

魚眼ステレオカメラと単眼深度推定を用いた路面平面推定と障害物検出

Road Surface Estimation and Obstacle Detection
using Fisheye Stereo Camera and Monocular Depth Estimation

○学 筑後 光 (中央大学) Sarthak Pathak (中央大学)

正 梅田 和昇 (中央大学)

Hikaru CHIKUGO, Chuo University, chikugo@sensor.mech.chuo-u.ac.jp

Sarthak PATHAK, Chuo University

Kazunori UMEDA, Chuo University

In this study, we propose a method for road surface estimation and obstacle detection using a fisheye stereo camera. In obstacle detection using stereo cameras, obstacles are detected based on distance information obtained through stereo matching. However, due to processing load, this cannot be done quickly. And obstacle detection methods using deep learning cannot accommodate obstacles that are not included in the training data. Therefore, we first detect obstacles on the road surface using the relative depth obtained from monocular depth estimation. Then, by focusing only on the obstacle regions for distance measurement, we aim to quickly detect all obstacles. Experiments demonstrate the ability to detect only obstacles and the high processing speed.

Key Words: Fisheye Stereo Camera, Image Processing, Obstacle Detection, Autonomous Vehicle

1 序論

1.1 研究背景

近年、自動運転に関する技術の進歩は目覚ましく、すでに実用化に至っているケースも多数存在する [1, 2]。これらのシステムでは、障害物検知などを目的とした周囲の状況の把握が必要であり、主に距離情報が用いられる。距離情報を取得するための代表的なセンサとして、ステレオカメラや LiDAR、ソナーなどがある。しかし、従来のセンサでは計測範囲の狭さや距離計測の角度分解能の低さなどにより、細い障害物や真横にある障害物が検知しにくいことが問題となっている。また、複数のセンサや機構を組み合わせることにより計測範囲を増大させた例も存在するが、取り付け位置の拘束やメンテナンス性の悪化、コストの増大などが問題となっている。

そこで本研究では、魚眼ステレオカメラに注目する。魚眼カメラは視野角が約 180° と非常に広角なため、1台で広範囲に計測が可能となる。さらに、被写界深度が大きいため被写体が距離によってぼやけて撮像されることが少なく、画像認識において有利である。駐車時の自動運転システムもしくは移動ロボット用のカメラセンサを想定した際、これらの利点から広範囲計測可能な魚眼ステレオカメラは有用である。魚眼ステレオカメラを用いた3次元計測手法として、大橋らは魚眼画像から正距円筒画像へ変換することで魚眼画像の歪みを低減した3次元計測を提案している [3]。しかし、湾曲したエピポーラ線に沿ってテンプレートマッチングを行うことによる誤マッチングの増加によって距離計測精度が不十分である。一方、飯田らは、領域ベースのステレオ計測と特徴点ベースの Structure from motion (SfM) を融合する手法として擬似バイラテラルフィルタを提案している [4]。時系列画像を導入することで距離計測精度が大幅に向上したが、処理負荷のため高速に行えないといった課題を残している。そこで、新井らは魚眼カメラを縦に配置することでエピポーラ線を直線にし、ステレオマッチングを簡便に行うことで距離計測精度を向上させた魚眼縦ステレオカメラを提案した [5]。しかし、障害物検出を目的とした場合、全ての計測範囲において距離計測を行う必要はない。魚眼ステレオカメラを用いた障害物検出手法として、佐久田らは視差画像に対して複数平面をフィッティングすることにより路面形状の制約が少ない路面平面と障害物高さの推定手法を提案した [6]。しかし、路面推定を行う際、再帰的な処理を行っているため、高速に障害物を検出することが困難である。また、3次元計測において誤計測がある場合、路面を障害物として誤検出することがある。そこで、筑後らは輝度情報を考慮することで路面を誤検出せず、障害物のみを検出する手法を提案した [7, 8]。し

