

擬似バイラテラルフィルタによる魚眼ステレオカメラの性能向上*

飯田浩貴** 佐久田朝優*** 池 勇勳† 梅田和昇††
大橋 明††† 福田大輔††† 金子修造††† 村山純哉†††

Performance Improvement of Fisheye Stereo Camera by Bilateral-Like Filter

Hiroataka IIDA, Tomoyuu SAKUDA, Yonghoon JI, Kazunori UMEDA,
Akira OHASHI, Daisuke FUKUDA, Shuzo KANEKO and Junya MURAYAMA

The purpose of this study is to develop a wide-range measurement system for parking support systems. A fisheye stereo camera is used as a sensor for this purpose. The fisheye camera has a wide angle of view, so it is suitable for parking assistance sensors. In previous research, distance measurement was performed using a fisheye stereo camera with a general method called binocular stereo, but the accuracy was insufficient. Therefore, this research attempts to introduce motion stereo in addition to binocular stereo. Since these two measurement methods have different properties, they can be expected to reinforce each other's weaknesses by fusing them. In the motion stereo, three types of matching are introduced. In this method, three types of measurement results with different baseline directions can be obtained so that the accuracy of distance measurement can be expected to improve by fusing. In order to take advantage of each measurement method, we propose a fusion method called bilateral-like filter. This filter gives the continuity of the measurement values obtained from the dense measurement of binocular stereo and the weight that takes into account the position in the image where the measurement accuracy of each method is high.

Key words: fisheye stereo camera, equirectangular image, binocular stereo, motion stereo, bilateral-like filter

1. 序 論

近年、距離センサを用いた自動車の運転支援システムの研究開発が活発になっている¹⁾²⁾。その代表的なセンサとして、ステレオカメラ、レーザ、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の狭さや計測密度の低さといったデメリットが存在する。これに対し、本研究では魚眼カメラに着目する。魚眼カメラの画角は180°以上であり、そのサイズも比較的小さいことから車載に適している。従来研究として以下のような研究がある。Abrahamらは、魚眼ステレオカメラにステレオ平行化を適用することによって、ステレオマッチングを単純化した³⁾。Moreauらは、等立体角投影モデルの魚眼ステレオカメラを用いて、3次元環境復元を行う手法を構築した⁴⁾。Haneらは、plane-sweeping法を用いてリアルタイムでの3次元環境計測を実現した⁵⁾。これらの研究では、対応点探索を単純化するために、魚眼画像を透視投影画像に変換している。しかし、変換の際に画像の周辺部が引き伸ばされるため、周辺領域でのステレオマッチングが困難となってしまう。この問題に対し、大橋らは正距円筒画像への変換を用いた魚眼ステレオカメラを提案している⁶⁾。画像周辺部を引き伸ばさない画像変換で、魚眼画像の歪みの低減と対応点探索の単純化を実現してい

る。しかし、内部、外部パラメータの誤差の影響で誤マッチングや距離精度に課題を残している。そこで、本研究では同一フレームの画像を扱う2眼ステレオだけでなく、時系列画像を扱うモーションステレオ（Structure from Motion (SfM)）にも着目する。一般的に、2眼ステレオでは領域ベースマッチングが、モーションステレオでは特徴ベースマッチングが用いられる⁷⁾。領域ベースマッチングは画像全体で適用が可能なので密なデータが求まる一方、特徴ベースマッチングでは、画像中の特徴の強い点しか扱わないので疎なデータとなるが、誤マッチングの可能性は減るといったメリットも存在する⁸⁾。実際に飯田らは2眼ステレオにモーションステレオの融合を試み、計測密度の高さを維持したまま魚眼ステレオカメラの高精度化を実現した⁹⁾。本研究では上記の手法に計測手法の基線長の向きを考慮した重みを加えることで、魚眼ステレオカメラの更なる性能向上を目指す。図1に提案手法の流れを示す。まず環境中で魚眼ステレオカメラを光軸方向に直進させ、2フレームのステレオ画像を得る。取得した画像に対して、2眼ステレオ計測、モーションステレオ計測を各々行う。最後に、2つの計測結果を後述する擬似バイラテラルフィルタを用いて統合する。2章、3章で、それぞれ2眼ステレオ、モーションステレオに関して説明した後、4章で擬似バイラテラルフィルタによる統合を示す。5章で実験を行い、6章で結論を述べる。

