

3D 都市モデル詳細化に向けた 自転車設置スマートフォン撮影画像の建物マッチング自動化

佐藤 剛*・中村 遼斗**・前田 紘弥***・小川 芳樹****・関本 義秀****

Automated Building Matching of Bicycle-mounted Smartphone-photographed Images for 3D City Model Detailing

Go Sato*, Ryoto Nakamura**, Hiroya Maeda***, Yoshiki Ogawa****, Yoshihide Sekimoto****

Abstract: The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) is currently developing an open 3D city model named “Project PLATEAU”. However, most of the currently available building data in PLATEAU does not have wall texture information. If the building images taken with a smartphone are linked to the PLATEAU building data, they will be useful for the development of wall texture information. In this paper, we propose an algorithm for assigning PLATEAU building IDs to the building images taken with a bicycle-mounted smartphone, with the aim of improving the quality of open 3D city models. When the proposed IDs are compared with the answer PLATEAU building IDs, 82% Top-3 accuracy and 90% Top-5 accuracy are obtained.

Keywords: 3D 都市モデル(3D city model), スマートフォン(smartphone), 建物マッチング(building matching)

1. はじめに

近年、世界各国において 3D 都市モデルの整備が進められている。日本においても 3D 都市モデルの整備が進められており、現在国土交通省において行われている Project PLATEAU はそうした 3D 都市モデル整備プロジェクトの 1 つである。PLATEAU は日本全国の都市において 3D 都市モデルをオープンデータとして整備するプロジェクトであり、整備されたデータは都市活動モニタリング・防災・まちづくり等に活用されることが見込まれている。

PLATEAU において、建物モデルのほとんどは壁面テクスチャ情報を持たない白箱型の建物モデルとして整備されている。建物の壁面画像を撮影し、PLATEAU の建物データに結びつけることができ

ば、壁面テクスチャ情報付きの 3D 都市モデルを作成することができる。壁面テクスチャ付きの 3D 都市モデルは防災・まちづくりシミュレーションを行った際の結果の分かりやすさを高めるほか、窓やドアの位置を特定することで自動運転・ドローン配送にも活用可能であるため、整備される価値は大きい。

都市に存在する膨大な数の建物の壁面画像を撮影するにあたり、汎用スマートフォンを使用できれば低コストでの画像入手が可能となる。近年の深層学習技術の発展により、画像内から自動で建物を取り出すことが可能となっている。そこで本論文ではスマートフォン撮影画像に対しあらかじめ準備された建物領域に PLATEAU 建物 ID を自動付与するアルゴリズムを作成し、評価を実施した。

* 学生会員 東京大学工学系研究科社会基盤学専攻
(Department of Civil Engineering, the University of Tokyo)

Email: gosato@iis.u-tokyo.ac.jp

** 学生会員 東京大学工学系研究科社会基盤学専攻
(Department of Civil Engineering, the University of Tokyo)

*** 正会員 東京大学生産技術研究所
(Institute of Industrial Science, the University of Tokyo)

**** 正会員 東京大学空間情報研究センター
(Center for Spatial Information Science, the University of Tokyo)



図1 テクスチャ付都市モデル生成における本研究の位置付け

はじめにPLATEAUでCityGMLファイルとして公開されている3D都市モデルをQGISにより建物頂点位置と建物IDを含むCSVリストに変換した。続いてスマートフォンのGPS記録から撮影に利用したカメラの方向を計算した上で、走行した地域に存在する建物を走行道路の左右どちらに写るのかという基準で分類した。その上で、スマートフォンのGPS記録、カメラ方向と建物の画像内位置から建物の地図上位置を推定し、推定位置から近い順に建物のPLATEAU IDを返すプログラムを作成した。この際、画像内で建物が左(右)に写っている場合には進行方向左(右)の建物のみを探索した。本手法の精度は2回の自転車走行によって得られた232枚の画像に含まれるのべ401個の建物領域に対し検証した。

2. 既存研究の紹介

2.1 テクスチャ付都市モデル生成の全体フロー

白箱型都市モデルからテキスト付都市モデルを生成するにあたっては、図1で示すように大きく分けて都市画像撮影・建物検出・ID付与・画像貼付の4つの手続きが存在する。本研究は4つの手続きの中のID付与に注目した研究であるが、ID付与を行う上では都市画像撮影・建物検出も行った上で、画像貼付との接続も考慮する必要がある。ここではこれら4つの手続きに関連した既存研究について述べ