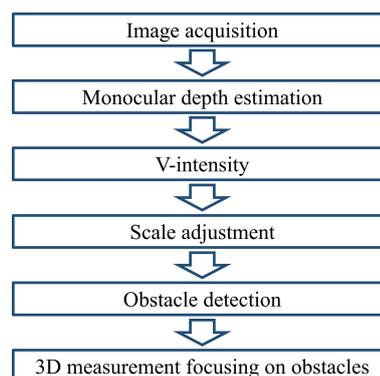


Fig.1 The flow of the proposed method

かし、3次元計測を基に推定された障害物領域以外の頻度値が高い輝度を用いて障害物を再検出しているため、白線や日差しがある部分などを障害物として誤検出することがある [7]。また、入力画像のエッジと3次元計測で得られた障害物領域が重なった領域の輝度を用いて障害物を再検出しているため、白線が誤検出されている場合、白線を障害物として誤検出してしまうことやテクスチャレスな障害物には対応出来ないといった課題がある [8]。そこで、単眼深度推定を用いることで路面上の障害物を検出し、その障害物だけに注目した3次元計測を行うことで高速かつテクスチャレスな環境に対応した障害物検出手法を目指す。

1.2 関連研究

周辺環境を把握するための手法として、推定された路面から障害物を検出する手法と色情報などから障害物を検出する手法の大きく2つに分けることが出来る。推定された路面から障害物を検出する手法として UV-disparity を用いる手法がある [9, 10]。しかし、これらの手法では視差情報を得るためにステレオマッチングを用いている。そのため、テクスチャレスな障害物がある場合、障害物を検出出来ないことがある。関らは射影行列を用いた手法を提案している [11]。しかし、この手法では単一の平面を仮定しているため、斜面の傾斜が変化する際に問題となる。色情報などから障害物を検出する手法として、深層学習を用いた手法がある [12, 13, 14]。しかし、深層学習は学習時の環境に依存するため、学習データとは違う環境に対して精度が保証されないという課題や説明可能性が低いという課題もある。

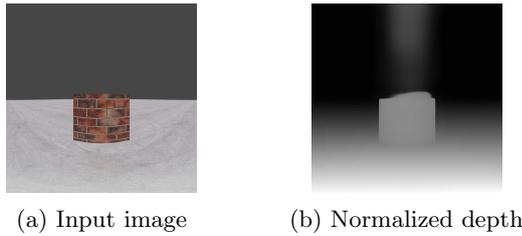


Fig.2 Results of MiDaS

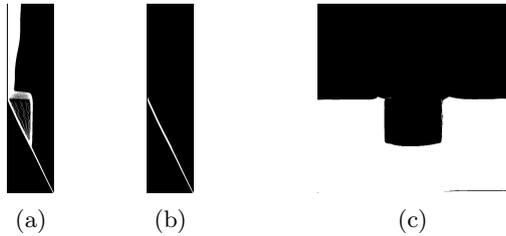


Fig.3 V-intensity.(a)V-intensity image (b)Road region of V-intensity image (c)Road region on input image

2 提案手法

2.1 概要

提案手法の流れを Fig. 1 に示す。最初に2つの魚眼カメラから魚眼画像を取得し、魚眼画像の歪が低減された正距円筒画像への変換を行う。そして、チェッカーボードを用いたステレオ平行化を行うことで正距円筒画像の補正を行う。次に単眼深度推定モデルの MiDaS [15] を用いて相対距離を取得する。取得した相対距離を用いて V-intensity 画像を生成することで、大まかに路面を推定する。ここで、V-intensity とは V-disparity [16] で用いる視差を輝度に置き換えて行う。その次に信頼度の高い点のみステレオマッチングを行い、その点と MiDaS とのスケール調整を行う。最後に路面上の障害物を検出し、その障害物領域のみに注目した距離画像計測を行う。

2.2 画像の取得

魚眼カメラで得られた魚眼画像には特有の歪があるため、この歪を低減した正距円筒画像への変換を行う [3]。また、2台のカメラで距離画像計測を行う際、単純化のため光軸と基線を垂直にすることが一般的である。実際には、カメラ取付時によるずれなどによって完全に垂直にはなっていない。そのため、チェッカーボードを用いたステレオ平行化により補正を行う [5]。

