

視覚障がい者のためのデプスカメラを用いた 下り階段検知・通知システムの開発

齋藤 勇磨^{1,a)} 塩野目 剛亮^{1,b)}

概要: 本研究では、視覚障がい者の下り階段からの転落を防止するために、デプスカメラで撮影した深度画像を入力として下り階段を検知し、その方向と距離を通知するシステムを提案する。検知には、CNNの深層学習を用いる。システムは、下り階段・上り階段・平坦な床を学習したモデルにより、各カテゴリに分類される確率を出力する。下り階段を検知した際には、階段までの距離を測定し音声で通知する。本稿では、システムの概要について述べ、データセットとモデルの作成、および下り階段検知の試行結果について報告する。

キーワード: 視覚障害者, 歩行支援, 深層学習, デプスカメラ, 階段認識

Development of Down Stairs Detection and Notification System Using Depth Camera for Blind People

SAITO YUMA^{1,a)} SHIONOME TAKEAKI^{1,b)}

Keywords: Blind people, walking support, deep learning, Depth camera, stairs recognition

1. はじめに

2016年度の厚生労働省の調査によると、身体障害者手帳を所持している視覚障がい者は31.2万人である[1]。視覚障がい者は移動に困難があり、下り階段から転落し怪我をする事故が発生している[2]。

視覚障がい者の事故防止のために、先行研究では、歩行者の頭頂部に配置した単眼視カメラによる周辺環境の認識[3]や、Kinect白杖システムによる下り階段の認識[4]など、視覚障がい者の歩行支援に関する研究が行われている。これらの研究は階段の認識や階段までの距離情報の取得を目的としており、周辺環境の情報をフィードバックはしていない。

本研究では、視覚障がい者の下り階段転落防止を目的として、デプスカメラを用いて下り階段を検知し、階段までの距離、および方向を音声で通知するシステムの開発を行う。

2. 関連研究

視覚障がい者の歩行支援に関する研究は数多く存在する。山口らは、Kinectを用いた白杖システムによる下り階段認識の手法を提案している[4]。空間をセルで分割した上で物体の影を考慮する手法により、下り階段を正確に認識することができ、提案手法の有効性を確認している[4]。

大木らは、視覚障がい者の歩行を支援するために、ウェアラブル端末の単眼カメラで撮影したRGB画像を入力として、バー型障害物を検出する手法を提案している[5]。

速水らは、2台のステレオカメラで、撮影した階段画像のペアから、階段の特徴を抽出し、それを利用したマッチング法を提案している[6]。

¹ 帝京大学理工学部
Teikyo University, Toyosatodai, Utsunomiya-shi, Tochigi
320-8551, Japan

a) 1831234r@stu.teikyo-u.ac.jp

b) shionome@ics.teikyo-u.ac.jp

諏訪部らは、視覚障がい者誘導システムの中で、階段を含めた誘導方法について検討し、昇降開始判定と昇降終了判定の2つのシステムを開発している [7].

島川らは、段差検出の方法として、RGB 画像と Depth 画像の両方を補完的に利用するとともに、特徴点追跡手法を導入した手法を提案している [8].

原田らは、マルチチャンネル化と焼きなまし法に基づく特徴画像の選択、深層学習によって椅子を認識する手法を提案している [9].

今井らは、RGB-D カメラで撮影した深度画像を元に、凹凸の無い歩行可能な平面領域を検出し、安全に歩ける範囲を認識する方法を提案している [10].

工藤らは、5G 環境に適応した視覚障がい者向けの外出歩行支援デバイスを提案している。視覚障がい者が持ち歩くカメラで捉えた画像を、クライアントサーバーモデルを用いて処理用デバイスに送信し、障害物検知、偏軌^{*1}の検出、段差の検出を行うシステムである [11].

田中らは、赤外線を多方向に照射する距離画像センサを用いて障害を検出し、歩行時の注意喚起をリアルタイムに行うことで、視覚障がい者の単独歩行を支援するシステムを提案している [13].

