

見立て表現の解釈に基づく道具操作の観察学習

○出村聡規（信州大学大学院総合理工学研究科），長濱虎太郎，山崎公俊（信州大学工学部）

要旨：本研究の目的は、見立て表現の解釈に基づく道具操作の新たな教示方法の実現である。本稿では、教示者の持つ任意形状物体の姿勢推定をおこなう手法を提案する。本手法では、教示者が持つ任意形状物体の各フレーム間での姿勢変化を、Optical Flowを活用して推定する。攪拌・台拭き動作での任意形状物体の姿勢推定により、本手法の有効性を評価した。

キーワード：見立て、画像処理、観察学習

1. 緒言

我々が普段生活している中で行う家事などの作業は、高齢者・要介護者にとっては負担の大きい作業である。生活支援ロボットの普及によって、そのような作業の負担の軽減や、労働力の補填につながるのではないかと考えられる。しかし家事作業一つをとっても、作業の内容や方法は各家庭によって異なる。したがって各家庭でのふるまいをロボットへ教えることのできる機能が鍵技術となるであろう。ただし、生活支援ロボットを利用するユーザは誰しもがロボットに関する知識を持つとは限らない。よって、ロボットの動作を直観的かつ容易に修正・拡張できる方法があれば有用であると考えられる。

従来、直観的な動作教示手法として大きく分けて視覚教示法と直接教示法の研究が進められている。そのうち視覚教示法[1]では、物体操作[2]や道具操作[3]を視覚情報から学習させる研究が行われている。この手法では大きな力を必要とせず、また人から人へ教示する場合と同様な方法であるため、幅広いユーザが活用できると考えられる。しかし従来研究の課題として、作業中の動作修正が難しい、複数のセンサを利用しての認識や同じ道具を用いての教示など、家庭で教示環境を用意するのが難しいといった点が挙げられる。これらの課題に対し、著者らは「見立て」に着目した動作教示法の研究をおこなってきた[4]。

「見立て」とは「ある物事に他の物事の像を重ね合わせる事」[5]であり、人が人へ物体の姿勢や動作を伝達する際に伝統的に用いられてきた技法である。文献[4]の手法では、教示に使用する物体や環境によらず、それぞれの作業空間に合った動作の生成が可能となった。この手法では物体の軌道の教示はできるものの、物体の姿勢変化の認識は行っておらず、フライ返しで裏返す動きのような、道具の姿勢が変化する複雑な動きの教示には未対応である。

そこで本研究では、教示者が「見立て」表現により操作する任意形状物体の動きの解釈を行い、より複雑な作業の教示をも可能とする新たな教示方法を構築する。本稿では特に、教示者の持つ任意形状物体の姿勢推定法を提案する。

2. 見立て解釈に基づく視覚動作学習

人から人へ、見立てを用いて道具操作を伝達する

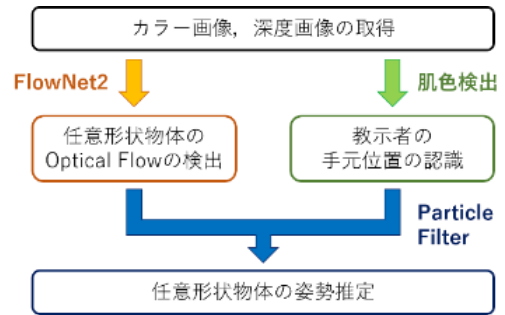


図1 教示者の持つ任意形状物体の動き推定の流れ

状況として、学習者の持つ道具や環境と、教示者が持つ物体や環境が異なる場合を考える。例として、調理中の学習者が調理道具で何かをかき混ぜている際、教示者が持つ道具を調理道具に見立てて動かして見せ、より良いかき混ぜ方を教えるといった状況が挙げられる。

この時、かき混ぜるという動作一つをとっても、鍋の底をかき混ぜるように、道具の姿勢を変化させず手元の動きに合わせて道具を動かす場合、また泡立てるように手元の位置をあまり変化させず、道具の姿勢を変化させるよう動かす場合などが考えられる。このようにより詳細な道具操作を教示する場合には、手元の移動量とともに、教示者の持つ物体の姿勢変化の認識を行う必要があると考えられる。

2.1. 教示者の持つ任意形状物体の動きの認識

本研究で提案する、教示者が持つ任意形状物体の姿勢変化推定の流れを図1に示す。まず教示者の手元の位置を検出するため、人間の姿勢認識手法の一つである OpenPose[6]を用いる。OpenPoseにより得られた教示者の右手首座標と深度画像を照らし合わせることで三次元位置を獲得し、それをもとに教示者の手元周辺に物体の認識領域を設定する。

次に、教示者の手元位置、任意形状物体の動きの認識を行う。教示者の手元位置の認識は、認識領域内で肌色検出によって手元の二次元位置を取得し、深度画像と照らし合わせ三次元位置を認識する。また物体の動きの認識には、移動物体の検出・解析などに用いられる Optical Flow 推定を用いる。推定には深層学習を利用した手法である FlowNet2[7]を用いた。得られた Optical Flow に対し、手元位置の認識と同様な方法により三次元 Optical Flow の推定を行

