

摘要

近年の建築では BIM (Building Information Modeling) と呼ばれるコンピュータ上に現実と同じ建物の立体モデルを再現し、可視化するワークフローが注目されている。従来の配管 BIM は高精度な Lidar センサを用いて配管モデルの推定を行なわれていたが、振動に弱く高価である。そのため、Lidar センサより安価である RGB-D カメラを使用し、従来の点群データのみを用いた 3D 再構築を行わず、取得画像と関連する点群データに基づき配管のアイソメ図を作成を目標とする。

謝辞

本研究の遂行にあたり、ご指導くださった立命館大学理工学部ロボティクス学科馬書根教授、田陽助教に深く感謝の意を表します。また、研究室合同セミにおいて貴重なご意見を頂いた同学科野方誠教授に深く感謝の意を表します。最後に、日頃から研究に対するご指摘、ご協力頂いた生物知能機械研究室の皆様、特にソフトウェア班の皆様に深く感謝の意を表します。

目次

图 目 次

表 目 次

序論

配管は気体、液体、粉粒対などの流体を輸送や配線の保護などを目的とする管のことである。配管は電気配線やケーブルを保護する電気配管や、生活に必要な水を家庭や学校などに輸送する水道管など様々な場面で使用されており、私たちの生活において重要な役割を担っている。そのため、配管を運用するにあたって常に耐久性と安全性を保ち続ける必要性がある。

0.1 研究背景

BIM とは、Building Information Modeling の略称で、コンピュータ上に建築物や土木構造物などの立体モデルを形成し、設計から維持管理までのプロセスをデジタル化する新しいワークフローの一環である。この BIM モデリングはこれまでの 3D モデリングとは大きく異なる。従来の 3 次元モデリングでは平面図などの 2 次元上で作成した図面を元に別途 3 次元のモデルを作成していた。そのため、図面と 3 次元モデルが連動しておらず、設計変更がある度に図面と 3D モデルの両方を修正する必要があり効率的ではなかった。しかし、この BIM 手法は一つのデータを修正すると全てのデータが連動し、関係する図面の該当箇所が自動修正され、従来の方法よりも高効率で作業を行うことが可能になる。

配管は建築物の中でも日常生活に欠かせない存在である。生活に必要な物資を運送したり電線やケーブルを保護するために使用されるなど幅広い面で活用されているため常に耐久性と安全性が求められている。その配管の図面を作成する際にはアイソメトリック（アイソメ）図と呼ばれる立体を斜めから見た視点で表示した等角図が用いられる。このアイソメ図の最大の特徴が図面を見るだけで配管のルートを一目でイメージしやすくなることだ。設計図には平面図や立体図、系統図など様々な種類の図面を使用するが、配管の場合、配管同士が複数にも重なり合っているため左右上下からの視点では見分けることが困難である。そのため、アイソメ図は図面を立体的に描画する手法を扱えるだけでなく、配管のルートや交差する配管の前後関係をイメージしやすくなる。

アイソメ図を取得するためにはこれまでに Light Detection and Ranging (LIDAR) センサーと呼ばれるレーザー光を使用して離れた場所にある物体の形状や距離を測定できるセンサーを使用していた。LIDAR センサーは距離情報を活用して 3 次元情報を取得するだけでなく、測定範囲の広さや取得データの精度が評価されている。しかしその反面、他のセンサーと比較すると高価であるというデメリットを抱えているため、たくさんの人が使用することは困難であるとされていた。このような背景から

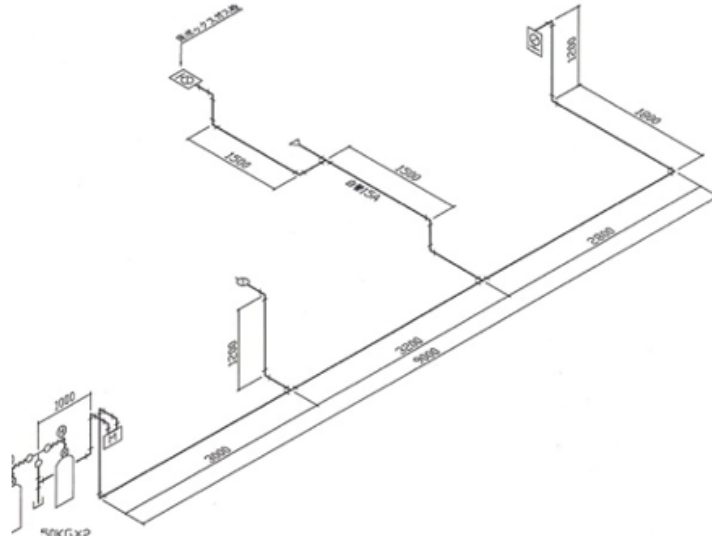


図 1: アイソメ図の例

近年 LIDAR センサーよりも安価な RGB カメラや RGB-D カメラを用いた認識手法が研究され始めている。近年の画像認識分野の研究では機械学習を用いた研究が多く取り上げられている。業務効率化や生産性向上、そして人手不足解消を実現できるというメリットがあり、人工知能の導入は加速していくことが予測されている。

