

# 屈折を用いたスケール復元可能な Structure from Motionのための 初期値フィルタリングによる全体最適化

○奥村有加里（東京大学） 藤井浩光（東京大学） 山下淳（東京大学） 浅間一（東京大学）

## 1. 序論

Structure from Motion は1台のカメラを用いて移動しながら取得した画像から、3次元計測を行う手法である [1]. この手法は、計測対象の3次元座標と同時に、カメラの移動情報、すなわちカメラの回転と並進移動の方向を推定する. 一方で、カメラの並進移動の大きさ（スケール）は算出することができず、計測対象の大きさを求めることができない問題がある.

計測対象のスケールまで含めた復元をするために、カメラの位置関係に関する定量的な値や、幾何学的な情報を与える手法が提案されている [1-4]. しかし、環境が完全に未知でスケールに関する情報が利用できない場合には、これらの手法の適用は不可能である. この問題を解決する方法として、柴田らは屈折を利用することで、スケールを含めて3次元復元が可能な Structure from Motion を提案した [5,6]. この手法は、カメラと計測対象の間に透明平板を配置し、透明平板を通過することで起こる屈折現象を利用している. 透明平板を通して得た2枚の画像から、スケールまで含めた復元が可能であることが示されている. しかし、この手法は計測において画像を取得する際の量子化誤差の影響を大きく受けることがわかっており、ピクセル精度の計測では十分な復元精度が得られないとの報告がなされている.

Structure from Motion において観測の誤差の影響を小さくし計測の精度を向上させるための手法として、バンドル調整による全体最適化が広く用いられている [7]. バンドル調整を行う上では、最適化の初期値およびその評価関数が極めて重要となる.

本研究では、屈折を用いた Structure from Motion における3次元復元の精度を向上させるために、計測系の幾何学的整合性を考慮した初期値フィルタリング手法を提案する. さらに、バンドル調整を用いた全体最適化のために屈折の影響を考慮した評価関数を提案する.

## 2. 計測の原理

屈折を用いた Structure from Motion における計測の原理について述べる [5,6].

### 2.1 カメラと透明平板の位置関係

本手法におけるカメラと透明平板の位置関係を図1に示す. カメラ中心を原点とし、カメラの光軸方向を  $z$  軸とした右手座標系をカメラ座標系とする. 透明平板は、カメラの光軸と垂直になるように配置する. このとき、光線は屈折により図1の実線のような光路を

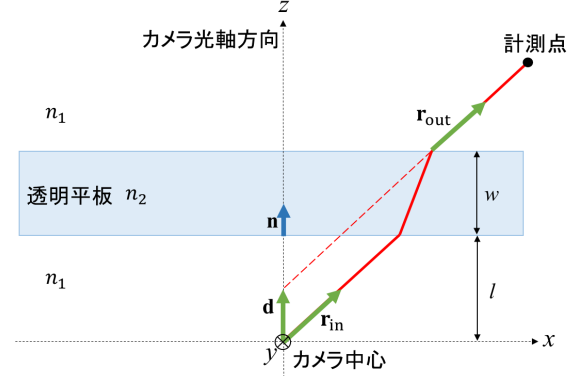


図1 カメラと透明平板の位置関係

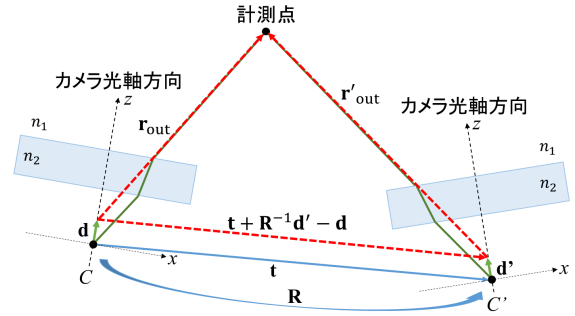


図2 2視点間での幾何学的関係

たどる.

