

固有値テンプレートによる Fast Graspability Evaluation の高速化

○真野航輔 長谷川昂宏 山内悠嗣 山下隆義 藤吉弘亘 (中部大学)
堂前幸康 川西亮輔 関真規人 (三菱電機 (株))

1. はじめに

産業用ロボットや生活支援ロボットの重要なタスクの1つとして、ボルトやリモコン等の物体把持が挙げられる。このタスクを実現するには、ロボットに搭載されたビジョンセンサを用いて RGB 画像や距離画像を撮影し、物体の最適な把持位置を検出する必要がある。このような物体把持位置検出法は、機械学習に基づく手法とモデル当てはめに基づく手法に大別できる。モデル当てはめに基づいた把持位置検出法は、Fast Graspability Evaluation [1] が提案されており、ロボットのハンドモデルを用いて把持位置を検出する。また、把持物体の3次元点群を円柱等の単純なモデルで当てはめることで、物体の把持位置を推定する手法も提案されている [2]。機械学習に基づく把持位置検出法は、最適な把持位置を検出するために、学習用画像から得られる特徴量を用いて物体の把持位置を SVM やニューラルネットワークで学習させる [3, 4]。また、畳み込みニューラルネットワークにより画像特徴量も学習過程で獲得することで、より最適な把持位置を検出する手法も提案されている [5, 6]。様々な把持位置検出法の中でも Fast Graspability Evaluation (FGE) では、対象物体領域にハンドモデルの2値画像を畳み込むことにより、高速に把持位置を検出することができ、産業用ロボットで利用されている。しかし、高精度な把持位置検出を実現するには、ハンドの開き幅や回転角度等の状態数を増やして2値画像を畳み込む必要がある。これによりハンドの状態数に応じて計算コストが増えるという欠点がある。

本研究では、固有値テンプレート法を導入した FGE により物体の把持位置を高速に検出する。提案手法は、オフライン処理でロボットハンドと物体が衝突する領域と接触する領域のそれぞれに対してハンドテンプレート群を生成し、生成したハンドテンプレート群に特異値分解を適用することでテンプレート群をコンパクトに表現し高速化を図る。

2. Fast Graspability Evaluation

2.1 FGE による把持位置検出

Domae 等により提案された Fast Graspability Evaluation (FGE) [1] は、図1に示すようにロボットのハンドモデルからハンドの衝突領域と接触領域をテンプレートとして作成し、物体の衝突領域と接触領域に畳み込むことで、ハンドと物体が衝突しない領域を求める。最後にガウシアンフィルタをかけた Graspability の値が最大となる座標を把持位置として検出する。

ロボットハンドが対象物と衝突する領域 H_c と接触

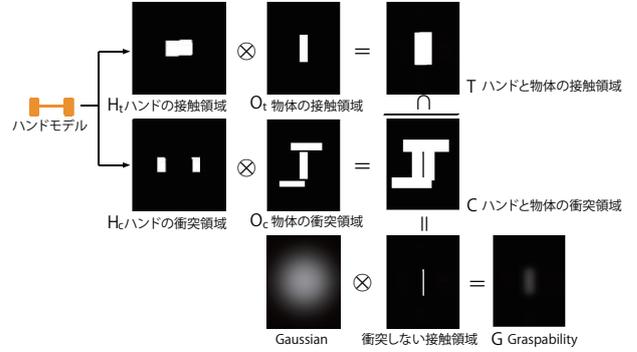


図1 Fast Graspability Evaluation の処理。

する領域 H_t の2値画像をそれぞれテンプレートとして保持する。物体の距離画像 O 、把持アプローチの際にハンドが進む深さ w と対象物の高さ h を用いて対象物とハンドの衝突領域 O_c と接触領域 O_t を定義する。位置 (x, y) における O_t の値を $O_t(x, y)$ とすると接触領域は式 (1) となる。同様に $O_c(x, y)$ は式 (2) となる。

$$O_t(x, y) = \begin{cases} 1 & (O(x, y) \geq h) \\ 0 & (\text{otherwise}) \end{cases} \quad (1)$$

$$O_c(x, y) = \begin{cases} 1 & (O(x, y) \geq h - w) \\ 0 & (\text{otherwise}) \end{cases} \quad (2)$$

