

選択的投光に基づく環境照明下の透明物体に対する ステレオ偏光カメラを用いた深度推定

○荻原 佑介, 樋口 寛, 伊賀上 卓也, 安琪, 山下 淳 (東京大学)

Depth Estimation of Transparent Objects under Ambient Illumination Using a Stereo Polarization Camera with Selective Projection

○ Yusuke OGIHARA, Hiroshi HIGUCHI, Takuya IGAUE,

Qi AN, Atsushi YAMASHITA (The University of Tokyo)

Abstract: This research proposes a method to estimate depth of glass using a stereo polarization camera and a projector. In indoor environments, there is a lot of glass in the form of fences, windows, and doors. Since glass is transparent and smooth, it is difficult to measure them with a LiDAR or a stereo camera. Polarization is useful for estimating the shape of an object. However, when measuring transparent objects, transmitted and reflected light cancel each other's polarization. This research proposes a method to acquire polarization information of reflected light by selectively illuminating only non-glass areas. In the indoor experiment, we successfully estimate the depth of glass.

1. 緒言

近年, 警備ロボットや案内ロボットといった用途のために, オフィスビルなどの屋内環境を走行するロボットの需要が増加している. このようなロボットに搭載したセンサを用いて, 自己位置推定や地図生成を行う場合, カメラから周囲の物体までの距離を表す深度画像を生成することが有用である.

一方で, 屋内環境においては柵や窓, ドアのような形で多くのガラスが存在するため, ガラスの位置を計測することがロボットの安全な走行に必要である¹⁾. ガラスは透明であり, 背後の物体が透過して見えることから, 可視情報を用いてステレオ計測を行った場合, ガラスの向こう側の物体の深度を計測する. Time of Flight 方式やアクティブステレオ方式をはじめとする, 計測対象に投光する形式の計測の際も, ガラス表面は滑らかであり, 拡散反射をほぼしないことから, ガラスそのものの深度を計測することは困難である.

透明物体の形状を計測するにあたって, 直線偏光情報を取得可能なカメラである, 偏光カメラが広く使用される. 透明物体に入射した非偏光は, 透明物体の表面を反射, 透過する際に直線偏光するため, 非偏光光源下で透明物体の偏光を計測することで物体の法線の推定が可能である. 偏光カメラを用いることで偏光の位相に対応する偏光角度と, 振幅に対応する偏光度を取得することができる.

Shao らは単一の偏光カメラで透明物体を計測することで, 透明物体の法線を推定する手法を提案した²⁾. この手法では, 非偏光がガラスに入射した際の偏光を透過と反射の重ね合わせとして定式化し, ガラスの法線推定を行った. このとき反射光の偏光角度と透過光の偏光角度は 90 度ずれるため, 打ち消し合うように働き偏光度が小さくなる. これは偏光情報の振幅が小さくなることに対応し, 偏光情報の S/N 比が低下する. したがって透過光の偏光の影響により, ガラス表面で反射した光の偏光情報を取得することは困難になる. このような透過の影響を統制するため, Shao らの手法では計測時に黒く平板な背景を用いることや光源に関し

ての制限が存在する. 加えて, Shao らの手法は法線推定を行うものであり, 移動ロボットの地図生成に必要な深度情報を取得するものではない.

偏光カメラを使用する深度推定手法として, 機械学習ベースのステレオ計測手法である DPSNet³⁾ がある. DPSNet は偏光画像, RGB 画像を入力として, 視差マップを出力とするネットワークである. この手法では, RGB 情報と偏光情報を組み合わせることでテクスチャレスな領域の深度が推定可能である. ただし, この手法においては全領域に対して RGB の情報を深度推定に用いるため, 透明物体の向こう側にテクスチャがある場合, RGB でのマッチングに基づき, 透過した先の深度を出力する.

そこで本手法では, ロボットが走行するような屋内照明灯環境下に適用可能な, ガラスの深度を含む深度画像の生成手法を提案する.