3. 提案するシステム

提案システムは、暗所においても物体認識が可能であるデプスカメラ (Intel 社製, RealSense D435i, 以下 RealSense) とモバイル端末 (Android) を用いて、下り階段を認識し、距離を通知する (図 1)。視覚障がい者が歩行する際に携行義務のある白杖に RealSense を装着する。さらに、Android 用のアプリを開発し、認識処理と通知をモバイル端末上で行うことで、システムの携行性向上を目指す。

下り階段の認識は、デプスカメラで周辺環境の深度画像をリアルタイム取得し、学習済みモデルを用いて、上り階段・平坦な床との判別によって行う。距離は RealSense に搭載されているセンサーでデプスカメラ-下り階段間の距離を測定する。下り階段を認識した場合、下り階段までの距離を通知アプリケーションを用いて「下り階段が前方 1m 先にあります」とモバイル端末から音声読み上げで通知する (図 1)。白杖は周辺環境の情報を得るために左右に動かす場合があるため、デプスカメラを装着した本体自体が動くことを前提とし、解像度 1280×720px, 最大 90fps で深度情報が取得できる RealSense を用いる。

先行研究では、健常者が平均歩行速度 (1.38[m/s]) で歩行した場合、平均停止時間は 2.04[sec], 減速開始から停止までの平均距離は 0.99[m] であった [14]。一方、視覚障がい者の平均歩行速度は 0.78[m/s] である [15]。大木らは視覚障がい者から 2.5[m] 以内のバー型障害物の認識を目標と

*1 偏軌とは、歩行中に、本人の意志とは関係なく、歩行コースの進行方向から右、または左に自然にそれてしまうことである [12].

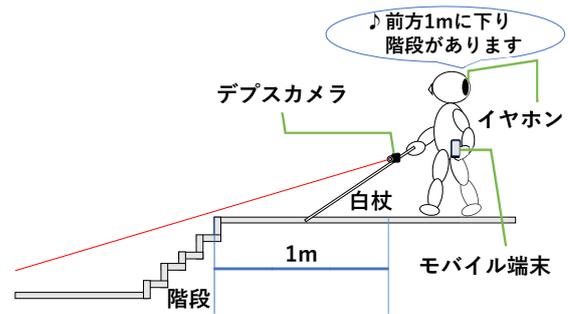


図 1 提案システム
Fig. 1 Proposed system.

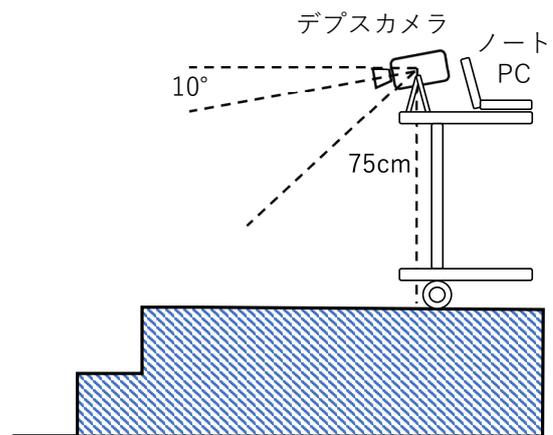


図 2 撮影環境
Fig. 2 Shooting environment.

していた [5] が、本研究では、視覚障がい者から比較的近い範囲 (1.0[m] 以内) に存在する下り階段の認識を目指す。

4. データセットの作成

4.1 データセットの撮影環境

本来は使用状況を考慮して、白杖にデプスカメラを取り付けて撮影することが望ましい。便宜上白杖にデプスカメラを取り付けた際と同程度の高さと同程度の角度 (およそ 75cm, 下向きに 10 度程度) で荷台の前方にデプスカメラを固定し、ノートパソコンを設置して撮影を行った。撮影環境を図 2 に示す。

4.2 学習用データの概要

データセットを作成するにあたって、帝京大学宇都宮キャンパス内の階段と平坦な床 (屋内・屋外) と深度画像を RealSense Viewer で撮影した。撮影時のパラメータを表 1 に示す。撮影した画像は解像度が 424×240px であり、上り階段、下り階段、平坦な床の 3 つにカテゴリ分けをした。また、機械学習をするにあたって、画像編集ソフトウェア Ralpha Image Resizer [16] (以下, Ralpha) を用いて、240×240px の正方形でトリミングした。この際、学習用データを増やすために、元画像の位置を x 軸方向に 50px

