

全方位カメラによる設備外形を考慮した 天井パノラマ画像合成手法

小濱 大輝¹ 永田 吉輝¹ 安武 和成² 浦野 健太¹ 片山 晋¹ 米澤 拓郎¹ 河口 信夫^{1,3}

概要：運用中の建物の天井改修において、天井の現況を示したデータを一から作るには人的・経済的コストがかかる。安価な全方位カメラで撮影した動画から天井を一様に写したオルソ画像を合成できれば、コストを抑えられる。本研究では、天井を一様に写す天井パノラマ画像が、忠実に実際の天井を再現できるように、合成時に天井パノラマ画像上の各地点が最も綺麗に写るような合成手法を提案する。また、天井設備種類特定に役立つように、設備外形が破綻なく写ることにも考慮する。設備外形が破綻なく写された天井パノラマ画像は、設備が1枚の画像から合成されるため、その特徴を確実に捉えられる。実験では、個室が複数ある約273m²の室内空間で全方位動画を撮影し、データを収集した。合成したパノラマ画像を精度及び同一設備の合同性の観点で評価したところ、提案手法はベースラインと比較して高精度で、視認性が高いことを確認した。

Panoramic Ceiling Image Synthesis Method Considering Fixture Outline using 360° Camera

Daiki KOHAMA¹ Yoshiteru NAGATA¹ Kazushige YASUTAKE² Kenta URANO¹ Shin Katayama¹
Takuro YONEZAWA¹ Nobuo KAWAGUCHI^{1,3}

1. はじめに

デジタルツインの活用や施設の維持管理に関するファシリティマネジメント [1] のために、建物の3次元モデルの需要が高まっている。建設時の効率化・高度化のために Building information Modeling (BIM)/Construction Information Modeling (CIM) などの3次元モデルを使用して建てた建物では、建設後もそれらのモデルを活かすための研究が多くなされている [2], [3], [4]。しかし、既存の建物の多くは、2次元図面を用いて建設され、10年以上運用された建物では、過去の設備改修時の内容が反映されていなかったり図面が残っていなかったりする場合も多く、現況を完璧に反映した図面を得られることは稀である。

建物の設備改修時に電気設備工事をする場合、工事前に

見積もりや部品発注、工事の計画立案の目的で工事箇所の現況を確認する。その際、現況を反映した天伏図が必要となる。天伏図とは、照明器具やスプリンクラー、天検口などの天井設備が記された図面である。正確な天伏図がない場合は、通常人が目視で確認をしながら天伏図を作成する。目視で作成する場合、天井設備の位置を全て記録するには多大な時間がかかる。図2に天井設備を目視で記録した場合と3Dレーザスキャナの測定結果から記録した場合の精度差について示す。目視で記録した時には、本来あるはずの照明やスプリンクラーを見落としている。3Dレーザスキャナを使用すれば、高精度に天井設備を測定できるが、高価なため限られた現場でしか使用されていない。

本研究では、図1に示すように、全方位動画のフレームから作成した天井画像から、設備外形を考慮して天井を一様に写す天井パノラマ画像合成手法を提案する。天井パノラマ画像は、広域な天井を一様に写すため、天伏図を作成する際に有用な情報源となる。ここで、天井画像とは、全方位画像を撮影中心から天頂方向に垂直な面に対して、透

¹ 名古屋大学大学院 工学研究科
Graduate School of Engineering, Nagoya University

² (株) 九電工 技術開発部技術開発課

³ 名古屋大学 未来社会創造機構
Institutes of Innovation for Future Society, Nagoya University

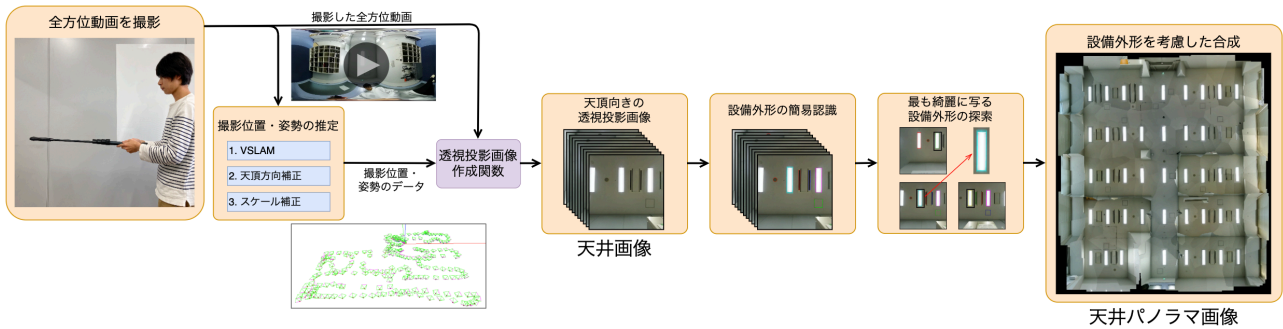


図 1 天井パノラマ画像合成の処理の流れ

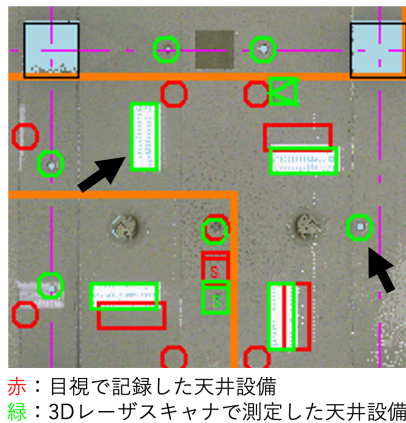


図 2 目視と 3D レーザスキャナで天井設備を測定した場合の精度差

視投影した画像である。また、天井パノラマ画像とは、天井画像同士を合成して天井面においてオルソ画像となるような広域を写す画像である。

本研究では、全方位カメラで撮影された全方位動画を使用して天井パノラマ画像を合成する。全方位カメラとは、撮影地点から全方向を撮像可能なカメラで、360 度カメラや全天球カメラとも呼ばれる。また、全方位動画とは、全方位カメラで撮影した撮影地点から全方向を動画として記録したものであり、360 度動画や全天球動画とも呼ばれる。提案手法では正距円筒図法で表される全方位動画を使用する。また、前処理の撮影位置・撮影姿勢の推定に、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。単眼カメラで天井を撮影すると、一般的に画像特徴点の少ない天井のみが記録されるのに対し、全方位カメラは 360 度全方向を撮影しているため、天井以外の場所からも画像特徴点を得ることができ、高精度な撮影位置・姿勢推定が期待できる。

