

# 視覚障害者のための買い物支援システムの開発

大久保 達矢<sup>\*1</sup> 磯山 直也<sup>\*1</sup> 酒田 信親<sup>\*1</sup> 清川 清<sup>\*1</sup>

## Development of a Shopping Support System for Visually Impaired

Tatsuya Okubo<sup>\*1</sup>, Naoya Isoyama<sup>\*1</sup>, Nobuchika Sakata<sup>\*1</sup> and Kiyoshi Kiyokawa<sup>\*1</sup>

**Abstract** – It is generally difficult for a visually impaired person to find an item at a shop that he or she wants to buy by himself or herself, and asking somebody for assistance is not always feasible. To address the problem, we have designed a shopping support system for visually impaired. Our system features three functions; (1) creation of an in-store environment map, (2) prediction of an in-store environment map for efficient navigation, (3) path-finding and navigation of a user. In this paper, we report on the detail of the system design and current progress.

**Keywords** : Assistive Technology, Visually Impaired, Shopping Support, Deep Learning

### 1 はじめに

厚生労働省によると、平成18年時点で国内の視覚障害者数は約31万人に上り、「外出する」や「日常の買い物をする」等の動作で介助を必要とする割合が高い[1]。同省の「視覚障害者の移動支援の在り方に関する実態調査」では、ガイドヘルパーとの主な外出は「団体等の会議・催し」および「買い物」であることから視覚障害者の単独での買い物は困難と考えられる[2]。しかし、市町村の条例により、ガイドヘルパーの利用は一定期間中で回数や時間に上限が定められており、ガイドヘルパーを気軽に利用できないという問題がある。また、筆者らが熊本県立盲学校の教員である視覚障害者2名（全盲1名と弱視1名）に買い物に関するヒアリング調査を行った結果、全盲の方は単独での買い物は不可能なため店員や介助者の助けが必要であると回答した。弱視の方は行き慣れたコンビニエンスストア程度の規模の店舗であれば単独での買い物が可能であるが、初めて訪れる店舗での買い物は難しいと回答した。また、両者から可能ならば単独で気軽に買い物をしたいという回答が得られた。

従来、視覚障害者の買い物を技術的に支援する研究がいくつか存在する。例えば、ShopTalk [3] はユーザがバーコードリーダーとコンピュータ等を携帯し、バーコードリーダーで商品棚のバーコードをスキャンし、目的の商品を探索する。また、RoboCart [4] はカートにRFIDタグを取り付け、ユーザの自己位置を特定し商品まで誘導する。しかし、これらの研究は商品を自ら探索する必要性や、RFIDリーダーの設置や店内地図の

提供などの店舗の協力の必要性があるために初めて訪れる店舗での利用は困難である。

そこで、我々は視覚障害者が初めて訪れる店舗でも他人の協力を要することなく、単独で買い物が可能な買い物支援システムの開発に取り組んでいる。本稿では、我々が提案する買い物支援システムの概要と、現在の進捗及び今後の計画について報告する。

### 2 提案システムの概要

我々は、店舗の協力なしに単独での買い物を実現する支援システムには、3つの段階が存在すると考えている。1段階目は目的とする商品が存在する商品棚付近までの誘導、2段階目は商品棚の目の前までの誘導、3段階目は商品棚から目的の商品を手取るまでの誘導である。現在、3段階目については、スマートグラスと触覚提示手袋に設置したカメラから一般物体認識により商品を特定し、手に装着した触覚提示手袋の振動刺激により、目的の商品を手取るまで誘導するシステムが存在する[5]。しかし、1, 2段階目についてのシステムは存在しない。そこで我々は、まず1段階目である目的とする商品棚付近までユーザを誘導することについて取り組む。具体的には、商品棚の位置とカテゴリを記録した店内環境地図の作成、店内環境地図の予測、経路誘導の3つの機能を持つシステムを設計し、解決を試みる。以下にそれぞれの機能について詳述する。

店内環境地図の作成は、図1(a)のようなレイアウトの店舗に対して、同図(b)のように商品棚の位置とそのカテゴリを記録した二次元の環境地図を作成する。この店内環境地図作成のために、商品棚のカテゴリ認