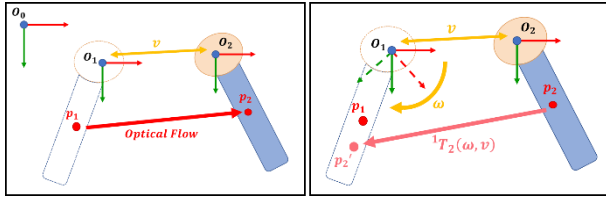


図2 Optical Flowを用いた尤度計算方法

う。ただし、移動量の大きすぎる・小さすぎる Flow は信頼できないデータとして、フィルタにより除外するものとした。

2.2. 物体の姿勢変化推定

前節で得られた Optical Flow を用い、教示者の持つ物体の姿勢がフレーム間にどう変化したのかを推定する。フレーム間の姿勢変化はパーティクルフィルタによって推定する。パーティクルフィルタの状態変数には任意形状物体の姿勢の変化量 ω [rad/frame] を用い、尤度計算は以下の手順で行うものとした。

基準座標系、直前・現在のフレームでの手元座標系を O_0, O_1, O_2 、フレーム間の手元移動量を v [m/frame]、基準座標系から見た Optical Flow の移動元、移動先の点をそれぞれ p_{prev}, p_{cr} [m] とする。まず、 p_{prev}, p_{cr} を O_1, O_2 から見た点 p_1, p_2 に変換する。同次変換行列 ${}^1T_2(\omega, v)$ により p_2 を p_2' へ変換し、 p_1 と p_2' の距離 d を計算する。以上の計算を検出された全ての Optical Flow に行い d の合計値を計算する。合計値を最も小さくする ω が、尤度が最も高いとする

3. 実験結果

提案手法を評価するために、攪拌動作、台拭き動作の二つの動作での物体姿勢認識を行った。教示者が用いる物体としては、家庭環境下によくあるものを選び、台拭き動作にはボックスティッシュ、攪拌動作にはペットボトルを利用した。また、教示者はカメラに正対するように座り、教示者の上半身と手元、教示者が持つ物体が写るようにセンサを設置した。認識結果を表現する座標系は、教示者の手元位置を原点とし、センサに対し前後方向を x 軸、左右方向を y 軸、上下方向を z 軸とした。

物体姿勢認識の結果を図3に示す。またオンラインでの認識結果を基にシミュレーション環境下でヒューマノイドロボット HIRO を用い、攪拌・台拭きの動作を行わせた結果を図4に示す。

4. 結言

図3の結果より、教示者が持つ任意形状物体の姿勢推定が正しく行われていることが分かる。また図4の結果より、任意形状物体の姿勢だけでなく手元の移動量も合わせてロボットへ伝達できていることがわかる。今後は、動きを模倣するだけでなく教示者の動きとロボット自身の対応付けを行い、自身の作業空間に合わせた姿勢の教示もできるよう改良を

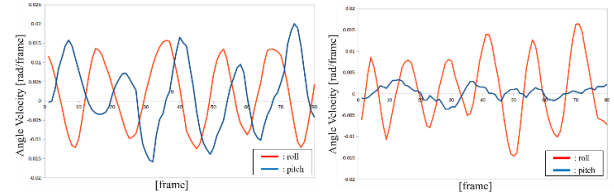


図3 教示者の持つ物体の姿勢推移の推定結果 (左：攪拌動作，右：台拭き動作)

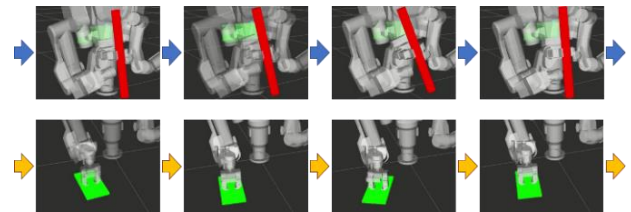


図4 オンラインでの動作教示結果 (上：攪拌動作，下：台拭き動作)

行っていく。また、作業中のロボットに対して道具操作を教示するだけでなく、強化学習の入力として利用することで、より効率良く学習できるのではないかと考えている。

参考文献

- [1] Y. Kuniyoshi, M. Inaba, and H. Inoue, "Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance", IEEE Transactions on Robotics and Automation, Vol. 10, No. 6, pp. 799--822, 1994.
- [2] M. Sieb, K. Fragkiadaki, "Data Dreaming for Object Detection: Learning Object-Centric State Representations for Visual Imitation", IEEE-RAS International Conference on Humanoid Robots, pp. 1--9, 2018.
- [3] Y. Liu, A. Gupta, P. Abbeel, S. Levine, "Imitation from Observation: Learning to Imitate Behaviors from Raw Video via Context Translation", IEEE International Conference on Robotics and Automation, pp.1118--1125, 2018.
- [4] S. Demura, Y. Mo, K. Nagahama and K. Yamazaki, "A Trajectory Modification Method for Tool Operation Based on Human Demonstration Using MITATE Technique", IEEE International Conference on Robotics and Biomimetics, pp.1915--1920, 2018.
- [5] Y. Hyakutome, "Nihon Bunka Ni Okeru MITATE To Nihongo", Kokugo Kyoiku Ronso (Domestic in Japanese), vol. 22, pp47--57, 2013.
- [6] CMU-Perceptual-Computing-Lab/openpose: <https://github.com/CMU-Perceptual-Computing-Lab/>
- [7] Eddy Ilg et al., "FlowNet2.0: Evolution of Optical Flow Estimation with Deep Networks", IEEE Conference in Computer Vision and Pattern Recognition (CVPR), 2017