固有値テンプレートによる高分解能把持姿勢推定

真野 航輔 † 長谷川昂宏 † 山内悠嗣 † 山下隆義 † 藤吉弘亘 † 堂前幸康 ‡ 川西亮輔 ‡ 関真規人 ‡

†中部大学 ‡三菱電機

E-mail: runrun70@mprg.cs.chubu.ac.jp

Abstract

物体の把持位置検出法として Fast Graspability Evaluation(FGE) が提案され、産業用ロボットに利用されて いる. FGE は対象物体領域にハンドテンプレートを畳 み込むことで最適な把持姿勢を推定する.しかし,高 精度に把持姿勢を推定するにはハンドの開き幅や回転 角度の分解能を高く設定して求める必要があるため,計 算量が増加する. そこで, 本研究では特異値分解によ りハンドテンプレートをコンパクトに表現し, 高速化 を行う手法を提案する. 特異値分解により, ハンドテ ンプレートは少数の固有値テンプレートと固有関数の 線形結合で表現できる. また, 固有関数は離散的な値 を取るため連続関数のフィッテングにより任意のパラ メータの応答値を計算することが可能となる. 評価実 験より、従来の FGE と同等の検出精度において、計算 時間を平行ハンドと3指ハンド共に約1/3に短縮でき ることを確認した.

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つに、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、対象物体に対して最適な把持位置を検出する必要がある。このような物体の把持位置検出法は、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、距離画像と3次元点群を用いた2種類がある。堂前らにより提案された Fast Graspability Evaluation [1] は、距離画像からロボットのハンドモデルを用いて把持位置を検出する。3次元点群からの把持位置検出には、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手法[2]や、把持物体の3次元モデルと作業エリア内3次元点群を用いる手法が提案されている[3,4,5]。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を

用いて学習する [6,7]. また,畳み込みニューラルネットワーク [8] を用いて,より最適な把持位置を検出する手法が提案されている [9,10].

これらの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、産業用ロボットで実利用されている. 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たハンドテンプレートを2値画像に畳み込む必要があり、ハンドの状態数に応じて計算コストが増加するという問題がある.

本研究では、固有値テンプレート法を導入した FGE により物体把持位置を高速に検出する手法を提案する. 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する. 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. また、固有関数に対して連続関数によりフィッテングし、任意の分解能でハンドモデルを近似することで、把持位置の検出精度の高精度化を行う.

2 関連研究

本節では、モデル当てはめに基づいた把持位置検出の従来法である Fast Graspability Evaluation について説明する.

2.1 Fast Graspability Evaluation

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み、ガウシアンフィルタを畳み込むことで Graspability マップを生成する. そして、Graspability マップを生成する. そして、Graspability マップの値がピークとなる位置を把持位置として検出する. ハンドの回転角度や開き幅といった分解能を変化させることにより、最適な把持姿勢を検出する. 以下にその処理過程を述べる.

ロボットハンドが対象物と衝突する領域 H_c と接触す

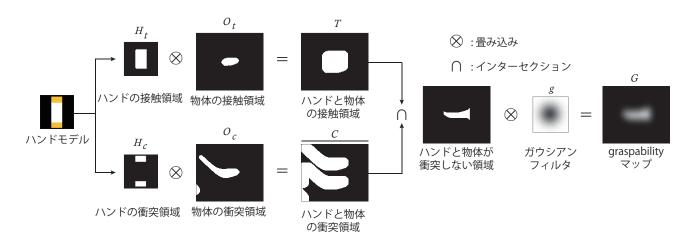


図 1 Fast Graspability Evaluation の処理

る領域 H_t を入力された距離画像から作成する. 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 から,式 (1) を用いて座標 (x,y) における接触領域 $O_t(x,y)$ と衝突領域 $O_c(x,y)$ を求める.

$$O_{t}(x,y) = \begin{cases} 1 & (O(x,y) \ge h) \\ 0 & (\text{otherwise}) \end{cases}$$

$$O_{c}(x,y) = \begin{cases} 1 & (O(x,y) \ge h - w) \\ 0 & (\text{otherwise}) \end{cases}$$
 (1)