本研究で提案する設備外形を考慮した天井パノラマ画像合成手法は、忠実に実際の天井面を再現できるように、合成時に天井パノラマ画像上の各地点が最も綺麗に写り、設備外形が破綻なく天井パノラマ画像上に写ることを目的としている。設備外形が 1 枚の画像から合成されるため、その特徴を確実に捉えられ、天井設備種類特定に役立つ。

実験では、約 273 m² の個室がある室内空間と、個室のない開けた約 86 m² の室内空間で、全方位動画を撮影して天井パノラマ画像を合成した。精度及び設備外形の形状を評価をしたところ、提案手法は既存の商用製品のベースラインと比較して、高精度であることを確認した。本研究の貢献を以下にまとめる。

- 全方位カメラを使用して天井面を一様に写す天井パノラマ画像合成手法を確立した。
- 天井設備外形を考慮した天井パノラマ画像の合成手法を提案した。
- 実際の室内で天井パノラマ画像を合成し精度及び解像度の評価を行った。

2. 関連研究

2.1 VSLAM を応用した研究

VSLAM を応用した研究として、電動車椅子の運転支援を行なった研究 [5] がある。車椅子の周辺状況をセンシングする際に、全方位動画から VSLAM を使用してカメラの姿勢と疎な 3 次元モデルを作成した。

Kayukawa らは複数階に跨って全方位カメラで撮影した動画の経路を VSLAM で特定し、経路上の 2 地点を指定するとその間のウォークスルーが生成される研究 [6] を行ない、科学館や空港施設で検証を行なった。

2.2 室内空間をモデリングする研究

全方位カメラを用いた室内空間をモデリングする技術として、1 枚の全方位画像からレイアウト推定を行う研究 [7], [8], [9], [10] が盛んに行われている。しかし、1 枚の全方位画像からモデリング可能な範囲は限られており、広い室内空間を測定したい場合には向かない。

Wang らはレイアウト推定精度向上のために、2 枚の全方位画像から撮影地点及び姿勢を考慮したレイアウト推定 [11] を行った。1 枚の全方位画像から空間モデリングを行う場合よりも、広い範囲で推定可能だが、商業施設やオフィスなど全方位画像 2 枚では、測定しきれない広い空間での測定は不可能である。また、狭い室内空間で推定できたとしても、解像度が高く視認性のある天井パノラマ画像

の合成は期待できない。

建築物のデジタル化やインデックス化を行う Matterport 社の商用ソフトウェア Matterport では、複数地点での全方位画像を用いて天井パノラマ画像を作成できるが、測定時に複数の画像から被写体を 3 次元に復元する技術の知識がないと測定がうまくいかない場合がある。また、Matterport の出力する天井パノラマ画像は天井設備が画像間の切れ目にあったり、解像度が低かったりするため、視認性が乏しい。

提案手法では、測定環境が広い空間であっても利用可能である。全方位動画を入力とするため、静止画を何枚も撮影するよりも手間も少ない。さらに、視認性向上のために、天井設備外形が破綻なく写るように考慮する。

3. 提案手法

図 1 に示したとおり、まず室内を全方位カメラで撮影した動画から、フレーム毎の撮影位置・撮影姿勢推定結果をもとに、天井画像を作成する。撮影位置・撮影姿勢推定のために、撮影開始時は天井にサイズが既知の矩形オブジェクトがある場所から、撮影を開始する。次に、天井画像上の設備外形の認識を行い、最も綺麗に写る設備の探索をする。最後に、認識した設備外形が破綻なく写るように考慮して、天井パノラマ画像を合成する。

3.1 節では、撮影姿勢・撮影位置の推定や天井画像作成時に使用する正距円筒図法の全方位画像から透視投影画像を作成する関数について説明する。3.2 節では、VSLAM を使用した撮影位置・撮影姿勢の推定及びその補正方法について述べ、3.3 節で、撮影位置・撮影姿勢を考慮した天井画像の作成方法について説明する。3.4 節では、OpenCV を用いた天井画像中の設備外形簡易認識手法について説明し、3.5 節で、最も綺麗に映る設備外形の探索手法について述べる。3.6 節では、設備外形を考慮した天井パノラマ画像合成手法について説明する。

3.1 透視投影画像作成関数

正距円筒図法の縦横の長さが $H[\text{pix}]$, $W[\text{pix}]$ の全方位画像からカメラ座標系の回転を考慮した透視投影画像を作成する関数について説明する。透視投影画像に写したい 3 次元空間上の点を、全方位画像上の点に対応付けて作成する。

Z 軸が光軸上にあるカメラ座標系から、天頂向きを Z 軸の正方向とする原点が同じ天頂座標系への回転行列を \mathbf{R} とする。透視投影画像に写す天頂座標系空間上の点の集合を P_z とする。縦横の長さが $H[\text{pix}]$ の透視投影画像に、天頂座標系での $Z = d[\text{mm}]$ となる平面を写す。この時、その平面を 1 ピクセルあたり $M[\text{mm}]$ で透視投影画像に写すとすると、 $P_z(x_z, y_z, z_z)$ は式 (1) を満たす。

