

# 計測点の信頼度を考慮した全天球ステレオカメラによる運動推定

○野田 純平 (東京大学), Pathak Sarthak (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)

## Motion Estimation of Spherical Stereo Camera Considering Confidence of Reconstructed Points

○Junpei NODA (The University of Tokyo), Sarthak PATHAK (The University of Tokyo),  
Hiromitsu FUJII (The University of Tokyo), Atsushi YAMASHITA (The University of Tokyo),  
and Hajime ASAMA (The University of Tokyo)

Abstract : We propose a new localization and 3D measurement method using a spherical stereo camera. The disparity is searched in highly textured regions in the image and pixels are triangulated to create a semi-dense 3D map. The geometric uncertainty of every 3D point is calculated to remove points with low confidence. The remaining points are registered to accurately localize the camera.

### 1. 緒言

人間が直接作業することの困難な環境では、ロボットを用いて作業を行うことが期待されている。このとき、作業位置を特定するなどの目的でロボットの運動を推定することが求められる。例えば、ロボットを用いた橋梁の点検では、点検箇所を把握することが必要であり、画像情報を用いたロボットの運動推定に関する研究が行われている [1]。橋梁点検などのように、作業対象物に近接した状態においてより多くの情報を取得するためには、視野の限られた通常のカメラよりも視野の広いカメラが適している。本研究では、周囲 360° の情報を一度に取得できる極めて広い視野を有する全天球カメラに注目する。

全天球カメラを用いた運動推定の研究として、Structure from Motion (SfM) による手法が提案されている [2]。SfM は、カメラ 1 台の移動によって異なる視点からの画像を取得し、3次元復元と運動推定を行う手法である [3]。橋梁の点検においては、点検箇所を把握するために絶対的な移動距離を推定することが求められる。しかし SfM には、スケールの不定性や、カメラの移動量が小さい場合に運動推定精度が低下するという問題がある。

その他の運動推定の研究として、複数台のカメラを用いたステレオカメラによる 3次元計測の結果を利用する運動推定の手法が提案されている [4]。基線長が固定されているステレオカメラによる運動推定の性質として、上述のスケールの不定性や、カメラの移動量が小さい場合の運動推定精度の低下という問題は生じない。しかし、[4] は通常のカメラを用いており、全天球カメラの広い視野を活用したステレオによる運動推定の手法は構築されていない。

視野の広いカメラを複数台用いた 3次元計測の研究とし

ては、複数台の魚眼カメラを用いた手法が提案されている [5]。しかし、[5] は対応点を曲線上で探索しており、また 180° 程度の範囲のみを取得する魚眼ステレオカメラを用いて運動推定を行う場合、移動前後での計測に共通部分が含まれずに運動が推定できない場合が生じうるといった問題がある。全天球カメラを用いた場合には、対応点の探索を直線上で行うことができ、またカメラがどのような移動をしても必ず共通部分が得られて運動を推定することができる。

本研究では、全天球ステレオカメラを用いた 3次元計測結果を利用する運動推定の手法の構築を目的とする。ステレオカメラとの幾何的な関係から計測における誤差が大きいような点を、信頼度が低いとして除去することで運動推定の精度を向上させる手法を提案する。

### 2. 概要

全天球カメラにおいて、空間上の点は球上に投影される。投影された点には、Fig. 1 に示すように球面画像と正距円筒画像の 2 つの表し方がある。球上の 3次元座標  $\hat{\mathbf{x}} = [x, y, z]^T$  と正距円筒画像上の 2次元座標  $\hat{\mathbf{U}} = [u, v]^T$  の関係は、正距円筒画像の高さ  $h$  を用いて式 (1)~(3) で表される。

$$x = \sin\left(\frac{\pi u}{h}\right) \sin\left(\frac{\pi v}{h}\right), \quad (1)$$

$$y = \cos\left(\frac{\pi u}{h}\right) \sin\left(\frac{\pi v}{h}\right), \quad (2)$$

$$z = \cos\left(\frac{\pi v}{h}\right). \quad (3)$$

Fig. 2 (a) に全天球カメラとその投影面及び座標系を、Fig. 2 (b) に全天球ステレオカメラを示す。この計測系を移動させて、その際の運動推定を行う。運動推定の手順と