表 1 RealSense Viewer の設定

Resolution	848 × 480
Frame Rate(FPS)	30
Depth	z16
Emitter Enabled	Laser
Depth Visualization	Dynamic
Color Scheme	White to Black

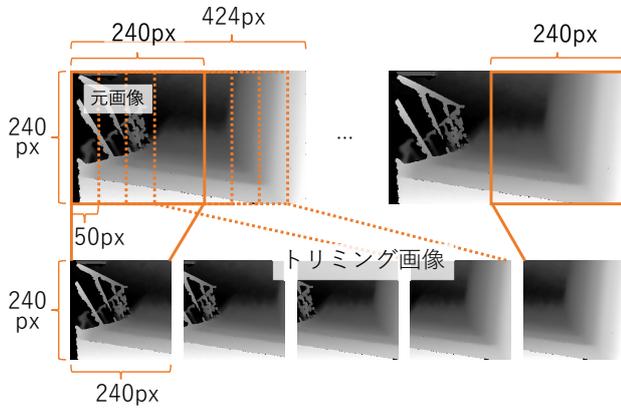


図 3 撮影画像のトリミング
Fig. 3 Cropping of the captured image.

Model: "sequential_14"

Layer (type)	Output Shape	Param #
conv2d_42 (Conv2D)	(None, 240, 240, 16)	800
max_pooling2d_28 (MaxPooling)	(None, 120, 120, 16)	0
conv2d_43 (Conv2D)	(None, 120, 120, 128)	100480
conv2d_44 (Conv2D)	(None, 120, 120, 256)	1605888
max_pooling2d_29 (MaxPooling)	(None, 60, 60, 256)	0
dropout_28 (Dropout)	(None, 60, 60, 256)	0
flatten_14 (Flatten)	(None, 921600)	0
dense_28 (Dense)	(None, 128)	117964928
dropout_29 (Dropout)	(None, 128)	0
dense_29 (Dense)	(None, 3)	387
Total params: 119,672,483		
Trainable params: 119,672,483		
Non-trainable params: 0		

図 4 シーケンシャルモデルの構成
Fig. 4 Sequential model.

ずつ移動させトリミングを行い、学習用データの増強をした。すなわち、元画像から5枚の学習用画像を生成した(図3)。上り階段の画像が940枚、下り階段の画像が789枚*2、平坦な床の画像が185枚、合計で1914枚の深度画像を用いて、データセットを作成した。

5. CNNによる機械学習

5.1 学習モデルの作成

本研究では、Jupyter Notebook上で機械学習を行う。

*2 平坦な床と区別できない画像1枚について削除した。

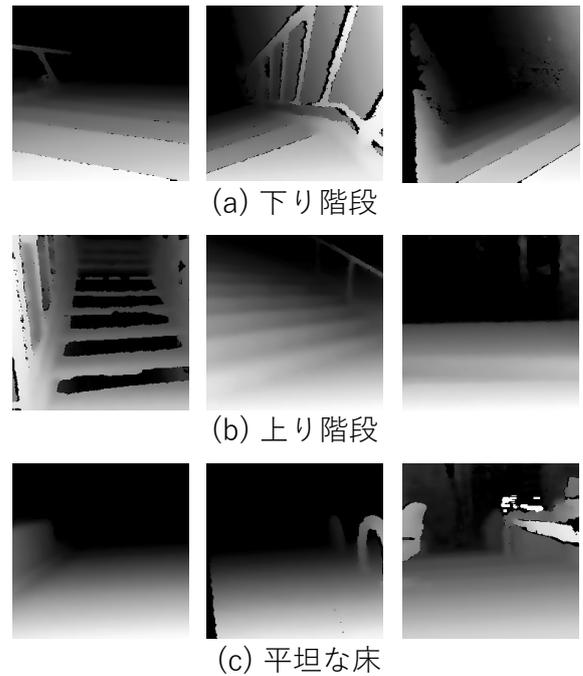


図 5 学習に用いた画像の例
Fig. 5 Examples of training image.

