

移動ロボットによる収納家具操作のためのモデル獲得 Model Creation for Furniture Handling by a Mobile Robot

学 山崎 公俊 (筑波大) 正 坪内 孝司 (筑波大) 正 友納 正裕 (東洋大)

Kimitoshi YAMAZAKI, Takashi TSUBOUCHI (Univ. of Tsukuba) and Masahiro TOMONO (Toyo Univ.)
yamazaki@roboken.esys.tsukuba.ac.jp

In this paper, a modeling method to handle furniture is proposed. In real environment, there is a lot of furniture such as drawer, cabinet and so on. If mobile robots can handle these furniture autonomously, it is expected that multiple daily tasks, for example, storing a small object in a drawer, can be performed by the robots. We propose direct teaching based modeling method which can easily give not only how to handle the furniture but also an appearance and 3D shape of it simultaneously. The main topic is about the method of appearance and shape acquisition using external sensors.

Key words: Mobile Manipulator, IM-model, Line Matching

1 はじめに

近年、ヒューマノイド型を例として、複雑な肢体のロボットが盛んに開発されている。これらのロボットは、環境中を自由に動き回れる移動能力と、人間のような種々の動作を可能にする動きの自由度を併せ持つため、我々が日常的に行っている作業の代替が期待できる。

そのような作業の具体例として、引き出し等を開閉して物体を出し入れする動作がある。これは、環境中の移動や物体操作など、いくつもの技術要因を含んでおり、様々な作業の基本動作と考えることができる。そのため、従来のロボット研究でも、同様の作業を実現する研究が行われてきた。ただし、これらの研究は、以下のいずれかの問題設定を置くことがほとんどであった。(1) 操作に必要なほとんどの知識を人手で与える [4] [5]、(2) ID タグや人工マーカを操作対象に添付する [2] [3]、(3) 遠隔操作などの手段で操作方法を提示する [6]。これらのアプローチは、ロボットの操作対象を増やしたいときに、人手にかかる負担が多いか、特別な道具を要する。我々が普段生活する実環境には多くの物体が存在するため、上述のような事前準備は、なるべく軽減することが望ましい。

本研究の目的は、我々が日常的に行っている作業のうち、引き出しなどの収納場所へ物体を納める動作に着目し、これをロボットに代替させることである。ここでの収納場所とは、日常生活に存在するキャビネット、戸棚などの家具や、冷蔵庫などを想定する。これを実現するには、それらの操作に関わる知識をロボットに持たせる必要がある。そこで、人手による教示を通して、操作知識をモデル化する方法を提案する。提案手法の利点は、操作対象を特定せず、特別な道具も必要とせず、簡便な教示を通して操作モデルを獲得できることである。操作モデルとは、収納場所の見た目・形状・操作方法を電子データとして獲得した統合的なモデルである。本稿では、そのようなモデルを「IM モデル (Instructed Motion Model)」と称する。なお、同様の目的に基づく研究成果を文献 [7] で報告しているが、本稿では特に、ロボットに搭載したカメラと測域センサにより、収納場所の見た目・形状を獲得する方法について述べる。

2 問題設定とアプローチ

2.1 想定タスク

日常的によく行う動作として「小型の物体を運搬し、人間が指示した家具に納める」を、想定タスクとして設定する。ここでの家具とは、収納を行う際になにかしらの操作が必要である収納場所とする。例えば引き出しであれば、「取っ手を掴んで引き出すことで、収納領域を確保する」といった操作が必要である。

2.2 アプローチ

本研究では、ロボットの操作対象を簡便に拡張できる枠組みの構築を目指す。基本方針として、ダイレクトティーチング [1] を適用する。すなわち、人間がロボットを直接的に動かして、操作方法を教示する。これにより、特別な道具や、詳細な操作指示が不要になる。また、ロボットにカメラと測域センサを搭載し、教示の前後に取得したセンサデータから、操作対象の見た目・形状を自動で獲得させる。

この方針には、以下の課題がある。(1) 人手で指示された教示データから、操作方法獲得に必要な情報をいかに抽出するか、(2) 外界センサを用いて得たデータから、見た目・形状情報をいかに抽出するか。これらの課題に対し、本研究では以下のアプローチをとる。

教示データからの操作方法抽出

本研究では、我々の身近に存在する家具のほとんどは、Table.1 のいずれかに属する方法で操作可能であることに着目する。この想定から、それぞれの分類ごとに操作知識を一般化し、ロボットにあらかじめ与える。これらは人手で定義する必要があるが、その数は少数なため、手間も少ない。言い換えれば、このアプローチの利点は、一つの操作知識を複数の収納場所に適用できるため、汎用性が高いことである。