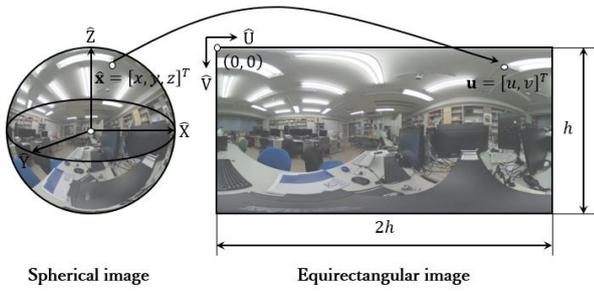


Fig. 1: Two formats of a spherical image

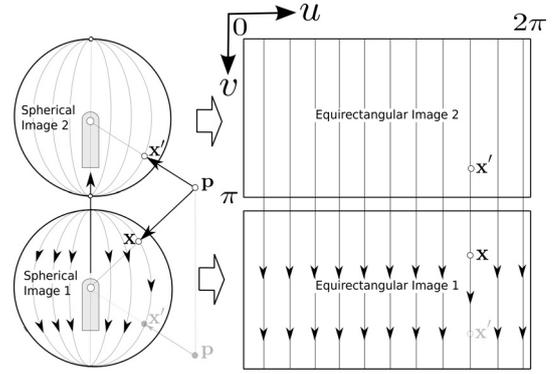


Fig. 3: Epipolar rectification of spherical stereo images

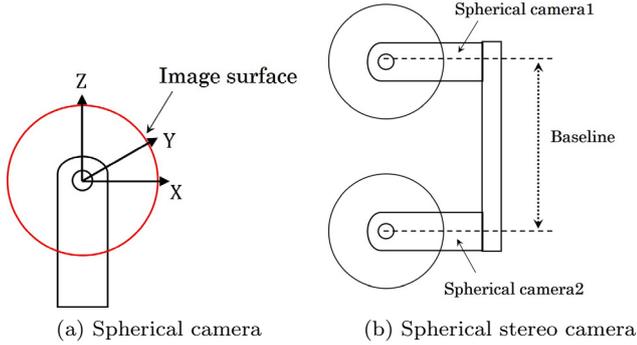


Fig. 2: Spherical camera and spherical stereo camera

しては、まず各地点で周囲の環境の3次元計測をし、信頼度の低い点を除去する。残った点群を地点間で位置合わせすることでカメラの運動を推定する。

### 3. 手法

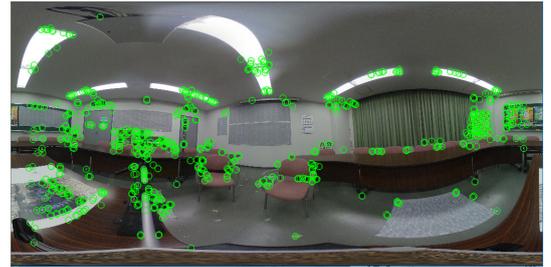
#### 3.1 3次元計測

3次元計測の際、異なる視点から取得した画像の組から対応点を求める必要がある。これには、局所特徴量を利用して2つの画像上で対応点を得る方法と、エピポーラ拘束を利用して局所特徴量を持つ点以外の画素に対しても対応点を探索する方法がある。復元した点群を精度良く位置合わせするためには、より多くの点群を復元して多くの情報を用いることが良いと考えられる。そこで、本研究ではエピポーラ拘束を利用して局所特徴量を持つ点以外の画素も計測する。

Fig. 3のように、空間上の点  $p$  がそれぞれの全天球カメラに投影される点  $x, x'$  は、点  $p$  と2つのカメラの投影中心の3点を通る平面がそれぞれの投影面と交わる大円上にある。空間上の各点に対応する全ての円は球上の2点で交わり、この2点がそれぞれ上端と下端になるようにそれぞれの画像を正距円筒画像に変換すると、対応する投影点は同じ  $u$  座標となる。この性質は全天球画像におけるエピポーラ拘束として知られており、これを利用することで



(a) Points used for semi-dense 3D map



(b) Feature points used for sparse 3D map

Fig. 4: Semi-dense map utilizes many more points as compared to sparse feature points.