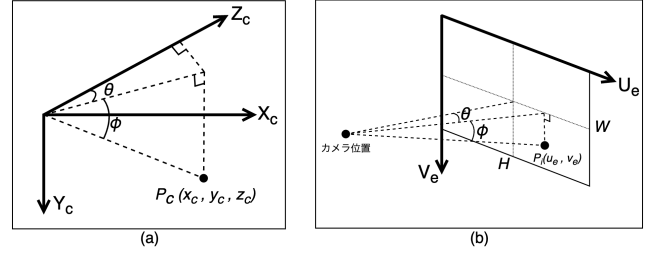


図 3 カメラ座標系 (a) と正距円筒座標画像 (b) の方位角と仰俯角

$$\begin{cases} -\frac{H}{2}M \leq x_z < \frac{L}{2}M \\ -\frac{H}{2}M \leq y_z < \frac{L}{2}M \\ z_z = d \end{cases} \quad (1)$$

P_z を天頂座標系からカメラ座標系に座標変換したものを P_c とすると、 P_c はこの回転行列 \mathbf{R} を用いて式 (2) で表される。

$$P_c = \mathbf{R}P_z \quad (2)$$

カメラ座標系上の $P_c(x_c, y_c, z_c)$ と全方位画像上の点 $P_i(u_e, v_e)$ を、図 3 に示す方位角 θ 、仰俯角 ϕ を介して対応づける。それぞれの θ, ϕ は式 (3), (4) で表される。

$$\begin{cases} \theta = \arccos \frac{x_c}{z_c} \\ \phi = \arccos \frac{y_c}{\sqrt{x_c^2 + z_c^2}} \end{cases} \quad (3)$$

$$\begin{cases} \theta = \left(\frac{u_e - W/2}{W} \right) 2\pi \\ \phi = \left(\frac{v_e - H/2}{H} \right) \pi \end{cases} \quad (4)$$

対応付けられた全方位画像上の点 P_i を射影して透視投影画像を作成する。

以下本論文では、全方位画像から透視投影画像を作成する際に、 (\mathbf{R}, d, M) のパラメータを与えることで作成できるとする。

3.2 VSLAM による撮影位置・姿勢推定及びその補正

本研究では、全方位動画を扱える OpenVSLAM [12] を使用する。室内を撮影した正距円筒図法の全方位動画の撮影位置・姿勢の推定を、OpenVSLAM で行う。OpenVSLAM の出力の内、提案手法で使用する値は以下の 3 つである。ただし、 k は一定間隔で挿入されるキーフレームのインデックス番号を表す。

- 3 次元復元の空間座標系 (以下、VSLAM 座標系) 上で、各キーフレームの撮影位置 p_{ck}
- VSLAM 座標系上で表される、各キーフレームのカメラ座標系から VSLAM 座標系への回転行列 \mathbf{R}_{ck}
- VSLAM 座標系上で表される、全方位動画中の特徴点から構成される 3 次元点群 PC

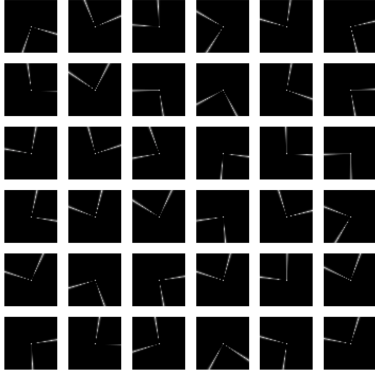


図 4 エッジフィルタ

3.2.1 VSLAM 出力の補正

VSLAM 座標系は、1 フレーム目のカメラの傾きに依存し、現実空間の座標系（以下、ワールド座標系）とのスケールは不明である。天井画像作成時に、透視投影画像に写すカメラ座標系上の P_c を求めるには、カメラ座標系から天頂座標系への回転行列 \mathbf{R} を求める必要がある。VSLAM 座標系では、天頂方向が不明なため、全方位動画の 1 フレーム目で天頂方向推定を行い、水平面に対する VSLAM 座標系の傾きを求める。推定された傾きから、VSLAM 座標系とワールド座標系の水平面が一致するように、各キーフレームの撮影位置 p_{ck} 、回転行列 \mathbf{R}_{ck} 、3 次元点群 PC を回転させる。

また、天井画像作成時、透視投影画像に写す平面を決めるために、各撮影中心から天井面までの距離 $d = h_k[\text{mm}]$ を求める必要がある。そのため、OpenVSLAM 出力のスケール推定を行った西田らの手法 [13] と同様に、VSLAM 座標系とワールド座標系の間のスケールの推定を行う。推定されたスケールから、VSLAM 座標系の一単位が $1[\text{mm}]$ となるように、各キーフレームの撮影位置 p_{ck} に倍率の適用を行う。

3.2.1.1 1 フレーム目の天頂方向推定

天頂方向推定は 2 段階の処理で行う。まず、全方位動画の 1 フレーム目から回転を考慮せずに透視投影画像を作成する。作成した透視投影画像から天井面上の矩形物体を認識する。認識した矩形の 4 頂点が真の矩形の位置関係となるような、透視投影画像作成時のパラメータ \mathbf{R} を探索する。探索した回転行列 $\mathbf{R} = \mathbf{R}_i$ で各キーフレームの撮影位置 p_{ck} 、回転行列 \mathbf{R}_{ck} 、3 次元点群 PC を回転させる。

次に、OpenVSLAM 出力の 3 次元点群 PC から天井面の推定を行う。推定した天井面の傾きから VSLAM 座標系の傾きを修正する。

画像認識を用いた天頂方向推定

全方位動画の 1 フレーム目から透視投影画像を作成する。作成時のパラメータは、カメラ座標系の回転を考慮しないために \mathbf{R} を単位行列とする。また、一般的なオフィスの天井高を $2800[\text{mm}]$ として撮影時に床から $1000[\text{mm}]$ の

