

環境の3Dモデルと全天球カメラ画像を用いた 色差最小化によるカメラの位置姿勢推定

陽 東旭 †, 樋口 寛 †, Sarthak PATHAK†, Alessandro MORO†, 山下 淳 †, 浅間 一 †

Dongxu YANG † Hiroshi HIGUCHI † Sarthak PATHAK † Alessandro MORO †
Atsushi YAMASHITA † and Hajime ASAMA †

†: 東京大学, yangdongxu@robot.t.u-tokyo.ac.jp

<要約> 環境の3Dモデルと全天球カメラで取得した画像を利用し、オクルージョンが多い環境でも使用可能なカメラの位置姿勢推定の手法を提案する。特定の特徴を抽出するのではなく、画像の全ピクセルの色情報を使って位置姿勢推定を行う。オクルージョンによる影響を軽減するために、ロバストな評価関数を推定に使用する。提案手法の有効性を実験により確認した。

<キーワード> 位置姿勢推定, 全天球カメラ, 環境の3Dモデル

1 序論

近年、インフラや工場の点検など、さまざまな目的でより一般的にドローンが使用され始めている。点検場所に到着するために、ドローンの位置姿勢を正確に知る必要がある。GPSを使用した位置推定方法は屋外で効率的に機能するが、屋内環境では信号が届かないという問題点がある。そこで、カメラを使用した屋内でのロボットの位置姿勢推定に関する研究が多く行われている。

ロボットの位置姿勢推定を行うには、環境の3D情報を使用するのが効果的である。先行研究では、深度カメラを使用して3D情報を取得した研究がある。たとえば、[1]は、RGB-Dカメラを使用し、深度画像とカラー画像の両方を利用し、リアルタイムの6自由度の位置姿勢推定を行った。[2]は、深度カメラを使用し、RANSACおよび三次元点のマッチングで位置姿勢推定をした。しかし、RGB-Dカメラの視野は比較的に狭いため、取得できる情報は少ない。ロボットの位置姿勢推定を行う際に、フレーム間の共通の視野がなければ推定ができないため、フレーム間の共通の視野は重要である。たとえば、ロボットの回転が非常に大きい場合、視野が狭いカメラで撮影された画像はまったく異なる可能性があり、位置姿勢推定ができなくなる。より広い視野を持つカメラを使用することにより、より

多い情報を取得でき、フレーム間の共通の視野は大きくなる。したがって、視野の広いカメラは、位置姿勢推定に効果的である。

本研究では、3D情報を取得する手段としては、事前に取得した環境の3Dモデルを使う。工場のような環境では、環境の3Dモデルは建設のCADデータから容易に入手できる。CADデータがない場合も、3Dスキャンをすることで比較的簡単に入手できる。従って、本研究では、事前に取得した環境の3Dモデルと360度の視野を持つ全天球カメラを用いた位置姿勢推定法を提案する。屋内では、障害物によるオクルージョンの問題が発生する。それを解決するため、特定の特徴を使用する代わりに、全天球カメラ画像のすべてのピクセルの色差を使用することにより、オクルージョン問題に対してよりロバストな位置姿勢推定手法を構築することを目的とする。

2 提案手法

2.1 問題設定

本研究では、カメラの初期位置姿勢は既知であると想定する。既知の位置姿勢で時刻 t で撮った画像をキー画像と呼ぶことにする。そのカメラの位置姿勢を使用して3Dモデルから生成したデプスマップ(t)を推定に使う。時刻 $t+1$ で撮った画像を推定対象画像とする。入力はキー画像、デプスマップ、推定対象画像であり、

