正距円筒画像への適用のための 深層学習手法による距離画像推定手法のスケール補正

○筑後光 †, Sarthak Pathak ‡, 梅田和昇 ‡

 \bigcirc Hikaru CHIKUGO, † Sarthak PATHAK ‡ and Kazunori UMEDA ‡

†:中央大学大学院理工学研究科, chikugo@sensor.mech.chuo-u.ac.jp

‡:中央大学理工学部, {pathak, umeda} @mech.chuo-u.ac.jp

<要約>本研究では,透視投影画像で学習された深層学習による距離画像推定手法を正距円筒画像に用いた際に 得られる距離のスケールを補正する手法を提案する.多くの深層学習手法では,透視投影画像を用いて学習を行っ ている.そのため,広視野を計測することが可能な魚眼カメラで撮影された投影方式の異なる正距円筒画像のよ うな画像に対応出来ない可能性がある.また,正距円筒画像を用いたデータセットは多くないため,多様な環境 に対応出来ない可能性がある.そこで,深層学習手法により得られた距離を非学習手法で得られた距離に近づく ようにスケールを補正することで,正距円筒画像に対応した高精度な距離画像計測を目指す. <キーワード>魚眼縦ステレオカメラ,正距円筒画像,スケール補正,3次元計測,広視野センシング

1 序論

近年,距離センサを用いた自動運転の研究が盛んに 行われている.代表的な距離センサとして,ステレオ カメラや LiDAR,ソナーなどがある.しかし,これら のセンサには距離計測範囲や計測密度などの問題があ る.そこで本研究では,一台で広範囲を計測出来る距 離センサとして魚眼ステレオカメラに注目する.

魚眼ステレオカメラを用いて距離画像計測を行う手 法として,大橋らは魚眼画像から正距円筒画像への変 換を行うことで魚眼画像の歪みを低減した距離画像計 測を実現した [1].しかし,誤マッチングや基線長の方 向などにより距離計測精度が不十分である.そこで,筑 後らは魚眼カメラを縦に配置する魚眼縦ステレオカメ ラを構築することでステレオマッチングを簡便にし,左 右端の距離計測精度を向上させた [2].魚眼縦ステレオ カメラを用いて距離画像計測を行う場合,セミグロー バルマッチングなどの非学習手法,もしくは RAFT な どのをはじめとする深層学習手法を用いる [3,4].一般 的なステレオカメラを用いる場合,距離精度は非学習 手法よりも深層学習手法の方が高いことが多い.その ため,魚眼縦ステレオカメラにおいても深層学習手法 を用いることで非学習手法より距離精度が高くなると



図1 提案手法の概要

考えられることが期待できる.しかし,これらの深層 学習手法では,正距円筒画像ではなく透視投影画像を 用いて学習を行っていることが多い.また,正距円筒画 像を用いた大規模なデータセットは多くないため,多 様な環境に対応出来ない可能性がある.そこで本研究 では,非学習手法で得られた距離を用いて深層学習手 法で得られた距離のスケールを補正することで正距円 筒画像に対応した高精度な距離画像計測を目指す.

2 提案手法

2.1 概要

提案手法の概要を図1に示す.魚眼縦ステレオカメ ラから正距円筒画像を取得し,テクスチャ領域のみ非 学習手法により距離画像計測を行う.その後,ノイズ



図 2 魚眼画像から正距円筒画像への変換





除去を行う.最後に深層学習手法で得られた距離のス ケールを補正する.

2.2 正距円筒画像の取得

魚眼画像には特有の歪みがある.そのため,図2に示 すように魚眼画像から正距円筒画像への変換を行う.そ の際,内部パラメータを考慮して行う.内部パラメータ の推定には Scaramuzza らの提案した OcamCalibToolbox を用いる [5].

2.3 非学習手法による距離画像計測

深層学習手法で得られた距離のスケールを補正する 際、非学習手法で得られた距離に近づけるように処理 を行う. そのため、非学習手法により信頼度の高い距 離を得る必要がある.上下カメラで得られた正距円筒 画像を用いてテクスチャ領域のみ非学習手法により距 離画像計測を行う. その後、ノイズ除去を行う. 本研 究では、非学習手法としてブロックマッチングを用い た. テクスチャ領域を抽出する際, エッジフィルタを 用いる.ノイズ除去を行う際、オクルージョンや繰り 返しパターンによるノイズを除去する. オクルージョ ンを考慮するためにもう片方のカメラ基準で対応点を 探索し、対応点が一致しなかった場合にはノイズとし て除去する.繰り返しパターンを考慮するためにウィ ンドウを用いる. そのウィンドウ内の最大距離と最小 距離の差がしきい値より大きい場合、そのウィンドウ 内の距離最大と最小距離の点をノイズとして除去する.