2.3 単眼深度推定

従来のステレオマッチングで得られた視差を基に V-disparity を用いる際、誤計測がある場合やテクスチャレスな環境の場合に路面や障害物を正確に推定または検出することが出来ない。また、不必要な部分を計測しているため、処理速度が遅くなる。そこで、ある程度正確に環境把握を行うことが出来る単眼深度推定モデルの MiDaS を用いる。入力画像である Fig. 2(a) に MiDaS を用いた結果を Fig. 2(b) に示す。MiDaS は画像全体に対して処理を行っているが、従来のステレオマッチングよりも高速に行える可能性が高いことや小さなノイズが少ないという利点がある。

2.4 V-intensity

MiDaS を用いて取得した相対距離は 0 から 255 に正規化しているため、画像上では相対距離を輝度として扱うことが出来る。そこで、相対距離画像における上下方向の輝度の頻度値を表した画像である Fig. 3(a) に示すような V-intensity 画像を生成する。ここで、Fig. 3(a) の斜線領域は路面、縦方向の領域は障害物もしくは空や計測地点から遠距離の建物などの背景領域である。路面領域を抽出するために斜線領域の底部に注目し、底部からしきい値以内の領域を路面領域とする。しきい値以上の場合、近傍の路面領域と同等の大きさとする。この処理によって得られた路面領域を Fig. 3(b) に示す。また、抽出された路面領域を入力画像に投影すると Fig. 3(c) のようになる。

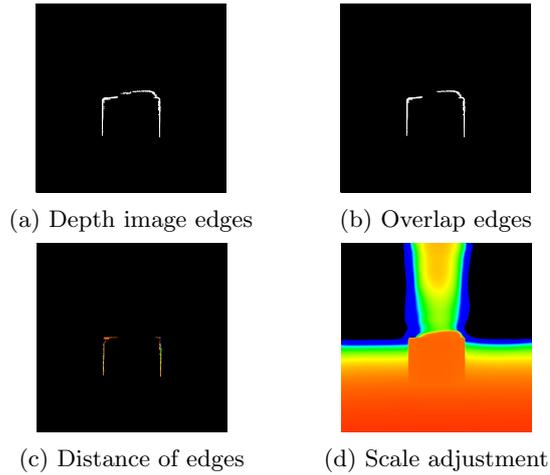


Fig.4 Scale adjustment between stereo and MiDaS

2.5 スケール調整

MiDaS で出力される距離は相対距離であるため、カメラから障害物までの絶対距離を計測することが出来ない。そこで、エッジに注目し、ステレオマッチングによって得られた信頼度の高い絶対距離と MiDaS で得られた相対距離のスケールを調整する。画像全体に対してステレオマッチングを行った場合、テクスチャレスな領域におけるノイズの影響によって正確にスケールを求めることが出来ない可能性がある。また、画像全体に処理を行うことで処理速度が遅くなる。そのため、相対距離画像と入力画像のエッジが重なる点のみにステレオマッチングを行う。障害物がある場合、相対距離画像上にエッジがあると考えられる。また、ステレオマッチングを行う際、エッジの近傍であれば信頼度の高い距離を得ることが出来ると考えられる。そこで、相対距離画像のエッジを用いる。なお、スケール調整に用いる相対距離画像のエッジは推定された路面に接しているエッジのみとする。

路面に接している相対距離画像のエッジ画像とステレオマッチングに用いるエッジ画像をそれぞれ Fig. 4(a), (b) に示す。ステレオマッチングを行う際、ブロックマッチングを用いる。また、テクスチャフィルタを用いたノイズ除去と誤マッチング処理を行う。誤マッチング処理では、2つのカメラから得られたピクセル間の視差を用いる。相対距離画像と入力画像のエッジの重なる点がしきい値以下であれば相対距離画像のエッジを用いてステレオマッチングを行う。エッジに注目したステレオマッチングを行った結果を Fig. 4(c) に示す。MiDaS とのスケールを求める際、ステレオマッチングを行った各画素のスケールの平均とする。この処理によって得られたスケールを MiDaS の相対距離に掛けることで絶対距離への変換を行う。スケール調整の結果を Fig. 4(d) に示す。ここで、Fig. 4(c), (d) は距離が近いほど赤く、遠いほど青いことを表している。

