

レーザーレンジデータと Google Maps 画像の位置合わせを用いた環境のモデリング

松村 未来[†] アヌラグ アグラワル[†] 中澤 篤志^{†,††} 竹村 治雄^{†,††}

[†] 大阪大学大学院 情報科学研究科 〒565-0871 大阪府吹田市山田丘 1-5

^{††} 大阪大学 サイバーメディアセンター 〒560-0043 大阪府豊中市待兼山町 1-32

E-mail: †{matsumura,anuraag}@lab.ime.cmc.osaka-u.ac.jp, ††{nakazawa,takemura}@ime.cmc.osaka-u.ac.jp

あらまし 本研究では、取得したレーザーレンジデータと Google Maps 画像を共に複数クラスに領域分割し、クラスの関係性を考慮した位置合わせを行うことによって、実環境をモデリングする手法を提案する。実環境のモデリングにおいて、近年レーザーレンジセンサなどの距離センサが広く用いられている。しかし市街地のような広域環境をモデリングするためには、環境中の様々な場所で獲得した複数の距離（レンジ）データに対して、位置合わせを行う必要がある。本手法ではまず最初に、得られたレンジデータとその周囲の地理情報をそれぞれ、あらかじめ設定したいくつかのクラスに領域分割する。その後、両データのクラスが一致する地点を探索することで、レンジデータ撮影地点の判別や、相互の位置合わせを実現する。

キーワード レンジデータ, Google Maps, モデリング, 位置合わせ

Modeling 3D Scenes with Registration between Laser Range Data and Google Maps Image

Miki MATSUMURA[†], Anuraag AGRAWAL[†], Atsushi NAKAZAWA^{†,††}, and Haruo

TAKEMURA^{†,††}

[†] Graduate School of Information Science and Technology, Osaka University

1-5, Yamadaoka, Suita-shi, Osaka, 565-0871 Japan

^{††} Cyber Media Center, Osaka University 1-32, Machikaneyama-cho, Toyonaka-shi, Osaka, 560-0043 Japan

E-mail: †{matsumura,anuraag}@lab.ime.cmc.osaka-u.ac.jp, ††{nakazawa,takemura}@ime.cmc.osaka-u.ac.jp

Abstract In this research, we propose a method for taking expansive 3D models of environments using highly-accurate laser range data. Laser range data is taken using distance sensors called laser range sensors, which have grown in use over recent years, including such practical applications like Google's street-view. However, trying to model a large-scale environment like a city requires taking data from various locations around the environment, and this data must be aligned with respect to each other. In our method, we initially decide on a location to model, taking range and map data for it and classifying it into multiple regions. We then find the location and orientation on the map that allows the two results to match, giving the location and orientation of the captured range data.

Key words Range Data, Google Maps, Modeling, Registration

1. はじめに

近年、カーナビゲーションシステムや、Google Earth [1], Virtual Earth [2] などのシステムにおいて、地図上に建物等の三次元モデルを重畳表示できるシステムの利用が広まりつつある。しかしこれらは、建物等の概形を単純なポリゴンモデルで表現したものであり、より正確かつリアルティのある形状モデルを距離（レンジ）データを用いて獲得する研究が注目を集め

ている。

街などの広域環境をモデリングするためには、局所的に撮られたレンジデータ群に対し、互いの位置関係を求める必要がある。2枚のレンジデータ同士を位置合わせする手法はICP(Iterative Closest Point)法 [3] がよく知られているが、ここでは近接した限られた数のデータに対して位置関係を決定するものであり、データ数や計測範囲が大規模になる環境では、累積誤差などの影響で適用できないことが知られている。

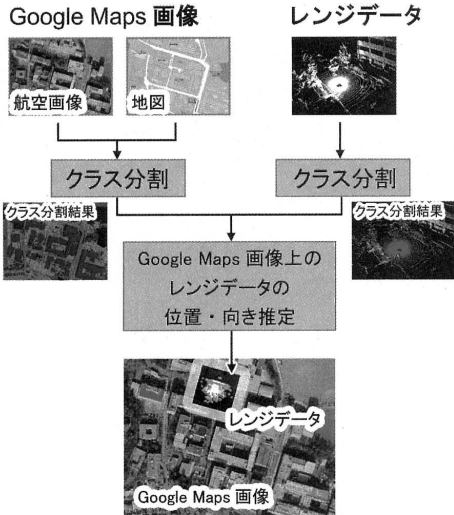


図 1 提案手法概要

この問題を解決するため、GPS や INS、上空から撮影された高度データなど、レンジデータ以外の外部の情報を利用する研究がなされている。H. Zhan らの手法では、車両にレーザレンジセンサならびにライン CCD カメラ、また GPS/INS の両機を搭載し、三次元モデルを取得する。その上で、航空機搭載のレーザセンサから取得した DSM (Digital Surface Model) に対し、手動でいくつかの "tie-point" を指定することで、さらに誤差を軽減することができる [4]。また C. Früh らは、車両に垂直と水平のレーザレンジセンサを搭載することで簡易な自己位置推定を行い、同様に取得した DSM に対し、モンテカルロ法を用いて正しい位置を推定、建物の天井も含めた完全な三次元モデルの生成を提案している [5]。このように、レンジデータ以外の外部情報と組み合わせることで、より正確な位置を得ることができる。しかしながら GPS や INS といった機器は単体では誤差が大きく、DSM はまだまだ普及しているとは言い難い。

