

# 移動計測で取得した点群と画像からの道路周辺地物の自動抽出

Automatic Extraction of Roadside Features Using Point-Clouds and Images Captured by Mobile Mapping Systems

電気通信大学 増田 宏, 森 悠真

車載の移動計測装置(MMS)により取得される点群や画像には様々な物体が含まれているため、それらを種別ごとに自動で抽出することが求められる。本稿では、点群と画像の両方を用いて、物体の抽出と分類を行う方法を解説する。

## ○ はじめに

道路や道路周辺の3次元情報を取得するために、車載の移動計測装置 (Mobile Mapping System, MMS) が用いられている。MMSは、車両の上にレーザスキャナ、デジタルカメラ、GPS、IMUを搭載し、走行しながら点群と画像を取得する(第1図)。レーザスキャナは、レーザ光を照射して、反射波が戻るまでの時間を計測することで、対象物までの距離を算出する。レーザスキャナの地球上での位置はGPSから計算できるので、レーザの照射方向と距離から、対象物の3次元座標を算出することができる。レーザ光を高速回転(この機種では毎秒100回転)させることで、MMSの周囲の3次元座標を大量に取得することができる。このデータを点群と呼ぶ。MMSにはデジタルカメラも搭載しており、走行中の連続画像を得ることができる。点群と画像には、取得時の時刻が記録されている。

MMSで取得された点群には様々な物体が含まれている。点群データを活用するためには、点群を対象物ごとに分割することが求められることが多い。この作業を人

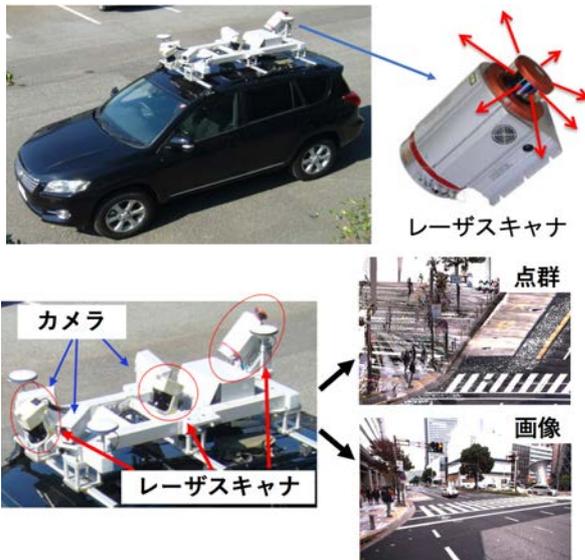
手で行うと多大な時間と労力が必要となるため、できる限り自動で行うことが望ましい。

しかし、点群だけを用いた分類では、形状が似ている物体や、少ない点しか乗らない小物体を正しく分類することは難しい。一方、画像による分類では、撮影角度や天気などの環境の影響を受けやすいという問題がある。

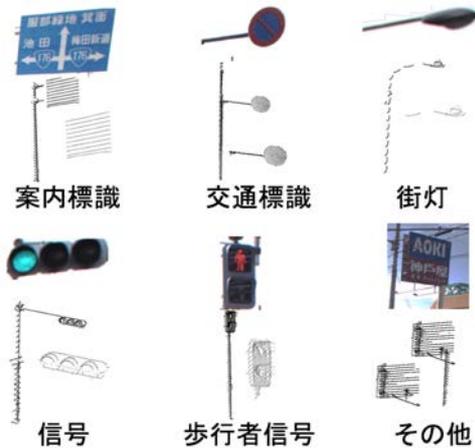
本稿では、MMSで取得される点群と画像の両方から抽出される特徴量を用いて、道路周辺の物体の抽出と分類を行う手法について解説する。

## ○ 点群からの柱状物体の抽出

道路周辺には、電柱、信号、標識、街灯、樹木など様々なものが存在するが、人工物に関しては、人や車両との衝突を防ぐために、柱状物体の付属物として、高所に設置されることが多い。そこで、本稿では、道路周辺の物体として、柱状物体に設置される、信号、歩行者信号、交通標識、案内標識、街灯を検出することを考える(第2図)。なお柱状物の付属物には、看板や電柱の変圧器、樹木の枝葉などもあるが、検出対象以外は「その他」として分類するものとする。