た上で、既存研究と本研究の関連について述べる。

2.2 都市画像撮影に関連した研究

360度カメラを用いて全世界の地上画像を入手したGoogle Street Viewに関する論文としてAnguelov et al. (2010)が存在する。Anguelov et al. (2010)における地上画像撮影においては360度カメラ・レーザー・GPS付き四輪車両のほか、歩行者のみが進入可能な道路の撮影のために三輪車両を利用している。

都市には膨大な数の建物が存在するため、画像入手にあたっては低コストの方法が好ましい。都市において膨大な数の画像を低コストで入手した研究の事例としてMaeda et al. (2018)がある。Maeda et al. (2018)では高価な専用車両を用いずに道路損傷を検出するため、自動車のダッシュボード上に設置したスマートフォンで1秒間隔で画像を撮影することで163,664枚の路面画像を入手している。

2.3 建物検出に関連した研究

画像から特定の物体を検出する問題については深層学習を用いた研究が数多く存在する。深層学習を用いた物体検出を建物壁面画像に用いた研究として小川らの研究(2020)がある。小川らの研究(2020)ではセマンティックセグメンテーションを用いて空や地面を含む画像から建物壁面領域のみを抽出している。

2.4 建物 ID 付与に関連した研究

建物 ID 付与問題は、画像内から抽出された物体の地図上位置を推定し、地図上の地物と対応させる問題として考えることができる。画像が車両側面方向について撮影されたものであればこの手続きは比較的容易である。例として、Google Street View 画像から建物側面画像を取り出しセマンティックセグメンテーションを行った Femiani et al. (2018)においては、Google Street View 画像の GPS 座標と直線近似された GIS 建物データからなる三角形から、ある建物が 360 度カメラ内で写る角度範囲を推定することで、建物検出と建物特定を同時に行っている。

画像が車両前面のものである場合、画像内から抽出された物体の地図上位置を推定するためには画像内において物体が写った位置とサイズから GPS 位置と地物の位置の差分を推定する必要がある。深層学習を用いた物体検出と車両の進行方向データを組み合わせ、360 度カメラにおいて車両前面方向に写った物体の地図上位置を求めた研究として Campbell et al. (2019) がある。Campbell et al. (2019) では Google Street View 画像内の道路標識を深層学習により矩形領域として検出し、矩形の中心位置座標から道路標識の存在方向を画像の中心からの相対角 θ として求めている。車両の進行方向方位角に θ を足すと車両に対して道路標識が存在する方向の方位角 ϕ が求まる。道路標識の実際のサイズ、画像内サイズとカメラ焦点距離を使うと道路標識までの距離 D が計算できる。画像が撮影された位置を表すベクトルに D と ϕ から求まる道路標識の相対位置ベクトル $(D\cos\phi, D\sin\phi)$ を足せば道路標識の地図上位置を求めることができる。

2.5 建物画像貼付に関連した研究

車載カメラ撮影画像をテクスチャマッピングにより建物 3 次元ポリゴンに貼り付けた研究として桑野らの研究(2008)がある。桑野らの研究(2008)では建物ポリゴンの壁面全体が撮影画像に写っている状況に対して実験が行われ、画像の貼付においては透視変換が行われている。建物テクスチャ付与において建物画像の色や窓やドアなどの特定の要素のみに注目

した研究に Kelly et al. (2017) がある。Kelly et al. (2017) では Google Street View の車両側面画像から Edge スコアを用いて個々の建物を取り出し、建物の代表色を最頻値により 1 色に決定した上で、建物側面画像から窓・バルコニー・ドアなどの位置を推定しテクスチャ付き 3 次元都市モデルを作成している。

2.6 既存研究と比較した本研究の位置付け

都市画像撮影について、本研究では Maeda et al. (2018)と同様に低コストで画像入手を行うため、器具としてスマートフォンを用いた。本研究の対象地においては街路幅が狭く、車両の小回りがきくことが好ましいため、本研究においては Anguelov et al. (2010) において三輪車両が用いられたことを参考に、スマートフォンは自転車に設置した。

建物検出について、小川らの研究(2020)のように、画像内から自動で建物を取り出すことは深層学習技術により既に可能となっている。本論文では建物検出を行った前提で PLATEAU 建物 ID 付与アルゴリズムを構築することを目指したため、建物壁面ラベルについては画像の目視により手動で設定した。

