

平成 23 年 4 月 30 日現在

機関番号：32641

研究種目：基盤研究(C)

研究期間：2008 ～ 2010

課題番号：20500184

研究課題名(和文) 準動的環境における移動ロボットのマルチモーダル環境地図生成の研究

研究課題名(英文) A study on mapping of multi-modal environment for mobile robot in semi-dynamic environment

研究代表者

坂根 茂幸 ( SAKANE SHIGEYUKI )

中央大学 理工学部 教授

研究者番号：10276694

研究成果の概要(和文)：本研究では、準動的環境を対象に、移動ロボットが自律的に作業を行う上で必要なマルチモーダル環境地図の生成手法の開拓を目的として以下の3つのサブシステムを構成した。(1) 環境物体のポーズ推定精度の向上のために、距離データに加えてカラー画像データを利用するサブシステム、(2) 環境内の物体の意味的情報を地図に付加するため、環境教示者の指差し方向を推定するサブシステム、(3) 教示者とロボットの音声対話により物体の情報を教示するサブシステムである。そして、これらのサブシステムの有効性を確認した。

研究成果の概要(英文)：In order to explore multimodal mapping techniques for mobile robot to work in semi-dynamic environment, we have developed three subsystems, (1) a subsystem to improve accuracy of object's pose by using color image data in addition to its range data, (2) a subsystem to estimate direction of finger pointing gesture of a tutor in teaching objects in the environment, and (3) a subsystem to teach semantic information of the objects, such as names and owners, using voice interaction between a tutor and a robot. And we validated these subsystems.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,600,000	480,000	2,080,000
2009年度	1,200,000	360,000	1,560,000
2010年度	700,000	210,000	910,000
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域→情報学→知覚情報処理・知能ロボティクス

科研費の分科・細目：基盤研究(C)

キーワード：移動ロボット、地図生成、マルチモーダル、準動的環境、無線 IC タグ

## 1. 研究開始当初の背景

近年、移動ロボットの環境地図を生成する技術の発展が著しい。人間が生活する環境、例えば、オフィス環境を例にとれば、環境内には、机、棚、壁等の(常にその位置・姿勢が不変な)静的物体以外に、可動キャビネット、椅子、ゴミ箱、部屋のドアなどのように、その位置・姿勢が時々変化する準動的物体、そして人間のように頻りに環境内を広範囲に

移動する動的物体が混在している。従来、移動ロボット技術の中では大部分は静的環境物体を地図生成の対象としており、準動的物体・動的物体が混在する環境下で地図を生成する技術は、その重要性にも係わらず未開拓の状況にある。

筆者らは、平成 17～19 年度科研費基盤研究において、準動的物体にも IC タグが付けられるユビキタス環境を想定して、IC タグに格納された U コードから物体の 3 次元幾何モ

デルがアクセスできる仮定の下で、レーザ距離センサで得た距離プロフィールに含まれている準動的物体の部分形状推定から、準動的物体のポーズ推定を行い、かつ地図生成と自己位置推定を同時に行う方法 (SLAM-SD) を提案した。このアルゴリズムでは、従来の SLAM 手法で行われる地図と自己位置の同時推定に加えて、IC タグのポーズ推定も行うための拡張を行っている。しかし、このシステムでは、(従来の地図生成手法の多くと同様に) 2 次元の距離プロフィールに基づいた占有グリッド地図であるため、環境物体の色、テクスチャ、3 次元立体形状などの情報については反映されていない。

従来研究では、レーザ距離センサと視覚を併用して移動ロボット用の地図作成を行う研究は既にあるが、準動的物体の取り扱いも含むマルチモーダルな情報を活用して地図生成を行う技術は未発達である。

## 2. 研究の目的

本研究では、準動的環境下における地図生成の研究成果を発展させる観点から、ロボットに搭載したレーザ距離センサから得る距離プロフィール、カメラから得る画像データ (カラー画像・距離画像)、そして、ユビキタス環境で物体に貼られた IC タグデータなど、多様な情報を反映させる「マルチモーダル環境地図」を SLAM の中で生成し、移動ロボットの諸タスクに利用するシステム化技術の開発を主な目的とした。

## 3. 研究の方法

筆者らがこれまでに開発した SLAM-SD のアルゴリズムを発展させる観点から、カラー画像処理の利用を図った。具体的には、Z 軸周りの回転角度が不確定な物体のポーズ推定の信頼性を向上させるサブシステムを構成し、その有効性を検証した。