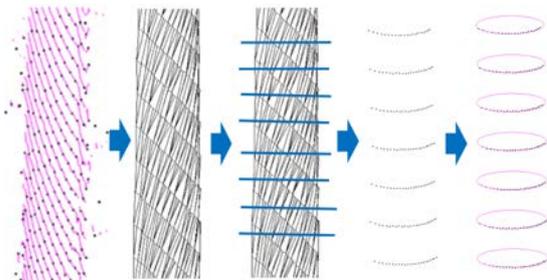
第1図 Mobile Mapping System



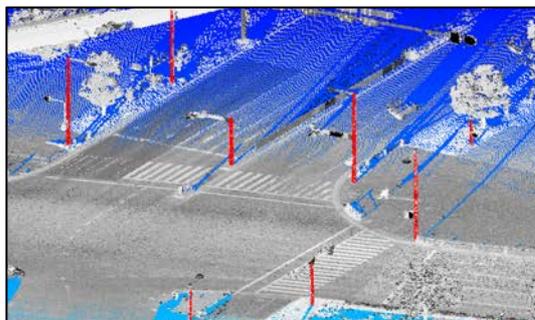
第2図 検出対象

検出対象の物体を検出するために、まず、点群から柱状物体を検出する<sup>1)</sup>。点群から柱状物体を検出するために、柱状物体の断面を計算する。

手順を第3図に示す。レーザスキャナから出力された点群は計測順に並んでいるので、それを順番に調べ、点間距離が十分小さければそれらを連結する。さらに、隣り合ったラインを調べて、近傍点を連結する。それにより、点群がワイヤフレームモデルに変換される。次に、ワイヤフレームモデルを水平面で切断して、切断点を求める。柱状物体であれば断面は円弧になるので、切断点から円の方程式を算出することができる。この操作を多数の水平面に対して行うと、柱状物体では、中心が垂直方向に並ぶ円の集合が得られる。そこで、そのような場合を検出することで、柱状物体を検出することができる。第4図に、点群から検出された柱状物体を示す。

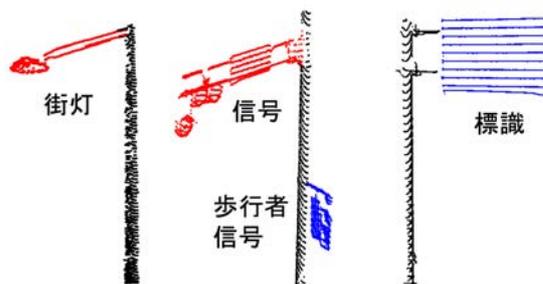


第3図 柱状物の検出



第4図 点群から検出された柱状物体

### ○ 点群特徴量の算出

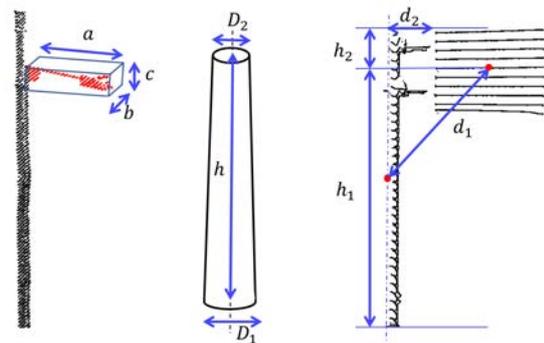


第5図 柱状物の付属物

検出対象の物体は、柱状物体の周辺にある。そこで、柱状物体が検出されたら、その周辺の点群を取り出し、近傍の点と同じグループになるように、セグメンテーションを行う(第5図)。これらの点群を柱状物体の付属物とし、検出対象の候補とする。

付属物の点群の種類を調べるためには、対象物の特徴をよく表す特徴量を計算する。まず、点群の特徴量として、付属物の点群を囲む直方体の大きさ( $a, b, c$ )、柱状物体に関して、高さ  $h$ 、底部と上部の直径  $D_1, D_2$ 、テーパ  $t$ 、付属物と柱状物体の相対位置  $d_1, d_2, h_1, h_2$  を算出する(第5図)。

