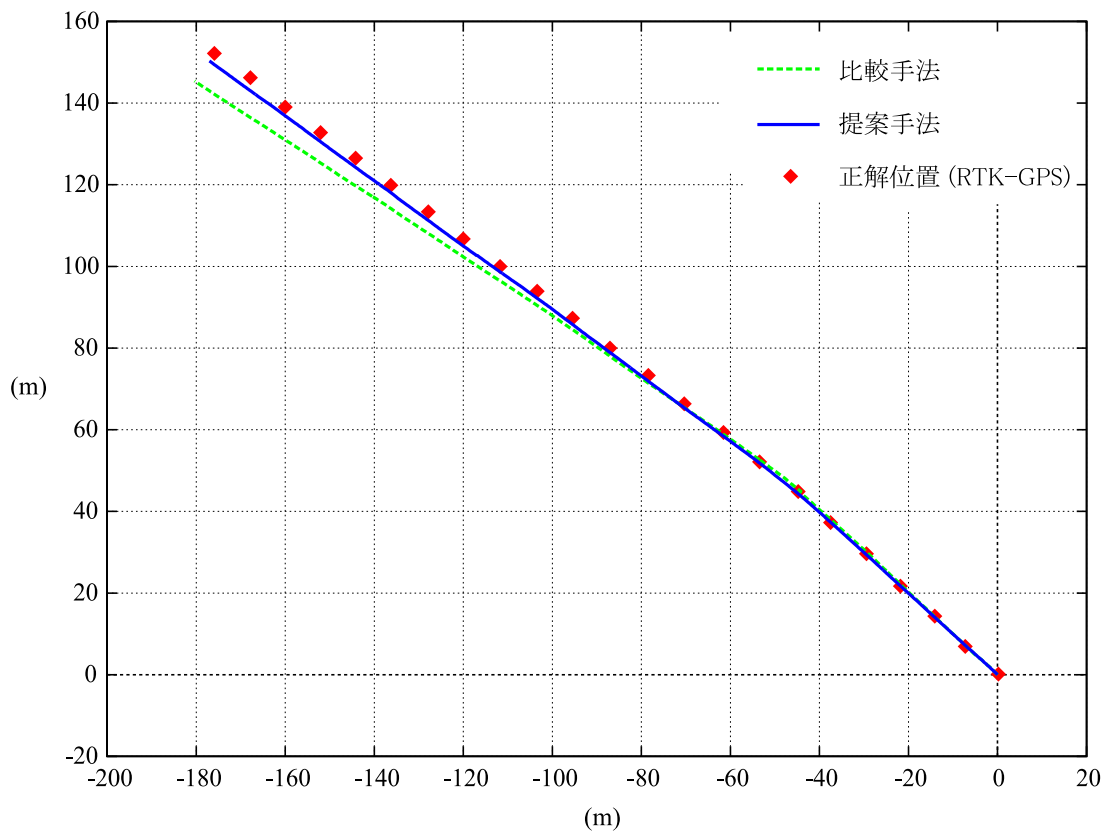


図 17 自車両の推定走行軌跡 (シーン 2)

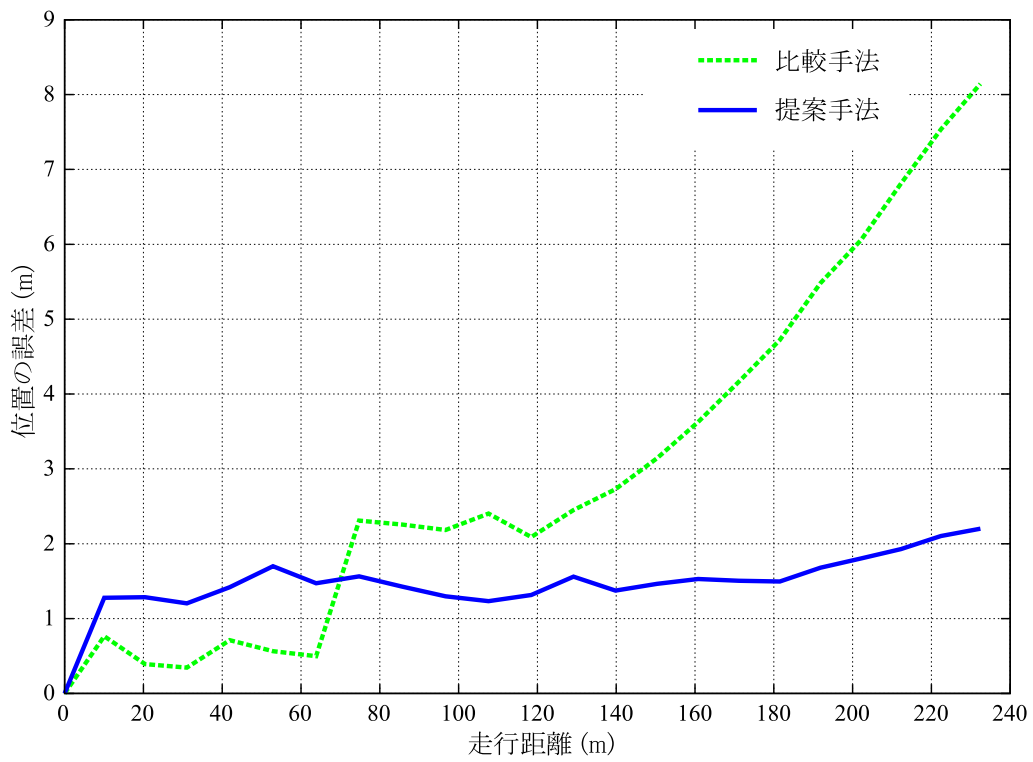


図 18 自車両の推定位置の誤差 (シーン 2)

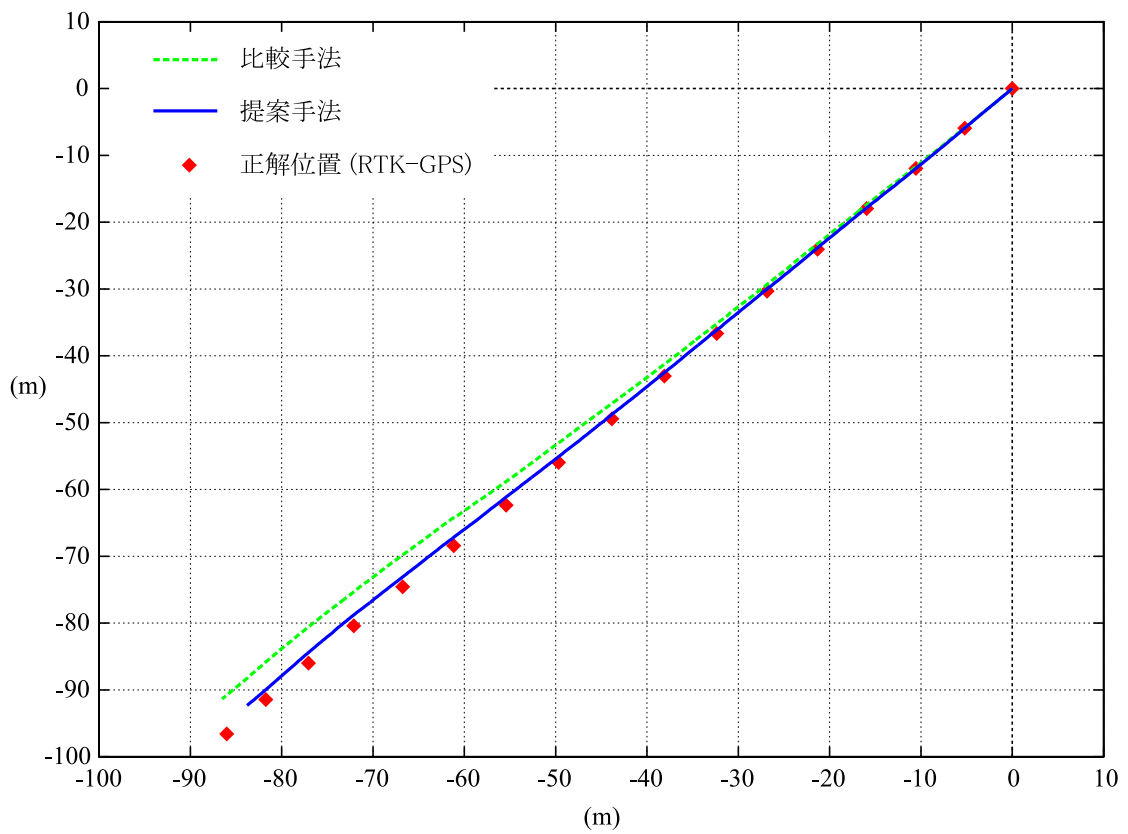


図 19 自車両の推定走行軌跡 (シーン 3)

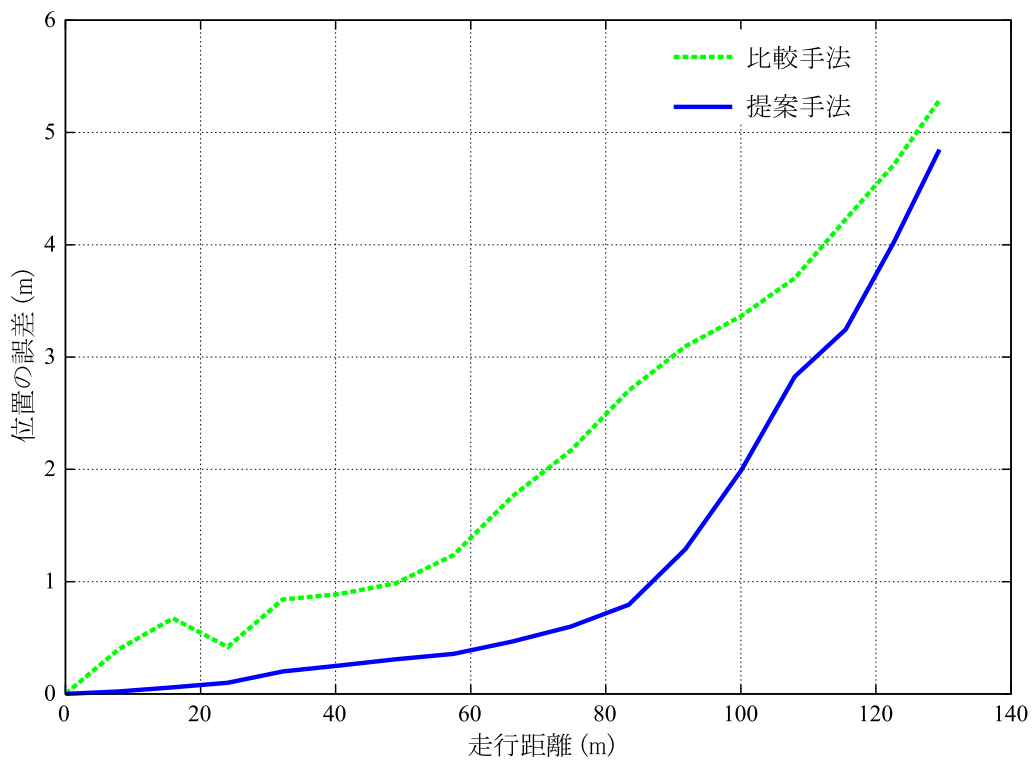


図 20 自車両の推定位置の誤差 (シーン 3)

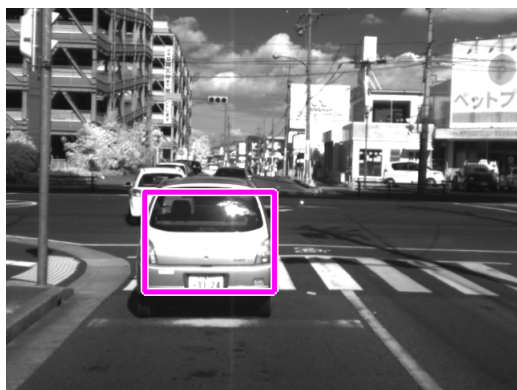


(a) すべての特徴点を使用した場合

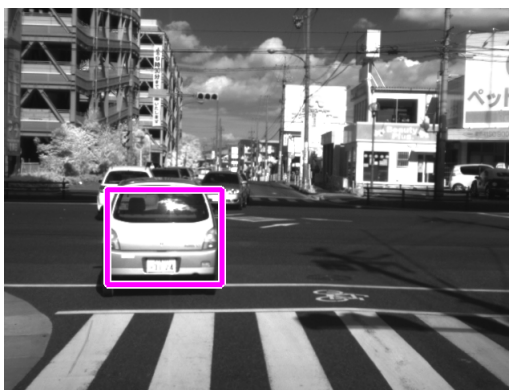


(b) 提案手法により特徴点選択を行った場合

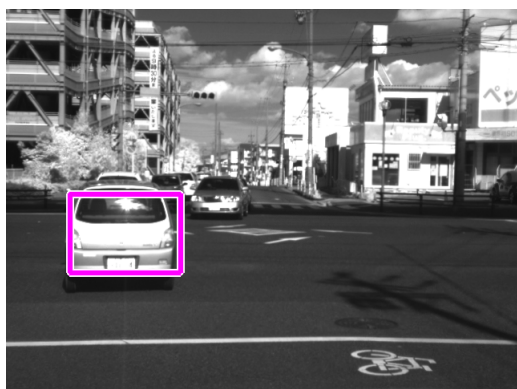
図 21 運動推定に用いた特徴点とインライア/アウトライア判定結果



frame 80



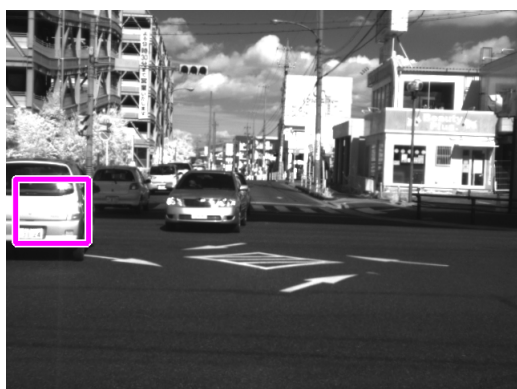
frame 115



frame 135



frame 160



frame 175



frame 200

図 22 移動物体の存在するシーンの評価実験画像と移動物体の検出結果

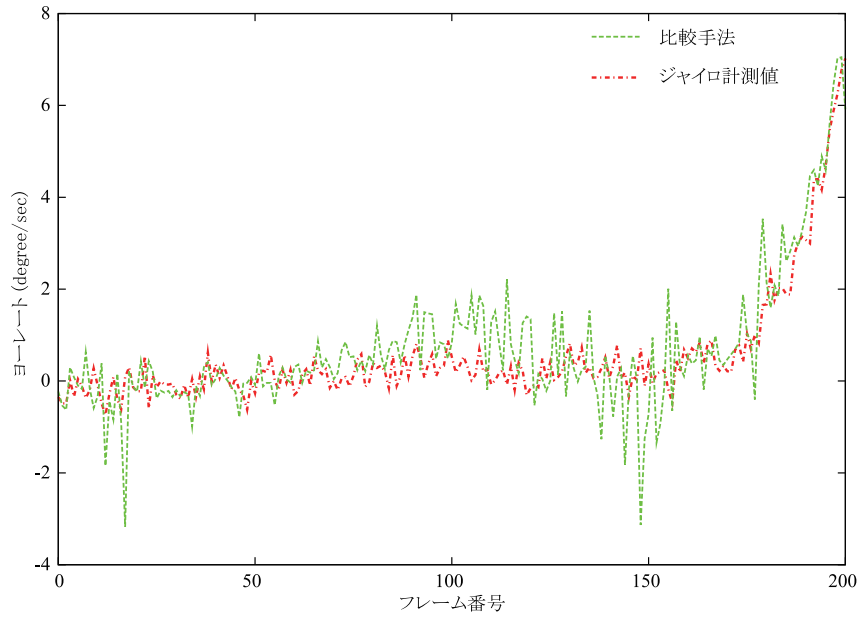


図 23 交差点での提案手法 (一般的な SfM) による推定ヨーレート

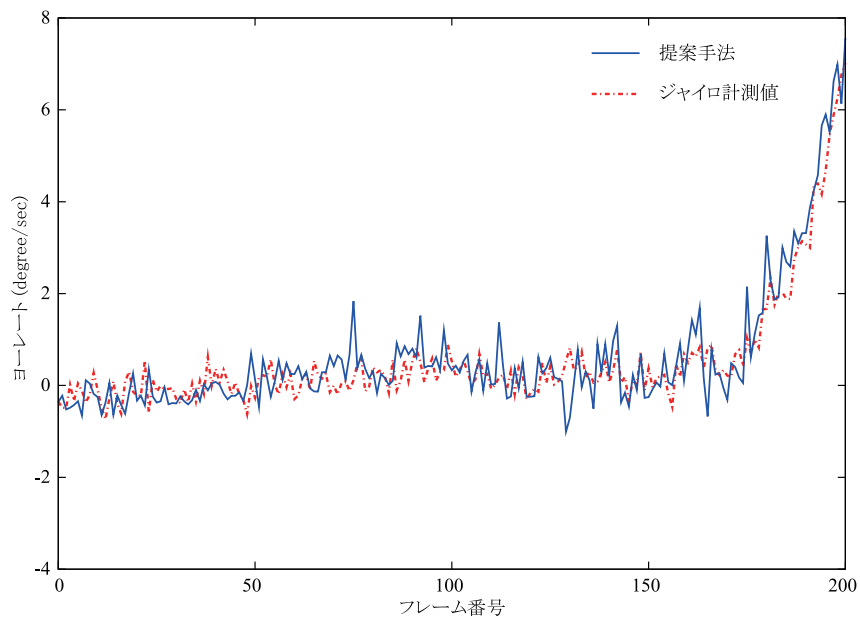


図 24 交差点での提案手法による推定ヨーレート

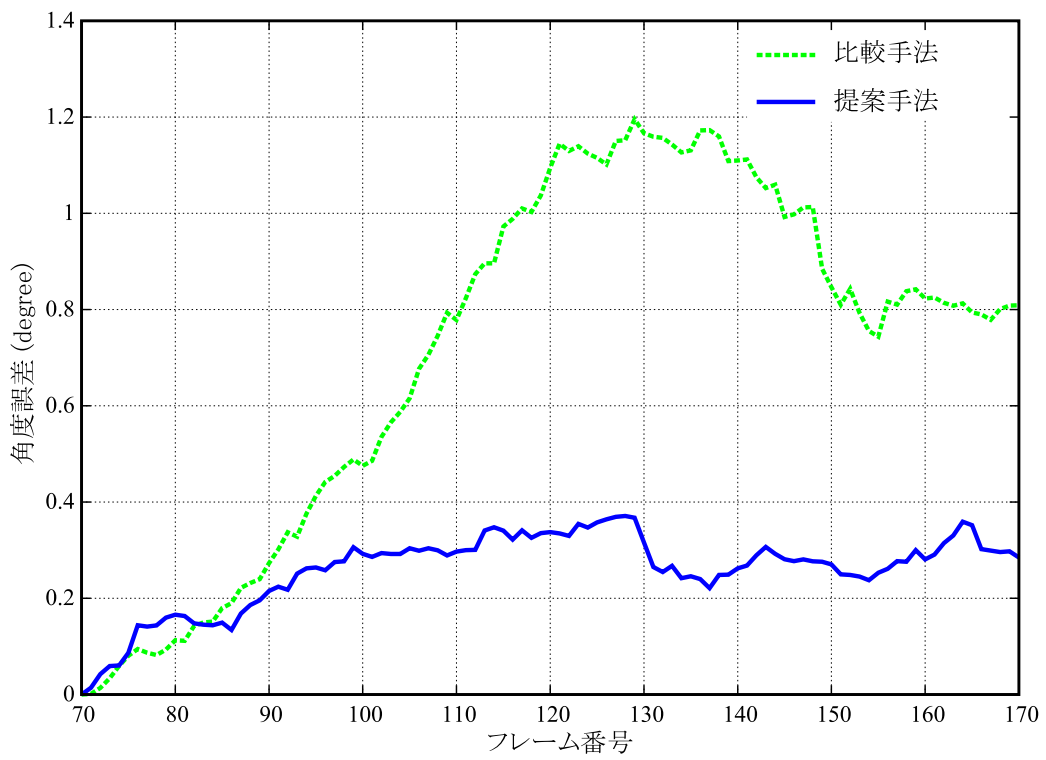


図 25 推定ヨー方向の誤差

2.8 結言

本章では，他の走行車両や歩行者等が存在する一般走行環境において車載単眼カメラで撮影した時系列画像から自車両の運動を推定する手法を提案した．提案手法は，SfM アルゴリズムに基づく方法であり，道路環境での課題に対応するため，移動物体を検出し，特徴点の分布を考慮して運動の推定に利用する特徴点の選択を行う．また，復元した 3 次元構造から道路平面を推定することにより，画像情報のみから平行移動のスケールを求め，6 自由度の 3 次元運動パラメータを決定することができる．実走行画像を用いた実験において，提案手法が自車両の運動を高精度に推定できることを確認した．