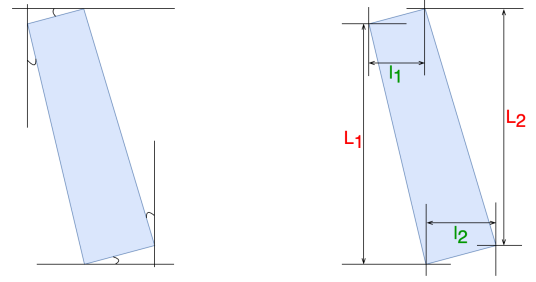


図 5 変換角度算出方法

位置にカメラを持ってくることを想定して $d = 1700[\text{mm}]$ とする。その際に天井設備が確実に透視投影画像に写るよう撮影地点 6m に、全方位動画の縦の長さ $H[\text{pix}]$ を用いて $M = 6000/H[\text{mm}/\text{pix}]$ とする。

作成した画像中から OpenCV を使用して天井面中の四角形の物体を認識し、そのうちのひとつを使用して天頂方向推定を行う。その四角形の各頂点 $v (= 0, 1, 2, 3)$ の周辺に微分処理が可能な prewitt フィルタ [14] を適用した画像を $I_v(x, y)$ とする。四角形の頂点を正確に捉えるため、図 4 に示すような縦横の長さが $H_e[\text{pix}]$, $W_e[\text{pix}]$ の 4800 個のエッジフィルタ $E_k (k = 1, 2, \dots, 4800)$ を各頂点周辺の全画素に式 (5) のように適用する。式 (6) のように、適用した演算結果が最も大きい画素 (\hat{x}, \hat{y}) を頂点とする。

$$f(I(x, y), E_k) = \sum_{i=-\frac{W_e}{2}}^{\frac{W_e}{2}} \sum_{j=-\frac{H_e}{2}}^{\frac{H_e}{2}} I(x+i, y+j) E_k(i, j) \quad (5)$$

$$\hat{x}, \hat{y}, \hat{k} = \underset{x, y, k}{\operatorname{argmax}} f(I(x, y), E_k) \quad (6)$$

頂点 4 点の座標の位置関係をもとに、注目している四角形物体が長方形となるような姿勢変換角度（オイラー角）を繰り返し求める。姿勢変換角度から式変換角度の差が閾値より小さくなるか一定回数繰り返したら、その時の姿勢を天頂方向とする。変換角度は、図 5 に示すような以下の項目をもとに算出する。

- 4 辺の各角度とそれぞれの理想的角度との差の平均値
- 対になる 2 辺の比（左側と右側）(L_1/L_2)
- 対になる 2 辺の比（上側と下側）(l_1/l_2)

ただし、姿勢変換を収束させるために、算出した各角度に対して繰り返し回数を $n(0, 1, \dots, 34)$ として $-0.98^{-n} + 2$ を乗算する。 $-0.98^{-n} + 2$ としたのは、線形的に収束させるよりも繰り返し回数が少ない時に変換角度を大きくできるからである。繰り返し毎の姿勢変換の推移を図 6 に示す。最初の姿勢から繰り返し処理を終えた時の姿勢への回転行列を \mathbf{R}_i とする。後の処理のために、 \mathbf{R}_i で各キーフレームの撮影位置 p_{ck} 、回転行列 \mathbf{R}_{ck} 、3 次元点群 PC を式 (7), (8), (9) のように回転させ、それぞれ p_{cik} , \mathbf{R}_{cik} , PC_i とする。

$$p_{cik} = p_{ck} \mathbf{R}_i, \quad (7)$$



図 6 姿勢変換の推移

$$R_{ci k} = \mathbf{R}_{c k} \mathbf{R}_i \quad (8)$$

$$PC_i = PC \mathbf{R}_i \quad (9)$$

3 次元点群を用いた天井方向推定

画像認識を用いた処理では、光のあたり方や歪みによって矩形の頂点を正確に捉えきれず、正確な天井方向を求められない場合がある。そのため VSLAM が出力する 3 次元点群からも天井方向推定を行う。

蛍光灯などの天井設備は、天井面と色差があることが多いため、画像上で特徴点となる。天井面に特徴点が多数あれば、3 次元点群 PC_i から天井面を抽出可能である。抽出処理の前に、ノイズ除去のために、3 次元点群 PC_i 上の全ての点に対して、その点を中心とした球の中の他の点の数を基準にノイズを判定する Radius outlier removal を行う。

3 次元点群 PC_i からの天井面の抽出には Random Sample Consensus (RANSAC) [15] を使用する。RANSAC では、最初のサンプルをランダムに選ぶため、天井面ではない面が抽出される可能性がある。そのため、3 次元点群 PC_i の天井側 1/4 を切り出した点群に対して、3 次元の平面式モデルを使用した RANSAC を適用する。推定した天井面をもとに、VSLAM 座標系を Z 軸が天井面の法線と平行になるように回転させる。この時の回転行列を \mathbf{R}_p とする。また、後の処理のために、回転行列 \mathbf{R}_p で各キーフレームの撮影位置 $p_{ci k}$ 、回転行列 $\mathbf{R}_{ci k}$ 、3 次元点群 PC_i を式 (10)、(11)、(12) のように回転させ、それぞれ $p_{cp k}$ 、 $\mathbf{R}_{cp k}$ 、 PC_p とし、 $R_z = \mathbf{R}_p \mathbf{R}_i$ と定義する。

$$p_{cp k} = \mathbf{R}_p p_{ci k}, \quad (10)$$

$$\mathbf{R}_{cp k} = \mathbf{R}_p \mathbf{R}_{ci k} \quad (11)$$

$$PC_p = \mathbf{R}_p PC_i \quad (12)$$

3.2.1.2 スケール推定

VSLAM 座標系はスケールが不明なため、OpenVSLAM 出力のスケール推定を行った西田らの手法 [13] と同様に、VSLAM 座標系とワールド座標系で対応する 2 地点間距離の比から、スケールの推定を行う。VSLAM 座標系上で、原点 (1 フレーム目の撮影地点) から推定した天井面までの距離を h_v とする。また、ワールド座標系で 1 フレーム目の撮影地点から天井面までの距離 h_0 [mm] を、透視投影画像中の天井設備の長さから、算出する。

h_0 を算出するために、全方位動画の 1 フレーム目から透視投影画像を作成する。作成時のパラメータは、 $\mathbf{R} = \mathbf{R}_z$ と