我々は、近年一般にもよく利用されつつある地図データベースである Google Maps を用い、取得したレンジデータの姿勢を Google Maps 画像上に配置することで、広域環境の三次元モデルを構築する。これにより、レンジセンサ以外に必要なセンサを必要とすることなく、街等の広域における三次元地図を構築することができる。

図 1 に提案手法の流れを示す。まず Google Maps 画像、レンジデータのそれぞれを、建物、道路、樹木等のクラスに分割する。両データをクラス識別結果を用いて比較し、同じクラスの部位同士が一致するように Google Maps 画像上でのレンジデータの位置と向きを求める。

以降では、2 章で Google Maps 画像のクラス分割について、3 章でレンジデータのクラス分割について述べ、4 章ではその結果を用いた Google Maps 画像上でのレンジデータの位置・

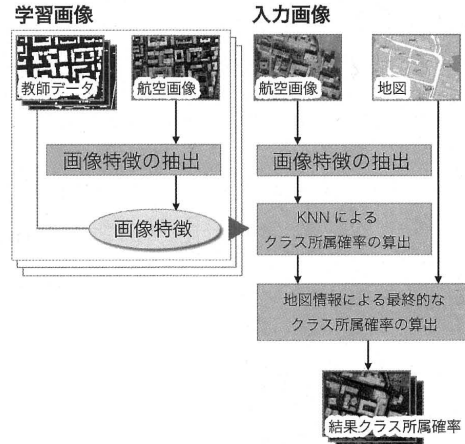


図 2 Google Maps 画像のクラス分割手法概要

向き推定について記述する。また 5 章では、本手法を用いた実験とその結果について説明する。

2. Google Maps 画像のクラス分割

本節では、Google Maps 画像の各画素に対し、クラス所属確率を求める手法について説明する。

本節で説明する処理の概要を図 2 に示す。入力データとして、Google Maps より得られる画像のうち、航空写真と地図の 2 種類の画像を併せて用いる。地図から得られる道や建物などの情報は正確であるように考えられるが、植物や細い道など記載されていない情報も多く、取得したレンジデータの位置を地図のみから特定するのは困難である。また航空画像からは逆に、より詳細な情報を得ることができるが、これらのクラスの判別は、場合によっては目視でも困難である。そこでこの双方の画像を用いることにより、欠点を補い合うことが可能となると考える。

このうち地図情報については基本的に、色とクラス所属確率の対応が 1 対 1 で決まると考える。そこで航空画像から後述する 2 種類の画像特徴を抽出し、特徴ベクトルを生成する。この特徴ベクトルを、予め用意した教師データと K-Nearest Neighbor を用いて比較し、画素ごとにクラス所属確率を計算する。これと地図から得られた確率を掛け合わせることで、最終的なクラス所属確率を計算する。

以降ではそれぞれ、用いる画像特徴、KNN によるクラス所属確率、地図情報による最終的なクラス所属確率について説明する。

2.1 画像特徴

画像から判別できる局所特徴は、テクスチャおよび平均色の 2 種を用いる。テクスチャは建物や道の境界部分のエッジ、規則的な物体の並びなどに、色は植物を初めとする自然物の認識に有用である。

そこで航空画像から得られる画像特徴として、SURF と Yxy 色情報の 2 つを用いる。以下に各々の説明と、詳しい採択理由

について述べる。

2.1.1 SURF

SURF は, H. Bay らによって提案された画像特徴であり, Haar ウェーブレットを用いてテクスチャを表現することができる [6]. 画像特徴としては SIFT [7] に類似しており, 回転や明度の変化に頑強である. そこで今回の航空画像のような, 照明条件や撮影方向が様々な変化する画像の認識に適切であると考える.

また SURF は Integral Image を用いることにより, 非常に高速な計算が可能である. 本研究では, SIFT や SURF の機能の 1 つである特徴点検出は行わず, 全画素に対し特徴ベクトルを求めるため, ある程度計算コストの低い画像特徴が望ましいと考える.