提案手法は，100m 程度の走行においては累積誤差が位置 1m，角度 0.3° 程度であり，高い精度で走行中の相対的な自車両位置を推定可能である．一時停止支援システムには，1m 以下の精度で自車両の絶対位置を推定することが必要とされているが [三谷 06]，数十 m の精度で絶対位置を測位できる一般的な GPS と組み合わせる方法 [小島 07] によりその要求を満たすシステムを実現できると考えられる．また雨天時，夜間への対応は課題であり，IMU (Inertial Measurement Unit) のような運動を計測する他のセンサと組み合わせたロバスト性向上手法などが必要と考えられる．また，衝突防止支援システムのためには，提案手法により推定される自車両の運動，3 次元構造を利用して道路上の障害物を検出する技術を開発することも必要である．

3. 車載単眼カメラを用いた道路面領域の推定

3.1 序言

走行環境において道路面領域は、自車両が走行可能な領域であり、その領域が検出できると走行可能領域と立体物領域を区別することができるため、運転支援システムにおいて有効な機能の1つとなる。単眼カメラを用いる従来手法 [SMS00, KK03] では、道路面領域の検出を自車両の運動の推定と同時に行っており、自車両の運動の推定に 1.1.1 節で述べた道路面基準の方法をとっている。この場合、自車両の運動パラメータ数を減らした近似運動モデルを用いる必要があるため、高精度に自車両の運動を推定することが難しい。また、道路面にパターンがほとんどない場合には精度が劣化する可能性があり、これらの運動推定に対する問題がそのまま道路面領域の推定に影響する。

本章では、前章で述べた自車両の運動推定手法を利用して、道路面領域を検出する手法を提案する。前章の手法により、高い精度で自車両運動が推定され、道路平面に関するパラメータも得られていることから、これらを用いて連続する2枚の画像間での道路平面に関する射影変換を求め、射影変換による変形と重ね合わせにより平面領域の検出を行う。さらにテクスチャのない領域で発生する平面領域の誤検出を除去するため、前章で述べた方法で推定された3次元構造から高さ画像を生成する。最終的に、射影変換に基づく平面領域の検出結果と3次元構造から生成された高さ画像を用いて道路面領域の推定を行う。

3.2 道路面領域の推定法の概要

提案手法では、車載単眼カメラから得られる画像を用いてカメラ運動と3次元道路平面のパラメータを推定し、推定したパラメータから計算される2次元射影変換行列と、3次元構造情報から生成される高さ画像を利用して道路面領域を検出する。なお、道路面は平面であると仮定とする。

図 26 に処理の流れを示す。まず、前章の手法を用いて、画像間のカメラ運動を推定し、3次元構造を復元する。そして、復元した3次元構造を利用して道路平面パラメータを推定する。カメラ運動と道路平面に関するパラメータが推定され

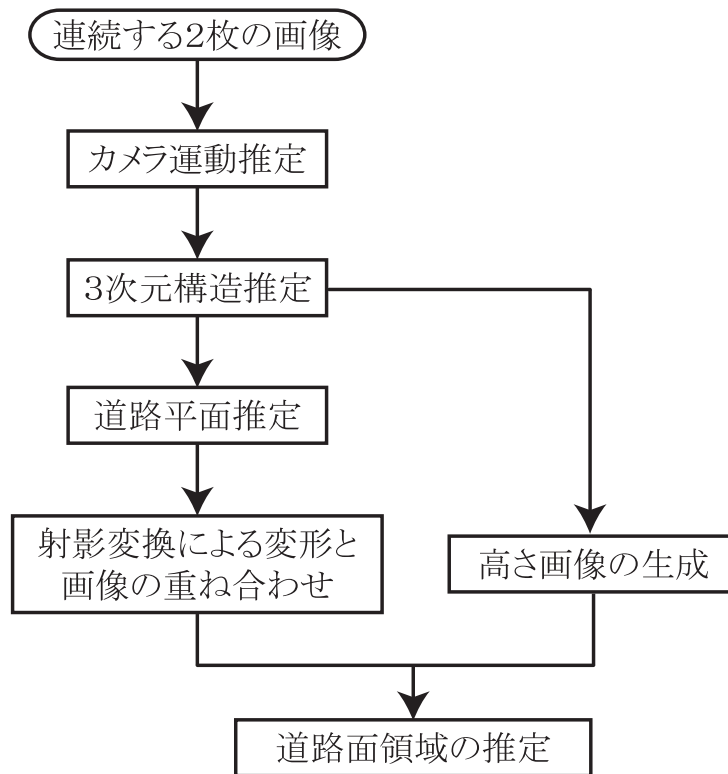


図 26 道路面領域の推定の処理の流れ

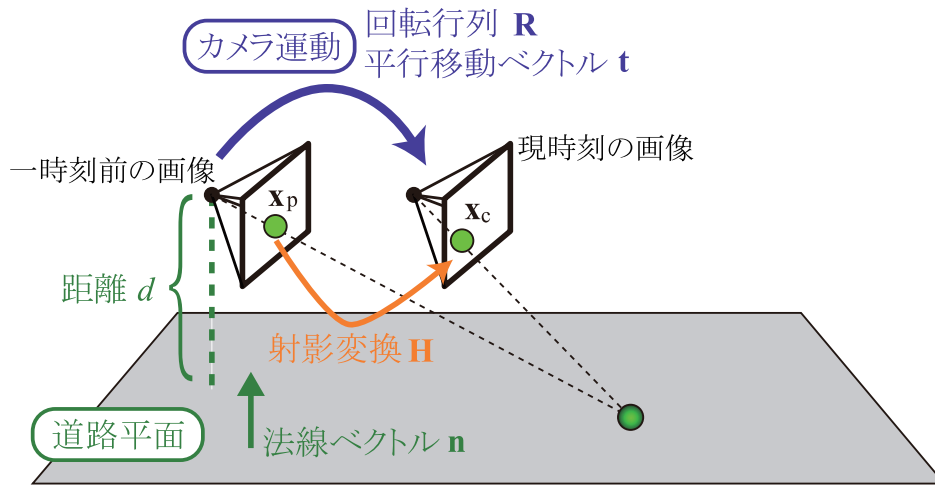


図 27 道路平面上の点の射影変換

ると、画像の重ね合わせ処理を行う。これは、2枚の画像間の道路平面に対する2次元射影変換行列を計算して一時刻前の画像を射影変換し、現時刻の画像との差分を計算することにより行う。道路平面に属する領域では差分値が小さくなることから、道路面領域が検出できる。しかし、テクスチャのない様な領域では、差分を計算する2枚の画像間でずれが生じていたとしても差分値が大きくならないため、道路面領域か否かを正しく判定することができない。そこで、復元した3次元構造を用いて、各画素において道路面からの高さ値を保持する高さ画像を生成する。さらに、車載カメラで走行中に撮影した画像を用いていることから、道路面領域の境界は1本の線で表すことができ立体物領域と道路面領域はその線の上下に二分されるという条件を加えた上で、射影変換画像との差分値と道路面からの高さ値の両方を用いて道路面領域を推定する。以上の処理により、一般道路環境において道路面領域を安定に検出することができる。

3.3 道路平面に関する射影変換に基づく平面領域の検出

図 27に示すように、道路平面上の点が2枚の画像へ投影される時、その2つの投影点は2次元射影変換行列 H により変換することができる。2枚の画像間でのカメラの回転行列を R 、平行移動ベクトルを t 、道路平面の法線ベクトルを

$\mathbf{n}, \|\mathbf{n}\| = 1$, カメラと道路平面の距離を d とすると, 画像間での道路平面に関する 2次元射影変換行列 \mathbf{H} は以下のように与えられる [FL88].

$$\mathbf{H} = \mathbf{K} \left(\mathbf{R} + t \frac{\mathbf{n}^T}{d} \right) \mathbf{K}^{-1}. \quad (14)$$

ただし, \mathbf{K} はカメラの内部パラメータから構成されるキャリブレーション行列である. そして, 道路平面上のある点が 2枚の画像に投影されるとすると, それぞれの画像へ投影された点は以下に示す 2次元射影変換により関係付けられる.

$$\mathbf{x}_c \simeq \mathbf{H}\mathbf{x}_p. \quad (15)$$

なお, \mathbf{x}_p は一時刻前の画像における投影点の同次座標ベクトル, \mathbf{x}_c は現時刻の画像における投影点の同次座標ベクトルであり, \simeq は定数倍の不定性を許して等しいことを示す.

射影変換による重ね合わせ処理の概要を図 28 に示す. 処理の手順は以下の通りである.

- (1) 前章で述べた自車両の運動の推定手法によりカメラ運動パラメータ \mathbf{R}, t と道路平面パラメータ \mathbf{n}, d を推定する.
- (2) 式 (14) で計算される 2次元射影変換行列 \mathbf{H} により一時刻前の画像 I_p を射影変換し, 射影変換画像 I'_p を生成する.
- (3) 射影変換画像と現時刻の画像の差分画像を生成する.

なお, 差分値として各画素を中心としたウィンドウでの SAD (Sum of Absolute Differences) を用いる. ウィンドウの大きさを $(2 \times w + 1) \times (2 \times h + 1)$ 画素としたとき, 差分画像の座標 (x, y) の画素値 $D(x, y)$ は以下のようにして計算する (実験では, $w = 5, h = 5$ の 11×11 のウィンドウを用いた).

$$D(x, y) = \sum_{u=-w}^w \sum_{v=-h}^h |I_c(x+u, y+v) - I'_p(x+u, y+v)| \quad (16)$$

なお, $I_c(x, y)$ は現時刻の画像の座標 (x, y) の画素値, $I'_p(x, y)$ は射影変換した一時刻前の画像の座標 (x, y) の画素値を表す.



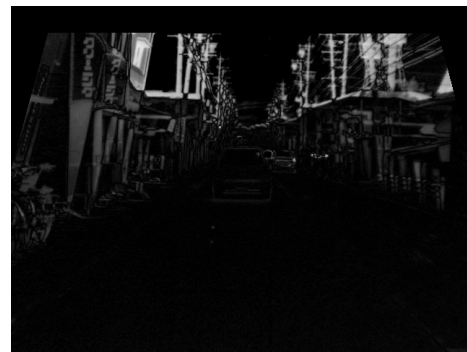
一時刻前の画像



射影変換画像



現時刻の画像



差分画像



図 28 射影変換による重ね合わせ処理

(4) 差分画像のしきい値処理により平面領域を検出する.

一時刻前の画像を射影変換した画像は、道路平面に関する2次元射影変換行列により変形して生成しているため、道路平面の領域では変換された一時刻前の画像と現時刻の画像が一致して差分値が小さくなり、それ以外の領域では変換画像と現時刻の画像でずれが生じて差分値が大きくなる。そこで、差分値がしきい値以下の画素を抽出することで画像の道路面領域を検出することができる。

しきい値の決定 以上で述べたように差分画像のしきい値処理で道路面を検出するためには適当なしきい値を設定することが必要となる。しかし、様々な道路環境で道路面領域を検出する場合には、あらかじめしきい値を決定しておくことは難しい。例えば、昼夜での明るさの違いによるS/N比の変化や道路領域のテクスチャの違いがあったときには、道路面領域での差分値が異なってくる可能性がある。そこで、各時刻での差分値に対するしきい値 θ を以下のように設定する。

$$\theta = \theta_0 \times (2 \times w + 1) \times (2 \times h + 1), \quad (17)$$

$$\theta_0 = \min(k \times D_{\text{ave}}, \theta_{\text{max}}). \quad (18)$$

θ_0 は、1画素あたりの許容差分値を意味しており、 D_{ave} は道路面候補領域内の画素の差分値の平均値、 θ_{max} は1画素あたりの最大許容差分値である。 θ_{max} は、あらかじめ設定された値である(実験では $\theta_{\text{max}} = 15$)。また、 k は許容差分値を制御するパラメータであり、 k を大きくすると、射影変換の推定誤差などによる微少なずれに影響を受けず、道路面領域と判定されやすくなるが、立体物領域を道路面と誤判定することも増える。後述する実験では、基本パラメータとして $k = 1$ を用いているが、道路面領域推定の再現率と適合率の関係を評価する際には k を変化させている。道路面候補領域は、推定した3次元道路平面の水平線以下の領域としている。以上のように、しきい値を各時刻で動的に変更することにより、様々な環境において適当なしきい値処理で道路面領域を推定することができる。

3.4 3次元構造からの高さ画像の生成

前節で説明した射影変換による重ね合わせ処理により道路平面領域の推定が可能となるが、テクスチャのない様な領域では射影変換画像とのずれが生じていたとしても差分値が大きくなるため、道路面領域か否かを正しく判定することができない。本研究では、射影変換行列を精度良く求めるために時刻間のカメラ運動を推定し、3次元構造を復元しているため、その復元した3次元構造を道路面領域の推定にも利用する。

前段までの処理で得られている3次元構造情報は、特徴点群の3次元位置であり、非常に粗な情報であるため、特徴点群の3次元位置情報を利用して各画素の高さを求め、高さ画像を生成して、道路面領域の推定に利用できるようにする。各画素の高さの計算は、以下の手順により行う。

(1) Delaunay 三角網の生成

現時刻の画像において、すべての特徴点 $\{\mathbf{x}_i, i = 1, \dots, n\}$ から、それらの特徴点を頂点とする Delaunay 三角網を生成する。生成された三角形集合を $\{T_m, m = 1, \dots, M\}$ 、各三角形 T_m の頂点を $\{\mathbf{v}_1^m, \mathbf{v}_2^m, \mathbf{v}_3^m\}$ で表す。ただし、 $\mathbf{v}_{\{1,2,3\}}^m \in \{\mathbf{x}_i\}$ 。

(2) 三角形の平面パラメータの推定

Delaunay 三角網によりシーンが多面体表現できるとし、各三角形 T_m の平面パラメータ $\{\mathbf{g}_m = (\alpha_m, \beta_m, \gamma_m)^T, \delta_m\}$ を、3つの頂点 $\mathbf{v}_1^m, \mathbf{v}_2^m, \mathbf{v}_3^m$ から求める。なお、平面パラメータは以下の式に従う。

$$\alpha_m X + \beta_m Y + \gamma_m Z = \delta_m, \quad \|\mathbf{g}_m\| = \sqrt{\alpha_m^2 + \beta_m^2 + \gamma_m^2} = 1 \quad (19)$$

(3) 各三角形の内部にある画素の高さの計算

三角形 T_m の内部にある画素に対して、道路面からの高さを計算する。ただし、この処理は三角形 T_m を構成する3頂点 $\mathbf{v}_1^m, \mathbf{v}_2^m, \mathbf{v}_3^m$ に対応する特徴点がすべてインライアの場合にのみ行い、1点でもアウトライアの特徴点が含まれる場合には、その三角形の内部の画素の道路面から高さは計算しないものとする。インライア・アウトライアの判定は、2.5.1節で述べた方法で

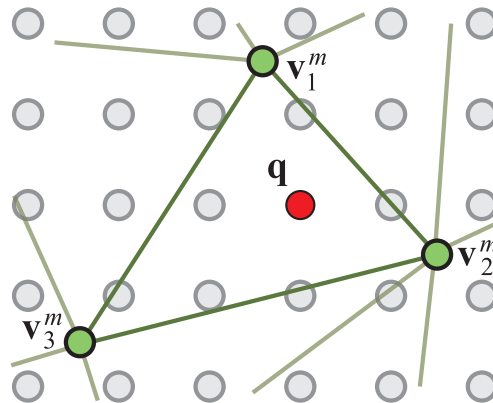


図 29 三角形と画素の関係

行う。3 頂点に対応する特徴点がすべてインライアの場合に三角形の内部の画素に対する高さの計算は次のようにして行う。

(a) 三角形の内部判定

図 29 に示すように、三角形の 3 頂点が時計回りの順に $\mathbf{v}_1^m, \mathbf{v}_2^m, \mathbf{v}_3^m$ で表されるとき、三角形の内部に含まれる画素 \mathbf{q} は以下の式を満たす。

$$(\mathbf{q} - \mathbf{v}_1^m) \times (\mathbf{v}_2^m - \mathbf{v}_1^m) > 0 \quad (20)$$

$$(\mathbf{q} - \mathbf{v}_2^m) \times (\mathbf{v}_3^m - \mathbf{v}_2^m) > 0 \quad (21)$$

$$(\mathbf{q} - \mathbf{v}_3^m) \times (\mathbf{v}_1^m - \mathbf{v}_3^m) > 0 \quad (22)$$

ただし、 \times はベクトルの外積を表す。

(b) 画素の 3 次元座標の計算

画素 \mathbf{q} が三角形 T_m の内部に含まれるなら、画素 $\mathbf{q} = (q_x, q_y)$ の 3 次元座標 $\mathbf{Q} = (Q_x, Q_y, Q_z)$ を計算する。3 次元座標 \mathbf{Q} は、画素 \mathbf{q} に対応する光線と三角形 T_m の平面との交点として、以下のように計算される。

$$\mathbf{Q} = t\mathbf{q}, \quad (23)$$

$$\bar{\mathbf{q}} = \mathbf{K}^{-1} \begin{bmatrix} q_x \\ q_y \\ 1 \end{bmatrix}, \quad (24)$$

$$t = \frac{-\delta_m}{\mathbf{g}_m \cdot \bar{\mathbf{q}}} \quad (25)$$

なお、 \mathbf{K} はカメラのキャリブレーション行列である。

(c) 画素の道路面からの高さの計算

道路平面が式 (13) に示すように $aX + bY + cZ = d$ で表され、画素 \mathbf{q} の 3 次元座標が $\mathbf{Q} = (Q_X, Q_Y, Q_Z)$ であるとき、画素 \mathbf{q} の道路面からの高さ $h_{\mathbf{q}}$ は以下のように計算される。

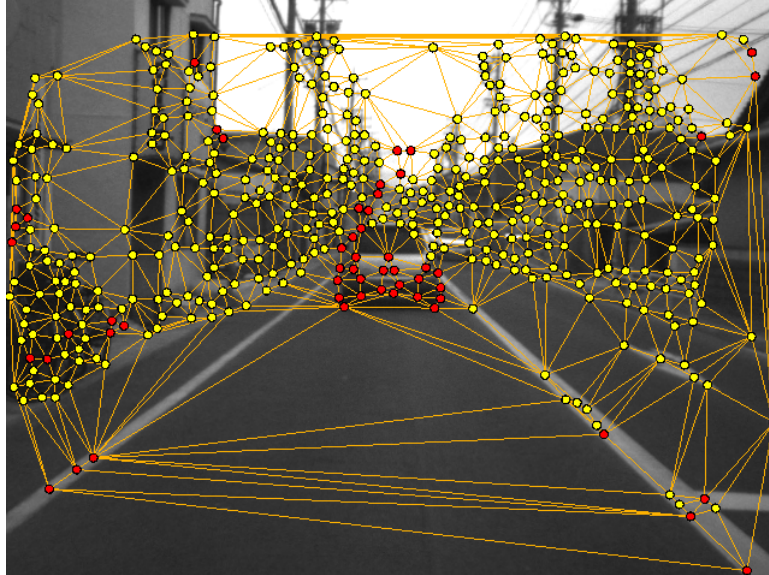
$$h_{\mathbf{q}} = |aQ_X + bQ_Y + cQ_Z - d| \quad (26)$$

以上のようにして、インライアの特徴点から作られる三角形内部の画素の高さをすべて計算し、高さを画素値とする高さ画像を生成する。

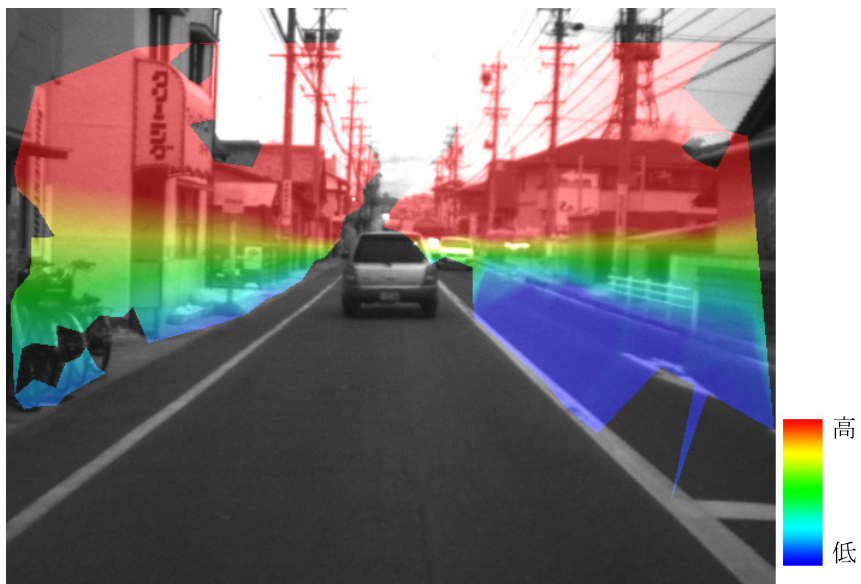
図 30 に、特徴点から生成された Delaunay 三角網と高さ画像の例を示す。図 30(a) において、黄色の点はインライアの特徴点、赤色の点はアウトライア (移動物体上または誤対応) の特徴点を示している。また、図 30(b) では元画像に重ねた高さ画像を示している。高さ画像は、高さが高いほど赤く、低いほど青くなるように表示している。また、高さの値がない画素には色を付けず元画像のまま表示している。このようにして生成した高さ画像を道路面領域の推定に利用する。

