

魚眼カメラを用いた運動立体視に関する基礎検討

光本尚訓(中央大学), 新垣洋平, 下村倫子(日産自動車), 寺林賢司, 梅田和昇(中央大学)

Basic Examination on Motion Stereo Vision Using a Fish-Eye Camera

*Hisanori MITSUMOTO (Chuo University), Yohei ARAGAKI(Nissan Motor Co., Ltd.),
Noriko SHIMOMURA(Nissan Motor Co., Ltd.), Kenji TERABAYASHI(Chuo University),
Kazunori UMEDA(Chuo University)

Abstract— It is important to understand wide range 3-dimensional environment at once in the mobile robot's autonomy and operation support system of a car. For this purpose, it is thought that sensing with a fish-eye camera is effective. In this paper we propose methods for 3D scene measurement using a fish-eye camera. Our method measures 3D scene and estimates camera's motion at the same time by the structure-from-motion approach. Experiments show that a 3D scene model can be obtained with a fish-eye camera.

Key Words: fish-eye camera, structure from motion(SFM), camera's motion, 3D model

1. 緒言

近年, 魚眼カメラを利用した研究が盛んに行われている. このカメラは視野角が 180 °程度あり, 被写界深度が深い. そのため一定姿勢で広範囲のシーンを撮影することができ, カメラの台数の削減にも繋がる. このような利点から魚眼カメラは監視用途や自動車の視界補助システム [1][2] に利用されている. また移動ロボットの自律化や自動車の運転支援システムについても一度に広範囲の 3 次元環境を把握することが重要であり, 魚眼カメラによるセンシングが有用であると考えられる. この際, 精度良く 3 次元環境をセンシングするにはカメラの内部パラメータを正しく推定する必要がある. さらに一台のカメラで 3 次元計測を行うには, カメラの移動前後の相対的な位置・姿勢情報を推定する必要がある. 本研究では魚眼カメラの内部パラメータを推定し, Structure from Motion(SFM)を用いて, 画像のみからカメラ間の位置・姿勢の推定と計測対象の 3 次元計測を同時に行う. さらに 3 次元計測データに三角メッシュを貼ることで色情報を含んだ 3 次元の環境モデルを生成する.

2. 魚眼カメラの概要

魚眼レンズの射影方式は, シーンからレンズへの投射線がレンズの光軸となす角 θ と, 画像上の像高(射影点の光軸からの距離) r_f [pixel] によって, 一般に記述される. 用途によって射影方式が異なるが, 代表的なものは以下の二つである.

$$r_f = \delta\theta \text{ (等距離射影)} \quad (1)$$

$$r_f = \delta \sin \theta \text{ (正射影)} \quad (2)$$

$\delta = f/w$, f : 焦点距離 [mm], w : 画素サイズ [mm]

しかし一般に魚眼レンズの投影は, 誤差などの影響のために, 式 (1)(2) の理想的な射影方式には従わない. 本研究では, 魚眼カメラモデルを以下のように 3 次多項式で定義する [3].

$$r_f = k_1\theta + k_3\theta^3 \quad (3)$$

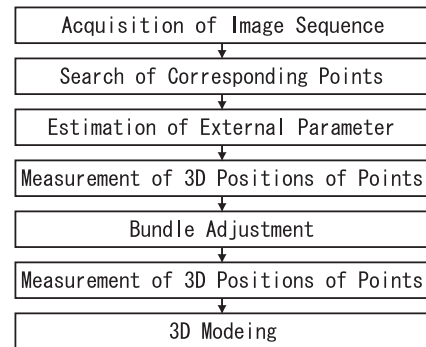


Fig.1 Process of 3D Measurement

ここで k_1, k_3 はカメラの内部パラメータである. さらに光軸のズレを考慮すると内部パラメータ I は

$$I = [k_1 \ k_3 \ c_u \ c_v]^T \quad c_u, c_v: \text{光軸点の位置} \quad (4)$$

となる. 内部パラメータ推定には中野の手法 [3] を利用する.

3. 手法概要

1 台の魚眼カメラより撮影位置の異なる 2 画像を取得し, 計測対象の 3 次元形状復元を行う.