2. 魚眼ステレオカメラの2眼ステレオ

2.1 魚眼カメラモデル

魚眼レンズの射影方式には等距離射影、立体角射影などがある。しかし、製造過程において発生する誤差などにより、実際の魚眼カメラは理想的な射影方式には従わない。そこで本論文では、Scaramuzzaらが提案した汎用的な全方位カメラモデル

* 原稿受付 令和2年4月12日

掲載決定 令和2年6月30日

** 中央大学大学院（現、パナソニック（株）、神奈川県横浜市都筑区池辺町4261）

*** 中央大学大学院（東京都文京区春日1-13-27）

† 中央大学（現、北陸先端科学技術大学院大学、石川県能美市旭台1-1）

†† 正 会 員 中央大学（東京都文京区春日1-13-27）

††† クラリオン（株）（埼玉県さいたま市中央区新都心7-2）

を用いる¹⁰⁾。チェッカーパターンを複数撮像した魚眼画像を用いてカメラの内部パラメータの算出を行い、魚眼画像の歪み補正に用いる。

2.2 正距円筒変換

本論文では、魚眼画像を図2に示す正距円筒変換を行ってからステレオマッチングを行う。正距円筒画像は横軸と縦軸を方位角 λ と仰角 ϕ にとった等間隔な座標に投影するため、引き伸ばしや歪みの少ない画像を生成できる。正距円筒画像変換の手順を以下に示す⁹⁾。解像度 $M \times N$ の正距円筒画像の注目位置 (u', v') を式(1)(2)により円筒モデル上の点 (λ, ϕ) に投影し、

$$\lambda = \frac{\pi}{M} \left(u' - \frac{M}{2} \right) \quad (1)$$

$$\phi = \frac{\pi}{M} \left(v' - \frac{N}{2} \right) \quad (2)$$

式(3)(4)により極座標上の点 (θ, α) へと変換する。

$$\theta = \arccos(\cos \phi \cos \lambda) \quad (3)$$

$$\alpha = \arctan \frac{\tan \phi}{\sin \lambda} \quad (4)$$

極座標上に変換した点 (θ, α) から魚眼画像への変換は式(5)(6)により行われる。ただし (u_0, v_0) は魚眼画像の画像中心とし ρ はカメラの内部パラメータ¹⁰⁾ $a_0 \sim a_4$ を用いて式(7)により計算される。

$$u = \rho \frac{\sin \lambda}{\sqrt{\tan^2 \phi + \sin^2 \lambda}} + u_0 \quad (5)$$

$$v = \rho \frac{\tan \phi}{\sqrt{\tan^2 \phi + \sin^2 \lambda}} + v_0 \quad (6)$$

$$a_0 + \left(a_1 + \frac{1}{\tan \theta} \right) \rho + a_2 \rho^2 + a_3 \rho^3 + a_4 \rho^4 = 0 \quad (7)$$

2.3 正距円筒画像のステレオマッチング

2眼で距離計測を行うとき、簡単化のため各々の光軸と基線長を垂直にすることが一般的である。これを平行ステレオという。実際には、各々の光軸と基線長は完全には垂直になっていないため、ステレオ平行化により、そのずれ分を計算し、補正

する。ステレオによって距離計測を行うには、左右の画像中から対応点を探索するステレオマッチングを行う必要がある。このとき、処理時間の短縮と誤マッチングの減少のためにエピポーラ線上のみを探索する。正距円筒画像のエピポーラ線は曲線になる。方位角 λ が0のときの仰角を ϕ_0 としたとき、正距円筒画像上のエピポーラ線の軌跡は次式のようにになる。

$$\phi = \arctan(\tan \phi_0 \cos \lambda) \quad (8)$$

よって、正距円筒画像においては、対応点探索を図3に示す曲線上で行う。このとき、リアルタイム性を重視し、領域ベースマッチングとしてブロックマッチングを適用し、その評価値としてSAD (Sum of Absolute Difference) を用いる。また、サブピクセル精度で視差を得るために等角直線フィッティングを行う¹¹⁾。