3.5 道路面領域の推定

射影変換による重ね合わせ処理から生成した差分画像と 3 次元構造情報から生成した高さ画像を用いて道路面領域の推定を行う。つまり、差分画像の値がしきい値以下で、かつ高さ画像の値が高さに対するしきい値以下である画素を道路面領域の画素として判定する。さらに、検出した領域に対して収縮、膨張のモルフロジー処理を行うことで、面積の小さい領域の除去を行う。しかし、一般的に白線などの路面標示の領域は輝度値が大きく、アスファルトの領域などは輝度値が小さいため、推定したカメラ運動、道路平面パラメータの誤差がある場合や実



(a) 特徴点と三角網



(b) 高さ画像

図 30 特徴点から生成された三角網と高さ画像

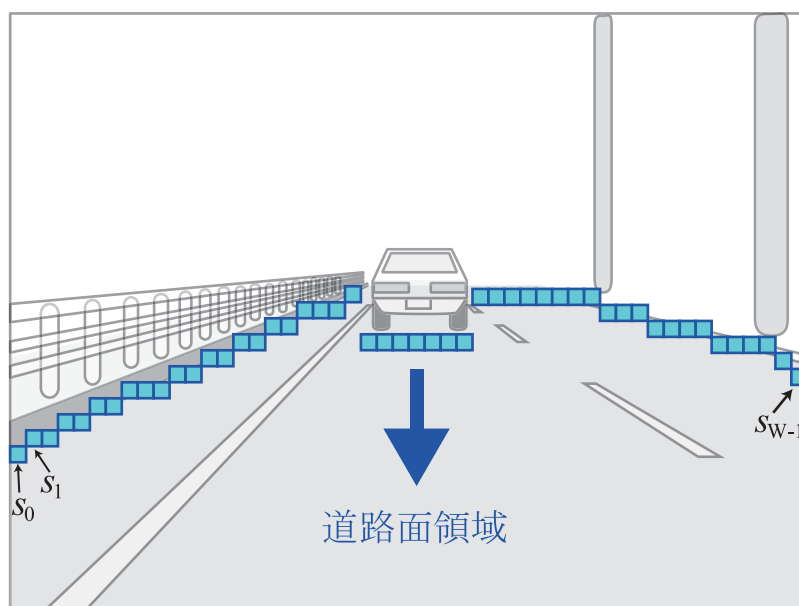


図 31 道路面領域の境界の条件

際の道路面が平面でない場合には射影変換した画像に微小なずれが生じて、路面標示領域とアスファルト領域が対応したとき、差分値が大きくなって道路面領域と判定されず、道路面領域に抜けが生じ、道路面領域が立体物領域と誤判定される問題が発生する。

そこでこの課題に対応するため、走行可能な道路面領域と立体物領域は画像上で完全に二分されるという仮定条件を適用する。これは、図 31に示すように、画像の各列において道路面領域と立体物領域の境界点となる点が 1 点存在し、その点より上は立体物領域、下は道路面領域に分離できるという条件である。つまり、画像の大きさを $W \times H$ (幅 \times 高さ) 画素とすると、道路面境界は各列の道路面境界点 $\{s_i\}, i = 0, \dots, W - 1$ で表される。このとき、道路面領域は、条件 $y \geq s_x$ を満たす画像座標 (x, y) の画素で構成される。ただし、画像の左上を画像座標の原点 $(0, 0)$ とし、下向きを画像 y 座標の正の方向とする。

この条件を用いた道路面領域の推定は以下のようにして行う。

- (1) 差分画像での差分値が差分のしきい値以下で、かつ高さ画像での道路面か

表 4 道路面領域の推定の評価実験に用いたカメラの仕様

カメラ	SONY XC-55
画角	32.6×24.8 度
解像度	640×480 画素
フレームレート	10 フレーム / 秒
カメラ設置高さ	1.28m
俯角	下向き 3 度

らの高さ値が高さのしきい値以下の画素を道路面領域として検出する。

- (2) 検出された道路面領域に対して収縮処理を適用し，そのあとに膨張処理を行うことで，孤立した微小領域を削除する。
- (3) 各列 i ごとに，最も上に存在する道路面領域に属する画素を検索し，その画素の垂直方向の位置を境界位置 s_i とする。
- (4) 各列 i において，境界位置 s_i より下の画素をすべて道路面領域の画素とする。

以上の処理により，安定に道路面領域を検出することができる。

3.6 実験

車載単眼カメラを用いて撮影した実走行画像に提案手法を適用する実験を行った。実験に用いた画像は (株) 豊田中央研究所が撮影したものである。用いたカメラの仕様を表 4 に示す。使用した計算機は，CPU Core2 Duo 3.0GHz，メモリ 4GB であり，計算時間は 1 フレームあたり 0.95 秒であった。道路面領域の推定精度を評価するため，手動で道路面領域の正解画像を作成した。正解画像では，車両の走行する道路部分を正解領域としている。正解画像は動画画像から 1 秒おきに抽出した昼間 107 枚，夜間 38 枚の合計 145 枚である。また比較手法として，射影変換による重ね合わせ処理のみから道路面領域を検出する手法を用いた。

昼間画像と夜間画像の結果例を図 32, 33にそれぞれ示す. 図 32, 33において, (a) は正解画像, (b) は射影変換による画像の重ね合わせ処理のみを用いた比較手法の結果画像, (c) は提案手法による結果画像を示す. なお, 道路面領域は水色で色づけした領域で表している. SfM アルゴリズムを用いた高精度な射影変換行列の推定と重ね合わせでのしきい値の適応的な決定により, 比較手法, 提案手法ともに昼夜で道路面領域が推定できていることが分かる. さらに提案手法では, 道路面領域の推定結果が比較手法より改善され, 道路面領域と誤って判定された立体物領域が減っていることが確認できる. 図 32, 33から, テクスチャの少ない建物や壁, 夜間の車両や照明が十分当たっていない比較的暗い部分に関して, 比較手法では誤って道路面領域と判定されているが, 提案手法では改善されている. これらは, 3次元構造情報から生成した高さ画像を利用した効果である. また, 図 32の1行目では, 推定した射影変換の誤差や道路面が完全に平面でないことから, 比較手法では道路面領域の未検出部分が発生しているが, 提案手法では未検出部分が少なくなっている. これは, 道路面領域の境界に対する条件を適用した効果である.

図 34に, 評価データ全体での道路面領域推定の再現率 (Recall) と適合率 (Precision) のグラフを示す. 再現率は (道路面領域と正しく判定された画素数)/(正解画像での道路面領域の総画素数), 適合率は (道路面領域と正しく判定された画素数)/(道路面領域と判定された総画素数) により計算される. 図 34のグラフは, 式 (18) の k を変化させて射影変換による重ね合わせ処理でのしきい値を変化させることにより得た結果である. 図 34より, 特に再現率が高くなったとき, 提案手法は比較手法より適合率が高いことが確認できる. 再現率が高くなるのは, 射影変換による重ね合わせ処理での差分値に対するしきい値を高く設定し, 道路面領域に判定される画素が増えた場合であるが, その場合, 射影変換による重ね合わせ処理のみでは, 誤って道路面領域と判定される立体物領域も増大することになり, 適合率が低下する傾向にある. しかし提案手法では, 図 32, 33に示したように, 高さ画像の利用により立体物領域を誤って道路面領域と判定することを軽減できるため, 再現率が高くなった場合にも適合率の低下を防ぐことができる. 以上の結果から, 実走行環境において提案手法は高い精度で道路面領域を検出すること

ができると言える。

3.7 結言

本論文では、車載単眼カメラを用いて撮影した画像から道路面領域を推定する手法を提案した。提案手法は、SfM アルゴリズムを道路環境へ適用する既存手法を利用することで、3次元構造と時系列画像間の道路平面に関する射影変換行列を高精度に推定する。そして、射影変換による重ね合わせに基づく処理と3次元構造復元の結果を両方利用することで道路面領域の推定精度を向上することができる。また、様々な環境においてロバストな推定を実現するため、道路面境界に対する条件を適用する処理を追加した。実走行画像を用いた実験で、提案手法は道路面領域を精度良く推定できることを確認した。特に、テクスチャの少ない壁などの領域を誤って道路面と判定することを低減することができる。

今後、さらに道路面領域の推定精度を向上させるためには、白線、路面標示検出手法やテクスチャ特徴を用いた識別手法と組み合わせることが考えられる。

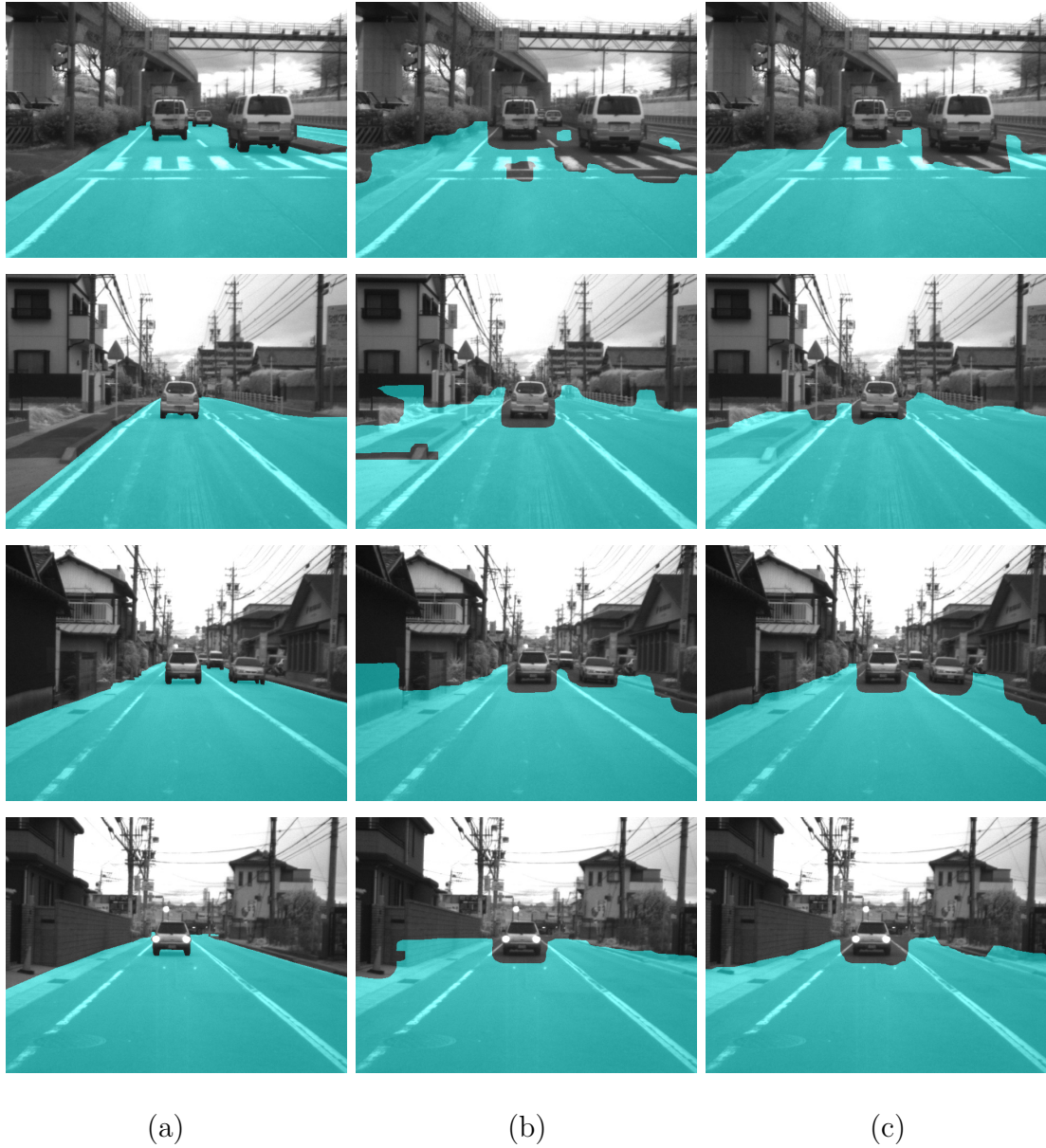
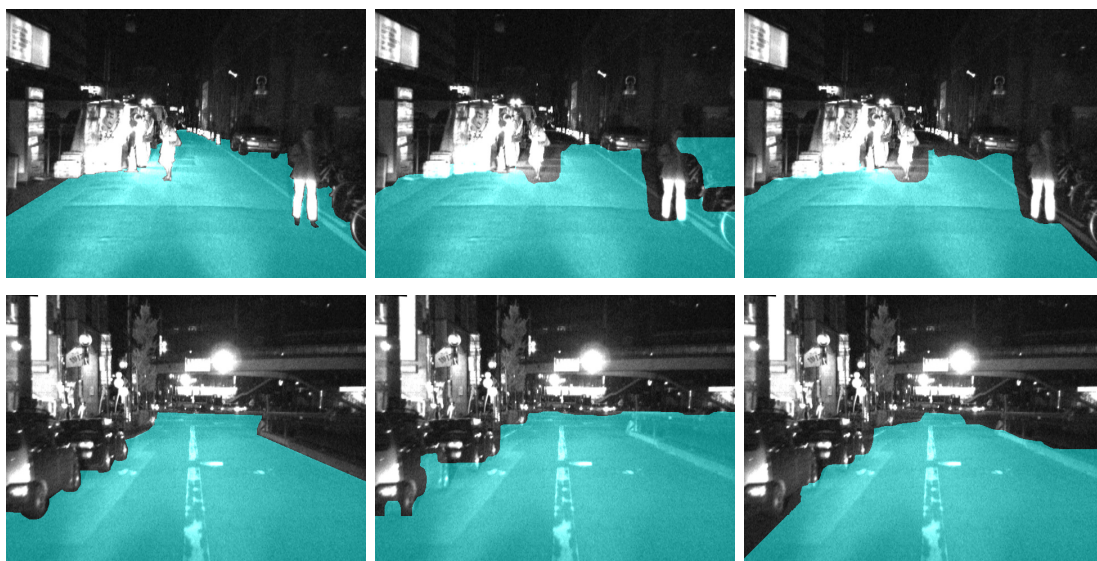


図 32 道路面領域の推定結果 (昼間): (a) 正解画像, (b) 射影変換による重ね合わせのみの結果, (c) 提案手法の結果



(a)

(b)

(c)

図 33 道路面領域の推定結果 (夜間): (a) 正解画像, (b) 射影変換による重ね合わせのみの結果, (c) 提案手法の結果

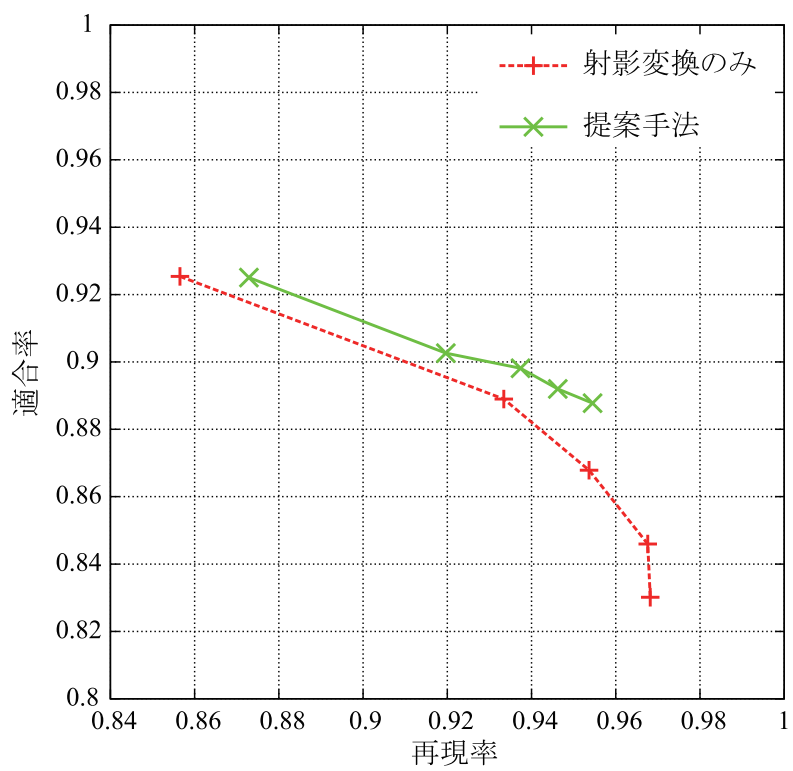


図 34 道路面領域推定の再現率と適合率の関係

4. 勾配方向の 2 次元ヒストグラム特徴を用いた歩行者検出

4.1 序言

現在、国内の交通事故による死亡者のうち、歩行者は3分の1以上を占め、最も高い割合となっている。交通事故による死傷者の低減を目的とする運転支援システムにおいて、歩行者を検出することは重要となる機能である。人体は電波の反射が強くないためレーダでの検出は難しく、レーザレーダを用いた場合にも離れた位置の歩行者は解像度の問題から検出が困難になる上に、物体が歩行者か否かを判定することは難しい。そこで、画像を用いて歩行者を検出する機能を実現することが期待されており、近年の一般物体認識技術 [EGW+] の発展とともに、人、歩行者を主な検出対象とする研究も数多く行われている [DWSP09, EG09]。現在提案されている多くの歩行者検出手法では、Sliding Window の枠組みを利用している。Sliding Window は、画像中のすべての位置、大きさのウィンドウを切り出し、切り出されたウィンドウごとに歩行者か否かを識別することにより、画像中の歩行者を検出する枠組みであり、切り出されたウィンドウから抽出する特徴量や各ウィンドウの識別を行う識別器に関して様々な手法が提案されている。歩行者検出に有効な特徴量として、濃度勾配情報を用いる特徴があり、中でも勾配方向のヒストグラムから特徴ベクトルを生成する HOG (Histograms of Oriented Gradients) 特徴 [DT05, Dal06] は高い識別性能を持つ特徴量として知られており、様々な検出手法で基本の特徴量として利用されている [MBM08, VGVZ09, FGMR10, WHY09, WWS09, WS08]。

本章では、濃度勾配情報による特徴量として、勾配方向の 2 次元ヒストグラムを用いる 2-Dimensional Histograms of Oriented Gradients (2D HOG) 特徴を提案する。2D HOG 特徴では、2 次元ヒストグラム特徴を用いる従来手法である GLAC [KO08] や CoHOG [WIY09] のように画素ごとのグリッド表現で 2 つの画素間の位置関係を表すのではなく、2 画素間の位置関係を極座標を用いて量子化して表すことにより、効果的に勾配の共起性を表現することができる。特徴量の計算時には、各画素に対してその周辺領域を極座標で分割し、分割した領域に対応する位置関係ごとに 2 次元ヒストグラムを生成することで、特徴ベクトルの次

元数を大きく増大させることなく、様々な位置関係にある画素間の共起性を効率的に表すことができる。

4.2 特徴量計算手法の処理の概要

画像から歩行者を検出するために Sliding Window の枠組みを用いることを考える。Sliding Window による歩行者検出の処理の流れを図 35に示す。Sliding Window では、ウィンドウの大きさを変化させながら画像を走査してウィンドウ画像を切り出し、切り出した画像を一定の大きさの画像にリサイズして、各画像の識別を行う。様々な大きさで切り出された画像をすべて一定の大きさの画像にリサイズして識別を行うことにより、共通の特徴量と識別器を用いることができ、学習時にもすべての大きさの歩行者画像を用いた多数の学習サンプルによる学習が可能となる。本節では、このような Sliding Window の枠組みにより、一定の大きさの画像が与えられたとして、その画像から勾配方向の 2 次元ヒストグラムを用いる特徴量 2D HOG を計算する手法について説明する。

提案手法では、入力画像の各画素の勾配の方向と大きさを計算し、特定の位置関係にある 2 つの画素の組の勾配から、図 36のような勾配方向の 2 次元ヒストグラムを生成する。ヒストグラムでは、勾配方向を量子化して bin を設定し、量子化数が D であるとき、2 次元ヒストグラムの bin の数は $D \times D$ となる。次に、図 37に処理の流れを示す。入力画像から切り出され、リサイズされた識別対象のウィンドウ画像が入力されると、最初に各画素の勾配を計算する。次に、2 つの画素の勾配方向の組み合わせを投票して 2 次元ヒストグラムを生成する。このとき、例えば図 37に示すような左右に隣り合う 2 つの画素 (図 37の位置関係 1) や斜めに隣り合う 2 画素 (位置関係 2) または離れた位置にある画素 (位置関係 M) など、様々な位置関係にある 2 つの画素の組み合わせから複数の勾配方向の 2 次元ヒストグラムを生成する。また、入力画像を格子状に分割して生成したブロックごとにこれらの 2 次元ヒストグラムを計算する。そして、生成した 2 次元ヒストグラムをそれぞれ正規化し、すべての 2 次元ヒストグラムの要素を並べて特徴ベクトルを生成する。最後に得られた特徴ベクトルから線形 SVM により歩行者の識別を行い、歩行者 / 非歩行者の判定を出力する。