処理の手順を図 1 に示す. まず動画像を取得する. 1 枚目の画像から特徴点を抽出し, 対応点を求める. 次に対応点情報よりカメラ間の位置・姿勢を推定し, 推定された外部パラメータを用いて特徴点の 3 次元座標を初期推定する. その後, バンドル調整 [11] により外部パラメータをより精度良く推定し, 最後に特徴点の 3 次元座標を再推定する.

4. 計測対象の 3 次元計測手法

4.1 対応点探索

得られた 2 枚の画像間において対応点を抽出するために, 一枚目の画像から追跡しやすい点, すなわちコーナーやエッジを特徴点として抽出し対応点を求める. 屋

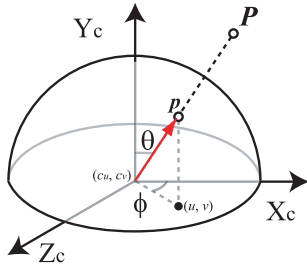


Fig.2 Sphere Model

内においては比較的，特徴点の対応点を求めやすいことから，KLT-Tracker[4]を用いる．屋外では，処理時間はKLT-Trackerよりかかるが，精度良く対応点を求めるためにSIFT[5]を用いる．

4.2 光線ベクトルの算出

図2に示すように，カメラのレンズ中心から特徴点の3次元空間上の点Pに向かう光線を光線ベクトル p とする．本研究で用いる魚眼カメラにおける光線ベクトル p は，天頂角 θ と方位角 ϕ を用いて以下のように表すことができる．

$$p = \begin{bmatrix} \sin \theta \cos \phi & \cos \theta & \sin \theta \sin \phi \end{bmatrix}^T \quad (5)$$

天頂角 θ と方位角 ϕ は特徴点の画像座標 $[u, v]^T$ より式(6)(7)を用いて求めることができる．

$$k_1 \theta + k_3 \theta^3 - \sqrt{(u - c_u)^2 + (v - c_v)^2} = 0 \quad (6)$$

$$\phi = \tan^{-1} \left(\frac{v - c_v}{u - c_u} \right) \quad (7)$$

天頂角 θ に関しては式(6)より解が複数個存在する可能性があるが， θ の範囲はおおよそ $0 \sim \pi/2$ であることから解を限定する． θ をBrent法[6]より求める．

4.3 基本行列Eの算出

二つの画像における光線ベクトル $p_i = [x_i, y_i, z_i]^T$ と $p'_i = [x'_i, y'_i, z'_i]^T$ を用いてカメラ間の位置・姿勢の情報からなる基本行列 E を求める[7]．基本行列と光線ベクトルの間には以下の関係式が成り立つ．

$$p_i'^T E p_i = 0 \quad (8)$$

式(8)を変形すると式(9)となる．

$$u_i^T e = 0 \quad (9)$$

ここで，

$$u_i = [x_i x'_i, y_i x'_i, z_i x'_i, x_i y'_i, y_i y'_i, z_i y'_i, x_i z'_i, y_i z'_i, z_i z'_i]^T$$

$$e = [e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23}, e_{31}, e_{32}, e_{33}]^T$$

である． e_{ij} は E 行列の i 行 j 列の要素である．

E の要素数は9であるが， E には定数倍の不定性が残るため，最低8組以上の対応点の光線ベクトルが得られれば，式(9)の連立方程式を解くことで E 行列を求めることができる．

4.4 外れ値除去

4.1節で求めた対応点データには誤対応のデータが含まれている可能性がある．誤対応のデータを含めしまうと基本行列 E を正しく算出できず，カメラ間の位置・姿勢の推定に悪影響を及ぼす．そこで本手法ではこのような悪影響を及ぼす点を外れ値として除去するためにRANSAC[8]を用いる．誤対応データ除去の処理の流れは以下の通りである．

