

短波長赤外線カメラと深層学習を用いた 物体の濡れを検出するビジョンシステムの開発

林元 辰己[†] 古賀 崇了[†]
[†] 近畿大学大学院産業理工学研究科

1. 背景

近年、短波長赤外 (SWIR) 帯の撮像を可能とする InGaAs カメラやハイパースペクトルカメラの様々な分野への応用が進んでいる。この波長帯は水の吸光波長を含むため、水分の検出などによく利用される。さらに、近年では民生用途のホームロボットが普及しており、高度なビジョンシステムを備えているが、水分や濡れの検出という観点に着目すると、その性能はまだ十分ではない。

2. 目的・手段

本研究ではホームロボットのビジョンシステムに、雨天時の廊下の水溜まりや机の上に水がこぼれた状況などを画像から正確に識別できる機能を付与することを目的とする。具体的には、深層学習によるシーン中の水分の検出と状況の把握という2つのタスクを高い精度で実現することを目指す。本研究は、RGB 画像だけでなく、SWIR 帯の画像を深層学習の入力情報として用いることで上記のタスクを解決しようとする点に特色がある。以下、手段を説明する。

2.1 画像位置合わせ

本研究では、RGB 画像と SWIR 画像の情報を融合させて学習させる。実験に用いる RGB カメラと InGaAs カメラは、別の筐体であり撮像位置が異なるため、RGB 画像と SWIR 画像にずれが生じる。これを解決するために、レジストレーションを行う必要がある。RGB 画像と SWIR 画像では、シーン中の特徴点の自動マッチングに基づく画像位置合わせが難しく、別の手段を用いる必要がある。本研究では、それぞれのカメラでチェッカーボードを撮像して、特徴コーナーを特徴点として検出し、RGB と SWIR のそれぞれの点を手動で対応させてホモグラフィ変換作業を行った。

2.2 深層学習

SWIR 帯の 1450nm の波長周辺で撮像した画像では水分が黒く見える。この特性を活かして RGB 画像 3 チャンネルに SWIR 画像 3 チャンネルを加えたマルチモーダルな機械学習を実施する。本研究では、Eitel ら[1]が開発した RGB-D (RGB 画像と Depth 画像) の深層学習モデルの入力画像である Depth 画像を SWIR 画像に置き換えて機械学習を行う。Eitel らのデュアルストリームモデルは、基本的に AlexNet の構造で出来ている(図1)。実験では、学習済みモデルとして VGG モデルを使用している。SWIR 画像の学習済みモデルはないため、VGG モデルに RGB 画像を学習させたものを用いることで精度を向上させる。

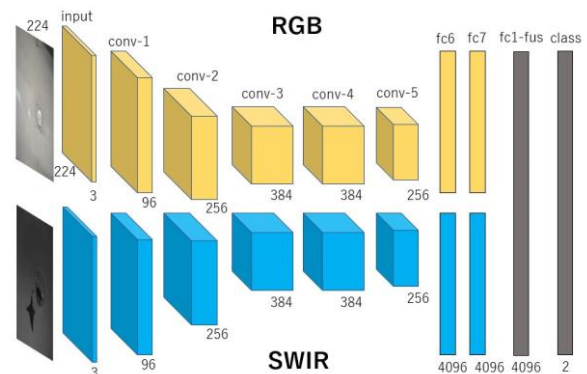


図1 RGB と SWIR のデュアルストリームモデル

学習方法は以下の通りになる。

- ① VGG モデルに RGB 画像を学習させる。
- ② ①の学習済みモデルに SWIR 画像を学習させる。
- ③ ①と②のモデルの両方を用いて RGB 画像を RGB ストリーム(図1黄色)、SWIR 画像を SWIR ストリーム(図1青色)にそれぞれ通したものを融合し、融合層のみを学習させる(図1中の fc1-fus)。

3. 実験結果

まずは、基本的な問題(水分の有無の検出)を解くために、水の入ったペットボトルと入っていないペットボトルの 2 クラス分類の実験を行った。学習画像の RGB 画像と SWIR 画像合わせて 1320 枚(訓練用 1040 枚検証用 280 枚)を用意し、ペットボトルに水が入っているかどうかの認識を行った。その結果、96.4%の精度を確認した。この結果から水分の有無が認識できることがわかった。

4. まとめと今後の展望

現段階で、従来の画像認識モデルに SWIR 画像情報を加えることで、水分の有無を認識し、クラス分類をすることが可能になった。今後は、ホームロボットがタスクをこなすのに有用な認識ができるようクラス数を増やして実験する。また、他のネットワークモデルを用いることで精度が向上するかについても実験を行う予定である

謝辞 本研究は科研費 19K12091 の助成を受けた。

参考文献

- [1] A. Eitel, *et al.*, "Multimodal deep learning for robust RGB-D object recognition," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.681-687, 2015.