光線に関して、2種類の光線ベクトルを考える. カメラから平板までの内側光線ベクトルを  $\mathbf{r}_{in}$  とし、平板から計測点までの外側光線ベクトルを  $\mathbf{r}_{out}$  とする. 透明平板の法線ベクトルを  $\mathbf{n}$  とすると、カメラの光軸と平行であることから、 $\mathbf{n} = (0, 0, 1)^T$  となる. 透明平板の法線方向への屈折による光線の変化量ベクトル  $\mathbf{d}$  を考え、 $\mathbf{d} = (0, 0, d)^T$  とする.  $d$  は幾何学的に算出することができ、スネルの法則を利用すると、以下のよう

$$d = w \left( 1 - \frac{\mathbf{r}_{in} \cdot \mathbf{n}}{\sqrt{\left(\frac{n_2}{n_1}\right)^2 - \|\mathbf{r}_{in} \times \mathbf{n}\|^2}} \right), \quad (1)$$

ここで、 $n_1, n_2$  はそれぞれ空気と透明平板の屈折率、 $w$  は透明平板の厚さである.

### 2.2 屈折を用いた Structure from Motion

2枚の画像を用いる屈折を利用した Structure from Motion の模式図を図2に示す. 2つのカメラ座標系を  $C, C'$  とし、カメラ座標系  $C$  は世界座標系と一致させる. このとき、回転行列  $\mathbf{R}$  を座標系  $C$  から座標系  $C'$  に姿勢を変換する行列とし、並進ベクトル  $\mathbf{t}$  を座標系

$C$  で表した座標系  $C'$  の原点の位置ベクトルとする。また、座標系  $C, C'$  における外側光線ベクトルと変化量ベクトルをそれぞれ  $\mathbf{r}_{\text{out}}, \mathbf{d}, \mathbf{r}'_{\text{out}}, \mathbf{d}'$  とおく。このとき、2つの外側光線ベクトルとそれぞれの始点を結んだベクトルは同一平面上に存在している。この幾何学的関係は

$$\{(\mathbf{t} + \mathbf{R}^{-1}\mathbf{d}' - \mathbf{d}) \times \mathbf{R}^{-1}\mathbf{r}'_{\text{out}}\}^T \mathbf{r}_{\text{out}} = 0, \quad (2)$$

と表すことができる。ここで各ベクトルの成分を  $\mathbf{r}_{\text{out}} = (x, y, z)^T$ ,  $\mathbf{d}, \mathbf{r}'_{\text{out}} = (x', y', z')^T$ ,  $\mathbf{d}' = (0, 0, d')^T$  とおき、式 (2) を既知数と未知数の積の形となるように整理すると、

$$\begin{pmatrix} xx' \\ yy' \\ zz' \\ xy' \\ xy' \\ yy' \\ yy' \\ xz' \\ yz' \\ zz' \\ dyx' + d'xy' \\ -dxx' + d'yy' \\ d'zy' \\ dyy' - d'xx' \\ -dxy' - d'yx' \\ -d'zx' \\ dyz' \\ -dxx' \end{pmatrix}^T \begin{pmatrix} r_{12}t_3 - r_{13}t_2 \\ r_{13}t_1 - r_{11}t_3 \\ r_{11}t_2 - r_{12}t_1 \\ r_{22}t_3 - r_{23}t_2 \\ r_{23}t_1 - r_{21}t_3 \\ r_{21}t_2 - r_{22}t_1 \\ r_{32}t_3 - r_{33}t_2 \\ r_{33}t_1 - r_{31}t_3 \\ r_{31}t_2 - r_{32}t_1 \\ r_{11} \\ r_{12} \\ r_{13} \\ r_{21} \\ r_{22} \\ r_{23} \\ r_{31} \\ r_{32} \end{pmatrix} = 0, \quad (3)$$

となる。ここで、 $r_{ij}$  は回転行列  $\mathbf{R}$  の  $i$  行  $j$  列の成分の値であり、 $t_i$  は並進ベクトル  $\mathbf{t}$  の  $i$  番目の成分の値である。式 (3) は各対応点について成り立つため、連立方程式を立て最小二乗法を用いることで、未知数からなるベクトルを求めることができる。その際、未知数ベクトルの第10–第17成分が回転行列の成分であることから、回転行列の正規直交性より、

$$r_{11}^2 + r_{12}^2 + r_{13}^2 = 1, \quad (4)$$

$$r_{21}^2 + r_{22}^2 + r_{23}^2 = 1, \quad (5)$$

$$r_{11}r_{21} + r_{12}r_{22} + r_{13}r_{23} = 0, \quad (6)$$

$$r_{12}r_{23} - r_{13}r_{22} = r_{31}, \quad (7)$$

$$r_{13}r_{21} - r_{11}r_{23} = r_{32}, \quad (8)$$

が成り立つ。

式 (4)~(8) を制約条件とした、ラグランジュの未定乗数法を用いた最小二乗推定により、式 (3) の未知数ベクトルを求める。得られた未知数ベクトルを分解することで、回転行列  $\mathbf{R}$  と並進ベクトル  $\mathbf{t}$  を求めることができる。式 (4)~(8) を制約条件とすることで、式 (3) の未知数ベクトルのノルムが一意に定まるため、スケール復元が可能となる。この結果をもとに復元した3次元点群座標およびカメラの相対姿勢・位置を表す  $\mathbf{R}, \mathbf{t}$  が次章の提案手法におけるバンドル調整の初期値候補となる。