2.1.2 Yxy 色情報

前節でも述べた通り, 航空画像は撮影した時間や天気によって, また影と日向で, 激しい照明環境の変化がある. そこで明度と, 色相・彩度の要素が分かれた表色系が望ましいという仮説の元, 複数地域の航空画像を用いて, Yxy 表色系, HSV 表色系, $L^*a^*b^*$ 表色系, $L^*u^*v^*$ 表色系のそれぞれに対し色値の分布とクラスの関係性を調べた.

2.2 KNN によるクラス所属確率

航空画像の全画素に対して得られた 67 次元 (SURF64 次元, Yxy 色情報 3 次元) の特徴ベクトルを元に, K-Nearest Neighbor 法により, 画素ごとのクラス所属確率を計算する. 特徴ベクトル f^1, f^2 間の距離 $d_{1,2}$ は, 各画像特徴ごとのマンハッタン距離を用いて, 下記の式で求める.

$$d_{1,2} = w_m \sum_{i=0}^{63} |f_i^1 - f_i^2| + w_s \sum_{i=64}^{66} |f_i^1 - f_i^2| \quad (1)$$

ここで w_s, w_c はそれぞれ, SURF, Yxy 色情報 に対する重みを表す. この $d_{1,2}$ を用いて, K 近傍となる, 教師データ内の画素群, ならびにそのクラスを得る. 対象画素のクラス C への所属確率は, $P_C^{KNN} = \frac{N_C}{K}$ で求められる. ただし N_C は, K 近傍に得られた教師データ内の画素群における, クラス C の画素数である.

2.3 地図情報による最終的なクラス所属確率

航空画像のみからでは, 局所的には判別の難しい点が存在する. 例えば, 建物の天井と道路の中央は, テクスチャ, 色共に非常によく似ている. そこで, より確実な情報として, 地図の色情報を利用する.

ただし, Google Maps の地図画像には, そのままでは建物や道路などの境界線や, 地名の文字等, 不要な情報が含まれる. そこで数回に渡り中間値フィルタをかけ, ノイズを除去した後, 各画素から得た RGB 値に対し, 予め設定した色ごとの確率を適用する. ここで, 完全に色が一致しなかった点に対しては, そのマンハッタン距離に応じて傾斜をかけ, 確率を用いる. こうして求めた確率 P_C^{MAP} を用いて, 最終的なクラス C への所属確率

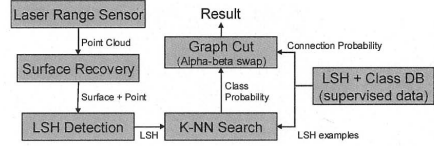


図 3 レンジデータのクラス分割手法概要

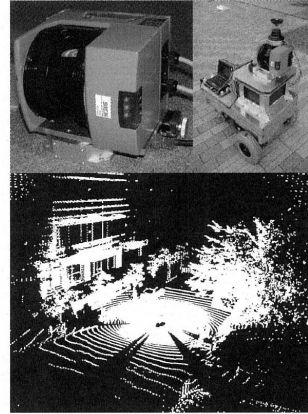


図 4 レーザレンジセンサー (左) と移動ロボット (右), 得られた点群データ (下). 本研究で用いたレーザレンジセンサーは 1 軸のものだが, 回転台と組み合わせることによって 2 軸とし, 全周画像を取得できる.

$$P_C = \frac{P_C^{KNN} P_C^{MAP}}{\sum_C P_C^{KNN} P_C^{MAP}}$$

を求める.

3. レンジデータのクラス分割

レンジデータのクラス分割には, 中澤らの手法を用いる [9]. 図 3 に手法の概要を示す. まず入力となる三次元データ点群から面を再構成し, 次の各面を中心に Local Shape Histogram (LSH) を計算する. 得られた LSH と, あらかじめ与えられた教師データとを K-Nearest Neighbor によって比較し, 各点のクラス所属確率を求める. 次に, 面の再構成結果を用いてシーン全体をグラフ構造で表し, グラフカット法により MRF を最適化する.

3.1 レンジデータ

本研究で用いるレンジデータは, センサーを一点に固定し, 周辺環境を 2 つの回転軸でスキャンしながら撮影するものである (図 4). レーザレンジセンサーはいくつかのスキャニング機構の種類があるが, 今回用いる 2 軸のものはもっとも一般的な物の一つである. 得られるデータの特徴として, 水平・垂直方向に隣接した距離データ間の, センサを中心とした角度は一定値なので, センサから距離が大きくなるほどデータの密度は疎となる.