対象物とハンドモデルの接触領域の値が最大かつ対象物とハンドモデルの衝突領域値が最小の把持座標は Graspability が高い座標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 T 及び対象物とハンドモデルの衝突領域 C は式 (2) により求める。

$$T = H_t \otimes O_t$$

$$C = H_c \otimes O_c$$
 (2)

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 (3) により求める。

$$G = (T \cap \overline{C}) \otimes g \tag{3}$$

ハンドモデルと対象物が衝突しない領域 G の最大となる座標と分解能から把持姿勢を推定する.

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 ハンドの開き幅と回転角度を考慮する必要がある。ハンドの開き幅 d を $\{20\text{mm},30\text{mm},\cdots,60\text{mm}\}$ とし、ハンドの回転角度 θ を $\{0^\circ,10^\circ,\cdots,170^\circ\}$ 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 回となる。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 $\{20\text{mm},25\text{mm},\cdots,60\text{mm}\}$ とし、ハ

ンドの回転角度 θ を $\{0^\circ,5^\circ,\cdots,175^\circ\}$ と状態数を増加する. このとき, 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. 畳み込み回数は 648 回となり, 計算量が大幅に増加することになる.

3 提案手法

本研究では、FGE で用いる大量のハンドテンプレート群に対して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.

3.1 特異値分解による固有値テンプレートの算出

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固有値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、ハンドの開き幅 d を $\{20,25,30,\cdots,60\}$ 、回転角度 θ を $\{0,5,10,\cdots,175\}$ とした際の 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 (= 2500 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ると,式 (4) のように 3 つの行列 U, S, V^T に分解することができる.

$$M_c = U_c S_c V_c^{\mathrm{T}}$$

$$M_t = U_t S_t V_t^{\mathrm{T}}$$
(4)

図 2 に示すように、行列 S_c , S_t は対角成分に特異値 S_{c_i} , S_{t_i}

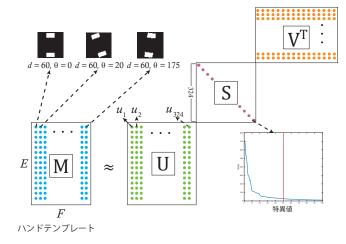


図2 ハンドテンプレート群の特異値分解

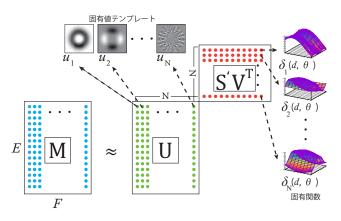


図3 寄与率を用いた行列のコンパクト化

を持つ対角行列であり、特異値は上位の要素のみ大きな値を持ち、下位の要素では0に近い値をとる.ここで、行列 S_c , S_t の特異値の値から求めた寄与率用いて、次元削減をした行列を S_c' , S_t' とすると、式(4) は式(5) となる.

$$M'_{c} = U_{c}S'_{c}V_{c}^{\mathrm{T}}$$

$$M'_{t} = U_{t}S'_{t}V_{t}^{\mathrm{T}}$$
(5)

 $u_{c_i} \in \mathbb{R}^E, u_{t_i} \in \mathbb{R}^E$ 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 $S'V_c^\mathsf{T}, S'V_t^\mathsf{T}$ の i 番目の行ベクトルとすると、パラメータ (d,θ) のテンプレート $m_c^{(d,\theta)}, m_t^{(d,\theta)}$ は式 (6) のように定義することができる.

$$m_c^{(d,\theta)} = \sum_{i=1}^{N} \delta_{c_i}(d,\theta) u_{c_i}$$

$$m_t^{(d,\theta)} = \sum_{i=1}^{N} \delta_{t_i}(d,\theta) u_{t_i}$$
(6)

図3に示すように、 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 δ_{c_i} 、 δ_{t_i} を"固有関数"と呼ぶ。平行 ハンドの固有値テンプレートを可視化した例を図4に、3指ハンドの固有値テンプレートを可視化した例を図5