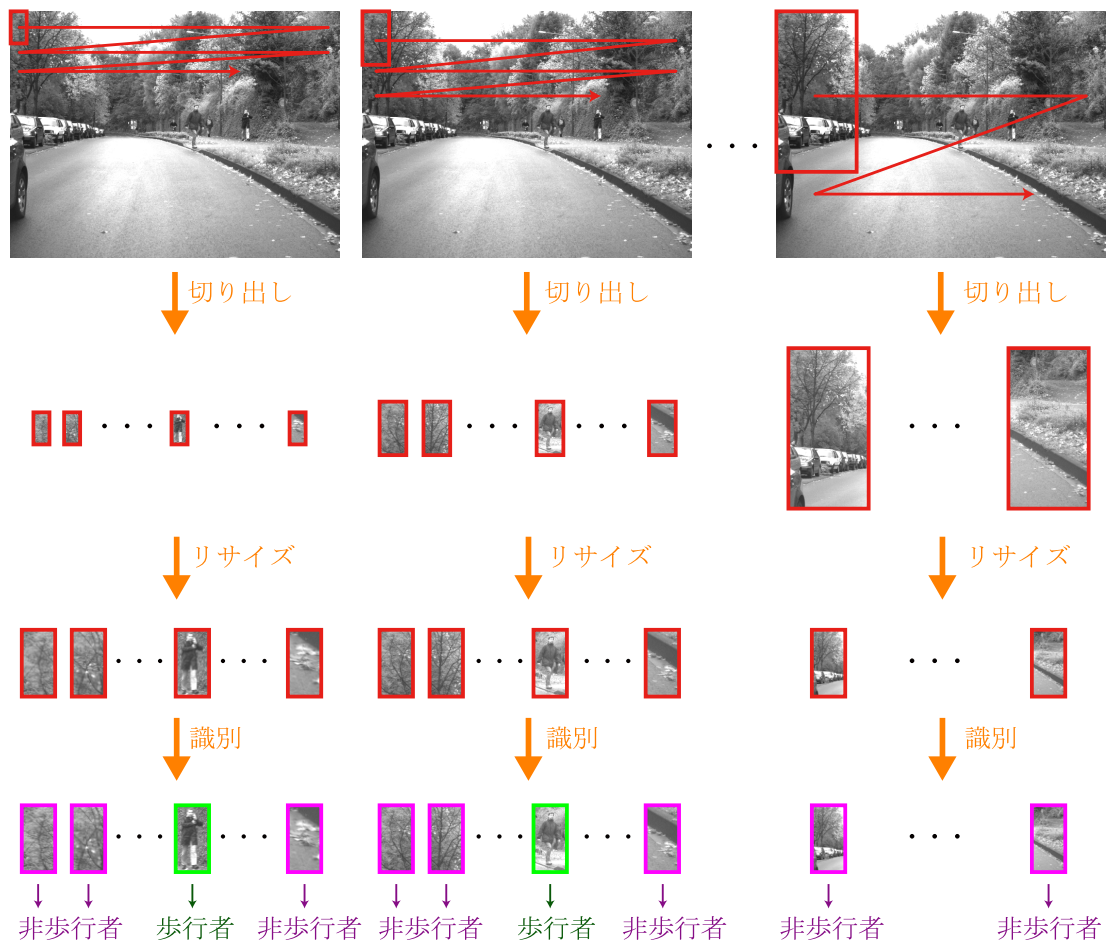


図 35 Sliding Window による歩行者検出処理の流れ

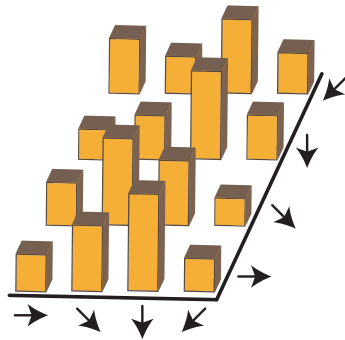


図 36 勾配方向の 2 次元ヒストグラム

4.3 勾配方向の 2 次元ヒストグラム特徴

この節では、濃度勾配方向の 2 次元ヒストグラム特徴 2-Dimensional Histograms of Oriented Gradients (2D HOG) の計算方法について説明する。2D HOG 特徴は、図 36 に示すような、2 つの画素の勾配方向の組み合わせを用いて生成した 2 次元ヒストグラムから構成される特徴量である。以下、勾配の計算、2 次元ヒストグラムの生成、正規化処理と特徴ベクトル生成について順に説明する。

4.3.1 勾配の計算

勾配の大きさ と 方向の計算 特徴量を抽出するウィンドウの画像 I が与えられたとき、各画素 (x, y) において勾配の大きさ $m(x, y)$ と方向 $\theta(x, y)$ を以下のように計算する。

$$m(x, y) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2}, \quad (27)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{\partial I}{\partial y} / \frac{\partial I}{\partial x}\right). \quad (28)$$

$\partial I / \partial x, \partial I / \partial y$ の計算は画像の微分オペレータで行う。使用する微分オペレータとしては、表 5 に示すオペレータが使われることが多い。よりコンパクトなオペレータとして、 $[-1, 1], [-1, 1]^T$ もあるが、 x 方向と y 方向で微分の中心位置が異

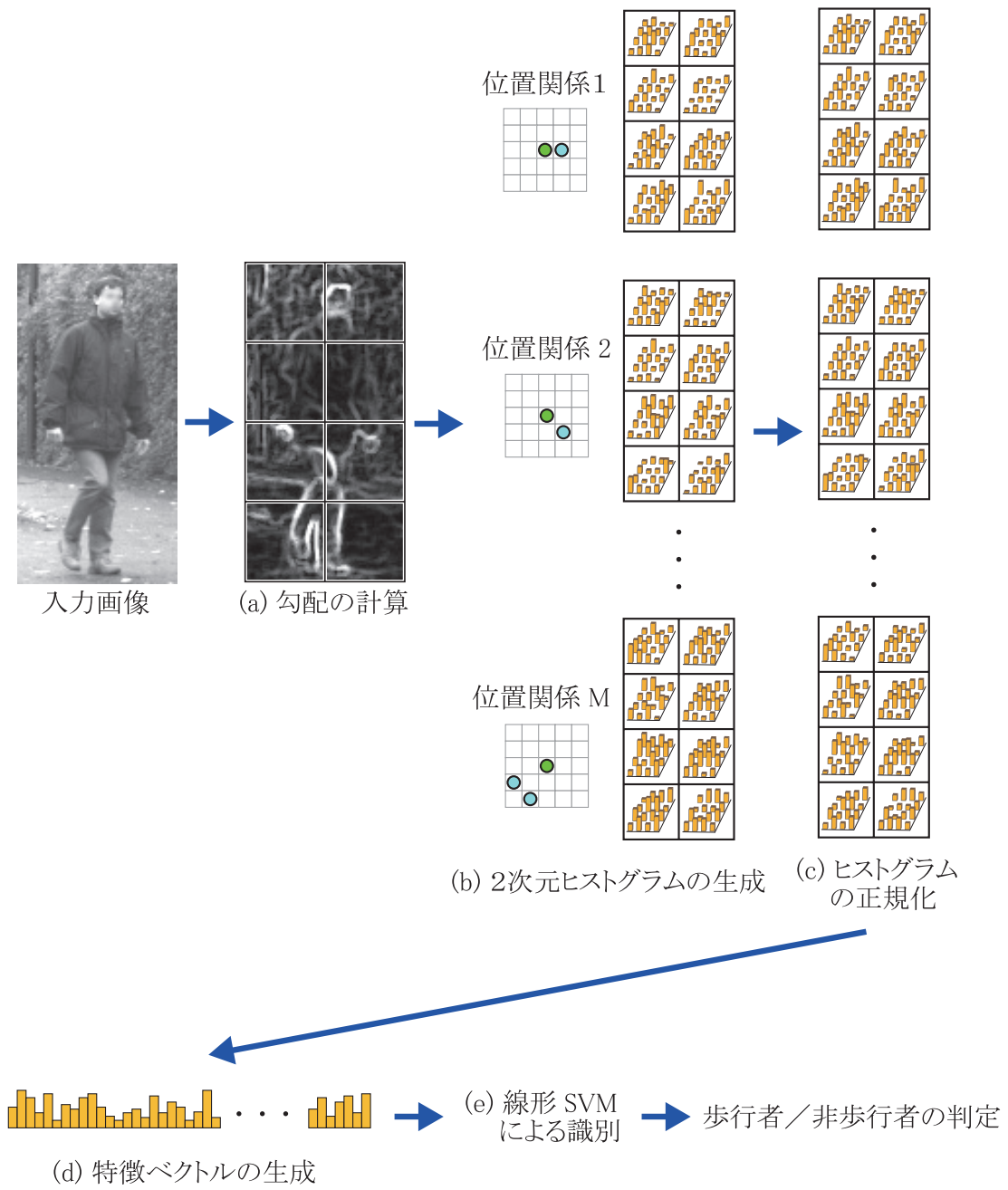


図 37 2D HOG 特徴による歩行者検出処理の流れ

表 5 微分オペレータ

1D Gradient	Roberts	Sobel
$[-1, 0, 1], \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

なるため、正しく勾配を計算することができず、HOG 特徴での評価において性能が劣化することが報告されている [Dal06].

G-O ベクトルの計算 方向 $\theta(x, y)$ は、ヒストグラムにおける勾配方向の量子化数 D と同じ次元数の G-O ベクトル (gradient orientation vector) $\mathbf{f} \in \mathbf{R}^D$ に変換する [KO08]. これは、図 38 に示すように勾配の方向を、量子化した D 方向に重み付けして表したベクトルであり、2 つまたは 1 つの要素のみが 0 でない、粗な D 次元ベクトルとなる. 勾配方向の量子化数 D は、ヒストグラムの bin 数を決定する値であり、1 次元ヒストグラムの場合には bin 数は量子化数 D と等しくなり、2 次元ヒストグラムの場合には bin 数は D^2 となる. HOG 特徴では、 $D = 8$ または 9 が一般的に用いられる. G-O ベクトル $\mathbf{f}(x, y)$ の i 番目の要素の値 $f_i(x, y)$ は以下のように定義される.

$$f_i(x, y) = \begin{cases} \frac{\Omega}{D} (b(\bar{\theta}(x, y)) + 1) - \bar{\theta}(x, y) & (i = b(\bar{\theta}(x, y)) \text{ のとき}) \\ \bar{\theta}(x, y) - \frac{\Omega}{D} b(\bar{\theta}(x, y)) & (i = b(\bar{\theta}(x, y)) + 1 \text{ のとき}) \\ 0 & (\text{それ以外の場合}) \end{cases} \quad (29)$$

$$b(\bar{\theta}(x, y)) = \left\lfloor \frac{D}{\Omega} \bar{\theta}(x, y) \right\rfloor \quad (30)$$

Ω は、勾配方向の範囲を表し、 $\Omega \in \{180^\circ, 360^\circ\}$ である. $\Omega = 180^\circ$ の場合には、勾配の方向 $\theta(x, y)$ は $\bar{\theta}(x, y) = [0, 180^\circ)$ で表され、逆方向の勾配と区別しない符

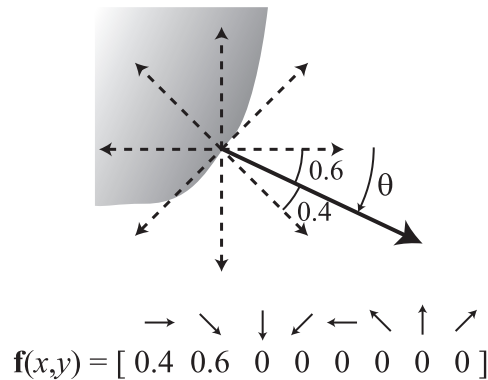


図 38 勾配と G-O ベクトル

号なし表現となる．一方， $\Omega = 360^\circ$ の場合には， $\bar{\theta}(x,y) = [0, 360^\circ)$ で勾配方向を表し，逆方向の勾配方向も別の方向として扱う．また， $b(\bar{\theta}(x,y))$ は勾配方向 $\bar{\theta}(x,y)$ に隣接する 2 つの量子化方向のうち，小さい方の番号を表す．このような重み付けで各画素の勾配方向を表すことで，勾配方向の量子化によるエイリアシングを防ぐことができる．

4.3.2 2次元ヒストグラムの生成

勾配方向の 2 次元ヒストグラムは，2 つの画素の勾配方向の組み合わせを投票して生成する．このとき，歩行者の形状情報をより詳細に表現するため，2 つの画素の位置関係ごとに 2 次元ヒストグラムを生成する．2D HOG では，入力画像を走査しながら，各画素においてその周辺領域を極座標で分割し，その画素と分割した領域内の画素の勾配の組み合わせを投票して 2 次元ヒストグラムを計算する．周辺領域を分割した領域をオフセット領域と呼ぶ．オフセット領域は，注目する画素の周囲に設定された周辺領域を，極座標で分割することにより定義される．以下，周辺領域の設定と分割方法について説明する．

周辺領域の設定 注目画素の周囲の円領域を周辺領域として設定し，注目画素と周辺領域内の画素のすべての組み合わせを 2 次元ヒストグラム生成のために

用いる。この周辺領域は、円の半径 R と形状 (全体または下半分) をパラメータとして決定する。まず、注目画素が $\mathbf{x} = (x, y)$ であるときの周辺領域内の画素の集合 $S(\mathbf{x})$ は、次の式により定義されるオフセット画素 $\mathbf{p} = (u, v)$ から構成される。

- $R = 1$ のとき

$$S(\mathbf{x}) = \{\mathbf{p} = (u, v) \mid |x - u| \leq 1, |y - v| \leq 1, (u, v) \neq (x, y)\} \quad (31)$$

- $R = 2, 3, 4, \dots$ のとき

$$S(\mathbf{x}) = \{\mathbf{p} = (u, v) \mid (x - u)^2 + (y - v)^2 < R^2 + 4, (u, v) \neq (x, y)\} \quad (32)$$

半径 R が 2 以上のときには、水平方向に R 画素、垂直方向に 2 画素 (または、垂直方向に R 画素、水平方向に 2 画素) だけ注目画素から離れた画素との距離を半径とした円の内側にある画素、つまり注目画素 $\mathbf{x} = (x, y)$ からの距離が $\sqrt{R^2 + 2^2}$ 未満となる画素 $\mathbf{p} = (u, v)$ を周辺領域内のオフセット画素としている。また、形状が全体の場合は、円領域内の画素全体を、下半分のときは、以下を満たす円領域内の画素 \mathbf{p} のみを用いる。

$$\{\mathbf{p} = (u, v) \mid (u > x \text{ and } v = y) \text{ or } v > y\} \quad (33)$$

図 39 に、周辺領域の設定例を示す。図 39 は、半径 $R = 4$ の例であり、中心の黒く塗りつぶされた画素が注目画素、丸をつけた画素が設定された周辺領域内の画素である。形状が全体の場合は、白丸と黒丸をつけた画素両方、下半分の場合は、黒丸の画素のみとなる。

距離, 方向による分割 設定された周辺領域を極座標で分割してオフセット領域を定義する。分割するためのパラメータには以下のものがある。

- 半径の分割数 n
- 分割する半径 $r_i, i = 1, \dots, n$

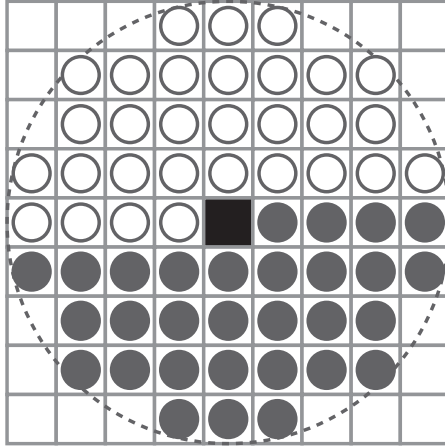


図 39 2次元ヒストグラム生成のための周辺領域

- 方向数 $d_i, i = 1, \dots, n$

分割パラメータに基づいて周辺領域を極座標で分割した領域をそれぞれオフセット領域と呼び、 $W_{r,d}$ ($r = 1, \dots, n, d = 1, \dots, d_r$) で表す。図 40に周辺領域を分割して生成したオフセット領域の例を示す。図 40(a)は、周辺領域が半径 $R = 4$ 、形状は全体で、分割数 $n = 2$ 、分割半径 $r_1 = 2, r_2 = 4$ 、方向数 $d_1 = 8, d_2 = 12$ のオフセット領域、(b)は、周辺領域が半径 $R = 4$ 、形状は下半分で、分割数 $n = 2$ 、分割半径 $r_1 = 2, r_2 = 4$ 、方向数 $d_1 = 3, d_2 = 3$ のオフセット領域、(c)は、周辺領域が半径 $R = 2$ 、形状は下半分で、分割数 $n = 1$ 、分割半径 $r_1 = 2$ 、方向数 $d_1 = 3$ のオフセット領域を示している。なお、このオフセット領域の設定パラメータを $(r\{r_1\}d\{d_1\}, r\{r_2\}d\{d_2\}, \dots, r\{r_n\}d\{d_n\})$ という形式で表すこととする。例えば、図 40(a)は $(r2d8, r4d12)$ 、(b)は $(r2d3, r4d3)$ 、(c)は $(r2d3)$ となる。

勾配方向の2次元ヒストグラムの生成 前節で述べた方法で設定されたオフセット領域ごとに勾配方向の2次元ヒストグラムを生成する。各オフセット領域 $W_{r,d}$ において、 $D \times D$ 次元の2次元ヒストグラム $H_{r,d}$ は次のようにして計算する。

$$H_{r,d}(o_0, o_1) = \sum_{\mathbf{x} \in I} \sum_{\mathbf{p} \in S(\mathbf{x})} \left\{ w[m(\mathbf{x}), m(\mathbf{p})] \cdot t_{r,d}(\mathbf{p}) \cdot \mathbf{f}_{o_0}(\mathbf{x}) \cdot \mathbf{f}_{o_1}(\mathbf{p}) \right\}, \quad (34)$$

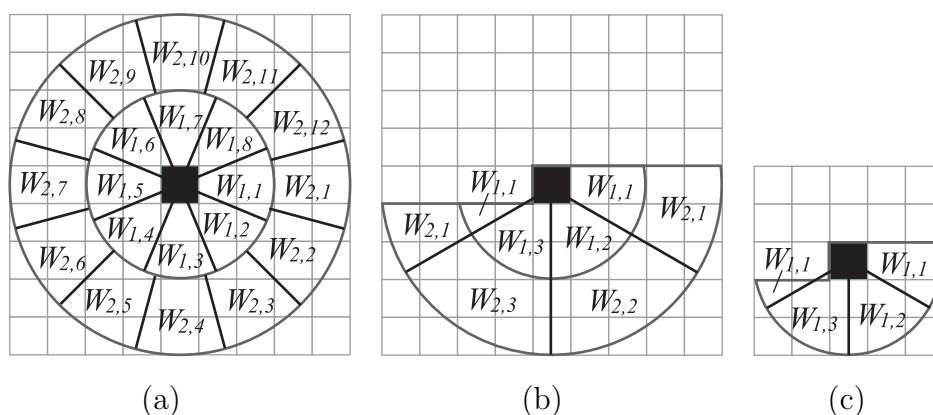


図 40 オフセット領域の設定例

$o_0, o_1 = 1, \dots, D$ は 2 つの勾配方向, $m(\mathbf{x}), m(\mathbf{p})$ はそれぞれ式 (27) で計算される注目画素 \mathbf{x} とそのオフセット画素 \mathbf{p} における勾配の大きさ, w は 2 つの勾配の大きさにより決定する重み関数, $t_{r,d}(\mathbf{p})$ はオフセット画素 \mathbf{p} に関するオフセット領域 $W_{r,d}$ の重み, \mathbf{f}_o は G-O ベクトルの o 方向の要素を表す. オフセット領域の重み $t_{r,d}(\mathbf{p})$ は, オフセット画素 \mathbf{p} の注目画素からの方向により隣接する 2 つのオフセット領域に重み付けする. 図 41 にオフセット領域の重みの設定例を示す. このオフセット領域は, 図 40(a) と同じパラメータで設定されたものであり, 白丸を付けたオフセット画素に関して, オフセット領域 $W_{2,1}$ と $W_{2,2}$ の 2 つのオフセット領域に方向に応じて重み付けされる. また, 重み関数 $w(\cdot)$ は, $\min(\cdot)$ を用いる.