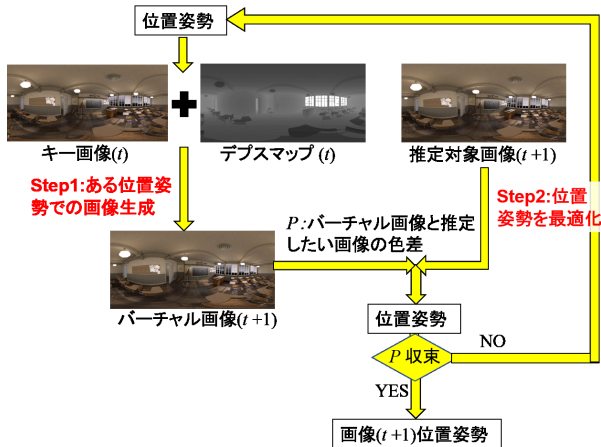


図 1 提案手法の全体像

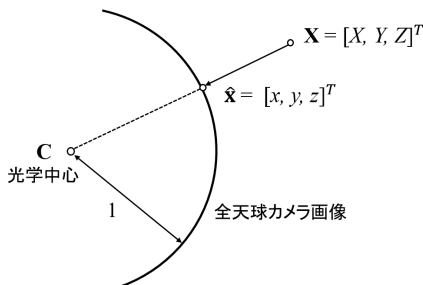


図 2 全天球投影モデル

出力は推定対象画像の位置姿勢である．本研究では，カメラの連続位置姿勢推定を想定するため，カメラの移動量は小さいとする．

2.2 アプローチ

提案手法の全体像を図 1 に示す．2 つの画像の色差が最小になるとき，同一視点から撮られた可能性が最も高いことを利用して位置姿勢推定を行う．時刻 t で撮ったキー画像とデプスマップから，時刻 $t+1$ の推定対象画像と色差が小さくなるように回転行列 R および並進ベクトル t で視点の変換をし，バーチャル画像を生成する．バーチャル画像と推定対象画像の色差が最小となる \hat{R} と \hat{t} を，推定対象画像の推定位置姿勢とする． \hat{R} と \hat{t} は，次式によって求める．

$$\hat{R}, \hat{t} = \operatorname{argmin}(\mathbf{I}_V(\mathbf{R}, \mathbf{t}) - \mathbf{I}_T), \quad (1)$$

ただし， $\mathbf{I}_V(\mathbf{R}, \mathbf{t})$ は \mathbf{R}, \mathbf{t} で変換した視点のバーチャル画像行列を 1 次元に展開したベクトルであり， \mathbf{I}_T は推定対象画像行列を 1 次元に展開したベクトルである．

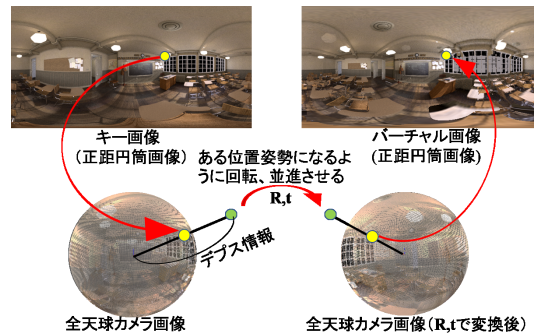


図 3 バーチャル画像生成

2.3 バーチャル画像生成

まず，全天球カメラの投影について説明する．全天球投影モデルを図 2 に示す．全天球カメラの投影面は，光学中心 c を原点を中心とする単位球面である．空間上の 3 次元点 $\mathbf{X} = [X, Y, Z]^T$ の色情報は単位ベクトル $\hat{\mathbf{x}} = [x, y, z]^T$ として投影面に投影される．全天球カメラ画像は三次元のベクトルで表したもので，画像処理で直接扱うのは難しい．そのため，本研究では，2 次元の正距円筒画像に変換して扱う．

次に，バーチャル画像生成について説明する．図 3 に示すように，キー画像（正距円筒画像）をまず全天球カメラ画像に変換する．各ピクセルに対し，カメラの光学中心からの距離をデプスマップから取得したデプスに変え， R と t で変換した後，ピクセルと光学中心の距離を正規化し，また全天球カメラ画像に変換する．その後，全天球カメラ画像を正距円筒画像に戻す．その結果，回転行列 R および並進ベクトル t に対応する視点からのバーチャル画像（正距円筒画像）が生成される [3]．

次に，図 3 に示した全天球カメラ画像の変換のプロセスを図 4 に示す．まず，各ピクセルのカメラの光学中心からの距離をデプスマップから取得したデプス情報に変換すると，3D の点群が得られる．カメラの位置姿勢に R と t の変換を加え，視点 1 から視点 2 に変える．このとき，各ピクセルと視点 2 の位置関係が計算できる．次に，各ピクセルから視点 2 までの距離を正規化することにより，各ピクセルを球面上に投影すると， R と t で変換後の全天球カメラ画像が得られる．