ディープラーニングの実装には機械学習のライブラリであるTensorFlow, およびKerasの2つを用いる。また、画像や動画認識でよく使用されるモデルである、Convolutional Neural Network(CNN)を採用した。ニューラルネットワークのモデルはシーケンシャルモデルであり、構造を図4に示す。学習に使用した計算機は、CPU: Intel core i5-1135G7 2.40[GHz]を搭載したノートパソコン(Windows10)である。データセットは第4章で述べたものを使用する。使用した深度画像を抜粋し、図5に示す。

深度画像は3カテゴリ(下り階段, 上り階段, 平坦な床)の場所について撮影したもので、図5(a)左が手すりがある屋外の下り階段, 中央が手すりがある屋内の下り階段, 右が手すりがない屋内の下り階段である。図5(b)左が手すりとの隙間がある屋内の上り階段, 中央が手すりがある屋外の上り階段, 右が手すりがない屋外の上り階段である。図5(c)左と中央が屋外で障害物がある平坦な床, 右が屋内で椅子や机が存在する環境である。

学習回数の値を[5, 10], バッチサイズの値を[8, 16, 32, 64, 128]に設定し、学習を行った。

5.2 学習モデルの精度

図6にモデルの精度, 図7に損失関数を示す。accuracyは訓練データの精度, val_accuracyは検証データでの精度を表している。lossは訓練データの損失関数, val_lossは検証データでの損失関数を表している。

本研究では、試行した中で精度が94.7%と最も高く、損失関数が0.18で最小であった、学習回数の値が5、バッチ

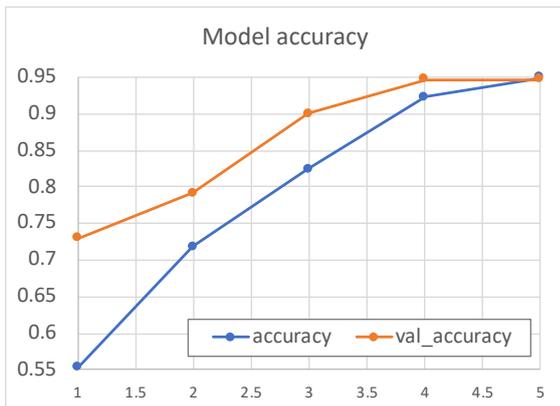


図 6 精度
Fig. 6 Model accuracy.

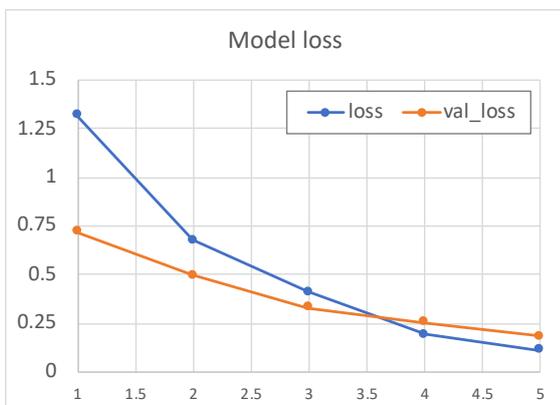


図 7 損失関数
Fig. 7 Model loss.

サイズの値が 32 の学習モデルを採用する。

6. 認識結果

6.1 静止画像を用いた認識結果

作成した学習モデルを用いて、Jupyter Notebook 上で下り階段、上り階段、平坦な床の静止画像を入力として認識を行った。認識に使用した深度画像と同じ場所のカラー画像を図 8～図 10 に示す。図 8 の深度画像を認識した結果、下り階段である確率が 99%であった。同様に、図 9 の深度画像を認識した結果、上り階段である確率が 90%と出力された。そして、図 10 の深度画像を認識した結果、平坦な床である確率が 58%、上り階段である確率が 33%、下り階段である確率が 7%であった。平坦な床である確率は高いとは言えないが、3 カテゴリの中では最大であった。

6.2 リアルタイム認識結果

6.1 と同じ学習モデルを用いて、リアルタイムで下り階段、上り階段、平坦な床の深度画像で認識を行った。RealSense を用いて入力画像 (640×480px, 15fps) を撮影し、学習した画像と同じ解像度になるように、240×240px にトリミン



図 8 下り階段の深度画像とカラー画像
Fig. 8 Images of downstairs.

