

# 把持経験の蓄積に基づく布製品の掴み上げ能力の獲得

○守屋佑亮（信州大学） 山崎公俊（信州大学） 竹下佳佑（トヨタ自動車）

## 1. はじめに

生活支援に係わる作業を自動化するとき、自動機械に物体操作能力を持たせることが必要な場合が多々ある。それを実現する伝統的な方法の一つは、操作対象物のモデルに基づいて物体認識をおこない、その結果を入力として、計画器により自動機械の動作列を生成するというものである。操作対象物のモデルは、形状等として手動もしくは半自動で定義することが多い。また、動作計画においては、操作対象物や周囲の事物を仮想的に再現したシミュレーション環境を用意することが多い。

上述の方針は、基本的に操作対象物が剛体であることを想定している。一方で、我々が日常生活を送るうえで布製品のような柔軟物もきわめて身近な存在であり、それらに係わる自動化の要望が多いことも事実である。そのような柔軟物を扱おうとする場合は、剛体と同様の方法により操作を自動化することは困難である。布製品の操作をおこなう従来研究では、吊り下げ状態へ移行させる前段動作があり、その後に適切な把持位置を決定する手法が提案されており、その際は布の角部分が把持箇所候補となることが多い [1][2]。無造作に置かれた状態からの把持点選出をおこなう方法も存在するが、検出率にはまだ課題がある [3][4]。

筆者らは、操作の第一段階である布製品の「把持」に着目し、自動機械により把持方法を自動で獲得させることを目指している。本稿では、折り畳まれた布製品を対象として、その形態をなるべく崩さずに適切な場所を掴んで持ち上げるタスクを考える。このとき、適切な持ち上げ方に関するルールを開発者が明示的に与えず、様々な把持方法をロボットに試行させる中から、必要な知識をロボットに自動で獲得させる。

本研究で特に着目するのは、柔軟物の持ち上げの状況把握に適したセンサデータの加工方法である。もし適切な加工方法が得られ、それにより良い特徴表現が生成できれば、知識獲得を効率的におこなうことができ、その知識を利用した柔軟物操作の精度も高まると考えられる。知識の獲得に関する最近の流れとして、畳みこみニューラルネットワークを利用する方法 [5] などがあるが、大量の経験データを必要とすることが問題である。例えば対象物が剛体であれば、ある程度の経験データ獲得をシミュレーションでおこなうことが容易であるが、折り畳まれた状態の柔軟物については適切なシミュレーションが難しい。すなわち、実際のロボットを用いて大量の持ち上げ経験データを収集することが必要になる。それに対し筆者らは、これまでの研究経験 [6][7] を生かし、布製品に適した特徴量表現を手手で与える。そして、知識獲得にあまり多くの経験データを要しない仕組みを構築することを目指す。

## 2. 問題設定と手法の概要

折り畳まれた布製品がテーブル上に置かれている。外界センサと多関節アームからなるロボットを利用して、その布製品の一部を掴み、持ち上げる作業をおこなう。持ち上げ後、畳まれた折り目を保持したままの状態が保たれていれば成功、折りが解消され、だらりと垂れ下がるようなことがあれば失敗とする。また、把持位置が適切でなく、そもそも把持できなかった場合や、持ち上げ中に落とした場合も失敗とする。

本研究では、外界センサとして三次元距離画像センサの利用を想定する。すなわち、持ち上げ中の様子がカラー画像と深度画像の時系列データとして得られるものとする。持ち上げ知識を獲得するためのデータ収集では、実際の持ち上げ作業をおこないながら、布製品を把持した時点での手先の位置座標と、持ち上げ作業の開始から終了までの様子を納めたカラー画像列・深度画像列を記録していく。

一方、持ち上げのための知識が獲得できたとしたとき、新たな持ち上げ作業は次のように実行する。まず、テーブル上に置かれた布製品のどこを掴むかを決定する。この決定では、現在のセンサデータと過去に把持に成功した時のセンサデータを照合し、過去の成功時に掴んだ場所を現在のセンサデータ上にマッピングする。ここで良い把持位置が選択でき、その後の把持も問題がなければ、持ち上げタスクは成功する確率が高い。しかしながら、マッピングのミス等により、把持動作後に持ち上げに失敗することも考えられる。そのため、持ち上げ中も常に把持状況を認識し、失敗検知をおこなうことを目指す。すなわち、もし持ち上げ中の布製品の形状状態が想定通りではない場合などは、作業を中断してやり直し動作へ移行することを想定する。その場合、やり直し動作の内容を現在の把持状態から決められるような認識ができる必要がある。