(対象物とハンドモデルの接触領域) の交点 T が最大かつ (対象物とハンドモデルの衝突領域) の交点 C が最小の把持位置は Graspability が高い位置とみなすことができる。(対象物とハンドモデルの接触領域) の交点 T は式 (3) のように定義できる。(対象物とハンドモデルの衝突領域) の交点 C は式 (4) のように定義できる。

$$T = H_t \otimes O_t \quad (3)$$

$$C = H_c \otimes O_c \quad (4)$$

ここで、 \otimes は畳み込み演算子を表す。 g をガウシアンフィルタとすると、ハンドモデルと対象物が衝突しない領域 G (Graspability) は式 (5) のように定義できる。

$$G = (T \cap \bar{C}) \otimes g \quad (5)$$

2.2 FGE の問題点

FGE により、平行ハンドの把持位置を検出する場合、ハンドの開き幅 d を $\{20\text{mm}, 30\text{mm}, \dots, 60\text{mm}\}$ とし、ハンドの回転角度 θ を $\{0^\circ, 10^\circ, \dots, 170^\circ\}$ としたときの

ハンドの衝突領域のテンプレートは 90 枚、接触領域のテンプレートは 90 枚となる。この場合の畳み込み回数は 180 回である。しかし、高精度に把持位置を求めるためにはハンドの開き幅 d を $\{20\text{mm}, 25\text{mm}, \dots, 60\text{mm}\}$ とし、ハンドの回転角度 θ を $\{0^\circ, 5^\circ, \dots, 175^\circ\}$ とする。すると、ハンドの衝突領域のテンプレートは 324 枚、接触領域のテンプレートは 324 枚となる。この場合の畳み込み回数は 648 回となり、非線形に計算量が増加する。

3. 提案手法

本研究では、Fast Graspability Evaluation (FGE) で用いる大量のテンプレート群を特異値分解を用いて近似計算することで、高速に物体の把持位置を検出する。ロボットハンドが物体に接触するテンプレート群と衝突するテンプレート群は特異値分解により、固有値テンプレートと呼ばれる数十枚のテンプレート画像で近似可能となる。

3.1 ハンドテンプレート群の分解

ロボットハンドの衝突領域と接触領域それぞれに対して特異値分解を適用することで少ないテンプレートの組み合わせで衝突領域と接触領域の畳み込み画像を近似する。これにより、324 枚のハンドテンプレート群は数十枚の固有値テンプレートと重み係数の線形結合で近似することができるため、処理時間の増加を抑制することができる。特異値分解を適用するために、衝突領域のハンドテンプレート 324 枚と接触領域のハンドテンプレート 324 枚をそれぞれ 2 次元の行列 M_c, M_t で表現する。行列 $M_c \in \mathbb{R}^{E \times F}, M_t \in \mathbb{R}^{E \times F}$ の各列ベクトルにはベクトル化した 324 枚のハンドテンプレート $m_c^{(d,\theta)} \in \mathbb{R}^E, m_t^{(d,\theta)} \in \mathbb{R}^E$ で構成される。図 2 に示すように、 E はハンドテンプレートのピクセル数 (= 10000 px), F はロボットハンドの状態数 (= 324 枚) である。行列 M_c, M_t に対して特異値分解を適用す

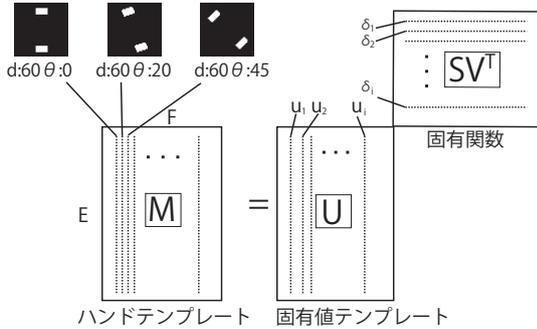


図 2 ハンドテンプレート群の特異値分解

ることで式 (6) のようにそれぞれ 3 つの行列に分解することができる。

$$\begin{aligned} M_c &= U_c S_c V_c^T \\ M_t &= U_t S_t V_t^T \end{aligned} \quad (6)$$