2.6 障害物検出

障害物領域を推定するために Fig. 3(c) と Fig. 4(b) を基に行う。Fig. 4(b) の点数がしきい値以上の場合、エッジから路面までの領域を障害物領域とする。点数がしきい値よりも少ない場合、Fig. 4(a) と Fig. 3(c) において同様の処理を行う。また、Fig. 4(a) の点数もしきい値より少ない場合、Fig. 4(d) の色付けされている領域内の Fig. 3(c) の路面ではない領域を障害物領域とする。この処理によって検出した障害物領域を Fig. 5(a) に示す。しかし、Fig. 5(a) 右下において路面を障害物領域として検出している。そこで、検出された各障害物領域に対して近傍の路面パラメータを用いたノイズ除去を行う。各障害物領域の近傍の路面パラメータはスケール調整後の MiDaS の距離から最小 2 乗法を用いて求める。最小 2 乗法によって得られた路面パラメータを用いることで各障害物領域近傍の路面平面を推定し、路面平面からしきい値以上離れている領域は障害物とし、それ以外の領域は路面としてその領域を障害物領域から除去する。各障害物領域近傍の路面を推定することで、傾斜がある場合においても対応することが出来ると考えられる。この処理を行った結果を Fig. 5(b) に示す。

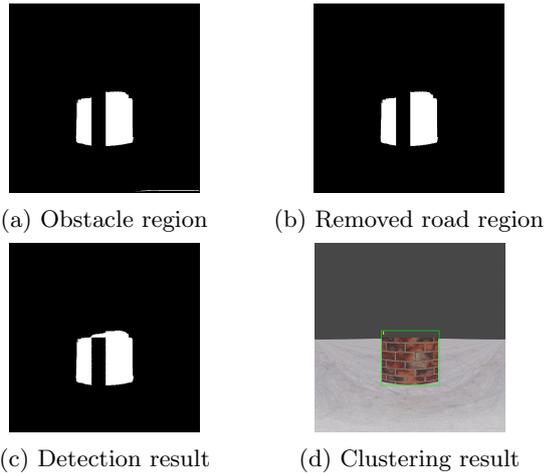


Fig.5 Obstacle detection

す。Fig. 5(b) 右下において障害物領域として検出されてしまっていた路面が除去されていることが分かる。ノイズ除去を行った後、障害物領域近傍のスケール調整後の MiDaS の距離と入力画像の輝度に注目し、類似していればその点を障害物領域に追加する。この処理の結果を Fig. 5(d) に示す。最後に 2 値画像における連結性を考慮した障害物領域のクラスタリングを行う。

2.7 障害物に注目した 3 次元計測

クラスタリングされた各障害物に対して、カメラからの距離と障害物高さを推定する。カメラからの距離を求める際、ステレオマッチングで得られた距離と MiDaS で得られた距離を比較する。そして、カメラからの距離と路面パラメータを用いて高さを推定する。ステレオマッチングで得られた距離を求める際、Fig. 4(c) を用いる。Fig. 4(c) において求めた距離が近い順から 5 つ選び、その 5 つの距離の平均をステレオマッチングで得られたカメラからの距離とする。MiDaS で得られた距離を求める際、Fig. 4(b), (d) を用いる。Fig. 4(b) で得られたエッジ部分の距離が近い順から 5 つ選び、その 5 つの平均を MiDaS で得られた距離とする。2 つの距離を比較する際、2 つの距離の差がしきい値以下である場合、両方法の距離の信頼度は高いと考えられるため、安全性を考慮し、距離が近い方を選ぶ。障害物回避を考慮した場合、近い方の距離とすることにより迅速に対応出来ると考えられる。2 つの距離の差がしきい値より大きい場合、どちらかの信頼度が低いと考えられる。ステレオマッチングで得られたカメラからの距離を求める際、5 点の平均を求めているため、誤マッチングによってノイズがある場合、MiDaS で得られたカメラからの距離との差が大きくなる可能性がある。MiDaS で得られたカメラからの距離を求める際、Fig. 4(b) のエッジ部分を用いていることや小さなノイズがないため、ステレオマッチングにおいて誤マッチング点を含んでいる場合より信頼度は高いと考えられる。そこで、2 つの距離がしきい値より大きい際、ステレオマッチングで得られたカメラからの距離が MiDaS で得られたカメラからの距離よりも小さい場合、カメラからの距離は MiDaS で得られた距離とする。