一方で、個々の収納場所は異なった大きさや見た目を有するため、ダイレクトティーチングで得た教示データを元に、それぞれの差異を吸収する。この内容は文献 [7] に譲り、本稿では概説に留める。

表 1: Variation of Furniture Handling

種類	操作方法	収納場所の例
引き出し	取っ手を掴み、前後方向に直線水平軌道で引く	引き出し、キャビネット
引き戸	取っ手を掴み、横方向に直線水平軌道で引く	戸棚
開き戸	取っ手を掴み、鉛直な軸回りの回転軌道で引く	冷蔵庫、ドア

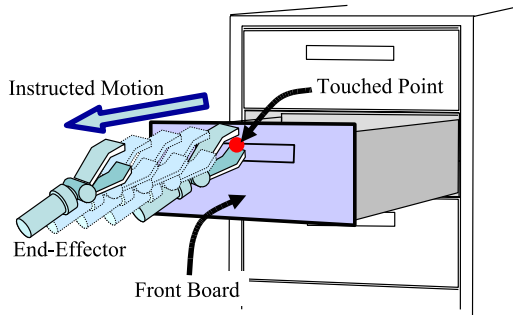


図 1: IM model

外界センサを用いた見た目・形状獲得

ロボットに搭載した単眼カメラと測域センサから、画像と三次元スキャンデータを取得し、それらを元に操作モデルを獲得する。具体的には、収納場所の前板について、その三次元輪郭と内部画像を抽出する。前者は、三次元スキャンデータに画像エッジを重畳することで得る。後者は、画像データから該当部分を切り出すことで得る。この詳細は次章で述べる。

本研究では、上述のように、人間による教示を通して得られる収納場所の操作モデルをIMモデル (Instructed Motion Model) と称する。

3 IMモデル獲得の概要

本章では、IMモデルの生成法 [7] を概説する。

3.1 教示手順

操作教示は下記の手順で行われる。

1. ロボットを収納場所の近くに設置する。ロボットは、その場で取得した画像とスキャンデータから、IMモデルの初期値を生成する。
2. ダイレクトティーチングを通して、操作方法を教示する。
3. 教示後、画像を一枚撮影する。この画像と初期値、そして操作軌跡のデータから、IMモデルを獲得する。

以下で、これらの詳細を説明する。

3.2 初期値の生成

初期値とは、教示の開始直後、家具を動かす前のセンサデータから選出する前板の候補である。前板 (Fig.1) は、操作の状態を直接的に反映する部位であり、本研究で想定するほとんどの家具に存在する。これを以下の手順で生成する。まず、測域センサから得た点列データに三角

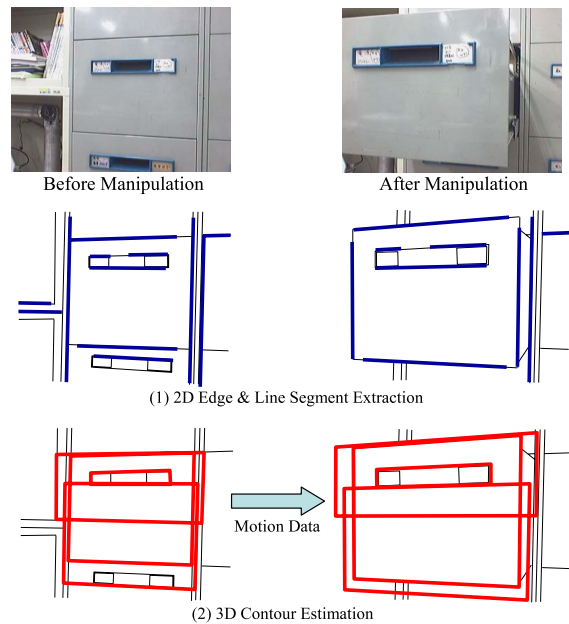


図 2: Front board estimation

パッチを張り、環境の三次元形状を得る。一方で、教示の直前に撮影した画像内からエッジを抽出し、そのうち所定の長さ以上の直線成分を得る (Fig.2,(1) 左)。そして、それらの直線成分のうち、四角形を形成する線分の組合せを選び、前板輪郭の候補とする (Fig.2,(2) 左)。最後に、画像内の輪郭の候補を三次元の三角パッチ上へ投影することで、それぞれの三次元位置を求める。