3. 魚眼ステレオカメラのモーションステレオ

3.1 特徴点抽出

時系列画像間での対応点探索では、画像の明るさや見えの変化が激しいので、領域ベースよりも特徴ベースのマッチングが適している。そこで、本研究では特徴点抽出アルゴリズムであるAKAZEによって特徴点を検出する¹²⁾。AKAZEは画像の回転、スケール、輝度の変化に頑健な特徴量の抽出が可能であり、移動するカメラから得られた画像に対して有効である。また、AKAZEは矩形のウィンドウ上で特徴検出器の極値を探索するため、魚眼画像と比較して画像中の位置による変化が少ない正距円筒画像を用いることが適している。

3.2 3種類のマッチング

特徴ベースマッチングでは、画像中の特徴の強い点しか利用しないため、疎なデータになる。本研究では画像全体で特徴点を得るために、図4に示すように、2フレームのステレオ画像(計4枚)の時刻 t における右画像を基準とした3種類のマッチングを行う。このマッチングにより、特徴ベースのデメリットである計測密度の低さをある程度補強することができる。

3.3 運動パラメータ

2画像間でのカメラの運動パラメータは、回転行列 R と並進ベクトル t で表される。まず、3.2節で得られた対応点群から5点アルゴリズムを用いて基本行列 E を計算する¹³⁾。このとき、RANSACにより外れ値に対してロバストに計算が行われる。続いて、 E を特異値分解することで R と t を求める。な

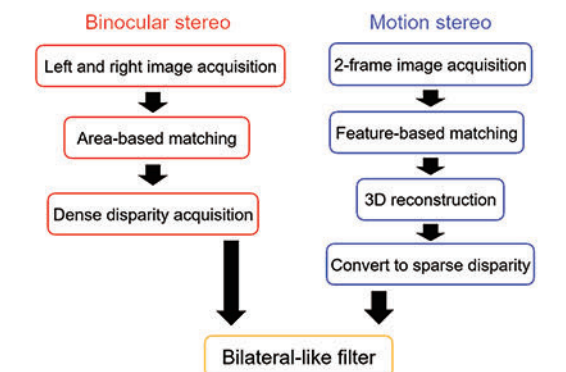


Fig.1 Flow chart of the proposed method

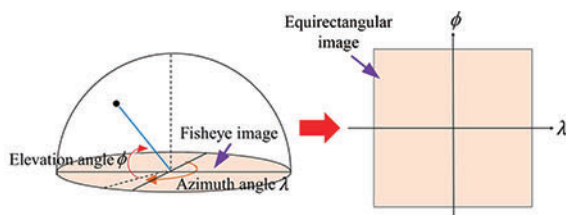


Fig.2 Transformation from fisheye image to equirectangular image

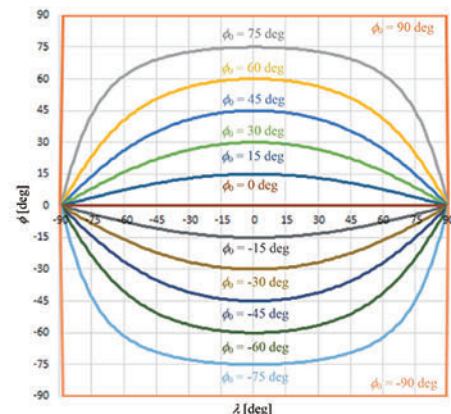


Fig.3 Epipolar line on equirectangular image

お, 3.1 節で得られる特徴点は正距円筒座標 (λ, ϕ) であるため,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \tan \lambda \\ \frac{\tan \phi}{\cos \lambda} \\ 1 \end{pmatrix} \quad (9)$$

