

街路の全方位画像と建物 GIS データの結合手法

小川芳樹*・沖拓弥**・陳聖隆***・関本義秀*

A Method for Combining Omnidirectional Street Images and Building GIS Data

Yoshiki Ogawa*, Takuya Oki**, Shenglong Chen***, Yoshihide Sekimoto*

Abstract: This paper proposes a method to combine the attributes of building GIS data with each building in an omnidirectional image of a street taken by a vehicle-mounted camera. Specifically, we extract buildings by instance segmenting the buildings in the omnidirectional images and identify the building coordinates in the image. Next, we estimate the building coordinates on the image, which are estimated from the building vertices in the building GIS data and the vehicle location history, and combine the building attributes corresponding to each building image. This method makes it possible to maintain building image data sets with building attributes.

Keywords: 深層学習 (deep learning), 建物抽出 (building detection), 異種データ (different types of data), 画角 (angle), インスタンスセグメンテーション (instance segmentation)

1. はじめに

近年、都市のデジタルツインの整備を念頭に車両に搭載されたカメラやセンサーにより位置情報付きの街路の全方位画像データが蓄積されている。全方位街路画像に建物 GIS データを結合することができれば、建物画像に建物 GIS データの属性情報（築年代、構造）などが付与され、建物画像に対する意味付け（アノテーション）が可能となる。そうしたデータを整備することは、建物画像を用いた深層学習による建物属性（例えば、構造や築年代など）の学習データセットの整備も同時に行うことを意味し、学習データセット整備の効率化を可能にする。しかしながら、全方位街路画像と建物 GIS データを用いた個々の建物情報の画像との紐づけは、建物以外の地物が含まれており、また複数の建物が映り込んでいるケースが多いため、簡単にデータ結合することが難しく、その手法は確立されていない。

以上の背景から本研究では、全方位街路画像と建

物 GIS データを用いて自動でデータ結合する手法の構築を行うことを目的とする。本研究の全体の流れを図 1 に示す。まず、全方位街路画像データを使用して、深層学習アプローチの 1 つであるインスタンスセグメンテーションにより、画像内の建物クラスを個別建物単位で抽出し、画像上の位置（画角）を算出する。また建物 GIS データと車両に搭載された GPS により観測された各全方位街路画像の撮影地点情報（緯度、経度、方位）を用いて、撮影地点からの各建物の理論的な画角を推定する。その上で、2 つの画角を比較して最近傍の建物データの結合を行う。各建物画像には、結合の際に建物 GIS データの建物 ID を付与するため、それと同時に対象の建物の GIS データ情報が付与される。

2. データ

本研究で用いる街路の全方位画像は株式会社ゼンリン（2013 年）の神戸市の 195,106 枚を用いる。車

* 正会員 東京大学空間情報科学研究センター (The University of Tokyo)
〒153-8505 東京都目黒区駒場 4-6-1 Tel : 03-5452-6406