ここで、 $u_{c_i} \in \mathbb{R}^E, u_{t_i} \in \mathbb{R}^E$ を行列 U_c, U_t の i 番目の列ベクトル、 $\delta_{c_i} \in \mathbb{R}^F, \delta_{t_i} \in \mathbb{R}^F$ を行列 $S V_c^T, S V_t^T$ の i 番目の行ベクトルとすると、ハンドパラメータ (d, θ) のテンプレート $m_c^{(d,\theta)}, m_t^{(d,\theta)}$ は式 (7) のように定義す

ることができる。

$$\begin{aligned} m_c^{(d,\theta)} &= \sum_{i=1}^F \delta_{c_i}(d, \theta) u_{c_i} \\ m_t^{(d,\theta)} &= \sum_{i=1}^F \delta_{t_i}(d, \theta) u_{t_i} \end{aligned} \quad (7)$$

u_{c_i}, u_{t_i} は 2 次元のテンプレート画像とみなせるため、ここでは“固有値テンプレート”と呼び、重み係数となる $\delta_{c_i}, \delta_{t_i}$ を“固有関数”と呼ぶ。固有値テンプレートを可視化した例を図 3 に示す。固有関数 $\delta_{c_i}, \delta_{t_i}$ は離散的

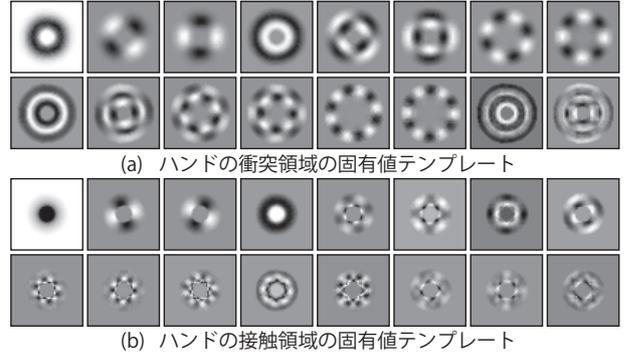


図 3 固有値テンプレートの可視化

な値をとる関数で、分解前のロボットハンドのテンプレート枚数分の値を持つ。ここで、行列 S_c, S_t は対角成分に特異値 s_{c_i}, s_{t_i} を持つ対角行列であり、 i が大きくなるにつれて特異値は 0 に近い値となる。そのため、式 (8) に示すように主要な N 枚の固有値テンプレートと固有関数で元のハンドテンプレートを十分に近似することができる。

$$\begin{aligned} m_c^{(d,\theta)} &\approx \sum_{i=1}^N \delta_{c_i}(d, \theta) u_{c_i} \\ m_t^{(d,\theta)} &\approx \sum_{i=1}^N \delta_{t_i}(d, \theta) u_{t_i} \end{aligned} \quad (8)$$

特異値分解より得られた固有値テンプレートと固有関数を用いて衝突領域と接触領域の畳み込み画像 $T^{(d,\theta)}, C^{(d,\theta)}$ は式 (9) のように物体の衝突領域と接触領域 O_t, O_c とハンドテンプレートの畳み込みで計算することができる。

$$\begin{aligned} C^{(d,\theta)} &\approx O_c \otimes \sum_{i=1}^N \delta_{c_i}(d, \theta) u_{c_i} \\ T^{(d,\theta)} &\approx O_t \otimes \sum_{i=1}^N \delta_{t_i}(d, \theta) u_{t_i} \end{aligned} \quad (9)$$

