

## ステレオ画像から推定した 3 次元情報と属性情報に基づく障害物地図の生成 2D Drivable Grid Map Based on 3D Points and 2D Semantic Labels from Stereo Images

佐野 雄磨<sup>†</sup>  
Yuma Sano

渡辺 友樹<sup>†</sup>  
Tomoki Watanabe

### 1. はじめに

自動車の安全運転支援システムや自動運転システムは、周囲の環境を正確に認識する必要がある。周囲環境を認識する方法の 1 つとして障害物地図があり、周囲環境を格子状に分割し、各分割領域において障害物が存在する確率を算出することで、障害物の存在を認識することが出来る。従来は、主に LiDAR やステレオカメラなどのセンサを使用して計測した 3 次元情報に基づいて障害物地図が生成されてきたが、入力情報が 3 次元情報だけでは立体物以外の走行不可能な領域(歩道など)と走行可能な領域(車道)を識別することが出来ないという問題があった。そこで、本稿では、ステレオ画像から推定した属性情報(立体物、車道、歩道など)と 3 次元情報を統合して障害物地図を生成する方法を提案する。

### 2. 従来技術

障害物地図[1][2]は、周囲環境を格子状に分割し、各分割領域において障害物が存在する確率を算出することで、障害物の存在を認識する。従来は、主に LiDAR やステレオカメラなどのセンサを使用して計測した 3 次元情報に基づいて障害物地図が生成されてきた。そのため、従来の障害物地図は走行の支障となる立体物が計測された領域については、障害物が存在して走行不可能であることを示すことが出来るが、立体物が存在しない領域については、歩道などの車が走行してはいけない領域であっても、障害物の存在確率を低く算出し走行可能と示してしまう問題点がある。歩道や芝生などの領域は、立体物が存在しなくても障害物地図上で走行不可能であると示されるべきである。この問題への対策として、セマンティックセグメンテーションによる属性情報を使用する方法がある。[3]はカメラ画像に対するセマンティックセグメンテーションのラベルが示す属性情報を、平面であると仮定した地面に投影することで、自車周辺の分割領域の属性情報を認識している。しかしながら、地面が平面でない場合は、障害物地図の位置の正確さが低下するという問題がある。[4]や[5]は、幾何情報と属性情報の両方を使用して 3 次元障害物地図を生成している。[4]と[5]は、歩道や芝生といった走行してはいけない平面領域を走行不可能と認識することは可能だが、3 次元形状を再構成するため計算コストが高い、という課題がある。高度な安全運転システムや自動運転の実現へ向けでは、低い計算コストで正確な周囲環境認識が望まれる。そこで、本稿では、3 次元点群と 2 次元属性情報を統合した 2 次元障害物地図を生成する方式について提案する。カメラ画像に対する属性情報を 3 次元点群に対応付けて、属性を持った 3 次元点群を水平面に投影することで、3 次元形状を再構成することなく、幾何情報と属性情報を反映した 2 次元障害物地図を生成する。