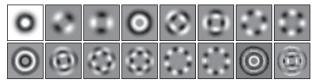
に示す. 固有関数 δ_{c_i} , δ_{t_i} は離散関数であり、分解前の



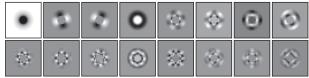




(a) ハンドモデルにおける衝突領域と接触領域



b) ハンドの衝突領域における固有値テンプレート



(c) ハンドの接触領域における固有値テンプレー

図4 平行ハンドにおける固有値テンプレートの 可視化







a) ハンドモデルにおける衝突領域と接触領域



(b) ハンドの衝突領域における固有値テンプレート



図 5 3 指ハンドにおける固有値テンプレートの 可視化

ロボットハンドのテンプレート枚数分の値を持つ.

衝突領域と接触領域の畳み込み画像 $C'^{(d,\theta)}, T'^{(d,\theta)}$ は,式 (7) のように物体の衝突領域と接触領域 O_t, O_c と近似により求めた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.

$$C'^{(d,\theta)} = O_c \otimes \sum_{i=1}^{N} \delta_{c_i}(d,\theta) u_{c_i}$$

$$T'^{(d,\theta)} = O_t \otimes \sum_{i=1}^{N} \delta_{t_i}(d,\theta) u_{t_i}$$
(7)

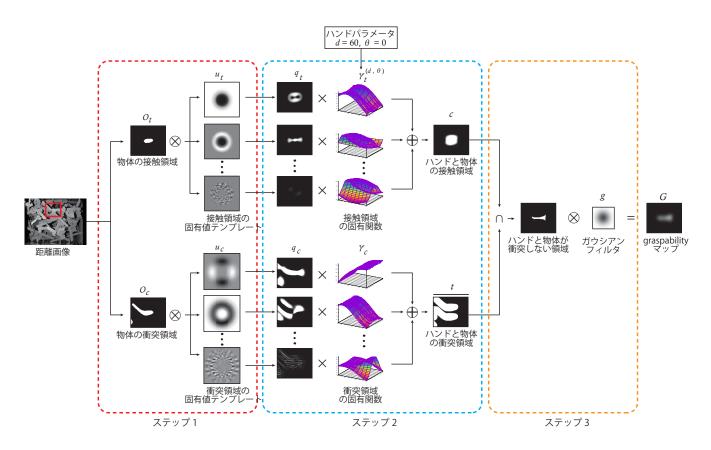


図6 提案手法の把持位置検出の流れ

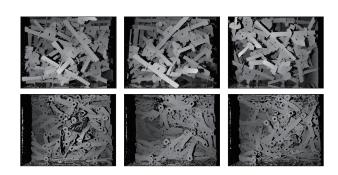


図7 評価に用いる距離画像の例

3.2 連続関数による固有関数のフィッテング

特異値分解から得られる固有関数 δ_{c_i} 、 δ_{l_i} は離散的な値しか持たない.そのため,ハンドの分解能は特異値分解に使用したパラメータでしか再構成することができない.そこで,SVD から得られた離散的な固有関数 δ_{c_i} 、 δ_{l_i} を線形補間もしくは,連続関数によるフィッティングを行う.線形補間と連続関数によるフィッテングはそれぞれオフラインで計算を行う.固有関数の線形補間では,開き幅と回転角度それぞれに線形補間を行うことで任意の分解能を近似する.固有関数の連続関数によるフィッテングは,以下のような関数モデル $\gamma_{c_i}(d,\theta)$ 、 $\gamma_{l_i}(d,\theta)$ を定義する.

$$\gamma_{c_{i}}(d,\theta) = \sum_{j=0}^{J} \sum_{k=0}^{K} \alpha_{j,k} d^{j} \cos(k\theta) + \sum_{j=0}^{J} \sum_{k=1}^{K} \beta_{j,k} d^{j} \sin(k\theta)
\gamma_{i_{i}}(d,\theta) = \sum_{j=0}^{J} \sum_{k=0}^{K} \alpha_{j,k} d^{j} \cos(k\theta) + \sum_{j=0}^{J} \sum_{k=1}^{K} \beta_{j,k} d^{j} \sin(k\theta)$$
(8)