### 3. 提案手法

#### 3.1 計測の精度向上のためのアプローチ

バンドル調整では、推定した点群を画像上に再投影し、元の画像上の点との距離の誤差などを評価関数と

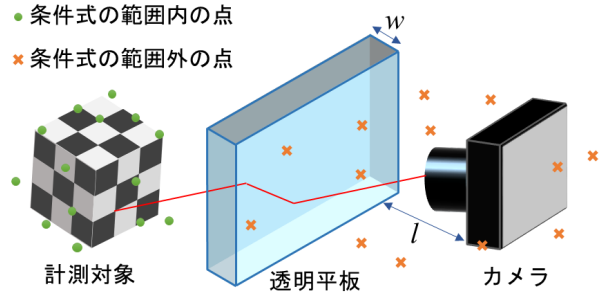


図3 推定された点群と幾何的な整合性

して最適化を実施することで、未知パラメータの調整を行う [7]。しかし、調整すべきパラメータが多く、また一般的に評価関数が非線形であることから、最適化の初期値が最適解への到達に大きく影響を与える。そのため、バンドル調整の初期値を適切に選択する必要がある。さらに、屈折を伴う計測系のバンドル調整においては、屈折の影響を考慮して復元精度を評価する必要がある。

以下、多視点から取得した画像組を幾何的な制約に基づいて検証することでカメラ視点を選抜し、適切な初期値として与えることで、屈折を用いた Structure from Motion における3次元復元の精度を向上させる手法を提案する。さらにバンドル調整においても、最適化の過程で生じ得る幾何的な例外条件を考慮した評価関数を提案する。

#### 3.2 幾何的な制約を考慮した初期値フィルタリング

本研究では、バンドル調整の初期値には、第2章で述べた2枚の画像による Structure from Motion により復元された点群とカメラの位置姿勢を用いる。

しかし、2枚の画像による Structure from Motion の際には、量子化誤差の影響で、最小二乗推定の結果に幾何的に整合性のない点群が得られる場合がある。整合性のない点群が初期値に含まれている場合、バンドル調整を行ったとしても正しい解を得られない場合が多いことが実験的に確認できている。そのため、幾何的な整合性を検証する必要がある。ここで考慮する整合性とは、本研究では復元結果の点群が透明平板に対してカメラと反対側に存在することは明らかであり、計測対象の推定位置が平板の中やカメラと同じ側にあるべきではないことを意味する (図3)。したがって、点群位置が平板に対して平板内を含むカメラと同じ側に復元される視点における復元結果は不適切であると、初期値から除く。

具体的には、以下のように初期値のフィルタリングを行う。2つのカメラ座標系を  $C$  と  $C'$ 、それぞれにおける奥行き方向の座標値を  $c_z$  と  $c'_z$  とする。図1のように、透明平板の厚さを  $w$ 、カメラの中心から平板までの距離を  $l$  とすると、復元結果について

$$c_z > w + l \text{ and } c'_z > w + l, \quad (9)$$

の条件を点群の推定結果のすべてに対して満たす結果のみ初期値に利用する。

初期値に使用する点群の決定の手順を以下に示す。まず、基準とするカメラ座標系を決定する。異なる  $n$  視点から計測した  $n$  枚の入力画像それぞれについて、カ

メラ座標系を考える．そのうち1つの画像のカメラ座標系は，一般性を失うことなく世界座標系と一致させることができる．この世界座標系と一致させたカメラ座標系を基準座標系とする．続いて，基準座標系の画像とそれ以外の  $(n - 1)$  枚の画像との間で，それぞれ屈折を用いた Structure from Motion を行う．ここでは合計  $(n - 1)$  回の Structure from Motion を実施することになる．その後，推定結果の点群が式 (9) の条件を満たしているかを， $(n - 1)$  回の推定結果ごとに検証する．そして，すべての点が式 (9) を満たしている点群を求め，それらの位置の平均をとる．この平均の点群の位置をバンドル調整の初期値とする．