<sup>\*1</sup>奈良先端科学技術大学院大学

<sup>\*1</sup>Nara Institute of Science and Technology

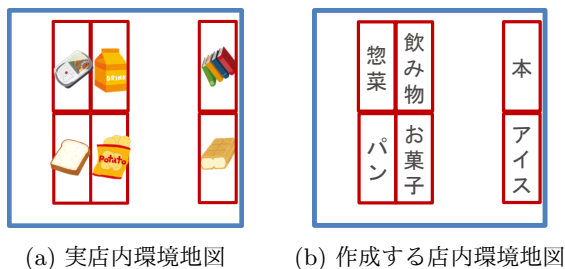


図 1: 店内環境地図の例

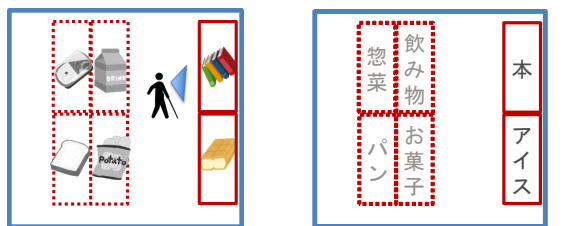


図 2: 店内環境地図の予測

識と、ユーザの自己位置および商品棚や壁の位置の記録を行う必要がある。そこで、我々は機械学習を用いた商品棚画像からのカテゴリ認識と、Microsoft 社の HoloLens に搭載された SLAM (Simultaneous Localization and Mapping) を用いた自己位置および商品棚や壁の位置の認識を行うことで解決する。HoloLens は MR (Mixed Reality) 用の光学透過型 HMD であるが、今回はディスプレイ機能は用いず、画像の取得や SLAM を行うためのセンサとしてのみ利用する。また、機械学習によるカテゴリ認識と二次元の店内環境地図の作成については、HoloLens への処理負荷が大きいため、サーバ上で処理する。

店内環境地図の予測は、例えば、図 2(a) のように本とアイスのカテゴリを観測すると、観測結果をもとに同図 (b) のように未観測領域のカテゴリの予測を行う。店内環境地図の予測により、ユーザの探索負荷を軽減し、効率的な経路誘導を目指す。予測は、実店舗における各カテゴリ間の距離関係を調査し、カテゴリを予測したい位置を入力として与えると、予測カテゴリを求めるためのコスト関数を用いて予測する。

経路誘導は、目的のカテゴリが存在する商品棚付近までユーザを誘導する。カテゴリごとの誘導地点の地図は、観測と予測の店内環境地図を組み合わせて作成する。誘導地点は観測結果を優先するが、カテゴリが観測されていない場合は、予測結果を利用する。この地図を用いて、ユーザから目的のカテゴリの誘導地点までの経路探索を行い、振動ベルトもしくは HoloLens の音声案内を用いて誘導する [6]。

以降、店内環境地図の作成を 3 節、店内環境地図の予測を 4 節、商品棚付近までの経路誘導を 5 節で詳述



図 3: 各クラスの画像例

する。

### 3 店内環境地図の作成

#### 3.1 商品棚画像からのカテゴリ認識

##### 3.1.1 データセット

本システムでは、商品棚画像から商品カテゴリを認識するために CNN を利用する。CNN を学習するために、商品棚画像のデータセットを作成する必要がある。そこで、我々は学内のコンビニエンスストアの協力を得てデータセットを作成した。今回、商品カテゴリを本 (book), パン (bread), インスタント食品 (instant food), 日用品 (daily), 飲み物 (drink), アイス (ice), 惣菜 (side dish), お菓子 (snack) の 8 クラスに設定した。各クラスの画像例を図 3 に示す。データセット撮影は HoloLens 内蔵のカメラを用いて行った。撮影解像度は  $1216 \times 684$  pixel であり、これを  $50 \times 50$  pixel に縮小して学習に用いた。各クラスの画像データ数は、本: 579, パン: 945, 日用品: 892, 飲み物: 2477, アイス: 848, インスタント食品 1019, 惣菜: 1186, お菓子: 2798 である。学習方法については、文献 [7] に示す通りである。

##### 3.1.2 学習結果

CNN である 18 層の ResNet を用いて学習した結果、精度: 99.81%, 損失値: 0.06 と非常に高い精度の認識モデルが得られた [8]。本システム構築のために、100% の認識精度を持つモデルを作成することが理想である。

しかし、商品棚を通り過ぎる際に複数回カテゴリ認識を行い、投票数の最も多いカテゴリに決定するなどの処理を行うことで、今回得られた精度でも十分利用できると考えている。