** 正会員 東京工業大学環境・社会理工学院 (Tokyo Institute of Technology)

*** 学生会員 東京大学生産技術研究所 (The University of Tokyo)

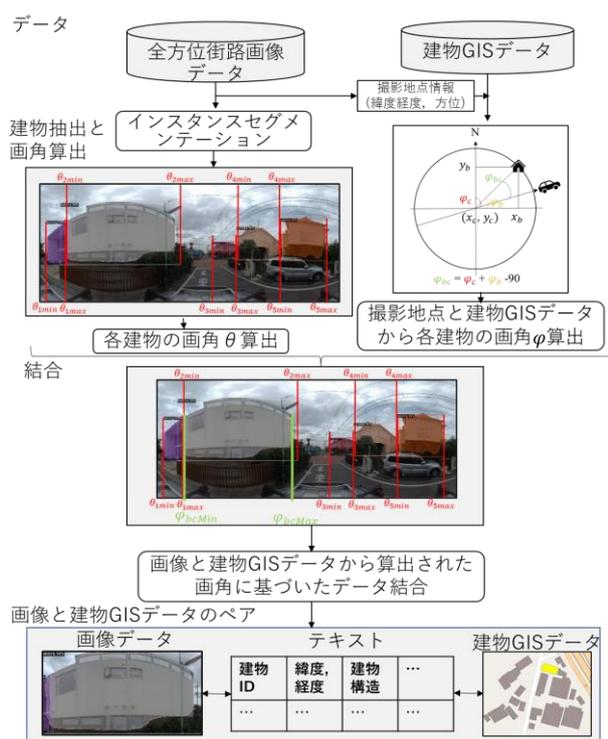


図1 本研究の全体像

表1 対象とする全方位街路画像データの仕様

項目	仕様
フォーマット	パノラマ化した Jpeg tar フォーマット
サイズ	横 5,400 × 縦 2,700 画素
撮影高	200cm
撮影ピッチ	250cm
位置情報	緯度, 経度, 方位
対象地域	兵庫県神戸市



図2 本研究で用いる街路の全方位画像の例



図3 学習データの建物ラベル付けの例

両の屋根上に搭載された 360 度カメラが 2.5m 間隔で街路を走行しながら撮影したものである。このデータには、画像ごとに GPS により測位された緯度・経度、車両方向、撮影日時などのテキスト情報が付属している。オリジナル画像は、パノラマ化された縦 2,700 pixel、横 5,400pixel の jpeg tar フォーマットである (表 1)。また画像の下部約 35%もしくは 45%の領域は、撮影車両の屋根部分が映り込んでいる。全方位街路画像の例を図 2 に示す。また建物 GIS データは、G 空間情報センターより提供いただいた神戸市の都市計画基礎調査 (建物用途・階数別現況調査 (2015 年)) を用いる。この建物 GIS データは建物単位のポリゴンデータであり、属性には用途や階数などの現況データが含まれる。建物 GIS データの前処理として建物ポリゴンから頂点を抽出した。

3. 手法

3.1. インスタンスセグメンテーションによる建物抽出と画角推定

全方位街路画像内の複数の建物を判別し、1 棟 1 棟の建物のバウンディングボックスを抽出する。インスタンスセグメンテーションは、画像の中の同じオブジェクトクラスに属する部分をピクセル単位でクラスタリングし、1 棟 1 棟の建物を検出することが可能である。本研究では、既存のモデルをファインチューニングすることで全方位街路画像に適した新たなモデルを構築する。そのため、ファインチューニング用の学習データには、ランダムに抽出した合計 1000 枚の街路の全方位画像を手動で建物をアノテーションしたものを用いる (図 3)。その上で全方位街路画像をアノテーションしたデータを用いて、COCO データセットで事前に学習したインスタンスセグメンテーションモデル (Mask R-CNN) をファインチューニングする (He et al. 2017)。バックボーンには、ResNet-101 の特徴を持つピラミッド型ネットワークモデル (X-101-FPN) を用いた。Mask R-CNN のアーキテクチャを図 4 に示す。Mask R-CNN を用いることで、各建物の最終的なピクセル単位の予測値のバウンディングボックスを得ることができる。最後に建物ごとのバウンディングボックスの両端のピクセル値から建物の画角 (左端: θ_{\min} , 右端: θ_{\max})

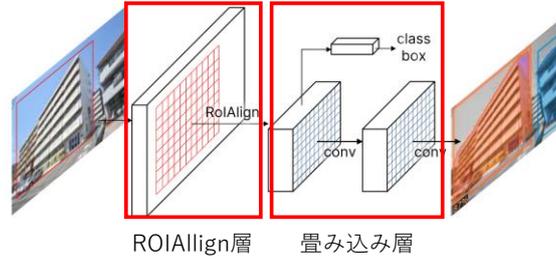


図4 Mask R-CNN のアーキテクチャ

を推定する.

$$\varphi_{bc} = \varphi_c + \frac{\pi}{2} - \varphi_b \quad (x_i < x_c, y_i > y_c) \quad (11)$$

3.2. 建物データと撮影地点に基づく建物画角推定

撮影地点における緯度経度を (x_c, y_c) , 進行方位(北を0とし, 南を π とする)を φ_c とする. また建物の頂点 i の緯度経度を (x_i, y_i) とする. これらを用いて撮影地点から半径 r 以内に位置する建物

$((x_i, y_i) \ni (x_r, y_r))$ を対象に画角の理論値を以下の

(1) から (3) の手順で計算する. なお本稿では, 計算コストの観点から $r = 30m$ に設定した.

