# Random Forest による車載カメラ映像からの路面領域検出の検討

三品 陽平 † Abdelaziz Khiat‡ 下村 倫子 ‡ 藤吉 弘亘 † †中部大学 ‡日産自動車 (株) 総合研究所 E-mail: mishi@vision.cs.chubu.ac.jp

### Abstract

自動車の安全運転支援システムでは、自車の周囲状況を理解し、自車を制御しているため周囲の環境を高精度に認識する必要がある。そこで、本研究は車載カメラ映像からの路面領域検出に、統計的学習手法である Random Forest を用いて高精度かつ高速に実現することを目的とする。車載システムが動作するハードウェアは処理能力が低いため、特徴量の算出コストが高いとリアルタイム処理が不可能となる。そこで提案手法では、画素値を直接利用できるカラーヒストグラム特徴量と、車載カメラと路面の幾何学的な特徴として消失点とパッチ座標の関係性を捉えることで、高精度かつ高速に路面検出を行う。実験結果より、朝、昼、夜などの時間的変化や晴れ、雨、霧など気象変化のある、市街地走行シーンにおいて「値で 0.74 の識別精度を実現し、30fps での動作を確認した。

# 1 はじめに

自動車の安全運転支援システムは、障害物に衝突す る前に自車を制御することで、衝突回避や事故被害の 軽減を目的としている. 安全運転支援システムは高い 信頼性が必要不可欠であり、カメラやレンジセンサを用 いた自車の周囲の高精度な環境認識を実現する必要が ある。回避制御を行う場合、走行可能な路面領域を認識 する必要があるため、これまでに多くの路面検出の研 究が行われている[2]-[4]. 従来の路面検出のアプローチ は、大別するとステレオカメラを用いる手法と単眼カ メラを用いる手法に分けられる. ステレオカメラを用 いる手法は、ステレオカメラから得られる距離情報か ら,路面や物体等の3次元空間の構造を認識する[2][3]. しかし、高精度な距離推定には、正確なキャリブレー ションや画像間の対応を正確に求める必要があるため, 設置コストや演算コストが高く安価なシステムを構築 することができないという問題がある。一方、単眼カ メラを用いる手法では、 単眼カメラ画像のテクスチャ を利用して白線を検出し認識する[4]。まず、画像にエッ ジ保存平滑化フィルタをかけ、その後エッジ成分を抽 出しハフ変換により白線を検出する. 検出した白線の構造から,路面領域を認識する. しかし,白線が存在しない路面も多く存在することと,前方車両や障害物などの走行可能領域を判定することが難しいという問題が存在する.

そこで、本研究は車載カメラ画像から高精度かつ高速に路面領域を検出することを目的とする。提案手法は、識別器に Random Forest を用いる。Random Forest は木構造の識別器のため効率よく高速な識別処理が可能である。また、特徴量は抽出に処理コストの少ないカラーヒストグラムと車載カメラと路面の幾何学的な情報を利用することで、高精度でリアルタイム処理が可能な路面検出を実現する。

### 2 Random Forest

本章では、路面領域検出に利用する Random Forest について述べる。Random Forest は、複数の決定木構造を持ったマルチクラス識別器を構築する学習アルゴリズムである。Random Forest の特徴は、Bagging[6]と同様にブートストラップを取り入れ過学習を防ぐ点、Random Feature Selection[7]を取り入れ特徴ベクトルの次元数が大きくても高速に学習が可能な点である。このようなメリットを持つため、コンピュータビジョンの分野でも、Semantic Segmentation[5]、文字認識[8]、物体認識[9]、人体姿勢推定[10]で用いられている。

#### 2.1 学習

Random Forest は、学習サンプルからサブセットを作成し、複数の決定木構造を持つ識別器を構築する。各決定木は、分岐ノードと末端ノードにより構成され、分岐ノードを繰り返し作成し、一定の基準により分岐が不可能になった際に、末端ノードを作成する。学習アルゴリズムを Algorithm1 に示す。木の数 T、木の最大の深さ D とする。学習サンプル集合 T =  $\{\mathbf{x}_1, y_1\}, \dots, \{\mathbf{x}_N, y_N\}, \mathbf{x}_i \in \mathcal{X}, y_i \in \{1, 2, \dots, C\}$  を入力し、サブセットを作成する。サブセットは学習サンプル T からサンプルの重複を許容してランダムに選択する。サブセット一つを用いて決定木を構築する。決定木の構築は、分岐が不可能になるまで分岐ノードを繰り

