大規模点群におけるノイズ自動除去システムに関する研究
－深層学習による画像認識と 3 次元点群の幾何学的処理を組み合わせた高精度マッチング－

小村風我*・吉田大介・植田凌生*****

A Study on an Automatic Denoising System for Large-Scale Point Clouds
- High-Accuracy Matching through the Integration of Deep Learning-based Image Recognition and Geometric Processing of 3D Point Clouds -

Fuga Komura*, Daisuke Yoshida**, Ryosei Ueda***

Abstract: In recent years, acquiring 3D point cloud data has become easier, but manually removing noise such as people captured during measurement remains labor-intensive and poses a significant challenge for large-scale scans. To address this issue, this research developed a system that detects people outdoors using point clouds and images acquired by a terrestrial laser scanner, then removes them as noise.The proposed method first extracts people from images using an object detection model (YOLOv8) and associates them with the corresponding 3D point cloud. Next, applying Principal Component Analysis (PCA) to the point cloud clearly distinguishes people from flat structures like walls, significantly reducing false detections that are difficult to avoid with image recognition alone. By selectively and precisely removing only the noise, this system contributes to expanding the application range of 3D point cloud data.

Keywords: 3 次元点群（3D point cloud），ノイズ除去（noise removal），画像認識（image recognition），深層学習（deep learning），主成分分析（principal component analysis）

１．はじめに

近年，レーザースキャナの低価格モデルが普及してきたことによって，3 次元点群データの取得がこれまで以上に容易になっている．スマートフォンやタブレットへの LiDAR 搭載に加え，地上設置型レーザースキャナ（TLS）や車両搭載型の計測システム（MMS）の普及により，さまざまな場所で文化財保全やデジタルツインといった用途で 3 次元点群データが活用されるようになった（Remondino, 2011）．しかし，点群データの活用には，計測時に取得してしまう不必要な点群（例えば，人間や車など）の除去作業は大きな課題である．現在，これらのノイズ除去は主として人手に依存しており，大規模な 3 次元計測では作業効率が深刻な課題となっている．特に街や建物を対象とした数十億点規模の点群では，膨大な時間と労力を要する．点群のノイズ除去に関しては，センサ融合に基づくアプローチ（板倉ほか，2024）や統計・幾何的特徴に基づく自動除去（峰岸ほか，2022）が報告されている．一方で，画像認識の分野では Mask R-CNN（He et al., 2017）や YOLO （Jocher et al., 2023）といった高精度な検出モデルが発展しており，これらを点群処理に応用する試みも行われてきた．たとえば Frustum PointNets（Qi et al., 2018）は，2 次元の画像上の物体検出結果から 3 次元的な候補領域を切り出し，その内部の点群に対して 3 次元の検出ネットワークを適用する効率的な手法を提案した．また，PointPainting（Vora et al., 2020）は画像セグメンテーションの結果を点群に付与することで融合的に物体検出を実現している．これらは，2 次元の検出結果を点群に反映させるという点で本研究と方向性が通じているが，いずれも検出後の点群に対して幾何学的特徴を活用した誤検知抑制までは十分に踏み込んでいない．そこで本研究では，画像認識と幾何学的特徴分析に基づく厳密なフィルタリングを組み合わせることで，高精度なノイズ自動除去システムを開発し，点群データの取得から活用までの流れを省人化・効率化することを目的とする．

２．提案手法

本研究では，3 次元点群データにおけるノイズ除去の効率と精度を向上させるため，深層学習を用いた画像認識と，点群の幾何学的特徴解析を組み合わせた手法を提

*　学生会員　大阪公立大学大学院情報学研究科（Osaka Metropolitan University）
〒558-8585　大阪府大阪市住吉区杉本 3-3-138　E-mail：sd24608k@st.omu.ac.jp
**　正会員　大阪公立大学大学院情報学研究科（Osaka Metropolitan University）
*** 非会員　大阪公立大学現代システム科学域（Osaka Metropolitan University）

案する．多くの 3 次元レーザースキャナが点群と同時に画像を取得する特性に着目し，まず画像認識によってノイズ候補を大まかに特定し，画像と点群のマッチングを行う．次に，特定された候補点群に対し，主成分分析やクラスタリングといった幾何学的処理を適用することで，人間と背景構造物との識別精度を高め，誤検知を効果的に抑制する．本手法の処理フローの全体像を図 1 に示す．