<sup>†</sup>株式会社 東芝 研究開発センター Corporate Research and Development Center, Toshiba

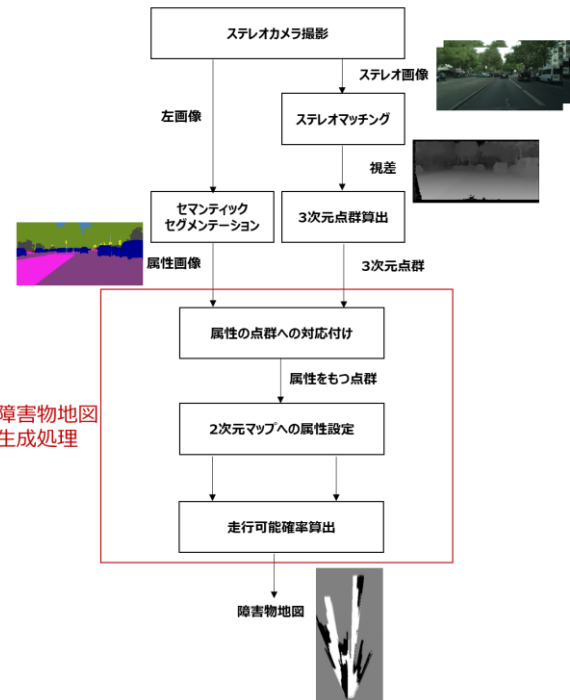


図 1 障害物地図生成の概要

### 3. 障害物地図の生成

#### 3.1 概要

図 1 に提案する障害物地図の概要を示す。提案方式では、障害物地図は格子状に分割された各領域において、車が走行可能か否かを示す確率(走行可能確率)を算出する。走行可能な領域は高い確率を示し、走行不可能な領域は低い確率を示す。提案方式は、3 次元点群情報と画素ごとに対象物の属性を持つ属性画像から 2 次元の障害物地図を生成する。本稿では、ステレオ画像から推定された視差をもとに算出した 3 次元点群と一方のカメラ画像に対するセマンティックセグメンテーションによって推定された属性画像を入力情報として障害物地図を生成する。

#### 3.2 点群への属性対応付け

まず、入力情報である 3 次元点群に対して属性を対応付ける。数式(1)に基づいて、3 次元点群を属性画像に投影し、3 次元点群が投影された画素の属性を当該点群に割り当てる。数式(1)において、 $\mathbf{X}=[X, Y, Z]^T$  は 3 次元点群の座標を示し、 $\mathbf{K}$  はカメラの内部パラメータを表す行列であり、 $\tilde{\mathbf{x}}$  は同次表現で表された属性画像上での画素位置を示す。

$$\tilde{\mathbf{x}} = \mathbf{KX} \quad (1)$$

なお、本稿では属性情報を 1. 立体物、2. 走行可能領域、3. 走行不可能領域、4. その他の 4 種類の属性に分類する。立体物は、車両や建物や電柱など、車両の走行の障害となる対象物を示す。走行可能領域は、平面な領域でかつ規則的にも走行してよい領域を示す。具体的には車道を示す。走行不可能領域は、平面な領域だが規則的に走行してはならない領域を示す。具体的には歩道や芝生などを示す。その他は、立体物でも平面領域でもない属性であり、例えば空などを示す。なお、属性がその他を示す点群は以降の障害物地図生成では使用されない。

### 3.3.2 次元マップへの属性設定

自車を中心とする仮想的な水平面を小さな領域(分割領域)に分割した 2 次元マップを生成し、各分割領域に対して属性を設定する。そのために、属性が対応付けられた 3 次元点群を 2 次元マップ上に投影し、投影された位置に相当する分割領域に割り当てる。次に、2 次元マップの分割領域に割り当てられた点群が持つ属性を、当該領域の属性として設定する。なお、1 つの分割領域に異なる属性を持つ点群が割り当てられている場合は(立体物 > 走行不可能領域 > 走行可能領域)の優先度に従い、優先度が最も高い点群の属性を当該領域の属性として設定する。図 2、は、カメラ画像の画素に対応する点群の分割領域への割り当てを示している。図 2 では、画素に対応する点群 8 つのうち 5 点は属性が「走行可能領域」であり 3 点は属性が「立体物」である。分割領域 1~3 に割り当てられた点は全て属性が「走行可能領域」であるため、分割領域 1~3 の属性を「走行可能領域」に設定する。一方で、分割領域 4 には属性が「立体物」の点と「走行可能領域」の点とが混在するため、前述の優先度に従い分割領域 4 の属性を「立体物」に設定する。