し、 d, M は 3.3.1.1 項と同様に 1700 [mm], 6000/ H [mm/pix] とする。作成した透視投影画像に写る天井設備の画像上の長さ D_p [pix] と実際の長さ D_r [mm] を測定する。式 (13) からワールド座標系上での、 h_0 [mm] を算出する。

$$h_0 = d \frac{D_p}{D_r M} \quad (13)$$

次に、式 (14) でスケール S を計算する。

$$S = \frac{h_0}{h_v} \quad (14)$$

天井画像作成のために、各キーフレームの撮影位置 $p_{cp k}$ を推定したスケールをもとに式 (15) に従って正規化し、 $p_{cs k}$ とする。

$$p_{cs k} = p_{cp k} \times S \quad (15)$$

3.3 天井画像の作成

各キーフレームに対して、推定された撮影位置 $p_{cs k} (x_{cs k}, y_{cs k}, z_{cs k})$ と回転行列 $\mathbf{R}_{cp k}$ をもとに天井向きの透視投影画像を作成する。各キーフレームの撮影地点から天井面までの距離 h_k は、式 (16) で表される。

$$h_k = h_0 - z_{cs k} \quad (16)$$

また、作成する天井画像の 1 ピクセルあたりの長さを M_c とする。よって、透視投影画像作成時のパラメータを、 $\mathbf{R} = \mathbf{R}_{cp k}$, $d = h_k$, $M = M_c$ として、各キーフレーム毎に天井画像を作成する。

3.4 天井設備外形の簡易認識

各天井画像に対して、天井設備の簡易認識を行う。本研究で行う設備認識の対象は、矩形と円形の設備である。OpenCV でグレースケール化した 8 ビット階調の天井画像に対して、30 から 230 まで 10 刻みの 21 段階の閾値で二値化した画像を得る。各二値化画像に対して OpenCV で輪郭を取得し、その最小外接矩形と最小外接円を求める。求めた矩形と円形の中から設備候補となるものを抽出し、同一画像の異なる二値化閾値で同一の設備候補とみなされたものを最終的に設備とみなす。複数の閾値で抽出を行い確認することで、光の当たり方や色味などに対する頑健性を向上させるためである。

最小外接矩形に対しては、以下の条件を 1 つでも満たすものを除外し、残ったものを設備候補とする。

- 矩形の面積が天井画像の面積の 0 から 0.005 倍または、0.6 倍から 1.0 倍までのもの (ノイズ除去)
- 矩形の設備が見切れている可能性があるもの (設備の全貌が写っていない可能性のものを除去)
- 矩形の周の長さとは二値画像の輪郭線の長さの比が、0.8 以下または、1.2 以上のもの (実際には矩形でないものを除去)

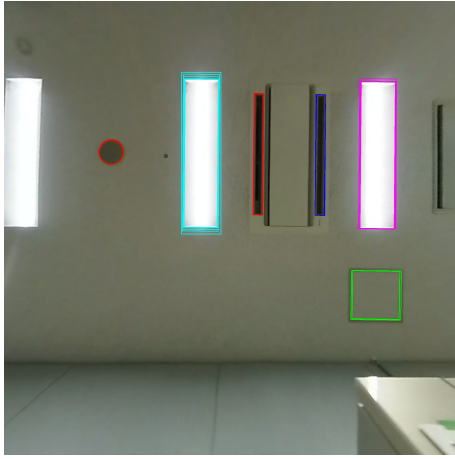


図 7 簡易天井設備認識例

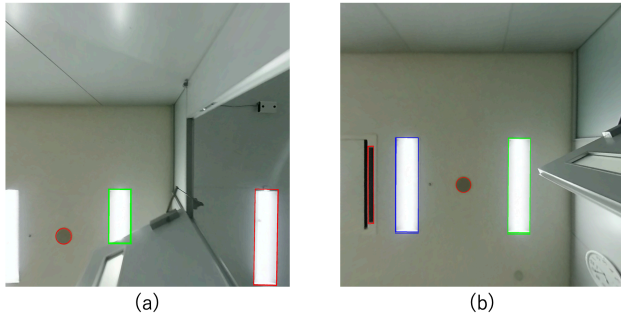


図 8 天井設備の全容が見えない画角 (a) と見える画角 (b)

- 矩形の面積と輪郭線の内側の面積の比が、0.8 以下または、1.2 以上のもの（実際には矩形でないものを除去）
 - 矩形の内側にある、黒または白の画素の数の合計値で、大きい方の値と矩形の面積の比が、0.8 以下または、1.2 以上のもの（矩形の内側が一様でないものを除去）
- 最小外接円に対しては、以下の条件を 1 つでも満たすものを除外し、残ったものを設備候補とする。