### 3.2 店内環境地図のリアルタイム作成

#### 3.2.1 店内環境地図の概要

店内環境地図は二次元のグリッド状の地図で、グリッドは縦横 0.2m 間隔で店内全体を分割している。グリッド内の各区画をセルと呼ぶ。提案システムは各セルに対してカテゴリを割り当てることで店内環境地図を作成する。今回、カテゴリは先述した商品棚画像から認識を行う 8 カテゴリに加え、レジ、入口、歩行可能領域、ユーザ、未計測領域の計 13 カテゴリとする。レジは今後、商品棚画像から認識を行うカテゴリ（以降、認識カテゴリ）に加える予定である。入口はシステム起動時の位置とし、原点とする。歩行可能領域は、障害物が存在しない歩行が可能な領域である。ユーザ位置は SLAM により得られた自己位置を示す。未計測領域はまだ観測されていない、どのカテゴリが存在するか未知な領域である。

各セルは、各カテゴリの投票数を保持しており、投票数が最も多いカテゴリとして分類される。各セルの投票数の初期値は、未計測領域の投票数を 0.1 とし、他のカテゴリの投票数は 0 とする。カテゴリの認識により投票が行われた場合は、そのカテゴリの投票数に 1 を加算する。ただし、ユーザが位置するセルは、ユーザが存在する間、そのセルのユーザへの投票数を他のカテゴリの投票数よりも一時的に多くすることで、必ずユーザ位置としてカテゴリ分類されるようにする。

次に店内環境地図の作成手順について述べる。まず、ユーザが装着した HoloLens から正面方向にレイをキャストする [9]。レイが環境地図に衝突すると、HoloLens はカメラから RGB 画像を取得し、SLAM によって得られる自己位置およびレイが衝突した座標と共に処理サーバに送信する。処理サーバでは、受信した画像のカテゴリ認識を行った後、レイが衝突した座標に対応する店内環境地図上のセルに投票を行う。このとき、ユーザ位置からレイが衝突した座標までの空間は物体が存在しないため、その区間のセルは歩行可能領域に投票を行う。

#### 3.2.2 仮想店舗での動作検証

作成した仮想店舗（図 4）を用いて、店内環境地図のリアルタイム作成の動作検証を行う。同図は、壁面と棚面に商品棚画像のテクスチャを貼り付けた簡易的な仮想店舗である。この仮想店舗は図 5 のような、実際のコンビニエンスストアを模倣したレイアウトで、大きさは  $x$  方向 6.0m、 $y$  方向縦 7.3m とした。作成する店内環境地図は、システム起動時の位置である入口を

表 1: カテゴリと色の対応表

色	カテゴリ	色	カテゴリ
青	本	オレンジ	お菓子
緑	パン	紫	レジ
赤	インスタント食品	黒	入口
黄緑	日用品	茶	歩行可能領域
黄	飲み物	黒	ユーザ位置
紫	アイス	白	未計測領域
茶	惣菜		

原点とし、負の実数値に対応できるように、 $x, y$  方向に  $-1\text{m}$  から  $7\text{m}$  までの  $8\text{m}$  の領域を定義しており、仮想店舗よりも大きい領域とした。また、定義した領域を  $0.2\text{m}$  間隔で分割するため、セル数は  $40 \times 40$  個存在する。仮想店舗での店内環境地図作成の途中経過を図 6 に示す。図中においてセルの色は各カテゴリを示しており、色とカテゴリの対応を表 1 に示す（本稿では以降もこの色の対応を利用する）。同図において、未計測領域のセルは、レイの衝突もしくはレイの通過が一度も無く、投票が行われなかったセルである。したがって、仮想店舗外の領域のセルも未計測領域となる。仮想空間における店内環境地図の作成精度は、商品棚画像からのカテゴリ認識の精度に依存する。今回、仮想環境用のデータを再度作成し学習させ、精度: 97.45%、損失値: 0.28 となった認識モデルを使用した。図 6 から、仮想店舗上での店内環境地図は、図 5 と比較して、再現度高く店内環境地図が作成できていることが確認できる。