3.2 固有値テンプレートによる Graspability の効率的な計算

対象物の画像 O_t, O_c と固有値テンプレート u_{c_i}, u_{t_i} はあらかじめ畳み込みが可能である式 (10)。すると、固有関数の値のみを変化させて式 (10) との積をとること

で、全てのロボットハンドの状態の畳み込み画像を計算することができる。

$$\begin{aligned} q_{c_i} &= O_c \otimes u_{c_i} \\ q_{t_i} &= O_t \otimes u_{c_i} \end{aligned} \quad (10)$$

衝突領域または接触領域の畳み込み画像 $C^{(d,\theta)}, T^{(d,\theta)}$ は式 (11) のように効率的に計算することができる (図 4)。

$$\begin{aligned} C^{(d,\theta)} &\approx \sum_{i=1}^N \delta_{c_i}(d,\theta) q_{c_i} \\ T^{(d,\theta)} &\approx \sum_{i=1}^N \delta_{t_i}(d,\theta) q_{t_i} \end{aligned} \quad (11)$$

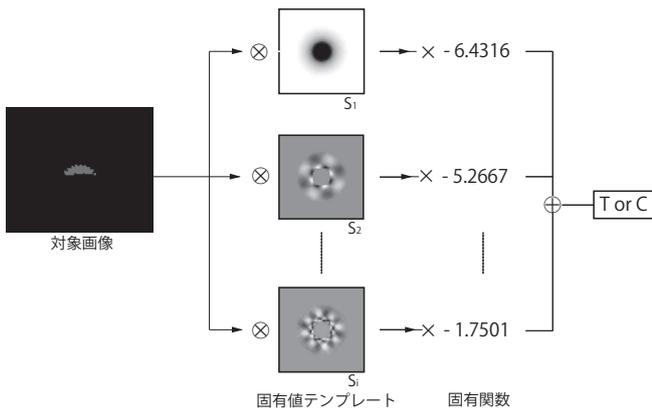


図 4 畳み込み画像の近似.

4. 評価実験

提案手法の有効性を確認するために評価実験を行う。比較手法には Fast Graspability Evaluation(FGE) を用いる。

4.1 実験概要

本研究では、評価実験用の距離画像として、工業用部品のバラ積み画像を用いる。

提案手法と従来法 (FGE) を用いて、図 5 に示す平行ハンドと多指ハンド (3 指) のモデルを使用した際の処理時間と把持位置の検出精度を比較する。従来法と提案手法の平行ハンドの開き幅を 20mm から 60mm まで 5mm 刻み、回転角度を 0° から 175° まで 5° 刻みで変化させる。多指ハンド (3 指) は開き幅を 20mm から 60mm まで 5mm 刻み、回転角度を 0° から 115° まで 5° 刻みで変化させる。

4.2 処理時間の比較

提案手法と FGE のハンドの状態数を変化させたときの把持位置検出処理時間を図 6 に示す。(a) の黄色は平行ハンドの従来法、橙色は提案手法、(b) の青色は多指ハンド (3 指) の従来法、緑色は提案手法の多指ハンド (3 指) の状態数を変化させたときの 1 箇所の検出にかかる平均時間を表す。図 6 より、ハンドの状態を (2mm, 2°) 刻みの時の提案手法は従来法の FGE に比べ平行ハンドでは約 2.4 秒、多指ハンド (3 指) では約 3.1 秒

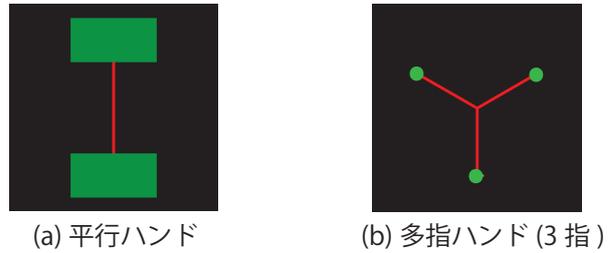


図 5 ハンドモデル

短縮し、(1mm, 1°) では、平行ハンドは約 9.7 秒、多指ハンド (3 指) では約 14.9 秒短縮した。従来法と比べ平行ハンドは約 1.7 倍、多指ハンド (3 指) は約 1.5 倍に処理時間を短縮できた。