により透視投影座標 (x, y, z) へ変換してから基本行列の計算を行う。

3.4 3次元位置の復元

特徴点の透視投影座標 (x, y) と 3 次元座標 (X, Y, Z) の関係はカメラの運動パラメータで構成される透視投影行列 P を用いて次式で表せる。

$$\tilde{x} \sim P\tilde{X} \quad (10)$$

ただし,

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (11)$$

このときの $\sim$ は同次座標として等しいことを表す。式 (11) にマッチングで得られた 2 組の透視投影座標と透視投影行列を代入すると, 以下のように整理される。

$$BX = b \quad (12)$$

ただし,

$$\begin{bmatrix} p_{31}x - p_{11} & p_{32}x - p_{12} & p_{33}x - p_{13} \\ p_{31}y - p_{21} & p_{32}y - p_{22} & p_{33}y - p_{23} \\ p'_{31}x' - p'_{11} & p'_{32}x' - p'_{12} & p'_{33}x' - p'_{13} \\ p'_{31}y' - p'_{21} & p'_{32}y' - p'_{22} & p'_{33}y' - p'_{23} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_{14} - p_{34}x \\ p_{24} - p_{34}y \\ p'_{14} - p'_{34}x' \\ p'_{24} - p'_{34}y' \end{bmatrix} \quad (13)$$

(x, y) と (x', y') は 2 枚の画像間でマッチングして得られた各々の特徴点の座標である。 p と p' は各々のカメラの透視投影行列 P の要素である。従って, 最適 3 次元座標 $\hat{X} = (\hat{X}, \hat{Y}, \hat{Z})$ は以下の最小 2 乗法によって求められる。

$$\hat{X} = B^+b \quad (14)$$

ただし, B^+ は擬似逆行列である。3.2 節で得られた対応点群は 3 組となるので, この計算を 3 回繰り返す。この 3 組は全て同じ画像 (時刻 t における右正距円筒画像) を基準に復元しているので, その結果を一緒に扱う。

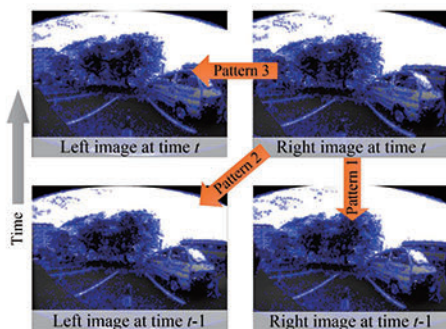


Fig.4 Three types of matching of stereo image pairs

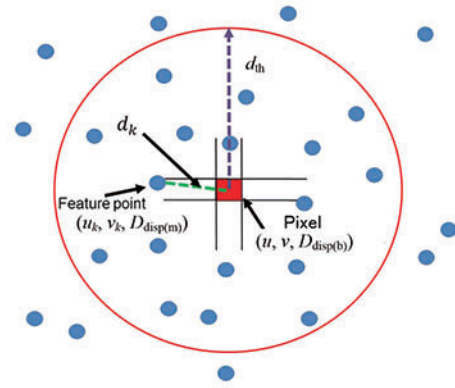


Fig.5 Calculating weighted average

以上により, 3 次元座標が得られるが, 画像のみを用いた 3 次元復元では実際のスケールは得られない。そこで, 本研究では路面平面推定によるスケール合わせを行う¹⁴⁾。

4. 擬似バイラテラルフィルタによる異なる計測手法の融合

3 章では密な視差, 4 章では特徴点ごとの視差が得られるので, この 2 つを上手く融合する手法を提案する。なお, 本手法はバイラテラルフィルタ¹⁵⁾に似ているため擬似バイラテラルフィルタと呼ぶ。

4.1 空間方向と視差方向の重み

注目画素 (u, v) とその周辺の特徴点 (u_k, v_k) までの距離を d_k , 2 眼ステレオで得た各画素の視差を $D_{disp(b)}$, モーションステレオで得た特徴点ごとの視差を $D_{disp(m)}$ とする。空間方向と視差方向に応じたガウス分布に従う重み w_{bk} を以下の式で定義する。

$$w_{bk} = \exp\left(-\frac{d_k^2}{2\sigma_1^2}\right) \exp\left[-\frac{(D_{disp(b)} - D_{disp(m)})^2}{2\sigma_2^2}\right] \quad (15)$$

$$d_k = \sqrt{(u_k - u)^2 + (v_k - v)^2} \quad (16)$$

ここで, σ_1 は空間方向の重み, σ_2 は視差方向の重みを表すガウス分布の標準偏差である。重み付き平均の概念図を図 5 に示す。注目画素からしきい値 d_{th} 以下の範囲に特徴点が存在しない場合, その画素の視差は不定とする。