2. 提案手法

2.1 提案手法の概要

物体の透過率と反射率は偏光方向によって異なり, この差により物体に入射した非偏光は偏光する. 光を入射面に垂直な成分と平行な成分に分離した場合の誘電体の反射率と透過率を考えると, 垂直な成分の反射率は平行な成分の反射率以上であり, 平行な成分の透過率は垂直な成分の透過率以上である⁴⁾. つまり, 反射の場合は入射面に垂直な方向に偏光し, 透過の場合は平行な方向に偏光する. このように, 屋内環境において偏光を計測する場合, 透過光と反射光の偏光方向が異なることから, 透過光と反射光の重ね合わせを計測する透明物体の場合は, 互いに打ち消し合い偏光度が低下し, 偏光情報の S/N 比が低下する. 本手法では, ガラス以外の領域にのみ選択的に光を当てることで透過光の寄与を減らし, 反射光の偏光情報を取得する手法を提案する.

本手法ではステレオ偏光カメラによる撮影を 2 回行う. 1 回目の撮影で取得した画像に対してガラスの存在する領域をセグメンテーションする. ガラスセグメンテーション結果をもとに, 画像平面上でガラスが存

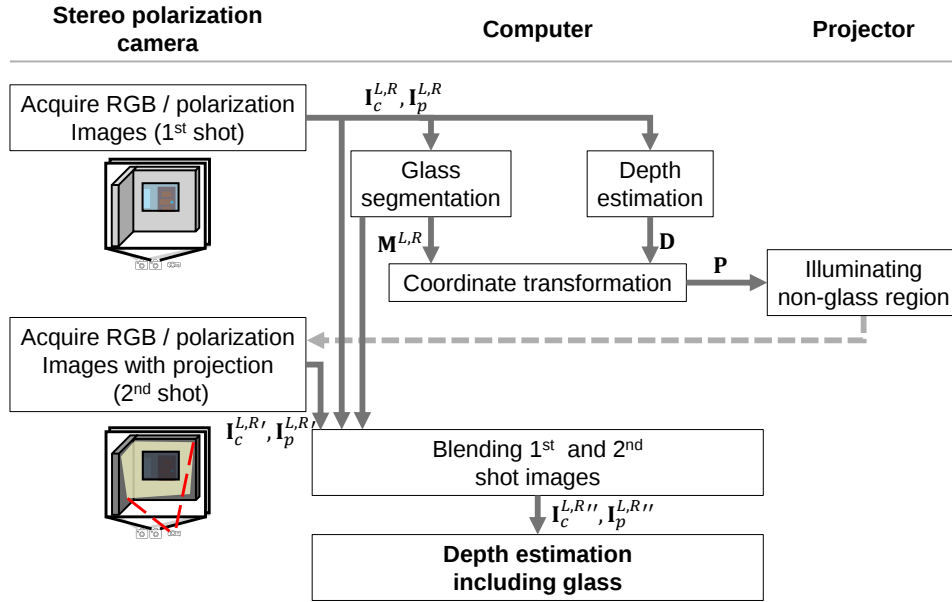


Fig. 1 Flow of the proposed system.

在する領域以外へプロジェクタで光を当てながら2回目の撮影を行う。ガラスの奥の物体には直接光が入射しないため、ガラス背後から透過し、カメラに入射する光が少なくなる。また、ガラスの周囲には光を当てることによって、壁や窓枠で反射した光がガラス表面で反射し、反射光の偏光情報をカメラで取得することができる。

提案手法の流れを Fig. 1 に示す。1回目の撮影ではプロジェクタによる投光を行わず、環境照明下で撮影を行う。このとき、左右それぞれの偏光カメラから偏光画像と RGB 画像を取得することができる。偏光画像と RGB 画像を用いて、ガラスのセグメンテーションを行い、画像平面上でガラスの存在する領域を推定する。