高さを推定する際、Fig. 4(b) で得られたエッジの上部の点の高さをそれぞれ求め、それらの平均を障害物の高さとする。高さを推定する際の路面平面の方程式が

$$ax + by + c = z \quad (1)$$

で与えられている場合、点 (x_0, y_0, z_0) の高さ H は

$$H = \frac{|ax_0 + by_0 + c - z_0|}{\sqrt{a^2 + b^2 + 1}} \quad (2)$$

を用いて求めることが出来る。ここで、 a, b, c はそれぞれ路面パラメータを表している。これらの路面パラメータは各障害物領域近傍の距離から最小 2 乗法を用いて求めている。

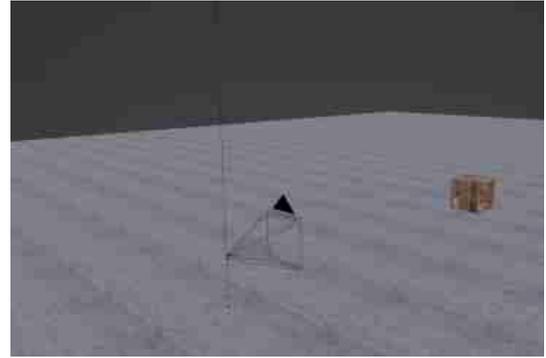


Fig.6 Experimental appearance

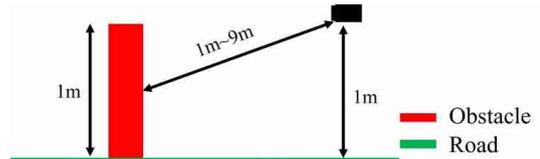


Fig.7 Experimental condition

3 精度評価実験

3.1 実験条件

本実験では 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、路面を誤検出することなく障害物を検出可能か確認した。また、カメラから障害物までの距離の誤差と障害物高さの誤差、処理速度を評価する。Blender においてレンダエンジンは Cycles を用いており、レンズをパノラマ状の正距円筒筒とした。カメラの解像度を 1048×1048 pixel とし、水平方向の画角を 180° 、垂直方向の画角を 180° 、カメラ間の基線長を 0.072m とした。また、Blender で得られた画像は理想的なものであるため、カメラの内部パラメータと外部パラメータには誤差がないものとする。実験環境を Fig. 6, 7 に示す。Fig. 6 は実験環境の外観を示しており、Fig. 7 は実験条件を示している。画像中の仰角を 0° で一定とし方位角 $-60^\circ, 0^\circ, 60^\circ$ の 3 か所を計測位置として正方形ブロックを置いた。ブロックの一辺を 1m とし、カメラからブロックまでの距離を $1\text{m}, 3\text{m}, 5\text{m}, 7\text{m}, 9\text{m}$ の 5 箇所計測した。