4.2 基線長の向きを考慮した重み

本研究では, 図 4 に示した 3 種類のマッチングで得られた点群を同一に扱う。よって, マッチングの種類ごとに基線長の向きが異なることになる。基線長の向きが変わると, 画像位置によって距離精度が大きく変わる。そこで重み w_{pk} をマッチングの種類 Pattern 1~3 ごとに以下のように定義する。

• Pattern 1

$$w_{pk} = -\exp\left[-\frac{(u_k - \frac{width}{2})^2}{2\sigma_3^2}\right] + 1 \quad (17)$$

• Pattern 2

$$w_{pk} = -\exp\left[-\frac{(u_k - \frac{width}{2} - \alpha)^2}{2\sigma_3^2}\right] + 1 \quad (18)$$

• Pattern 3

$$w_{pk} = \exp\left[-\frac{(u_k - \frac{width}{2})^2}{2\sigma_3^2}\right] \quad (19)$$

ここで, $width$ は画像の横幅, σ_3 は空間方向の重みを表すガウ

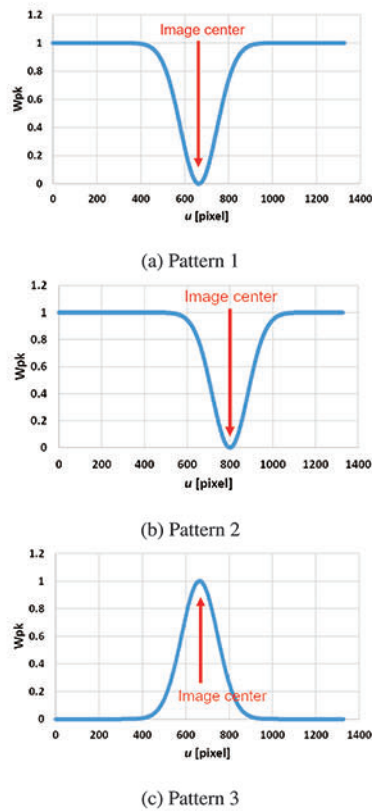


Fig.6 Weight considering the direction of the baseline

ス分布の標準偏差である． α は，図 4 中の基線長が斜めのときの光軸とのなす角度である．Pattern ごとの重み関数を図 6 に示す．

4.3 擬似バイラテラルフィルタ

特徴点を用いた補間により，各画素の視差を決定する．各画素の視差 $\hat{D}_{\text{disp}}$ は次式で計算される．

$$\hat{D}_{\text{disp}} = \frac{\sum_{k=1}^n w_{bk} w_{pk} D_{\text{disp}(m)k}}{\sum_{k=1}^n w_{bk} w_{pk}} \tag{20}$$

この処理により，計測精度の高い特徴ベースの計測結果を領域ベースの連続性を参照しつつ平均化することができる．

5. 距離精度評価

5.1 実験条件

提案手法の定量的評価を屋外環境で行った．カメラは FLIR 社の Flea3，魚眼レンズは SPACE 社の TV1634M を用いた．魚眼レンズの内部パラメータは MATLAB の OcamCalib Toolbox を利用して推定した¹⁰⁾．カメラの解像度は 1328×1048 pixel，基線長は 52 mm，画角は水平方向 165° ，垂直方向 132° である．モーションステレオは，カメラを光軸方向に 0.15 m 前進させたときの 2 フレームを用いた．この移動距離は，駐車時の平均時速 (10 km/h) と駐車支援システムに必要なフレームレート (20 fps) から導出した．この条件のとき，式 (18) における α は 18.43° になる．擬似バイラテラルフィルタで扱われる他の各パラメータは表 1 の通りである．実際に扱った魚眼ステレオカメラを図 7 に示す．実験環境と計測箇所を色分けした図を図 8 に示す．カメラを回転させて図 8(a) に示すように計測対象の画像中での位置を調整することで，図 8(b) に示す 5 箇所

Table1 Parameters for bilateral-like filter

Parameter	
Distance threshold d_{th} [pixel]	12
Standard deviation σ_1 [pixel]	6
Standard deviation σ_2 [pixel]	5
Standard deviation σ_3 [pixel]	83



Fig.7 Fisheye stereo camera.