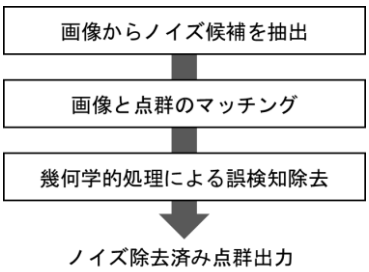


図 1　ノイズ自動除去システムの処理フロー

２．１　画像からノイズ候補の抽出

本手法の第一段階では，点群と同時に取得した 2 次元画像を活用し，ノイズとなりうる人間の領域をピクセル単位で精密に抽出する（図 2）．この処理には，インスタンスの輪郭まで識別可能なセマンティックセグメンテーションモデルである YOLOv8 を採用した．このモデルを画像全体に適用することで，人間の形状に沿ったセグメンテーションマスク（図 2 の赤色の領域）を生成し，対象となるピクセル座標群を正確に把握する．これにより，後続の点群マッチング処理において，矩形のバウンディングボックスを用いる手法よりも背景の誤検出を初期段階で大幅に抑制し，処理対象をより厳密に絞り込むことが可能である．



図 2　YOLOv8 によるセグメンテーション結果
（左：入力，　右：出力）

２．２　画像と点群のマッチング

点群のマッチング処理では物体検出モデルにより，画像上で特定された 2 次元の候補領域を，3 次元点群データと対応付ける（図 3）．この処理は，レーザースキャナの既知の内部・外部パラメータを用いた座標変換によって行う．具体的には，3 次元点群の各点（世界座標）を，式(2.2.1)を用いてカメラの撮像面におけるピクセル座標へと投影する．

$$s\begin{bmatrix}u\\v\\1\end{bmatrix} = K \cdot R \cdot \left(\begin{bmatrix}X\\Y\\Z\end{bmatrix} - T\right) \qquad (2.2.1)$$

ここで，K はカメラ行列，R は回転行列，T は並進ベクトルを示す．また，**u, v** は画像上のピクセル座標を，**s** はスケーリング係数を表す．この計算により，YOLOv8 が検出したピクセル座標に含まれる 3 次元点群を正確に特定し，後続の幾何学的処理の対象として絞り込める．



図 3　画像と点群のマッチング結果
（左：入力画像，　右：マッチングした点群）

２．３　幾何学的処理による誤検知除去

最後に候補として絞り込まれた 3 次元点群に対し，幾何学的な特徴に基づいた精密な検証を行い，誤検出を抑制する（Weinmann et al., 2017）．この処理は，主成分分析（PCA）とクラスタリングを組み合わせた多段階のフィルタリングで構成される．まず，候補点群に対して PCA を適用し，その形状が壁のような平面構造に近いか，あるいは人間のような立体的な構造に近いかを平面性スコアによって定量的に評価する．平面性が高い点群は誤検出と判断して除去する（図 4）．さらに，残った候補点群にクラスタリング手法（DBSCAN）を適用し，検出したい本体である人間から離れた位置にある、誤検知した無関係な点群をノイズとして除去する（Ester et al., 1996）．除去結果を図 5 に示す．