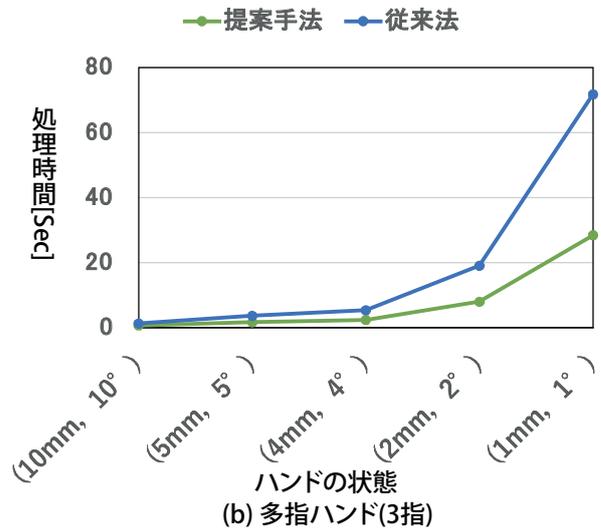
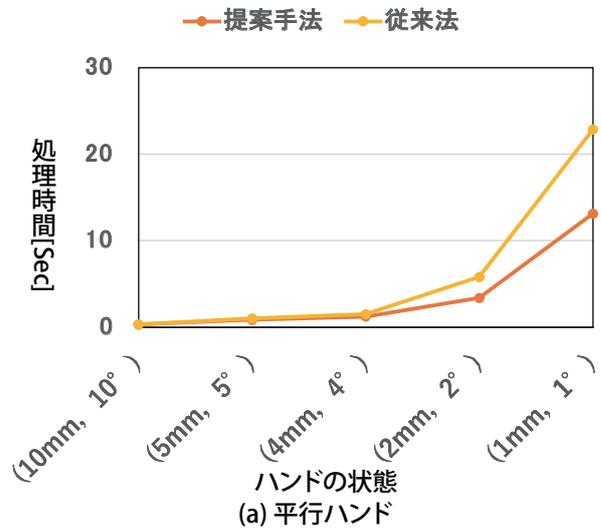