3.2 実験結果

実験結果を Table 1-4 に示す。Table 1-3 は方位角 $-60^\circ, 0^\circ, 60^\circ$ のそれぞれの位置で計測されたカメラから障害物までの距離と障害物高さの誤差を示している。また、Table 4 はそれぞれの処理における処理速度の平均と全体における処理速度の平均をそれぞれ示している。実験の結果より、全ての状況において路面の誤検出なく、計測対象物であるブロックを検出することが出来た。また、カメラから障害物までの距離の誤差は最小が 0m 、最大が 5.90m であった。カメラからの距離 1m 方位角 0° や 9m 方位角 0° の状況では計測誤差が 0m と正確に行えていることが分かる。しかし、一部の状況では誤差が 1.00m より大きくなるなどロバストに行えていないことが分かる。これは 2.7 節におけるカメラからの距離の求め方に問題があると考えられる。カメラからの距離を求める際、ステレオマッチングで得られた最も近い 5 つの距離の平均を求めている。そのため、1 点でも誤差の大きい距離がある場合、最終的な誤差は大きくなる。また、MiDaS は学習時に正距円筒筒画像ではなく透視投影画像を用いているため、カメラからの距離が MiDaS で得られた距離となった場合、誤差が大きくなる可能性があると考えられる。

障害物高さの誤差は最小が 0.11m 、最大が -0.90m であった。一部の状況では精度良く行えているが、ほぼ全ての状況において誤差が大きいことが分かる。これは障害物近傍の路面パラメータを最小 2 乗法によって求める際、その障害物の画像下側の距離を用いているためであると考えられる。Fig. 8 のような検出結果の場合、路面パラメータを求める際に障害物領域の距離も含まれてしまい、正確に路面を推定することが出来ない。そのため、

Table 1 Distance and height errors at 0° azimuth

	Error of distance [m]	Error of height [m]
1m	0.00	-0.19
3m	-0.02	-0.45
5m	-0.20	-0.19
7m	1.01	-0.78
9m	0.00	0.11

Table 2 Distance and height errors at -60° azimuth

	Error of distance [m]	Error of height [m]
1m	0.09	-0.45
3m	-0.48	-0.29
5m	-0.20	-0.45
7m	-1.00	-0.76
9m	-0.99	-0.90

ほぼ全ての状況における高さ推定結果が実際よりも低いと推定されていると考えられる。

Table 4 より、処理速度は全体で 2.35fps と自動運転を目的とした場合、十分な処理速度ではない。最も処理速度が遅い処理はスケール調整であり、その中でもステレオマッチングが要因であると考えられる。

4 結論

本論文では、魚眼ステレオカメラと単眼深度推定モデルの MiDaS を用いた路面上にある障害物に注目した障害物検出手法を提案した。実験によって、路面を誤検出することなく障害物を検出可能なことを示した。しかし、計測精度や処理速度において十分な精度を得ることが出来なかった。

今後の展望として、テクスチャレスな環境や実環境での実験を行う。また、従来手法との比較実験を行う。さらに、ステレオマッチングを行う際、特徴点を用いることで計測精度と処理速度の向上を行う。

参考文献

- [1] Jau-Woei Perng, Pei-Yu Liu, Kai-Quan Zhon and Ya-Wen Hsu, "Front object recognition system for vehicles based on sensor fusion using stereo vision and laser range finder," *2017 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*, pp.261–262, 2017.
- [2] Nobuo Sasaki, Naoyuki Iijima and Daiki Uchiyama, "Development of ranging method for inter-vehicle distance using visible light communication and image processing," *2015 15th International Conference on Control, Automation and Systems (ICCAS)*, pp.666–670, 2015.
- [3] 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," *Stereo rectification for equirectangular images*, pp.535–540, 2017.
- [4] Hirofuka Iida, Yonghoon Ji, Kazunori Umeda, Akira Ohashi, Daisuke Fukuda, Shuzo Kaneko, Junya Murayama and Yoshitaka Uchida, "High-accuracy Range Image Generation by Fusing Binocular and Motion Stereo Using Fisheye Stereo Camera," *2020 IEEE/SICE International Symposium on System Integration (SII)*, pp.343–348, 2020.
- [5] 新井健斗, 筑後光, Sarthak Pathak, 梅田和昇, "魚眼縦ステレオ法による高精度魚眼ステレオカメラの構築", 動的画像処理実用化ワークショップ講演論文集, pp.284–288, 2023.
- [6] Tomoyu Sakuda, Hikaru Chikugo, Kento Arai, Sarthak Pathak and Kazunori Umeda, "Estimation of Road Surface Plane and Object Height Focusing on the Division Scale in Disparity Image Using Fisheye Stereo Camera," *Journal of Robotics and Mechatronics*, vol.35-5, pp.1354–1365, 2023.
- [7] 筑後光, 新井健斗, Sarthak Pathak, 梅田和昇, "魚眼ステレオカメラを用いた視差情報と輝度情報による障害物の検出と高さ推定", ロボティクス・メカトロニクス講演会講演概要集 2023, pp.2A1–D07, 2023.