バンドル調整の初期値に用いるカメラの位置姿勢も同様に選抜する．すなわち，すべての点が式 (9) を満たしている点群を推定した際の視点のみを使用し，条件を満たさない点群を推定した際の視点での計測結果はバンドル調整には用いない．

### 3.3 屈折を用いたバンドル調整による最適化の変数

一般的なバンドル調整による最適化の変数は，推定した点群の3次元座標と各カメラ座標系の回転行列と並進ベクトルである．ここで，回転行列は直交性を満たす必要があり，最適化の過程で適切に評価をする必要がある．

そこで本手法では，最適化の繰り返し計算中での回転行列の直交性を保証するために，回転行列から回転角を算出し，回転角を最適化の変数として扱う．回転角は軸の方向とその軸回りの回転で表現することで，回転行列の直交性を保つ．すなわち，推定した点群の3次元座標と各カメラ座標系の回転角と並進ベクトルを最適化の変数とする．

### 3.4 屈折を用いたバンドル調整における評価関数

一般的なバンドル調整では，推定した点群を画像上に再投影した点と元の画像上の点との距離の誤差を評価関数とする．しかし，屈折を用いた Structure from Motion において再投影誤差を評価関数とした場合には，3.2 節で述べたような幾何学的な不整合状態において，再投影誤差を用いてカメラ座標系の位置姿勢を正しく評価できない．例えば，推定の結果， $\mathbf{r}_{\text{out}}$  の  $z$  座標が負になる場合，この範囲に推定された点は画像上に正しく再投影されないため，正しく評価することができない．初期値フィルタリングで整合性のない初期値は除外されているが，バンドル調整の過程で不整合状態になる場合がある．そのため，評価関数でも整合性のない推定値を考慮する必要がある．

そこで本手法では，内側光線ベクトル  $\mathbf{r}_{\text{in}}$  と外側光線ベクトル  $\mathbf{r}_{\text{out}}$  の同一性に注目し，以下のように定量化し評価値として用いる．まず，内側光線ベクトル  $\mathbf{r}_{\text{in}}$  は，対応点の画像座標から求めることができる．次に，推定される外側光線ベクトル  ${}^c\hat{\mathbf{r}}_{\text{out}}^p$  は，

$${}^c\hat{\mathbf{r}}_{\text{out}}^p = {}^c\hat{\mathbf{p}}^p - {}^c\mathbf{d}^p, \quad (10)$$

と計算できる． ${}^c\hat{\mathbf{r}}_{\text{out}}^p$ ， ${}^c\hat{\mathbf{p}}^p$ ， ${}^c\mathbf{d}^p$  はそれぞれ，カメラ座標系  $c$  から見た点  $p$  に関する外側光線ベクトルの推定値，位置ベクトルの推定値，変化量ベクトルを表す．ベクトル  ${}^c\hat{\mathbf{p}}^p$  はバンドル調整の変数であり，ベクトル  ${}^c\mathbf{d}^p$  は式 (1) から算出できる．以上のように導かれる内側

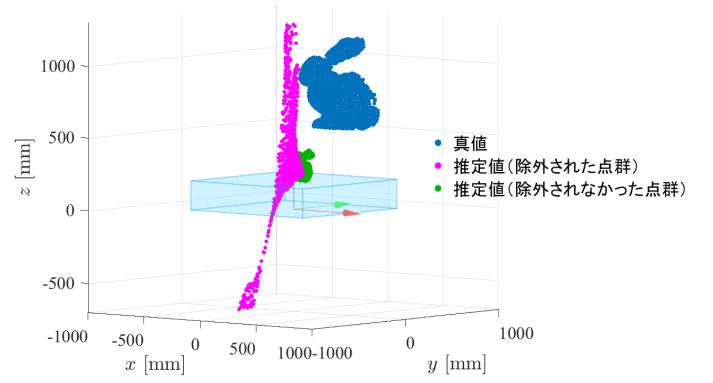


図4 初期値フィルタリングを適用した結果

光線ベクトルと外側光線ベクトルの方向を一致させるために，それぞれの正規化ベクトルの成分の比較を行う．すなわち， $\mathbf{r}_{\text{in}}$  と  $\hat{\mathbf{r}}_{\text{out}}$  の正規化ベクトルをそれぞれ  $\hat{\mathbf{r}}_{\text{in}}$ ， $\hat{\mathbf{r}}_{\text{out}}$  とすると，

$$e_1 = \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}} \left\| {}^c\hat{\mathbf{r}}_{\text{in}}^p - {}^c\hat{\mathbf{r}}_{\text{out}}^p \right\|^2, \quad (11)$$