0.2 既存研究

近年、機械学習を用いた物体認識分野では、6D 姿勢推定問題に取り組む研究が盛んに行われている。

物体認識の分野においては物体検出というタスクが存在している。特に、YOLO は高速かつ高精度な物体検出を実現している。このモデルは、入力画像を特定のサイズにリサイズした後、畳み込みニューラルネットワーク（CNN）を用いて特徴を抽出し、特徴マップに基づいて物体の存在確率、境界ボックス、そしてクラスを予測する。さらに、マルチスケール予測を採用することで、大小さまざまなサイズの物体に対しても高い検出性能を発揮する。また、学習時にデータ拡張やアンカー最適化を活用することで、ロバスト性と汎化性能が向上している。YOLO を用いることで、複数配管に対する高精度な物体検出を実現し、6D 姿勢推定の精度向上を図る基盤を構築することができた。この組み合わせにより、従来のモデルでは対応できなかった複数物体を扱う新たな可能性を提供する。

次に 6D 姿勢推定では、3 次元空間上の位置情報（X, Y, Z）だけでなく、回転や向き（Roll, Pitch, Yaw）も求める必要がある。物体の姿勢を正確に推定できれば、配管の向きを把握することが可能となり、アイソメ図作成において大きな利点をもたらす。

6D 姿勢推定には RGB カメラや RGB-D カメラを用いる方法があり、それぞれ異なる特徴と利点がある。RGB カメラを使用する手法の一例として、6D 姿勢推定ネットワーク「Generalizable Model-Free 6-DoF Object Pose Estimation from RGB Images

(Gen6D)」を紹介する。このモデルでは、データセットの作成に「Colmap」というソフトウェアが用いられている。Colmap は「Structure from Motion (SfM)」技術を用いて、複数の視点から撮影された画像を基に 3 次元形状を復元するものであり、その結果得られる点群データを基に 6D 姿勢推定を行う。

Gen6D の処理は大きく 3 つのステップに分かれている。まず「Detector」では、参照画像の情報を基に検出対象物体の領域を特定する。次に「Selector」では、検出された領域に対応する画像の中から、テスト画像に最も近い視点を持つ参照画像を 1 枚選び出す。この参照画像の視点を基に、物体の初期姿勢を形成するが、ここでは多少の誤差が生じる可能性がある。最後のステップでは、初期姿勢を改良するためにさらに 6 枚の近い視点画像を選択する。これらの画像間で平均と分散を計算し、それを基に姿勢情報を改善して最終的な推定結果を得る。ただし、このモデルには事前に SfM を用いて点群データを準備する必要がある、これが大きな手間となる課題である。

次に、RGB-D カメラを用いた 6D 姿勢推定手法の一例として「SAM-6D」を紹介する。SAM-6D では、画像内の物体をピクセル単位で特定する「Instance Segmentation (ISM)」を用いる。この手法では、次の 3 つの情報が抽出される。1 つ目は「Semantics」であり、「Segment Anything Model (SAM)」のゼロショット能力を活用し、未知の物体に対しても領域を特定する。2 つ目は「Appearance」であり、検出された物体の色やテクスチャを基に、既知の物体テンプレートとの類似性を評価する。3 つ目は「Geometry」であり、物体の形状やサイズを基に、バウンディングボックスの重なり具合を計算し、位置や向きに基づいた幾何学的整合性を評価する。

SAM-6D の姿勢推定では、物体の点とオブジェクトモデルの点を対応付ける「ポイントマッチング」を採用している。この手法により、物体の位置と向きを高精度に推定することが可能である。また、オクルージョン（遮蔽）による影響を軽減するため、「バックグラウンドトークン」と呼ばれる仮想点を導入し、欠損領域があっても正確な姿勢推定を実現している。

0.3 研究目的

本研究では、比較的安価な RGB-D カメラを活用し、計算効率に優れた深層学習ベースの手法を提案することで、一般的に利用可能なアイソメ図生成方法の確立を目指す。本研究の主な貢献は以下の通りである。1 つ目は、RGB-D カメラを利用した低コストで汎用性の高いアイソメ図作成方法を提案し、広く使用可能なアプリケーションを提示したことである。2 つ目は、アイソメ図作成に必要な 6D 姿勢推定モデルを比較検証し、本研究に適したモデルを選定したことである。3 つ目は、6D 姿勢推定情報を基に、アイソメ図の CAD 生成手法を提案し、本研究の最終目的であるアイソメ図作成に貢献した。

0.4 本論文の構成

本論文の構成は以下になる。第一章では研究背景、既存研究、研究目的について述べる。研究背景では、Building Information Modeling(BIM) についてや従来のア

イソメ図の取得方法について述べる.

