［研究助成（A）］

視覚環境を音声で表現可能な盲導犬ロボットに関する研究

Research on a Guide Dog Robot of Expressing Visual Environemnt by Voice

2211017

研究代表者　神奈川大学　工学部機械工学科　助　教　張　斌

［研 究 の 目 的］

視覚障害者の歩行ガイドを行う盲導犬は需要に対しては圧倒的に不足しており，代替手段が求められており，視覚障害者と協調歩行する盲導犬ロボットの研究開発が盛んに行われている。しかし，既存の盲導犬ロボットは道誘導機能を実現する際，ユーザの社会性のニーズへの考慮が不足で，視覚障碍者が周囲の環境や街中のシーンを楽しめることができていない。そこで，本研究では，ユーザ周囲の視覚環境をセンシングし，途中の風景や街中のシーンを理解し，音声で表現可能な盲導犬ロボットについて研究する。視覚障碍者で街中を散策し，楽しめることを目指す。盲導犬ロボットの社会受容性を向上することが期待できる。

［研究の内容，成果］

道誘導タスクを完成するとともに，視覚環境を音声で表現可能な盲導犬ロボットを開発し，その有効性を実機実験によって検証した。

（1）盲導犬ロボットの機構

開発した盲導犬ロボットを Fig. 1 に示す。移動ベース部（Vstone 社のメカナムローバー Ver2.1），制御部（PC 2 台），センシング部（距離センサと RGB-D センサ RealsenseD435 を 2 台）とハンドルで構成されている。ハンドルを上下 0° から 90° 及びと左右に 25° ずつ，動くように伸縮リードを使い，ベース部固定されている。機体の材料には軽量で強度の高いポリカーボネートとアルミフレームを使用している。さらに，ハンドルの高さは使用者の身長を考慮し，60[cm] から 80[cm] まで調節が可能である。

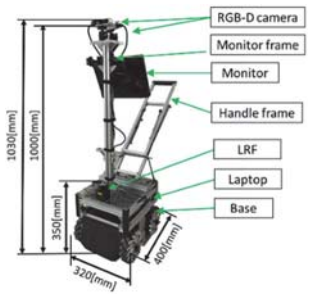


Fig. 1　Overview of guide dog robot

（2）地図生成

ロボットが環境を理解するためには，LRF により得られる距離情報を用い，周囲環境の 2 次元地図を生成する。地図の生成には，自己位置推定と環境地図生成を同時に行う SLAM（Simultaneous Localization and Mapping）という手法を使用する。3 次元環境を理解するためには，RGB-D カメラで取得した情報に対し，YOLO による物体認識を行い，認識した物体の 3 次元空間情報を 2 次元の地図に投影する。YOLO とはディープランニング手法に基づき，リアルタイムで物体を認識手法である。実際に生成した地図を Fig. 2 に示す。黒色の点はロ

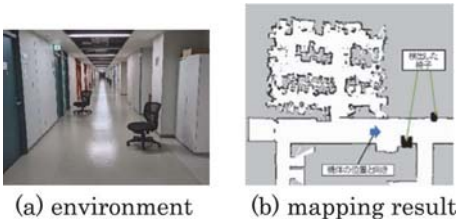


Fig. 2　Environmental mapping

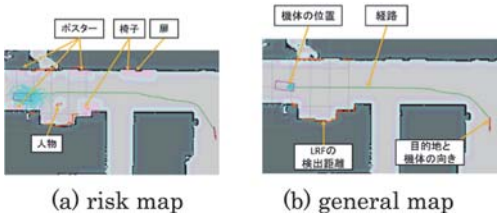


Fig. 5　Mapping with space risk

ボットが障害物として認識した不動領域，灰色は未知領域，白色は空領域を示している。認識した椅子の空間点群を地図に投影した結果は黒いブロックで示している。

(3) 空間リスクマップを基づいた移動制御

リスクマップとは領域の危険度を段階的に表した地図のことで，Fig. 3 の地図の水色はロボットが侵入した際衝突する高リスク領域を表し，紫色は衝突の可能性がある中リスク領域を表す。壁や人物など衝突が許されない領域に高いリスクを設定し，その近くには中等のリスク，物体の遠くにある何もない空間はリスクなしとする。人物を検知した場合は，人物の位置と向きの情報をもとにパーソナルスペースと会話状態の判断を行い，リスクマップに反映する。人物の体の向きの推定は，Realsense によって認識した人物のスケルトンモデルの右肩から左肩へのベクトルの法線ベクトルで判断する。位置座標は，スケルトンモデルを地面に投影される 2 次元座標系に準ずる。Fig. 4(a) に 1 人の場合を示す。人物の前方に長く，全長 117[cm]，最大幅 80[cm]，先端幅 32[cm] の卵形の領域になる。Fig. 4(b) に示す会話中の人物は，各人を向かい合わせに接続した形状である。空間のリスクは人物からの距離増加につれ減少していく。実験環境（Fig. 2 に示す）にある扉やポスターの前などの潜在的占有空間のリスクを指定し，生成したリスクマップを Fig. 5 に示す。