周辺領域の形状 これまでに説明したように周辺領域の形状として, 円領域内のすべての画素をオフセット画素とする「全体」と注目画素の下側にある画素のみをオフセット画素とする「下半分」がある. 従来の 2 次元ヒストグラム手法 [KO08, WIY09] では, 「下半分」に対応するオフセット画素のみが用いられているが, これは 1 つの 2 次元ヒストグラムを生成するために用いる画素間の位置関係は 1 つであるため「全体」に対応するオフセット画素を用いることは完全に冗長となるからである. 一方, 本手法では注目画素に対して, オフセット領域の

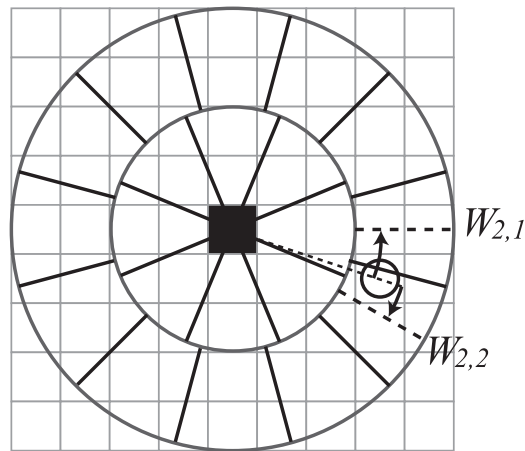


図 41 オフセット領域の重み

複数の画素を用いて1つの2次元ヒストグラムが計算されることから、周辺領域の形状を「全体」としたときにも完全に冗長とはならない。そのため、周辺領域の形状として「全体」と、従来手法で用いられている「下半分」の2種類を設けている。

周辺領域の形状が「全体」のときと「下半分」のときに違いについて簡単に述べる。図 42に注目画素をシフトしたときの、注目画素とオフセット画素の位置関係を示す。図 42において、注目画素を黒丸、オフセット画素を白丸で示している。従来手法のようにオフセット画素が1画素のみの場合(図 42(a)参照)、注目画素に対して右下の画素をオフセット画素とした組み合わせと、注目画素を右下にシフトして左上の画素をオフセット画素とした組み合わせは、同じ2画素の組み合わせとなるため、2次元ヒストグラムを生成する際には冗長な情報となる。一方、本手法のように複数のオフセット画素から成る場合(図 42(b)参照)、注目画素がシフトしても同一の画素の組み合わせ集合とならない。したがってこのような場合には、周辺領域の形状を「全体」としたときと「下半分」としたときには、大きな違いとはならない可能性はあるものの、生成される2次元ヒストグラムが異なる結果となる。

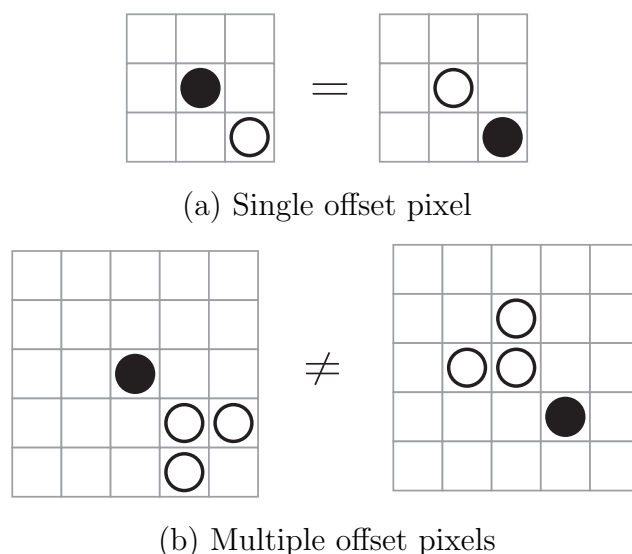


図 42 画素シフト時の注目画素とオフセット画素の関係

4.3.3 正規化と特徴ベクトルの生成

入力画像内の位置情報を特徴量に反映させるため，入力画像 I をブロックと呼ぶ小領域に分割し，各ブロックでオフセット領域ごとの 2 次元ヒストグラムを生成する．そして，各ブロック，オフセット領域の 2 次元ヒストグラムを正規化する．これは，2 次元ヒストグラムへの投票は勾配の大きさを重みとして用いているため，明るさ，コントラストの変化に対応するためであり，また 2D HOG ではオフセット領域 $W_{r,d}$ ごとに投票する画素の組の数が異なる可能性があるため，すべての 2 次元ヒストグラムの大きさを合わせるためでもある．正規化は，L2-Hys [Low04] と呼ばれる方法を用いる．

最後に，すべてのブロック，オフセット領域の 2 次元ヒストグラムの要素を並べたベクトルを特徴ベクトルとする．

以上の処理により計算される 2D HOG 特徴ベクトルの次元数は，ヒストグラムにおける勾配方向の量子化数が D ，オフセット領域の半径の分割数が n ，方向数が $d_i, i = 1, \dots, n$ ，ブロックの数が $b_w \times b_h$ のとき， $D^2 \times b_w \times b_h \times \sum_{i=1}^n d_i$ となる．

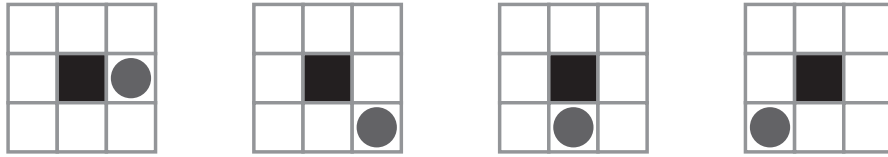


図 43 GLAC 特徴のオフセットパターン

4.4 従来研究との関係

提案する 2D HOG 特徴は、勾配方向の 2 次元ヒストグラムを用いた特徴量であるが、同様に 2 つの画素の勾配方向の組み合わせから 2 次元ヒストグラムを生成する従来手法である GLAC 特徴と CoHOG 特徴との関係について説明する。

- GLAC 特徴

GLAC 特徴で計算される勾配の 1 次自己相関は、式 (34) と同様の書式を用いた場合、以下のように定義される。

$$H(o_0, o_1, \mathbf{a}_i) = \sum_{\mathbf{x} \in I} \{w [m(\mathbf{x}), m(\mathbf{x} + \mathbf{a}_i)] \cdot \mathbf{f}_{o_0}(\mathbf{x}) \cdot \mathbf{f}_{o_1}(\mathbf{x} + \mathbf{a}_i)\}, \quad (35)$$

\mathbf{a}_i は、オフセットベクトルであり、1 次自己相関の場合、図 43 で示す 4 つのオフセットパターンを用いる。つまり GLAC 特徴で生成される特徴量は、上下斜めに隣接する 2 つの画素の勾配方向の組み合わせにより生成した 4 種類の 2 次元ヒストグラムであると言える。そしてこれは、2D HOG 特徴のオフセット領域を、分割数 $n = 1$ 、分割半径 $r_1 = 1$ 、方向数 $d_1 = 4$ としたものと等価である。

- CoHOG 特徴

CoHOG 特徴では、図 44 に示す 1~30 の番号が付けられた 30 種類のオフセットベクトルそれぞれを用いて勾配方向の 2 次元ヒストグラムを生成する。これは、2D HOG 特徴のオフセット領域を、分割数 $n = 4$ 、分割半径 $r_1 = 1, r_2 = 2, r_3 = 3, r_4 = 4$ 、方向数 $d_1 = 4, d_2 = 6, d_3 = 8, d_4 = 12$ とし

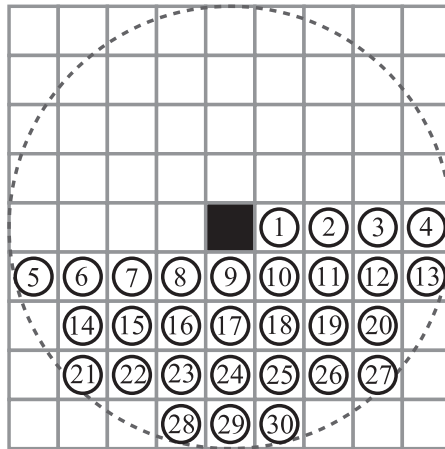


図 44 CoHOG 特徴のオフセットパターン

たものとはほぼ同じと考えることができる。(各半径において CoHOG のオフセットベクトルの方向は等間隔でないため、完全には同じではない。)

ただし、CoHOG 特徴は 2 次元ヒストグラムの生成において、勾配の大きさによる重み付け投票は行わず、定数を投票、つまり勾配方向の組み合わせの数をカウントする。また、各画素の勾配方向は最も近い量子化方向の 1 つのみに割り当てて表し、G-O ベクトルのように 2 つの方向の重み付け表現を行わない。さらに、勾配方向の大きさがしきい値以下の画素に対しては、勾配なしとして、2 次元ヒストグラム投票には使用しない。CoHOG 特徴では、30 種類の 2 次元ヒストグラムを用いるために膨大な次元数となるが、このように勾配方向を 1 つの量子化方向のみへ割り当て、勾配の大きさがしきい値以下の画素を用いないことにより、2 次元ヒストグラムのほとんどの要素が 0 となり、特徴ベクトルが非常に粗なベクトルとなる。そのため、次元数が膨大であるにもかかわらず、多数の学習サンプルを用いた学習や識別を効率的に実行することができるという利点を持つ。また、CoHOG 特徴では、30 種類の 2 次元ヒストグラムに加えて、1 次元ヒストグラムも連結して特徴ベクトルとしている。

表 6 歩行者データベースの仕様

データセット	Daimler Pedstrian	夜間近赤外画像
学習データ	歩行者画像: 4,800 枚 ×3	歩行者画像: 5,000 枚
	非歩行者画像: 5,000 枚 ×3	非歩行者画像: 5,000 枚
テストデータ	歩行者画像: 4,800 枚 ×2	歩行者画像: 5,000 枚
	非歩行者画像: 5,000 枚 ×2	非歩行者画像: 5,000 枚
画像サイズ	18×36 画素	30×60 画素

以上で述べた関係から，2D HOG 特徴は従来手法である GLAC 特徴，CoHOG 特徴を一般化した特徴量と言える。

4.5 歩行者認識性能の評価実験

提案手法の歩行者識別性能の評価実験を行った。実験では，歩行者と非歩行者の切り出し画像から構成される 2 種類のデータベースを用いた。1 つは，公開データベースである Daimler Pedestrian Classification Benchmark [MG06] で，本論文で提案する 2D HOG 特徴と従来手法の性能比較および 2D HOG のパラメータと性能の関係の評価を行った。もう 1 つは，(株) 豊田中央研究所において，車載近赤外線カメラによる夜間の歩行者検出技術の開発のため構築した歩行者データベースで，従来手法との性能比較を行った。2 つのデータベースの仕様を表 6 に示す。Daimler データベースは，それぞれ歩行者が 4,800 枚，非歩行者が 5,000 枚から構成される学習データセットが 3 つ，テストデータセットが 2 つあり，異なる学習データセットとテストデータセットの組の 6 種類の組み合わせによる評価結果の平均を最終的な評価結果として用いる [MG06]。また，夜間近赤外画像データベースは，夜間近赤外画像から Viola-Jones の手法 [VJ04] を適用し，誤検出したウィンドウ画像を非歩行者サンプルとしている。Daimler データベース，夜間近赤外画像データベースの歩行者，非歩行者サンプルの例をそれぞれ図 45，46 に示す。

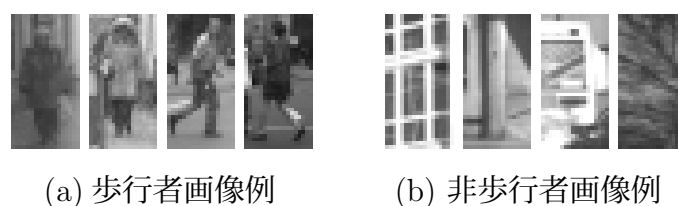


図 45 Daimler データベースの歩行者，非歩行者画像例

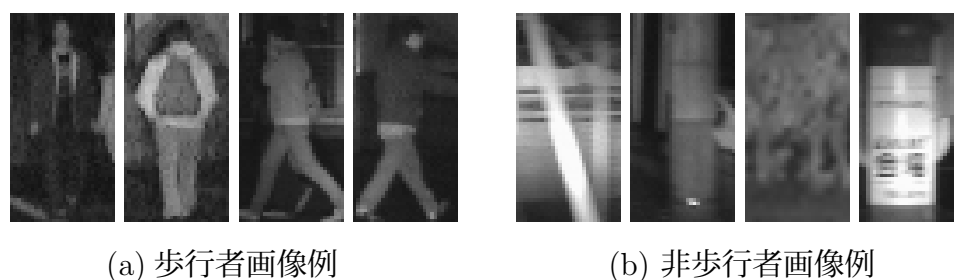


図 46 夜間近赤外画像データベースの歩行者，非歩行者画像例

実験では，2D HOG 特徴の他に HOG，GLAC，CoHOG 特徴を用いて評価を行った．比較評価では，これらの4つの特徴に対して同じパラメータ設定を用いた．画素の勾配は Roberts フィルタにより計算，勾配方向は $0\sim 360^\circ$ を8方向に量子化した．これらは，Daimler データベースで識別性能が高くなるように選択したパラメータであり，夜間近赤外データベースでも同様の性能の傾向を示した．HOG は，Daimler データベースではセルサイズ 3×3 ピクセル，ブロックサイズ 2×2 セル，夜間近赤外データベースではセルサイズ 5×5 ピクセル，ブロックサイズ 2×2 セルとした．また，GLAC，CoHOG，2D HOG では，入力画像を 3×6 ブロックに分割した．2D HOG では，周辺領域が半径 $R = 4$ ，形状が下半分，オフセット領域が $(r2d3, r4d3)$ を基本パラメータとし (図 40(b) 参照)，評価結果においてなんらかのパラメータを変更している場合にも，表記していないパラメータに関しては基本パラメータに従う．各特徴量の特徴ベクトルの次元数は，HOG が 1,760 次元，GLAC が 4,608 次元，CoHOG が 34,704 次元，2D HOG は基本としたパラメータのとき，6,912 次元である．なお，すべての評価において，識別器は線形 SVM [Vap98](付録 B 参照) を適用し，線形 SVM の学習，識別には，

LIBLINEAR ライブラリ [FCH⁺08] を使用した。使用した計算機は、CPU Core2 Quad 2.67GHz、メモリ 8GB であり、1 枚のウィンドウ画像の特徴量計算に要した時間は、HOG が 0.17 ミリ秒、GLAC が 0.44 ミリ秒、CoHOG が 0.47 ミリ秒、2D HOG が 1.28 ミリ秒であった。

Damiler データベースでの評価結果を図 47~54 に示す。図 47~54 は、ROC (Receiver Operating Characteristic) カーブであり、横軸は非歩行者を歩行者と誤って判定した割合 False Positive Rate、縦軸は歩行者を歩行者として正しく判定した割合 Detection Rate である。ROC カーブが左上にいけばいくほど識別性能が高いことを示す。

図 47 に従来手法と提案手法の比較結果を示す。勾配方向の 2 次元ヒストグラムを用いる特徴 GLAC、CoHOG、2D HOG は、HOG と比べて識別性能が高いことが確認でき、その中でも 2D HOG は、最も高い識別性能であることが分かる。

図 48 にオフセット領域の数を増やした場合の結果を示す。オフセット領域の数を増やすことで識別性能の向上が確認でき、(r1d4, r2d4, r4d12) の場合には、False Positive Rate が 0.1 のとき、Detection Rate が 0.99 となった。しかしこの場合、特徴量の次元数は大きく増大し、23,040 次元となる。

図 49 に周辺領域を全体 (Full) と下半分 (Half) に変更したときの結果を示す。(r2d3, r4d3) の Full と Half または (r2d6, r4d6) の Full と Half のようにそれぞれ同じオフセット領域の数では、形状が下半分の方が全体より識別性能が高い。これは、周辺領域を分割する方向数が同じでも、全体の場合は 360° を、下半分の場合は 180° を分割することになるため、1 つのオフセット領域の占める角度が広がるためであると考えられる。周辺領域を全体にしたとき、分割する方向数を倍にすれば、1 つのオフセット領域の占める角度を下半分のときと同じにすることができ、その場合、Half の (r2d3, r4d3) と Full の (r2d6, r4d6) の結果のように周辺領域が全体の方が性能が高くなる。しかしこのとき、特徴量の次元数が増大することになるため、同じ次元数であれば、周辺領域を下半分とする方が効果的と考えられる。

図 50 に周辺領域を半径 $R = 4$ 、下半分の形状としたときの、半径の分割数を変更したときの結果を示す。(r2d3, r4d3) と (r4d6) は、周辺領域が同じ (半径 $R = 4$,

下半分の形状) であり, またオフセット領域の総数も同じであるため特徴量は同じ次元数となるが, 方向数を増やした (r4d6) よりも, 半径の分割数を2つにした (r2d3, r4d3) の方が性能が高く, さらに (r4d6) より方向数を増やし, 特徴量の次元数も増大している (r4d9) よりもわずかに識別性能がよい. この結果から, 画素間の距離に応じて別の2次元ヒストグラムを生成することが識別に有効な特徴となることが分かる.

図 51に周辺領域の半径を $R = 2, 4, 6$ と変化させたときの結果を示す. それぞれ, オフセット領域は半径2ごとに分割し, 方向数はすべて3方向である. 半径を大きくした方が性能が向上することが確認できるが, 半径を $R = 2$ から $R = 4$ にしたときの性能向上と比べると, $R = 4$ から $R = 6$ にしたときの性能向上は小さいことが確認できる. このことから, ある程度以上半径を大きくしても効果は大きくないと考えることができるが, 図 51の評価では方向数をすべて一定としており, その影響がある可能性もある. そこで, 図 52に, 周辺領域の半径を $R = 4, 6$ としたときに方向数を増やした場合の結果を示す. (r2d3,r4d6,r6d12) のように半径と方向数をともに増やしたとき, (r2d3,r4d3) と比べると性能が大きく向上しているが, (r2d6,r4d12) とはほぼ同じ性能であり, 半径を大きくした効果はほとんどないことが確認できる. 以上のことから, ある程度以上半径を大きくしても性能向上にはあまり寄与せず, 適当な半径を設定することが望ましいと言える.

図 53に画素の勾配方向の量子化方法の違いによる結果を示す. 画素の勾配方向を重み付けで表す G-O ベクトルでなく, 最も近い1つの量子化方向のみで表した場合には, 大きく識別性能が低下していることが分かる. これは, 2D HOG 特徴が CoHOG 特徴より性能が高くなった1つの要因であると言える. また, 図 54に2次元ヒストグラムに加えて1次元ヒストグラムを用いた場合の結果を示す. 2次元ヒストグラムを用いる場合には, 1次元ヒストグラムを加えても性能向上には寄与しないことが確認でき, これは [KO08] で示されている結果と同様である.

次に, 夜間近赤外画像データベースでの従来手法との比較結果を図 55に示す. このデータにおいても, 2D HOG 特徴が最も性能が高いことが確認できる.