図 6 提案手法と従来法の処理時間

4.3 把持位置検出例

提案手法と従来法の検出結果の例を図 7 に示す。緑色と赤色のハンドモデルが検出結果である。赤色の線

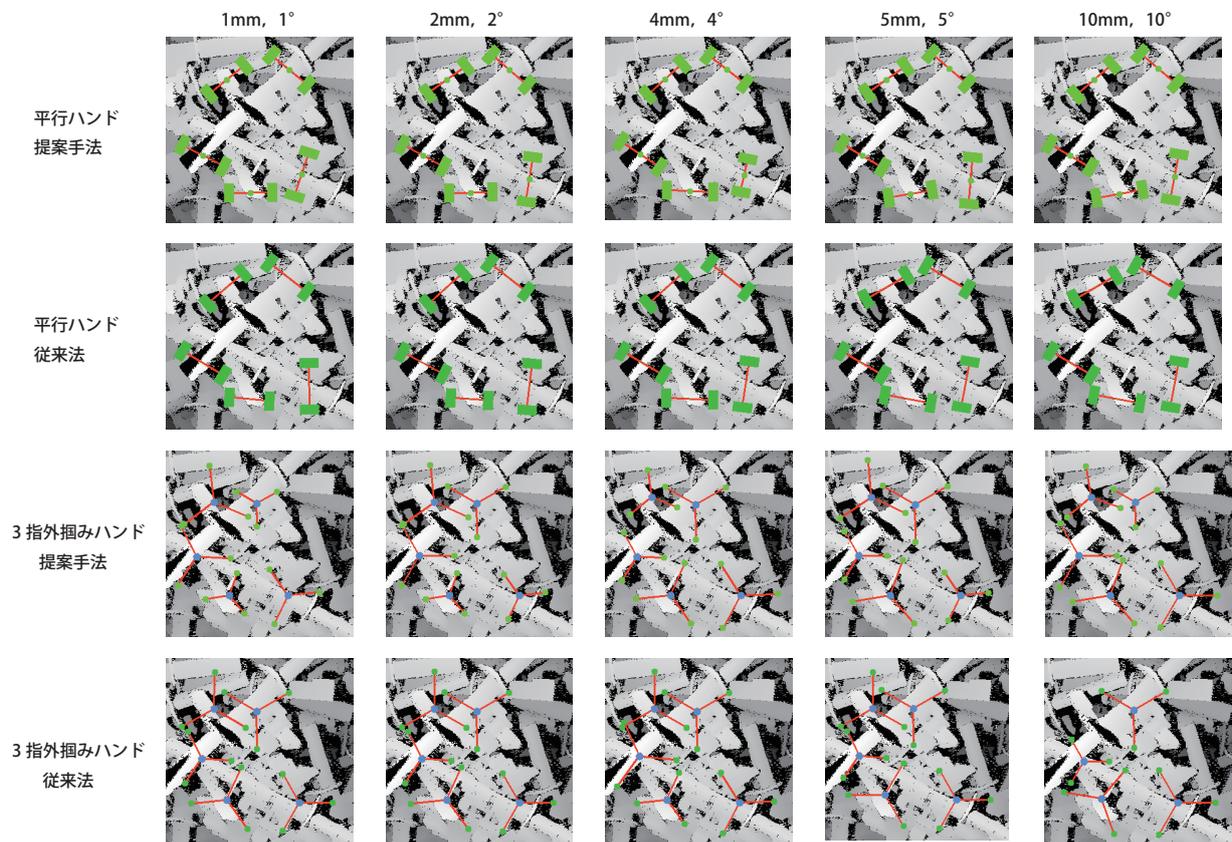


図7 提案手法とFGEの把持位置検出の例

はハンドの接触領域にあたり、緑色はハンドの衝突領域を示している。提案手法の平行ハンドと多指ハンド(3指)は従来法とほぼ同じ把持位置を検出しており、処理の高速化を実現できた。

5. おわりに

本研究では、Fast Graspability Evaluationに固有値テンプレート法を導入した高速かつ高精度な物体把持位置検出を提案した。Fast Graspability Evaluationのハンドの状態数により計算コストが増加する問題を固有値テンプレートにより近似計算を用いることで、精度を維持したまま計算コストを抑えることができた。今後は、Graspabilityを連続関数を用いて計算することで効率的な把持位置探索と、ロボットマニピュレータを用いた評価を行う予定である。

参考文献

- [1] Y. Domae, H. Okuda, Y. Taguchi, K. Sumi, and T. Hirai, "Fast graspability evaluation on single depth maps for bin picking with general grippers", International Conference on Robotics and Automation, pp.1997–2004, 2014.
- [2] K. Harada, K. Nagata, T. Tsuji, N. Yamanobe, A. Nakamura, and Y. Kawai, "Probabilistic approach for object bin picking approximated by cylinders", International Conference on Robotics and Automation, pp. 3742–3747, 2013.
- [3] Y. Jiang, S. Moseson, and A. Saxena, "Efficient grasping from rgb-d images: Learning using a new rectangle representation", International Conference on Robotics and Automation, pp. 3304–3311, 2011.

- [4] I. Lenz, H. Lee, and A. Saxena, "Deep Learning for Detecting Robotic Grasps", International Journal of Robotics Research, vol. 34, no. 4-5, pp. 705–724, 2015.
- [5] J. Redmon, and A. Angelova, "Real-time grasp detection using convolutional neural networks", International Conference on Robotics and Automation, pp. 1316–1322, 2015.
- [6] 荒木諒介, 長谷川昂宏, 山内悠嗣, 山下隆義, 藤吉弘亘, "Graspabilityを導入したDCNNによる物体把持位置検出", 2016.