1. 8個の特徴点をランダムに選択
2. 4.3節の手法で基本行列 E_{rand} を計算
3. 画像1,2の光線ベクトル p_i, p'_i に対して式(10)を満たす特徴点の数 k をカウント

$$|p_i'^T E_{rand} p_i| < q \quad q: \text{閾値} \quad (10)$$

4. 1~3を適当な回数繰り返す
5. k が最大となった時の式(10)を満たさなかった点を外れ値とする
6. 外れ値を除き基本行列 E を求める．

4.5 カメラ間の位置・姿勢の推定

基本行列 E からカメラ間の位置・姿勢を求める．カメラ間の並進ベクトルを $t = [t_x, t_y, t_z]^T$ とすると， E と t の間には

$$E^T t = 0 \quad (11)$$

が成り立つ．式(11)より EE^T の最小の固有値に対応する固有ベクトルとして t の方向の単位並進ベクトル \bar{t} が求まる．また， \bar{t} にスケールを合わせた E を \bar{E} とすると

$$\bar{E} = [\bar{t}]_{\times} R, \quad (12)$$

$$[\bar{t}]_{\times} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}. \quad (13)$$

が成り立つ．よって回転行列 R は

$$c_1 = \|\bar{E} - [\bar{t}]_{\times} R\|^2 \quad (14)$$

を最小とする解として求められる． $\|\cdot\|$ は，行列のFrobeniusノルムを表す．本手法では式(14)を4元数を用いて解く[9]．

4.6 3次元形状の復元

カメラ間の位置・姿勢を表す単位並進ベクトル \bar{t} と回転行列 R が復元できれば，2画像間の対応づけを利用して，特徴点の3次元座標 $X_i = [X_i, Y_i, Z_i]^T$ を算出することができる．図3にその概念図を示す．移動前と移動後のカメラのレンズ中心から特徴点へ向かう光線ベクトルが空間上で交わる点を特徴点の3次元座標とする．実際にはノイズの影響で二つの光線ベクトルは空間上で交わらない．そこで， X_i から二直線までのユークリッド距離の2乗和 d が最小となるように3次元座標を求める．最終的に3次元座標 X_i は式(15)

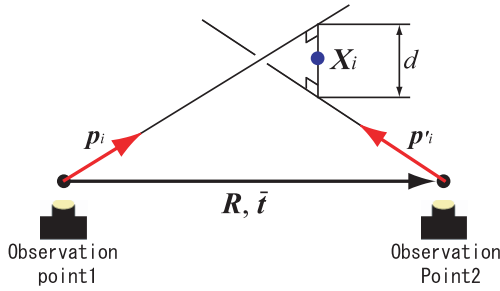


Fig.3 Measurement of 3D Position

~(17) で表される [9] .

$$X_i = B^{-1}b, \quad (15)$$

$$B = 2I - \frac{p_i p_i^T}{p_i^T p_i} - \frac{R p_i' p_i'^T R^T}{p_i'^T p_i'}, \quad (16)$$

$$b = \left(I - \frac{R p_i' p_i'^T R^T}{p_i'^T p_i'} \right) \bar{t}. \quad (17)$$

4.7 計測精度の悪い点の除去

計測精度は一般にカメラの移動方向に近いほど、またカメラから遠いものほど悪い。計測結果には精度の良いものと悪いものが存在することになる。ここで計測精度の良い対応点を使用するため、計測結果 X_i を2画像上の特徴点座標 $[u_i, v_i]^T$, $[u_i', v_i']^T$ で偏微分したものから式(18)で誤差を評価するベクトル g を求める[10] .

$$g = \left| \frac{\partial X_i}{\partial u_i} \right| + \left| \frac{\partial X_i}{\partial v_i} \right| + \left| \frac{\partial X_i}{\partial u_i'} \right| + \left| \frac{\partial X_i}{\partial v_i'} \right| \quad (18)$$

ベクトル g が式(19)を満たさない点は計測精度の悪い点として除去する .

$$\|g\| < h \quad (19)$$

h は誤差評価のための閾値である .