J, K は連続関数モデルの時数, $\alpha_{j,k}$, $\beta_{j,k}$ は未知係数であり,以下の最小化問題により未知係数を算出する.

$$\underset{\alpha,\beta}{\arg\min} (\sum_{d} \sum_{\theta} (\delta_{c_{i}}(d,\theta) - \gamma_{c_{i}}(d,\theta))^{2}), \qquad (9)$$

$$d = \{20,25,30,\dots,60\},$$

$$\theta = \{0,5,10,\dots,175\}$$

本研究ではJ=4, K=12 とすることで元の固有関数を近似することができた. 固有関数を連続関数モデルで表現することにより、任意の分解能のハンドテンプレートを再構成することができる.

3.3 固有値テンプレートによる **Graspability** の効率的 か計算

図 6 に提案手法による把持位置検出の流れを示す.ステップ 1 では,対象物の 2 値画像 O_t, O_c と固有値テンプレート u_{c_i}, u_{t_i} は, $q_{c_i} = O_c \otimes u_{c_i}, q_{t_i} = O_t \otimes u_{c_i}$ のようにあらかじめ畳み込めるため,式 7 は式 11 のように置き換えることができる.ステップ 2 では,衝突領域または接触領域 $C'^{(d,\theta)}, T'^{(d,\theta)}$ は式 (10) のように固有関数

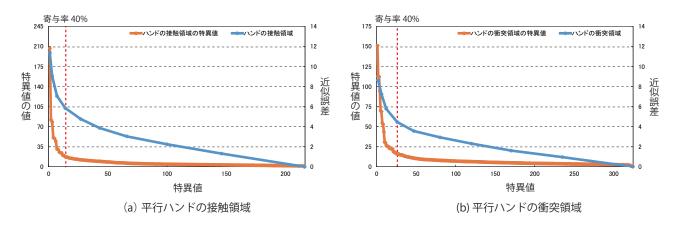


図8 平行ハンドの近似誤差と特異値の関係

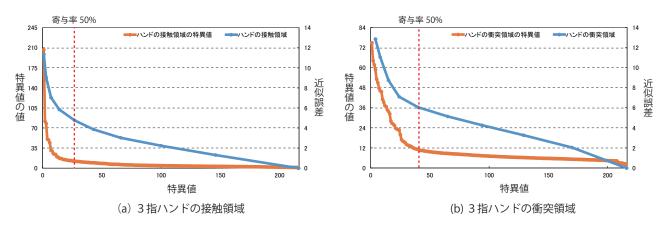


図9 3指ハンドの近似誤差と特異値の関係

の値のみを変化することで任意の開き幅 d と回転角度 θ における出力を効率的に計算することができる.

$$C'^{(d,\theta)} = \sum_{i=1}^{N} \gamma_{c_i}(d,\theta) q_{c_i}$$

$$T'^{(d,\theta)} = \sum_{i=1}^{N} \gamma_{t_i}(d,\theta) q_{t_i}$$
(10)

ステップ 3 では,graspability マップ $G^{(d,\theta)}$ を,次式のようにガウシアンフィルタ g の畳み込みで求める.

$$G^{(d,\theta)} = (T'^{(d,\theta)} \cap \overline{C'}^{(d,\theta)}) \otimes g \tag{11}$$

graspability マップの最大値の位置を把持位置とする.

4 評価実験

提案手法の有効性を確認するために FGE と検出誤差 及び処理時間の比較を行う.