さらに、付属物の点群に関して、主成分分析 (PCA) を行う。PCA は、点群が3次元空間でどのように分布するかを調べるもので、3個の主成分方向と固有値が算出される。付属物の点群の特徴量としては、PCA の固有値  $\lambda_1, \lambda_2, \lambda_3$  ( $\lambda_1 \geq \lambda_2 \geq \lambda_3$ ) と第1主成分方向、点群の反射強度、また、近傍点の形状特徴についても特徴量に加える(第1表)。近傍点の形状特徴とは、各点での近傍

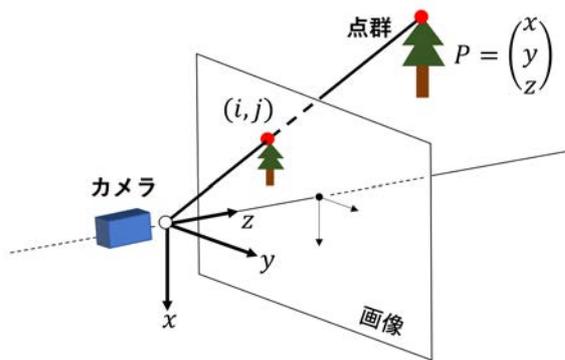


第6図 点群特徴量

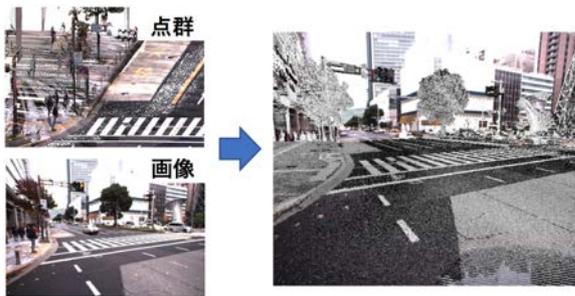
第1表 柱状物体の点群特徴量

	特徴量の種類
付属物	Bounding Box の大きさ ( $a, b, c$ )
	PCA の固有値 ( $\lambda_1, \lambda_2, \lambda_3$ )
	第1主成分方向と水平面との角度 $\theta$
	直線, 平面, 球状の点の比率 ( $r_l, r_p, r_s$ )
	点群の平均反射強度 $I$
柱状部	高さ $h$
	底面と上面の円の半径 ( $D_1, D_2$ )
	テーパ $t = (D_1 - D_2)/h$
相対位置	重心間距離 $d_1$
	最小距離 $d_2$
	柱状部底面からの距離 $h_1$
	柱状部上面からの距離 $h_2$

点に PCA を施して計算されるものである<sup>2)</sup>。点  $p_i$  の近傍点の固有値を  $\lambda_{i1}, \lambda_{i2}, \lambda_{i3}$  ( $\lambda_{i1} \geq \lambda_{i2} \geq \lambda_{i3}$ ) とするとき、 $L = (\lambda_{i1} - \lambda_{i2}) / \lambda_{i1}$ ,  $P = (\lambda_{i2} - \lambda_{i3}) / \lambda_{i1}$ ,  $S = \lambda_{i3} / \lambda_{i1}$  を計算する。このとき、 $L$  が最大ならば点  $p_i$  は直線状、 $P$  が最大なら平面状、 $S$  が最大なら球状と定義する。ここでは、直線状、平面状、球状と判定される点の個数の比率  $r_l, r_p, r_s$  を付属物の特徴量とする。



第7図 ピンホールカメラモデル



第8図 点群の画像上への投影



付属物の点群

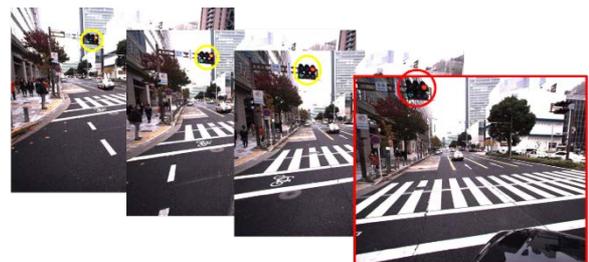
## ○ 画像の抽出と画像特徴量の算出

MMS では、車両上にレーザスキャナとカメラが固定されている。そのため、それらの相対位置を測定しておけば、レーザスキャナで取得した3次元座標に対応する画像上の位置を求めることができる。