3.3 手先軌跡の分化

ダイレクトティーチの間、ロボットは一定時間ごとに自分自身の動きデータを記録する。このデータには家具の操作情報が含まれているため、全データの中からその部分のみを抽出する。Fig.3は、引き出しを引く手先軌道を横から見た例である。(1) 初期姿勢からロボットの手先を誘導し、(2) 引き出しの取っ手に手先を差し込み、(3) 引き出す。それが終わった後、(4) 取っ手から手先を外し、(5) 初期姿勢に戻す。このうちIMモデルに必要な情報は(2)~(4)であるため、その抽出方法を事前知識としてロボットに与えておき、モデル化に利用する。

3.4 前板候補の淘汰

初期値生成で得た前板候補は、まだ家具が動かされる前のものであるため、本来の前板以外の部分も候補として持ちうる。そこで、教示後に得たデータから、真の前

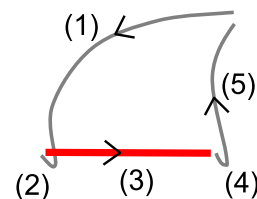


図 3: Hand trajectory in drawer case

板部分に存在する候補を選択する。

手順は以下の通りである。まず、教示が終わりロボットの先手が元の位置に戻された直後、収納場所の画像を取得する。一方で、それまでに得られた操作軌跡を基に、現在の輪郭候補の三次元的な姿勢を推定する。そして、その推定結果を画像上に重畳し (Fig.2,(2) 右)、その結果が画像エッジ (Fig.2,(1) 右) と良く重なる輪郭のみを残す。

以上の方法で IM モデルを獲得する。本手法では、前板情報を画像と三次元形状で獲得できる。そのため、このモデルをデータベースに登録しておけば、教示時と異なる環境下でも、ロボットが見た目情報を基に自動で収納場所を発見し、作業を行うことが期待できる。

4 前板モデリング

本章では、前板のモデル化について詳細を述べる。

4.1 前板モデリングにおける課題とアプローチ

3.2 で述べた初期値生成では、前板輪郭の候補が三次元情報として多数得られる。これらは、画像エッジを直線分化し、そのうち四角形を構成する組合せを選び、三次元空間に投影することで得られるものである。これらの候補には、前板以外の部分も含まれている可能性があるため、収納場所を操作した後に取得した画像から、前板らしさを検証する必要がある。

この検証処理では、二つの課題がある。

前板の対応付け

前板のモデルを正確に得るためには、教示前に生成した前板候補を、教示後に撮影した画像上で対応付ける必要がある。しかし、収納場所の操作のように二枚の画像間で特定の部分のみが動かされ場合には、初期に得られたのと同じ状態で線分を得ることは困難である。

カメラ姿勢と三次元形状の誤差

カメラの姿勢変化は、基本的にロボットの動きデータから取得できるが、ここには誤差が入りうる。また、画像エッジと測域センサから得た三次元情報にも、誤差が含まれる。この誤差を完全になくすことは不可能である。

これらに対処するため、対応付けは画像エッジベースで行う。初期値として得た前板輪郭の候補を教示後の画像上に投影し、それを参照線分とする。そして、これらと画像エッジと重畳し、その重なり具合から対応付けの判断を行う。ただし、このときロボットの姿勢やセンサデータの誤差が問題となり、エッジと参照線分は完全には一致しない。そこで、対応付けを二段階で行う。

4.2 線分と画像エッジの対応付け

初期値生成では、前板輪郭の候補が三次元線分の集まりとして得られる。また、教示を通して前板の動きが判る。これらを利用すると、前板輪郭の候補を構成する線分が、教示後に撮影した画像上にどのように写り込むかが推定できる。以後、この投影線分を参照線分と呼ぶ。

対応付けは、参照線分と画像から抽出されるエッジの重なりを調べることで行う。対応付けで問題となるのは、(1) 初期画像のエッジを直線分化する時点で生じるずれ、(2) カメラ姿勢の誤差、(3) 外界センサの誤差、などの影響でなる。誤差の原因が (1) のみであれば、参照線分と画像エッジはほぼ問題なく対応づけられるが、移動ロボッ

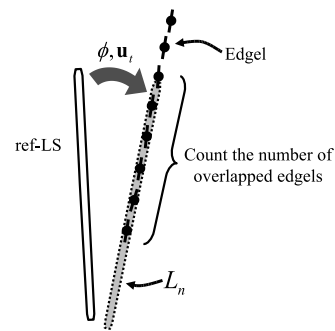


図 4: Line segment vs. edgels

トでは (2)、(3) が避けられない。そこで、以下の手順で対応付けを行う。

第一段階：参照線分と edgel の比較による対応候補選出

Fig.4 は、対応付けの概念図である。以後の説明では、参照線分を *ref-LS*、参照線分に微小な回転 ϕ と平行移動 u_t を加えた線分を L_n と称する。