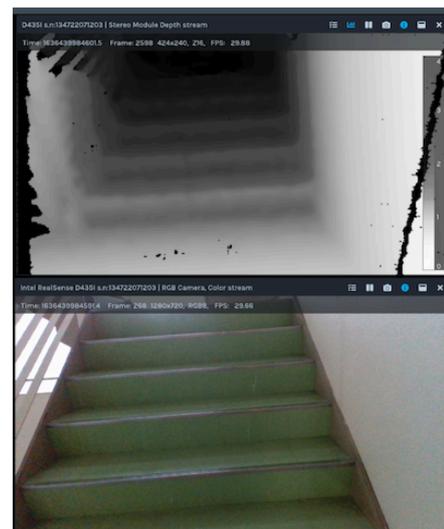


図 9 上り階段の深度画像とカラー画像
Fig. 9 Images of upstairs.

グ・リサイズして認識を行った。また、Anaconda Prompt にカテゴリごとの確率を出力した。リアルタイム認識をしている瞬間のスナップショットを図 11～図 13 に示し、判定不能時のスナップショットを図 14 に示す。

下り階段の認識で得られた画像は、図 11(a) が正しく認識された画像、(b) が不正解の画像である。図 11(a) の結果は、下り階段の確率が 100%であった。図 11(b) の結果は、下り階段の入力に対して、上り階段である確率が 99%～100%と出力された。

上り階段の認識で得られた画像は、図 12(a) が正しく認識された画像、図 12(b) が不正解の画像である。図 12(a) の結果は上り階段の確率が 92%～98%であった。図 12(b) の結果は、上り階段の入力に対して、下り階段である確率が 100%と出力された。

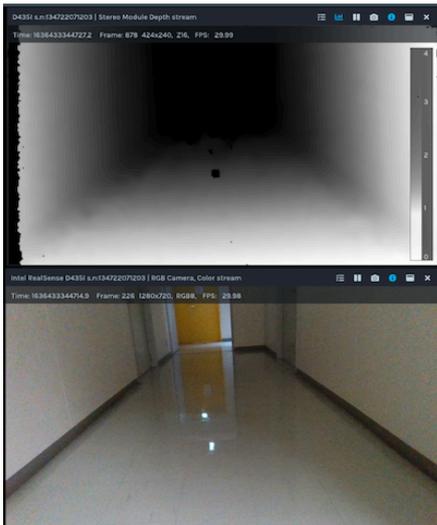


図 10 平坦な床の深度画像とカラー画像
Fig. 10 Images of corridor.

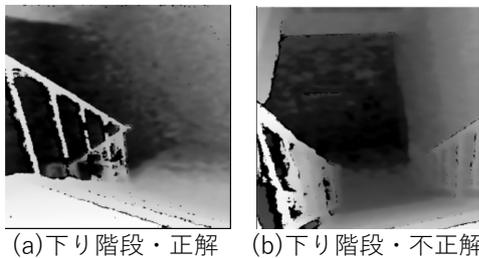


図 11 リアルタイム認識時の下り階段深度画像
Fig. 11 Depth image of downstairs.

平坦な床の認識で得られた画像は、図 13(a) が正しく認識されたの画像、図 13(b) が不正解の画像である。図 13(a) の結果は、平坦な床である確率が 89%~97%であった。図 13(b) の結果は、平坦な床の入力に対して、下り階段である確率が 99%と出力された。

また、図 14 の認識結果は、カメラを固定しているが、上り階段の確率が最大 95%、下り階段である確率が最大 72%であり、認識結果にばらつきが見られた。

全体的に学習用データと撮影角度や手すりの有無など、同様の条件によってリアルタイム認識をした場合、確率がおよそ 90%以上に達するが、学習用データに含まれていない映像における認識は、誤認識する確率が多い結果が得られた。平坦な床の認識においても、一部の学習用データと同様な画像に対しては床と認識をしたが、そうではない場合、床である確率が 0%となり、他カテゴリに誤認識する結果となった。

