

Random Forest を用いた道路周辺地物の分類

東京大学 ○敷地 琢也, 増田 宏

Classification of road objects using Random Forest

The University of Tokyo: Takuya Shikichi, Hiroshi Masuda

It is important to efficiently maintain infrastructure such as traffic network and road signs. In this paper, we discuss a method for automatically recognize types of features in point-clouds captured by the mobile mapping system (MMS), which is a vehicle on which GPSs, IMUs, laser scanners and cameras are mounted. To utilize point-clouds in maintenance tasks, it is necessary to segment a point-cloud into each object and recognize what the object is. We propose a method for classifying objects by using the random forest classification method, which is one of machine learning methods.

1. はじめに

交通網や道路標識などの社会的なインフラの管理やメンテナンスに対する需要が高まっている。メンテナンスを効率的に行うためには、移動計測による点群データの自動処理が有効である。移動計測では、車両に GPS, IMU, レーザスキャナを搭載し、走行しながら道路や周辺の点群データを取得する。

ただし、取得されるデータは、離散的な点集合であるため、それらを有効に利用するためには、点群を地物ごとに分割し、その物体がなんであるかを正しく認識する必要がある。そこで、本研究では、移動車両計測装置により取得した点群データから、道路周辺の地物を自動で分類する手法を提案する。

2. 手法の概要

本研究では、道路周辺の地物を対象とするため、図 1 に示すような Mobile Mapping System (MMS)を用いる。レーザスキャナやカメラ等を搭載しており、道路周辺の点群データを取得することが可能である。本研究では、道路周辺の地物を対象とする。分類対象は、電柱、ガードレール、街灯、標識、信号機、木の 6 つとする。

まず、MMS で取得された点群データに関して、路面データを検出して路面を除去した後、点群を路面上の連結な地物成分に分割する。次に、個々の連結な点群の特徴量を計算し、機械学習の分類手法である Random Forests[1]を用いて、どのような種類の地物であるかの判定を行う。

3. 地物の分類手法

3.1 地物点群の抽出

地物を分類するためには、MMS から得られた点群データから個々の物体に分ける必要がある。まず点群データから、[2]に示す方法を用いて路面を除去する。さらに、地物同士を分離させるために、個々の点の近傍点をエッジで連結することで、ワイヤフレームモデルを作成する。近傍の点の検出には、kd-tree を用いた。分割された地物を図 2 に示す。

3.2 Random Forest

Random Forest 法は、機械学習における手法の一つであり、識別や、回帰に用いられる手法である。学習データからランダムに変数をサンプリングした後に複数の決定木を生成し、最終的にそれらの結果を統合することによって精度を向上さ



Fig.1 Mobile Mapping System



Fig.2 セグメンテーションされた地物

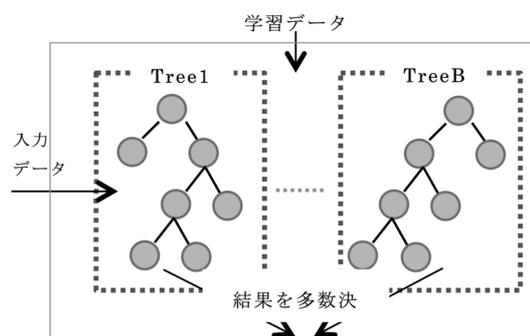


Fig3. Random Forest 概念図

せるアルゴリズムである。

以下に Random Forest 法のアルゴリズムを示す。図 3 は、Random Forest の概念図である。

- (1) 学習データから説明変数をランダムサンプリングすることにより、B 個の部分集合(ブートストラップサンプル)を生成する。この際、用いる学習データの 3 分の 1 をテスト用として取り除き、残りを学習用とする。テスト用

として取り除いたデータは、OOB (Out-Of-Bag)データと呼ばれる。

- (2) 各々のブートストラップサンプルを用いて、決定木を生成し、木の生成に用いていない OOB データを用いてテストを行う。その時の誤り率を OOB 推定値と呼ぶ。決定木の各ノードの分岐に関しては、M 個の特徴量の中から m 個 ($m < M$) をランダムに選択し、その中からもっとも分岐がよい変数を用いて各分岐の関数を決定する。
- (3) 次に先ほど生成した全決定木に対して分類したいデータを入力し、すべてのブートストラップサンプルの OOB 推定値に基づいて多数決をとることによって分類器を生成する。

Random Forest 法は精度が高くない複数の結果を組み合わせることで学習し最終的に精度を向上させる手法、集団学習法の一つであり、代表的な手法としてバギング法や、ブースティング法が挙げられる。他の手法と比較すると、特徴として、計算速度が速いことや、外れ値やノイズに関して比較的強いことが挙げられる。