- 円形の面積が天井画像の面積の 0 から 0.001 倍または、0.6 倍から 1.0 倍までのもの（ノイズ除去）
- 円形の設備が見切れている可能性のあるもの（設備の全貌が写っていない可能性のものを除去）
- 円形の周の長さとの輪郭点を結ぶ長さの比が、0.8 以下または、1.2 以上のもの（実際には円形でないものを除去）
- 円形の面積と輪郭点の内側の面積の比が、0.8 以下または、1.2 以上のもの（実際には矩形でないものを除去）
- 円形の内側に黒または白の画素の数の合計値で大きい方の値と円形の面積の比が、0.8 以下または、1.2 以上のもの（円形の内側が一様でないものを除去）

同一画像の異なる二値化閾値で、設備候補の重心距離が 50[mm] 以内ならば、同一の設備候補とみなす。図 7 に認識した設備の画像例を示す。

3.5 最も綺麗に写る設備外形の探索

認識した天井設備から、最も綺麗に写る設備の探索を行う。まず、異なる天井画像間で同じ天井設備であるものを探索する。各キーフレームの撮影位置 $p_{csk}(x_{csk}, y_{csk}, z_{csk})$ から、各天井画像の中心座標は (x_{csk}, y_{csk}) で表される。これら天井画像同士の位置関係から、認識した各天井設備位置を比較し、以下の条件に当てはまる各天井設備を同じ天井設備であるとみなす。

- それぞれの天井設備の重心の距離が 300[mm] 以内
- それぞれの天井設備の面積比が 0.8 から 1.2

また、ある天井設備に別の天井設備の重心が内包されている場合、2 つの天井設備の面積を比較して小さい面積の方を天井設備から除外する。図 8 のように画角などの理由で天井設備の一部のみが認識された場合に、天井設備の全容が写っているもののみを天井設備とみなすためである。

次に、それぞれ同じ天井設備とみなされた物の中から最も綺麗に写る設備外形を探索する。天井画像は中心に近いほど、歪みの影響が小さくなるため、天井画像の中心に写っている天井設備ほど綺麗であると判断できる。同じ天井設備とみなされたものの中で、天井画像の中心から天井設備の重心までの距離が最も短い天井画像に天井設備があると記録する。つまり、天井設備がより中心に写っている天井画像が記録される。

3.6 天井パノラマ画像の合成

各天井画像の中心座標 (x_{csk}, y_{csk}) の位置関係と認識した天井設備の位置から、天井パノラマ画像上の各地点が最も綺麗に写り、天井設備外形が破綻なく写された天井パノラマ画像を合成する。まず、合成する天井パノラマ画像の縦横の長さ [pix] を決める（この際、ワールド座標系からパノラマ画像の座標系の変換が定義される）。次に、各天井画像の中心座標 (x_{csk}, y_{csk}) を、パノラマ座標系に変換して (x_{cpk}, y_{cpk}) とする。その後、パノラマ画像の各ピクセルに対して、各天井画像の中心 (x_{cpk}, y_{cpk}) までの距離 [pix] を求め、最も短い距離となる天井画像から天井パノラマ画像上に画素値をコピーする。ただし、天井設備外形を破綻なく写すために、そのピクセルが天井設備外形及びその周辺領域に該当するなら、各天井画像の中心までの距離を無視して、その設備が写る天井画像から画素値をコピーする。

4. 実験と評価

2 つの異なる室内環境において、提案手法とベースラインでそれぞれデータ収集を行い、天井パノラマ画像を合成した。合成した天井パノラマ画像に対して、設備位置の測位誤差による精度と解像度の評価を行った。



図 9 (a) ワークスペースの室内環境 1, (b) 計算機室の室内環境 2

4.1 ベースライン

ベースラインとして、建築物のデジタル化やインデックス化を行う Matterport 社の商用ソフトウェア Matterport を使用する。Matterport は、3D レーザスキャナや全方位カメラ、スマートフォンなどのデバイスで室内空間のキャプチャが可能なツールであるが、正確にキャプチャするには 3 次元復元の基礎的な知識が必要である。本実験では、提案手法と同じ全方位カメラを用いて測定を行う。ただし、Matterport の全方位カメラ使用時の入力動画ではなく画像である。スマートフォン用アプリと全方位カメラを Wi-Fi で接続して、室内空間を撮影地点に連続性があるように順に撮影していく。撮影した全方位画像をスマートフォン用アプリを介してアップロードすると、クラウド上での処理が終わった後に、測定した室内のウォークスルーや 3 次元モデル、床面と天井面のパノラマ画像を見られる。

4.2 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、個室が複数ある約 273 m² の室内空間で、図 9(a) に示すようにデスクやモニター、椅子などが置かれたワークスペース環境である。2 つ目は、個室のない開けた約 86 m² の室内空間で、図 9(b) に示すようにパソコンがデスクの上に整列している計算機室である。

提案手法、ベースラインともに THETA Z1 を使用してデータを収集した。表 1 に収集したデータの詳細を示す。提案手法用の全方位動画撮影時は、図 1 中に示したように全方位カメラの片方のレンズを天井側に向け、室内をくまなく歩きながら撮影を行う。天井面を片方のレンズで撮影し、レンズ間での明るさや色味の違いの影響を受けないようにするためである。ベースライン用の全方位画像撮影時は、自立する長さ約 126 cm の一脚を使用して全方位カメラのレンズが水平方向を向くように撮影した。また、正解データとして FARO 製の 3D レーザスキャナ FocusM70 を用いて、室内の 3 次元点群を取得した。取得した 3 次元点群の天井部分を切り出し、天井パノラマ画像とした。

表 1 収集したデータの詳細

	測定環境	室内環境 1	室内環境 2
提案手法	動画の長さ [秒]	211	200
	フレームレート [fps]	29.97	29.97
	解像度 [pix]	3840×1920	3840×1920
	収集にかかった時間 [分]	3.5	3.3
ベースライン	全方位画像の枚数 [枚]	100	45
	解像度 [pix]	6720×3360	6720×3360
	収集にかかった時間 [分]	57	27