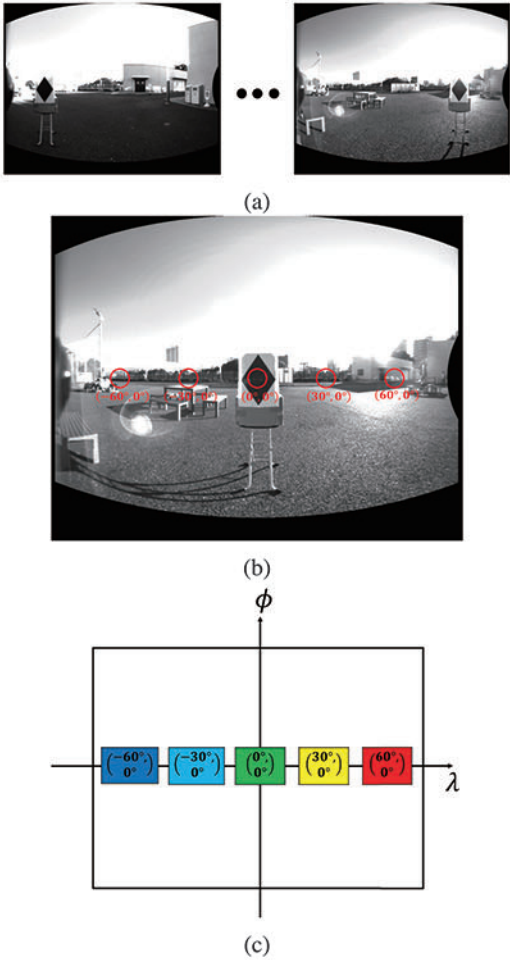


Fig.8 Experimental conditions: (a) captured images (b) outdoor environment and (c) measurement points

紙を計測対象とした．計測は，注目点と上下 1 点ずつの計 3 点の計測を 4 回繰り返して，得られた計 12 点の誤差の平均と標準偏差で評価した．

5.2 実験結果

図 9 に画像中心の計測距離 1 m における距離画像の比較を示す．提案手法では，誤マッチングの減少に加え，遠距離の距離精度の改善が見られる．これは，特徴点ベースのマッチングが領域ベースのマッチングより精度が高いこと，基線長の向きを考慮した重みを使用することにより各々が得意な画像中の位

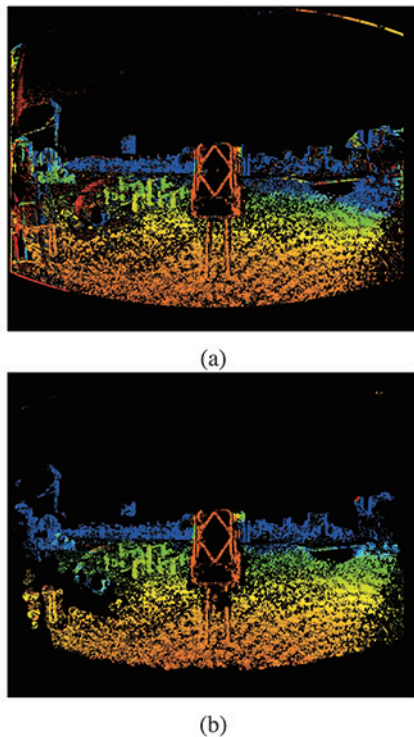


Fig.9 Comparison of range images: (a) without bilateral-like filter and (b) with bilateral-like filter. Color represents distance values from 0 m (red) to 10 m and above (blue)

