

天伏図作成補助システムにおける 全方位カメラを用いた天井設備の多視点画像抽出手法

○小濱大輝 (名古屋大学), 永田吉輝 (名古屋大学), 安武和成 ((株)九電工), 浦野健太 (名古屋大学),
片山晋 (名古屋大学), 米澤拓郎 (名古屋大学), 河口信夫 (名古屋大学)

Multi-View Image Extraction of Ceiling Equipment using 360 Camera in Ceiling Plan Support System
KOHAMA Daiki (Nagoya University), NAGATA Yoshiteru (Nagoya University), YASUTAKE Kazushige (Kyudenko Co.,Ltd),
URANO Kenta (Nagoya University), KATAYAMA Shin (Nagoya University), YONEZAWA Takuro (Nagoya University)
and KAWAGUCHI Nobuo (Nagoya University)

キーワード：天井設備, 天伏図, 全方位カメラ

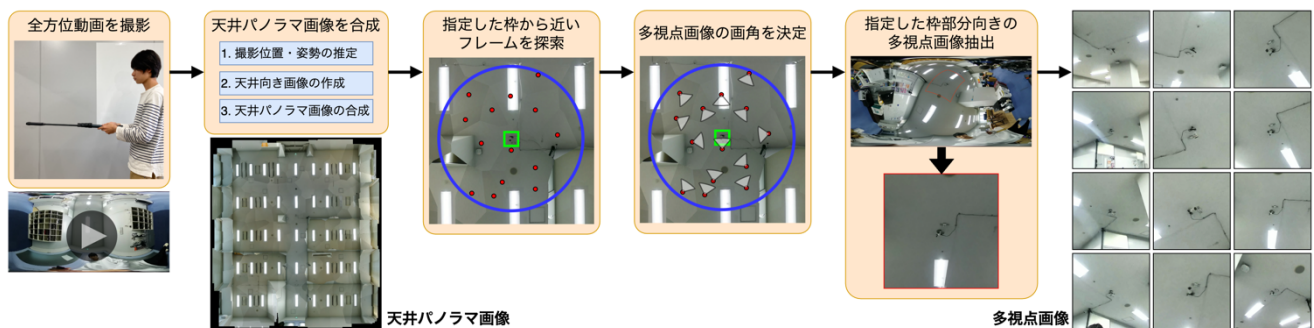


図 1 多視点画像抽出処理の流れ

1. はじめに

建物の設備改修時に電気設備工事を行う場合、工事前に計画立案の目的で、工事箇所の現況を確認する。その際、照明設備やスピーカなどの天井設備が記される天伏図が必要となる。10年以上運用された建物では、図面が残っていなかったり、残っていたとしても、過去の設備改修時の内容が反映されていなかったりする場合が多く、現況を反映した天伏図を得られることは稀である。現況の天伏図を作り直す場合、通常人が目で見て確認しながら作成するが、天井設備の位置を全て記録するには多くの時間を要する。3D レーザスキャナで記録した3次元点群からでも天伏図を作成可能だが、3D レーザスキャナは高価なため限られた現場でしか使用されていない。

我々は、天井面を一樣に写したオルソ画像（以下、天井パノラマ画像）を、室内空間で撮影した全方位動画から合成する技術を研究している。これは、天井面全体を一樣に見られる天井パノラマ画像を天伏図作成時に提示して、その作成を補助する。しかし、天井パノラマ画像は2次元平面に投影した画像なので、天井から突き出た誘導灯やカメラ、プロジェクタなどの設備を確認することは難しい。

本論文では、天伏図作成時に天井から突き出た設備を

様々な方向から見られるように、全方位動画から天井設備の多視点画像を抽出する手法を提案する。さらに、各多視点画像がどの地点から撮影された画像であるかを示すため、パノラマ画像上にそれぞれの撮影地点を示す。天井パノラマ画像上では底面しか見られない天井から突き出た設備を、様々な視点から見られるため、設備種類の特定やその信頼性向上が期待できる。