3.2 Local Shape Histogram

LSH は三次元点群から得られる局所特徴量であり, 注目点を

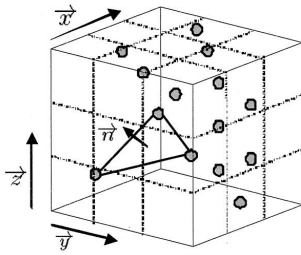


図5 Local Shape Histogram. LSH は各ポリゴンにおいて定義された三次元ヒストグラム空間で、与えられた上方向と面の法線方向から空間の方向が定義される。空間中で分割された投票空間（ビン）に周辺の点が投票されヒストグラム化される。

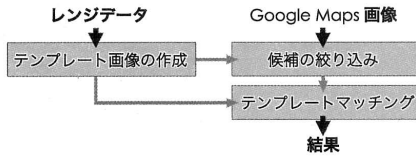


図6 Google Maps 画像上でのレンジデータの位置・向き推定処理概要

中心にした一定範囲の点の分布をヒストグラム化したものである。LSH は対象の局所的な形状（奥行き変化およびバウンダリ形状）を表現でき、また正規化されたヒストグラムであるので、点密度がセンサからの距離によって変化する場合にも対応できる（図5）。

3.3 LSH と KNN によるクラス所属確率

対象点から得られた LSH とあらかじめラベル付けされた教師データ群を元に、各点のクラス所属確率を K-Nearest Neighbor を使って計算する。すなわち、入力となる LSH(h1) と教師データの LSH(h2) の距離を Bhattacharyya 係数によって求め、K 近傍となる教師データ群を得る。対象点のクラス C への所属確率は $p_C = \frac{n_C}{n}$ で求められる。ただし n_C は、K 近傍の教師データ群中のクラス C の個数である。

3.4 LSH と MRF による屋外シーンの認識

前節の方法である点の所属するクラスが確定するわけではなく、点が所属する可能性のある複数のクラスの確率が求められる。これは、オブジェクトの認識は局所特徴のみで決定できるとは限らず、周囲の点が所属するクラス等によって決定される事も多いためである。この周囲の点の情報を考慮するため、シーン全体で Markov Random Field を構成し最適化を行う。この問題は、シーンを複数のラベル領域に分類するということであり、グラフカットによる複数のラベル領域分割手法である alpha-beta swap によって、高速に準最適解を得ることができ。

4. Google Maps 画像とレンジデータの位置合わせ

クラス分割された Google Maps 画像とレンジデータを位置

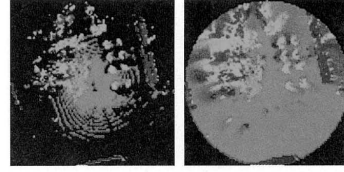


図7 レンジデータを画素上に投影した画像（左）と欠けている画素を補間した後の画像（右）。ここで赤は地面を、緑は植物を、青は建物を表している。

合わせし、レンジデータの姿勢推定を行う手法について述べる。

Google Maps 画像は鉛直方向から見下ろしたデータであり、また両者のスケールは既知であるため、位置合わせ問題は水平方向の位置、鉛直軸における回転成分を求めればよい。すなわち、レンジデータを地表と平行な平面の上に投影した画像（テンプレート画像）を作成し、Google Maps 画像とマッチングを行う。

この処理は回転を考慮したテンプレートマッチングを行う事で実現できるが、検索を効率化するため2段階の方法を用いる。まず初期マッチングとして、テンプレート画像からサイズの異なる複数のサブテンプレートを作成、これら全てに対して、含まれる各クラスのヒストグラムが似た点のみを、Google Maps 画像中の候補領域として選び出す。次に候補領域に対して回転を考慮した詳細な比較を行い最終的な結果を得る。

処理の概要を図6に示す。

4.1 レンジデータからのテンプレート画像の作成

レンジデータの性質として、センサに近いほど点の密度が高くなり、逆に遠い程疎になりクラス認識結果も不確実になる。また建物など、鉛直方向に伸びている物体は表面しか取得できず、そのまま画像に投影してしまうと、実際の占有面積とは著しく異なる結果となる。これらの問題を解決するため、下記の2つの手順でレンジデータからテンプレート画像を作成する。

(1) 点群をテンプレート画像に投影

(2) 補間処理

a) 点群をテンプレート画像に投影

レンジデータの各点はそれぞれ、レンジデータ座標系において、3次元の座標値を持つ。まずはレンジデータ座標系における地平面上に、テンプレート画像の画素に対応するグリッドを作成する。そのグリッドに投影する形で一定範囲内にある各点のクラスを投票し、最終的に投票された点の総数で正規化した値を、その画素のクラス所属確率として用いる。ここで作成されたテンプレート画像の一例を、図7左に示す。

b) 補間処理

この状態では、例えば地面はほぼ全点が広く散らばっている一方で、建物に対しては表面の点しか取得できないため、実際の面積に反して線状の画素のみが建物としてラベリングされるといった、アンバランスな割り当てが行われている。そこで、欠けている情報を復元するために、すでに確定している画素の情報から、不確定な画素の値を補間することを考える。