セグメンテーション結果をもとに、ガラス以外の領域に投光するための画像を作成する。セグメンテーション結果は画像座標系でのガラスの方向を表すため、プロジェクタで投光するためにはプロジェクタ座標系への座標変換が必要である。この座標変換には事前に計測した偏光カメラとプロジェクタの内部パラメータ、外部パラメータに加え深度情報が必要である。そのため、1回目の撮影時の左右の偏光画像と RGB 画像を入力として深度推定を行う。なお、この時点で推定する深度はガラスの向こう側の物体の深度を出力しており、ガラスそのものの深度は取得不可能である。このようにして推定した、ガラスの向こう側の物体の深度を用いて座標変換を行い、ガラス以外の領域への選択的投光のための画像を作成する。

2回目の撮影では、プロジェクタでガラス以外の領域への投光を行いながら、ステレオ偏光カメラによる撮影を行う。1回目の撮影の画像と、2回目の撮影の画像を統合することで、深度推定アルゴリズムに入力するための偏光画像と RGB 画像を作成する。これらの画像を深度推定アルゴリズムに投入しガラスの深度を含んだ深度画像を取得する。本手法では、深度推定アルゴリズムとして機械学習ベースのステレオ計測手法である DPSNet³⁾ を用いる。

2.2 透明物体のセグメンテーション

ステレオ偏光カメラでは、左右それぞれの偏光カメラにおいて RGB 画像と偏光画像を取得することができる。左カメラで取得した RGB 画像と偏光画像をそれぞれ $I_c^L \in \mathbb{R}^{h \times w \times 3}$, $I_p^L \in \mathbb{R}^{h \times w \times 2}$ 、右カメラで取得した RGB 画像と偏光画像をそれぞれ $I_c^R \in \mathbb{R}^{h \times w \times 3}$, $I_p^R \in \mathbb{R}^{h \times w \times 2}$ と呼称する。ここで $h, w \in \mathbb{N}$ は取得した画像の縦横の画素数である。本手法では偏光画像と RGB 画像を用いてガラスのセグメンテーションを行うアルゴリズムとして PGSNet⁵⁾ を用いる。PGSNet は、偏光画像と RGB 画像を入力として、画像平面上でガラスの存在する領域を出力するネットワークである。

このアルゴリズムを利用することで左右の偏光カメラ座標系におけるガラスの存在する方向を計測することが可能である。左右のカメラにおけるセグメンテーション結果 $M^L, M^R \in \mathbb{R}^{h \times w}$ は2値画像であり、

$$M^{L,R}(i, j) = \begin{cases} 1 & \text{if glass is included in the pixel} \\ 0 & \text{else} \end{cases}, \quad (1)$$

と表せる。

2.3 プロジェクタによる選択的投光

ガラス以外の領域への選択的な投光を実現するために、偏光カメラ座標系での透明物体のセグメンテーション結果を、プロジェクタ座標系での画像へと座標変換を行う。そのために、事前のキャリブレーションで取得した、偏光カメラとプロジェクタの内部パラメータと外部パラメータに加えて、偏光カメラ座標系での深度画像が必要になる。ここで、1回目の撮影でステレオ偏光カメラで取得した画像を用いて深度推定を行い深度画像 $D \in \mathbb{R}^{h \times w}$ を取得する。ただし、深度画像 D はガラスのある領域ではガラスを透過した向こう側の物体の深度を計測する。

セグメンテーション結果の2値画像をプロジェクタ座標系に座標変換しプロジェクタ座標系でのガラスセグメンテーション結果 $M^P \in \mathbb{R}^{h \times w}$ を取得する。その後、0と1を反転しガラスのない領域を1、ある領域を

0 とする 2 値画像 $\mathbf{P} \in \mathbb{R}^{h \times w}$ を作成する．この画像をプロジェクタで投影することで、ガラス以外の領域に選択的に投光を行い、ステレオ偏光カメラで 2 回目の撮影を行い、左右の RGB 画像 $\mathbf{I}_c^{L,R'} \in \mathbb{R}^{h \times w \times 3}$ と左右の偏光画像 $\mathbf{I}_p^{L,R'} \in \mathbb{R}^{h \times w \times 2}$ を取得する．