2. 関連研究

多視点から対象物体を見られる技術として、Structure from Motion(SfM)やMulti-view Stereo(MVS)があり、オープンソースとして広く使用されている [1] [2] [3] [4]。対象物体を様々な角度から撮影した画像より、特徴点マッチングやパラメータ推定などを用いて、物体の構造を3次元的に復元する技術である。しかし、事前に対象物を決めた上で多視点撮影する必要があり、撮影後に多視点から見る対象物を決められる提案手法とは異なる。

対象物が写る複数の画像から、機械学習を用いて多視点画像を生成する技術がある [5] [6] [7]。事前学習したモデルを使用して、入力画像にない視点から見た画像を生成する。この手法も事前に対象物を決めた上で多視点撮影が必要がある。

複数のビデオカメラで撮影した動画から、視聴するカ

メラ視点をシームレスに切り替えられるマルチアングル技術 [8]や、3次元復元結果をもとに新たな視点の映像を作り出すボリュメトリックビデオシステム [9]がある。しかし、これらは事前に複数のカメラを対象環境内に設置する必要があるため、機材費用がかさんだり、別の施設に適用する際に準備の手間がかかったりする。

3. 提案手法

図 1 に天井設備の多視点画像抽出手法の処理の流れを示す。まず、全方位カメラで室内空間を撮影した全方位動画から各フレームの撮影位置・姿勢を推定し、天井パノラマ画像を合成する。次に、天井パノラマ画像上の指定された任意の枠の中心（以下、枠中心）から、一定距離以内の地点で撮影されたフレームを探索する。さらに、多視点画像の画角が一定となるように、枠中心から最も近い地点で撮影されたフレームより焦点距離を定める。最後に、探索された各フレームから指定された天井の枠部分が写る画像を抽出する。

3.1. 天井パノラマ画像の合成

全方位動画の各フレームの撮影位置・姿勢を推定する。各フレームから、推定結果を用いて天井向き画像（以下、天井画像）を作成する。作成した天井画像とその位置関係から、天井パノラマ画像を合成する。

3.1.1. 全方位動画の撮影位置・姿勢推定

全方位動画の撮影位置・撮影姿勢推定には、動画から自己位置推定と環境地図作成が可能な Visual Simultaneous Localization and Mapping(VSLAM)を使用する。本研究では、全方位動画に対して VSLAM が可能な OpenVSLAM [10]を用いる。OpenVSLAM は、画像特徴点によって構成される撮影した環境の 3 次元点群と、一定間隔で挿入されるキーフレームに対しての相対的な撮影位置と撮影姿勢を出力する。

OpenVSLAM の出力である 3 次元点群や撮影位置の座標系（以下、VSLAM 座標系）は、動画の 1 フレーム目の撮影姿勢によって傾くため、画像処理と 3 次元点群解析の結果から天井方向推定を行い、傾きを補正する。また、VSLAM 座標系と現実世界の座標系のスケールは不明であるため、対応する 2 点間の距離を比較して、VSLAM 座標系の 1 単位が 1[mm]となるように補正する。

3.1.2. 天井向き画像の作成

天井画像作成のために、撮影位置・撮影姿勢の推定結果を用いて、 n 番目のキーフレームにおける、撮影位置から天井までの距離 h_n [mm]を求める。また、Z 軸が光軸となるカメラ座標系から、Z 軸が天井方向で原点が同じ天井座標系への回転行列 R_n を求める。さらに、作成する天

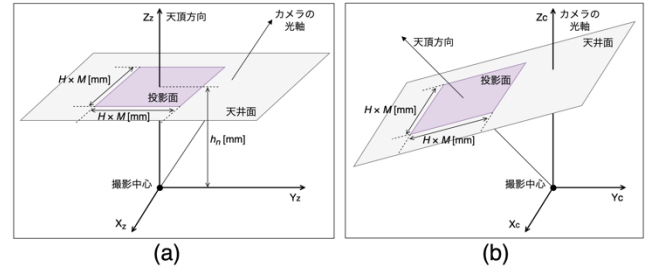


図 2 (a)天井座標系と(b)カメラ座標系における投影面

井画像の一辺の長さを H [pix]、その画像の 1 ピクセルあたりの大きさを M [mm]とする。天井面が天井方向にあると仮定すると、図 2(a)のように、撮影中心から天井方向に距離 h_n [mm]の面に、縦横 $H \times M$ [mm]の天井画像の投影面を天井座標系上で考えられる。投影面上の各点を全方位画像上の点に対応付けるために、その投影面を回転行列 R_n に従って回転させ、図 2(b)のようにカメラ座標系で表す。カメラ座標系上の任意の点は、撮影中心からその点に対する向きによって全方位画像上の点に対応付け可能なので、天井座標系上に投影した天井画像の投影面上の各点も、全方位画像上の点に対応付け可能となる。その対応付けをもとに、天井画像を作成する。