全天球ステレオカメラでも通常のステレオカメラと同じように対応点を探索することができる。

Fig. 4 (a) で黒く塗られていない領域における全画素は、エピポーラ拘束を利用した探索により求められた準密 (semi-dense) な対応点であり、これら各点において視差を求めることができる。この領域全ての復元により、Fig. 4 (b) に緑の円で示すような疎 (sparse) な特徴点のみの復元と比較してより多くの点を計測することが可能である。Fig. 4 (a) では視差の推定にステレオマッチングライブラリ ELAS [6] を、Fig. 4 (b) では特徴点検出に AKAZE 特徴量 [7] を用いている。

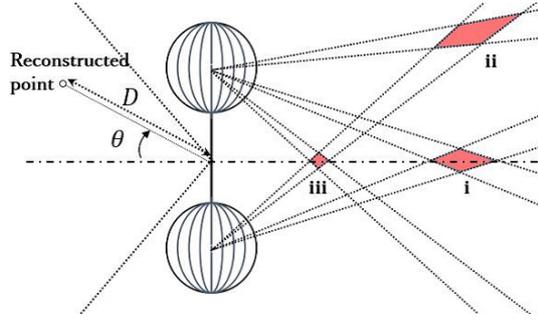


Fig. 5: Uncertainty calculation



Fig. 6: RICOH THETA S

### 3.2 計測点の信頼度

計測点は、ステレオカメラとの幾何的な関係から信頼度が異なり、信頼度が低い点ほど大きい誤差を含む可能性がある。Fig. 5 に赤く示す領域の範囲で、幾何的な関係から誤差が生じる。領域 ii よりも領域 i のほうが小さいことから、カメラ中点からの仰角  $\theta$  が小さい点ほど信頼度は高い。また、領域 i よりも領域 iii のほうが小さいことから、カメラ中点からの距離  $D$  が小さい点ほど信頼度は高い。信頼度が低い点として、閾値  $\theta_{th}$ ,  $D_{th}$  を用いて式 (4), (5) を満たす点を除去する。

$$\theta > \theta_{th}, \quad (4)$$

$$D > D_{th}. \quad (5)$$

### 3.3 位置合わせ

移動前後においてそれぞれ 3 次元計測によって復元した点群を位置合わせし、点群の剛体変換を表す変換行列を求めてカメラの運動を推定する。位置合わせには、信頼度の低い点をそれぞれ除去した移動前後の点群に対して、Iterative Closest Point (ICP) アルゴリズム [8] を用いる。 $n$  個の点から成る式 (6) の点群  $C_A$  と、 $m$  個の点から成る式 (7) の点群  $C_B$  をそれぞれ移動後と移動前の点群とする。

$$C_A = \{\mathbf{p}_i^A = (x_i^A, y_i^A, z_i^A)^T \mid 1 \leq i \leq n\}, \quad (6)$$

$$C_B = \{\mathbf{p}_j^B = (x_j^B, y_j^B, z_j^B)^T \mid 1 \leq j \leq m\}. \quad (7)$$

$C_A$  の点  $\mathbf{p}_i^A$  に対する  $C_B$  の最近傍の点が  $\mathbf{p}_k^B$  であるとき、 $C_A$  に対する剛体変換によって  $\mathbf{p}_i^A$  が点  $\mathbf{p}_i^{A'}$  に移動するとして、各  $i$  に対して式 (8) で表される距離を用いて式 (9) の条件を満たすように変換する。

$$d(i) = \sqrt{(x_i^{A'} - x_k^B)^2 + (y_i^{A'} - y_k^B)^2 + (z_i^{A'} - z_k^B)^2}, \quad (8)$$

$$\min \sum_{i=1}^n d(i). \quad (9)$$

移動によって得られた点群  $C_{A'}$  に対しても同様の処理を次々と繰り返すことで、 $C_A$  と  $C_B$  の中にそれぞれ含まれる同じ位置関係をもつ点群が一致するように  $C_A$  が移動する。

ICP では、適切な初期値を設定しなければ局所解に陥ってしまう。今回は、移動前後の同一カメラの画像間で局所特徴量を用いて対応点を検出し、8 点法 [9] で基本行列を求めることによって推定したカメラの運動を初期値とする。

ICP によって得られた点群の変換行列より、X, Y, Z 軸まわりの回転量  $\alpha$ ,  $\beta$ ,  $\gamma$  (deg) 及び X, Y, Z 軸方向の並進量  $t_x$ ,  $t_y$ ,  $t_z$  (m) を求められるので、これをカメラ運動の推定値とする。