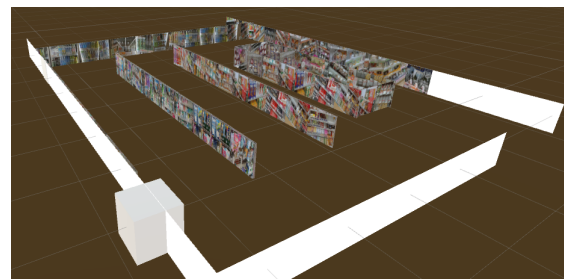


図 4: 仮想店舗

#### 3.2.3 カテゴリの重心の算出

本システムでは、作成途中の店内環境地図から定期的に各認識カテゴリの重心を求める。求めた重心は、ユーザの誘導地点の決定や店内環境地図の予測に利用する。カテゴリの重心は、同一カテゴリが連続する領域の中心位置とする。重心を求めるために、同一カテゴリが連続して存在することが望まれるが、実際にはカテゴリ認識の誤りを含むため、同一カテゴリが断続して認識される場合がある。そこで、縦方向と横方向にノイズ除去を適用した地図から重心を求める。例え

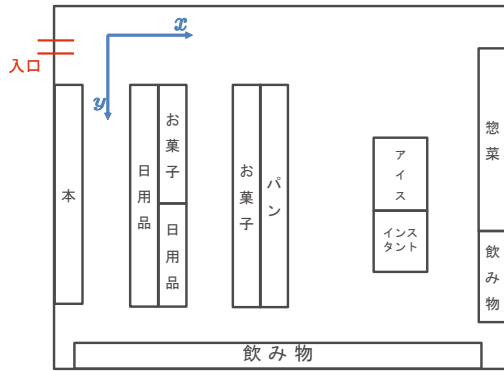


図 5: 仮想店舗のレイアウト

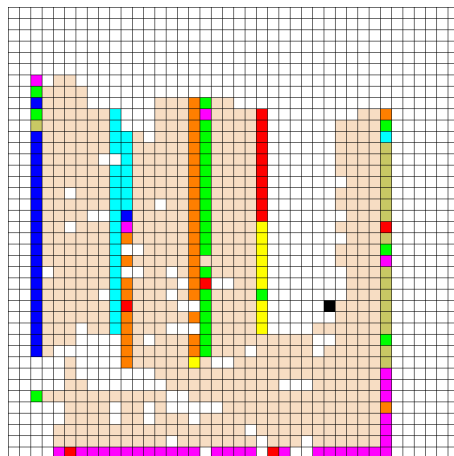
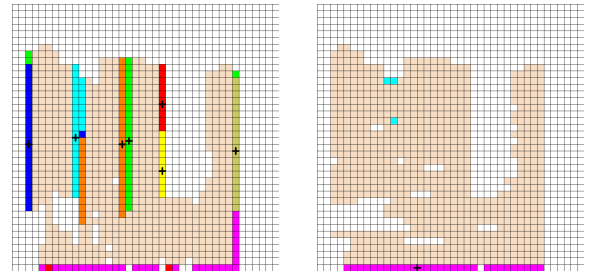


図 6: 店内環境地図作成の途中経過

ば、縦方向のノイズ除去は、基準のセルを設定し、上下方向に  $M$  セル内のカテゴリを調べ、計  $2M+1$  個のセルの中で最も多く含まれるカテゴリに基準のセルを修正する。横方向は左右方向に同様の処理を行う。その後、縦方向にノイズ除去した地図は縦方向に、横方向にノイズ除去した地図は横方向に、最も同一カテゴリが連続する領域を求める。これを、認識カテゴリ分行い、両地図ごとに各カテゴリの重心を調べる。店舗には同一カテゴリは複数存在しないという制約を設けているため、両地図でカテゴリが重複する場合は、連続する領域が大きい方の重心を、そのカテゴリの重心とする。縦方向のノイズ除去の結果を図 7(a)、横方向のノイズ除去の結果を同図 (b) に示す。両図とも、求めた各カテゴリの重心を十字で可視化しており、図 8 はそれらを合わせた結果である。図 7(a) で、飲み物のカテゴリが横方向に存在しているのは、縦方向のノイズ除去の際に、基準のセルから上下に  $M$  個のセルが存在しない場合は、ノイズ除去を行っていないためである。