以上の結果から, 2D HOG 特徴は, 1次元勾配方向ヒストグラムから構成され

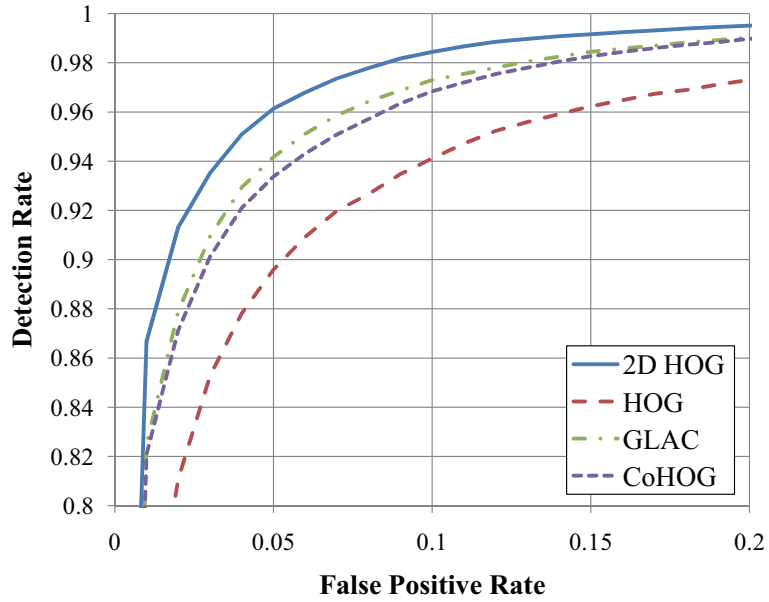


図 47 Daimler データベースの評価結果: 従来手法との性能比較

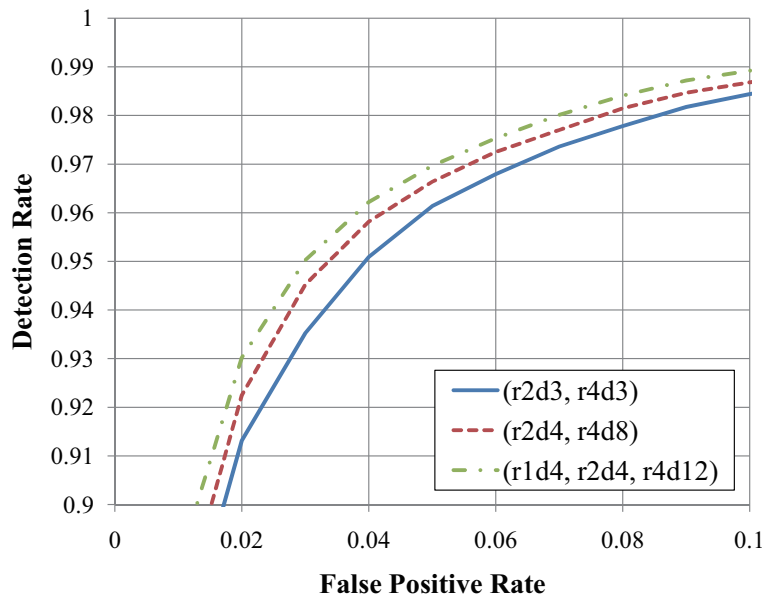


図 48 Daimler データベースの評価結果: オフセット領域

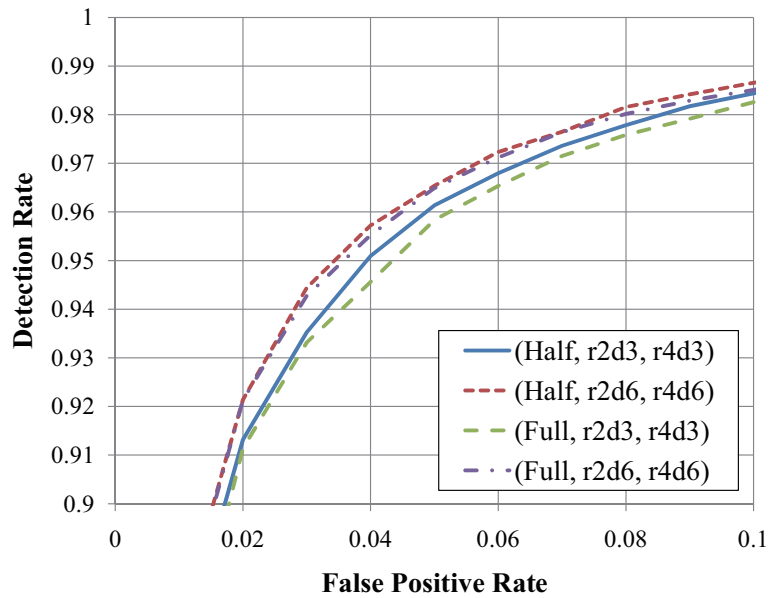


図 49 Daimler データベースの評価結果: 周辺領域の形状

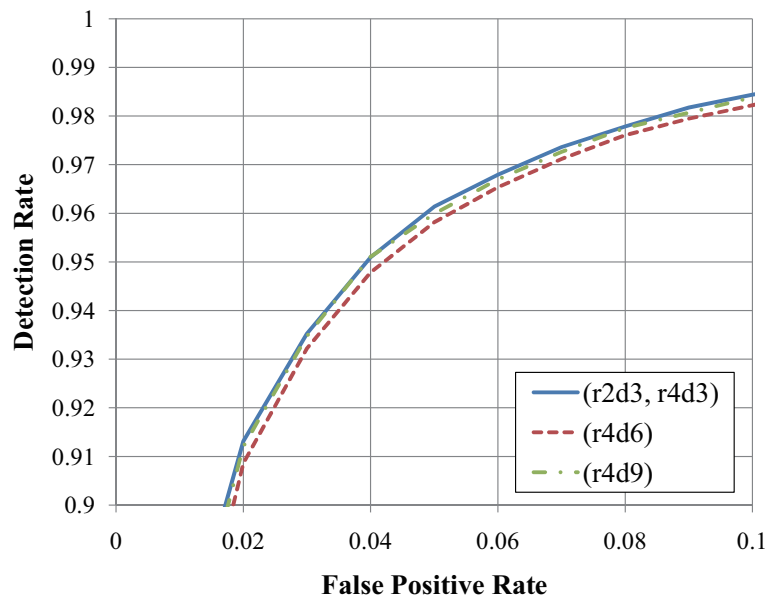


図 50 Daimler データベースの評価結果: 半径の分割数

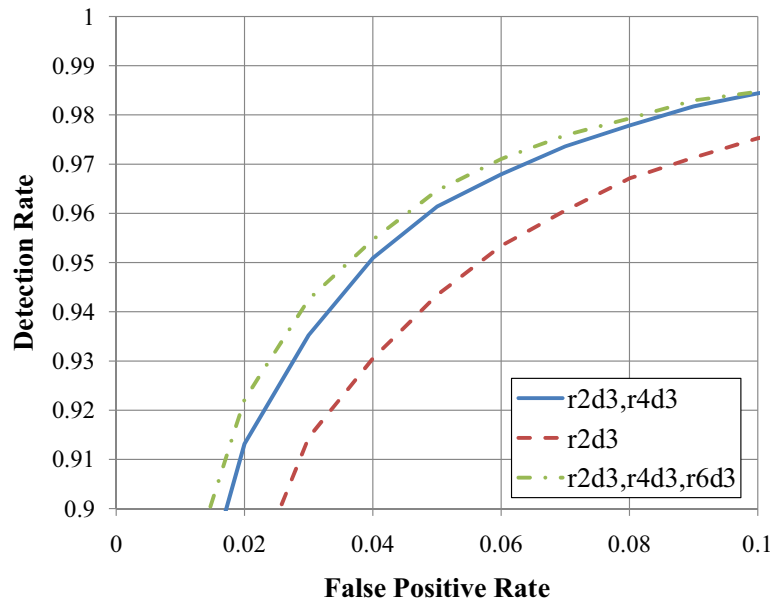


図 51 Daimler データベースの評価結果: 周辺領域の半径

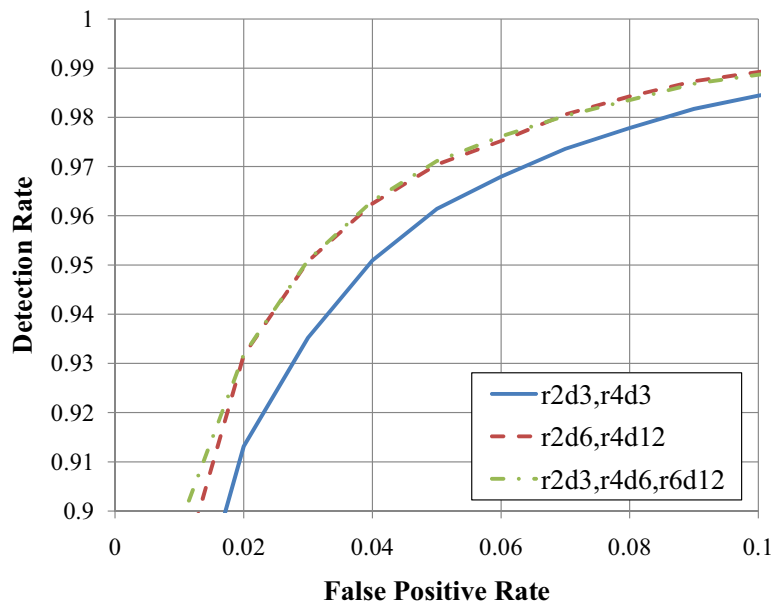


図 52 Daimler データベースの評価結果: 周辺領域の半径とオフセット領域の方向数

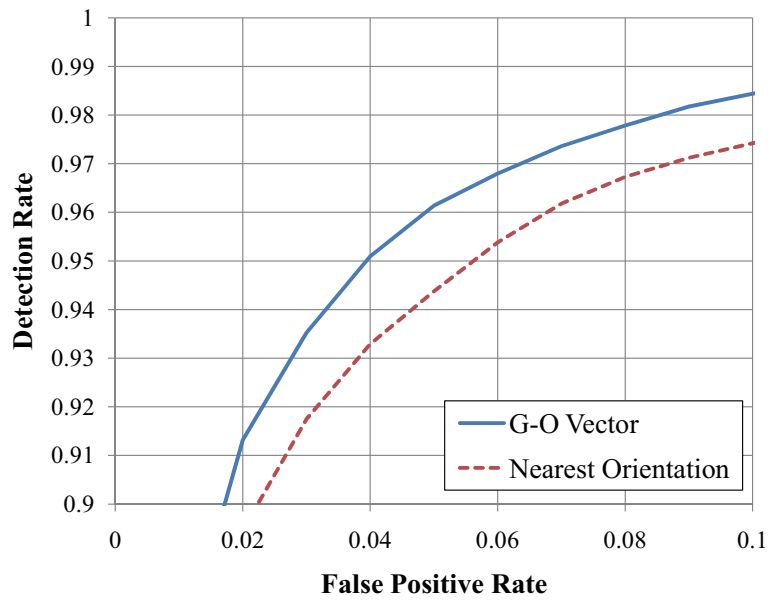


図 53 Daimler データベースの評価結果: 勾配の量子化方法

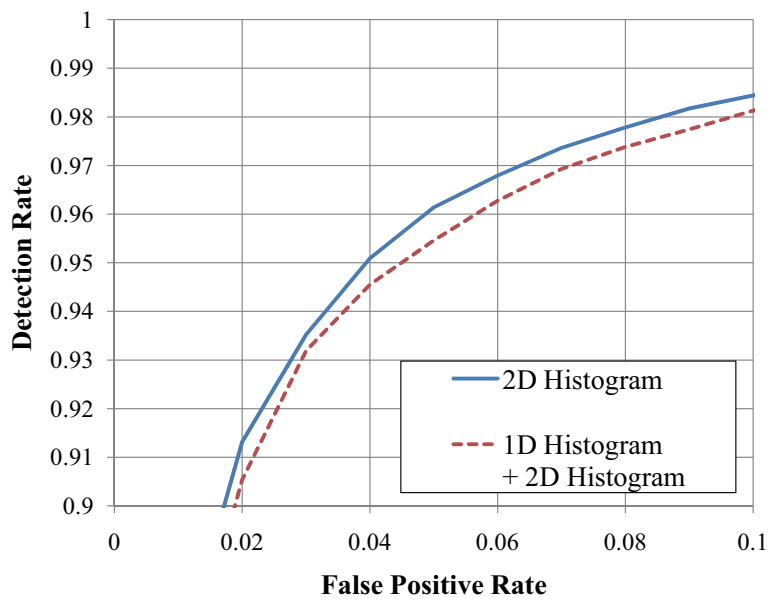


図 54 Daimler データベースの評価結果: 1次元ヒストグラムと2次元ヒストグラム

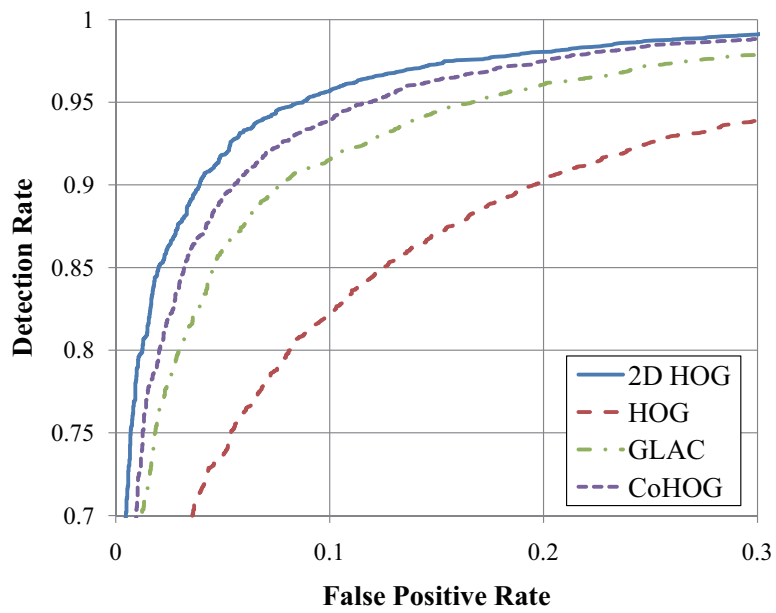


図 55 近赤外画像データベースでの評価結果

る HOG より識別性能を大きく向上でき、2次元ヒストグラムを用いた従来手法である GLAC, CoHOG より高性能であることが確認できる。

4.6 結言

本章では、歩行者検出のための特徴量として勾配方向の2次元ヒストグラムを用いる 2-Dimensional Histograms of Oriented Gradients (2D HOG) 特徴を提案した。2D HOG 特徴では、2つの画素間の位置関係を極座標表現を用いて表すことにより、様々な位置関係にある画素間の勾配の共起性を効果的に表現することが可能である。歩行者画像データベースを用いた評価によって、HOG 特徴や2次元ヒストグラムを用いる GLAC, CoHOG 特徴より高い識別性能を持つ特徴量であることを確認した。2D HOG 特徴では、次元数を大きく増大させることなく、識別性能を向上させることができるため、大量の学習サンプルデータによる学習や他の特徴量との組み合わせが可能となる。

提案した特徴量により、従来手法と比べて高い歩行者検出性能を実現することができたが、歩行者検出技術を衝突警報などを行うシステムとして実用化するためには、まだ検出性能が十分ではない。例えば、近赤外画像データベースでは、誤検出率 (False Positive Rate)0.1 で、検出率 (Detection Rate)0.95 以上となったが、警報を行う運転支援システムへ適用するためには、検出率を低減させずに誤検出率を最低限 1/10 以下にすることが要求される。したがって今後は、他の特徴との組み合わせによる性能向上について検討する予定である。また、提案手法は計算量が従来手法より増大したため、カスケード構成による識別対象の絞り込みや、特徴量の効率的な計算手法についても取り組む予定である。

5. 結論

5.1 本論文のまとめ

本論文では、自動車運転支援のための単眼カメラを用いた走行環境の認識手法として、自車両の運動を推定する手法、道路面領域を推定する手法、歩行者検出のための特徴量の生成手法について提案を行った。

自車両の運動の推定においては、特徴点追跡に基づく手法を、一般道路環境へ適用可能とするため、運動推定に利用する特徴点の選択法を導入した。そのため、移動物体の検出を行い、その検出結果を特徴点の選択に利用した。また、単眼カメラでは原理的に復元できない平行移動量の大きさ(スケール)を推定するため、道路平面の位置を推定し、それを基にスケールを求め、3次元空間での自車両の運動の6次元のパラメータをすべて推定可能とした。一般道路環境で撮影した実画像を用いた実験では、道路面にパターンがないシーン、移動物体が存在するシーンにおいて、自車両の運動を高い精度で推定できることを確認した。提案手法の推定精度は、100m程度の走行においては累積誤差が位置1m、角度 0.3° 程度であり、GPS計測値を補間して自車両位置を推定する位置推定システムのために利用できる技術であると考えられる。

道路面領域の推定においては、自車両の運動推定手法で得られる自車両の運動情報、特徴点の3次元位置情報、道路平面の位置と傾きの情報を利用し、射影変換による平面領域の検出と高さ画像から画像中の道路面領域を推定する手法を提案した。射影変換による平面領域の検出では、自車両の運動推定手法で得られる自車両の運動と道路平面の平面のパラメータを利用することにより、正しい射影変換行列を生成することができる。また、射影変換による重ね合わせにおけるしきい値を動的に設定することで、道路面のテクスチャの変化や昼夜の違いなどによる明るさの変化に対応可能である。さらに、特徴点の3次元位置から生成する高さ画像により、射影変換による平面領域検出では道路面領域であるか立体物領域であるかの判定が困難な領域の判定を可能とする。実画像を用いた実験では、昼夜ともに道路面の領域が推定できること、およびテクスチャのない壁などの立体物の誤検出を低減できることを確認した。道路面の領域が推定できると、自車

両が物理的に走行可能である領域が分かり、立体物領域との区別が可能となるとともに、車両、歩行者などの特定物体を検出するときの探索範囲の限定や道路面パターンの誤検出除去にも利用することが可能である。

歩行者検出のための特徴量においては、濃度勾配情報から勾配方向の 2 次元ヒストグラムを用いる特徴量 2-Dimensional Histograms of Oriented Gradients (2D HOG) 特徴を提案した。2D HOG 特徴では、各画素に対してその周辺領域を極座標で分割し、分割した領域に対応する位置関係ごとに 2 次元ヒストグラムを生成することで、特徴ベクトルの次元数を大きく増大させることなく、様々な位置関係にある画素間の共起性を効率的に表すことができる。歩行者画像データベースを用いた評価によって、HOG 特徴や 2 次元ヒストグラムを用いる GLAC, CoHOG 特徴より高い識別性能を持つ特徴量であることを確認した。2D HOG 特徴では、次元数が大きく増大させることなく、識別性能を向上させることができるため、大量の学習サンプルデータによる学習や他の特徴量との組み合わせが可能となる。

以上、本論文で提案した手法は、夜間や天候の環境変化への対応が行われていないものの、昼、晴れの基本的な環境では、自動車の運転支援システムのために有効に利用できる技術である。さらに、性能向上や環境変化に対するロバスト性向上により、将来の自動車運転支援システムに適用可能な基本技術となり得る。また、他のセンサと組み合わせることにより、さらなる性能向上が期待でき、より早く実用化することが可能となることも考えられる。

5.2 今後の検討課題と展望

本論文では、単眼カメラを用いた走行環境の認識技術を提案したが、実用化へ向けてさらに性能向上を図るためには検討すべき課題がある。手法を発展させるための検討課題について以下に述べる。

他のセンサとの組み合わせ

本研究では、単眼カメラのみを用いる手法を提案したが、走行環境においてロバストに認識を行うためには、他のセンサとの組み合わせが考えられる。自車両の運動の推定では、IMU と組み合わせる手法 [CLCG06, 武山 09] も

提案されており、IMU 自体の精度は十分でないが、画像による誤推定を防ぐことが可能となる。また、レーザスキャナと単眼カメラを組み合わせる手法 [山口 07] では、レーザスキャナで得られる 3 次元点を利用することにより、自車両の運動と 3 次元構造の推定の両方の推定精度を向上させている。このように、車両で利用できる他のセンサを積極的に利用し、各センサの特性を考慮した組み合わせ手法を検討することで、実環境への適用が容易になることが期待できる。

複数の画像認識技術の組み合わせ

本論文では、画像を用いた走行環境認識技術を、自車両の運動の推定技術、走行環境の構造の推定技術、特定物体の検出技術の 3 つに分類したが、それらを個々に行うのではなく、互いに結果をフィードバックしながら、あるいは同時に最適化を行って解く手法の実現が望まれる。例えば、本研究では自車両の運動推定のために、特徴点の対応関係から移動物体領域を検出したが、走行環境での移動物体である車両、二輪車、歩行者を特定物体の検出手法で検出できれば、その情報を自車両の運動推定に利用することができる。また、構造情報が得られていれば、車両検出において、高さのない領域を車両として誤検出することを防ぐことができる。

これまでに、SfM から推定される距離情報を利用してテクスチャ特徴を抽出するウィンドウサイズを制御し、画像の領域分割と物体の種類の識別を行う手法 [KYNN09] や、自車両の運動推定と歩行者検出、追跡を組み合わせ、それぞれの結果を互いに利用する方法 [ELSVG09]、物体検出と画像の各領域の面の向きの推定およびオクルージョン領域の推定を組み合わせ、それぞれの結果を利用して再び各処理を繰り返し行うことにより推定結果を向上させる方法 [HEH08] などが提案されている。

画像からの走行環境認識における今後の方向として、画素あるいは Super-Pixel [RM03, FH04] のような小領域ごとに、物体の種類、距離、動きをすべて推定することが考えられ、そのためには SfM のような幾何学的手法と物体検出で用いられるパターン認識手法を効果的に組み合わせ、シーン全体を認識する手法の実現が望まれる。

以上のように、運転支援システムへの実用化のためには検討すべき課題も多いが、今後開発、実用化するシステムには様々な種類が考えられる。例えば、できるだけ多くの車種へ展開するには、先行車両の検出や走路の検出などに特化したシステムも有用であり、一方、将来的に自律走行に近づけるためには高度な認識を行うシステムが必要である。そこで、画像を用いて要求される機能をできるだけ高速、安定に実現する技術、および高度な認識を行う先端的な技術両方の開発を進めていくことが重要である。

謝辞

本論文は、筆者が(株)豊田中央研究所 走行環境認識研究室にて行った研究をまとめたものです。本研究は、多くの方々の御協力を得て進めることができました。

はじめに、本論文の審査委員であり、筆者が博士前期課程在学時から現在に至るまで御指導頂いた奈良先端科学技術大学院大学 情報科学研究科 視覚情報メディア講座 横矢 直和教授に深く感謝致します。これまで長きにわたって温かい御指導と多くの御助言を頂いたこと、また博士後期課程へ入学し、本研究を博士論文としてまとめる機会を与えて頂いたことに心から感謝致します。そして本論文に関する有益な御助言と御鞭撻を頂いた、本論文審査委員であるインタラクティブメディア設計学講座 加藤 博一教授に厚く御礼申し上げます。本論文を作成するにあたり、多くの御助言と多大な御支援を頂いた視覚情報メディア講座 山澤 一誠准教授に感謝致します。そして有益な御助言を頂いた環境知能学講座 神原 誠之准教授、視覚情報メディア講座 佐藤 智和助教に御礼申し上げます。本論文をまとめるにあたり、御協力頂いた視覚情報メディア講座 牧田 孝嗣氏(現、産業技術総合研究所)に感謝致します。

本研究を進めるにあたり、御指導、御鞭撻頂いた(株)豊田中央研究所 安全・情報システム研究部 二宮 芳樹次長、走行環境認識研究室 内藤 貴志室長に深く感謝致します。本研究を行う機会を与えて頂き、御指導頂いたおかげで研究を進めることができました。また、博士後期課程への入学に対して御理解頂き、本論文をまとめる際には多大な御配慮を頂いたことに深く感謝致します。そして、御指導と御助言、御協力を頂いた走行環境認識研究室の諸氏に感謝致します。

最後に、本研究を進め、博士論文としてまとめるにあたり、ここに紹介できなかった視覚情報メディア講座および豊田中央研究所の先輩、後輩の方々、友人からも多くの御助言、御支援を頂きましたこと厚く御礼申し上げます。

参考文献

- [AKB08] M. Agrawal, K. Konolige, and M. R. Blas. CenSurE: Center surround extremas for realtime feature detection and matching. In *Proc. European Conf. on Computer Vision*, pp. 102–115, Oct. 2008.
- [BB98] M. Bertozzi and A. Broggi. GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Trans. Image Processing*, Vol. 7, No. 1, pp. 62–81, Jan. 1998.
- [BBFN00] M. Bertozzi, A. Broggi, A. Fascioli, and S. Nichele. Stereo vision-based vehicle detection. In *Proc. IEEE Intelligent Vehicles Symposium*, pp. 39–44, Oct. 2000.
- [BETG08] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. SURF: Speeded-up robust features. *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346–359, Jun. 2008.
- [BKY+08] O. Barinova, V. Konushin, A. Yakubenko, K. Lee, H. Lim, and A. Konushin. Fast automatic single-view 3-d reconstruction of urban scenes. In *Proc. the 10th European Conf. on Computer Vision, Part II*, pp. 100–113, Oct. 2008.
- [BPDB08] B. Babenko, Z. Tu P. Dollár, and S. Belongie. Simultaneous learning and alignment: Multi-instance and multi-pose learning. In *Faces in 'Real-Life' Images*, Oct. 2008.
- [CLCG06] N. Cornelis, B. Leibe, K. Cornelis, and L. Van Gool. 3D city modeling using cognitive loops. In *Proc. the Third Int. Symposium on 3D Data Processing*, pp. 9–16, Jun. 2006.
- [CYNN09] H. Cao, K. Yamaguchi, T. Naito, and Y. Ninomiya. Pedestrian recognition using second-order hog feature. In *Proc. 9th Asian Conf. on Computer Vision, Part II*, pp. 628–634, Sep. 2009.