7. 考察

静止画像で認識をした結果、下り階段と上り階段は十分な精度が得られた。平坦な床は、確率が 58%と低いが認識できた。

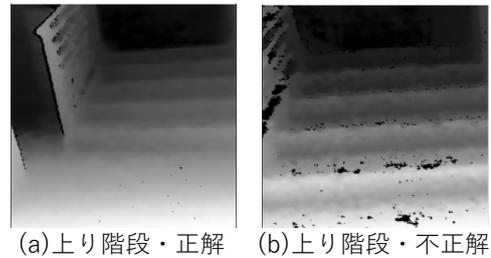


図 12 リアルタイム認識時の上り階段深度画像
Fig. 12 Depth image of upstairs.

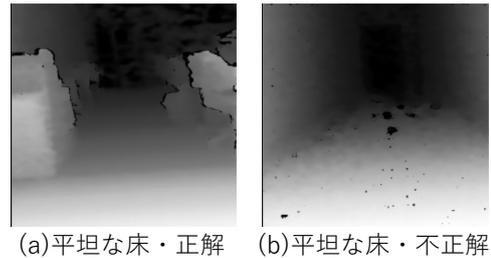


図 13 リアルタイム認識時の平坦な床深度画像
Fig. 13 Depth image of corridor.

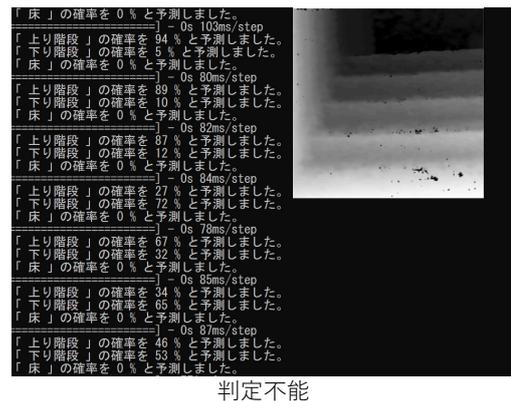


図 14 判定不能時の深度画像
Fig. 14 Unrecognizable depth image.

リアルタイム認識においては、学習用データと類似した深度画像の場合、下り階段と上り階段はおおよそ 90%以上の確率を得ている。しかし、撮影時の角度や手すりの有無など、学習用データに含まれていないような深度画像で認識を行った結果、誤認識が発生するため、精度が高いとは言えない。また、平坦な床の認識において、下り階段と上り階段のどちらかが 3 カテゴリのうち、最も確率が高いため、誤認識をしていることがわかる。すなわち、静止画像を用いた場合は認識ができていたが、リアルタイム認識では認識が正しく行われていないことがわかる。

以上の結果を踏まえて、学習用データと学習モデルに問題があると考えられる。学習用データは、1 枚の画像からトリミング位置をずらして 5 枚生成し、水増ししている。そのため、学習用データと似たような深度画像でしか認識しな

かったと推測する。深度画像は、階段までの距離、段の高さ、撮影時の角度、手すりの有無、壁の有無など、様々な条件で撮影し、学習用データのバリエーションを増やす必要がある。また、本研究で使用した学習モデルは、学習回数が5回と10回で最も精度の高いモデルを使用した。他の研究では学習回数を80回、100回として学習をしている。適切に学習回数とバッチサイズの値を設定する必要がある。

8. おわりに

本稿では、視覚障がい者の下り階段での転落防止のため、デブスカメラを用いた下り階段検出・通知システムを提案した。上り階段、下り階段、平坦な床の深度画像を用いてCNNによる学習モデルを作成した。作成したモデルによる静止画の認識を試みたところ、十分な精度で下り階段、上り階段、平坦な床の認識が可能であることがわかった。