2.4 画像の統合と深度推定

プロジェクタによる投光なしで、ステレオ偏光カメラで撮影した偏光画像 $\mathbf{I}_p^{L,R}$ と RGB 画像 $\mathbf{I}_c^{L,R}$ と、ガラス以外の領域への投光をしながらステレオ偏光カメラで撮影した偏光画像 $\mathbf{I}_p^{L,R'}$ と RGB 画像 $\mathbf{I}_c^{L,R'}$ を統合して、深度推定アルゴリズムへの入力とする偏光画像と RGB 画像 $\mathbf{I}_c^{L,R''} \in \mathbb{R}^{h \times w \times 3}$, $\mathbf{I}_p^{L,R''} \in \mathbb{R}^{h \times w \times 2}$ を作成する．

$\mathbf{I}_c^{L,R''}$ においては RGB 情報によるマッチングによってガラス背後の物体の深度を算出してしまふことを防ぐため、ガラスの存在するピクセルの値を 0 とし、それ以外は 1 回目の撮影の画像 $\mathbf{I}_c^{L,R}$ の値を利用する．つまり、

$$\mathbf{I}_c^{L,R''}(i, j) = \mathbf{I}_c^{L,R}(i, j)(1 - \mathbf{M}^{L,R}(i, j)), \quad (2)$$

となる．

$\mathbf{I}_p^{L,R''}$ においては、ガラスの存在するピクセルの値は、プロジェクタで投光しながら撮影した 2 回目の撮影の画像 $\mathbf{I}_p^{L,R'}$ の値を利用し、それ以外の領域においては投光による誤差を防ぐため 1 回目の撮影の画像 $\mathbf{I}_p^{L,R}$ の値を利用する．つまり、

$$\begin{aligned} \mathbf{I}_p^{L,R''}(i, j) = & \mathbf{I}_p^{L,R'}(i, j)(\mathbf{M}^{L,R}(i, j)) \\ & + \mathbf{I}_p^{L,R}(i, j)(1 - \mathbf{M}^{L,R}(i, j)), \end{aligned} \quad (3)$$

となる．

このようにして投光時の情報と非投光時の情報を統合した画像を深度推定アルゴリズムの入力とし、ガラスの深度を含んだ深度画像を出力する．

3. 実験

3.1 実験手順

提案手法によるガラスの深度推定が可能であることを示すために実験を行った．本実験では、Fig. 2 に示すように、偏光カメラとして LUCID 社の VP-PHX050S-Q を、プロジェクタとして EPSON 社の EB-X36 を用いた．実験は屋内、蛍光灯環境下で行った．計測対象は縦 295 mm 横 300 mm 厚み 4.8 mm の平板ガラスを用いた．

本実験では、提案手法を用いてガラスを含む環境の深度推定を行った後、ガラスの 4 隅に AR マーカを貼付し、AR マーカの 4 点からガラス平面の深度の真値を取得し、深度推定結果の精度を評価した．AR マーカの生成および検出には、ArUco⁶⁾ を利用した．

本実験では比較手法として、提案手法を用いることなくガラスを撮影した RGB・偏光画像を DPSNet に入力し、深度推定を行った．

3.2 結果

実験環境において偏光カメラから得られた RGB 画像を Fig. 3(a) に示す．画像内でガラスの存在している領域を赤枠で囲った．ガラスはアルミフレームで固定し、計測を行った．

この環境において、ガラスのセグメンテーションを行った結果を Fig. 3(b) に示す．ガラスの存在する領域

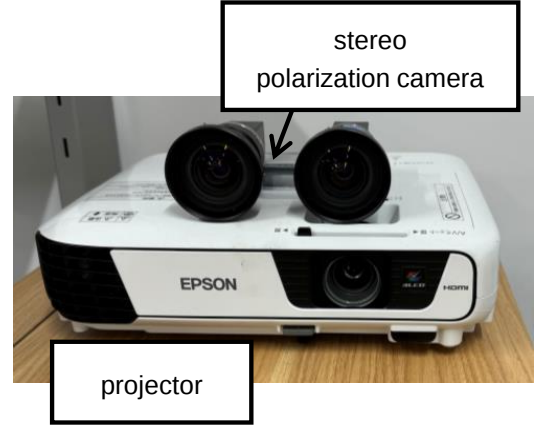


Fig. 2 Experimental equipments.