深層学習による配管6D姿勢推定

従来のアイソメ図作成方法では、3次元点群を取得可能なLIDARセンサーを利用して図面を作成していた。しかし、LIDARセンサーは他のセンサーと比較して高価であり、一般的な利用にはコストが課題となっていた。本研究では、LIDARセンサーよりも安価で入手しやすいRGB-Dカメラを用い、データ収集からアイソメ図作成までを効率的に行うための配管6D姿勢推定手法を提案する。本章では、RGB画像およびRGB-D画像を用いた6D姿勢推定手法について詳細を述べる。

0.5 RGB画像を用いた配管6D姿勢推定

RGB画像を用いた深層学習による配管6D姿勢推定の手順は、データ収集、データ処理、認識、データ統合の4つのステップに分類できる。以下に、各ステップについて詳述する。

0.5.1 物体検出

本手法では、物体検出にYOLOv8を使用する。一般的な6D姿勢推定モデルであるGen6Dには物体検出機能が内蔵されているが、複数物体を同時に扱うことができない。そのため、まずYOLOv8を用いて配管を含む複数の物体を検出することとした。YOLOv8の出力は、検出した物体ごとのバウンディングボックス座標およびスケール情報である。これにより、複数の配管を効率的に識別可能とした。

0.5.2 6D姿勢推定

物体検出の結果を基に、Gen6Dを用いて配管の6D姿勢を推定する。具体的には、YOLOv8が出力したバウンディングボックスを基に対象物体の画像を切り取り、これをGen6Dに入力することで姿勢推定を行う。Gen6Dは、Selectorを用いて検出領域内の画像から、事前に用意された参照画像と最も類似する画像を選択する。この参照画像は姿勢データを持ち、これを基に物体の初期姿勢を取得する仕組みである。

さらに、6枚の近似視点を持つ画像を選択し、それらの平均と分散を計算することで初期姿勢を補正し、最終的な姿勢を推定する。最終的に得られる6D姿勢は、オブジェクトの位置(X, Y, Z)および姿勢(Yaw, Pitch, Roll)の情報を含む。

本手法は、従来の単一物体に限定されていた6D姿勢推定を複数物体に拡張することで、産業用途での応用可能性を大きく向上させている。

0.6 RGB-D 画像を用いた配管 6D 姿勢推定

RGB-D 画像を用いた配管 6D 姿勢推定も同様に、データ収集、データ処理、認識、データ統合の 4 つのステップに分けて説明する。本手法では、セグメンテーションと 6D 姿勢推定に SAM-6D を採用している。

0.6.1 セグメンテーション

セグメンテーションは、配管画像をピクセル単位で分類する手法である。物体検出が物体のバウンディングボックスを取得するのみに留まるのに対し、セグメンテーションでは物体の形状をより正確に特定することが可能である。RGB-D 画像を用いることで、RGB 画像からの色情報に加え、Depth 画像から取得したピクセルごとの距離情報を参照できる。この結果、物体の形状を詳細に把握し、精度の高い姿勢推定を実現する。

0.6.2 6D 姿勢推定

SAM-6D による 6D 姿勢推定では、物体表面の点とオブジェクトモデルの点に対応付けるポイントマッチングを用いる。また、オクルージョン（視野遮蔽）が存在する場合でも精度を維持するため、バックグラウンドトークンと呼ばれる仮想的な点を導入している。これにより、欠損領域に対しても正確に姿勢推定を行うことが可能となる。

SAM-6D は、高度なセグメンテーション技術とポイントマッチング手法を統合し、従来の RGB 画像ベースの手法では対応が難しかった複雑なシーンにも対応可能としている。

アイソメトリック図生成

0.6.3 概要と生成手法

本研究では、配管 6D 姿勢推定の結果を基に、アイソメトリック図を生成する手法を提案する。最初に、各配管の接続部の姿勢情報を利用して配管間の接続関係を推定する。アイソメ図を作成するには、接続部のペアを特定し、正確に配管を接続する必要がある。これには、6D 姿勢推定によって得られた位置と姿勢の情報を用いる。次に、推定された接続関係に基づいて、配管を線で結び、アイソメ図を生成する。