MMS から得られるデータには、車を走らせながら計測を行うためノイズが多い。また、障害物によって一部分の情報が欠落することがある。よって、一つの変数に依存しない Random Forest 法が地物の分離に向いていると考える。

3.2 抽出する特徴量

Random Forest 法を地物の分類に適用するためには、地物間で大きな差が出るような特徴量を抽出する必要があるため、有効な特徴量を考える。

本研究で使用する特徴量 $f_i (i=1,2,\dots,20)$ を以下に示す。

1. $f_1 \sim f_3$: 点群サイズ (X, Y, Z)
2. $f_4 \sim f_6$: 主成分分析の固有値 ($\lambda_1, \lambda_2, \lambda_3$)
3. $f_7 \sim f_9$: 主成分分析による固有値の比 ($\lambda_1/\lambda_2, \lambda_2/\lambda_3, \lambda_1/\lambda_3$)
4. f_{10} : 主成分分析による第一主成分方向と Z 軸の成す角度
5. $f_{11} \sim f_{15}, f_{16} \sim f_{20}$: 高さ方向、横方向に 5 つにスライスし、その際の各スライスに含まれる点群の比。
6. $f_{20} \sim f_{22}$: ワイヤフレームに変換してスムージングしたとき、辺の方向と Z 軸との角度を 3 つに分けカウントした比。

特徴量 6 の例を図 4 に示す。スムージングをかけると人工物は水平か垂直に近い辺が多くなるに対して、木のような自然物にはこのような規則性がないことを利用している。

4. 評価実験

4.1 Random Forest による認識結果

地物の点群データに対して本手法を適用した。実験 1 では、点群サイズと、主成分分析の値 $f_1 \sim f_9$ のみを使用し、実験 2 では、先ほど述べたすべての特徴量 $f_1 \sim f_{20}$ を用いて、Random Forest 法を実装した。それぞれの実験には同じ学習データと入力データを用いて、それぞれ用いた個数を表 1 に示す。また、認識結果を表 2 に示す。

4.2 考察

実験 1 と実験 2 を比較すると、電柱や信号機の精度が大幅に向上しているため、 $f_{10} \sim f_{20}$ 有効な特徴量だったと言える。また、Random Forest 法では、すべての変数に対して各変数が精度にどのくらい貢献しているかを示す寄与度を計算することができる。実験 2 の寄与度が高い上位 3 つと低い下位 3 つの変数を図 5 に示す。

Z 軸方向の傾きが一番影響しており、主成分分析の固有値の比も強く影響している一方で、横方向と高さ方向のスライスの点群の比の一部が寄与していないことがわかる。

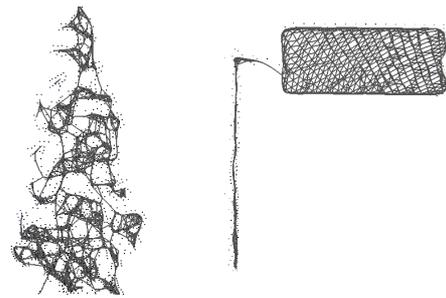


Fig.4 木(左)、標識(右)に対するスムージング処理

Table.1 学習データ数と入力データ数

地物	学習データ数	入力データ数
電柱	29	29
ガードレール	15	16
街灯	18	18
標識	27	27
信号機	11	12
木	19	20

Table.2 地物の分類精度

地物	実験 1 の精度	実験 2 の精度
電柱	79.3%	96.6%
ガードレール	80.0%	81.3%
街灯	100%	100%
標識	96.3%	100%
信号機	66.7%	96.6%
木	100%	100%

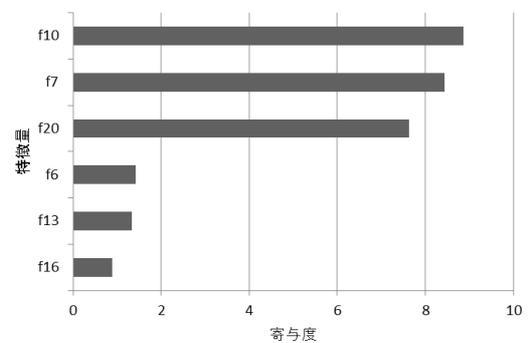


Fig.5 各変数の寄与度(一部)

5. まとめ

本研究では、MMS より取得した点群データから Random Forest 法を用いて道路周辺の地物を分類する手法を提案し、使用した特徴量が有効であることを示した。

今後、様々なシーンにおいてどの程度分類できるのかをより詳細に評価する必要がある。また、変数の寄与度が計算できるためそれを用いて様々な特徴量の中から効果的な特徴量を探していきたい。

参考文献

- 1) L. Breiman, "Random Forests", Machine Learning, Vol. 45, pp.5-23, (2001).
- 2) J. He, H. Masuda, Reconstruction of Roadways and Walkways Using Point-Clouds from Mobile Mapping System, ACDDE 2012