建物 ID 付与を行うにあたり、本研究で使用したスマートフォン撮影画像には車両前面方向の建物が含まれる。車両前面方向に写った建物に対し建物 ID を推定するため、本研究における建物 ID 付与アルゴリズムは原則として Campbell et al. (2019) の方法により画像内の建物の地図上位置を推定し、推定位置から近い順に PLATEAU 建物 ID を返すものとしている。ただし、建物壁面は実際のサイズが不明であるため、カメラと建物の距離 D を推定するにあたっては Campbell et al. (2019)のように実際の物体サイズと画像内サイズ、カメラ焦点距離を使うのではなく、建物の存在方向の画像中心からの相対角 θ と道路幅を用いた。また、単純に推定位置から近い順に PLATEAU 建物 ID を返すと道路を挟んだ反対側の建物の PLATEAU 建物 ID が返されることが多くなるため、あらかじめ PLATEAU 建物データは進行方向左右に分類し、画像内で建物が左(右)に写っている場合には進行方向左(右)の建物のみを探索した。

本研究では 3 次元都市モデルへのテクスチャ付与

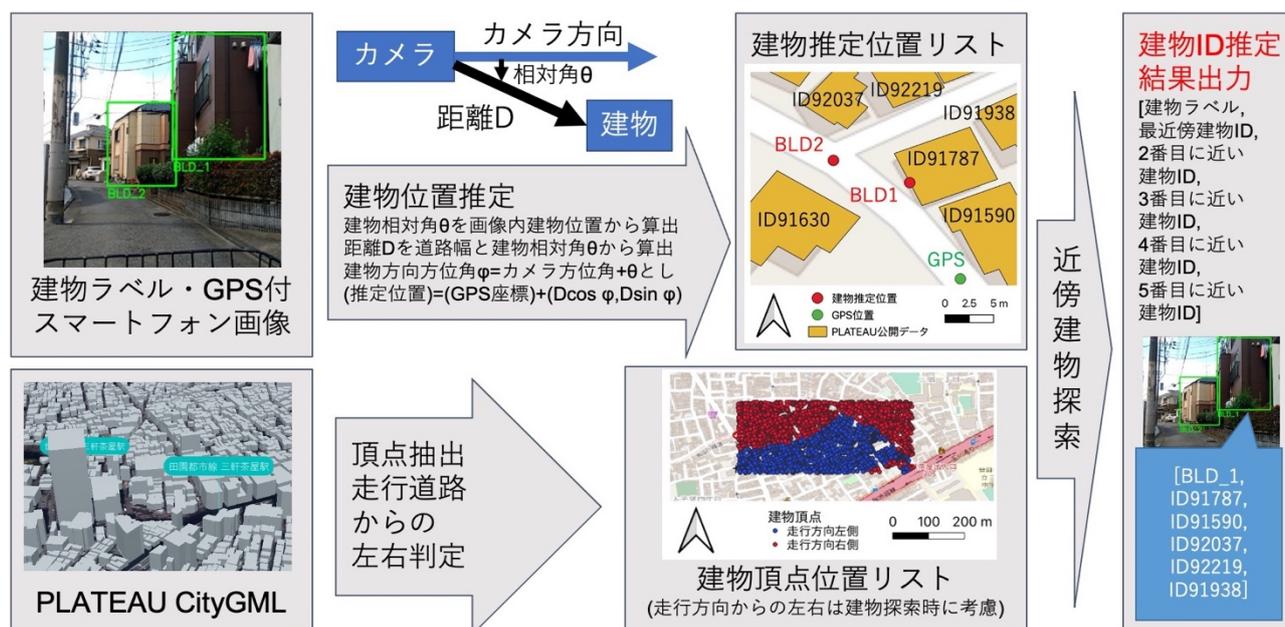


図2 本手法の概要

実装は行わなかったが、建物 ID が付与された建物側面画像の使用用途としては画像から色や窓やドアなどの特定の要素を Kelly et al. (2017)のように取り出し3次元都市モデルに付与することを想定した。