以上の流れについて、認識関連で肝要となるのは、把持位置を決定する部分と持ち上げ中の状態認識である。次章からはそれぞれを実現するための知識の獲得方法を説明する。

## 3. 把持位置の決定手法

### 3.1 学習データの取得

前述したように、持ち上げ知識の獲得のために学習データの収集をおこなう。まず、布製品がテーブル上に置かれている状況について、ロボットが動き出す直前のカラー画像と深度画像を記録しておく。また、人手による指示もしくはロボットが自動で選択した把持個所の座標も記録しておく。そして、これらのデータに対してその後の持ち上げ動作が成功したか否かも記録しておく。把持位置の推定に用いるのは、それらの

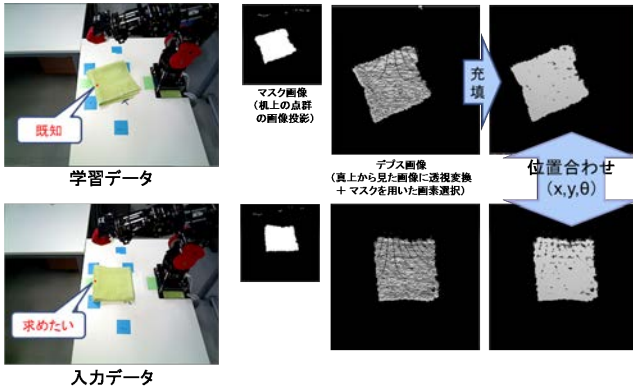


図1 把持点決定の流れ

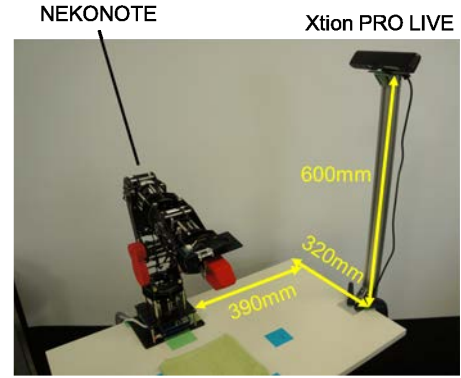


図2 実験環境

うち持ち上げに成功したデータのみである。

### 3.2 把持位置の決定

テーブル上に置かれた布製品について、カラー画像と深度画像を取得する。深度画像から三次元のポイントクラウドを生成し、平面検出によりテーブル天板の平面方程式を算出する。その後、その平面よりもセンサ側にある三次元点のみを選択し、それらを二次元平面に射影する。その二次元平面は、テーブル天板を垂直上方から観測した場合の画像平面である。以上により、布製品に所属すると考えられる点群の投影像を得る。この過程では、原画像から投影画像への画素位置の変換パラメータも得られるため、カラー画像についての投影像を得ることも容易である。

把持位置の決定は、学習データと入力データのそれぞれから投影画像を生成し、この2枚の画像を照合することでおこなう。もし照合度が高ければ、そのときの学習データに記録された把持位置を入力画像上にマッピングし、把持位置が決定できたとする。このとき問題となるのは、学習データを取得したときの布製品の置き方が、入力データを取得した時のそれとは同一にならないことである。そこで、入力データによく重なるように、学習データを位置合わせする。このために粒子フィルタを利用する。すなわち、学習データに多数の姿勢変位仮説を与え、最もよく重なるような姿勢変数を探索する。姿勢変数は本来であれば6種類であるが、前述のように鉛直視点からの変換を加えた後であれば、平面上の平行移動と回転の計3自由度と考えることができる。よって、一つの粒子が $(x, y, \theta)$ の3パラメータを持つものとして、学習データに平行移動・回転を加えたのち、当てはまりの評価をおこなう。図1に上述の流れを示す。