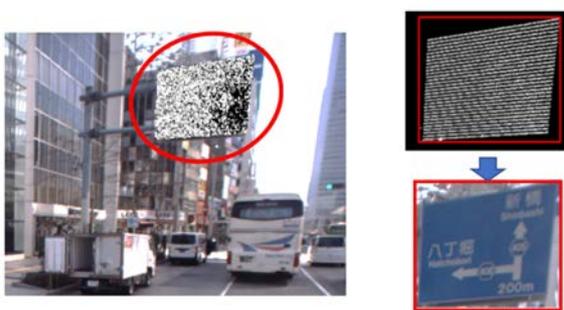
点群の画像上での位置は、ピンホールカメラモデルを用いることで計算できる(第7図)。この図のように、カメラ原点から3次元座標までの直線を考え、直線と投影面との交点を求めることで画像上の位置が算出できる。ただし、レンズの歪みなどのために、直線との交点と画像上の画素位置との対応がずれるため、歪み補正を行う必要がある。ここでは、Zhangの方法<sup>3)</sup>を用いた。点群を画像上に投影した例を第8図に示す。

この方法を用いて、付属物の点群に対応する画像を得る。点群と画像には、取得時刻が記録されているので、点群の取得時刻に近い画像に対して、点群を画像上に投影する。第9図の例では、標識の点群を画像上に投影している。投影された領域が点群に対応する画像となる。ここでは、点群が投影された領域を囲む矩形領域を取り、その範囲の画像を取り出して、検出対象の画像を得る。

MMS では連続的に画像を取得しているため、付属物は複数の画像に写っている。そこで、取得時刻に近いすべての画像に対して点群を投影し、付属物の点群がすべて画像に乗っており、かつ投影された面積が最大となる画像を選択する(第10図)。



第10図 点群を投影する画像の選択



第9図 付属物の点群の抽出

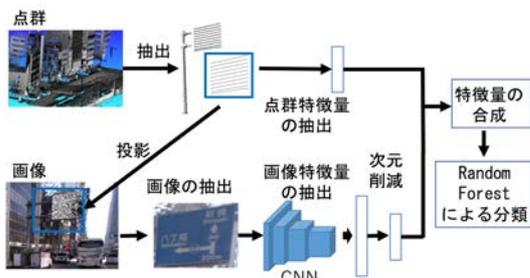
すべての付属物点群に対応する画像を得られたら、それらの画像から画像特徴量を算出する。画像特徴抽出には、Convolutional Neural Network (CNN) を用いる。ここでは、ImageNet<sup>4)</sup>の訓練用画像を用いて事前学習したAlexNet<sup>5)</sup>を用いた。AlexNetは畳み込み層に5層、全結合による特徴抽出に2層用いている。ここでは、出力層によるクラス分類の前の、全結合層の2層目から出力される4096個の画像特徴量を付属物の画像特徴量とする。

CNN から抽出される特徴量は4096次元のベクトルであり、3次元特徴量の次元と比較して圧倒的に大きい。また、多くが0の疎なベクトルである。そこで、画像特徴量の次元削減を行う。画像特徴量の削減は、機械学習手法の一つである Random Forest<sup>6)</sup>を用いて付属物画像の分類器を作成し、その際に計算される寄与度を用いて行う。寄与度とは、特徴量が分類に与える影響度を示す指標として提案されているものである<sup>9)</sup>。

ここでは、画像特徴量の次元削減のために、分類対象の画像をそれぞれ約500枚ずつ Random Forest に入力し、4096個の特徴量のうち、寄与度が高い順に64個の特徴量を抽出して、分類に用いる画像特徴とした。

## ○ 点群と画像を用いた物体認識

付属物の点群の種類を同定するために、物体認識には、点群特徴量と画像特徴量の両方を用いる。処理の流れをまとめると、第11図のようになる。まず、点群から柱状物体の付属物を抽出し、それらを画像に投影することで、対応する画像を得る。次に、点群と画像の両方から特徴量を抽出する。付属物の特徴量は、それらの特徴量を合成したものとする。合成した特徴量を用いて、第2図に示した6個のクラスに分類する。分類には、Random Forest を用いる。



第11図 点群と画像を用いた物体認識の手順

評価実験では、828個の柱状物体を2分割し、一方を学習データとして用い、もう一方を評価用のデータとして用いた。分類の評価では、適合率(分類が正しい比率)、再現率(全データから検出できた比率)、F値(適合率と再現率の調和平均)を求めた。