2.4 オクルージョン問題

バーチャル画像を生成する際に，図 5 に示したようなオクルージョンによる問題が発生する．環境にオク

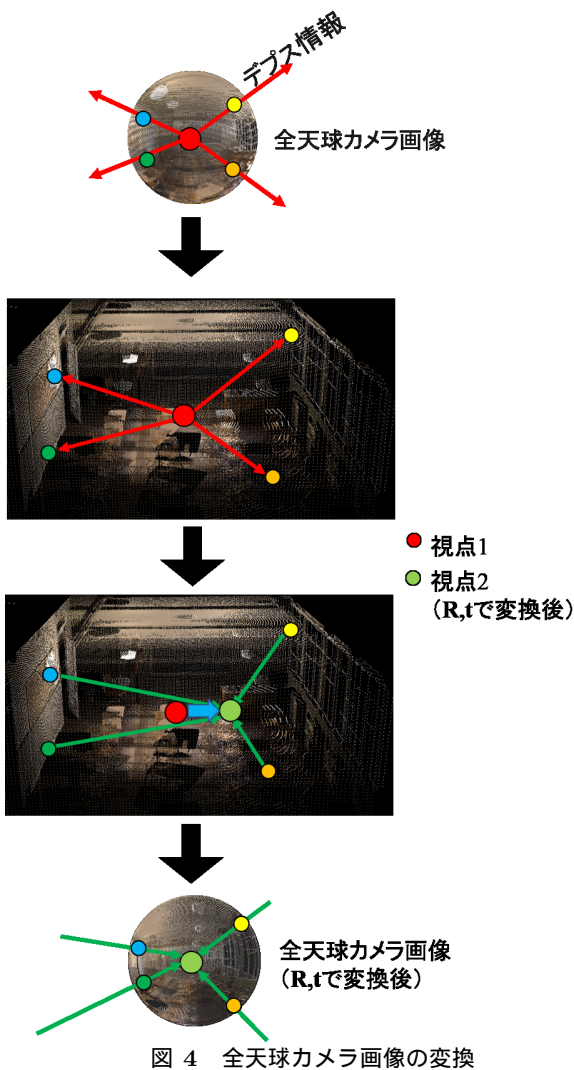


図 4 全天球カメラ画像の変換

ルージョンがあるとき、カメラが移動すると、最初の視点で見えなかったところが次の視点で見えるようになる。バーチャル画像を生成するとき、そのような隠されて見えなかった領域の色情報を正しく推定することができない。

オクルージョン問題の例のを図 6 に示す。キー画像では、机の後ろの部分が見えなかったため、バーチャル画像を作るとき、その部分の色を正しく推定できない。

2.5 最適化を使った推定

以下のように定義された評価関数 P を使用し、レーベンバーグ・マーカート法でバーチャル画像と推定対象画像を比較する。

$$P = \frac{1}{2} \|\mathbf{a}\|^2, \quad (2)$$

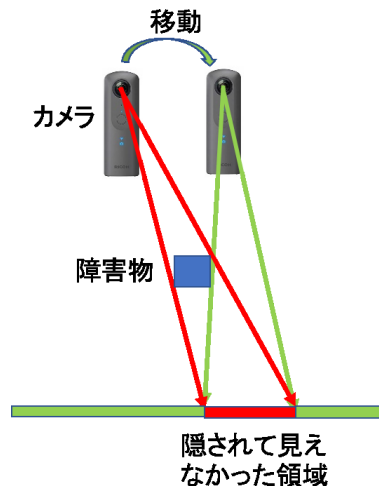


図 5 オクルージョン問題

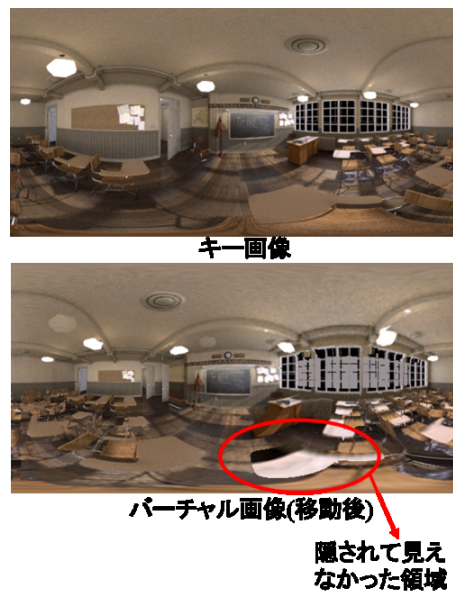


図 6 オクルージョン問題の例

ただし、

$$\mathbf{a} = \mathbf{I}_V(\mathbf{R}, \mathbf{t}) - \mathbf{I}_T, \quad (3)$$

であり、バーチャル画像と推定対象画像の色差である。 P が最小となる出力位置姿勢を推定対象画像の推定位置姿勢とする。

2.6 評価関数の重み付け

2.4 節で説明したようなオクルージョンによる影響を減らすために、評価関数の重み付けをする。バーチャル画像を生成するとき、オクルージョンがある領域の色が正しく推定できないため、その領域の色差は、他の領域より大きい傾向がある。そこで、式 (1) を次のように定義される Huber 損失関数 [4] に変更すること