4.8 バンドル調整

4.7節までに推定されたカメラ間の位置・姿勢や3次元計測データは、特徴点に含まれる誤差を考慮していないため、必ずしも良い推定結果であるとは限らない。そこで、4.5節で推定した外部パラメータと4.6節で推定した3次元データを初期値として、最適化を行う。この際、特徴点の逆投影誤差の2乗和を最小化するバンドル調整[11]と呼ばれる手法を用いる。

ここで逆投影誤差とは、推定された特徴点の3次元座標を画像平面に逆投影したときの座標と、元の画像座標との差である。観測点1,2における特徴点の逆投影点 u_{b1} , u_{b2} は式(20)(21)で表される。

$$u_{b1} = \begin{bmatrix} u_{b1} \\ v_{b1} \end{bmatrix} = \frac{k_1 \theta + k_3 \theta^3}{\sqrt{X_i^2 + Y_i^2}} \begin{bmatrix} X_i \\ Y_i \end{bmatrix} \quad (20)$$

$$u_{b2} = \begin{bmatrix} u_{b2} \\ v_{b2} \end{bmatrix} = \frac{k_1 \theta' + k_3 \theta'^3}{\sqrt{X_i'^2 + Y_i'^2}} \begin{bmatrix} X_i' \\ Y_i' \end{bmatrix} \quad (21)$$

ここで、

$$\theta = \tan^{-1} \left(\frac{X_i^2 + Y_i^2}{Z_i} \right), \quad \theta' = \tan^{-1} \left(\frac{X_i'^2 + Y_i'^2}{Z_i'} \right)$$

である。 $X_i = [X_i, Y_i, Z_i]^T$ と $X_i' = [X_i', Y_i', Z_i']^T$ はそれぞれ観測点1,2における特徴点の3次元座標であり、 X_i' は X_i を並進ベクトル \bar{t} と回転行列 R で座標変換したものである。さらに逆投影誤差の2乗和は式(22)となる。

$$c_2 = \sum_{i=1}^n \sum_{j=1}^2 (u_j - u_{bj})^2 \quad (22)$$

c_2 は逆投影誤差の2乗和、 u_1, u_2 は観測点1,2における元の特徴点座標、 n は逆投影する特徴点の数である。 c_2 の最小化にはLevenberg Marquardt法[6]を用いる。

4.9 テクスチャマッピング

3次元計測データから3Dモデルを生成するために、以下のモデリングを行う。得られた3次元計測データにドロネー分割を用いて三角メッシュを生成する。生成された三角メッシュの各面にテクスチャーを貼り付けることで、色情報をもつ3Dモデルを生成する。

5. 実験

5.1 実験環境

提案した手法を用いて、異なる位置から撮影した2枚の画像から計測対象の3Dモデリングを行った。CCDカメラはPointGreyResearch社のDragonfly2を、魚眼レンズはスペース社のTV1634Mを用いた。TV1634Mの内部パラメータは実験により推定した $k_1 = 383.778$, $k_3 = -22.272$, $c_u = 17.534$, $c_v = 17.688$ を用いる。画像サイズは1024[pixel] × 768[pixel]である。屋内、屋外の計測対象をそれぞれ図4の(a)(b)に示す。屋内環境では廊下を、屋外では建物を計測対象とした。移動前のカメラの光軸とカメラの移動方向とのなす角を θ_c (光軸を基点に反時計回りを正とする)として、カメラの撮影位置は以下のようにした。屋内では基線長 $b_c = 18[\text{cm}]$, $\theta_c = -\pi/4$ とし、屋外では $b_c = 108[\text{cm}]$, $\theta_c = 0$ とした。姿勢は一定とした。

5.2 実験結果

屋内の実験結果を図5,6に、屋外の実験結果を図7,8に示す。図5(a)は3次元計測結果の正面図、図5(b)と図7は3次元計測結果の平面図を示している。図6,8は生成された3Dモデルを示している。図5~8より、屋内や屋外の計測対象の大きさが復元できていることがわかる。しかし、正確な3次元形状が復元できたとは言いがたい。この主な原因として、以下の2点が挙げられる。一つ目は特徴点の誤対応を除去しきれていない。二つ目は計測点が疎である領域では、うまく三角メッシュが貼れていない点である。この問題点に対しては、観測する画像のデータを2枚から増やしていくことで対応していきたいと考えている。