## 4. 実験

### 4.1 実験設定

全天球カメラは、Fig. 6 に示す RICOH 社製の THETA S [10] を使用し、2 台の THETA S を基線長 0.48 m のステレオリグに取り付けて、Fig. 7 (a) に示すステレオカメラを構成した。任意の角度回転することができるカメラ台を用い、2 台の THETA S の中点が回転軸の真上になるようにステレオカメラをカメラ台に固定した。実験は Fig. 7 (b) に示す幅 6.6 m × 奥行 6.6 m × 高さ 2.7 m の室内で行い、Fig. 7 (c) に示すようにステレオカメラを、中点を通り Z 軸と平行な向きを軸として角  $\phi$  回転させた。このとき、Fig. 8 のように回転前後で信頼度が高い領域が変わる。それぞれの領域が重なる領域の大きさは角  $\phi$  によって異なり、 $\phi = 90$  deg のときに重なり大きさは最小となる。様々な回転角に対する運動推定精度を、信頼度の低い計測点を除去する場合としない場合で比較する。

X, Y, Z 軸まわりの回転量の真値  $\alpha_{tr} = 0$ ,  $\beta_{tr} = 0$ ,  $\gamma_{tr} = \phi$  として、 $\Delta\alpha = \alpha - \alpha_{tr}$ ,  $\Delta\beta = \beta - \beta_{tr}$ ,  $\Delta\gamma = \gamma - \gamma_{tr}$  をクォータニオンでの回転量に変換した値を推定の誤差  $err$  (deg) とする。

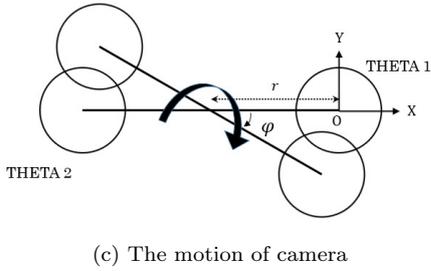
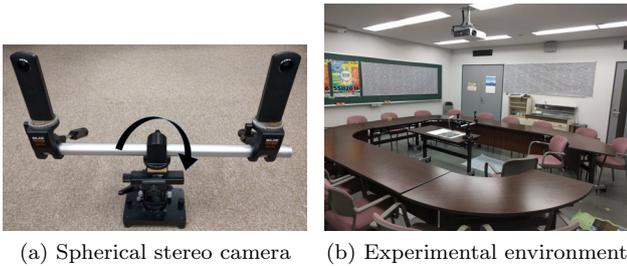


Fig. 7: Experimental conditions

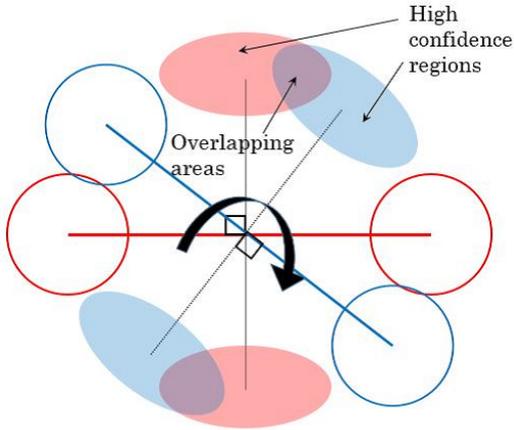


Fig. 8: When the stereo camera rotates, the regions of high confidence change and their overlapping areas become smaller.

#### 4.2 信頼度の低い計測点の除去

エピポーラ拘束を利用して、ステレオマッチングライブラリ ELAS [6] を用いて視差を求めた。Fig. 9 (a) に信頼度の低い計測点を除去する前の復元結果を示す。部屋全体が復元できており、全天球カメラの広い視野が活かされていることが確認できるが、ノイズが多く含まれている。Fig. 9 (b) に信頼度の低い計測点を除去した復元結果を示す。閾値  $\theta_{th}$ ,  $D_{th}$  をそれぞれ 75 deg, 5.8 m として、 $\theta > 75$  deg,  $D > 5.8$  m を満たす点を信頼度の低い点として除去した。特に、Fig. 9 (a) で赤い円で囲ったカメラの基線方向のノイズが除去されている。

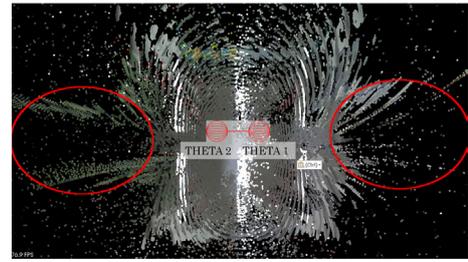


Fig. 9: Reconstruction results before and after removing points with low confidence