3. 実験の方法

3.1 街路画像と建物領域の準備

2021年7月10日、自転車のハンドルにスマートフォンを取り付け、東京都世田谷区三軒茶屋駅近辺の2区間の街路を1秒間隔で撮影し、1回目の走行(走行1)で97枚、2回目の走行(走行2)で135枚、計232枚の画像を得た。走行する街路は全て幅が狭くセンターラインが存在しない街路とした。画像撮影時にはGPSデータを同時記録した。

続いて、画像に鮮明に写った建物に対し矩形ラベルを手動で与え、走行1について158個、走行2に対して243個、計401個の建物領域を得た。

3.2 PLATEAU公開データとGPS位置情報の前処理

本論文では建物 ID 推定問題を点同士の距離や方位角を求める問題に分解して処理している。3D都市モデルを点群として扱うために、PLATEAUでCityGMLファイルとして公開されている3D都市モ

デルについて、QGISで頂点抽出操作を行うことで建物頂点位置と建物IDを含むCSVリストに変換した。さらに、3.1で得たGPSデータや本節で得た建物頂点の緯度経度は全て距離や方位角を求めることに適した座標系である平面直角座標系に変換した。

3.3 自転車進行方向算出と建物頂点の左右分類

時刻 $t=t_0-1, t_0, t_0+1$ でのGPS記録 $(X_{\text{before}}, Y_{\text{before}})$, (X_0, Y_0) , $(X_{\text{after}}, Y_{\text{after}})$ を考えると、 $t=t_0-1$ から t_0+1 へのGPS移動移動を表すベクトル $(X_{\text{after}} - X_{\text{before}}, Y_{\text{after}} - Y_{\text{before}})$ の方位角を時刻 t_0 でのカメラ方向とみなすことができる。また、全ての建物頂点に対し最近傍のスマートフォンGPS座標を求め、その座標におけるカメラ方向と建物頂点位置を比較することで建物頂点が行進方向左右どちらに存在するのかを判定することができる。この考え方にに基づき、3.2で得た建物頂点のうち、走行した地域から大きく外れたものを取り除いたものを進行方向左右の2クラスに分類した。

3.4 建物の地図上位置推定と建物ID推定

Campbell et al. (2019)と同様の方法を用いることで、3.1で準備した建物矩形ラベルの中心位置座標から建物の存在方向を画像の中心からの相対角 θ として

求めることができる。3.3 で求めたカメラ方位角にこの相対角 θ を足すことで建物の方向を表す方位角 ϕ を計算した。道路幅を 6m と仮定し、カメラと建物の距離 D を $(3 / \tan\theta)$ m とした上で、 D と ϕ から建物の相対位置ベクトル($D\cos\phi, D\sin\phi$)を決定した。 D が 20m を超えた建物については画像中心部に写ったものとして別途記録し、 D の値を 20m に訂正した。最後に GPS 座標(X_0, Y_0)と($D\cos\phi, D\sin\phi$)を足すことで画像内に写った建物の位置を推定した。

推定された建物位置について建物頂点との距離を計算し、距離が小さい順に PLATEAU 建物 ID を 5 種類求めた。この際、画像中心部に写ったものとして記録された建物については左右両方の建物頂点について距離計算を行い、その他の建物については画像内で建物が左(右)に写っている場合には進行方向左(右)の建物頂点のみについて距離計算を行った。

建物 ID 推定結果について、予測した 5 つの建物 ID の中において正解建物 ID が出現した位置をもとに定量評価を行った。予測が成功した画像と、5 番目の予測 ID までに正解 ID が入らず、予測に失敗した画像を目視で確認することで定性評価も行った。予測に失敗した建物の位置については QGIS による図示も実施した。

4. 実験の結果と考察

建物 ID を予測した全建物のうち、 k 番目までに正解建物 ID が現れる建物の比率である Top-k Accuracy を表 1 に示した。全ての建物領域を対象に計算を実施したところ、Top-3 Accuracy は 82%、Top-5 Accuracy は 90%となった。3.2 でダウンロードした三軒茶屋駅周辺の 3次元都市モデルには 5104 個の建物が含まれていることと、一般にスマートフォンの GPS にはメートル級~十メートル級の誤差が含まれることを考えれば、この値は十分に高く、本論文で考案した建物 ID 推定法の有効性を示していると考えられる。

画像の目視結果は以下の通りであった。図 3(a)で示したような大きなマンションについては 1 番目に予測された建物 ID が正解であることが多かった。一方 5 番目の予測 ID までに正解 ID が入らず、予測に失敗した建物には画像中心部に写りカメラと建物