置で重みを大きくしていること、モーションステレオの移動距離がステレオカメラの基線長の3倍ほどあることが寄与していると考えられる。図10~14に誤差の平均と標準偏差を示す。棒グラフの色は図8(c)の計測箇所に対応している。距離1 mでは、2眼ステレオ単体の方が距離精度が高かった。このことから、近距離においてはステレオカメラの基線長サイズが適しており、モーションステレオの移動距離は逆に大き過ぎると考えられる。すなわち、モーションステレオの場合、近距離においては移動前後で画像の見えが大きく変化するため、AKAZEを用いた特徴点のマッチングに悪影響が生じると考える。距離3~7 mでは、提案手法の精度が高かった。また、距離10 mにおいても提案手法の距離精度が全体的に高かったが、画像中心の精度は低かった。これは、距離10 mではステレオカメラの基線長サイズでは小さ過ぎるためと考える。距離3~10 mにおいてバイラテラルフィルタを適用しない場合には類似した偏りのある結果となった。これは、魚眼レンズの非等方性による誤差の影響を強く受けてしまったためであると考えられる。

提案手法の屋外環境での処理時間は490~730 msであった。最も処理時間を要しているのは、AKAZEにおける特徴記述である。本手法では、特徴の少ない路面で特徴点を抽出するためAKAZEのパラメータを変更し点数を多く取得できるようにしているが、路面でない特徴が顕著な箇所において多く点を取得してしまい処理時間が点数依存である特徴記述において多くの時間がかかってしまう。この処理時間は、運転支援アプリケーションとしては長過ぎるため改善が必要である。

6. 結 論

本論文では、魚眼ステレオカメラにおいて、2種類のステレオ計測結果を融合した距離計測を実現した。融合手法として、

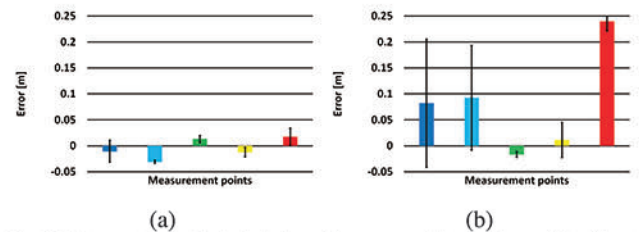


Fig.10 Mean and standard deviation of errors at distance 1 m: (a) without bilateral-like filter and (b) with bilateral-like filter

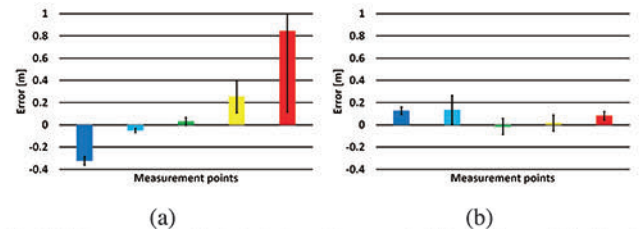


Fig.11 Mean and standard deviation of errors at distance 3 m: (a) without bilateral-like filter and (b) with bilateral-like filter

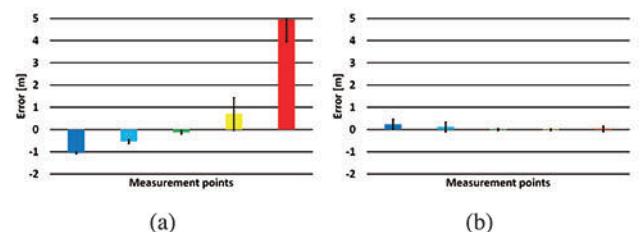


Fig.12 Mean and standard deviation of errors at distance 5 m: (a) without bilateral-like filter and (b) with bilateral-like filter

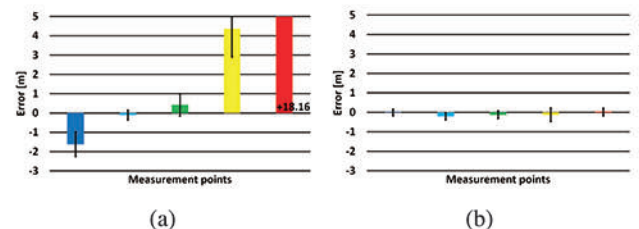


Fig.13 Mean and standard deviation of errors at distance 7 m: (a) without bilateral-like filter and (b) with bilateral-like filter

