

全方位カメラと認識技術を用いた 視覚障害者支援システム

岩村 雅一¹ 南谷 和範² 井上 慶彦¹ 程 征¹
内海 ゆづ子¹ 黄瀬 浩一¹

¹ 大阪府立大学 大学院工学研究科

² 大学入試センター

1 はじめに

近年のディープラーニングの発展により、文字や物体の認識・検出などの精度が著しく向上し、人と同程度の認識性能に近づいてきている。そこで、これらの技術を視覚障がい者の「目」として扱う試みが盛んになっている。

例えば、OrCam MyEye 2.0¹は眼鏡に取り付けるウェアラブルデバイスで、文字の読み取り、顔認識などの機能を備える。しかし、このデバイスや他のスマートフォンアプリなどは、認識対象が目の前であることを想定している。すなわち、対象の場所（where）は既知として、何か（what）を言い当てることに特化していると言える。この前提が成り立つ場合も多々あるが、視覚障害者にとって、そうでない場合も少なくない。

そこで、where と what を同時に解決できる方法として、本研究ではユーザの周囲360度が撮影できる「全方位カメラ」を用いることを提案する。

2 提案システム

図1に提案するシステムのプロトタイプを示す。このシステムは小型PCに接続された全方位カメラから成る。全方位カメラ（RICOH THETA V）でユーザの周囲360度の画像を撮影して、PCに送る。PC（Raspberry Pi 3 Model b+）では、ディープラーニングを利用した物体検出手法（MobileNet-SSD）により、学習してある物体を検出する。検出した物体の情報は、ユーザに音声で伝えることができる。ただし、認識した周囲の情報をむやみにすべて伝えることは、かえって混乱

¹<https://www.orcam.com/ja/myeye2/>



図 1: 提案システム



図 2: 正距円筒図法での物体検出例

を招く。そのため、認識技術で得られた周囲の情報の中から、ユーザにとって重要な情報を選択するシステムを別途研究中である。

全方位カメラでは、図 2 に示すような正距円筒図法という歪んだ画像が得られる。通常の物体検出手法は歪んだ画像では検出性能が低下するため、正距円筒図法の一部を通常の画像（透視投影図）に変換してから物体を検出する。現在の実装では、360度を45度ずつ8回に分けて物体を検出する。物体検出はディープラーニングに基づく手法であるため、CPUを使用して実行するには処理が重い。そのため、ポータブルコンピュータ向けのディープラーニング専用デバイスである Intel Neural Compute Stick 2 を 4 本使用して、処理を高速化する。

検出された物体のことをユーザに音声で伝える場合、物体が検出される度に何度も伝えられたら迷惑である。そのため、検出された物体を追跡することで、ある時刻で検出された物体と別の時刻で検出された物体が同じかどうかを判定する。

3 この先実現したいこと

今回紹介したのは、ユーザの周囲の物体を検出して、ユーザに伝えられるシステムのプロトタイプである。このシステムのユーザビリティの向上もさることながら、このシステムの物体検出結果を取捨選択したり、要約したりすることで、なるべく少ない言葉で、なるべく短時間に周囲の状況を伝えられることが今後の目標である。

謝辞 本研究は、JSPS 科研費#17H01803 の補助による。