まず単純な手法として、センサ（中央）からその視線方向に

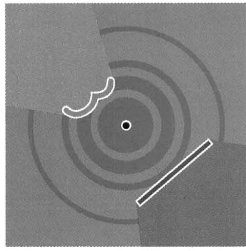


図 8 単純に視線方向のデータで補間した場合

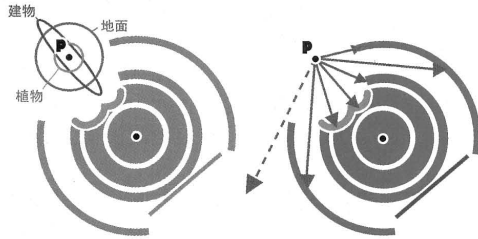


図 9 クラスごとに異なるカーネルを用いたガウス確率場に基づき、センサ方向の画素を参照する。左：ガウシアンカーネルの例 右：補間画素から一定範囲を参照

探索し、不確定な画素に関しては、直前の確定している画素のクラスで補間するという方法が考えられる。しかしこの場合、図 8 に示したように、木で遮られた領域の後ろが全て植物となる、建物の角の向こう側が正しく認識されない、などの不適当な補間が行われる。

そこで視線方向のみでなく、より領域的な情報を参照することを考える。参照する情報の重み付けとして、補間画素 p を中心とし、視線方向に長い非等方なガウス確率場を形成する。ただしこのとき、観測された点がどれほど遠くまで影響を及ぼすかは、クラスごとの性質に依存すると予想される。例えば地面は直線的に伸びていくが、木は通常、ある一定の大きさしか持たない。そのため、クラスごとに異なった大きさのガウシアンカーネルを用いる (図 9 左)。

また、補間対象の画素よりセンサから離れている画素は、参照すべきではない。建物など高さ成分を持つ物体は、観測された地点より遠方のみ存在し、センサ側には影響しないからである。同様にセンサ側の領域に対しても、補間画素から見て 2 つ以上のクラスが同じ方向に観測された場合、より補間画素に近い方のみを参照すべきである。建物の壁に遮られた地点において、観測された建物表面の前方にある地面のクラスを参照するのは不適切であると考えられるからである。

そこで、補間画素からセンサ方向の一定範囲に対し、補間画素に近い方から、参照すべき画素の探索を行う。その上で、前述のガウス確率場に基づいた重み付けを行い、補間画素のクラス所属確率を決定する。

こうして補完された画像の例を、図 7 右に示す。

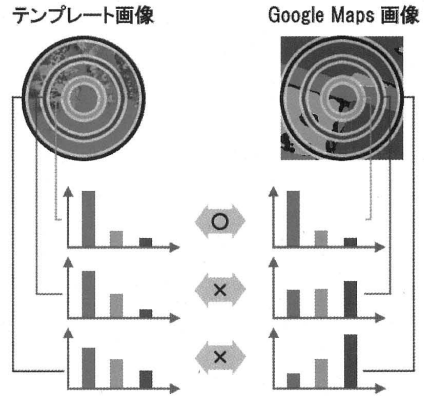


図 10 サブテンプレートごとにヒストグラムを作成し、比較する。

4.2 候補の絞り込み

続けて、テンプレートマッチングを行う前になるべく候補を絞り込むことを目的に、簡易なマッチングを行う。

マッチングにはクラスごとの割合を表したヒストグラムを用いるが、テンプレート画像全体のヒストグラムでは曖昧性が残り、効果的に候補を絞り込めない。そこでテンプレート画像を用いて、中心からの距離に応じた複数のサブテンプレートを作成し、その全てに対して一致率の高い点のみを、候補領域として検出する (図 10)。ヒストグラム同士の距離としては、Bhattacharyya 距離を用いる。

4.3 テンプレートマッチング

候補領域の各画素において、テンプレート画像を回転させながら、テンプレートマッチングを行う。このとき、画素同士の一致率の計算には同様に、Bhattacharyya 距離を用いる。

ここで最も一致率が高かった角度を、その画素地点におけるレンジデータの向きとして、またその角度における一致率を、その位置である確率として用いる。

5. 実 験

提案手法の有効性を確認するため、提案手法の 3 つに対し、それぞれ実験を行った。

- 実験 1: Google Maps 画像のクラス分割
- 実験 2: レンジデータのクラス分割
- 実験 3: Google Maps 画像上でのレンジデータの位置・向き推定

5.1 実験 1: Google Maps 画像のクラス分割

Google Maps より航空画像ならびに地図画像を取得し、正しくクラス所属確率が求められるかどうかを確かめた。画像サイズは約 1000 (幅) × 600 (高さ)、ズームレベルは 18 (1 ピクセルあたり約 50cm) のスケールの画像を使用した。