3.1.3. 天井パノラマ画像の合成

全方位動画中のキーフレームに対して推定された撮影位置から、天井画像の解像度 M [mm/pix]を考慮して、天井パノラマ画像の座標系における各天井画像の位置[pix]を算出する。上端と下端に位置する画像の座標及び左端と右端に位置する画像の座標から、それぞれ天井パノラマ画像の縦横の長さ[pix]を計算する。天井パノラマ画像の各ピクセルに対して、天井画像の中心までの距離[pix]を求め、最も距離の短い天井画像から天井パノラマ画像上のそのピクセルに画素値をコピーする。

3.2. 指定した枠から近いフレームの探索

天井パノラマ画像上で指定された枠中心 (x_c, y_c) [pix]から各天井画像の中心 (x_n, y_n) [pix]までの距離を求める。枠中心から水平方向に距離 l [mm]以内の撮影位置となるフレームを、式(1)に従って探索する。ここで、天井画像の中心とその撮影位置の x, y 座標は一致する。また、後の処理のために、最も距離の短いキーフレーム番号 n_s も求める。

$$\sqrt{(x_c - x_n)^2 + (y_c - y_n)^2} \leq \frac{l}{M} \quad (1)$$

各多視点画像がどの地点から撮影されたかを示すために、探索された各天井画像の中心から枠中心に向けてのびる矢印を、天井パノラマ画像上に描画する。

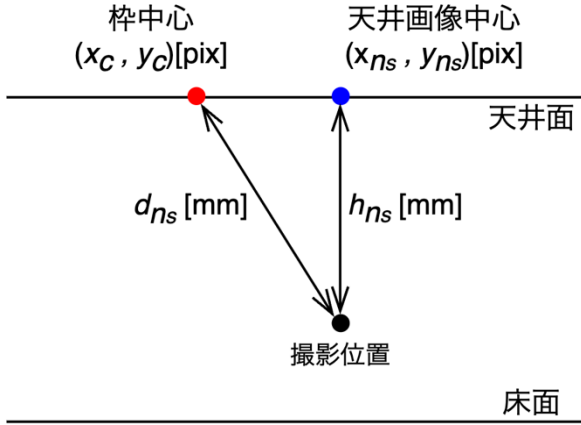


図 3 水平方向から見た天井画像中心、撮影位置、桙中心の位置関係

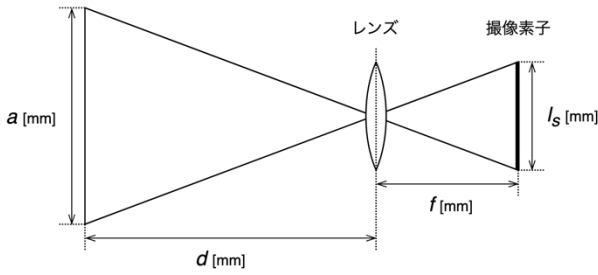


図 4 レンズモデル

3.3. 多視点画像の画角の決定

多視点画像が同一の単眼カメラで各位置から撮影されたように見せるために、多視点画像の画角を一定にする。そのために、焦点距離を定める。焦点距離は、撮影位置が桙中心に近くても桙全体を多視点画像に写せるように、桙中心から撮影位置までの水平距離が最も短い、前項で求めたキーフレーム番号 n_s のフレームより決定する。

キーフレーム番号 n_s の撮影位置から桙中心までの距離 d_{n_s} [mm]は、図 3 より式(2)で表せる。

$$d_{n_s} = \sqrt{M^2 \left\{ (x_c - x_{n_s})^2 + (y_c - y_{n_s})^2 \right\} + (h_{n_s})^2} \quad (2)$$