を最小化する．ここで， $\mathcal{C}$  は  $(n - 1)$  個のカメラ座標系  $c$  の集合， $\mathcal{P}$  は推定した点群  $p$  の集合とする．また， ${}^c\hat{\mathbf{r}}_{\text{in}}^p$ ， ${}^c\hat{\mathbf{r}}_{\text{out}}^p$  の上付き添え字は，それぞれカメラ座標系  $c$  から見た点  $p$  に関する値であることを示す．

ここで，2つの光線ベクトルの評価では，点群の位置が平板内や平板とカメラの間に復元されてしまう不整合状態を考慮することができない．そのため，式 (9) の条件を満たす方向に最適化が進行するように，幾何的に整合性のない点群にはペナルティを与える．具体的には，以下の外点ペナルティ関数  $e_2$  を評価関数に加える．

$$e_2 = \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}} \left( \max(0, w + l - {}^c\hat{z}^p) \right)^2. \quad (12)$$

${}^c\hat{z}^p$  は点  $p$  の推定された位置ベクトル  ${}^c\hat{\mathbf{p}}^p$  の  $z$  座標である．点  $p$  が式 (9) の条件を満たさない場合，その距離に応じてペナルティが与えられる．最終的に  $e_1 + e_2$  を評価値として，レーベンバーグ・マルカート法で最適化を行う．

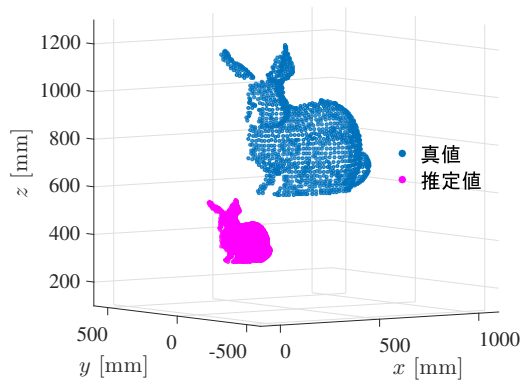
## 4. シミュレーション実験

### 4.1 Stanford Bunny を用いた実験設定

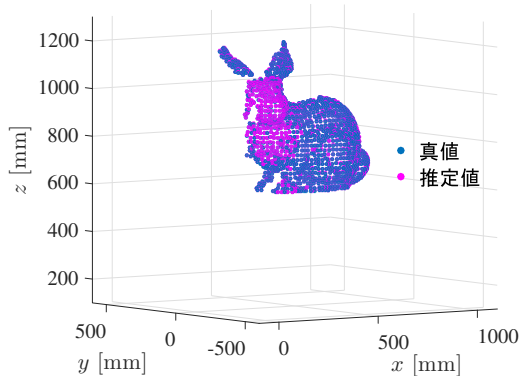
提案手法の有効性をシミュレーション実験により検証した．実験では，1,428 点の点群からなる Stanford Bunny を計測対象とした．異なる 10 視点から取得した 10 枚の画像を生成し，各視点間での点群中の各点の対応は正しく与えた上で，提案手法による 3次元復元を行った．実験条件は，空気と透明平板の屈折率  $n_1$  と  $n_2$  をそれぞれ 1.00, 1.49 とし，平板の厚さ  $w$  を 200 mm，カメラから平板までの距離  $l$  を 50 mm とした．計測点の量子化誤差については，計測対象の点群をピクセル精度で計測し，量子化誤差が大きい場合にも提案手法により精度よく復元可能であることの確認を目的とした．

### 4.2 実験結果

3.2 節で述べた初期値フィルタリングの結果を図4に示す．図中の座標系は世界座標系と一致させたカメラ



(a) 再投影誤差を評価関数とし、初期値フィルタリング未適用の場合



(b) 提案手法

図5 Stanford Bunny の復元結果 (1,428 点)