- [CZ07] O. Chum and A. Zisserman. An exemplar model for learning object classes. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, June 2007.
- [Dai08] DaimlerAG. Speed Limit Assist. <http://media.daimler.com/dcmedia/0-921-658892-1-1147690-1-0-0-0-0-1-11702-854934-0-1-0-0-0-0-0.html?TS=1281678107688>, 2008.
- [Dal06] N. Dalal. *Finding people in images and videos*. PhD thesis, Institut National Polytechnique de Grenoble, Jul. 2006.
- [Dav03] A. J. Davison. Real-time simultaneous localisation and mapping with a single camera. In *Proc. the Ninth IEEE Int. Conf. on Computer Vision*, pp. 1403–1410, Oct. 2003.
- [DdFMR00] A. Doucet, N. de Freitas, K. P. Murphy, and S. J. Russell. Rao-blackwellised particle filtering for dynamic bayesian networks. In *Proc. the 16th Conf. on Uncertainty in Artificial Intelligence*, pp. 176–183, Jul. 2000.
- [DT05] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 886–893, Jun. 2005.
- [DTS06] N. Dalal, B. Triggs, and C. Schmid. Human detection using oriented histograms of flow and appearance. In *Proc. European Conf. on Computer Vision*, pp. 7–13, May 2006.
- [DWSP09] P. Dollár, C. Wojek, B. Schiele, and P. Perona. Pedestrian detection: A benchmark. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 304–311, Jun. 2009.
- [EG09] M. Enzweiler and D. M. Gavrila. Monocular pedestrian detection:

- Survey and experiments. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 31, No. 12, pp. 2179–2195, Dec. 2009.
- [EGW⁺] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge 2009 (voc2009) results. <http://www.pascal-network.org/challenges/VOC/voc2009/workshop/index.html>.
- [ELSVG09] A. Ess, B. Leibe, K. Schindler, and L. van Gool. Robust multiperson tracking from a mobile platform. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 31, No. 10, pp. 1831–1846, Oct. 2009.
- [FB81] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. of the ACM*, Vol. 24, No. 6, pp. 381–395, Jun. 1981.
- [FCH⁺08] R.E. Fan, K.W. Chang, C.J. Hsieh, X.R. Wang, and C.J. Lin. L1-BLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, Vol. 9, pp. 1871–1874, Aug. 2008.
- [FGM10] P. Felzenszwalb, R. Girshick, and D. McAllester. Cascade object detection with deformable part models. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 2241–2248, Jun. 2010.
- [FGMR10] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 32, No. 9, pp. 1627–1645, Sep. 2010.
- [FH04] P. F. Felzenszwaalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *Int. Journal of Computer Vision*, Vol. 59, No. 2, pp. 167–181, Sep. 2004.

- [FL88] O. Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. Technical Report 856, INRIA, Jun. 1988.
- [FS97] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, Vol. 55, No. 1, pp. 119–139, Aug. 1997.
- [GHTC03] X. S. Gao, X. R. Hou, J. Tang, and H. F. Cheng. Complete solution classification for the perspective-three-point problem. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 25, No. 8, pp. 930–943, Aug. 2003.
- [GMM83] D. Gavrilu, P. Marchal, and M. M. Meinecke. SAVE-U, Derivable 1-A: Vulnerable road user scenario analysis. Technical report, Information Soc. Technology Programme of the EU, Feb. 1983.
- [Har97] R. I. Hartley. In defense of the eight-point algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, pp. 580–593, Jun. 1997.
- [HEH07] D. Hoiem, A. A. Efros, and M. Hebert. Recovering surface layout from an image. *Int. Journal of Computer Vision*, Vol. 75, No. 1, pp. 151–172, Oct. 2007.
- [HEH08] D. Hoiem, A. A. Efros, and M. Hebert. Closing the loop on scene interpretation. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Jul. 2008.
- [HHH07] G. Hager, M. Hebert, and S. Hutchinson. Special issue on vision and robotics. *Int. Journal of Computer Vision*, Vol. 74, No. 3, pp. 217–218, Sep. 2007.

- [HJ88] B. Horn and E. Weldon Jr. A direct method for recovering motion. *Int. Journal of Computer Vision*, Vol. 2, No. 1, pp. 51–76, Jun. 1988.
- [HS88] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conf.*, pp. 147–151, Aug. 1988.
- [HS97] R. I. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, Vol. 68, No. 2, pp. 146–157, Nov. 1997.
- [HZ04] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [IB98] M. Isard and A. Blake. CONDENSATION – conditional density propagation for visual tracking. *Int. Journal of Computer Vision*, Vol. 29, No. 1, pp. 5–28, Aug. 1998.
- [Jon02] W. Jones. Building safer cars. *IEEE Spectrum*, Vol. 39, No. 1, pp. 82–85, Jan. 2002.
- [KAS07] K. Konolige, M. Agrawal, and J. Solà. Large scale visual odometry for rough terrain. In *Proc. Int. Symposium on Research in Robotics (ISRR)*, Nov. 2007.
- [KC08] T. K. Kim and R. Cipolla. MCBoost: Multiple classifier boosting for perceptual co-clustering of images and visual features. In *Advances in Neural Information Processing Systems (NIPS)*, Dec. 2008.
- [KK03] Q. Ke and T. Kanade. Transforming camera geometry to a virtual downward-looking camera: Robust ego-motion estimation and ground-layer detection. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 390–397, Jun. 2003.
- [KM07] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *Proc. Sixth IEEE and ACM Int. Symposium on Mixed and Augmented Reality (ISMAR'07)*, Nov. 2007.

- [KO08] T. Kobayashi and N. Otsu. Image feature extraction using gradient local auto-correlations. In *Proc. European Conf. on Computer Vision, Part I*, pp. 346–358, Oct. 2008.
- [KSN08] K. Kanatani, Y. Sugaya, and H. Niitsuma. Triangulation from two views revisited: Hartley-Sturm vs. optimal correction. In *Proc. the 19th British Machine Vision Conference*, pp. 173–182, Sep. 2008.
- [Kue91] A. Kuehnle. Symmetry-based recognition for vehicle rears. *Pattern Recognition Letters*, Vol. 12, No. 4, pp. 249–258, Apr. 1991.
- [KYNN] Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya. Multiband image segmentation and object recognition for understanding road scenes. *IEEE Trans. Intelligent Transportation Systems*, (to appear).
- [KYNN09] Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya. Texture segmentation of road environment scene using SfM module and HLAC features. *IPSJ Trans. Computer Vision and Applications*, Vol. 1, pp. 220–230, Sep. 2009.
- [Lap09] I. Laptev. Improving object detection with boosted histograms. *Image and Vision Computing*, Vol. 27, No. 5, pp. 535–544, Apr. 2009.
- [LBH09] C. H. Lampert, M. B. Blaschko, and T. Hofmann. Efficient subwindow search: A branch and bound framework for object localization. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 31, No. 12, pp. 2129–2142, Dec. 2009.
- [Lex06] Lexus. プリクラッシュセーフティシステム. <http://lexus.jp/models/ls600h/safety/safety/precrash.html>, 2006.
- [LHJ07] Y. Li, K. He, and P. Jia. Road markers recognition based on shape information. In *Proc. IEEE Intelligent Vehicles Symposium*, pp.

117–122, Jun. 2007.

- [LHP80] H. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proc. the Royal Soc. of London B*, Vol. 208, No. 1173, pp. 385–397, Jul. 1980.
- [LK81] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Int. Conf. on Artificial Intelligence*, pp. 674–679, Aug. 1981.
- [LKK04] F. Lindner, U. Kressel, and S. Kaelberer. Robust recognition of traffic signels. In *Proc. IEEE Intelligent Vehicles Symposium*, pp. 49–53, Jun. 2004.
- [LLS08] B. Leibe, A. Leonardis, and B. Schiele. Robust object detection by interleaving categorization and segmentation. *Int. Journal of Computer Vision*, Vol. 77, No. 1-3, pp. 259–289, May 2008.
- [Low04] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. Journal of Computer Vision*, Vol. 60, No. 2, pp. 91–110, Nov. 2004.
- [LSP06] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 2169–2178, Oct. 2006.
- [MACH96] N. Matthews, P. An, D. Charnley, and C. Harris. Vehicle detection and recognition in greyscale imagery. *Control Engineering Practice*, Vol. 4, No. 4, pp. 473–479, Apr. 1996.
- [Mar63] D.W. Marquardt. An algorithm for least-squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics*, Vol. 11, No. 2, pp. 431–441, Jun. 1963.

- [MBM08] S. Maji, A.C. Berg, and J. Malik. Classification using intersection kernel support vector machines is efficient. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 1–8, Jun. 2008.
- [MG06] S. Munder and D. M. Gavrila. An experimental study on pedestrian classification. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 28, No. 11, pp. 1863–1868, Nov. 2006.
- [MLD⁺06] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd. Real time localization and 3D reconstruction. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 363–370, Jun. 2006.
- [Nis04] D. Nistér. An efficient solution to the five-point relative pose problem. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 26, No. 6, pp. 756–770, Jun. 2004.
- [NNB04] D. Nistér, O. Naroditsky, and J. Bergen. Visual odometry. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 652–659, Jun. 2004.
- [OTFO03] R. Okada, Y. Taniguchi, K. Furukawa, and K. Onoguchi. Obstacle detection using projective invariant and vanishing lines. In *Proc. IEEE Int. Conf. on Computer Vision*, pp. 330–337, Oct. 2003.
- [PP00] C. Papageorgiou and T. Poggio. A trainable system for object detection. *Int. Journal of Computer Vision*, Vol. 38, No. 1, pp. 15–33, Jun. 2000.
- [RD06] E. Rosten and T. Drummond. Machine learning for high-speed corner detection. In *Proc. European Conf. on Computer Vision*, pp. 430–443, May 2006.

- [RM03] X. Ren and J. Malik. Learning a classification model for segmentation. In *Proc. IEEE Int. Conf. on Computer Vision*, pp. 10–17, Oct. 2003.
- [SBM02] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection using gabor filters and support vector machines. In *Proc. Int. Conf. on Digital Signal Processing*, pp. 1019–1022, Jul. 2002.
- [SBM06] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection: A review. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 28, No. 5, pp. 694–711, May 2006.
- [SEL07] R. Sim, P. Elinas, and J. J. Little. A study of the rao-blackwellised particle filter for efficient and accurate vision-based slam. *Int. Journal of Computer Vision*, Vol. 74, No. 3, pp. 303–318, Sep. 2007.
- [SK99] T. Suzuki and T. Kanade. Measurement of vehicle motion and orientation using optical flow. In *Proc. IEEE Int. Conf. on Intelligent Transportation Systems*, pp. 25–30, Oct. 1999.
- [SMS00] P. Stein, O. Mano, and A. Shashua. A robust method for computing vehicle ego-motion. In *Proc. IEEE Intelligent Vehicle Symposium*, pp. 362–368, Oct. 2000.
- [SO06] A. Seki and M. Okutomi. Ego-motion estimation by matching de-warped road regions using stereo images. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pp. 901–907, May 2006.
- [Sri02] N. Srinivasa. A vision-based vehicle detection and tracking method for forward collision warning. In *Proc. IEEE Intelligent Vehicles Symposium*, pp. 626–631, June 2002.
- [SSN09] A. Saxena, M. Sun, and A. Y. Ng. Make3D: Learning 3D scene structure from a single still image. *IEEE Trans. Pattern Analysis*

and Machine Intelligence, Vol. 31, No. 5, pp. 824–840, May 2009.

- [ST94] J. Shi and C. Tomasi. Good features to track. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 593–600, Jun. 1994.
- [TMHF99] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon. Bundle adjustment - a modern synthesis. In *Proc. the Int. Workshop on Vision Algorithms: Theory and Practice*, pp. 298–372, Sep. 1999.
- [Vap98] V.N. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, 1998.
- [VGVZ09] A. Vedaldi, V. Galshan, M. Varma, and A. Zisserman. Multiple kernels for object detection. In *Proc. Int. Conf. on Computer Vision*, Sep. 2009.
- [VJ04] P. Viola and M. J. Jones. Robust real-time face detection. *Int. Journal of Computer Vision*, Vol. 57, No. 2, pp. 137–154, May 2004.
- [WHY09] X. Wang, T. X. Han, and S. Yan. An hog-lbp human detector with partial occlusion handling. In *Proc. IEEE Int. Conf. on Computer Vision*, pp. 32–39, Sep. 2009.
- [WIY09] T. Watanabe, S. Ito, and K. Yokoi. Co-occurrence histograms of oriented gradients for pedestrian detection. In *Proc. 3rd Pacific Rim Symposium on Advances in Image and Video Technology*, pp. 37–47, Jan. 2009.
- [WMSS10] S. Walk, N. Majer, K. Schindler, and B. Schiele. New features and insights for pedestrian detection. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 1030–1037, Jun. 2010.

- [WN05] A. Watanabe and M. Nishida. Lane detection for a steering assistance system. In *Proc. IEEE Intelligent Vehicle Symposium*, pp. 158–163, Jun. 2005.
- [WS08] C. Wojek and B. Schiele. A performance evaluation of single and multi-feature people detection. In *Proc. 30th DAGM symposium*, pp. 82–91, Jun. 2008.
- [WWS09] C. Wojek, S. Walk, and B. Schiele. Multi-cue onboard pedestrian detection. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 794–801, Jun. 2009.
- [YJ97] B. Yu and A. K. Jain. Lane boundary detection using a multiresolution hough transform. In *Int. Conf. on Image Processing*, Vol. 3, pp. 26–29, Oct. 1997.
- [ZMLS07] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *Int. Journal of Computer Vision*, Vol. 73, No. 2, pp. 213–238, Jun. 2007.
- [ZYCA06] Q. Zhu, M. C. Yeh, K. T. Cheng, and S. Avidan. Fast human detection using a cascade of histograms of oriented gradients. In *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 1491–1498, June 2006.
- [内村 00] 内村圭一, 脇山慎也, 藤野麻衣子. 限定色表示を用いた円形道路標識の抽出. 電子情報通信学会論文誌 (D-II), Vol. J83-D-II, No. 2, pp. 855–858, Feb. 2000.
- [奥富 02] 奥富正敏, 中野勝之, 丸山純一, 原智章. ステレオ動画像を用いた視覚誘導のための平坦部の連続推定. 情報処理学会論文誌, Vol. 43, No. 4, pp. 1061–1069, Apr. 2002.

- [木村 06] 木村文香, 高橋友和, 目加田慶人, 井手一郎, 村瀬洋. 安全運転支援のための多様な撮影環境における信号認識. 画像の認識・理解シンポジウム (MIRU2006) 講演論文集, pp. 618–623, Jul. 2006.
- [木村 07] 木村好克, 山口晃一郎, 加藤武男, 太田充彦, 二宮芳樹. 安全のための画像センサ技術. 映像情報メディア学会誌, Vol. 61, No. 12, pp. 1697–1700, Dec. 2007.
- [警察 10] 警察庁交通局. 平成 21 年中の交通死亡事故の特徴及び道路交通法違反取締り状況について, Jan. 2010.
- [国土] 国土交通省先進安全自動車 (ASV) 推進検討会. ASV(先進安全自動車). <http://www.mlit.go.jp/jidosha/anzen/01asv/>.
- [小島 07] 小島祥子, 高橋新, 城殿清澄, 山口晃一郎, 二宮芳樹. 車載外界監視センサを用いた高精度自車運動推定および自車位置推定. 電子情報通信学会 技術研究報告, ITS2007-11, Jun. 2007.
- [実吉 97] 実吉敬二, 埜圭二, 十川能之, 荒井一真. ステレオ画像を用いた運転支援のための前方状況認識システム. 電子情報通信学会 技術研究報告, PRMU97-41, May 1997.
- [高木 09] 高木雅成, 藤吉弘亘. SIFT 特徴量を用いた交通道路標識認識. 電気学会論文誌 (C), Vol. 129, No. 5, pp. 824–831, May 2009.
- [武山 09] 武山洪二郎, 小島祥子, 山口晃一郎, 二宮芳樹. 移動点除去による単眼カメラを用いた運動推定の高精度化. 電子情報通信学会総合大会講演論文集, No. D-12-53, Mar. 2009.
- [トヨタ 04] トヨタ自動車. レーンキーピングアシスト. <http://www2.toyota.co.jp/jp/tech/safety/technologies/active/lka.html>, 2004.

- [二宮 03] 二宮芳樹, 高橋新, 太田充彦. 高速パターン照合手法を利用したレーン認識システム. 電子情報通信学会論文誌 (D-II), Vol. J86-D-II, No. 5, pp. 625–632, May 2003.
- [野田 09] 野田雅文, 高橋友和, 出口大輔, 井手一郎, 村瀬洋, 小島祥子, 内藤貴志. 車両の姿勢・速度情報を利用した車載カメラ画像からの路面標示認識. 画像の認識・理解シンポジウム (MIRU2009) 講演論文集, pp. 614–621, Jul. 2009.
- [富士 08] 富士重工業. Eyesight. <http://www.subaru.jp/eyesight/eyesight/>, 2008.
- [本田 02] 本田技研工業. レーンキープアシストシステム. <http://www.honda.co.jp/safety/technology/car-safety/main-technology/hids/index.html>, 2002.
- [三谷 06] 三谷哲雄, 山中英生. 無信号交差点における一時停止支援システムの警告判定法の開発. 第 5 回 ITS シンポジウム論文集, pp. 391–398, Dec. 2006.
- [山内 09] 山内悠嗣, 山下隆義, 藤吉弘亘. Boosting に基づく特徴量の共起表現による人検出. 電子情報通信学会論文誌 (D), Vol. J92-D, No. 8, pp. 1125–1134, Aug. 2009.
- [山口 07] 山口晃一郎, 加藤武男, 二宮芳樹. 単眼カメラとレーザーレーダによる自車両の運動推定と 3 次元構造推定. 画像の認識・理解シンポジウム (MIRU2007) 講演論文集, pp. 696–701, Jul. 2007.
- [山口 08] 山口晃一郎. 車載カメラによる運動推定及び周辺構造推定技術の現状. 第 14 回画像センシングシンポジウム (SSII08) 講演論文集, OR1-02, Jun. 2008.

研究業績

学術論文誌

1. 山口晃一郎, 山澤一誠, 竹村治雄, 横矢直和, “全方位動画像を用いた両眼ステレオ画像の実時間生成によるテレプレゼンス,” 電子情報通信学会論文誌 (D-II), Vol. J84-D-II, No. 9, pp. 2048-2057, Sep. 2001.
2. K. Yamaguchi, Y. Aburakawa, Y. Nakayama, and T. Otsu, “Frequency Channel Allocation Scheme in Autonomous Distributed Control Wireless Base Station Network,” IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences, Vol. E86-A, No. 7, pp. 1627-1635, Jul. 2003.
3. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, “Texture Segmentation of Road Environment Scene Using SfM Module and HLAC Features,” IPSJ Trans. Computer Vision and Applications, Vol. 1, pp. 220-230, Sep. 2009.
4. 山口晃一郎, 加藤武男, 二宮芳樹, “車載単眼カメラを用いた自車両の運動推定,” 電気学会論文誌 C, Vol. 129, No. 12, pp. 2213-2221, Dec. 2009. (第2章に関連)
5. 浅井俊弘, 山口晃一郎, 小島祥子, 内藤貴志, 二宮芳樹, “構造物の知識に基づく車載カメラ画像からの走行環境の三次元直線復元,” 電子情報通信学会論文誌 (D), Vol. J93-D, No. 7, pp. 1236-1247, Jul. 2010.
6. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, “Integrating Motion and Segmentation for Road Scene Labeling,” IPSJ Trans. Computer Vision and Applications, Vol. 3, Oct. 2010. (採録予定)
7. 山口晃一郎, 内藤貴志, “歩行者検出のための勾配方向の2次元ヒストグラム特徴,” 電子情報通信学会論文誌 (D). (採録予定) (第4章に関連)
8. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, “Multiband Image Segmentation and Object Recognition for Understanding Road Scenes,” IEEE Trans. Intelligent Transportation Systems. (採録予定)

レター

1. H. Cao, K. Yamaguchi, M. Ohta, T. Naito, and Y. Ninomiya, "Feature Interaction Descriptor for Pedestrian Detection," *IEICE Trans. Information and Systems*, Vol. E93-D, No. 9, pp. 2656-2659, Sep. 2010.