5.1.1 学習データとクラス

同様に Google Maps より 9 地点の航空画像と地図画像を取得し、学習データとして用いた。うち 5 地点が大学キャンパスを、4 地点が都市部を写したものであり、ズームレベルは入力

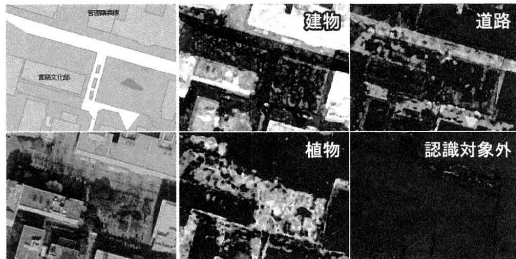


図 11 大阪大学豊中キャンパスにおける、Google Maps 画像のクラス分割結果。左 1 列が入力画像であり、右 2 列が順に、建物 (左上)、植物 (右上)、道路 (左下)、認識対象外 (右下) を表す。輝度が高い画素ほど、確率が高い。



図 12 同じく Google Maps 画像のクラス分割結果。得られたクラス所属確率から、最も確率の高いクラスを割り当てた。なお、色分けは、建物 (青)、植物 (緑)、道路 (赤)、認識対象外 (紫)

データと等しい。これらに対し、{建物、植物、地面、認識対象外} の 4 クラスに分け、正解部位をラベル付けした画像を用意した。

5.1.2 実験結果

実験を行った大阪大学豊中キャンパスの結果の一部を図 11, 12 に示す。図 11 はクラスごとの所属確率を輝度で示したものであり、左 1 列が入力データとして与えた地図と航空画像、右 2 列がその結果となっている。図 12 はこれら結果から最も確率の高いクラスを求め、色分けしたものであり、入力データのほぼ全域が示されている。

これらの結果より、細かなノイズはあるものの、おおむね正しく認識されていることが確認された。特に、地図の色からは判別できない道路、植物の領域や、多少地図とずれがある部分についても、航空画像が正しい情報を得ることに成功している。例えば図 11 の地図上にある細い道については、木々の中を通過しており、実際にレンジデータ取得時にも、道路よりも木が多く撮影されている。こういった情報は地図からは得ることができないが、航空画像を用いることによって、正しく植物を認識することができている。

一方、横断歩道や影と影の間、ベージュ色をした歩道等において、道路が建物や植物等、別クラスに認識された例もあった。これらが実際の道路上に占める割合は少ないため、学習データを用意する際に、別途多目に収集するなど工夫が必要である。また、グラウンドや駐車場が本来分類されるべきクラスに分類されていないという問題もある。これについては、航空画像を分類するにあたって適切な粒度のクラスを設定できていない可能性が考えられる。そこで、より適したクラスを模索すること

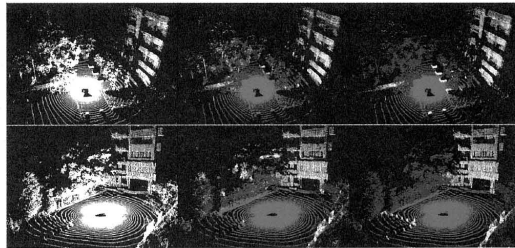


図 13 屋外シーンに対する 5 クラスの認識結果 (2 シーン)。左列: 入力データ、中列: 認識結果、右列: 正解データ (Ground Truth)。なお、色分けは地面 (赤)、樹木 (青)、建物・構造物 (黄)、段差 (紫)、柱・幹 (茶)

が、今後の課題として挙げられる。

5.2 実験 2: レンジデータのクラス分割

実際の屋外シーンを撮影した距離データ 10 シーンを準備し、5 クラスの領域に分類する実験を行った。対象とするシーンは、大阪大学豊中キャンパス内で取得したものであり、基本的に移動物体は含まれていない。使用データは SICK LMS-200 レーザレンジセンサ (奥行き精度約 $\pm 5\text{cm}$) で取得したものである。LMS-200 センサは、内部的にはレーザーの回転軸は 1 軸しか持たないが、我々のグループではこのセンサ自体を回転させることで、全周の距離データを撮影している。また、各回転軸でのスキャンステップは 1 度であり、シーン全体を 64800 点 (180×360) で表現する。LSH の次元は 512 次元 ($8 \times 8 \times 8$) であり、ポリゴンの近傍 ($\pm 40\text{cm}$, $\pm 40\text{cm}$, $\pm 40\text{cm}$) の範囲からヒストグラムを求める。

