

ウェアラブルカメラを用いた視覚障害者の行動認識と支援 Action recognition and support for visually impaired people using wearable cameras

村井 保之[†] 巽 久行[‡] 太田 友三子[†] 徳増 眞司^{*} 宮川 正弘^{*}
Yasuyuki Murai Hisayuki Tatsumi Yumiko Ota Shinji Tokumasu Masahiro Miyakawa

1. はじめに

本研究は、ウェアラブルカメラを用いて視覚障害者の行動を認識し、その行動の支援を自動化することを目標とする[1-3]。視覚障害者の場合、カメラの映像から行動の意味を推定することは難しい。理由は視覚に障害があるために、カメラの映像が必ずしも行動認識に結びつくような映像になっていないからである。人による行動支援では、支援者が視覚障害者の手を取って行動を支援する方法がある。例えば、視覚障害者に対して机上の物の位置を示す場合、単に方向や位置を口頭で伝えるだけでなく、視覚障害者の手を取ってその方向や位置を示すと効果的である。そこで本報告では、支援者が手を取る代わりに、指先に PC と接続した小型カメラを装着し、指先のカメラが捉えた画像を事前学習済みの畳み込みニューラルネットワークで認識し音声で利用者に知らせる。それが目的とする物である場合その位置を記憶し、その後の動作の支援を行うシステムの試作を行った。

2. 関連研究

視覚障害者向けの画像認識サービスが数多く登場している。例えば、専用のウェアラブルデバイスを用いる OrCam や、スマートフォンアプリの Microsoft Seeing AI などである。また、指先にカメラを装着するシステムとして EyeRing が Nanayakkara, S.ら[4]により開発されている。EyeRing で使用されるカメラは独自に開発されたもので制御装置を内蔵し指先にはめることができる。このシステムは、指に付けたカメラを対象物に向けることで対象物を認識し利用者に音声で示す。対象を指差すのは自然な動作であり利用者にとって無理なく行う事ができる。本研究でも指差し動作を利用者インターフェースとして利用する。

これらのサービスは、スマートフォンのカメラや専用のデバイスで捉えた画像を認識し、文字の読み上げ、貨幣の識別、周辺状況などを視覚障害者に示すものである。特に、Seeing AI は Microsoft 社の Azure Cognitive Services が使われており、アプリケーションの開発も可能なので本研究でも使用を検討している。

従来の視覚障害者の屋外歩行を支援する研究の多くは、歩行環境画像を解析し、看板等の文字を認識し移動位置を特定し横断歩道や信号の識別を行うものであった。これらの研究では、歩行中に十分な情報が取得できない状況では危険を回避することが困難であった。その原因の一つに、視覚情報が得られない場合の歩行時の偏軌傾向(真っ直ぐ歩けずに曲がってしまう特性)がある。このような歩行時における視覚障害者の偏軌傾向は駅のホームでの転落事故



図 1 実行結果

の要因でもある。本研究では、ウェアラブルカメラによる画像認識を用いて偏軌傾向の補正を目的の一つとする。

3. システムの試作

視覚障害者の行動支援システム開発の前段階として、ウェアラブルカメラを用いた周辺環境の認識プログラムと偏軌傾向検出プログラムを試作した。

周辺環境認識プログラムは、利用者が指で指し示した先にある物の認識を行う。まず、指先に PC と接続した小型カメラを装着する。実用的にはスマートフォンを手に持ち撮影する方法も考えられる。次に、指先のカメラが捉えた画像を、事前学習済みの VGG16 畳み込みニューラルネットワークで認識させる(図 1)。VGG16 は、機械学習で良く用いられる、ImageNet データベースの 100 万枚を超える画像で学習済みの畳み込みニューラルネットワークである。ネットワークの深さは 16 層あり、画像を 1000 個のカテゴリ(時計、机、PC、マウス、多くの動物など)に分類できる。試作プログラムの開発環境を以下に示す。

ソフト	ハード
Python 3.7	MacBook Air (Retina, 13-inch, 2018)
OpenCV 4.1	1.6 GHz Dual Core Intel Core i5
Keras 2.2.4 VGG16	8 GB 2133 MHz LPDDR3
	Intel UHD Graphics 617 1536 MB

Python で作成したプログラムは約 100 ステップ。指先の USB カメラの画像(1280x720 ピクセル)から、中心位置の 400x400 ピクセルを切り出す。切り出した画像を VGG16 で認識し、認識結果の上位 5 をカメラ画像上部に表示する。開発したプログラムを用いた実験の結果を以下に示す。

使用した PC は CPU の処理能力もそれほど高くなく、GPU も使用できないためフレームレートは 1.8 フレーム/秒程度であった。図 2 は、壁の時計を認識した結果である(図 1 の左上の拡大)。図左の赤枠の画像が認識すべき画像である。この画像は、カメラ画像の中心位置(図 1 中央

[†] 日本薬科大学 Nihon Pharmaceutical University

[‡] 筑波技術大学 Tsukuba University of Technology

^{*} 神奈川工科大学 Kanagawa Institute of Technology

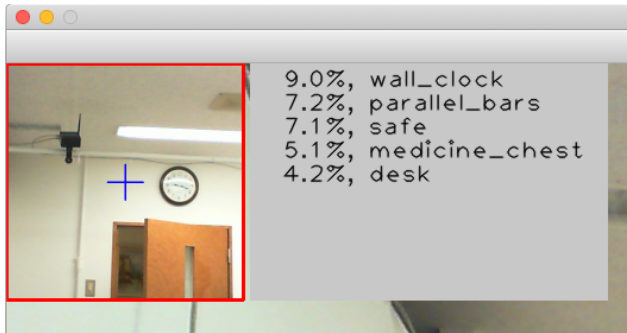


図 2 認識結果

の十字：指先の位置)の周囲 400x400 ピクセルを切り出したものである。図 2 右は、認識結果の上位 5 位までを示したものである。数値は認識率で隣は認識した物を表す。結果は、時計まで約 5m と距離が遠いため 9.0% と認識率は低いが壁の時計 (wall_clock) を認識できた。