また、物体の意味的情報を教示するために、実時間で距離画像を処理する距離カメラを導入し、サブシステムを構成して教示者の指差し方向を推定する手法を開発した。さらに、教示者とロボットの間で音声認識・対話を可能にするために、音声認識の Julius、形態素解析の MeCab、データベースの PostgreSQL 等のソフトウェアを用いたサブシステムを構成し、物体の名称や所有者等の意味的情報を地図に付加する基礎技術の検討を行った。

## 4. 研究成果

本研究では、レーザレンジファインダに基づく距離情報に加えて、(1) カラー画像情報を利用した対象物体のポーズ推定、(2) ロボッ

トに環境内の物体を教示者が指差し動作で教示する場合の指差し方向の推定を行うサブシステム、そして、(3) 教示物体の意味的情報を教示者が音声で教示する場合の音声情報の処理、を中心に取り組んできた。

### (1) カラー画像情報の利用

本研究では、レーザ距離センサにより得る距離プロフィール、準動的物体の 3 次元幾何形状の情報に加えて、搭載したカラーステレオ・カメラから得られる画像情報 (カラー画像・距離画像) も反映させた準動的環境の「マルチモーダル環境地図」生成により SLAM の信頼性向上を図ることを目指した。

具体的には、準動的物体の一例として、椅子 (軸回りに回転の自由度があるもの) の姿勢推定を画像情報に基づいて行うサブシステムを構成した [1]。これは距離センサ情報のみでは椅子の回転軸の位置のみが計測され、回転軸周りの角度は不明なため、準動的物体のポーズ推定が困難となる。マルチモーダルな情報の利用が必要となる例である。この場合の処理手順は、以下ようになる。

① 椅子の 3 次元幾何モデルに基づき軸回りの回転の自由度を推定するために、見えのモデル (2 次元画像) 生成する。② 無線 IC タグの検出に基づき、ロボットは対象の方向にカラーカメラ (USB カメラを使用) を向けて椅子の画像情報を入力する。そして、③ 椅子の色情報に基づく領域 (実験例では青色領域) の 2 値画像と、② で撮像した画像から抽出する青色領域との交差領域の面積を求める。この値を、② の実画像における青色領域の面積値で割り、マッチングの評価関数とする。この評価値が最大となる回転角度を求めることで椅子の姿勢角を推定する。

実験により、本方法で椅子の軸周りの回転姿勢について推定が可能であることを確認した。図 1 は、Z 軸回転角度を一定角度で変更して表示した、椅子の 3 次元幾何モデル表示の例である。

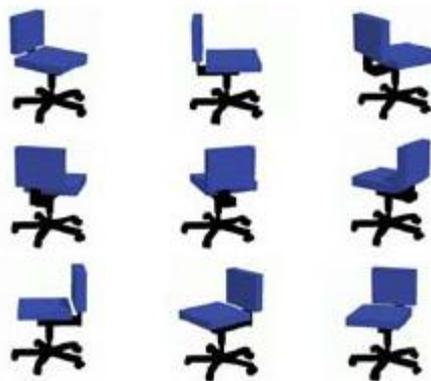
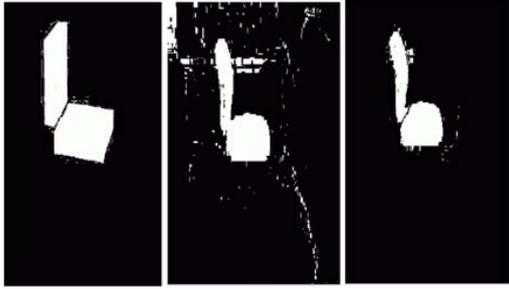


図 1 回転角度を一定角度ずつ変化させて表示した椅子の 3 次元幾何モデル



(a) (b) (c)

図2 (a) は椅子の 3 次元幾何モデル画像の例, (b) は実画像を 2 値化した画像の例, (c) は (a) と (b) の積演算で得た画像。

本サブシステムで構成したカラー画像情報に基づく対象物体のポーズ推定は比較的単純な方法であるため、今後は、よりロバスト化を図る必要がある。本システムでは、利用したレーザレンジファインダとカラー画像データの組み合わせで実画像と 3 次元幾何モデルとのマッチングを行っている。これに加えて、次節で用いるような距離画像データも利用することにより、ポーズ推定の信頼性をさらに向上させることが可能になる。