Require: 学習サンプル  $\{\mathbf{x}_1, y_1\}, \dots, \{\mathbf{x}_N, y_N\};$  $\mathbf{x}_i \in \mathcal{X}, y_i \in \{1, 2, \dots, M\}$ Run: for t = 1 : T do 学習サンプルからサブセット エを作成.  $\Delta E_{max} \Leftarrow -\infty.$ for k = 1 : K do 特徴量からランダムに  $f_k$  を 1 次元選択する. for h = 1 : H do しきい値 Th をランダムに選択する.

Algorithm 1 Random Forest の学習アルゴリズム

 $f_k$  と  $\tau_h$  によりサンプル集合  $\mathcal{I}_n$  を  $\mathcal{I}_l$  と  $\mathcal{I}_r$  に 分割.

情報利得  $\Delta E$  を算出:

$$\Delta E = E(\mathcal{I}_n) - \frac{|\mathcal{I}_l|}{|\mathcal{I}_n|} E(\mathcal{I}_l) - \frac{|\mathcal{I}_r|}{|\mathcal{I}_n|} E(\mathcal{I}_r).$$

if  $\Delta E > \Delta E_{max}$  then

$$\Delta E_{max} \Leftarrow \Delta E$$

end if

end for

end for

if  $\Delta E_{max} = 0$  または最大の深さ D に達した then 末端ノードにクラス確率 P(c|l) を保存.

else

分岐を続ける.

end if

end for

返し作成する. 分岐ノードはある特徴量  $f_k$  としきい値 Th を用いて左もしくは右への分岐を決める分岐関数が 保存されている。そのため、 $f_k$ と  $\tau_h$  は様々な候補の中 から最適な組み合わせを選択する. 特徴選択回数 K と しきい値選択回数 H から、分岐関数の候補を  $K \times H$  個 ランダムに選択する. 最適な分岐関数を選択する評価 基準に情報利得  $\Delta E$  を利用し、 $\Delta E$  が最も大きくなる 組み合わせを分岐関数として選択する。ある分岐ノー ドnにたどり着いたサンプル集合 $\mathcal{I}_n$ , 分岐関数の候補  $f_k$ と $\tau_h$ を用いて式(1)により $\mathcal{I}_l$ と $\mathcal{I}_r$ に分割する.

$$\mathcal{I}_{l} = \{ i \in \mathcal{I}_{n} | f_{k}(\mathbf{x}_{i}) < \tau_{h} \} 
\mathcal{I}_{r} = \mathcal{I}_{n} \setminus \mathcal{I}_{l}$$
(1)

この  $I_l$  と  $I_r$  を用いて、式 (2) により情報利得  $\Delta E$  を 算出する.情報利得は、現在のノードのエントロピー から子ノードのエントロピーの和を引いたものであり, 分岐関数によりどの程度情報量が減少したかを表して いる。子ノードのエントロピーが小さくなると情報利 得は大きくなり、クラスをよく分割する分岐関数と表 される.

$$\Delta E = E(\mathcal{I}_n) - \frac{|\mathcal{I}_l|}{|\mathcal{I}_n|} E(\mathcal{I}_l) - \frac{|\mathcal{I}_r|}{|\mathcal{I}_n|} E(\mathcal{I}_r)$$
 (2)

ここで、関数 E(I) は情報エントロピーを表し、クラス の出現確率により算出される。これらの処理を繰り返 すことによりサンプル集合を分割し、情報利得が0に なった場合や最大の深さ D に達した場合に末端ノード 1を作成し、到達したサンプル集合から各クラスの出現 確率 P(c|l) を計算する. このように、各決定木が構築 される.