認識結果は、カメラを少し移動すると次々と変化する。その中から、利用者にとどれをどの様に提示するかその方法を現在検討中である。

4. 偏軌傾向の補正

本研究で開発するシステムは、利用者の指差し動作とウェアラブルカメラの画像から周囲の状況を認識し行動支援を行う。また、偏軌傾向が見られた場合それを補正することを目指す。そのため、偏軌傾向検出のためのプログラムを試作した。偏軌傾向検出手順を以下に示す。

まず、システムが認識した物の中から利用者は次の行動に必要な目標を選択する。次に、選択した目標をカメラの最新画像で検出し数フレーム前の位置との差を求める。求めた差が偏軌となる。図 3 は検出のイメージで、左図のマウス上の赤い十字が設定した目標である。右図の赤い十字はカメラの現在の視点で、指が目標 (マウス) から離れたことを示す。マウスから伸びる赤い矢印が偏軌となる。この偏軌が少なくなるように利用者に指示を出す。

偏軌傾向を検出するために必要なカメラの移動を検出するために OpenCV のオプティカルフロー関数を使い物の移動を検出するテストプログラムを Python で作成した。プログラムの概要を以下に示す。

- USB カメラから画像をキャプチャする
- キャプチャ画像の中心部分 400x400 ピクセル切り出す
- 切り出した画像をグレースケールにする
- グレースケールを calcOpticalFlowFarneback で処理しオプティカルフローを求める
- 求めたオプティカルフローで移動した方向を図示する

オプティカルフローとは物体やカメラの移動によって生じる隣接フレーム間の物体の動きの見え方のパターンで、



図 3 偏軌傾向の検出

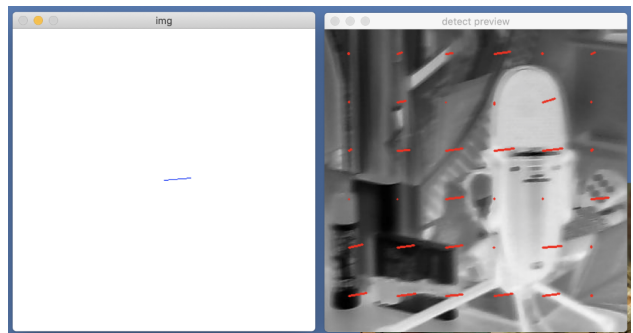


図 4 オプティカルフロー表示

各ベクトルが 1 フレーム目から 2 フレーム目への変位ベクトルを表す 2 次元ベクトル場で表現される。

テストプログラムの実行結果を図 4 に示す。OpenCV の関数ではオプティカルフローは画像の全画素に対して求められる。しかし、全画素について表示したのでは見にくくなるので、本プログラムでは画像を 6x6 の 36 分割しその中心の画素のみオプティカルフローを図示した。図 4 はカメラが左から右に移動した後の画像である。図右の赤い線の多くが右に伸びておりカメラが右に移動した事がわかる。左図は 36 点の平均値を図示したもので、線が右に伸びており、全体として右に動いていることを示している。

OpenCV の関数を用いてカメラの移動方向を検出する事が確認出来た。今後は検出した移動方向から利用者の偏軌傾向を求め利用者の行動を補正する方法を開発する。開発した方法を今回試作したシステムに実装し実証実験を行う予定である。

5. おわりに

試作したプログラムで指先のカメラにより指のさし示す先に何があるかを認識できた。オプティカルフローを用いてカメラの移動方向が検出できた。今後はウェアラブルカメラから得られた歩行環境の分析と視覚障害者自身の指差喚呼に基づき偏軌を修正するための方法の開発し実証試験を行いたい。構築するシステムは視覚障害者の視覚を代行し、歩行時の偏軌を修正して視覚障害者が単独歩行できるものを目指す。

謝辞

本研究は JSPS 科研費 JP17K01092 の助成を受けたものである。

参考文献

- [1] 村井保之, 巽久行, 徳増真司, 宮川正弘, “光学式モーションキャプチャによる触指動作の記録と分析”, 第 18 回情報科学技術フォーラム(FIT2019), Vol. 3, No. K-028, pp.363-364 (2019).
- [2] Yasuyuki Murai, Hisayuki Tatsumi, Masahiro Miyakawa, “Recording of Fingertip Position on Tactile Picture by the Visually Impaired and Analysis of Tactile Information”, Springer LNCS 10897 Part II, pp.201-208 (2018).
- [3] Yasuyuki Murai, Masaji Kawahara, Hisayuki Tatsumi, Tomoyuki Araki, Masahiro Miyakawa, “Congestion Recognition for Arm Navigation, --Aids for the Visually Impaired--”, Proc. 2010 IEEE Int. Conf. on Systems, Man and Cybernetics, No.539, pp.1530-1535 (2010).
- [4] Suranga Nanayakkara, Roy Shilkrot, K. S. Yeo, Pattie Maes, “EyeRing: a finger-worn input device for seamless interactions with our surroundings”, In Proceedings of the 4th Augme Human International Conference, pp.13-20 (2013).