(1) 撮影地点を基準とした建物の各頂点の角度は,

$$\varphi_b = \tan^{-1}(x_i - x_c, y_i - y_c) \quad (1)$$

で求まる.

(2) 撮影地点からの建物画角 φ_{bc} は, 進行方位 φ_c に
応じて以下の通り計算される.

i. 進行方位が北東 ($0 < \varphi_c \leq \frac{\pi}{2}$) のとき,

$$\varphi_{bc} = \varphi_c + \varphi_b - \frac{\pi}{4} \quad (\varphi_b > \frac{\pi}{4} - \varphi_c, x_i > x_c, y_i > y_c) \quad (2)$$

$$\varphi_{bc} = \frac{3\pi}{2} + \varphi_c + \varphi_b \quad (\varphi_b < \frac{\pi}{2} - \varphi_c, x_i > x_c, y_i > y_c) \quad (3)$$

$$\varphi_{bc} = \frac{3\pi}{2} + \varphi_c - \varphi_b \quad (x_i > x_c, y_i < y_c) \quad (4)$$

$$\varphi_{bc} = \frac{\pi}{2} + \varphi_c - \varphi_b \quad (x_i < x_c, y_i < y_c) \quad (5)$$

$$\varphi_{bc} = \frac{\pi}{2} + \varphi_c + \varphi_b \quad (x_i < x_c, y_i > y_c) \quad (6)$$

ii. 進行方位が南東 ($\frac{\pi}{2} < \varphi_c \leq \pi$) のとき,

$$\varphi_{bc} = \varphi_c + \varphi_b - \frac{\pi}{2} \quad (x_i > x_c, y_i > y_c) \quad (7)$$

$$\varphi_{bc} = \varphi_c - \varphi_b - \frac{\pi}{2} \quad (\varphi_c - \frac{\pi}{2} > \varphi_b, x_i > x_c, y_i < y_c) \quad (8)$$

$$\varphi_{bc} = 270 + \varphi_c - \varphi_b \quad (\varphi_c - \frac{\pi}{2} < \varphi_b, x_i > x_c, y_i < y_c) \quad (9)$$

$$\varphi_{bc} = \varphi_c + \frac{\pi}{2} + \varphi_b \quad (x_i < x_c, y_i < y_c) \quad (10)$$

iii. 進行方位が南西 ($\pi < \varphi_c \leq \frac{3\pi}{2}$) のとき,

$$\varphi_{bc} = \varphi_c - \frac{\pi}{2} + \varphi_b \quad (x_i > x_c, y_i > y_c) \quad (12)$$

$$\varphi_{bc} = \varphi_c - \frac{\pi}{2} - \varphi_b \quad (x_i > x_c, y_i < y_c) \quad (13)$$

$$\varphi_{bc} = \varphi_c - \frac{3\pi}{2} + \varphi_b \quad (\frac{3\pi}{2} - \varphi_c < \varphi_b, x_i < x_c, y_i < y_c) \quad (14)$$

$$\varphi_{bc} = \varphi_c + \frac{\pi}{2} + \varphi_b \quad (\frac{3\pi}{2} - \varphi_c > \varphi_b, x_i < x_c, y_i < y_c) \quad (15)$$

$$\varphi_{bc} = \varphi_c + \frac{\pi}{2} - \varphi_b \quad (x_i < x_c, y_i > y_c) \quad (16)$$

iv. 進行方位が北西 ($\frac{3\pi}{2} < \varphi_c \leq 2\pi$) のとき,

$$\varphi_{bc} = \varphi_c + \varphi_b - \frac{\pi}{2} \quad (x_i > x_c, y_i > y_c) \quad (17)$$

$$\varphi_{bc} = \varphi_c - \varphi_b - \frac{\pi}{2} \quad (x_i > x_c, y_i < y_c) \quad (18)$$

$$\varphi_{bc} = \varphi_c + \varphi_b - \frac{3\pi}{2} \quad (x_i < x_c, y_i < y_c) \quad (19)$$

$$\varphi_{bc} = \varphi_c - \varphi_b - \frac{3\pi}{2} \quad (\varphi_c - \frac{3\pi}{2} > \varphi_b, x_i < x_c, y_i > y_c) \quad (20)$$

$$\varphi_{bc} = \varphi_c + \frac{\pi}{2} - \varphi_b \quad (\varphi_c - \frac{3\pi}{2} < \varphi_b, x_i < x_c, y_i > y_c) \quad (21)$$