さらに、リアルタイムでの下り階段認識を試みたところ、学習用データと類似した画像の場合、およそ90%以上の確率で認識できることがわかった。

デブスカメラを白杖に装着しての運用を考慮すると、システム全体の重量を見直す必要であると考えられる。白杖にデブスカメラと取り付け器具の重量が加わるため、使用時に利用者が負担を感じるケースが考えられる。軽量な取り付け器具を採用することで重さの改善ができる。

今後の課題としては、認識率の向上があげられる。学習データの見直しと適切な学習回数とバッチサイズを設定して学習モデルを作成する必要がある。特に、平坦な床の深度画像を用いた際に、平坦な床である確率が低い傾向があったため、学習用データを増やす必要がある。また、距離の通知をするために、RealSense から下り階段の一段目までの距離を測定する手法を検討する必要がある。

参考文献

- [1] 厚生労働省社会・援護局障害保健福祉部：平成28年生活のしづらさなどに関する調査(全国在宅害児・者等実態調査)結果，入手先(https://www.mhlw.go.jp/toukei/list/dl/katsu_chousa_c_h28.pdf) (2021).
- [2] 安部信行，橋本典久：視覚障害者の歩行環境整備のための歩行事故全国調査，八戸工業大学紀要，24巻，pp.81-92(2005).
- [3] 二宮睦樹，松島俊明：歩行者上の単眼視カメラによる周辺環境の認識，情報処理学会第74回全国大会，pp.2-427-2-428(2012).
- [4] 山口翔太郎，滝沢穂高，江崎修央，水野慎士：Kinect 白杖システムによる下り階段の認識，第11回情報科学技術フォーラム，pp.631-634(2012).
- [5] 大木郁登，穴戸英彦，亀田能成：歩行の安全確保に向けた単眼RGB画像からのバー型障害物検出，信学技報，vol.121，no.52，WIT2021-10，pp.48-53(2021).
- [6] 速水正樹，ヒルドミヒヤエル：視覚障害者の安全歩行を目的とするステレオ画像を用いた階段の検知：平成22年度情報処理学会関西支部支部大会，(2010).

- [7] 諏訪部純，大矢 晃久：3次元測域センサを用いた視覚障害者誘導システムの開発-階段を含めた誘導-，ロボティクス・メカトロニクス講演会2017，2A1-J02(2017)
- [8] 島川学，村上聡一郎，清田公保：RGB-Depth画像を用いた階段検出に関する研究，第31回ファジィシステムシンポジウム，pp.699-702(2015).
- [9] 原田佑二，滝沢穂高，大矢晃久，小林真，青柳まゆみ：深層学習とマルチチャンネル化を用いた視覚障がい者のための椅子認識に関する基礎的検討，信学技報，vol.121，no.52，WIT2021-8，pp.1-4(2021).
- [10] 今井健太，北原格，亀田能成：RGB-Dカメラと加速度センサを併用した歩行可能な平面領域の検出，第20回画像の認識・理解シンポジウム，pp.1-4(2017).
- [11] 工藤啓史郎，竹藤輝，根本昇，柳原康佑，Pornthep Sarakon，河野英昭：5G環境に適応した視覚障がい者向けの外出歩行支援デバイスの開発，第33回バイオメディカル・ファジィ・システム学会年次大会，pp.49-53(2020).
- [12] 門脇弘樹，菊池志乃，牟田口辰己：歩行軌跡における偏軌の特徴と予測-近似曲線を用いた分析，視覚リハビリテーション研究，第7巻，第1号，pp.3-12(2017).
- [13] 田中成典，安彦智史，塚田義典，平松祐樹：距離画像センサを用いた視覚障がい者の歩行支援システムに関する研究開発，映像情報メディア学会誌，vol.68，no.7，pp.J265-J27(2014).
- [14] 萩原晃，朝倉智之，和田直樹，白田 滋：歩行停止動作における二重課題の影響：理学療法科学，vol.30，no.5，pp.679-682(2015).
- [15] 交通エコロジー・モビリティ財団：視覚障害者用誘導警告ブロックに関する研究報告書(2000).
- [16] Nilposoft：Ralpha Image Resizer，<http://nilposoft.info/ralpha/index.html>(2021.11.5).