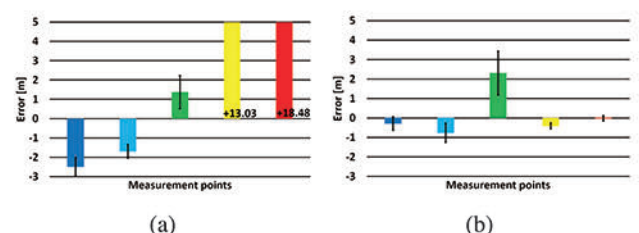


Fig.14 Mean and standard deviation of errors at distance 10 m: (a) without bilateral-like filter and (b) with bilateral-like filter

空間方向と視差方向に加え、基線長の向きを考慮する重み付き平均化を行う擬似バイラテラルフィルタを提案した。擬似バイラテラルフィルタにより、誤マッチングの減少と中~遠距離精度の改善が可能であることを示した。

今後の展望としては、計測距離に応じて手法の使い分けを行う予定である。また、処理時間の短縮を行い、運転支援向けの実用的なアプリケーションの開発を目指す。

謝 辞

本研究は JSPS 科研費 19H04191 の助成を受けたものである。

参 考 文 献

- 1) J. W. Perng, P. Y. Liu, K. Q. Zhong, and Y. W. Hsu, "Front object recognition system for vehicles based on sensor fusion using stereo vision and laser range finder," in *Proceedings of the 2017 IEEE International Conference on Consumer Electronics*, (2017) 261.
- 2) N. Sasaki, N. Iijima, and D. Uchiyama, "Development of ranging method for inter-vehicle distance using visible light communication and image processing," in *Proceedings of the 2015 15th International Conference on Control Automation and Systems*, (2015) 666.
- 3) S. Abraham and W. Forstner, "Fish-eye stereo calibration and epipolar rectification," *Journal of Photogrammetry and Remote Sensing*, **59**, 5, (2005) 278.
- 4) J. Moreau, S. Ambellouis, and Y. Ruichek, "Equisolid fisheye stereo-ovision calibration and point cloud computation," in *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (2013) 167.
- 5) C. Hane, L. Heng, G. H. Lee, A. Sizov, and M. Pelfeys, "Real-time direct dense matching on fisheye images using plane-sweeping stereo," in *Proceedings of the 2014 2nd International Conference on 3D Vision*, (2014) 57.
- 6) 大橋 明, 山野 史登, 増山 岳人, 梅田 和昇, 福田 大輔, 入江 耕太, 金子 修造, 村山 純哉, 内田 吉孝, "正距円筒画像への変換を用いた魚眼ステレオカメラの構築", 精密工学会誌, **83**, 12, (2017) 1095.
- 7) 後藤 真一, 川西 亮輔, 山下 淳, 金子 透, "二眼ステレオとモーションステレオを併用した全方位ステレオカメラによる 3 次元計測", ロボティクス・メカトロニクス講演会講演概要集, (2010) 1A2-E29.
- 8) N. Jayanthi and S. Indu, "Comparison of image matching techniques," *International Journal of Latest Trends in Engineering and Technology*, **7**, 3, (2016) 396.
- 9) 飯田 浩貴, 池 勇勲, 梅田 和昇, 大橋 明, 福田 大輔, 金子 修造, 村山 純哉, 内田 吉孝, "魚眼ステレオカメラの 2 眼ステレオとモーションステレオの融合による距離画像計測", 日本機械学会論文集, **85**, 875, (2019).
- 10) D. Scaramuzza, A. Martinelli, and R. Sliedright, "A toolbox for easily calibrating omnidirectional cameras," in *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (2006) 5695.
- 11) M. Shimizu and M. Okutomi, "Sub-pixel estimation error cancellation on area-based matching," *International Journal of Computer Vision*, **63**, 3, (2005) 207.
- 12) P. F. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," in *Proceedings of the British Machine Vision Conference*, (2013).
- 13) D. Nister, "An Efficient Solution to the Five-Point Relative Pose Problem," in *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**, 6, (2004) 756.
- 14) K. Yamaguchi, T. Kato, and Y. Ninomiya, "Moving Obstacle Detection using Monocular Vision," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, (2006).
- 15) C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the 1998 IEEE 6th International Conference on Computer Vision*, (1998) 839.