(3) 3.1 節から得られた全方位街路画像上の建物の
バウンディングボックスの画角 $(\theta_{min}, \theta_{max})$ と
3.2 節から得られた撮影地点と建物 GIS データ
から算出される建物画角 φ_{bc} から最近傍の建物
ID の紐づけを以下の式で行う.

$$ID = \arg \min\{(\varphi_{bcMax} - \theta_{Max} + \varphi_{bcMin} - \theta_{Min})\} \quad (22)$$

以上により, 全方位街路画像の建物に対して建物
GIS データの結合が完了する.

4. 結果

4.1. 建物抽出結果

建物外観画像にインスタンスセグメンテーションを適用し、建物クラスを建物単位で判別した画像の例を図5に示す。建物だけが個別によく識別されていることが確認できる。

また図6に示すように、学習曲線のトレーニングデータをプロットし、トレーニング中に対応する精度値を算出した。学習回数が増えていくことに伴い、モデルの学習損失が減少していることがわかる。モデルは約2000回の学習を繰り返すことで収束し、精度は最大で約0.86であった。また検証用データ（学習データの20%）を用いたピクセル単位の精度検証をした結果、35.7の mean average precision (mAP) を得た。mAPは、COCOデータセットで使用されている主要な評価指標で、ステップサイズ0.05で0.5から0.95までの Intersection over Union (IoU) の average precision (AP)の平均値をとっている。なおIoUは2つの領域がどれくらい重なっているかを評価する指標である。さらにバウンディングボックスの精度検証を行った結果、適合率は、0.941、再現率0.901の良好な精度を確認した（表2）。

その一方で、一部の画像では建物を背景クラスに識別した例や、高架を建物クラスに過大識別した画像が見られた。これらは、対象の全方位街路画像が学習データのパターンに含まれていないことに起因している可能性が高く、学習データのバリエーションを増やすことでモデルが学習し精度改善できる可能性がある。また今後はモデルが出力するバウンディングボックスが建物である確率を閾値として、識別することで建物に似た形状の地物の過大識別を減らすことを検討する。

4.2. 全方位街路画像と建物 GIS データの結合結果

図7に全方位街路画像と建物 GIS データの結合関係を線で結び地図に表示した例を示す。道路沿い1棟1棟の建物に対して、撮影地点データが紐づいていることが確認できる。本手法では、街路画像が2.5m単位で撮影されており、1棟の建物に対して複数の画像を取得することが可能である。またイン



図5 インスタンスセグメンテーションによる個別建物抽出結果の例(上図：市街地、下図：住宅街)

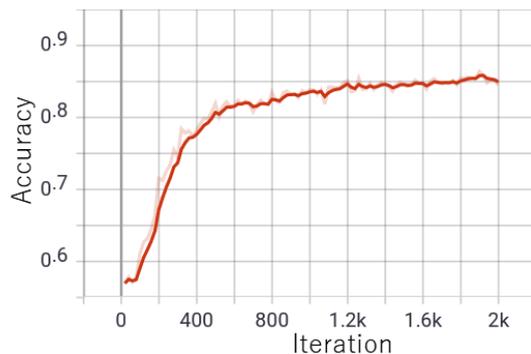
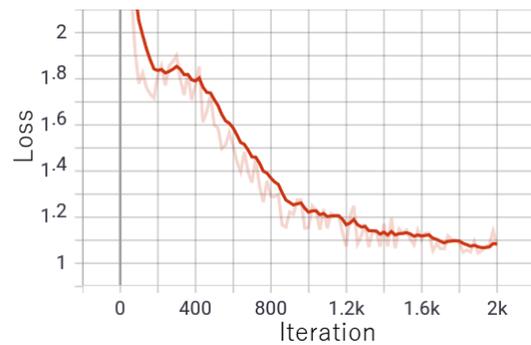