4.1 実験概要

実験では、2種類の工業用部品のバラ積み画像 34 枚を評価に用いる (7). FGE の分解能 $[0.1mm, 0.1^{\circ}]$ を基準とし、平行ハンド及び 3 指ハンドモデルの分解能の変化における検出誤差の評価を行う。平行ハンドの開き幅を 20mm から 60mm, 回転角度を 0° から 175° と

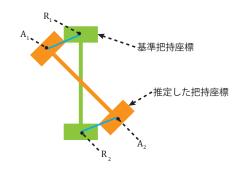


図 10 評価式の説明

し,3 指ハンドの開き幅を 20mm から 60mm,回転角度 を 0° から 115° とする.開き幅を 5mm,回転角度を 5° 間隔で変化する場合を [5mm, $5^{\circ}]$ とする.提案手法は [5mm, $5^{\circ}]$ の分解能で作成したハンドテンプレートを用いて特異値分解を行う.行列 S_c, S_t に対して寄与率を計算し,寄与率を変化させた際のハンドテンプレートの近似誤差と特異値から最適な寄与率を決定する.

近似誤差の算出は式 (12) に示すハンドテンプレート と近似により求めたハンドテンプレートに対してフロ ベニウスノルムを用いることで算出する.

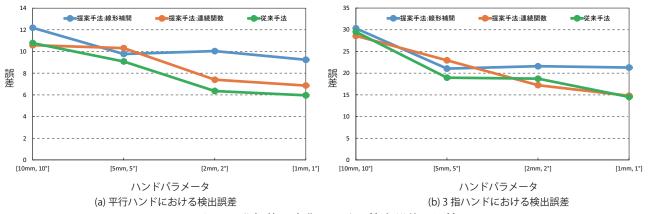


図 11 分解能の変化における検出誤差の比較

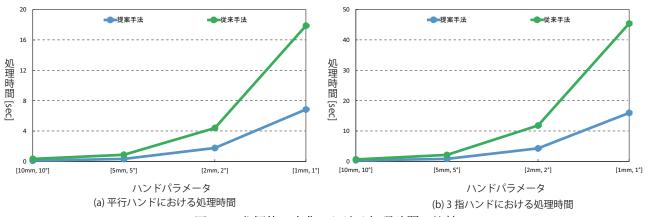


図 12 分解能の変化における処理時間の比較

衝突領域の近似誤差 =
$$\sqrt{\sum_{i}^{I} \sum_{j}^{J} (H_{c}^{(d,\theta)} - m_{c}^{(d,\theta)})^{2}}$$
接触領域の近似誤差 = $\sqrt{\sum_{i}^{I} \sum_{j}^{J} (H_{t}^{(d,\theta)} - m_{t}^{(d,\theta)})^{2}}$ (12)

把持位置の検出誤差の算出には式 (13) に示すように、基準座標 R_1 , R_2 と検出座標 A_1 , A_2 のユークリッド 距離により算出する.

Detection Error =
$$\sqrt{(R_1 - A_1)^2 + (R_2 - A_2)^2}$$
 (13)

4.2 寄与率の変化による近似精度

平行ハンドにおけるハンドテンプレートの近似誤差と特異値の関係を図8に示し、3指ハンドにおけるハンドテンプレートの近似誤差と特異値の関係を図9に示す。ここで、グラフの第1軸は特異値の値を表し、2軸は、ハンドテンプレートの近似誤差を表している。図8(a)より、平行ハンドの接触領域において近似誤差の小さい寄与率90%では特異値を235個必要とし計算コストが高い。近似誤差は7個目から近似誤差の増加量が増えるため特異値は16個を使用する。この時の寄与率は40%である。図8(b)より、平行ハンドの衝突領域も接触領域と同様に、近似誤差の変化量が12個から増

えるため 26 個を使用する. この時の寄与率は 40%である.

図 9(a) より、3 指ハンドの接触領域では、特異値の値の変化量が14 個から一定となり、近似誤差の増加量も14 個から増える.このことから、特異値27 個を使用する.この時の寄与率は50%である.図 9(b) より、3 指ハンドの衝突領域では、特異値の変化量が40 個から一定となり、近似誤差の増加量が24 個から増加するため特異値を40 個使用する.この時の寄与率は50%である.