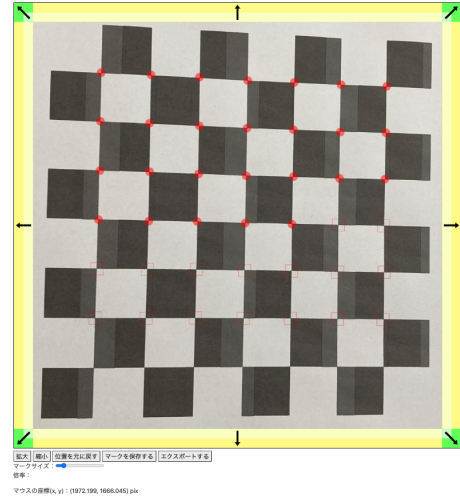


図 10 作成したマーキングツール

4.3 精度評価方法

提案手法、ベースラインそれぞれ天井パノラマ画像上の全ての蛍光灯の 4 つ角の位置誤差をもとに、天井パノラマ画像の精度評価を行う。天井パノラマ画像上の全ての蛍光灯の 4 つ角の位置の特定は目視で行い、図 10 のように自作した Web ツール上で記録した。

提案手法、ベースラインそれぞれ記録した点に対して回転や拡大縮小、平行移動の処理を行い、正解データとの位置誤差の平均が最小となる値で評価を行う。

4.4 設備外形の形状評価方法

合成した天井パノラマ画像上の複数の設備が、どの程度形状として破綻がないかを評価する。前節の精度評価の際に記録した蛍光灯の数を n 個として、 $4n$ 個ある蛍光灯の角の角度に対して、平均値と標準偏差を算出し、その値で形状を評価する。

4.5 結果

図 11 に提案手法、ベースラインで合成した天井パノラマ画像を示す。また、精度評価と形状評価の結果を表 2 に示す。図 11 より、ベースラインより提案手法で合成した天井パノラマ画像の方が、室内空間を正確に表し、設備外形を破綻なく写している。表 2 より、ベースラインより提案手法の方が精度が高かった。また、提案手法は形状評価の標準偏差の値が小さく、正確に蛍光灯の形状を天井パノラマ画像上に写している。

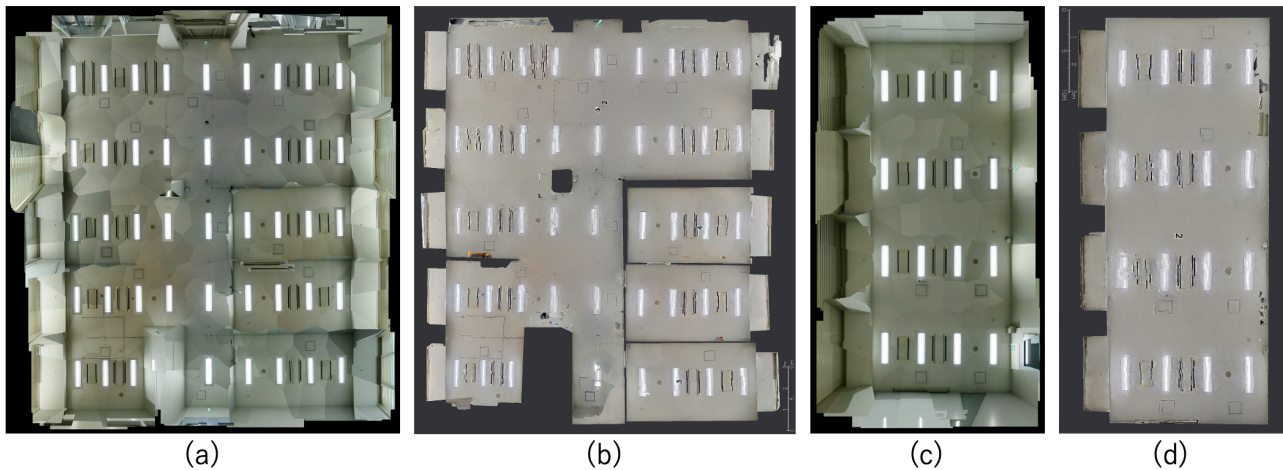


図 11 室内環境 1 の (a) 提案手法, (b) ベースライン, 室内環境 2 の (c) 提案手法, (d) ベースライン

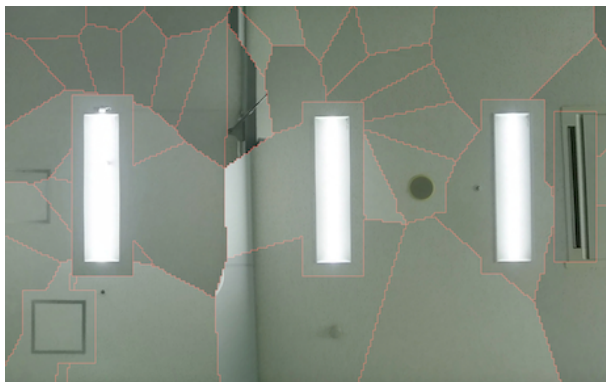


図 12 天井画像同士の切れ目を赤線で示した天井パノラマ画像

表 2 評価結果

	測定環境	室内環境 1	室内環境 2
提案手法	位置誤差平均 [mm]	53.09	14.47
	角度の平均値 [°]	90.00	90.00
	角度の標準偏差 [°]	0.95	0.68
ベースライン	位置誤差平均 [mm]	126.24	57.55
	角度の平均値 [°]	90.00	90.00
	角度の標準偏差 [°]	4.14	3.17

図 12 に天井画像同士の切れ目を赤線で示した天井パノラマ画像を示す。提案手法の目的である設備外形が切れ目なく天井パノラマ画像上に写っていることが分かる。

5. 考察

ベースラインよりも提案手法の方が優位な結果が出た。本実験では全方位カメラを用いたが、提案手法では入力に動画を使用し、ベースラインでは高解像度の静止画を使用した。高解像度の静止画を使用して天井パノラマ画像を合成するよりも、より連続性のある動画を使用して合成したために、精度と解像度が向上したと考えられる。