### 2.2 識別

Random Forest の識別のアルゴリズムについて説明 する. 未知入力サンプルxをすべての決定木に入力し、 分岐関数により左右に分岐させ決定木をトラバーサル する。そして、たどり着いた末端ノードに保存されて いるクラス確率  $P_t(c|\mathbf{x})$  を出力する。式 (3) に示すよう に、すべての決定木の出力の平均を算出する. 最終出力  $\hat{y}$ は、式(4)により最も確率の高いクラスに判別する.

$$P(c|\mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} P_t(c|\mathbf{x})$$
 (3)

$$\hat{y} = \arg\max_{c} P(c|\mathbf{x}) \tag{4}$$

# **3** 提案手法

路面領域の検出問題は、どの領域がどのクラスに属 するかを識別する問題であり、領域分割とクラス識別 を合わせた問題[5]である。本研究では、マルチクラス 識別器である Random Forest[1] を用いて路面領域の検 出を行う。ピクセル単位の識別は、抽出される情報量 が少ないため、高精度に識別することが困難であるた め、パッチ単位で学習、識別、評価を行う、図1に提案 手法の流れ図を示す.

# 3.1 テクスチャ情報の利用

本研究では、高速な検出を行うために特徴量の算出 コストが低いテクスチャ情報を利用する。特徴量は、カ ラーヒストグラムと勾配方向ヒストグラムを検討する. 以下に各特徴量の詳細を示す.

### 3.1.1 カラーヒストグラム

カラーヒストグラムの表色系は、事前実験の結果に より最も高精度に認識できる HSV 表色系を使用する. カラーヒストグラムpは、式(5)に示すように、パッチ から各ピクセルの HSV 表色系における量子化した値 I(u,v,r) の頻度を累積する.

$$p(r,n) = \sum_{u=1}^{M} \sum_{v=1}^{M} \delta[I(u,v,r), n]$$
 (5)

ここで、r は HSV 表色系の各要素 (r=H,S,V)、n は 量子化番号 (n=1,2,...,N),  $\delta[\cdot]$  はクロネッカーのデル タ関数を表しており、二つの要素が等しい場合は1、そ

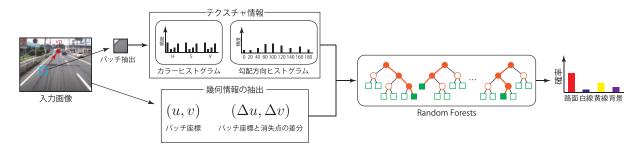


図1 提案手法の流れ

れ以外は 0 を出力する関数を表している。HSV 空間のそれぞれを 4 量子化 (N=4) するため,12 次元の特徴量となる.

### 3.1.2 勾配方向ヒストグラム

勾配方向ヒストグラムは、パッチ内のピクセルの勾配方向 $\theta$ ごとに勾配強度mを累積する。あるピクセエルの輝度値L(u,v)であるときの、勾配方向 $\theta$ 、勾配強度mは式(6)~式(8)より算出する。

$$\theta(u,v) = \tan^{-1} \frac{L_v(u,v)}{L_u(u,v)} \tag{6}$$

$$m(u,v) = \sqrt{L_u(u,v)^2 + L_v(u,v)^2}$$
 (7)

$$\begin{cases}
L_u(u,v) = L(u+1,v) - L(u-1,v) \\
L_v(u,v) = L(u,v+1) - L(u,v-1)
\end{cases}$$
(8)

算出した勾配方向  $\theta$  は  $0^\circ$  ~ 360° の値で算出されるが, 180° より大きくなる方向は 180° 引いて  $0^\circ$  ~ 180° とする.これにより,輝度の明暗関係に依存しない勾配方向を得ることができる.次に,算出した勾配強度 m と勾配方向  $\theta$  を用いて,式 (9) によりパッチ  $(M\times M$  ピクセル)における勾配方向ヒストグラム $\mathbf{v}=\{v(1),v(2),...v(N)\}$ を作成する.

$$v(f(\theta)) = \sum_{u}^{M} \sum_{v}^{M} m(u, v) \delta[\theta, f(\theta)]$$
 (9)

ここで, $f(\theta)$  は勾配方向  $\theta$  を N 方向に量子化した値である.このように,パッチ毎にヒストグラム化することにより,局所領域内の微小な幾何学的変化に対して 頑健な特徴量となる.今回の実験は,N=9 とし 9 次元の特徴量とする.