また、移動ロボットの能動性を利用して、ポーズ推定のために、より信頼性の高い観測位置へ能動的に移動してデータを入力するアプローチも有効と考えられる。

## (2) 物体の意味情報の教示における、 教示者の指差し方向の推定

自律移動ロボットの構築において、環境地図生成は重要な課題である。従来の環境地図の多くは幾何学的情報で構成されており、意味的な情報が含まれていない場合が多い。意味的な情報は、自律ロボットの様々なタスクにおいて必要である。例えば、人間がロボットにある物体を取ってくるよう指示した場合に、環境地図上のどの位置に物体があるのかが判らなければ、ロボットはそのタスクを実行できない。

そこで、本研究では、マルチモーダル環境地図の生成という観点から、環境地図に意味的な情報を統合するために必要な技術の一環として、人間の教示者(チューター)がロボットに様々な情報を指さして教示する課題を取り上げ、人体にマーカーを付けずに指さし方向を推定するサブシステムを構成した[2]。

具体的には、距離カメラで撮像した距離画像を処理することで、指さし方向を推定するプロトタイプシステムを構成し、実験によりその有効性を確認した。また、その改良システムとして、3次元関節モデルに基づく推定と追跡を行うシステム化の検討を行った。

本サブシステムでは、3次元の距離画像を

実時間で撮像できる TOF (Time-of-flight) 方式の距離カメラを用いている。サブシステムは 1 台の移動ロボット上でシステムが完結するため、教示者や環境に特別な装置を準備する必要がなく、多様な照明条件の影響を受けにくい。また、実時間で動作可能である。指差し方向の推定は、以下の手順で行う。

(a) 前景と背景の分離：明暗画像から前景(人領域)と背景を分離する。始めに、画像中から人の顔を Adaboost に基づくアルゴリズムで検出する。次に、検出した顔領域の距離値を平均する。この平均値と距離値との差が一定のしきい値以下で、かつ、顔領域と隣接する画素を前景(人領域)とする。



図3 人領域の抽出処理



図4 収縮処理の結果



図5 膨張処理の結果

### (b) 指差し方向の推定

2次元画像処理に基づく追跡アルゴリズムでは、細線化処理やモルフォロジー演算などを利用して、人の手先、肘、肩の3点の推定を毎フレーム実施する。肘を原点として手先方向に伸びる直線を指差し方向とする。

#### (b-1) 腕領域の分離

モルフォロジー演算で人領域から腕領域を分離する。これには次の4つの処理を用いた:

- ① 人物領域を2値画像に変換する (図3)
- ② ①の画像に収縮処理を行う (図4)
- ③ ②の画像に膨張処理を行う (図5)
- ④ ①の画像(図3)と③の画像(図5)の差分をとる(結果は図6)

収縮・膨張処理により腕領域が残る。処理前後の画像間の差演算から腕領域を分離する。

#### (b-2) 指差し方向の計算

以下の手順で分離した腕領域から3次元空間における指差しの方向を推定する。

- ①腕領域の画像に細線化処理を行う (図7)
- ②細線化後の画像から指先・肩(線の両端)を検出する。
- ③指先と肩を通る直線から最も遠い点を肘の位置と推定する。(図8)
- ④指先と肘の位置座標から方向ベクトルを計算する。

#### (b-3) 指差し方向の推定

2次元画像処理ベースの手法では、殆どの処理が2次元画像で行っているため隠蔽などの条件によっては、正しく推定できない場合がある。この問題の解決のため、3次元関節モデルに基づくトラッキング手法を検討した。始めに、対象の形状を表現する3次元モデルを準備する。このシステムとしては、単純な形状の3次元関節モデルを用い、観測したデータと関節モデルとのマッチングを行い、ポーズを推定する。マッチング手法については、パーティクルフィルタの利用が考えられる。



図6 差分演算の結果

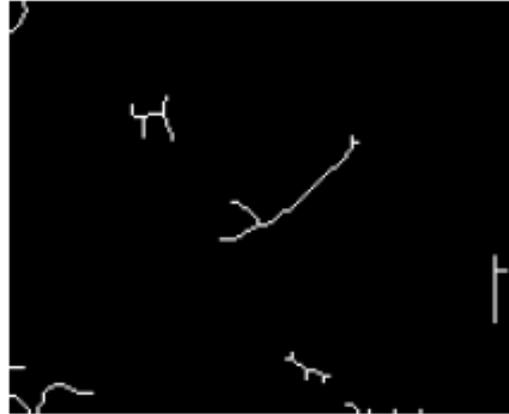


図7 細線化処理による腕領域の抽出

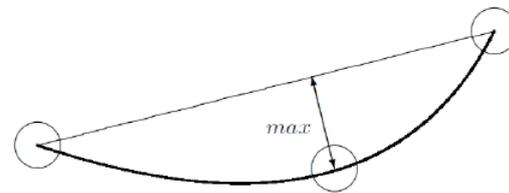


図8 肘の位置の推定

### (3) 環境地図における意味情報を教示するための音声情報処理

物体を教示者がロボットに教示をする際に、音声でインタラクションを可能にするために、Julius 音声認識システムと MeCab 形態素解析システムを結合した[3]。