5.2.1 学習データとクラス

上記 10 シーンとは別に、学習用に撮影した 5 シーンに対し、教師データとして手動でラベル付けを行った。ラベルの分類は {地面、樹木、建物・構造物、段差、柱・幹} の 5 種であり、この中に含まれないクラスの物体に関しては、予め除去し、学習に用いた。

5.2.2 実験結果

本手法によりクラス分けした結果を図 13 に示す。なお、左列が入力データ、中列が認識結果、右列が比較対象として人が与えた正解データである。実験結果から、本手法がおおむね良好な認識結果を示していることが確認できた。一方で、局所特徴であるパッチベース法に起因する問題点もいくつか見られた。特に多く見られたのが、建物の細かな部分や、高い位置にある木の葉を相互に間違えるという例である。これらは、部分形状を見る限りでは双方とも似通っており、クラスの所属確率としても双方に一定の分布が見られるが、結果的に誤判別となった部位である。グラフカットによる MRF は、これらの問題を解決するべく導入したが、広い部分に誤ったクラスが分布した場合には MRF による改善は期待できないため、局所特徴量そのものを改善する必要がある。これに対しては、LSH のサポート (周辺) 領域をより広くすることが考えられるが、それによって細かな形状変化がとらえられないという問題も生じるため、

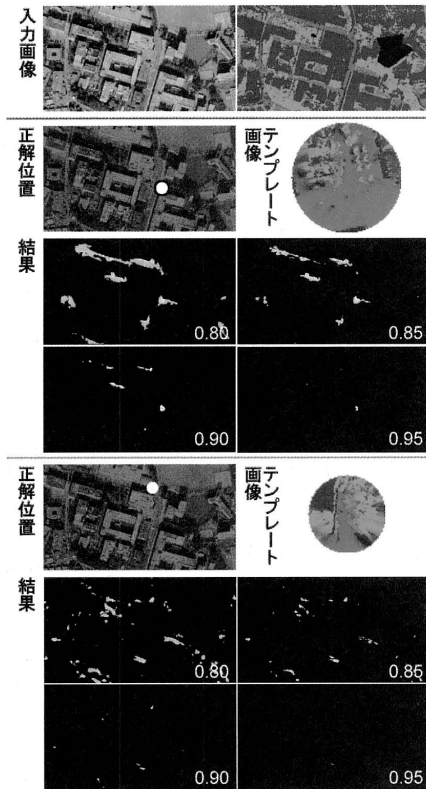


図 14 実験 3-1 結果 結果画像は閾値が 0.80 (左上), 0.85 (右上), 0.90 (左下), 0.95 (右下) のときに検出された領域を表している。

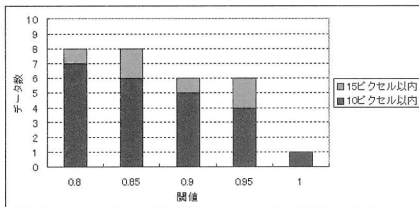


図 15 実験 3-1 結果 閾値と検出性能との関係。閾値 1 は、最も一致率の高かった点のみを検出領域として用いたもの。

LSH のサポートを段階的に縮小 (拡大) しながらかサーチするなどの手法を検討中である。

5.3 実験 3: Google Maps 画像上でのレンジデータの位置・向き推定

実験 1 の結果として図 12 左に示した、大阪大学豊中キャンパスの Google Maps 画像認識結果と、実験 2 の結果として得た 10 シーンのレンジデータ認識結果を用いて、以下の 2 種類の実験を行った。