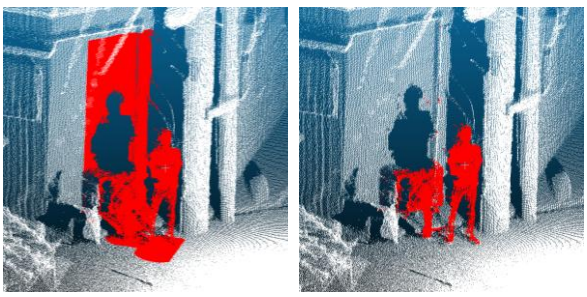


図 4　PCA により壁・地面を除去した結果
（左：PCA 処理なし，　右：PCA 処理あり）

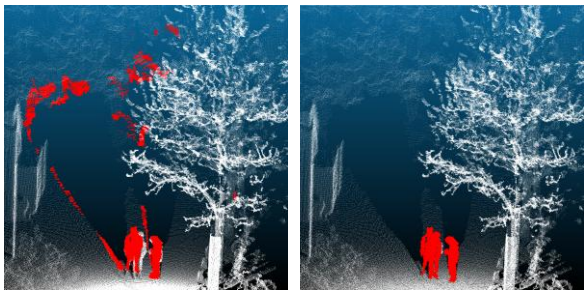


図 5　幾何学的処理により誤検知を除去した結果
（左：誤検知除去なし，　右：誤検知除去あり）

３．手法の検証

３．１　実験環境とデータセット

本検証で用いる 3 次元点群および画像データは，Matterport Pro3 を用いて取得した. 評価用データセットとして，特性の異なる 2 つの環境を選定した．一つは，大阪公立大学杉本キャンパス（60 地点, 80 シーン）である．杉本キャンパスは日常的に多くの歩行者が往来するため，ノイズ源となる人間の検出性能を評価するのに最適であると判断した．もう一つは，重要伝統的建造物群保存地区である大阪府富田林市寺内町（55 地点，68 シーン）である．この場所は壁面などの平面構造物が多く，文化財保全といった実用的な活用場面を想定した誤検出の抑制性能を検証するために適切であると考え選定した．

３．２　評価指標

本研究では，ノイズ除去性能を定量的に評価するため，適合率（Precision），再現率（Recall），そして IoU（Intersection over Union）の 3 つの指標を用いる．各指標は，システムが検出したノイズ点群（Positive）と，手動で付与した正解ラベル（Ground Truth）との比較に基づき，それぞれ式(3.2.1)，式(3.2.2)，式(3.2.3)を用いて算出される．具体的には，システムがノイズとして正しく検出した人間点群を真陽性（TP: True Positive），ノイズとして誤って検出した壁などの背景点群を偽陽性（FP: False Positive），そしてシステムが検出しきれなかった人間点群を偽陰性（FN: False Negative）と定義する．

$$Precision = \frac{TP}{TP + FP} \tag{3.2.1}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.2.2}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{3.2.3}$$

３．３　主成分分析におけるパラメータ設定の検証

人間と壁面を識別する高精度なフィルタを設計するため，まず両者の局所的な幾何学的特徴の差異を定量的に評価した．事前にラベリングした人間 10 体と壁面 10 箇所の点群に対し，各点を中心とする近傍点群に主成分分析（PCA）を適用し，共分散行列の固有値($\lambda_1 \geq \lambda_2 \geq \lambda_3$)から平面性スコア $P = (\lambda_2 - \lambda_3)/\lambda_1$ を算出した．このスコアは，点群が平面に近いほど 1 に，立体的なほど 0 に近づく．

各サンプルの平均平面性スコアに関する記述統計量を図 6 に示す．人間と壁面のスコア分布には明確な分離帯が存在し，重なりが見られなかった．人間の最大スコアが 0. 506, 壁面の最小スコアが 0.681 であったことから，両者を分離する閾値を 0.6 と設定することが統計的に適切であると結論付けた．

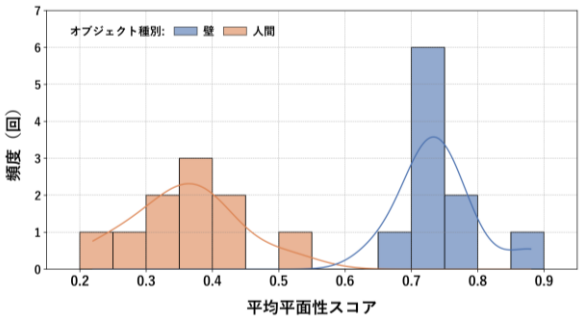


図 6　人間と壁面の平均平面性スコア分布

３．４　ノイズ除去性能の総合評価

前述の 2 つのデータセットを用いてシステム全体のノイズ除去性能を評価し，「画像認識のみの手法」と「提案手法」の性能を比較した．評価結果をそれぞれ図 7（大阪公立大学杉本キャンパス）および図 8（大阪府富田林市寺内町）に示す．