(a) 縦方向のノイズ除去結果 (b) 横方向のノイズ除去結果

図 7: ノイズ除去結果

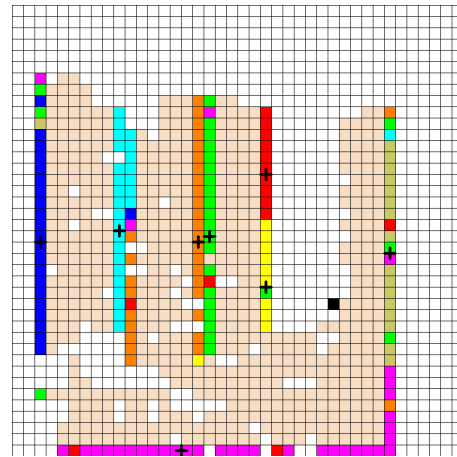


図 8: カテゴリの重心算出結果

#### 4 店内環境地図の予測

##### 4.1 コスト関数の定義

店内環境地図の各セルのカテゴリを予測するために、各カテゴリ間の距離を用いたコスト関数を定義する。例えば、図 5 の店舗レイアウトの場合、本と惣菜は  $x$  方向への距離の差は大きい、 $y$  方向への距離の差は小さい。このような  $x, y$  方向への 2 つの距離を用いたコスト関数を定義し、各セルのカテゴリ予測に用いる。具体的には、各セルについて各認識カテゴリのコストを計算し、最小のコストとなるカテゴリをそのセルの予測カテゴリとする。

コスト関数を定義する前に、各変数を定義する。まず、コストを算出する 2 つのカテゴリを  $i, j$  とし、 $i, j$  の  $x$  方向への一般的な重心間距離を  $\bar{x}_{ij}$ 、 $y$  方向への一般的な重心間距離を  $\bar{y}_{ij}$  とし、それぞれ 0 から 1 の間に正規化した実数値とする。 $\bar{x}_{ij}, \bar{y}_{ij}$  は、5 店舗のコンビニエンスストアの各カテゴリの重心間距離を実際に歩測で調査し、平均した。また、システムがリアルタイムに作成する店内環境地図から求めたカテゴリ  $j$  の重心を  $g_{x_j}, g_{y_j}$  とする。カテゴリ  $i$  と  $j$  のコストに関する信頼度を  $w_{ij}$  とする。 $w_{ij}$  は、 $i, j$  間の距離関係が店舗ごとで変化が小さく、距離関係が保たれている場合は、値を大きくする。店舗ごとで変化が大きい場合は、値を小さくする。重みの求め方については次項

で説明する。これらの変数を用いたコスト関数を (1) 式のように定義する。  $N$  は認識カテゴリ数である。同式の  $X, Y$  をそれぞれ, (2), (3) 式に示す。 (2), (3) 式に含まれる,  $W, H$  はそれぞれ, 店舗の  $x, y$  方向への長さを示す。したがって,  $|\frac{x-g_{x_j}}{W}|, |\frac{y-g_{y_j}}{H}|$  はそれぞれ, 0 から 1 の実数値となる。  $cost_i$  をレジを除く認識カテゴリ分算出し, 最小となるカテゴリ  $i$  を  $x, y$  に対応するセルの予測カテゴリとする。ただし,  $i = j$  の場合は, 同カテゴリの位置関係のため,  $w_{ij} = 0$  とすることでコストの計算に影響を与えないようにする。

$$cost_i(x, y) = \sum_{j=0}^{N+1} w_{ij} \sqrt{X^2 + Y^2} \quad (1)$$

$$X = \overline{x_{ij}} - \left| \frac{x - g_{x_j}}{W} \right| \quad (2)$$

$$Y = \overline{y_{ij}} - \left| \frac{y - g_{y_j}}{H} \right| \quad (3)$$

#### 4.2 重みの算出

$i, j$  の重心間距離の分散が大きいほど, 店舗によって位置関係にばらつきがあるため, 信頼度である重みの値を小さくする必要がある。  $\overline{x_{ij}}, \overline{y_{ij}}$  の分散を用いることで, カテゴリ間距離の店舗によるばらつきを調べることができ, これを重み  $w_{ij}$  に用いる。具体的には,  $\overline{x_{ij}}, \overline{y_{ij}}$  の分散を, それぞれ  $V_{\overline{x_{ij}}}, V_{\overline{y_{ij}}}$  とすると, 重みは  $w_{ij} = 2^{-(V_{\overline{x_{ij}}} + V_{\overline{y_{ij}}})}$  で求める。歩測した 5 店舗の各カテゴリの重心間距離の分散を計算し, 求めた重みを行ごとの和が 1 になるように正規化した結果を表 2 に示す。同表では, 全体的に「インスタント食品」, 「アイス」は重みが小さい。実際, これらのカテゴリは店舗ごとに配置が大きく異なっているため, 予測の際に算出される「インスタント食品」と「アイス」の  $X, Y$  は信用できない。したがって, 考案した店舗ごとの各カテゴリ間距離の分散を用いた重みの算出方法により, 目的通り重みが小さくなっている。