次に、点群が割り当てられず属性が設定されていない領域(未設定領域)へ属性を補間する。自車に対して未対応領域が位置する方向に存在する分割領域に対して、未対応領域から自車に近い方向と自車から遠い方向の 2 方向について、属性が設定された最も近い領域(設定領域)を探索する。2 方向で設定領域が存在し、両設定領域の属性が立体物である場合は、未設定領域の属性を立体物と設定し補間する。2 方向で設定領域が存在し、両設定領域の属性が平面を示す属性(走行可能領域あるいは走行不可能領域)である場合は、未設定領域からより近い方の設定領域の属性を未設定領域の属性に設定し補間する。これらは、立体物と立体物の間は立体物である可能性が高く、平面と平面の間は平面である可能性が高い、という前提に基づいている。なお、2 方向で設定領域が存在するが、一方の設定領域の属性が平面を示す属性で、もう一方の設定領域が立体物の場合は、未設定領域の属性は補間しない。また、いずれかの方向で設定領域が存在しない場合には未設定領域の属性は補間しない。図 3、4 に未設定領域の属性補間の例を示す。図 3 では、分割領域 6、8~10、12 が未設定領域となり、そのうち 6、8~10 は補間により属性が「走行可能領域」と設定される。一方で、図 3 の分割領域 12 は自車から遠い方向に設定領域を見つけられないため属性は補間されない。図 4 でも、図 3 と同様に分割領域 6、8~10 の属性が補間されるが、図 4 の分割領域 10 は「走行不可能領域」を示

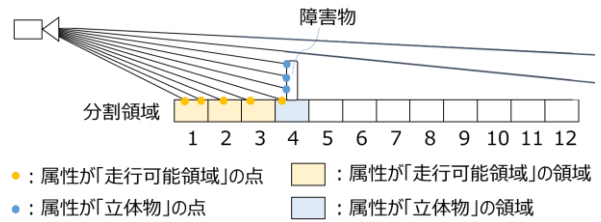


図 2 分割領域への点群割り当て

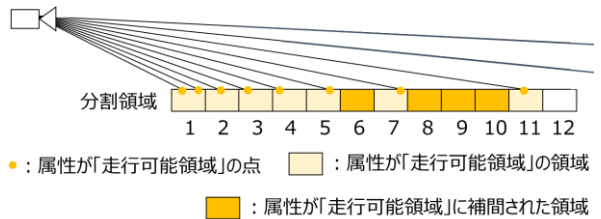


図 3 未設定領域の属性補間(1)

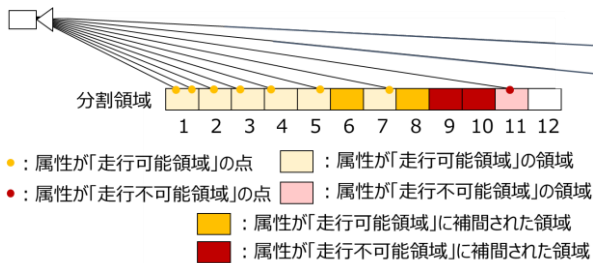


図 4 未設定領域の属性補間(2)

す設定領域 11 の方が近いため、属性は「走行不可能領域」が設定される。図 4 の分割領域 9 は同距離に「走行可能領域」を示す設定領域 7 と「走行不可能領域」を示す設定領域 11 が存在するが、優先度の高い「走行不可能領域」が設定される。なお、未設定領域に対する設定領域の探索は、予め設定する閾値よりも近くに存在する分割領域のみを探索する。

### 3.4 走行可能確率の算出

各分割領域に対して、設定された属性に基づいて、数式(2)に従って走行可能確率を算出する。属性が立体物や走行不可能領域である領域は低い走行可能確率を設定し、属性が走行可能領域である領域は高い走行可能確率を設定し、属性を持たない領域は中間値を設定する。

$$P = \begin{cases} 0 & \text{属性 = 立体物、走行不可能領域} \\ 0.5 & \text{属性 = なし} \\ 1 & \text{属性 = 走行可能領域} \end{cases} \quad (2)$$

さらに、各分割領域に走行可能確率を設定した 2 次元マップは、数式(3)に基づいて、走行可能確率を過去の走行可能確率と統合され、最終的に障害物地図として出力される。

$$\frac{P_t'(x,y)}{1-P_t'(x,y)} = \frac{P_t(x,y)}{1-P_t(x,y)} \cdot \frac{P_{t-1}'(x,y)}{1-P_{t-1}'(x,y)} \quad (3)$$

なお、数式(3)で  $P_t(x,y)$  は位置  $(x,y)$  の分割領域の時刻  $t$  の走行可能確率で、 $P_t'(x,y)$  は  $P_t(x,y)$  と過去の確率を統合した