- 実験 3-1: Google Maps 画像上でのレンジデータの位置・向き推定

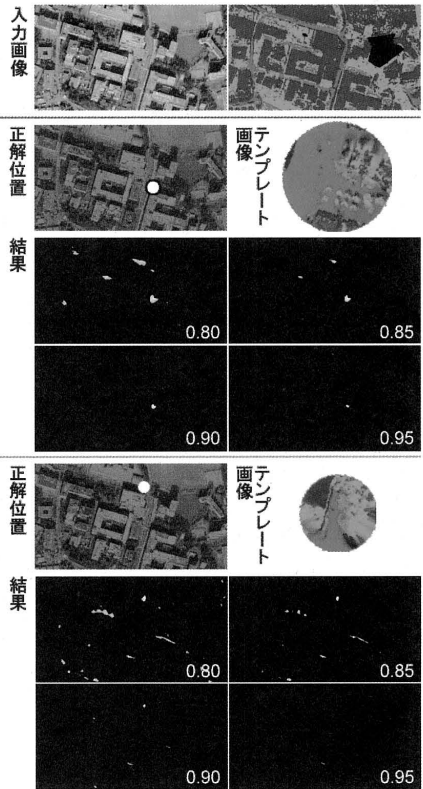


図 16 実験 3-2 結果 3-1 と比べ、検出範囲が狭くなっている。

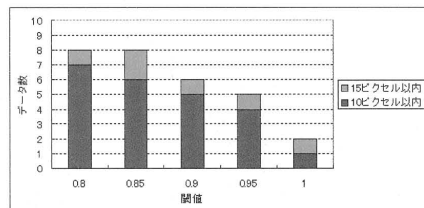


図 17 実験 3-2 結果 あまり 3-1 との差異は見られない。

- 実験 3-2: レンジデータの向きを既知のものとして与え、位置のみを推定

5.3.1 実験 3-1

Google Maps 画像の画素全体に対して一致率を計算、最も高い一致率を 1 とし、閾値以上の点のみを抽出した。閾値を 0.80 から 0.95 まで変化させたときの結果を図 14, 15 に示す。また閾値 0.80 のときに検出されなかった 2 点を除いて、全てのデータにおいて、正しい向きが認識された。

これらの結果より、閾値が低い場合には、多くの場合においてほぼ正しい位置・向きが検出されていることが分かった。その一方で、関係のない場所についても一致率が高くなっており、Google Maps 画像とレンジデータのみから場所を絞り込むのは非常に困難であると言える。

5.3.2 実験 3-2

そこで、比較的容易に取得できる情報として、方位情報を既知のものとして、実験 3-2 を行った。同様の結果を、図 16, 17 に示す。

本実験では方位情報を用いることで、検出範囲を実験 3-1 の 3 分の 1 以下に絞り込むことに成功した。しかし正答率に関しては、大きな変化は生まれなかった。このことから、向き情報は検出領域を絞り込むのに十分に有効であるが、それだけでは取得位置を決定するまでの効果は持たないと考えられる。

5.3.3 考察

実験結果より、改善すべき点は大きく 2 点に分けられる。

- 検出されない点を、検出されるようにする
- 検出領域を絞り込み、データの位置を唯一決定できるようにする

1 点目は主に、Google Maps 画像のクラス分割精度の問題である。実験 3-1, 3-2 共に検出されなかった地点が 1 点、ずれがあった地点が 2 点存在するが、これらは全て、Google Maps 画像のクラス分割に誤りがあったことが原因である。特に道路であるべきところが道路と認識されていない場合、検出される確率は非常に低くなる。そこで、この Google Maps 画像のクラス分割精度を上げるのがまず第 1 の課題である。

2 点目については、そもそも似たような特徴を持つ箇所というのは多くの場合複数存在し、レンジデータと Google Maps 画像のみの情報からでは、3-1 の結果でも述べた通り、特定は非常に困難であると考えられる。そこで現在は、レンジデータ撮影の際にオドメトリの情報を取得し、複数データ間のおおよその位置関係を用いて正確な位置を推定することを考えている。

6. おわりに

本論文では、レンジデータと Google Maps 画像を複数クラスに分割、比較することによって、Google Maps 画像上でのレンジデータの位置・向きを推定し、広域環境の三次元モデルを構築する手法を提案した。提案手法に基づき作成したシステムの動作実験を行ったところ、一致率の高い点として、正しい位置・向きが検出されることを確認した。検出率を改善するため、Google Maps 画像のクラス分割精度の向上、また効果的な追加情報の調査・連携が今後の課題として挙げられる。

7. 謝辞

本研究は、総務省 SCOPE 若手 ICT 研究者育成型研究開発 (ICT 安心・安全技術)「三次元データ認識による災害状況自動計測システムの研究開発」の支援を受けた。

文献

- [1] Google. "Google Earth". <http://earth.google.co.jp/>.
- [2] Microsoft. "Virtual Earth". <http://www.microsoft.com/japan/virtualearth/>.
- [3] C. Chao and I. Stamos. "Semi-automatic range to range registration: a feature-based method". In The 5th International Conference on 3-D Digital Imaging and Modeling, pp. 254-261, 2005.
- [4] H. Zhao and R. Shibasaki. "Updating a digital geographic database using vehicle-borne laser scanners and line cam-

eras". In International Archives of ISPRS 2004 Congress, 2004.

- [5] C. Früh and A. Zakhor. "Constructing 3D city models by merging ground-based and airborne views". In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, 2003.
- [6] H. Bay, T. Tuytelaars, and L. Van Gool. "SURF: Speeded Up Robust Features". Lecture Notes in Computer Science, Vol. 3951, p. 404, 2006.
- [7] D. Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.
- [8] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. "Visual categorization with bags of keypoints". In Workshop on Statistical Learning in Computer Vision, ECCV, pp. 1-22, 2004.
- [9] 中澤篤志. "グラフカットと Local Shape Histogram を用いた屋外シーンレーザーレンジデータのセグメンテーション". 画像の認識・理解シンポジウム (MIRU2008), pp. 1632-1637, July 2008.