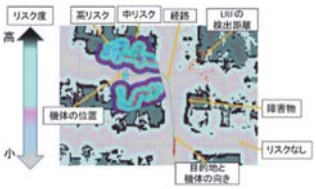


Fig. 3　Risk map of objects

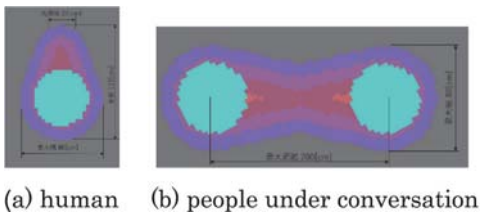


Fig. 4　Risk map of human beings

その後，ダイクストラ法を用いて最短経路を生成する。ダイクストラ法とはノードとエッジで構成されたグラフ上において，2 つのノード間の最短経路を計算するアルゴリズムである。ロボットの移動可能な空間に配置されたノードに対してリスクマップのリスク情報に応じたコストを設定し，そのコストの合計が最小となる経路を計算をする。ロボットが最短経路に従い移動する。移動制御はポテンシャル場を用いて実現する。ポテンシャル場法とは，目的地と障害物に引力と斥力のポテンシャル場を定義し，統合ポテンシャル場に従い移動させることで，障害物を回避しつつ，目的地へと移動可能な制御手法である。ロボットはユーザから一定の距離を維持するため，ユーザ前にソーシャル距離のところで引力場を設定する。人物のリスクマップも斥力場の一種として，統合ポテンシャル場に導入し，ロボットの移動速度を制御する。

Fig. 6　Guide scene using robot

実験中では成人男性が目を隠してロボットのハンドルを持ち，ロボットの案内に従い目的時まで移動を行った。案内中のシーンを Fig. 6 に示す。成人男性は潜在的占有空間や壁などを適切に避けながら目的地まで移動することができた。

(4) 視覚環境の音声表現システム

ロボットの周辺の視覚環境を音声で視覚障碍者に表現するシステムは Fig. 7 に示す。

① RGB-D カメラを起動し，環境のカラー画像を取得し，視覚環境の認識を始める。説明対象物は，取得したカラー画像からディープラーニング手法 YOLOv3 を用いて，物体の種類を認識する。YOLOv3 は，MS-COCO というデータセットで事前に学習した 80 種類の物体を認識し，バウンティボックスという物体領域の外接四角形を出力する。本研究では，視覚障碍者に案内する際の説明対象物を「person」「bicycle」「motorbike」「car」「bus」「truck」「traffic light」とする。距離情報に基づき，YOLO v3 で出力したバウンディングボックスの中心座標に対応した距離を測定する。そして，認識した物体との距離が 5 m 以内の場合のみを説明する対象とする。

② その後，事前に作成した地図と経路をもとに視覚障碍者を目的地に誘導するための案内を開始する。案内を開始する際，視覚障碍者に盲導犬ロボットが移動することを知らせるために「go straight」などの音声によりアナウンスをする。移動途中でユーザの安全を確保するため，曲がるべき場所（Fig. 8 の①～⑥のところ）で

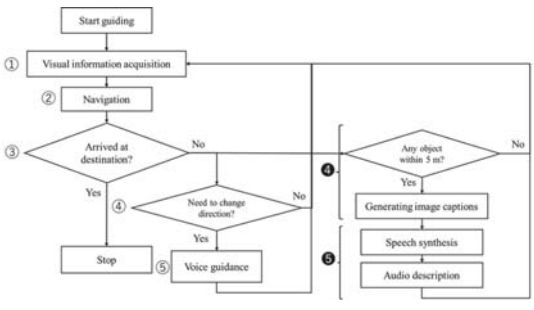


Fig. 7 Voice expression system of visual environment for the robot

Fig. 8 Moving path and motion changing points during guide task

音声により道案内をする。

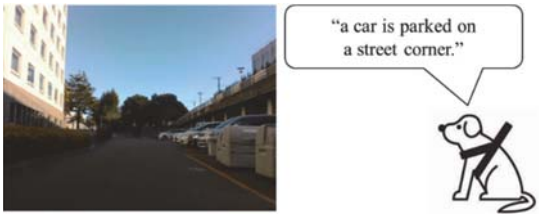
③ 案内開始後，ロボットが目的地に着いたかを判断する。目的地に到着している場合は，「We have arrived at the destination」と音声によりアナウンスをして案内を終了する。また，到着していない場合は，④（道案内をする処理）と ❹（視覚環境の説明）に分かれて並行処理で実行していく。

④ 事前に作成した地図をもとにロボットの移動状況を判断する。例えば，曲がるべき場所の場合，地図情報より判断し「Turn right」や「Turn left」などと道案内のアナウンスを行う。