走行可能確率を示す。時系列の走行可能確率の統合によって、属性情報や点群情報の突発的なエラーの影響を受けにくいロバストな障害物地図が生成される。

## 4. 性能評価

### 4.1 周囲環境の認識性能

提案方式によって生成した障害物地図を評価するため、公開データセットである *cityscapes dataset*[10] を使用して障害物地図を生成した。 *cityscapes* データセットは、ステレオカメラ画像と視差画像と属性の教示画像と車の位置情報を公開している。提案方式では、障害物地図に入力する 3 次元情報として、 *cityscapes* データセットが公開している視差画像から算出した 3 次元点群情報を使用した。得られた点群情報のうち、カメラから 1m 以上高い位置に存在する点群は車の走行に支障はないとして障害物地図生成の入力情報から除外した。さらに、2 次元の属性情報は、既存のセマンティックセグメンテーション方式である PSPNet[6] を用いて生成した。PSPNet によって推定される 19 種類の属性情報を図 5 に示す。図 5 の属性情報に対して、提案方式の障害物地図では、属性「road」を走行可能領域と設定し、属性「sidewalk」と「terrain」を走行不可能領域と設定し、その他の属性のうち「sky」、「traffic light」、「traffic sign」を除く属性を立体物として設定した。なお、障害物地図の時系列統合に際しては、 *cityscapes* データセットが公開する自車両の位置情報を利用した。障害物地図のサイズは  $80\text{m} \times 20\text{m}$  で、カメラの前方 6m から 86m の範囲で、カメラの左右 10m ずつの範囲である。セルのサイズは  $0.25\text{m} \times 0.25\text{m}$  である。提案方式の障害物地図は、各セルに走行可能確率を設定し、走行可能な道路には高い走行可能確率を設定し、立体物を含む走行不可能な領域には低い走行可能確率を設定する。

提案方式の障害物地図を、2 種類の既存方式で生成した障害物地図と比較する。1 つは、セマンティックセグメンテーションによる属性情報を、路面が平面でカメラの光軸方向と平行であるという前提の下で路面上に投影し、属性に基づいて路面上の各領域の走行可能確率を設定する方式であり、以降では属性方式と呼ぶ。2 つ目は、立体物に相当する点群のみに基づいて各光線方向の領域に障害物が存在する確率を設定する障害物地図であり、以降では立体物方式と呼ぶ。なお、障害物地図には真値が存在しないため、各障害物地図はカメラ画像と比較することで障害物の位置や走行可能領域の範囲について妥当性を評価する。

図 6、図 7 に各方式で生成した障害物地図を示す。各図において、(a)はカメラ画像、(b)は視差画像、(c)は属性画像を示し、(d)は提案方式、(e)は属性方式、(f)は立体物方式で生成した障害物地図を示している。いずれの障害物地図においても図中で白い領域は走行可能であることを示して

road	sidewalk	building	wall	fence
pole	traffic light	traffic sign	vegetation	terrain
sky	person	rider	car	truck
bus	train	motorcycle	bicycle	

図 5 セマンティックセグメンテーションによる属性情報

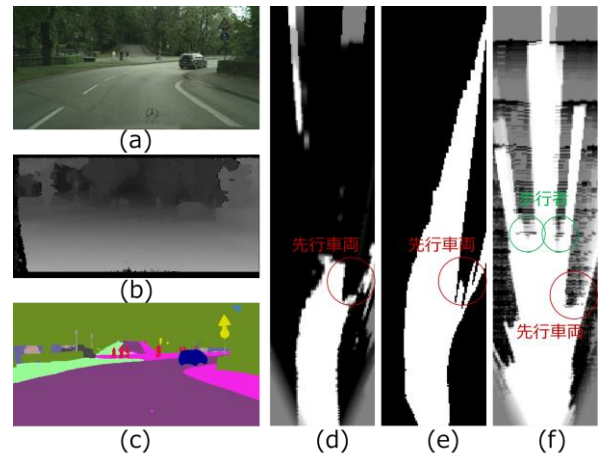


図 6 各方式の障害物地図の比較 (1)