第2表に分類結果を示す。すべてのクラスにおいて、高い成功率が得られた。また、比較のために、点群特徴量のみを用いた分類と、画像特徴量のみを用いた分類についても評価を行った。比較した結果を第3表に示す。この表では、F値が最も高いケースを色をつけて示して

いる。この評価では、すべての場合において、点群と画像の両方の特徴量を用いた分類が、最も良い結果となった。

次に、点群特徴量と画像特徴量の寄与度について調べた。上位10個の特徴量を第12図に示す。点群特徴量は第1表に示した記号を用い、P:をつけて示した。CNNで用いられた画像特徴量の意味を人が理解するのは困難なので、ここでは先頭にI:を付け、4096までの番号を付けて示した。上位1~4位と9~10位は点群特徴量、5~8位は画像特徴量であり、いずれの寄与度も高いことから、3次元の点群特徴量と2次元の画像特徴量は補完的であり、両方の特徴量が分類に有効であると言える。

## ○ おわりに

本稿では、車載移動計測装置 MMS で取得される道路周辺の点群と画像を用いて物体認識を行い、検出対象の物体の点群を抽出する手法について解説した。

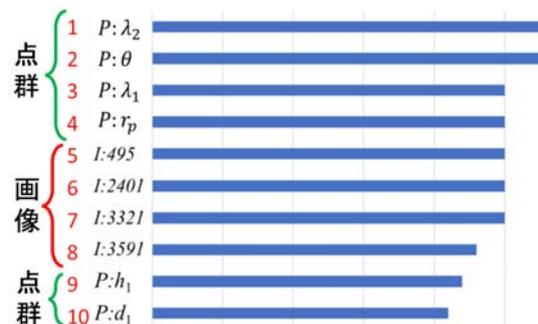
近年では、画像を用いた深層学習が飛躍的な進歩を見せており、また点群を用いた深層学習手法<sup>7)</sup>も提案されている。一方で、点群と画像は性質の異なるデータであ

第2表 点群と画像による分類結果

	案内 標識	歩行者 信号	信号機	街灯	規制 標識	その他
適合率	100%	92.8%	98.3%	92.7%	100%	90.2%
再現率	96.6%	92.8%	83.8%	100%	87.8%	95.1%
F値	98.3%	92.9%	90.5%	96.2%	93.5%	92.6%

第3表 分類結果の比較 (F 値)

	案内 標識	歩行者 信号	信号機	街灯	標識	その他	全体
点群のみ	96.6%	89.3%	93.9%	83.5%	91.4%	89.9%	90.6%
画像のみ	90.9%	92.9%	91.1%	81.2%	82.7%	93.1%	88.4%
点群+画像	98.3%	92.9%	96.2%	90.5%	93.5%	93.5%	93.7%



第12図 寄与度の高い特徴量

る。点群は夜間でも取得できるため、MMS での計測のように、それらが同時に取得できるならば、点群と画像を組み合わせる方が、多様な条件下でのロバストな物体認識が行えることが期待できる。

## 文 献

- 1) K. Fukano, and H. Masuda, Detection and Classification of Pole-Like Objects from Mobile Mapping Data, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. I-3/ W5, pp. 57-64 (2015)
- 2) J. Demantké, C. Mallet, N. David, and B. Vallet, Dimensionality Based Scale Selection in 3d Lidar Point Clouds, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-5/W12, pp. 97-102 (2011)
- 3) Z. Zhang, A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), pp.1330-1334 (2000)
- 4) K. Simonyan, and A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv technical report (2016)
- 5) A. Krizhevsky, I Sutskever, and GE Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems, pp.1097-1105 (2012)
- 6) L. Breiman, Random Forest, Machine Learning, 45, pp. 5-23 (2001)
- 7) C. R. Qi, H. Su, K. Mo, L. J. Guibas, Pointnet: Deep Learning on Point Sets for 3d Classification and Segmentation, the IEEE Conference on Computer Vision and Pattern Recognition. pp. 652-660 (2017)
- 8) Y. Zhou and T. Oncel, Voxelnet: End-to-End Learning for Point Cloud Based 3d Object Detection, the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4490-4499 (2018)

筆者紹介：

増田 宏 電気通信大学 大学院情報理工学研究科 教授  
森 悠真 電気通信大学 大学院情報理工学研究科