### 3.2 幾何学的情報の利用

本研究では、テクスチャ情報に加え、二つの幾何学的な情報を利用することで、識別精度の高精度化を検討する。幾何学的な情報は、画像座標系と世界座標系を検討する。以下にその詳細を述べる。

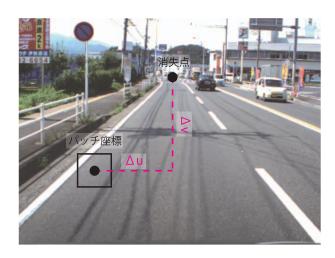


図 2 パッチ座標と消失点

### 3.2.1 画像座標系

カメラが車に固定して設置されている場合は、画像中に路面領域が出現する範囲が限られる。そのため、幾何学的情報の一つとして、パッチの中心座標 (u,v) をそのまま幾何情報特徴として用いることが考えられる。しかし、学習に用いた画像を取得したカメラ位置と異なるカメラから取得した画像では、識別精度が低下すると考えられる。そこで、本研究に利用する幾何学的な情報に、画像の消失点を利用することを検討する。自動車が路面を走行している場合は、道路は消失点方向に延びている。その関係を捉えることで、高精度に背景領域と路面領域の分離が可能になると考えらる。式(10)のように、画像中の消失点  $(VP_u, VP_v)$  と各パッチの中心座標 (u,v) との差分  $(\Delta u, \Delta v)$  を特徴量とする。

$$\Delta u = VP_u - u$$

$$\Delta v = VP_v - v \tag{10}$$

消失点は、画像のエッジ成分から直線を検出し、多くの直線が交わる点を消失点として検出している[11]. 図2に、画像座標系における画像座標上の幾何情報を示す。

#### 3.2.2 世界座標系

幾何情報を利用する際に、カメラ位置に依存した情報を利用して学習するとカメラ位置の異なる車に設置

した場合に識別性能が低下する可能性がある。そこで、画像座標系をホモグラフィ行列 H により世界座標系に変換するとで、カメラ位置が異なる画像座標系の座標を同一の特徴として表現できると考えられる。ホモグラフィ行列は、カメラ位置ごとにキャリブレーションすることで求まる。再学習せずにカメラ位置の異なる車種にシステムを搭載できる可能性が高い。ホモグラフィ行列によりパッチの中心座標は  $(x_w,y_w,1)^{\rm T}=H(u,v,1)^{\rm T}$ 、消失点とパッチ座標の差分は  $(\Delta x_w,\Delta y_w)$  と表される。二つの幾何情報を画像座標系と世界座標系において識別性能を比較する。

#### 3.3 路面領域の検出

提案手法の検出の流れは、車載カメラからの入力画像からパッチをラスタスキャンにより抽出する。各パッチから特徴量を抽出し、Random Forest に入力する。Random Forest では、特徴量を各決定木に入力しトラバーサルする。たどり着いた末端ノードからクラス確率を出力し、すべての決定木からの出力を式(3)により統合し、式(4)により確率が最大のクラスに判別する。

# 4 評価実験

提案手法の有効性を確認するための評価実験をする. 評価実験では、データセットを 3 グループに分割し、3-hold cross varidation により評価する. f 値により検出精度を評価する. f 値は、式 (11) により算出する.

$$f = \frac{2 \times precision \times recall}{precision + recall}$$
 (11)

ここで、precision は適合率を示し、ポジティブパッチと認識したパッチに対する正解率を示す。recall は再現率を示し、ポジティブパッチを正しくポジティブパッチと認識した割合である。f 値はこれらの値の調和平均であり、1 に近いほど認識精度が高いことを示す。