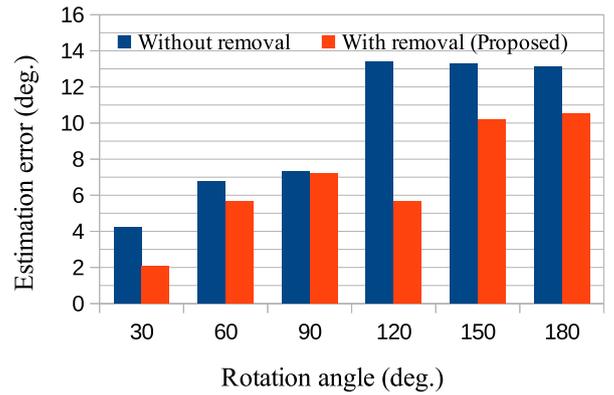


Fig. 10: Estimation errors (err) vs. Rotation angles ( $\phi$ )

#### 4.3 点群の位置合わせによる運動推定精度の評価

回転角  $\phi$  を 30 deg から 180 deg まで 30 deg ずつ変化した各状態に対する、姿勢推定の誤差を Fig. 10 に示す。全体的にみて信頼度の低い点を除去することによって誤差は低下している。 $\phi = 90$  deg のとき提案手法による精度向上の程度は小さいが、これは運動前後のカメラの姿勢の関係において信頼度の高い領域の重なりが小さいことが原因と考えられる。以上の結果より、運動推定において計測点の信頼度を考慮することが有効であると考えられる。

## 5. 結言

本研究では、ステレオカメラとの幾何的な関係から計測点の信頼度を考慮して全天球ステレオ計測を行い、運動推定精度の評価を行った。信頼度の低い計測点を除去することにより精度が向上し、計測点の信頼度を考慮することが運動推定において有効であることが示された。

ステレオカメラの並進移動を含む運動推定の精度評価や、計測点を除去せずに信頼度を用いて重み付けをして運動推定をすることで精度を高めることが今後の課題である。

## 謝辞

本研究の一部は、内閣府総合科学技術・イノベーション会議の戦略的イノベーション創造プログラム (SIP) 「インフラ維持管理・更新・マネジメント技術」(管理法人: 国立研究開発法人新エネルギー・産業技術総合開発機構) によって実施された。

## 参考文献

- [1] Y. Hada, M. Nakao, M. Yamada, H. Kobayashi, N. Sawasaki, K. Yokoji, S. Kanai, F. Tanaka, H. Date, S. Pathak, A. Yamashita, M. Yamada and T. Sugawara, “Development of a Bridge Inspection Support System Using Two-Wheeled Multicopter and 3D Modeling Technology”, *Journal of Disaster Research*, Vol. 12, No. 3, pp. 593–606, 2017.
- [2] S. Pathak, A. Moro, A. Yamashita and H. Asama, “Dense 3D Reconstruction from Two Spherical Images via Optical Flow-based Equirectangular Epipolar Rectification”, *Proceeding of the 2016 IEEE International Conference on Imaging Systems and Techniques (IST2016)*, pp. 140–145, 2016.
- [3] T. S. Huang and A. N. Netravali, “Motion and Structure from Feature Correspondences: A Review”, *Proceeding of the IEEE*, Vol. 82, No. 2, pp. 252–268, 1994.
- [4] J. Engel, J. Stückler and D. Cremers, “Large-Scale Direct SLAM with Stereo Cameras”, *Proceeding of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2015)*, pp. 1935–1942, 2015.
- [5] N. Kita, F. Kanehiro and Y. Kita: “3D Shape Measurement of a Large Cloth Close to a Fish-eye Stereo”, *Proceeding of the 2012 IEEE/SICE International Symposium on System Integration (SII2012)* pp. 895–900, 2012.
- [6] A. Geiger, M. Roser and R. Urtasun, “Efficient Large-Scale Stereo Matching”, *Proceeding of the 10th Asian Conference on Computer Vision (ACCV2010)*, pp. 25–38, 2010.
- [7] P. F. Alcantarilla, J. Nuevo and A. Bartoli, “Fast Explicit Diffusion for Accelerated Features in Non-linear Scale Spaces”, *Proceeding of the British Machine Vision Conference (BMVC2013)*, pp. 1–11, 2013.
- [8] P. J. Besl and N. D. McKay, “A Method for Registration of 3-D Shapes”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239–256, 1992.
- [9] R. I. Hartley, “In Defense of the Eight-Point Algorithm”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, pp. 580–593, 1997.
- [10] “RICOH THETA”, <https://theta360.com>