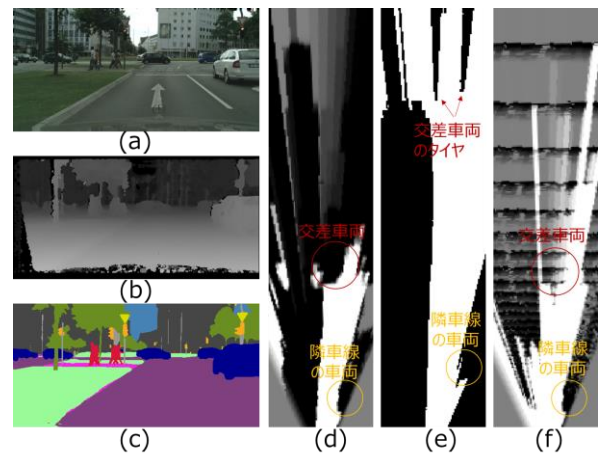


図 7 各方式の障害物地図の比較 (2)

おり、黒い領域は走行不可能であることを示している。図 6、図 7 より、(f)の立体物方式は、他車両、歩行者などの立体物をおおむね正しい位置に認識し、立体物よりも自車に近い範囲をフリースペースとして認識することが出来ている。一方で、芝生や歩道などの自車両が走行する上で障害にはならないが走行してはならない領域に対して、高い走行可能確率(図中で白)で表現しており、走行可能であると認識していることがわかる。

次に、(e)の属性方式は、芝生や歩道などの領域を低い走行可能確率で表現し、走行不可能と認識することが出来ている。しかしながら、図 6、図 7 の(e)の属性方式における走行可能な道路の領域は、(a)のカメラ画像から推測される形状と大きく異なって見える。図 6、図 7 では、(e)の属性方式の走行可能な道路は実際よりも広がっており、特にカメラ位置から遠い領域ほどその傾向が強い。これは、属性方式の路面の形状がカメラの光軸方向に対して平行であるという前提と周囲環境が異なっていることが原因と考えられる。具体的には、図 6、図 7 のシーンでは、路面がカメラの光軸方向に対して上向きの勾配を持っていることが原因と考えられる。また、同様の理由で、図 6 の(e)属性方式では、先行車両の車体が大きく広がっており、図 7 の(e)属性方式では、前方の横断車両のタイヤが大きく広がり車体については障害物地図の表示範囲から外れている。

これらに対して、(d)の提案方式による障害物地図は、まず、図 6、図 7 より芝生や歩道などの領域を走行不可能と認識することが出来ている。さらに、図 6～図 7 より走行可能な道路の領域を適切な形状で表現していることがわかる。また、図 6 において先行車両の形状を大きく広がることなく適切な大きさを表現しており、図 7 においても横断車両を適切な位置で適切な大きさを表現出来ていることがわかる。これらの効果は、提案方式が、属性情報を点の 3 次元情報に対応付けた上で障害物地図の各領域に設定し、走行可能確率を算出したためと考えられる。以上より、提案方式による障害物地図は、走行してはならない平面領域を走行不可能と認識可能であることに加えて、障害物を適切な位置に適切な大きさを表現可能であることを確認した。

## 4.2 計算コスト

提案方式の障害物地図を生成するために必要な処理時間について評価を行った。提案方式の障害物地図の処理としては図 1 の赤枠で囲った範囲の処理について処理時間を計測した。セマンティックセグメンテーションによる属性画像の取得やステレオマッチングに基づく 3 次元情報取得については、障害物地図の処理には依らず別の方式を使用することも可能なため、提案方式の処理時間計測からは除外した。処理時間の計測に使用した PC のスペックは Intel Core i7-3970X CPU@3.50GHz である。表 1 に障害物地図のセルサイズを変化させた時の処理時間の比較を示す。障害物地図の範囲は 20m×80m である。表 1 よりセルサイズが小さくなればなるほどセル数が増加するため障害物地図全体の処理時間は増加する。障害物地図生成の処理のうち一部の処理はセル数ではなくて点群数に依存する処理であるため、セルのサイズに対して、処理時間が線形に変化しないことがわかる。