### 4.1 データセット

評価実験には、カメラ取り付け位置が異なる二車種のデータセットを利用する。図3の上段に、車高の高い車にカメラを設置した ELGRAND データセットの例を示し、下段に車高が低い SKYLINE データセットの例を示す。これらの画像は、車内に単眼カメラを設置し、市街を走行した際に前方を撮影した VGA(640×480)画像である。撮影したシーンは、朝、昼、夜などの時間的変化や晴れ、雨、霧など気象変化がある一般道と高速道である。ELGRAND データセットは、一般道と高速道の各シーンに対して72枚づつあるため、144枚のデータセットである。SKYLINE データセットは、一般道シーン 16枚、高速道シーン8枚の合計24枚のデータセットである。ラベル画像は、路面領域を赤(255.0.

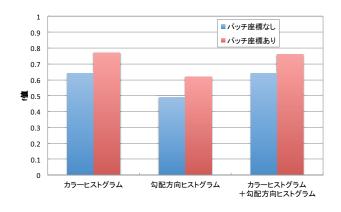


図 4 各特徴量の識別結果

表 1 各データセットの識別性能 (画像座標系)

学習データセット	EG	SKY	EG
評価データセット	EG	SKY	SKY
パッチ座標	0.77	0.79	0.56
消失点	0.76	0.74	0.63

0), 白線の領域を青(0,0,255), 黄色の線領域を黄色(255,255,0)として手動で色付けした.

#### 4.2 特徴量の検討

Random Forest を用いた路面検出に適した特徴量を検討する。比較する特徴量は、カラーヒストグラム、勾配方向ヒストグラム、カラーヒストグラムと勾配方向ヒストグラムそれぞれにパッチ座標を併用する場合と併用しない6パターンを比較する。各特徴量のf値を図4に示す。図よりテクステャ情報は、カラーヒストグラムが有用であり、勾配方向ヒストグラムの識別性能は低い。また、全てのテクステャ情報とパッチ座標を併用する場合の方が識別精度が高かった。識別結果例を図5に示す。上段に、一般道シーンの識別例、下段に高速道シーンの結果例を示す。図より、パッチ座標を利用することで、背景の誤検出を抑制している。これは、カメラは固定されているため、画像上に路面領域が出現する範囲を限定することで、精度を向上することができた。

### 4.3 幾何学的情報の検討

幾何学的な情報の検討として,画像座標系,世界座標系におけるパッチ座標と消失点を用いた特徴量を比較する.カメラ位置のことなるデータセットを学習と評価に用いて,識別性能の違いを比較する.

### 4.3.1 画像座標系での評価

表1に画像座標系における各データセットを学習,評価に用いた場合の結果を示す。図6にSKYLINEデータセットを用いて学習,評価を行った結果を示す。特に,画









ELGRAND データセット









(a) 一般道シーン

(b) 一般道シーンの正解ラベル

(c) 高速道シーン

(d) 高速道シーンの正解ラベル

SKYLINE データセット

図 3 データセット例

像下部に写り込んだ自車のボディを背景と認識するなど、車種特有の幾何構造を捉えている。図7に ELGRAND データセットにより学習し、SKYLINE データセットを評価した結果例を示す。各特徴量において学習用と識別用のデータセットが異なると識別性能が低下しているのが、パッチ座標をそのまま用いた場合よりも消失点との差分を用いた方が背景からの誤検出が抑制されている。

# 4.3.2 世界座標系での評価

表2に世界座標系における各データセットを学習,評価に用いた場合の結果を示す。ホモグラフィ行列による世界座標系への変換は、すべてのパターンにおいて識別性能が低く、カメラの取り付け位置の変化を吸収することはできなかった。

表 2 各データセットの識別性能 (世界座標系)

			(
学習	EG	SKY	EG
評価	EG	SKY	SKY
パッチ座標	0.63	0.30	0.43
消失点	0.64	0.30	0.44

# 5 おわりに

Random Forest と幾何情報を用いることで高精度かつ高速に路面検出できることを確認した。消失点と各パッチの画像座標の関係を用いることで、カメラの取り付け位置が異なる場合にも学習結果を有効に利用できることを確認した。また、消失点算出の処理コストを入れても30fpsを実現した。今後は、認識精度の高精度化に向けて車速やステアリング角度など車からの情報を利用することを検討する。また、大量の学習、評

価画像を用いた大規模な実験によるモデルの詳細な分析を検討する.