#### 4.3 予測結果

予測実行時にカテゴリの重心が求まっている重みを用いて, 表 2 の各行ごとに再度正規化を行う。図 9 のように, 本の重心位置が観測された際の予測を図 10 に示す。また, 図 6 のレジを除く全カテゴリの重心観測後の予測結果を図 11 に示す。両予測結果では, 図 10 の方が店舗レイアウト (図 5) をより正確に再現できており, 観測初期段階での予測結果としては精度が高い。また, 観測が進むことによって店内環境地図が出来上がってくると, 予測結果は実際の店舗レイアウトと異なる地図になっている。しかし, 観測済みのカテゴリについては観測で得られたカテゴリの重心位置へ誘導するため, システムの動作に悪影響は与えない。

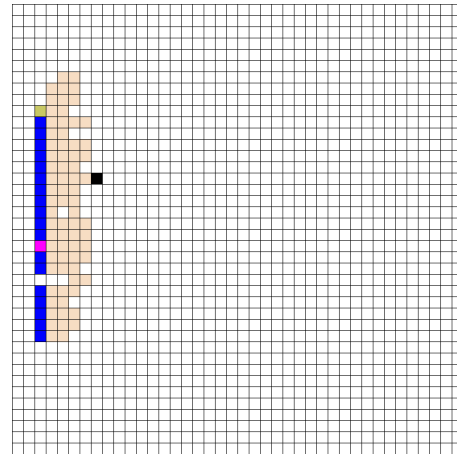


図 9: 観測序盤の店内環境地図

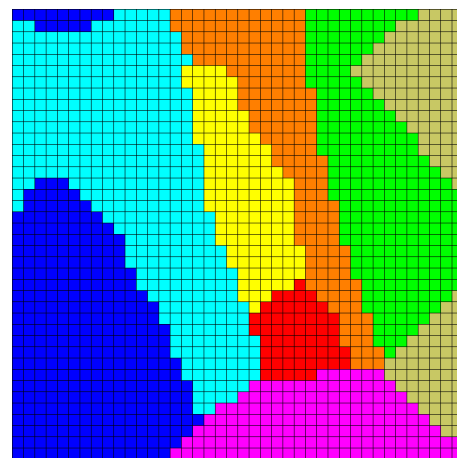


図 10: 観測序盤の予測結果

## 5 商品棚付近までの経路誘導

### 5.1 経路探索

ユーザの現在地から目的とする商品が存在する棚付近までの経路を探索する問題点として, 店内環境地図の通路である歩行可能領域と, 障害物や商品棚などの歩行できない領域が未知であることが挙げられる。そこで, 現在地と目的地が既知でかつ通路と障害物が未知な経路探索アルゴリズムとしてマイクロマウス競技で有名な足立法の利用を検討する [10]。

### 5.2 経路誘導

経路の誘導は前述した, 音声案内もしくは振動ベルトによる手を塞がない直感的な誘導を考えている。振動ベルトでの振動による誘導と開放型イヤホンでの音声による誘導については, 今後検討する必要がある。

## 6 おわりに

本稿では, 視覚障害者が単独での買い物が難しいという問題を解決するために, 目的とする商品が存在する棚付近までの誘導を実現する買い物支援システムを提案し, 進捗について報告した。商品棚画像からの商品カテゴリ認識モデルを作成し評価を行った結果, 優