次に、提案方式の障害物地図と既存方式の障害物地図との処理時間を比較した。比較対象の方式としては、提案方式と同様に自車両が走行する上で障害にはならないが走行してはならない領域を走行不可能と認識可能であり、かつ、立体物を適切な位置に適切な大きさを認識可能な **Semantic 3D Map** を選んだ。これは、提案方式と同様の効果を得られる方式との処理時間の比較を行うためである。なお、**Semantic 3D Map** もステレオ画像から視差画像を算出するが、提案方式と同じ条件で処理時間を比較するため、**Semantic 3D Map** の処理から視差画像算出は除外して処理時間を比較した。両者の処理時間の比較を表に示す。障害物地図の条件は **Semantic 3D Map** に合わせており、障害物地図の範囲は 25m×25m でセルのサイズは 0.1m×0.1m である。表 2 より、提案方式は **Semantic 3D Map** と比較して処理時間が約 1/9 となることがわかる。これは、**Semantic**

表 1 セルサイズによる処理時間の比較

Cell Size [m]	Time [s]
0.1	0.241
0.2	0.114
0.5	0.070
1.0	0.060

表 2 提案方式と既存方式の処理時間の比較

方法	Time [s]
Semantic 3D Map	0.892
提案方式	0.098

**3D Map** が、属性情報を持つ 3 次元ボクセルマップを生成した上で、上面図に相当する 2 次元障害物地図を出力するのに対して、提案方式は 3 次元マップを生成することなく、点群の 3 次元座標と属性から直接上面図に相当する 2 次元障害物地図を生成しているためと考えられる。

## 5. おわりに

本稿では、3 次元点群と 2 次元属性情報を統合した 2 次元障害物地図を生成する方式を提案した。カメラ画像に対する属性情報を 3 次元点群に対応付けて、属性を持った 3 次元点群を水平面に投影することで、3 次元形状を再構成することなく、幾何情報と属性情報を反映した 2 次元障害物地図を生成した。提案方式による障害物地図は、走行してはならない平面領域を走行不可能と認識可能であることに加えて、障害物を適切な位置に適切な大きさを表現することが可能となった。また、計算コストについてもリアルタイム動作に十分な速度で動作可能であることを確認した。

## 参考文献

- [1] A. Elfes, "Using occupancy grids for mobile robot perception and navigation". IEEE Computer, vol. 22, pp. 46–57, June 1989.
- [2] S. Thrun, W. Burgard, and D. Fox, "Probabilistic Robotics". MIT Press, 2005.
- [3] D. Nuss, M. Thom, A. Danzer, and K. Dietmayer, "Fusion of laser and monocular camera data in object grid maps for vehicle environment perception," in International Conference on Information Fusion, July 2014, pp. 1–8.
- [4] Abhijit Kundu, Yin Li, Frank Dellaert, Fuxin Li, and James M Rehg. Joint semantic segmentation and 3d reconstruction from monocular video. In European Conference on Computer Vision, pp 703–718. Springer, 2014.
- [5] S. Yang, Y. Huang, and S. Scherer, "Semantic 3d occupancy mapping through efficient high order crfs," IEEE/RSJ International Conference on Intelligent Robots and Systems, 2017.
- [6] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia. "Pyramid Scene Parsing Network," IEEE Conference on Computer Vision and Pattern Recognition, 2017
- [7] T. Weiss, B. Schiele, K. Dietmayer, "Robust driving path detection in urban and highway scenarios using a laser scanner and online occupancy grids," IEEE Intelligent Vehicles Symposium, pp 184–189, 2007.
- [8] F. Himm, N. Kaempchen, J. Ota, and D. Burschka, "Efficient occupancy grid computation on the GPU with lidar and radar for road boundary detection," IEEE Intelligent Vehicles Symposium, pp 1006–1013, 2010.
- [9] F. Andert, "Drawing stereo disparity images into occupancy grids: measurement model and fast implementation," IEEE International Conference on Intelligent Robots and Systems, pp 5191–5197, 2009.
- [10] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. enenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," IEEE Conference on Computer Vision and Pattern Recognition, 2016
- [10] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. enenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," IEEE Conference on Computer Vision and Pattern Recognition, 2016