前述した通り、教示の前後で撮影した画像間ではエッジの再現性が低いため、線分同士の場合はもちろん、線分とエッジを対応付けることも困難である。そこで、edgel を利用した対応付けを行う。edgel とは、画像エッジを画素単位に分割した各要素のことである。それぞれの edgel は、画素の座標と、勾配の大きさ・方向を保持する。

対応付けの手順は以下の通りである。まず、画像エッジを抽出し、一方で *ref-LS* を求める。次に、*ref-LS* に微小変化を加えた線分 L_n を生成し、 L_n に重畳している edgel を数える。このとき、edgel の合計が L_n の長さと比較して十分な比率となっていれば、その L_n を *ref-LS* の対応候補として記録する。ただし、この対応が真であるとは限らないため、この段階では、ひとつの *ref-LS* に対して、複数の対応候補を持たせる。

第二段階：対応付けの信頼性評価

前述した対応候補が、三次元情報として整合しているかを調べる。具体的には、既に判っているカメラ姿勢と、投影前の参照線分の三次元情報を用いて、線分の逆投影誤差を調べる。評価式は、以下のようである。

$$C = \sum_{i=1}^m \left(\rho_n - f \frac{\{\cos \theta_n (\mathbf{r}_1 \mathbf{x} + t_1) + \sin \theta_n (\mathbf{r}_2 \mathbf{x} + t_2)\}}{\mathbf{r}_3 \mathbf{x} + t_3} \right), \quad (1)$$

ここで、 ρ_n 、 θ_n は、候補線分 L_n をハフ変換して得た変数である。 \mathbf{x} は *ref-LS* の元となる三次元線分の中心座標を表す。 \mathbf{r}_i 、 $t_i (i = 1, 2, 3)$ はカメラの回転と平行移動を表す変数である。 f はカメラの焦点距離、 m はランダムに選択された L_n の数である。

この式は、*ref-LS* の対応候補 L_n について、元の三次元線分との逆投影誤差を求めるものである。一枚の画像には多数の *ref-LS* があり、それぞれの *ref-LS* は複数の対応候補を有する。そこで、一つの *ref-LS* に対して一つの L_n をランダムに選択する処理を、前板輪郭を構成する m 本の *ref-LS* で行い、選んだ L_n を式 (6.1) に代入する。そして、選択した逆投影誤差 C を、 L_n それぞれにスコアとして与える。この流れを所定の回数行い、各

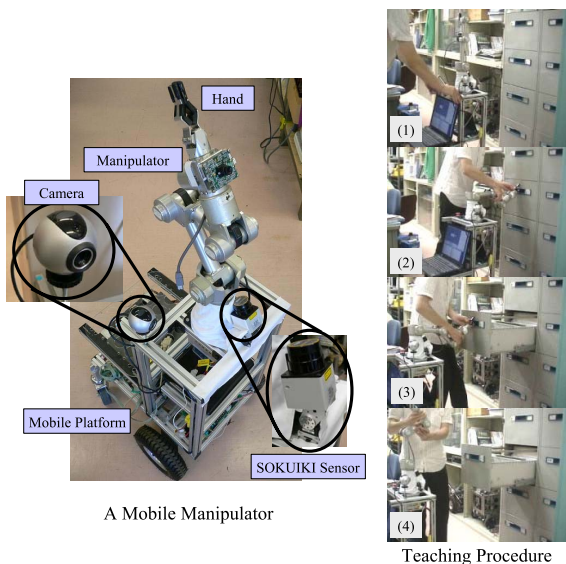


図 5: Experiment

ref-LS ごとにスコアの一番良い L_n を選ぶ。さらに、 L_n のスコアが所定の閾値以下であれば、その ref-LS は破棄する。以上の方法で、明らかに対応が合わない前板輪郭の候補は除去され、さらにそれぞれの前板候補を画像上で対応付けることができる。

4.3 前板モデリングの性能向上

式 (1) を用いた対応付けは、線分の候補が多くなるほど計算量が増える。そこで、二つの効率化手法を提案する。これにより、明らかに対応が見えない候補を事前に除去でき、処理が高速化する。さらに、対応付けの確実性も向上する。