座標系を示している。点群が式 (9) を満たしていない場合、初期値から除外できていることが確認できる。

次に、復元結果を図5に示す。(a)は再投影誤差を評価関数とし、初期値フィルタリングを適用していない場合、(b)は提案手法での結果である。提案手法では10視点から5視点が選抜され、バンドル調整が行われた。提案手法を用いることで、計測対象が復元できていることが確認できる。各点における真値と推定値の誤差の平均値は(a)705.15 mm, (b)0.73 mmであった。以上より、量子化誤差が大きい場合にも提案手法を用いることで精度のよい復元に成功することが確認できた。

### 4.3 ランダム点群を用いた提案手法の評価

計測対象の形状が異なり、カメラの視点が変わった場合にも精度よく復元可能であることを検証するために、3種類のランダム点群に対して実験を行った。画像は異なる20視点から取得した画像を生成して使用し、点数は200点とした。それぞれに関し、異なる100種類のカメラ配置で実験を行った。屈折率 $n_1$ ,  $n_2$ , 平板の厚さ $w$ , カメラから平板までの距離 $l$ , 量子化誤差に関する条件は4.1節と同様とした。

精度の検証のために、計測対象点群の各点に対して真値と推定値の誤差の平均値が1 mm以下となった試行の回数を求めた。結果を表1に示す。3種類の点群すべてにおいて、精度のよい復元に成功する回数が増えたことが確認できる。提案手法により、計測対象の形状や視点の変化に対してより頑健に、3次元復元の精度を改善できることが明らかとなった。

表1 3種の点群について精度が1 mm以下になった回数

	提案手法なし	提案手法あり
ランダム点群1	13 / 100	93 / 100
ランダム点群2	15 / 100	93 / 100
ランダム点群3	9 / 100	89 / 100

## 5. 結論

本論文では、屈折を用いた Structure from Motion による3次元復元の精度を向上させるために、計測系の幾何学的整合性を考慮した初期値フィルタリング手法を提案した。さらに、バンドル調整を用いた全体最適化のために屈折の影響を考慮した評価関数を提案した。また、提案手法を用いてシミュレーション実験を行うことで、3次元復元の精度が向上することを確認し、その有効性を示した。

実測実験による検証が、今後の課題である。また、本研究ではカメラの光軸と垂直になるように透明平板を配置したが、透明平板の傾きが3次元復元の精度に与える影響の解析も今後の課題である。

## 謝辞

本研究の一部は、総合科学技術・イノベーション会議により制度設計された革新的研究開発促進プログラム (ImPACT) 「タフ・ロボティクス・チャレンジ」の援助を受けた。また、実験に用いた点群形状モデルの Stanford Bunny は、The Stanford 3D Scanning Repository より提供されたものである。

## 参考文献

- [1] Richard Hartley and Andrew Zisserman: Multiple View Geometry in Computer Vision, Cambridge University Press, Second Edition, 2004.
- [2] Davide Scaramuzza, Friedrich Fraundorfer, Marc Pollefeys and Roland Siegwart: "Absolute Scale in Structure from Motion from Single Vehicle Mounted Camera by Exploiting Nonholonomic Constraints", Proceedings of the 2009 IEEE International Conference on Computer Vision, pp. 1413–1419, 2009.
- [3] Ashwin P. Dani, Nicholas R. Fischer and Warren E. Dixon: "Single Camera Structure from Motion", IEEE Transactions on Automatic Control, Vol. 57, No. 1, pp. 238–243, 2012.
- [4] Manolis Lourakis and Xenophon Zabulis: "Accurate Scale Factor Estimation in 3D Reconstruction", Proceedings of the International Conference on Computer Analysis of Images and Patterns, Vol. 1, pp. 498–506, 2013.
- [5] Akira Shibata, Hiromitsu Fujii, Atsushi Yamashita and Hajime Asama: "Scale-Reconstructable Structure from Motion Using Refraction with a Single Camera", Proceedings of the 2015 IEEE International Conference on Robotics and Automation, pp. 5239–5244, 2015.
- [6] Akira Shibata, Hiromitsu Fujii, Atsushi Yamashita and Hajime Asama: "Absolute Scale Structure from Motion Using a Refractive Plate", Proceedings of the 2015 IEEE/SICE International Symposium on System Integration, pp. 540–545, 2015.
- [7] 岡谷貴之: "バンドルアジャストメント", 情報処理学会研究報告 2009-CVIM-167 (37), pp. 1–16, 2009.