表 2: カテゴリ間の重み

	本	パン	インスタ タント 食品	日用品	飲み物	アイス	惣菜	お菓子	レジ	入口
本	0.000	0.001	0.000	0.212	0.491	0.000	0.272	0.010	0.003	0.011
パン	0.000	0.000	0.001	0.030	0.005	0.000	0.005	0.647	0.297	0.014
インスタント食品	0.000	0.019	0.000	0.018	0.839	0.031	0.022	0.059	0.011	0.000
日用品	0.194	0.040	0.001	0.000	0.285	0.000	0.037	0.079	0.052	0.312
飲み物	0.111	0.002	0.008	0.071	0.000	0.000	0.238	0.109	0.073	0.387
アイス	0.000	0.000	0.081	0.022	0.000	0.000	0.000	0.897	0.000	0.000
惣菜	0.155	0.004	0.001	0.023	0.599	0.000	0.000	0.031	0.057	0.130
お菓子	0.004	0.371	0.001	0.034	0.193	0.006	0.022	0.000	0.345	0.024
レジ	0.001	0.135	0.000	0.018	0.102	0.000	0.032	0.273	0.000	0.439
入口	0.003	0.005	0.000	0.091	0.455	0.000	0.061	0.016	0.369	0.000

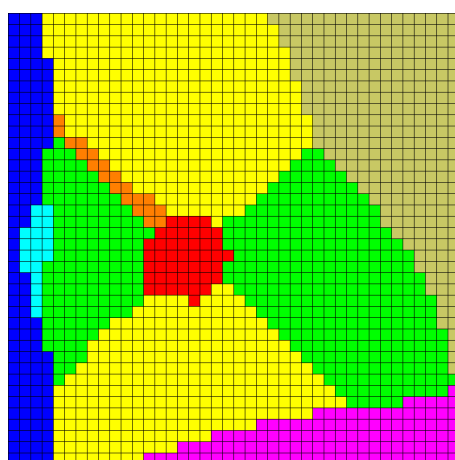


図 11: 観測終盤の予測結果

れた認識精度が得られた。また、仮想店舗においてリアルタイムで高精度な店内環境地図の作成を実現した。今後は、実環境での動作検証を行う必要がある。さらに、各カテゴリ間の距離関係を用いたコスト関数による店内環境地図の予測では、観測初期段階から精度の高い予測が行える感触が得られた。

今後は、各カテゴリの誘導地点を明確化した地図を作成し、足立法を用いた経路探索により、予測が有効であるか検証を行う。システム完成後に、仮想環境および実環境において、提案する買い物支援システムの有効性を検証する被験者実験を行う。

参考文献

[1] 厚生労働省, 平成 18 年身体障害児・者実態調査結果, <https://www.mhlw.go.jp/toukei/saikin/hw/shintai/06/dl/01.pdf> (last access: 2019 年 12 月 16 日)

[2] 厚生労働省, 厚生労働省平成 26 年度障害者総合福祉推進事業, 視覚障害者の移動支援の在り方に関する実態調査, <https://www.mhlw.go.jp/file/06-Seisakujouhou-12200000-Shakaiengokyokushougaihokenfukushibu/0000099353.pdf> (last access: 2019 年 12 月 16 日)

[3] John Nicholson, Vladimir Kulyukin, and Daniel Coster, “ShopTalk: Independent Blind Shopping

through Verbal Route Directions and Barcode Scans,” The Open Rehabilitation Journal 2.1, pp. 11–23, 2009.

[4] Vladimir Kulyukin, Chaitanya Gharpure, and John Nicholson, “Robocart: Toward Robot-Assisted Navigation of Grocery Stores by the Visually Impaired,” 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005.

[5] Peter A. Zientara, Sooyeon Lee, Gus H. Smith, Rorry Brenner, Laurent Itti, Mary B. Rosson, John M. Carroll, Kevin M. Irick, and Vijaykrishnan Narayanan, “Third Eye: a shopping assistant for the visually impaired,” Computer, Vol. 50, No. 2, pp. 16–24, 2017

[6] 大久保 達矢, 三好 正純, “屋外ナビゲーションシステムにおける視覚障害者用情報提示機器の開発”, 第 1 回日本福祉工学会九州支部大会, 2016.

[7] 大久保 達矢, 酒田 信親, 清川 清, “視覚障害者のための買い物支援システムの提案”, システム制御情報学会研究発表講演会講演論文集 63, pp. 116–119, 2019

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Deep Residual Learning for Image Recognition,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, 2016.

[9] Scott D Roth, “Ray Casting for Modeling Solids,” Computer Graphics and Image Processing, Vol. 18, No. 2, pp. 109–144, 1982.

[10] 井谷 優, “マイクロマウスの歩んだ路”, 日本ロボット学会誌 Vol. 27.9, pp. 979–982, 2009.

© 2020 by the Virtual Reality Society of Japan (VRSJ)