を白色、それ以外を黒色で表示している．ガラスセグメンテーションにより画像平面上でガラスの存在する領域を取得し、この 2 値画像をもとにガラス以外の領域への投光を行った．

このセグメンテーション結果をプロジェクタ座標系に変換し、ガラス以外の領域に選択的に投光を行った結果が Fig. 3(c) である．ガラスの周囲の壁や天井のみ投光し、ガラスには光が入射していないことが確認できた．

提案手法による深度推定精度を評価する．ガラスの深度の真値を Fig. 4(a)，提案手法による深度推定結果を Fig. 4(b)，比較手法として DPSNet のみで深度推定を行った結果を Fig. 4(c) に示した．これらの画像においては色が深度を表し、深度が小さいほど青色寄りの色に、大きいほど赤色よりの色になるように可視化している．ガラスのある部分の深度が、比較手法では奥の壁と同程度の深度を出力しているのに対し、提案手法では、より実際のガラスの深度に近い値を出力している．

透明物体の深度推定精度評価として、提案手法と比較手法のピクセルごとの真値からの誤差の平均絶対値誤差と Root Mean Squared Error (RMSE) を評価する．提案手法の平均絶対値誤差は 0.062 m，RMSE は 0.064 m となり、比較手法の平均絶対値誤差は 0.28 m，RMSE は 0.28 m となった．この環境では、ガラスと背後の壁までの距離が 0.3 m 程度であり、比較手法では奥の壁の深度を出力していると考えられる．一方で、提案手法ではガラスそのものの深度を出力できていることを確認した．

3.3 考察

本研究では蛍光灯照明環境下において、プロジェクタを用いてガラス周囲の不透明物体に投光することで、ガラス表面で反射した光の偏光情報の S/N 比を向上した．DPSNet のみを使用した場合、Fig. 4(c) のように、奥側のガラスの深度を出力する環境であっても、提案手法においては Fig. 4(b) に示すように、ガラスの深度を推定することに成功した．

今回使用した深度推定アルゴリズムである DPSNet は、深度推定に RGB の情報と偏光の情報の両方を使用している．提案手法においては、2 回の撮影で取得した画像を、RGB と偏光のそれぞれで合成し、深度推定アルゴリズムへの入力画像を作成している．そのため、提案手法は RGB 画像においては、ガラスセグメンテー

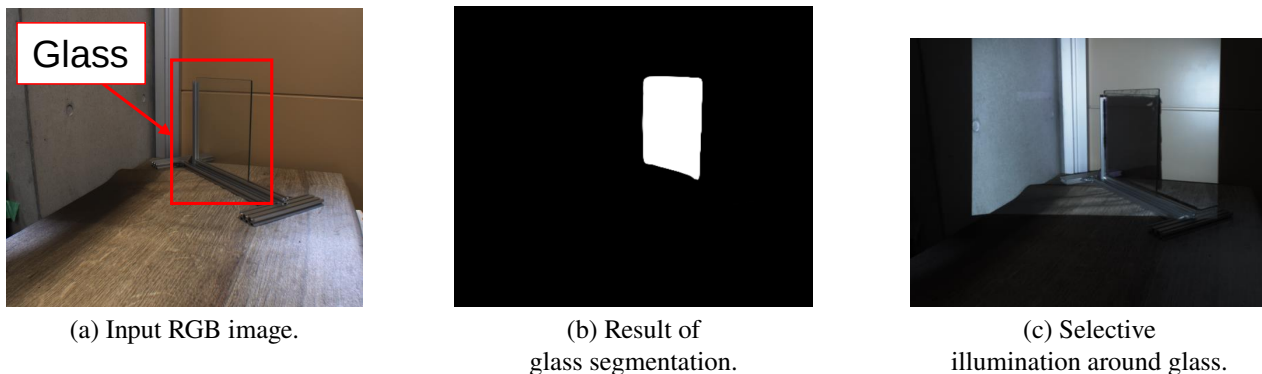


Fig. 3 Glass segmentation and selective illumination with a projector.