Table 3 Distance and height errors at 60° azimuth

	Error of distance [m]	Error of height [m]
1m	0.62	0.63
3m	-0.03	-0.44
5m	5.90	-0.55
7m	1.01	-0.84
9m	-0.99	-0.88

Table 4 Processing speed

	Processing speed [fps]
MiDaS	7.20
V-intensity	10.2
Scale adjustment	5.79
Obstacle detection	92.1
3D measurement	193
All	2.35

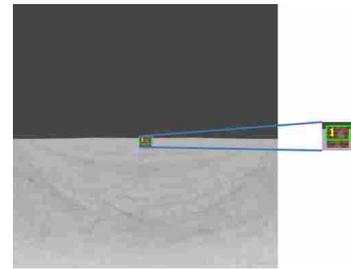


Fig.8 Detection result at 7m and 0°

- [8] Hikaru Chikugo, Tomoyu Sakuda, Sarthak Pathak and Kazunori Umeda, "Obstacle Detection and Height Estimation Using Fisheye Stereo Camera Considering Intensity Information," *2024 IEEE/SICE International Symposium on System Integration (SII)*, pp.147–152, 2024.
- [9] Mingguo Liu, Chenxing Shan, Haofeng Zhang and Qingyuan Xia, "Stereo Vision Based Road Free Space Detection," *2016 9th International Symposium on Computational Intelligence and Design (ISCID)*, pp.272–276, 2016.
- [10] Wenjie Song, Yi Yang, Mengyin Fu, Fan Qiu and Meiling Wang, "Real-Time Obstacles Detection and Status Classification for Collision Warning in a Vehicle Active Safety System," *IEEE Transactions on Intelligent Transportation Systems*, vol.19-3, pp.758–773, 2018.
- [11] Mingguo Liu, Chenxing Shan, Haofeng Zhang and Qingyuan Xia, "Robust obstacle detection in general road environment based on road extraction and pose estimation," *2006 IEEE Intelligent Vehicles Symp*, pp.437–444, 2006.
- [12] Peiliang Li, Xiaozhi Chen and Shaojie Shen, "Stereo R-CNN based 3D Object Detection for Autonomous Driving," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.7636–7644, 2019.
- [13] Buyu Li, Wanli Ouyang, Lu Sheng, Xingyu Zeng and Xiaogang Wang, "GS3D: An Efficient 3D Object Detection Framework for Autonomous Driving," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1019–1028, 2019.
- [14] Changxin Zhou, Yazhou Liu, Quansen Sun and Pongsak Lasang, "Vehicle Detection and Disparity Estimation Using Blended Stereo Images," *IEEE Transactions on Intelligent Vehicles*, vol.6-4, pp.690–698, 2021.
- [15] Reiner Birkel and Diana Wofk and Matthias Müller, "MiDaS v3.1 – A Model Zoo for Robust Monocular Relative Depth Estimation," arXiv preprint arXiv:2307.14460, 2023.
- [16] Raphael Labayrade, Didier Aubert, and Jean-Philippe Tarel, "Real time obstacle detection in stereo vision on non flat road geometry through "v-disparity" representation," *Intelligent Vehicle Symposium, 2002. IEEE*, vol.2, pp.646–651, 2002.