図6 インスタンスセグメンテーションにおける損失関数と精度の学習曲線（赤線：スムーズ曲線）

表2 インスタンスセグメンテーションの精度検証結果

適合率	再現率	F 値	IoU
0.941	0.901	0.914	0.849

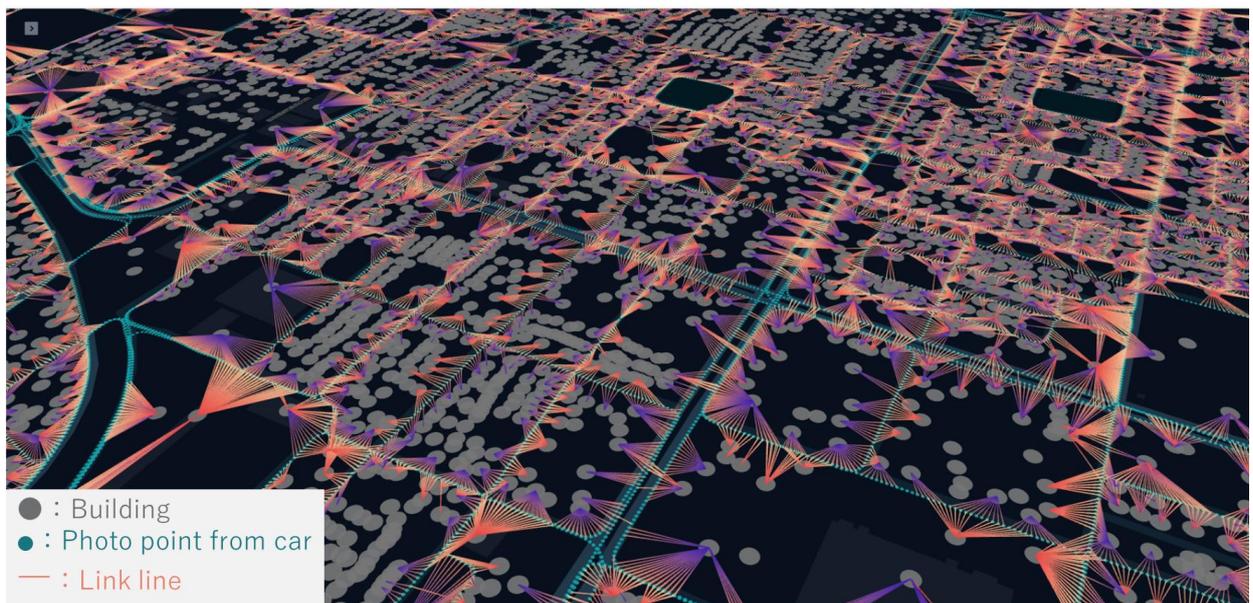


図7 全方位街路画像と建物 GIS データの結合関係

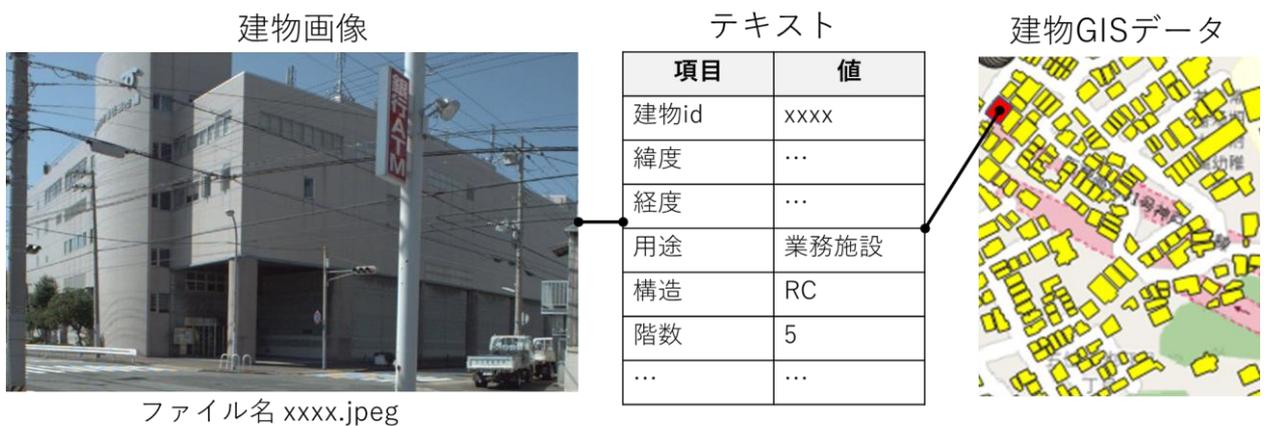


図8 結合したデータの例

スタンスセグメンテーションにより抽出した建物画像に建物 GIS データが建物 ID に基づいて結合されているため、単なる建物 ID 付の画像ではなく、建物階数や構造などの建物属性がアノテートされた建物画像になっている (図 8)。