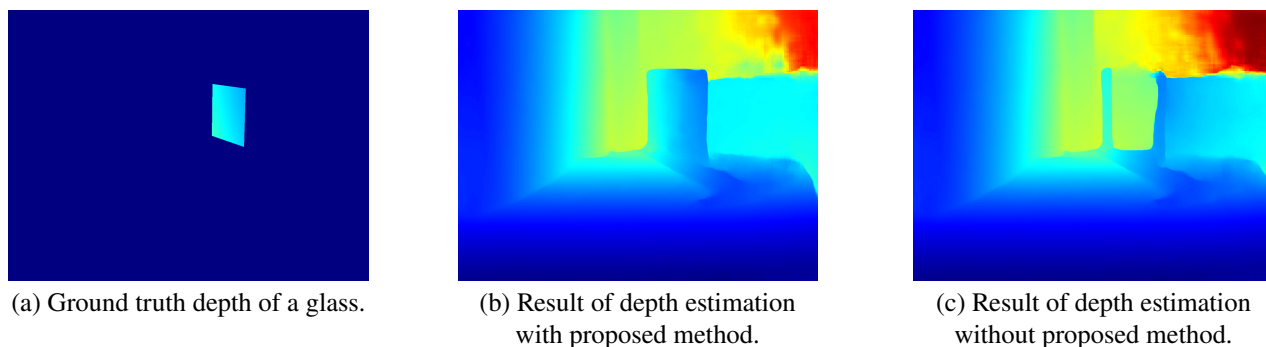


Fig. 4 Evaluation of depth map.

ション結果を活用することで、ガラス領域の RGB 情報を使用しないことでガラス背後の物体の深度を推定することを防ぐことができたと考えられる。

偏光画像においては、選択的投光により反射光の偏光情報の S/N 比を高めて、取得することができた。RGB 画像でガラス領域の情報を使用していないことから、この部分の深度を直接計測することはできないが、ガラスの端や周囲の物体の深度と、ガラスの法線に対応する偏光角度を組み合わせることで、ガラスの深度を推定することに成功したと考えられる。

4. 結言

本研究では、ステレオ偏光カメラとプロジェクタを用いることで、カメラの画角中のガラス以外の領域に選択的に投光し、ガラス表面で反射した偏光成分を取得することでガラスの深度推定を行う手法を提案した。実験においては、従来手法ではガラスの奥の深度を推定してしまうような環境であっても、ガラス自体の深度を推定することに成功した。

今後の展望としては、提案システムを移動ロボットに搭載し、本手法を SLAM に応用することが挙げられる。

謝辞

本研究の一部は JSPS 科研費 22H03666 の援助を受けた。

参考文献

- [1] P. Foster, Z. Sun, J. J. Park, and B. Kuipers. Visagge: Visible Angle Grid for Glass Environments. Proceedings of the 2013 IEEE International Conference on Robotics and Automation, pp. 2213–2220, (2013).
- [2] M. Shao, C. Xia, Z. Yang, J. Huang, and X. Wang. Transparent Shape from a Single View Polarization Image. Pro-

ceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9277–9286, (2023).

- [3] C. Tian, W. Pan, Z. Wang, M. Mao, G. Zhang, H. Bao, P. Tan, and Z. Cui. DPS-Net: Deep Polarimetric Stereo Depth Estimation. Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3569–3579, (2023).
- [4] 宮崎大輔, 池内克史. 偏光の基礎理論とその応用. 情報処理学会論文誌コンピュータビジョンとイメージメディア 1.1, pp. 64–72, (2008).
- [5] H. Mei, B. Dong, W. Dong, J. Yang, S.-H. Baek, F. Heide, P. Peers, X. Wei, and X. Yang. Glass Segmentation Using Intensity and Spectral Polarization Cues. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12622–12631, (2022).
- [6] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez. Automatic Generation and Detection of Highly Reliable Fiducial Markers under Occlusion. Pattern Recognition 47.6, pp. 2280–2292, (2014).