当てはまりの評価は、深度情報を用いておこなう。評価式は以下のとおりである。

$$eval = \sum_{u,v} \frac{1}{\{d_{ref}(u', v') - d_{input}(u, v)\}^2 + C} \quad (1)$$

ここで、 $(u, v)$ 、 $(u', v')$ は画像座標である。 $d_*(u, v)$ は座標 $(u, v)$ の深度である。 $(u', v')$ は、 $(u, v)$ を姿勢変数 $(x, y, \theta)$ により座標変換した結果である。 $C$ は定数

であり、 $ref$ は学習データの一つ、 $input$ は入力データを意味する。上式では、姿勢変換をおこなったのちにテーブル上に存在する学習データのすべての点に対して、入力データとの差分を取っている。多くの点が小さな差分で重なるほど、よい評価が得られる。

このほかにも、入力データとよく重なる可能性がある学習データをあらかじめ選び出す処理があれば、把持点の決定は効率的に進められる。ただし、必須ではないため本稿では扱わず、今後の課題とする。

## 4. 持ち上げ中の状態記述

### 4.1 状況の分類

持ち上げタスクにおいて想定できる状況は、大きく分けると以下の4通りになる。

**成功見込みー成功：**布製品から適切な把持位置が抽出され、それに応じた把持をおこない、持ち上げに成功する。

**成功見込みー失敗：**布製品から適切な把持位置が抽出されたが、持ち上げには失敗する。

**失敗見込みー成功：**布製品から不適切な把持位置が抽出されたが、持ち上げには成功する。

**失敗見込みー失敗：**布製品から不適切な把持位置が抽出され、それに応じた把持をおこない、持ち上げに失敗する。

これらの分別が可能となるようなセンサデータの加工方法を検討する。なお、失敗についても、把持できなかったり、持ち上げ中に落としたりなどの分類が可能であるが、本稿では扱わない。

### 4.2 特徴量の表現

持ち上げ作業の最中に、時系列のカラー画像と深度画像が得られるものとする。持ち上げ中の布の輪郭やシワには、成功か失敗かによってある程度の傾向が出ると考えられる。また、布の動き変化も重要な情報になると考えられる。本研究では、カラー画像から特徴表現を得て、それを用いて状態記述をおこなう。カラー画像に対しては、HOG、SIFTなどのエッジベースの

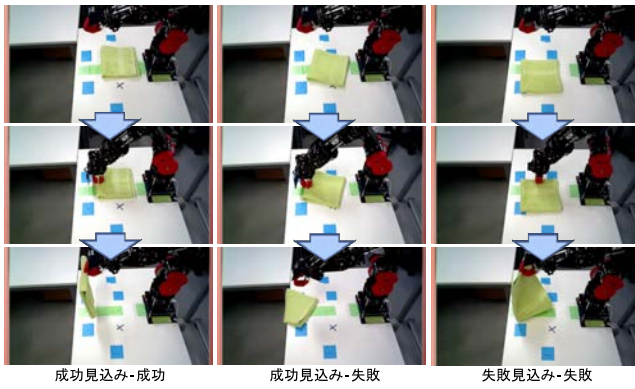


図3 持ち上げ作業の様子

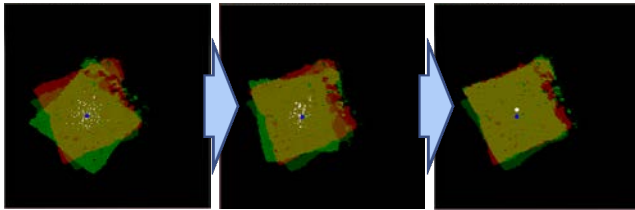


図4 把持点決定のための位置合わせ

特徴量	識別率 (%)	特徴量	Phase1 識別率	Phase2 識別率	Phase3 識別率
HOG	68	HOG	68	62	78
SIFT	61	SIFT	51	53	74
YI [7]	56	YI [7]	50	42	75
フロー方向	37	フロー方向	31	40	36
フロー強度	35	フロー強度	38	32	32

図5 各特徴表現での識別率

特徴表現や、テクスチャに基づく特徴表現 [7] を検討する。また、画像間での動きの情報であるオプティカルフローも検討する。オプティカルフローについては、フローの向きと大きさについての頻度ヒストグラムを生成する方針を採る。