0.6.4 接続関係推定

配管間の接続関係を特定するために、接続部の方向ベクトルを計算し、そのベクトル間の角度差が小さい接続部をペアとして認識する。曲管では出口が2つあるため2つの方向ベクトルを、T字管では3つの方向ベクトルを算出する。これらのベクトルの角度差が小さく、かつ距離が最も短い接続部をペアとして認識することで、接続関係を推定する。ただし、地面に接続している配管のようにペアが存在しない場合には、ペアマッチングを適用しない。

0.6.5 配管経路の探索

接続された配管の経路を効率的に探索するため、深さ優先探索 (Depth First Search, DFS) アルゴリズムを使用する。DFS は、グラフ構造内のノードを一つの経路で可能な限り深く進み、行き止まりに達した際に戻って別の経路を探索するアルゴリズムである。この手法により、すべての接続部を効率的に訪問し、再帰的に探索を進めることが可能となる。探索の起点は最も左端に位置する配管とし、接続情報に基づいて配管を順次描画する。

0.6.6 アイソメ図の描画

アイソメ図の描画には、Python の ezdxf ライブラリを使用する。ezdxf は DXF ファイルの作成に特化しており、直線や円弧などの描画が可能である。配管の寸法情報は、カメラ座標系で取得した3次元位置データを基に計算される。配管経路探索で得られた接続情報をもとに、ezdxf を用いて正確なアイソメ図を描画する。この方法により、配管システムの構造を視覚的に表現する図面を効率的に作成することができる。

参考文献

- [1] Author(V. Ferrari, T. Tuytelaars, and L. Van Gool): "Simultaneous object recognition and segmentation from single or multiple model views," *International Journal of Computer Vi-sion* ,vol.67, no.2, pp. 159–188, 2006.
- [2] Author(A. Collet, M. Martinez, and S. S. Srinivasa): " The moped framework: Object recognition and pose estimation for manipulation," *The International Journal of Robotics Re-search* ,vol.30, no.10, pp. 1284–1306, 2011.
- [3] Author(M. Aubry, D. Maturana, A. A. Efros, B. C. Russell, and J. Sivic): " Seeing 3d chairs: Exemplar part-based 2d-3d align- ment using a large dataset of cad models," in *Proceeed- ings of the IEEE Computer Vision and Pattern Recognition (CVPR)* , pp. 3762–3769, 2014.
- [4] Author(M. A. Fischler and R. C. Bolles): " Random sample consen- sus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM* ,vol.24m no.6, pp. 381–395, 1981.
- [5] Author(S. Tulsiani and J. Malik): " Viewpoints and keypoints," in *Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR)* , pp. 1510–1519, 2015.
- [6] Author(M. Schwarz, H. Schulz, and S. Behnke): " Rgb-d object recognition and pose estimation based on pre-trained convolutional neural network features," in *Robotics and Au- tomation (ICRA), 2015 IEEE International Conference on, IEEE* , pp. 1329–1335, 2015.
- [7] Author(Guilhem Cheron, Ivan Laptev, Cordelia Schmid): " P-CNN: Pose-Based CNN Features for Action Recognition," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* , pp. 3218–3226, 2015.
- [8] Author(WANG, Chen): "Densefusion: 6d object pose estimation by iterative dense fusion," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* , pp. 3343–3352, 2019.
- [9] Author(LIU Yuan): "Gen6D: Generalizable model-free 6-DoF object pose esti- mation from RGB images," In: *Computer Vision–ECCV 2022: 17th European Conference* , pp. 298–315, 2022.

- [10] Author(REDMON, Joseph; FARHADI, Ali): "Yolov3: An incremental improvement," *In: Computer Vision–ECCV 2022: 17th European Conference* , 2018.
- [11] Author(FISHER, Alex): "ColMap: A memory-efficient occupancy grid mapping framework," *Robotics and Autonomous Systems* , 2021.
- [12] Author(GIRSHICK, Ross): "Fast r-cnn," *In: Proceedings of the IEEE international conference on computer vision* , pp. 1440–1448, 2015.
- [13] Author(Xian, Yongqin and Choudhury, Subhabrata and He, Yang and Schiele, Bernt and Akata, Zeynep): "Semantic Projection Network for Zero- and Few-Label Semantic Segmentation," *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* , pp. 8256–8265, 2019.
- [14] Author(ZHOU, Tao): "RGB-D salient object detection: A survey," *Computational Visual Media* , 7; 37-69, 2021.
- [15] Author(Hinterstoisser, S., Lepetit, V., Ilic, S., Holzer, S., Bradski, G., Konolige, K., Navab, N.): "Model based training, detection and pose estimation of texture-less 3d objects in heavily cluttered scenes," *In: Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision* ,pp. 548–562, 2013.