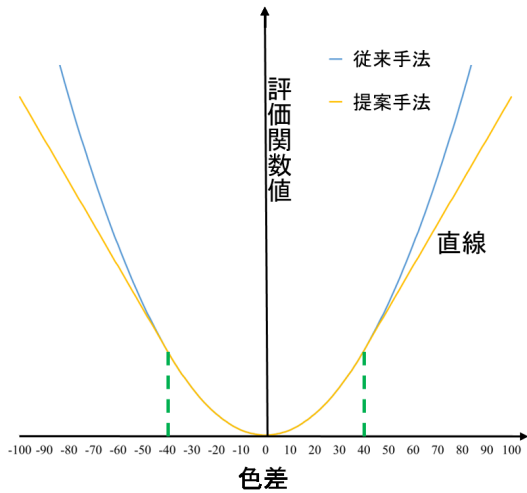


図 7 評価関数

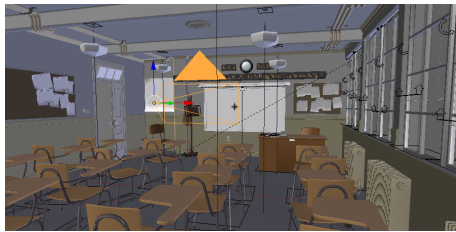


図 8 Blender の教室モデル

により、色差が比較的に大きい領域の重みを下げる．

$$P_i = \begin{cases} \frac{1}{2} \|a_i\|^2 & (\|a_i\| \leq \delta) \\ \delta (\|a_i\| - \frac{1}{2} \delta) & (\text{otherwise}) \end{cases}, \quad (4)$$

$$P = \sum_{i=1}^n P_i, \quad (5)$$

ただし、 a_i は a の i 番目の成分であり、 n は画像のピクセル数かける 3 (RGB チャンネル) である．本研究では、 δ を 40 に設定した．式 (2) で表した二乗エラー損失関数と、式 (4)、(5) で表した Huber 損失関数をとともに図 7 に示す．

3 実験と結果

本研究では、画像とデプスマップを生成可能なソフトである Blender を用いた．Blender を用いた実験では、カメラ位置の真値は正確にわかっているため、結果を正確に評価することができる．この実験では、教室の 3D モデル (図 8) を使用し、カメラの x, y, z 軸の並進と回転の 6 つのパラメータを同時に推定した．さまざまな位置姿勢から撮った 50 枚の画像を使用した．カ



オクルージョンなし



オクルージョンあり

図 9 実験環境

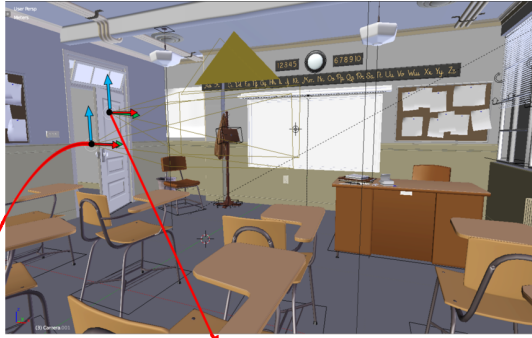
メラの移動範囲は、並進 $\pm 0.25\text{m}$ 、回転 $\pm 2.5\text{deg}$ に設定した．教室のモデルに椅子などを追加し、オクルージョンの多い環境 (図 9) を作った．評価関数として式 (2) を使った従来手法と式 (4)、(5) を使った提案手法についてそれぞれ調べ、4 種類の実験を行った．

推定結果の例を図 10 に示す．キー画像、推定対象画像、パッチャル画像の位置姿勢の真値を表 1 に示す．

実験では、最適化がローカルミニマムで終了し、位置姿勢が正しく推定されない場合があった．そのような場合は、失敗した場合と見なされ、四分位範囲 (IQR) [5] を使用して除外した．各実験においての成功回数の比率を表 2 に示す．失敗した場合を含まないデータの推定誤差の平均を図 11、12 に示す．