## 5. 実験

図2に示すように、矩形のタオルを4つ折りにし、テーブルの上に置いた。6軸のアーム（アールティ社製、NEKONOTE）をそのテーブルに設置し、タオルとアームを俯瞰できる視点に三次元距離画像センサ（ASUS社製、Xtion PRO LIVE）を設置した。センサからは  $640 \times 480$  [pixel] の大きさのカラー画像と深度画像を取得した。

知識獲得のための学習用データとして、前章で述べた各状況ごとに30の持ち上げデータを収集した。持ち上げ作業中の例を図3に示す。なお、4通りの状況を想定してデータの収集をおこなったところ、3.の状況がほとんど起きなかったため、本実験では1., 2., 4.の3クラスのみを扱うこととした。

把持位置の決定では、成功時の時系列学習データから最初の時刻のセンサデータを取得し、それと現

在の最初のセンサデータとの照合をおこなうことで、把持位置の推定結果を得た。図4に処理結果の一例を示す。赤色部分が入力データ、緑部分が学習データであり、二つのデータが重なり合った部分は黄色で表されている。多数の白い点は粒子のうちの位置座標  $(x, y)$  を可視化したものである。粒子フィルタについては、粒子数250、予測時の粒子の標準偏差はそれぞれ  $(\sigma_x, \sigma_y, \sigma_\theta) = (10[mm], 10[mm], 5[deg])$  とした。

一方で、持ち上げ中の状態記述のために、学習時の時系列データを5時刻ごとに分割し、各分割ごとに特徴ベクトルを算出した。それらを用いて線形SVMにより識別器を生成した。そして、得られた識別器を用いて、入力データの各フレームに対して識別をおこなった。図5は、持ち上げ中の識別結果を示している。左側の表は、一つの持ち上げシーケンスが終わったあと、いくつかのフレームで識別が成功したかを算出したものである。右側の表は、持ち上げシーケンスを3つのフェーズに分割し、各フェーズでの識別率を算出したものである。例えばHOGでは、布製品に手をアプローチし（phase 1）、中ほどまでつかみあげる（phase 2）までは6割台で失敗・成功が判別され、持ち上げきる段階（phase 3）では8割弱の正答率が得られている。

## 6. まとめ

本稿では、畳まれた布製品を崩さずに持ち上げるための知識獲得とその活用方法について述べた。今後はまず識別率の向上に注力し、その後はロボットが自動で知識を獲得するための手法提案と実装をおこなう。

## 参考文献

- [1] Cuen-Rochin, J. Andrade-Cetto and C. Torras: "Action Selection for Robotic Manipulation of Deformable Planar Objects," in Proc. of Frontier Science Conference Series for Young Researchers: Experimental Cognitive Robotics, pp. 1-6, 2008.
- [2] A. Doumanoglou, A. Kargakos, T-K. Kim, S. Malasiotis: "Autonomous Active Recognition and Unfolding of Clothes using Random Decision Forests and Probabilistic Planning," in Proc. of IEEE ICRA, 2014.
- [3] A. Ramisa, G. Alenya, F. Moreno-Noguer and C. Torras: "Using Depth and Appearance Features for Informed Robot Grasping of Highly Wrinkled Clothes," in Proc. of IEEE Int'l Conf. on Robotics and Automation, pp. 1703-1708, 2012.
- [4] K. Yamazaki: "A Method of Grasp Point Selection from an Item of Clothing Using Hem Element Relations," Advanced Robotics, Vol. 29, No. 1, pp.13-24, 2014.
- [5] I. Lenz, H. Lee, A. Saxena: "Deep Learning for Detecting Robotic Grasps," International Journal of Robotics Research, vol. 34 no. 4-5, pp. 705-724, 2015.
- [6] K. Yamazaki, R. Oya, K. Nagahama, K. Okada and M Inaba: "Bottom Dressing by a Dual-arm Robot Using a Clothing State Estimation Based on Dynamic Shape Changes," International Journal of Advanced Robotic Systems, ISSN 1729-8806, 2016.
- [7] 山崎, 稲葉: 「布地, しわ, 布の折れ重なりに着目した画像特徴量による無造作に置かれた布製品の個体識別」, 計測自動制御学会論文集, Vol.49, No.7, pp. 661-669, 2013