# 参考文献

- [1] L. Breiman, "Random Forests", Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.
- [2] 二宮芳樹, "走行環境認識のためのステレオ視の一方法", 電子情報通信学会論文誌, Vol. J91-D, No. 3, pp. 803-812, 2008.
- [3] 田中健一, 奥富正敏, "ステレオ画像を用いた道路シーン中の直線検出", 電子情報通信学会論文誌, Vol. J89-D, No.8, pp. 1892-1896, 2006.
- [4] 山口直人,田森信行,塩見彰睦,"適応エッジ保存平滑化を 用いた白線検出手法",電子情報通信学会誌,Vol. J89-D, No.8, pp. 1421-1431, 2005.
- [5] J. Shotton, M. Johnson and R. Cipolla, "Semantic Texton Forests for image categorization and segmentation", Computer Vision and Pattern Recognition, 2008.
- [6] L. Breiman, "Bagging predictors", Machine Learning, vol.26, no.2, pp.123-140, 1996.
- [7] Ho, T. K., "The random subspace method for constructing decision forests", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.20 no.8, pp.832-844, 1998.
- [8] Y. Amit and D. Geman, "Shape Quantization and Recognition with Randomized Trees", Neural Computation, No.9, pp.1545-1588, 1997.
- [9] Gall, J. and Yao, A. and Razavi, N. and Van Gool, L. and Lempitsky, V., "Hough forests for object detection, tracking, and action recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.33, no.11, pp.2188-2202, 2011.
- [10] J. Shotton, and A. Fitzgibbon, and Cook, M. and Sharp, T. and Finocchio, M. and Moore, R. and Kipman, A. and Blake, A., "Real-time human pose recognition in parts from single depth images", Computer Vision and Pattern Recognition, 2011.
- [11] Rasmussen, Christopher, "Texture-based vanishing point voting for road shape estimation", British machine vision conference, 2004.

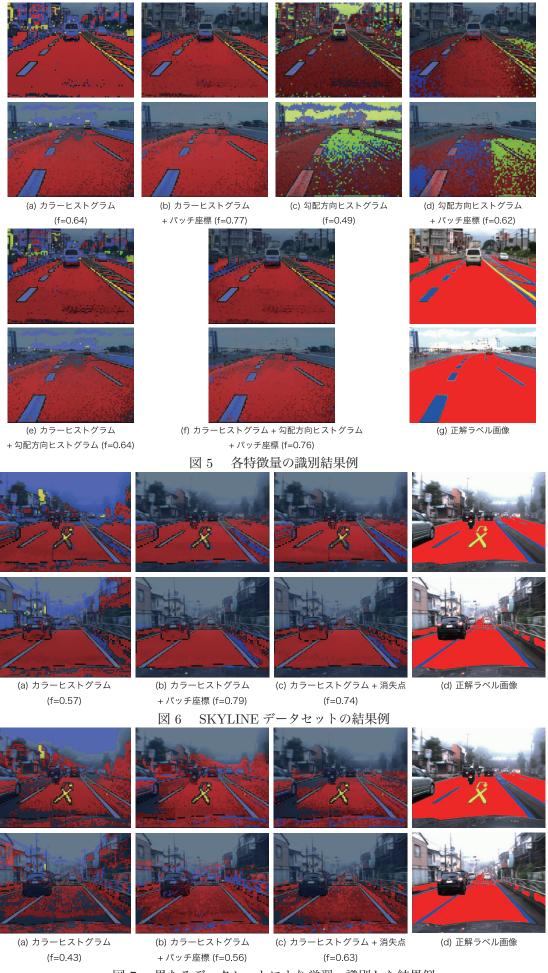


図7 異なるデータセットにより学習,識別した結果例