表 1 Top-k Accuracy 計算結果

	走行 1	走行 2	全データ
Top-1 Accuracy	47%	59%	54%
Top-3 Accuracy	77%	85%	82%
Top-5 Accuracy	88%	91%	90%

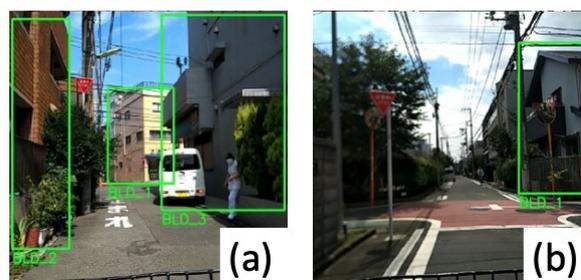


図 3 建物 ID 予測成功画像と失敗画像の例
(a) 大きなマンションの写った成功画像例
(b) 交差点における失敗画像例

の距離 D を一律に 20m とした建物や図 3(b)のように交差点奥部に存在し走行方向左(右)側にあるにも関わらず画像右(左)側に写った建物が多かった。

5 番目の予測 ID までに正解 ID が入らず、ID 付与に失敗した建物の位置を走行経路軌跡と合わせて図 4 に示した。ID 付与に成功した画像と失敗した画像の両方が当該建物に存在する場合についても図 4 においては ID 付与に失敗した建物として記載した。

図 4 中において東部に位置する走行 1 によって撮影された建物については、建物が疎である場所や緩いカーブにおいて画像中央付近に写った建物が ID 付与に失敗した建物となることが多かった。一方、図 4 中において西部に位置する走行 2 によって撮影された建物については、建物の走行道路からの左右分類が交差点において失敗した事例が多かったため、ID 付与に失敗した建物は交差点付近に集中した。

以上の結果から、本アルゴリズムは大きな建物に対しては有効であるものの、画像中央や交差点奥部の建物に対しては適用が難しいと考えられた。

5. おわりに

本論文ではスマートフォン撮影画像から取り出した建物領域に対し PLATEAU 建物 ID を自動付与す



図4 ID付与に失敗した建物位置と走行経路軌跡

るアルゴリズムを作成し、評価を実施した。結果における Top-3 Accuracy や Top-5 Accuracy の高さから、本論文の手法が建物 ID 推定においてある程度の有効性を持つことが示された。一方、本論文ではスマートフォンの画角を 52 度と仮定したうえ、道路幅も 6m と仮定したため、他のカメラを使用する場合や広い街路を走行する場合においては精度低下が予想される。さらに、本アルゴリズムは画像中央や交差点奥部の建物に対しては適用が難しい。今後はカメラの画角や道路幅に関する仮定を取り除く方法の検討や、画像中央や交差点奥部の建物に対するアルゴリズム改良が必要であるといえる。

参考文献

Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L. & Weaver, J. (2010) Google Street View: Capturing the world at street level. *Computer*, **43**(6), 32-38.

Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018) Road damage detection and classification using deep neural networks with smartphone images. *Computer Aided Civil and*

Infrastructure Engineering, **33**(12), 1127-1141.

小川芳樹・沖拓弥・陳聖隆・関本義秀.(2020)セマンティックセグメンテーションによる不動産物件外観画像からの建物抽出: 建物構造・築年代推定への応用,第29回地理情報システム学会講演論文集

Femiani, J., Para, W. R., Mitra, N., & Wonka, P. (2018) Facade segmentation in the wild. arXiv preprint arXiv:1805.08634.

Campbell, A., Both, A., & Sun, Q. C. (2019) Detecting and mapping traffic signs from Google Street View images using deep learning and GIS. *Computers, Environment and Urban Systems*, **77**, 101350.

桑野寛人・河野英昭・入江徹・前田博・生駒哲一。(2008) 車載カメラ画像を用いた 3D ポリゴン地図へのテクスチャマッピング: 第 24 回ファジィシステムシンポジウム講演論文集,428-431.

Kelly, T., Femiani, J., Wonka, P., & Mitra, N. J. (2017) BigSUR: large-scale structured urban reconstruction. *ACM Transactions on Graphics*, **36**(6). 204.

国土交通省

(2020)PLATEAU.<<https://www.mlit.go.jp/plateau/>>.