(a) ノイズ除去前
(b) ノイズ除去後
図 4 非学習手法による距離画像計測の結果





(b) スケールの補正後

(a) 深層学習手法

図 5 スケールの補正を行った結果

ウィンドウ内の最大距離と最小距離の差がしきい値以 下になるまでこの処理を繰り返す.図3と図4にこれ らの処理を行った結果を表す.図3(a)は3DCGソフト のBlender で得られた正距円筒画像を表しており,図 3(b)は図3(a)にエッジフィルタを用いた結果を表して いる.図4(a)はノイズ除去を行う前の距離画像を表し ており,図4(b)はノイズ除去を行った後の距離画像を 表している.図4は距離を表しており,近いほど赤く, 遠いほど青いことを表している.図4より,窓の領域 などにおいてノイズが除去されていることが分かる.

2.4 スケールの補正

非学習手法により得られた距離を用いて深層学習手 法で得られた距離の補正を行う.補正を行う際,非学 習手法と深層学習手法で得られた距離の誤差と再投影 誤差を用いた誤差関数

$$E_{(\mathbf{a},\mathbf{b},\mathbf{c})} = w_1 \cdot E_{\mathbf{d}(\mathbf{a},\mathbf{b},\mathbf{c})} + w_2 \cdot E_{\mathbf{r}(\mathbf{a},\mathbf{b},\mathbf{c})} \tag{1}$$

を最小化するパラメータ a,b,c を Levenberg-Marquardt 法 [6] により求める.ここで, w_1,w_2 はそれぞれの誤差に対する重み, $E_{d(a,b,c)}$ は距離の誤 差を表しており, $E_{r(a,b,c)}$ は再投影誤差を表している. 図5に処理を行った結果を示す.図5(a)は深層学習手 法である RAFT[4] を用いて得られた距離画像を表し

表 1 各手法による距離画像計測の結果

	RMSE (m)
ブロックマッチング	0.54
RAFT	0.48
提案手法	0.27

ており,図 5(b)は RAFT で得られた距離に上記の処 理を行った結果を表している.図5は距離を表してお り,近いほど赤く,遠いほど青いことを表している.

2.4.1 距離の誤差

 $E_{\rm d(a,b,c)}$ kt

$$E_{\rm d(a,b,c)} = \sum_{i=1}^{N} (d_i - D_i)^2$$
 (2)

を用いる.ここで, *d_i* は非学習手法で得られた距離, *D_i* はスケールの補正を行った深層学習手法で得られた距離を表しており.

$$D = a \cdot d' + b + c \cdot x \cdot d' \tag{3}$$

とする. ここで, d' は深層学習手法で得られた距離を 表している. また, x は

$$x = \left|\frac{u - W/2}{W}\right| \tag{4}$$

とし,画像中心の横座標を原点としている.そのため, 画像端になるにつれパラメータ*c*の影響が大きくなる. ここで,*u*は横座標を表し,*W*は画像横のサイズを表 している.

2.4.2 再投影誤差

E_{r(a,b,c)} はウィンドウを用いた処理を行い,

$$E_{\rm r(a,b,c)} = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=-m}^{m} \sum_{l=-m}^{m} \left(I_{\rm d(k,l)} - I_{\rm u(k,l)} \right)^2 \quad (5)$$

を用いる.ここで、 $I_{d(k,l)}$ は

$$I_{d(\mathbf{k},\mathbf{l})} = I_d(u_i + k, v_i + l) \tag{6}$$

と表し, *I*_{u(k,l)} は

$$I_{u(k,l)} = I_u(u'_i + k, v'_i + l)$$
(7)

と表す. $I_{d(k,l)}$ は下カメラにおける輝度を表しており, $I_{u(k,l)}$ は上カメラにおける輝度を表している.また, u_i, u'_i は横座標, v_i, v'_i は縦座標を表している.

表 2 各パラメータによる距離画像計測の結果

	RMSE (m)
a, b, c	0.27
а	0.31
a, b	0.31
a, c	0.43
b, c	0.47

2.4.3 各パラメータの初期値

式 (1) を最小化する際,パラメータa, b, cの初期値が 必要となる.それぞれのパラメータの初期値を a_0, b_0, c_0 とすると

$$a_0 = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{d_i}{d'_i} \right) \tag{8}$$

$$b_0 = \frac{1}{N} \sum_{i=1}^{N} \left(d_i - a_0 \cdot d'_i \right) \tag{9}$$

$$c_0 = \frac{1}{N} \sum_{i=1}^{N} \left(\left(d_i - \left(a_0 \cdot d'_i + b_0 \right) \right) \cdot x \right)$$
(10)

と表せる.式(8)-式(10)で得られた初期値を用いて式(1)を最小にしたパラメータ*a*,*b*,*c*よりスケール補正された深層学習手法の距離を式(3)により求める.