表 2 から、環境にオクルージョンがない場合もある場合も、従来手法より提案手法を用いたほうが成功回数の比率が高いことがわかった．

図 11、12 から、A と C の結果を比べると、オクルージョンがあると、並進と回転の両方の推定誤差が増加することがわかる．C と D を比べると、オクルージョンがある場合に提案手法を使用することにより、並進誤差と回転誤差の両方が小さくなることがわかる．提案手法を用いると、オクルージョンがあるときでも、推定誤差はオクルージョンなしのときとほぼ同じレベルに下ることがわかった．



キー画像



推定対象画像



バーチャル画像

図 10 推定結果の例

4 結論

環境の 3D モデルを使用して，全天球カメラの位置姿勢を推定する手法を提案した．オクルージョンを考慮し，特定な特徴ではなく，画像の全ピクセルの色情報を利用した推定方法を提案した．さらに，最適化の評価関数に工夫を加え，オクルージョンによる影響を下げる事ができた．実験により，環境にオクルージョンがあるときの提案手法の有効性を確かめた．提案手法を用いると，実験の成功回数の比率が上がる事が

表 1 推定結果の例の位置姿勢

画像	位置 [mm]			姿勢 [deg]		
	x	y	z	x	y	z
キー画像	0	0	0	0	0	0
推定対象画像	40	210	250	2.3	1.3	-2.3
バーチャル画像	45	200	249	2.5	1.5	-2.3

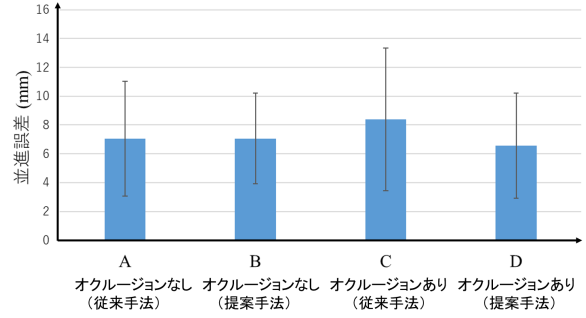


図 11 並進誤差

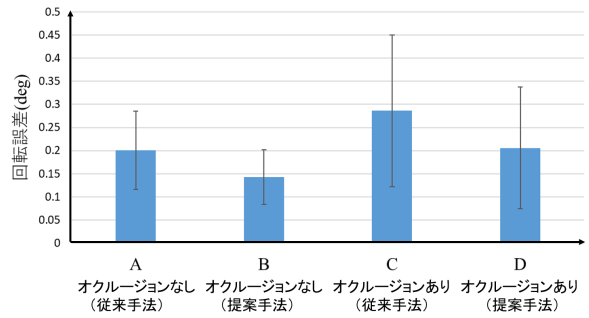


図 12 回転誤差

わかった．また，環境にオクルージョンがあると，並進誤差と回転誤差が増加するが，提案手法を用いると両方とも減少することがわかった．

謝辞

本研究の成果の一部は株式会社ポケット・クエリーズとの共同研究によるものである．

参考文献

- [1] Z. Fang and S. Scherer: "Real-time Onboard 6DoF Localization of an Indoor MAV in Degraded Visual Environments Using a RGB-D Camera," Proceedings of the 2015 IEEE International Conference on Robotics and Automation, pp. 5253-5259, 2015.
- [2] H. Du, P. Henry, X. Ren, M. Cheng, D. B. Goldman, S. M. Seitz and D. Fox: "Interactive 3D Modeling of Indoor Environments with a Consumer Depth Camera," Proceed-

表 2 成功回数の比率

実験	成功回数の比率 (%)
A. オクルージョンなし (従来手法)	88
B. オクルージョンなし (提案手法)	100
C. オクルージョンあり (従来手法)	82
D. オクルージョンあり (提案手法)	90

ings of the 13th International Conference on Ubiquitous Computing, pp. 75-84, 2011.

[3] S. Pathak, A. Moro, H. Fujii, A. Yamashita and H. Asama: "Distortion-Robust Spherical Camera Motion Estimation via Dense Optical Flow," Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP2018), pp. 2381-8549 (2018)

[4] P. J. Huber: "Robust Estimation of a Location Parameter,"

Annals of Mathematical Statistics, Vol. 35, No. 1, pp. 73-101 (1964)

[5] G. Barbato, E. M. Barini, G. Genta and R. Levi: "Features and Performance of Some Outlier Detection Methods," Journal of Applied Statistics, Vol. 38, No. 10, pp. 2133-2149 (2011)