天井パノラマ画像合成時に設備外形を考慮したことで、提案手法の精度向上に寄与したと考えられる。歪みの少ない天井画像から設備外形を一様に写して、蛍光灯の形状を

正確に捉えられた。図 12 に示したとおり設備外形が破綻なく写されているため、定性的に視認性が良いと言える。

また、表 1 に示したとおり、提案手法はベースラインよりもデータ収集にかかる時間が短く済むため、より手軽に測定を行えた。測定の手軽さという観点からも、提案手法に優位性があると言える。

6. まとめと展望

全方位カメラで撮影した全方位動画を用いて、設備外形を考慮した天井パノラマ画像合成手法を提案した。提案手法では、各フレームの撮影位置・撮影姿勢を VSLAM を使用して推定後、その推定結果を補正した。補正された撮影位置・撮影姿勢から天頂向きの天井画像を合成し、認識した設備外形が破綻なく写るように考慮して、天井パノラマ画像を合成した。

異なる 2 つの室内環境で天井パノラマ画像を合成し、提案手法で合成した天井パノラマ画像は、ベースラインの天井パノラマ画像より、精度及び視認性、測定の手軽さの点で優位性があると分かった。

設備外形を考慮して天井パノラマ画像を合成可能であると分かったため、今後は天井設備の認識技術を確立する必要がある。天井設備は種類が多いため、その大きさや形は様々である。そのため、未知の天井設備でも認識可能な認識技術の確立が今後の課題である。また、天伏図作成補助のために、天井設備の分類手法の確立も課題である。本研究では天井面に注目したが、天井以外の床面や壁面への本技術の応用も今後行う予定である。

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] 公益社団法人日本ファシリティマネジメント協会. ファシリティマネジメント (FM) とは. <http://www.jfma.or>.

- jp/whatsFM/index.html. (Accessed on 05/05/2023).
- [2] Lidia Pinto, Ricardo Codinhoto, and Serena Bonelli. A Review of Building Information Modeling (BIM) for Facility Management (FM): Implementation in Public Organisations. *Appl. Sci.*, Volume 12, Issue 3, 2022.
 - [3] You Cheng Lin, Ya Ting Hsu, and Sin Tau Hu. BIM Model Management for BIM-Based Facility Management in Buildings. *Adv. Civ. Eng.* 2022, 2022.
 - [4] Jonatan Villavicencio Moreno, Rita Machete, Ana Paula Falcão, Alexandre B. Gonçalves, and Rita Bento. Dynamic Data Feeding into BIM for Facility Management: A Prototype Application to a University Building. *Buildings*, Volume 12, Issue 5, 2022.
 - [5] Fabio Morbidi, Louise Devigne, Catalin Stefan Teodorescu, Bastien Fraudet, Émilie Leblong, Tom Carlson, Marie Babel, Guillaume Caron, Sarah Delmas, François Pasteau, Guillaume Vailland, Valérie Gouranton, Sylvain Guégan, Ronan Le Breton, and Nicolas Ragot. Assistive Robotic Technologies for Next-Generation Smart Wheelchairs: Codesign and Modularity to Improve Users' Quality of Life. *IEEE Robotics & Automation Magazine*, 2023.
 - [6] Seita Kayukawa, Keita Higuchi, Shigeo Morishima, and Ken Sakurada. 3DMovieMap: An Interactive Route Viewer for Multi-Level Buildings. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 2023.
 - [7] Cheng Sun, Chi-Wei Hsiao, Min Sun, and Hwann-Tzong Chen. HorizonNet: Learning Room Layout with 1D Representation and Pano Stretch Data Augmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1047–1056, 2019.
 - [8] Zhigang Jiang, Zhongzheng Xiang, Jinhua Xu, and Ming Zhao. LGT-Net: Indoor Panoramic Room Layout Estimation with Geometry-Aware Transformer Network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1654–1663, 2022.
 - [9] Pooya Fayyazsanavi, Zhiqiang Wan, Will Hutchcroft, Ivaylo Boyadzhiev, Yuguang Li, Jana Kosecka, and Sing Bing Kang. U2RLE: Uncertainty-Guided 2-Stage Room Layout Estimation. *arXiv:2304.08580*, 2023.
 - [10] Mizuki Tabata, Kana Kurata, and Junichiro Tamamatsu. Shape-Net: Room Layout Estimation from Panoramic Images Robust to Occlusion using Knowledge Distillation with 3D Shapes as Additional Inputs. *arXiv:2304.12624*, 2023.
 - [11] Haiyan Wang, Will Hutchcroft, Yuguang Li, Zhiqiang Wan, Ivaylo Boyadzhiev, Yingli Tian, and Sing Bing Kang. PSMNet: Position-Aware Stereo Merging Network for Room Layout Estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8616–8625, 2022.
 - [12] Shinya Sumikura, Mikiya Shibuya, and Ken Sakurada. OpenVSLAM: A Versatile Visual SLAM Framework. In *MM '19: Proceedings of the 27th ACM International Conference on Multimedia*, pp. 2292–2295, 2019.
 - [13] 西田龍平, 内海ゆづ子, 三輪由佳, 岩村雅一, 岩瀬浩一. ブドウ摘房支援のための画像における房計数領域の推定. 研究報告自然言語処理 (NL), pp. 1–7. 一般社団法人 情報処理学会, 2021.
 - [14] J. M. S. Prewitt. Object Enhancement and Extraction. In *Picture Processing and Psychopictorics*, pp. 75–149. New York: Academic Press, 1970.
 - [15] Martin A. Fischer and Robert C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 1981.