また、多視点画像の縦横の長さをそれぞれ H' , W' [pix]とする。指定された長方形の桙の中全てが写るように、キーフレーム番号 n_s の撮影位置から写す多視点画像の解像度を、桙の対角距離 L [pix]と W' から式(3)のように求め、 M' [mm/pix]とする。

$$M' = \frac{LM}{W'} \quad (3)$$

さらに、焦点距離 f [mm] は、撮像素子の長さを l_s [mm]とすると、図 4 より式(4)のように表せる。

$$f = \frac{l_s}{a} d = \frac{l_s d}{M' H'} \quad (4)$$

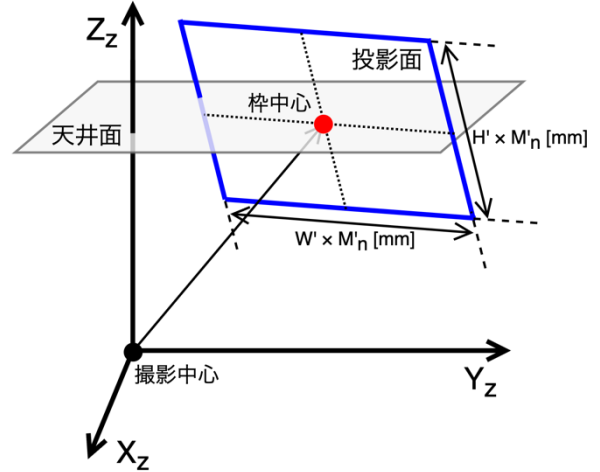


図 5 天頂座標系における多視点画像の投影面

3.4. 指定した桙部分向き多視点画像抽出

天井画像作成時と同様に、天頂座標系上で多視点画像に写す投影面を定める。その投影面と全方位画像中の画素の対応付けから多視点画像を抽出する。

まず、探索済みの各キーフレームに対して、多視点画像の解像度を決定するために、その撮影位置から桙中心までの距離 d_n [mm]及び焦点距離 f を用いて、解像度 M'_n [mm/pix]を式(5)から求める。

$$M'_n = \frac{l_s d_n}{f H'} \quad (5)$$

次に、各キーフレームに対して、桙中心を自然と見上げたような多視点画像を抽出するために、図 5 に示す、上下の辺が水平面に並行で、撮影中心と桙中心を結ぶ線分が法線となる長方形の投影面を、天頂座標系上で考える。天井画像作成時と同様に、回転行列 R_n によって天頂座標系上の投影面をカメラ座標系に変換する。カメラ座標系上の任意の点は、全方位画像上の点に対応付け可能なため、多視点画像の投影面上の各点は、全方位画像上の点に対応付け可能である。その対応付けから、指定した桙部分向き多視点画像を全方位画像から抽出する。

4. 実験

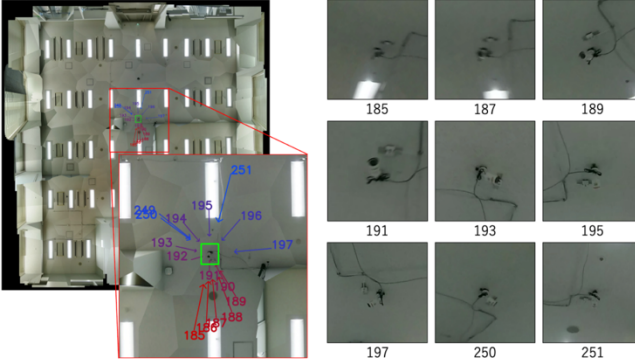
2 つの異なる室内環境で撮影した全方位動画から、合成した天井パノラマ画像上で天井設備を桙で囲み指定する。指定した桙部分向き多視点画像の抽出と、天井パノラマ画像上にその撮影地点の描画を行う。桙中の天井設備が、抽出した多視点画像中に写っているかを確認する。