手先の接触位置に基づく前板候補の淘汰

多く家具で共通していることの一つとして、前板と取っ手の位置関係がある。ほとんどの場合、前板領域の内部に取っ手が付けられている。そこで、この仮定に基づいて前板候補を淘汰する。すなわち、画像から抽出された前板候補のうち、その内部に手先の接触点の投影値を含んでいないものは、候補から外す。

見た目情報を利用した前板候補の淘汰

教示前後の画像間では、家具の部分的な変化が撮影されている。もし、候補が前板部分から抽出できていれば、それらの輪郭内部に写り込んでいる画素情報は、二枚の画像間で変わらないはずである。

そこで、この仮定に基づいて前板候補を淘汰する。具体的には、各候補の輪郭内部の画素についてヒストグラムを求め、それを教示の前後で比較する。その二つのヒストグラムがほぼ一致していれば、前板部分が正確に取得できていれるとみなす。

5 実験

Fig.5 左は、5 自由度のマニピュレータを搭載した車輪型移動ロボットである。本実験では、キャビネットを操作対象として選んだ。Fig.5 右は、ダイレクトティーチングの経過を示している。

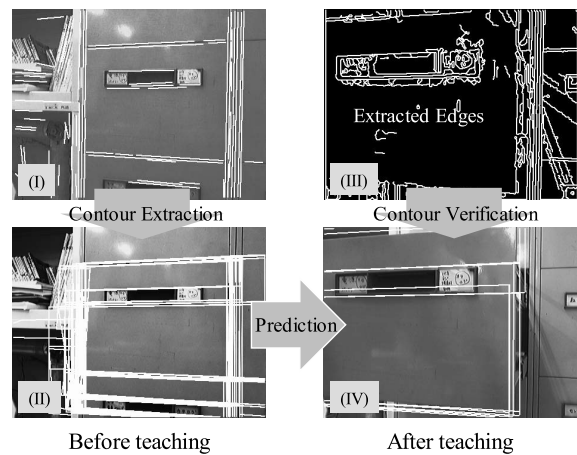


図 6: Image Processing

Fig.6 は、輪郭候補を示している。初期値生成では 81 の輪郭候補が、本来の前板とは異なる場所からも抽出されたが、提案手法を適用することで、5 の輪郭候補を正しく得ることができた。処理時間は 4.6 sec であった。

また、同様の実験を他の引き出しなどについても試行し、前板部分が正しく抽出できることを確認した。

6 まとめ

本稿では、日常環境に存在する引き出しなどの操作モデルを、ロボットに獲得させる方法を述べた。特に、収納場所の見た目・形状情報を、ロボットに搭載した外界センサから獲得する手法を提案した。本研究で提案した枠組みによれば、特別な道具や詳細な操作を必要とせずに、ロボットの操作対象を容易に増やすことが期待できる。

今後の課題として、より多くの家具操作に本手法を適用し、改良を加えることがある。

参考文献

- [1] H. Asada and Y. Asari: "The Direct teaching of tool manipulation skills via the impedance identification of human motion", Proc. of IEEE Int. Conf. on Robotics and Automation, pp.1269-1274, 1988
- [2] N. Y. Chong and K. Tanie: "Object Directive Manipulation Through RFID", Proc. Int. Conf. on Control, Automation, and Systems, Gyeongju, Korea, Oct. 22-25, 2003.
- [3] R. Katsuki, J. Ohta, T. Mizuta, T. Kito, T. Arai, T. Ueyama and T. Nishiyama: "Design of Artificial Marks to Determine 3D Pose By Monocular Vision", Proc. 2003 IEEE Int. Conf. Robotics and Automation, pp.995-1000, 2003.
- [4] J. Miura, Y. Shirai and N. Shimada: "Development of a Personal Service Robot with User-Friendly Interfaces", 4th Int. Conf. on Field and Service Robotics, pp.293- 298, 2003.
- [5] L.Petersson, P. Jensfelt, D.Tell, M.Strandberg, D. Kragic and H.I. Christensen: "Systems Integration for Real-World Manipulation Tasks", Proc. 2002 IEEE Int. Conf. Robotics and Automation, pp.2500-2505, 2002.
- [6] N.Sien, O.Stasse, Y.Kawai, T.Sakaguchi and K.Yokoi: "A Unified On-line Operation Interface for Humanoid Robots in a Partially-Unknown Environment", Proc. of IEEE Int. Conf. on Intelligent Robots and Systems, pp.4437-4439, 2006.
- [7] K. Yamazaki, T. Tsubouchi and M. Tomono: "移動マニピュレータへの直接的教示に基づく物体収納作業のための動作計画", 第 12 回ロボティクスシンポジウム, 2007.