両データセットにおいて，提案手法は「画像認識のみ」の場合と比較して，適合率（Precision）と IoU が大幅に向上した．特に壁面が多い寺内町では，幾何学的特徴に基づいたフィルタが壁の誤検出を効果的に抑制し，適合率が顕著に改善された．これは，PCA による事前検証に基づいたフィルタ設計が，多様な環境下で機能することを示している．

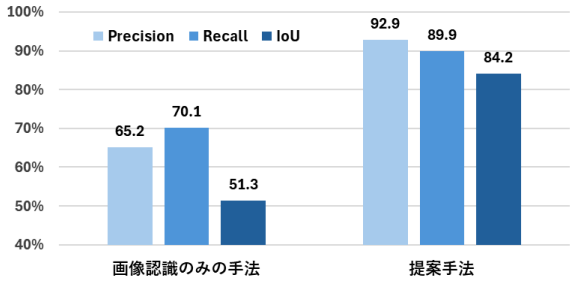


図 7　大阪公立大学杉本キャンパスにおける評価結果

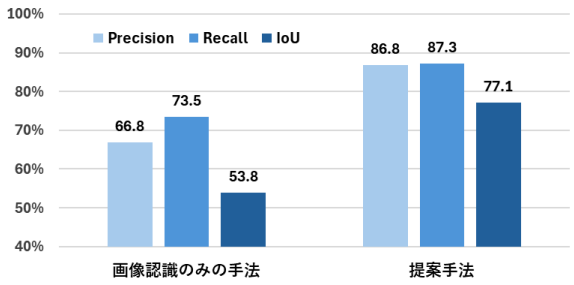


図 8　大阪府富田林市寺内町における評価結果

４．考察

実験結果は，画像認識と，幾何学的検証に裏付けられた点群フィルタリングを組み合わせる提案手法の有効性を示している．画像認識で検知した人間を点群とマッチングさせるのみでは，ずれによる背後の誤検知が多く，これが適合率低下の主な原因であった．提案手法では，PCA やクラスタリングによる形状フィルタが，これらの誤検出（False Positive）を効果的に除去することに成功した．その結果，除去すべきでない構造物を保護しつつ，ノイズのみを選択的に除去することが可能となり，IoU の大幅な向上に繋がったと考えられる．

５．まとめと今後の課題

本研究では，3 次元点群におけるノイズを高精度に自動除去する新たな手法を提案した．本手法は、YOLOv8 による画像認識と，PCA やクラスタリングを用いた幾何学的分析による多段階の点群フィルタリングを組み合わせたものである．評価実験の結果，提案手法による高精度な人間検出が可能であることを確認した．今後の課題としては，静止体のみを対象とした本研究を移動体へ拡張することや，人間以外の多様なノイズ除去への対応が挙げられる．

謝辞

本研究の遂行にあたり，実験場所の提供をしていただいた，大阪府富田林市，岸和田市，和歌山県田辺市に深く感謝します．また本研究は，JST 共創の場形成支援プログラム JPMJPF2115 の支援を受けたものです．

参考文献

板倉健太・林 拓哉・上脇 優人・全邦釘（2024）LiDAR とカメラのセンサーフュージョンによる点群からのノイズ除去．「AI・データサイエンス論文集」，**53**，757-768．

峰岸樹・江守央・佐田達典（2022）点群データに含まれるノイズの統計的・幾何的手法を用いた自動的除去に関する研究．「土木学会論文集 F3（土木情報学）」，**78**（2），149-155．

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, 226-231.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017) Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2961-2969.

Jocher, G., Chaurasia, A., & Qiu, J. (2023) YOLO by Ultralytics. Zenodo. https://doi.org/10.5281/zenodo.7347926.

Qi, C. R., Liu, W., Wu, C., Su, H., & Guibas, L. J. (2018) Frustum PointNets for 3D Object Detection from RGB-D Data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 918-927.

Remondino, F. (2011). Heritage recording and 3D modeling with photogrammetry and 3D scanning. *Remote Sensing*, 3(6), 1104-1138.

Vora, S., Lang, A. H., Helou, B., & Beijbom, O. (2020) PointPainting: Sequential Fusion for 3D Object Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4604-4612.

Weinmann, M., Jutzi, B., & Mallet, C. (2017) Geometric features and their relevance for 3D point cloud classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **IV-1/W1**, 157-164.