3 精度評価実験

3.1 実験条件

本実験は、3DCG ソフトの Blender を使用し、仮想 環境において行った.提案手法によって、深層学習手法 の距離精度が向上するか RMSE を用いて確認した.ま た、一部のパラメータを用いずに深層学習手法で得ら れた距離のスケール補正を行うことで距離精度に最も 影響を与えるパラメータを確認した.本実験では、非 学習手法として、ブロックマッチング、深層学習手法と して、RAFT[4] を用いた.

3.2 実験結果

実験結果を表1,2と図6-9に示す.表1はBlenderで 得られる真値の距離と各手法で得られた距離の RMSE を比較した結果を表している.表2はBlenderで得ら れる真値の距離と各パラメータによる処理で得られた 距離の RMSEを比較した結果を表している.図6は各 手法により得られた点群を表している.図7-9は真値 の点群と各手法の点群を比較している.赤の点群は真 値を表しており,緑は RAFT,青は提案手法,桃色は 2.3節で得られたブロックマッチングの点群を表してい る.表1より,提案手法により得られた距離が最も距離



(a) RAFT



(b) 提案手法



(c) 真値 図 6 各手法により得られた点群

精度が高いことが分かる.表2より,全てのパラメー タ a,b,c を用いることで得られた距離が最も距離精度 が高いことが分かる.また,各パラメータごとに深層 学習手法で得られた距離のスケール補正を行った結果, パラメータ a を用いた際とパラメータ a,b を用いた際 に最も距離精度が高いことが分かる.これより,パラ メータ a が距離精度に最も影響を与えていると考えら れる.図7-9より,RAFT で得られる点群より提案手 法で得られる点群の方が真値の点群に近いことが定性 的に分かる.また,図7(c)や図8(c),図9(c)より,テ クスチャ領域のみブロックマッチングにおいて距離画 像計測を行った結果,ブロックマッチングにおいて距離画 像計測を行った結果,ブロックマッチングにおいて得 られた点群が真値の点群に近いことが定性的に分かる. 提案手法の距離精度が向上した原因として,非学習手 法により得られた信頼度の高い距離になるべく近づけ



(a) RAFT



(b) 提案手法



(c) ブロックマッチング図 7 前方から見た際の真値の点群との比較

るようにパラメータ *a*,*b*,*c* を求めたことが考えられる.

4 結論

本論文では,透視投影画像を用いて学習された距離 画像推定手法によって得られた距離のスケールを正距 円筒画像に対応させる手法を提案した.実験により,提 案手法の距離精度が向上した.また,提案手法の距離 精度に最も影響を与えるパラメータを検証した.

今後の展望として,他の深層学習手法を用いて実験 を行うことや実環境での実験を行うことなどが挙げら れる.

参考文献

 Akira Ohashi, Fumito Yamano, Gakuto Masuyama, Kazunori Umeda, Daisuke Fukuda, Kota Irie, Shuzo Kaneko, Junya Murayama and Yoshitaka Uchida, "Development of ranging method for inter-vehicle distance using visible light communication and image processing," 2017 IEEE/SICE International Symposium on System Integration (SII), pp.535-540 (2017)



(a) RAFT



(b) 提案手法



(c) ブロックマッチング図 8 横方向から見た際の真値の点群との比較

- [2] Hikaru Chikugo, Kento Arai, Sarthak Pathak and Kazunori Umeda, "Fisheye Stereo Camera using Fisheye Vertical Stereo Method," 2024 IEEE International Conference on Image Processing (ICIP2024), pp.3512-3518 (2024)
- [3] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 807-814, (2005)
- [4] Lipson Lahav, Teed Zachary and Deng Jia, "RAFT-Stereo: Multilevel Recurrent Field Trans-



(a) RAFT



(b) 提案手法





forms for Stereo Matching," International Conference on 3D Vision (3DV), (2021)

- [5] Davide Scaramuzza, Agostino Martinelli and Roland Siegwart, "A Toolbox for Easily Calibrating Omnidirectional Cameras," 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.5965-5701 (2006)
- [6] Donald W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," Journal of the Society for Industrial and Applied Mathematics, vol.11, no.2, pp.431-441, (1963)