国際会議

1. K. Yamaguchi, K. Yamazawa, H. Takemura, and N. Yokoya, "Real-time Generation and Presentation of View-dependent Binocular Stereo Images Using a Sequence of Omnidirectional Images," *Proc. 15th Int. Conf. on Pattern Recognition (ICPR2000)*, Vol. 4, pp. 589-593, Sep. 2000.
2. K. Yamaguchi, K. Yamazawa, H. Takemura, and N. Yokoya, "Image-based Construction of 3D Virtual Environments from a Sequence of Omnidirectional Images," *Proc. IEEE Int. Conf. on Industrial Electronics, Control and Instrumentation (IECON-2000)*, pp. 2565-2570, Oct. 2000.
3. K. Yamaguchi, Y. Aburakawa, Y. Nakayama, and T. Otsu, "Frequency Channel Allocation Scheme Using Broadcast Signals in Wireless Multi-hop Network," *Proc. 8th Asia Pacific Conf. on Communications (APCC2002)*, pp. 31-34, Sep. 2002.
4. K. Yamaguchi, T. Kato, and Y. Ninomiya, "Moving Obstacle Detection Using Monocular Vision," *Proc. of IEEE Intelligent Vehicle Symp. (IV2006)*, pp.288-293, Jun. 2006. (第2章に関連)
5. K. Yamaguchi, T. Kato, and Y. Ninomiya, "Vehicle Ego-Motion Estimation and Moving Obstacle Detection Using a Monocular Camera," *Proc. 18th Int. Conf. on Pattern Recognition (ICPR2006)*, Vol. IV, pp. 610-613. Aug. 2006. (第2章に関連)
6. S. Ikeda, T. Sato, K. Yamaguchi, and N. Yokoya, "Construction of Feature Landmark Database Using Omnidirectional Videos and GPS Positions," *Proc. 6th Int. Conf. on 3D Digital Imaging and Modeling (3DIM)*, pp. 249-256, Aug. 2007.

7. T. Asai, K. Yamaguchi, Y. Kojima, T. Naito, and Y. Ninomiya, "3D Line Reconstruction of a Road Environment Using an In-Vehicle Camera," Proc. 4th Int. Symp. on Visual Computing, Part II, pp. 897-904, Dec. 2008.
8. K. Yamaguchi, A. Watanabe, T. Naito, and Y. Ninomiya, "Road Region Estimation Using a Sequence of Monocular Images," Proc. 19th Int. Conf. on Pattern Recognition (ICPR2008), Dec. 2008. (第3章に関連)
9. H. Cao, K. Yamaguchi, T. Naito, and Y. Ninomiya, "Pedestrian Recognition Using Second-Order HOG Feature," Proc. 9th Asian Conf. on Computer Vision (ACCV2009), pp. 628-634, Sep. 2009.
10. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, "Road Scene Labelling Using SfM Module and 3D Bag of Textons," Proc. 2nd Int. IEEE Workshop on 3D Representation for Recognition (3dRR-09), Sep. 2009.

研究会・シンポジウム

1. 山口晃一郎, 山澤一誠, 竹村治雄, 横矢直和, "全方位動画像を用いた任意方向両眼ステレオ画像の実時間生成と提示," 電子情報通信学会 技術研究報告, PRMU99-159, Nov. 1999.
2. 山口晃一郎, 山澤一誠, 竹村治雄, 横矢直和, "全方位動画像を用いた両眼ステレオ画像の実時間生成システム," 画像の認識・理解シンポジウム (MIRU2000) 講演論文集, Vol. I, pp. 45-50, Jul. 2000.
3. 山口晃一郎, 油川雄司, 中山雄二, 大津徹, "無線マルチホップ基地局ネットワークにおける報知信号を用いた帯域割り当て法," 電子情報通信学会 技術研究報告, RCS2002-64, May 2002.
4. 山口晃一郎, 加藤武男, 二宮芳樹, "車載単眼カメラによる車両前方の障害物検出," 情報処理学会 研究報告, CVIM-151-10, Nov. 2005. (第2章に関連)
5. 山口晃一郎, 加藤武男, 二宮芳樹, "単眼カメラとレーザレーダを用いた走行環境の3次元構造推定," 第11回パターン計測シンポジウム講演論文集, pp. 43-48, Nov. 2006.
6. 木村好克, 山口晃一郎, 加藤武男, 太田充彦, 二宮芳樹, "ステレオビジョン

- による立体物検出技術,” 動的画像処理実利用化ワークショップ (DIA2007), O4-1, Mar. 2007.
7. 小島祥子, 高橋新, 城殿清澄, 山口晃一郎, 二宮芳樹, “車載外界監視センサを用いた高精度自車運動推定および自車位置推定,” 電子情報通信学会 技術研究報告, ITS2007-11, Jul. 2007.
 8. 山口晃一郎, 加藤武男, 二宮芳樹, “単眼カメラとレーザレーダによる自車両の運動推定と3次元構造推定,” 画像の認識・理解シンポジウム (MIRU2007) 講演論文集, pp.696-701, Jul. 2007.
 9. 池田聖, 佐藤智和, 山口晃一郎, 横矢直和, “全方位動画像とGPS位置情報からのランドマークデータベースの構築,” 画像の認識・理解シンポジウム (MIRU2007) 講演論文集, pp. 1129-1134, Jul. 2007.
 10. 姜有宣, 山口晃一郎, 内藤貴志, 二宮芳樹, “道路環境におけるSfMを用いた3Dテクスチャセグメンテーション,” 情報処理学会 研究報告, 2007-CVIM-161(24), Jan. 2008.
 11. 浅井俊弘, 小島祥子, 山口晃一郎, 二宮芳樹, “車載単眼カメラを用いた走行環境の直線構造の三次元位置推定,” 電子情報通信学会 技術研究報告, PRMU2007-231, Feb. 2008.
 12. 山口晃一郎, “車載カメラによる運動推定及び周辺構造推定技術の現状,” 第14回画像センシングシンポジウム (SSII08), OR1-02, Jun. 2008.
 13. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, “3D Image Segmentation for Road Environment Scenes Using SfM Module and HLAC Features,” 画像の認識・理解シンポジウム (MIRU2008) 講演論文集, pp. 317-324, Jul. 2008.
 14. 浅井俊弘, 山口晃一郎, 小島祥子, 内藤貴志, 二宮芳樹, “車載カメラを用いた走行環境の直線構造の三次元復元,” 画像の認識・理解シンポジウム (MIRU2008) 講演論文集, pp. 1088-1091, Jul. 2008.
 15. 山口晃一郎, 渡邊章弘, 内藤貴志, 二宮芳樹, “車載単眼カメラを用いた道路面領域の推定,” 画像の認識・理解シンポジウム (MIRU2008) 講演論文集, pp. 1318-1323, Jul. 2008. (第3章に関連)

16. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, "Segmentation and Object Recognition in Road Scenes Using Multi-scale Semantic Bag of Textons Method," 電子情報通信学会 技術研究報告, PRMU2009-41, Jun. 2009.
17. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, "Integration of Motion and Segmentation for Road Scene Labeling Using Near Infrared and Color Image," 画像の認識・理解シンポジウム (MIRU2009) 講演論文集, pp. 245-251, Jul. 2009.
18. 浅井俊弘, 山口晃一郎, 小島祥子, 内藤貴志, 二宮芳樹, "構造物の知識に基づく車載単眼カメラを用いた走行環境の三次元直線復元," 画像の認識・理解シンポジウム (MIRU2009) 講演論文集, pp. 1578-1585, Jul. 2009.
19. 亀田能成, 藤吉弘亘, 羽下哲司, 山口晃一郎, 中村裕一, "'First Person Vision'への挑戦," 電子情報通信学会 技術研究報告, PRMU2010-7, May 2010.

全国大会

1. 山口晃一郎, 山澤一誠, 竹村治雄, 横矢直和, "全方位動画像を用いた任意方向の両眼ステレオ画像の生成," 情報処理学会全国大会予稿集, No. 2M-05, Sep. 1999.
2. 山口晃一郎, 油川雄司, 中山雄二, 大津徹, "ワイヤレスマルチホップネットワークにおけるシステム所要周波数帯域の評価," 電子情報通信学会ソサイエティ大会講演論文集, No. B-5-232, Sep. 2001.
3. 山口晃一郎, 油川雄司, 中山雄二, 大津徹, "無線マルチホップ基地局網における報知信号を用いた帯域制御法," 電子情報通信学会総合大会講演論文集, No. B-5-271, Mar. 2002.
4. 山口晃一郎, 油川雄司, 中山雄二, 大津徹, "無線マルチホップネットワークにおけるトラフィック変動を考慮した動的帯域割り当ての検討," 電子情報通信学会ソサイエティ大会講演論文集, No. B-5-176, Sep. 2002.
5. 山口晃一郎, 油川雄司, 中山雄二, 大津徹, "無線マルチホップネットワーク用帯域制御における変調方式適応選択の効果," 電子情報通信学会総合大会講演論文集, No. B-5-304, Mar. 2003.

6. K. Yamaguchi, T. Kato, and Y. Ninomiya, "Estimation of the 3D Structure of Road Scenes from Monocular Images and Range Data," Proc. of SICE Annual Conf. 2007, pp. 1465-1470, Sep. 2007.
7. 山口晃一郎, 渡邊章弘, 内藤貴志, 二宮芳樹, "単眼カメラによる道路面領域の検出," 電子情報通信学会総合大会講演論文集, No. D-12-88, Mar. 2008. (第3章に関連)
8. 武山洪二郎, 小島祥子, 山口晃一郎, 二宮芳樹, "移動点除去による単眼カメラを用いた運動推定の高精度化," 電子情報通信学会総合大会講演論文集, No. D-12-53, Mar. 2009.
9. 武富貴史, 佐藤智和, 山口晃一郎, 横矢直和, "AR カーナビゲーションのためのランドマークデータベースを用いた高精度な自車位置・姿勢の推定," 電気学会 電子・情報・システム部門大会 講演論文集, pp. 917-920, Sep. 2009.
10. Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, "Road Scene Segmentation and Recognition Using Hierarchical Bag of Textons Method," 日本ロボット学会学術講演会論文集, JSAE20094538, Sep. 2009.

解説

1. 木村好克, 山口晃一郎, 加藤武男, 太田充彦, 二宮芳樹, "安全のための画像センサ技術," 映像情報メディア学会誌, Vol. 61, No. 12, pp. 1697-1700, Dec. 2007.
2. 浅井俊弘, 山口晃一郎, 小島祥子, 二宮芳樹, "車載単眼カメラを用いた走行環境の直線構造の三次元復元—走行環境の三次元地図生成に向けて—," 画像ラボ, Vol. 19, No. 8, pp. 6-10, Aug. 2008.

表彰

1. IEEE Intelligent Vehicle Symposium, Outstanding Papers in Interactive Sessions Award, 2006年6月.
2. 第11回パターン計測シンポジウム, 優秀論文賞, 2006年11月.

付録

A. 8点アルゴリズム

撮影位置の異なる2枚の画像間で8点以上の対応点 $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\} (i = 1, \dots, n, n \geq 8)$ が与えられると、8点アルゴリズムにより、以下に示すエピポーラ拘束式を満たす基礎行列 (Fundamental Matrix) \mathbf{F} を計算することができる。

$$\mathbf{x}'_i{}^T \mathbf{F} \mathbf{x}_i = 0. \quad (36)$$

なお、対応点の画像座標は同次座標ベクトルとして表されており、 $\mathbf{x}_i = (x_i, y_i, 1)^T$ 、 $\mathbf{x}'_i = (x'_i, y'_i, 1)^T$ である。

8点アルゴリズムの処理は以下の通りである。

(1) 対応点の画像座標の正規化

対応点 $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}, i = 1, \dots, n$ のそれぞれの画像での画像座標を次の式により変換する。

$$\hat{\mathbf{x}}_i = \mathbf{T} \mathbf{x}_i, \quad \hat{\mathbf{x}}'_i = \mathbf{T}' \mathbf{x}'_i \quad (37)$$

\mathbf{T}, \mathbf{T}' は、それぞれの画像において、対応点の画像座標の中心が原点、原点からの平均距離が $\sqrt{2}$ になるように変換する行列である。

(2) 正規化対応点に関する基礎行列の計算

正規化された対応点集合 $\{\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i\}$ に対する基礎行列 $\hat{\mathbf{F}}$ を次のようにして求める。 $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i, 1)^T$ 、 $\hat{\mathbf{x}}'_i = (\hat{x}'_i, \hat{y}'_i, 1)^T$ で、求める基礎行列 $\hat{\mathbf{F}}$ の u 行 v 列の要素を \hat{F}_{uv} で表すとすると、 $\hat{\mathbf{x}}'_i{}^T \hat{\mathbf{F}} \hat{\mathbf{x}}_i = 0$ より、1組の対応点に対して以下の式が成立する。

$$\hat{x}_i \hat{x}'_i \hat{F}_{11} + \hat{y}_i \hat{x}'_i \hat{F}_{12} + \hat{x}'_i \hat{F}_{13} + \hat{x}_i \hat{y}'_i \hat{F}_{21} + \hat{y}_i \hat{y}'_i \hat{F}_{22} + \hat{y}'_i \hat{F}_{23} + \hat{x}_i \hat{F}_{31} + \hat{y}_i \hat{F}_{32} + \hat{F}_{33} = 0. \quad (38)$$

すべての対応点に対する上記の式をまとめると、以下の式が得られる。

$$\mathbf{A} \hat{\mathbf{f}} = 0. \quad (39)$$

ただし,

$$\mathbf{A} = \begin{bmatrix} \hat{x}_1 \hat{x}'_1 & \hat{y}_1 \hat{x}'_1 & \hat{x}'_1 & \hat{x}_1 \hat{y}'_1 & \hat{y}_1 \hat{y}'_1 & \hat{y}'_1 & \hat{x}_1 & \hat{y}_1 & 1 \\ \hat{x}_2 \hat{x}'_2 & \hat{y}_2 \hat{x}'_2 & \hat{x}'_2 & \hat{x}_2 \hat{y}'_2 & \hat{y}_2 \hat{y}'_2 & \hat{y}'_2 & \hat{x}_2 & \hat{y}_2 & 1 \\ & & \vdots & & & & & & \\ \hat{x}_n \hat{x}'_n & \hat{y}_n \hat{x}'_n & \hat{x}'_n & \hat{x}_n \hat{y}'_n & \hat{y}_n \hat{y}'_n & \hat{y}'_n & \hat{x}_n & \hat{y}_n & 1 \end{bmatrix}, \quad (40)$$

$$\hat{\mathbf{f}} = (\hat{F}_{11}, \hat{F}_{12}, \hat{F}_{13}, \hat{F}_{21}, \hat{F}_{22}, \hat{F}_{23}, \hat{F}_{31}, \hat{F}_{32}, \hat{F}_{33})^T. \quad (41)$$

したがって, 式 (39) を解くことにより, 基礎行列 \hat{F} の要素を求めることができるが, 式 (39) の解 $\hat{\mathbf{f}}$ は定数倍の不定性を持つため, ここでは $\|\hat{\mathbf{f}}\| = 1$ という条件を追加する. $n \geq 8$ のとき, 式 (39) の解 \mathbf{f} は, 行列 \mathbf{A} を特異値分解して計算される最小特異値に対応する特異ベクトルとして得られる.

(3) 基礎行列の復元

正規化対応点から求められた基礎行列 $\hat{\mathbf{F}}$ から, 元の対応点に対する基礎行列 \mathbf{F} を計算する.

$$\mathbf{F} = \mathbf{T}'^T \hat{\mathbf{F}} \mathbf{T}. \quad (42)$$

\mathbf{T}, \mathbf{T}' は, 式 (37) の画像座標の正規化で用いた変換行列である.

以上により, 2 枚の画像間の対応点 $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$ に対する基礎行列 \mathbf{F} を計算することができる.

B. サポートベクターマシン

サポートベクターマシン (SVM) は, 2 クラスの線形識別器であり, 識別性能の高い手法として知られている. 通常, 学習型の識別器は, 学習サンプルを用いて識別器を構築するが, 学習サンプルに含まれていない未学習サンプルに対する高い識別性能 (汎化性能) を持つことが望まれる. SVM では, 汎化性能を向上させるための工夫として, 「マージン最大化」という基準を用いて識別器を構築する.

線形 SVM は, 入力特徴ベクトル \mathbf{x} に対して, 次の式により 2 値の出力値 y を計算する.

$$y = \text{sign}(\mathbf{w}^T \mathbf{x} + b). \quad (43)$$

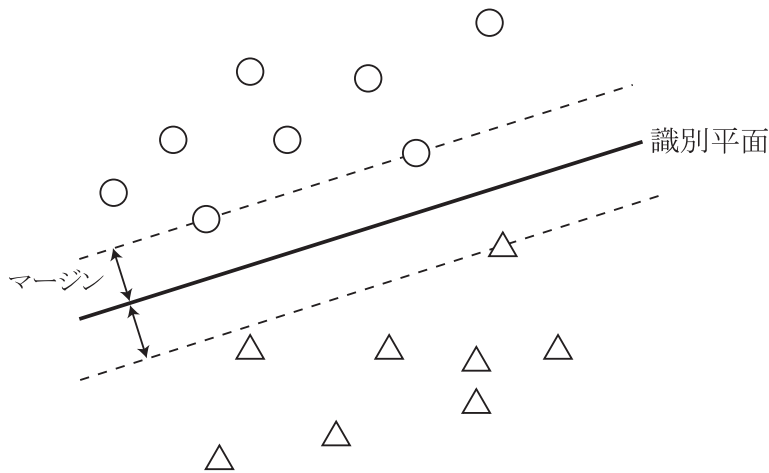


図 56 SVM の識別平面とマージン

\mathbf{w} は重みパラメータ, b はバイアス値である. また, 関数 $\text{sign}(u)$ は $u > 0$ のとき 1 を, $u \leq 0$ のとき -1 を出力する関数である. つまり, 線形 SVM の式 (43) は, $\mathbf{w}^T \mathbf{x} + b$ で表される超平面により特徴空間を 2 つに分けることに相当する.

学習サンプルから特徴空間を線形に 2 つに分離できるとしても, 一般にはその分け方は一意でなく, 特徴空間を分離する識別平面が複数存在する. SVM では, 図 56 に示すように, 識別平面から学習サンプルまでの距離 (マージン) が最大となる識別平面を求める. なお, 図 56 において, 丸印は Positive クラス (出力が $+1$ のクラス), 三角印は Negative クラス (出力が -1 のクラス) を表す. このとき, 識別平面を決定するサンプル (図 56 の波線上に存在する丸印と三角印) をサポートベクターと呼ぶ.

学習サンプル集合が $\{(\mathbf{x}_i, y_i)\} (i = 1, \dots, n)$ で与えられるとする. なお, \mathbf{x}_i は特徴ベクトル, $y_i \in \{+1, -1\}$ はクラスのラベルである. 図 56 において, 識別平面が $\mathbf{w}^T \mathbf{x} + b = 0$ であるとき, サポートベクター上を通る 2 本の波線はそれぞれ $\mathbf{w}^T \mathbf{x} + b = 1$, $\mathbf{w}^T \mathbf{x} + b = -1$ となる. また, マージンは $\frac{1}{\|\mathbf{w}\|}$ で表されるため, 学習サンプルを正しく識別し, マージンを最大化する識別平面を求める問題は, 次

のように表すことができる.

$$\min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|^2 \quad (44)$$

$$\text{subject to } y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, \dots, n \quad (45)$$

しかし実際には, 学習サンプルが線形分離できない場合がほとんどであるため, 学習サンプルに対する誤識別を許す代わりにペナルティを課す「ソフトマージン」と呼ばれる方法が用いられる. ソフトマージン法では, 以下の問題を解くことにより識別平面を求める.

$$\min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \quad (46)$$

$$\text{subject to } y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \quad i = 1, \dots, n \quad (47)$$

C はパラメータで, マージンの最大化と学習サンプルの誤識別の程度のバランスを決定する. 変数 ξ_i は, i 番目の学習サンプルが識別平面の反対側に入り込んだ距離を表す. この制約条件付き最小化問題は, Lagrange の未定乗数法を用いることで制約条件が簡単化された双対問題が得られ, その問題を解くことにより, SVM の識別平面を決定することができる.