4.1. データ収集

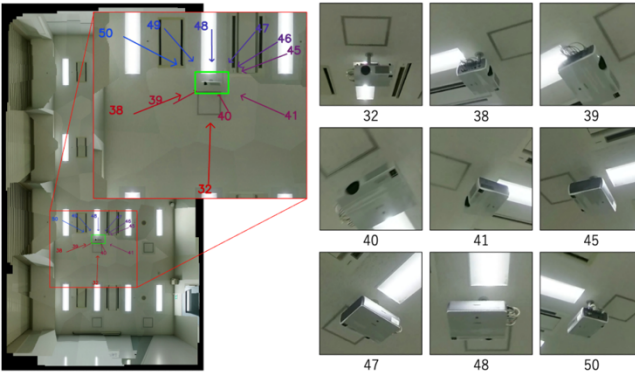
データ収集は、2 つの異なる室内環境で行った。1 つ目は約 273m²の室内空間で、デスクやモニター、椅子などが置かれたワークスペースのような環境で、2 つ目は約 86m²の室内空間で、デスクの上にパソコンが整列し

表 1 収集したデータ

	室内環境 1	室内環境 2
動画の長さ[秒]	211	91
フレームレート[fps]	29.97	29.97
解像度[pix]	3840×1920	3840×1920



(a)



(b)

図 6 室内環境 1(a)と室内環境 2(b)における
多視点画像抽出結果

ている計算機室である。全方位動画は、RICOH 社製の THETA Z1 を使用して撮影した。表 1 に収集したデータの詳細を示す。

4.2. 実験結果

図 6 に、それぞれの環境において、天井パノラマ画像上で 1 箇所ずつ指定した枠向きの多視点画像の一部及び、その撮影地点を記載した天井パノラマ画像を示す。室内環境 1 については天井に備え付けのカメラを、室内環境 2 については天井から吊るされたプロジェクタを、それぞれ枠で囲み指定した。枠中心から近いフレームを探索するときのパラメータ l は、2000[mm]とした。

室内環境 1, 2 両方で、枠で囲った天井設備を多視点画像上で写せている。特に、室内環境 2 においては、天井パノラマ画像上で囲んだ枠部分にある設備がプロジェクタであると判断し難い。しかし、多視点画像では、枠部分の設備がプロジェクタであると確実に判断できた。

5. まとめと今後の展望

合成した天井パノラマ画像中の任意の枠で囲った設備を、全方位動画から抽出して多視点画像とする手法を提案した。本実験では多視点画像を抽出する各撮影位置までの水平距離を 2000[mm]としたが、この値を大きくすると壁によって区切られた撮影地点のフレームが探索されてしまい、本来写したい枠部分が壁で遮られた多視点画像を抽出してしまう。そのような枠部分を写せないフレームを除くフレーム探索手法を考案する必要がある。

謝辞

本研究の一部は、JST CREST JPMJCR22M4, NICT 委託研究 22609 に支援いただいています。

参考文献

- [1] M. Pierre, M. Pascal, P. Romuald and M. Renaud, "OpenMVG: Open multiple view geometry," in *International Workshop on Reproducible Research in Pattern Recognition*, 2016, pp. 60-74.
- [2] Mapillary, "OpenSfM," [オンライン]. Available: <https://github.com/mapillary/OpenSfM>.
- [3] S. L. Johannes and F. Jan-Michael, "Structure-from-Motion Revisited," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [4] C. D., "OpenMVS: Open Multiple View Stereovision," [オンライン]. Available: <https://github.com/cdcseacave/openMVS/>.
- [5] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," 2020.
- [6] P. Keunhong, S. Utkarsh, B. T. Barron, B. Sofien, G. B. Dan, S. M. Steven, M.-B. Ricardo, "Nerfies: Deformable Neural Radiance Fields," 2021.
- [7] Y. Alex, Y. Vickie, T. Matthew, K. Angjoo, "pixelNeRF: Neural Radiance Fields From One or Few Images," 2021.
- [8] Swipe Video, [オンライン]. Available: <https://swipevideo.jp/>. [アクセス日: 29 5 2023].
- [9] キヤノン, "キヤノン ボリュームトリックビデオシステム," [オンライン]. Available: <https://global.canon/ja/vvs/>. [アクセス日: 29 5 2023].
- [10] S. Shinya, S. Mikiya, S. Ken, "OpenVSLAM: A Versatile Visual SLAM Framework," 著: *Proceedings of the 27th ACM International Conference on Multimedia*, 2019.