6. 結言

魚眼カメラを用いて3次元環境モデルを生成する手法を提案し、実装した。実験結果より計測対象の大き



(a) Indoor Scene

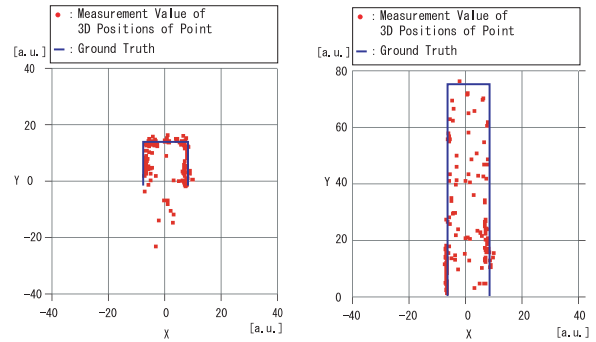


(b) Outdoor Scene

Fig.4 Measurement Scenes

かな形状を復元することができた．今後の展望としては，観測する画像データを増やすことで，より詳細な3Dモデルを生成させることと，生成された3Dモデルの定量的評価および誤差解析が挙げられる．またカメラの内部パラメータの推定精度を高めることで，計測精度の向上を図る．

- [1] 鈴木政康, 知野見聡, 河野照久: “俯瞰ビューシステムの開発”, 電子情報通信学会論文誌 D-II, Vol.J88-D-II, No.9 (2005) pp.1847-1856.
- [2] K. Oizumi: “Development of “All-Around View” System,” SAE Technical Paper Series(2003).
- [3] 中野誠士, 李仕剛, 千葉則茂: “球面画像獲得のための魚眼カメラの校正”, 電子情報通信学会論文誌 D-II, Vol.J88-D-II, No.9 (2005) pp.1847-1856.
- [4] J. Shi and T. Kanade: “Good Features to Track,” Proc. of CVPR, 1994.
- [5] D. G. Lowe: “Object Recognition from Local Scale Invariant Features,” Proc. of IEEE International Conference on Computer Vision (ICCV), pp.1150-1157, 1999.
- [6] 丹慶勝市, 奥村晴彦, 佐藤俊郎, 小林誠: “NUMERICAL RECIPES in C”, 技術評論社.
- [7] 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).
- [8] M. A. Fischler and R. C. Bolles: “Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography, ”Communications of the ACM, Vol.24, pp.381-395, 1981.
- [9] 徐剛: “写真から作る3次元CGイメージ・ベースド・モデリング&レンダリング”, 近代科学社 (2001).
- [10] 山下 淳, 原田 知明, 金子 透: “全方位カメラ搭載移動ロボットによる Structure from Motion を用いた3次元環境モデリング”, 日本機械学会論文集 (C 編), Vol.73, No.726, pp.512-519, 2007.
- [11] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon: “Bundle Adjustment -A Modern Synthesis,” Vision Algorithms: Theory & Practice,” Springer-Verlag LNSC 1883, 2000.



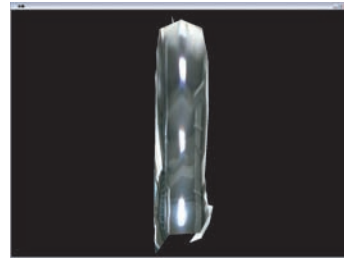
(a) Front Elevation(x-y)

(b) Ground Plan(x-z)

Fig.5 Result of 3-D Measurement (Indoor Scene)



(a) Front Elevation (x-y)



(b) Ground Plan (x-z)

Fig.6 Texture Mapping (Indoor Scene)

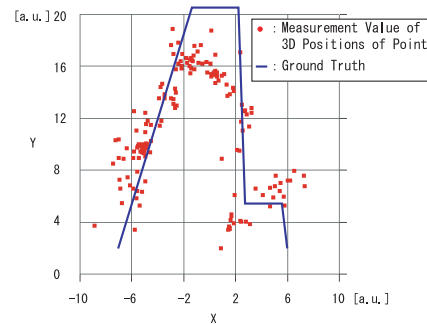


Fig.7 Result of 3-D Measurement (Outdoor Scene: Ground Plan(x-z))



Fig.8 Texture Mapping (Outdoor Scene)