4.3. 精度検証

以下では、開発した全方位街路画像と建物 GIS データの結合手法の精度と建物網羅率を評価した (表 3)。まず実験対象エリアの建物数は 36,969 棟あり、その内、撮影のために車両が通行した道路沿いの建物数は 21,098 棟である。本手法によって結合できた建物数は 18,421 棟であり、その網羅率は

対象エリア全建物数に対して 49.8%であった。また、本研究が目的とする道路沿いの建物数に対する網羅率は、87.3%であった。結合出来た建物画像枚数は、33,055 枚であり、建物 1 棟当たり約 1.8 枚の画像を取得出来ていることが確認できた。データ結合が出来なかった原因としては、道路から距離のある建物が多くあり、検索対象 (本実験では、車両の撮影地点から半径 $r = 30\text{m}$ を検索対象とした) に含まれなかったことが原因として挙げられる。これは、車両の撮影地点からの建物 GIS データの検索範囲を広げることで解決できる可能性があるが街路に面していない建物に関しては、衛生画像を用いる

ことが考えられる。他の原因としては、インスタンスセグメンテーションにより建物抽出が出来ていない建物もあった。インスタンスセグメンテーションの検出精度を改善することで網羅率の改善が見込まれる。次に、実際に結合したデータの精度評価を行う。これは、インスタンスセグメンテーションで抽出した建物画像が正しい建物 GIS データの建物と結合出来ているかを目視により評価するものである。延べ 1000 棟を対象に検証した。その結果、総合精度は 90.5%であり本手法の有効性を示している。

表 3 データ結合の網羅率と精度

項目	値
結合した建物数	18,421 棟
結合した建物画像	33,055 枚
道路沿いの建物数	21,098 棟
対象地域全体	36,969 棟
網羅率(対全建物数)	49.8%
網羅率(対道路沿い建物数)	87.3%
結合精度	90.5%

4.4. 今後の展望

全方位街路画像のインスタンスセグメンテーションの精度については、データ結合の精度にも影響を与えるため複数の建物として抽出される場合や並行して走行するトラックが建物として誤検出される例も見られたため、さらに高精度に建物輪郭を抽出することが課題である。そのためには、今回の実験で用いた学習データに見られないバリエーションの画像を多数用意して学習を行うことで、誤検出を少なくできる可能性がある。また今後は整備したデータを応用して Oki et al. (2021)の手法を用いて画像から構造築年代モデルを開発することで網羅的に建物構造・築年代をマッピングする予定である。さらには、抽出した建物画像と GIS データを組み合わせることによって 3 次元都市モデルへの応用も並行して進めていく (中村ら 2021)。

5. まとめ

本稿では、街路の全方位街路画像と建物 GIS データを用いて自動的に建物を画像抽出し、建物画像への建物 GIS データの結合手法を提案した。その結果、道路沿いの建物に対して良好な精度でデータ結合できることが確認できた。

謝辞

本研究を行うにあたり、JSPS20K15001 科研費の助成を受けた。株式会社ゼンリンには本研究で用いた全方位画像データを提供して頂いた。また、都市計画基礎調査は G 空間情報センターより提供して頂いた。全方位街路画像のアノテーションデータの作成にあたっては、株式会社バオバブにご協力頂いた。関係各位に謝意を表する。

参考文献

- T. Oki, and Y. Ogawa (2021) Model for Estimation of Building Structure and Built Year Using Building Facade Images and Attributes Obtained from a Real Estate Database. In *Urban Informatics and Future Cities*, 549-573, Springer.
- K. He, G. Gkioxari, P. Dollár, R. Girshick, R.(2017) "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*. pp. 2961-2969.
- 中村 遼斗・佐藤 剛・小川 芳樹・前田 紘弥・関本 義秀 (2021), 街路の全方位画像に基づくテクスチャ付き 3 次元建物モデルの自動生成手法の検討, 「地理情報システム学会論文集」, Vol. 30.