❹ 物体認識を行い，あらかじめ指定した説明対象物が 5 メートル以内にある場合のみを用いて，説明文を生成していく。カラー画像から説明文を生成するには，CNN（Convolutional Neural Network）で画像から特徴量を抽出し，LSTM（Long Short Term Memory）で説明文を生成する NIC（Neural Image Caption）を用いる。Fig. 9～Fig. 12 に盲導犬ロボットが街中で視覚障碍者に説明した場面と生成された説明文を示す。Fig. 9 は，人間が歩いている様子の場面である。この場面に対して「a man is walking down a street with a cell phone.」と説明文が生成された。Fig. 10 は，車が駐車されている場面である。この場面に対して「a car is parked on a street corner.」と説明文が生成された。Fig. 8 は，建物の前に車が駐車されている場面である。この場面に対して「a car is parked in front of a building.」と説明文が生成

(a) Scene of human walking　　(b) Description

Fig. 9　Sound explanation scene of human walking



(a) Scene of cars are parked　　(b) Description

Fig. 10　Explanation scene where cars are parked

(a) Scene of the traffic light　　(b) Description

Fig. 11　Explanation scene of the traffic light

(a) Scene of the bus stop　　(b) Description

Fig. 12　Sound explanation scene of the bus stop

された。Fig. 11 は，信号機がある風景である。この場面に対して「a traffic light on a pole near a street.」と説明文が生成された。Fig. 12 は，バス停のある風景である。この場面に対して「a man is standing in front of a bus stop.」と説明文が生成された。

❺ 「gTTS（Google Text-to-speech）」という Google Translate's text-to-speech API を利用し，❹で生成された説明文を音声合成により MP3 の音声ファイルを作成し，再生することで視覚障碍者に視覚環境を伝達する。

また，本研究の有効性を検証するために街中の風景とその説明文を男子大学生 24 人に見せ，生成された説明文が適切であったかの 5 段階によるアンケート調査を行う。アンケートの問いは，「ある方は，それぞれの場面をみてこのように回答しています。この解答は適切だと思いますか。」という問いで行った。すべてのシーンに対し，適切と回答した割合が 70% 以上と達しているため，本盲導犬ロボットは視覚障碍者に視覚環境を概ね適切に説明できたといえる。よって，視覚障碍者が普段見ることのできない視覚環境を把握でき道散策を充実させる可能性が期待できると言える。

本研究では，視覚環境を音声で表現可能な盲導犬ロボットを開発し，空間リスクマップに空基づいた動的環境下での自律移動ロボットの制御手法を提案した。ロボットが安全確保かつ自然に移動するため，物に近づきすぎないように，それから人に不快感を与えないようにパーソナルスペースを避けながら自律移動することができた。また，視覚障碍者を安全に誘導しながら，街中の風景を認識し，説明文を生成し音声で表現するシステムを提案した。道案内実験により，ロボットの自律移動，物体認識，説明文生成，音声合成機能を確認し，開発した盲導犬ロボットの有効性が検証された。

［今後の研究の方向，課題］

今後は物体の特性をさらに分析し，物体の種類によって周りの空間の使用状況，占有強度を定義し，ロボットの動きに影響することで，より人間らしい自律移動の実現を目指す。それから，人物の前に通過するようなケースが避けられない場合は人間のように「すみません」とか声をかけてから移動することで，人間に優しいロボットの実現ができる。また，より多くの日

常生活場面を認識し，音声で説明することで，盲導犬ロボットの実用に向けて検証を行う。

［成果の発表，論文等］

［1］ 張斌，仙石龍司，中村友昭，金子正秀，林憲玉，“空間リスクマップに基づく自律移動ロボットの移動制御，”ロボット学会学術講演会，1F4-04，Sep. 2021.

［2］ 張斌，金子正秀，林憲玉，“物体の潜在的な占有空間を考慮した自律移動ロボットの経路決定，”情報処理学会全国大会，1C-01，Mar. 2022.

［3］ 市川玲也，張斌，林憲玉，“盲導犬ロボットによる視覚環境情報をユーザに音声で伝達するシステムの開発，”ロボティクス・メカトロニクス講演会，2P1-G04，Jun. 2022.（予定）

［4］ 張斌，奥津幹也，落合凜，田山萌，林憲玉，“手持ちしやすい盲導犬ロボットの開発及び空間リスクマップを基づいた移動制御，”電気学会 電子・情報・システム部門大会，Aug. 2022.（予定）

［5］ R. Ichikawa, B. Zhang and H. Lim, “Voice Expression System of Visual Environment for a Guide Dog Robot,” The 8th International Symposium on System Security, Safety and Reliability, Oct. 2022.（accepted）

［6］ B. Zhang, M. Okutsu. R. Ochiai, M. Tayama and H. Lim, “Robust Motion Control for a Considerate Guide Dog Robot,” Electronics, 2022.（査読中）

［7］ 市川玲也，張斌，林憲玉，“視覚環境を音声で表現可能な盲導犬ロボットに関する研究，”電気学会論文誌（電子・情報・システム部門誌），2022.（査読中）