本サブシステムで用いた主なソフトウェアは、連続音声認識を行う Julius、形態素解析を行う MeCab、そして、オブジェクト・リレーショナルデータベース管理システムの PostgreSQL である。これらのソフトウェア間の統合化には、ロボット用の Lisp システムである EusLisp を用いた。

意味情報を教示する場面を想定し、物体の物体名 (object) とその物体の所有者名 (owner) の教示に限定した。そして、object と owner は、MeCab による形態素解析結果の素性により、教示者が発話した文章情報から、対象物教示の意味情報として抽出した。

MeCab の形態素解析の結果により意味情報である単語を抽出するために、object または owner となり得る品詞を予め決めておく。object は一般的な物体名の品詞とした。owner には、一般的な物体名の品詞の他に、「彼」、「私」、「君」などの品詞も含めた。入力された発話文章から物体名 object を抽出するために、発話文章を Julius でディクテーションした結果を MeCab への入力として、形態素解析を行う。この解析結果に基づいて

object を抽出した。

日本語の特性上、重要な内容は文章の後部に来るため、最後にある object の品詞を持つ形態素を object とした。ここで、MeCab では、「液晶テレビ」のような物体名称が、「液晶」と「テレビ」という2つの形態素に分けられてしまう。そこで、object として抽出した形態素の1つ前の形態素を調べる。その形態素も object の品詞を持っている場合には、その形態素と先の object の形態素を合成して1つの object とした。

実験結果では、被験者によるばらつきがあるが、意味情報となる owner と object に対応する単語は、日本語の文法上の特性と素性の情報を用いて基本的に抽出することが可能になった。しかし、Julius による音声認識と、意味情報を抽出するための MeCab による形態素解析だけでは信頼性が不十分である。今後は、単語の意味やグループなど、素性以外の情報の適切な取り扱いが必要である。

#### [参考文献]

[1]遠藤翼：移動ロボットの準動的環境におけるマルチモーダル地図生成 -画像情報の照合による準動的物体のポーズ推定-, 中央大学理工学研究科経営システム工学専攻修士論文, 2010.

[2]守山晃生：移動ロボットの地図生成における意味的情報の拡張 -教示者の指差しジェスチャの推定-, 中央大学理工学研究科経営システム工学専攻修士論文, 2010.

[3]永田久美, 安澤遥平, 蒲谷栄光：移動ロボットと環境教示者の音声対話, 中央大学理工学部経営システム工学科 2009 年度卒業論文, 2010.

#### 5. 主な発表論文等

[雑誌論文] (計 3 件)

① Hongjun Zhou, Shigeyuki Sakane, “A Hierarchical Approach for Active Mobile Robot Localization,” INFORMATION: An International Interdisciplinary Journal, 査読有, Vol.12, No.1, 2009, pp.87-115.

② Hongjun Zhou, Shigeyuki Sakane, “Localizing Objects During Robot SLAM in Semi-Dynamic Environments,” Proc. of 2008 IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics (AIM2008), 査読有, 2008, pp. 595-601.

③ Hongjun Zhou, Shigeyuki Sakane, “Sensor Planning for Mobile Robot Localization -A Hierarchical Approach Using a Bayesian Network and a Particle Filter,” IEEE Transactions on Robotics, 査読有, Vol. 24, Issue 2, 2008, pp. 481-487.

[学会発表] (計 1 件)

①守山晃生, 渡部大志, 坂根茂幸: 環境地図への意味的情報付加のためのヒューマン・ロボットインタラクション -対象物への指さしの推定 -, 日本機械学会ロボティクス・メカトロニクス講演会 2009 (ROBOMEC2009), 1A2-E01, 2009 年 5 月 25 日, 福岡.

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

特になし

#### 6. 研究組織

(1) 研究代表者

坂根 茂幸 (SAKANE SHIGEYUKI)

中央大学・理工学部・教授,

研究者番号: 10276694

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

周 洪鈞 (HONGJUN ZHOU)

同済大学 (中国)・准教授