4.3 検出精度と処理時間の比較

提案手法と従来手法 (FGE) の分解能を変化させたときの検出誤差を図 11, に示す. また, 1 アイテムあたりの把持姿勢推定時間を図 12 に示す. 図 11 より, 平行ハンドと 3 指ハンドで固有関数の線形補間における [10mm, 10°] と [5mm, 5°] は従来手法と同等の精度である. しかし, $[2mm, 2^\circ]$ と $[1mm, 1^\circ]$ では, 固有関数の線形補間における $[5mm, 5^\circ]$ と同等の検出精度であることがわかる. この結果より, 固有関数の線形補間は把持位置の高精度化にあまり貢献しないことがわかる. 一方で固有関数の連続関数によるフィッテングでは, 平行ハンドと 3 指ハンドにおいて $[2mm, 2^\circ]$ と $[1mm, 1^\circ]$ の検出誤差が従来手法と同等の検出誤差である. これにより, 固有関数の連続関数によるフィッテングを用

いることは有効であるといえる. 図 12(a) より,平行 ハンドにおける把持姿勢推定時間を提案手法は従来手 法に比べ $[5mm,5^\circ]$ では、約 62%, $[1mm,1^\circ]$ では、約 61%の処理時間を削減できることを確認した. 図 12(b) より、3指ハンドにおける把持姿勢推定時間を提案手法 は従来手法に比べ $[5mm,5^\circ]$ では、約 45%, $[1mm,1^\circ]$ では、約 64%削減できた. 以上の結果より、連続関数のフィッテングにより高速化と高精度化の両方の実現が可能である.

4.4 把持位置検出例

提案手法と従来法のバラ積み部品の把持検出の例を図 13 に示す.ここで,赤色と緑色は推定した把持姿勢である.提案手法の検出結果は従来手法とほぼ変わらない結果であることが確認できる.また平行ハンドでは,提案手法の連続関数のフィッテングと線形補間において $[10\text{mm}, 1^\circ]$ から $[1\text{mm}, 1^\circ]$ に変化するにつれ物体に対して垂直な把持位置を検出していることが確認できる.また, $[1\text{mm}, 1^\circ]$ において連続関数によるフィッテングで推定した検出結果は線形補間より,従来手法に近いことが確認できる.

3指ハンドの [10mm, 10°] では,線形補間の検出結果は,連続関数によるフィッテングより従来手法の検出結果に近いことが確認できる.しかし,[1mm, 1°] において,連続関数によるフィッテングは,線形補間より従来手法の検出結果と同等であることが確認できる.

5 おわりに

本研究では、固有値テンプレートを導入することで、高分解能把持姿勢推定を高速化する手法を提案した. 従来の Fast Graspability Evaluation において、分解能を高くすると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することで、精度を維持したまま計算コストを約1/3に抑えることができた. また、固有関数に対して連続関数のフィッテングを行うことで把持位置検出の高精度化を実現した. 今後は、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] B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3D object recognition", Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference, pp. 998–1005, 2010.
- [4] C. Choi, Y. Taguchi, O. Tuzel, M. Liu, and S. Ramalingam, "Voting-based pose estimation for robotic assembly using a 3D sensor", International Conference on Robotics and Automation, pp. 1724– 1731, 2012.
- [5] D. Chetverikov, D. Svirko, D. Stepanov, and P. Krsek, "The trimmed iterative closest point algorithm", International Conference on Pattern Recognition, pp. 545–548, 2002.
- [6] Y. Jiang, S. Moseson, and A. Saxena, "Efficient grasping from rgbd images: Learning using a new rectangle representation", International Conference on Robotics and Automation, pp. 3304–3311, 2011.
- [7] 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.
- [8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-Based Lerning Applied to Document Recognition", Proceedings of the IEEE, vol, 86, pp. 2278–2324, 1998.
- [9] J. Redmon, and A. Angelova, "Real-time grasp detection using convolutional neural networks", International Conference on Robotics and Automation, pp. 1316–1322, 2015.
- [10] 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", 日本ロボット学会, 2016.

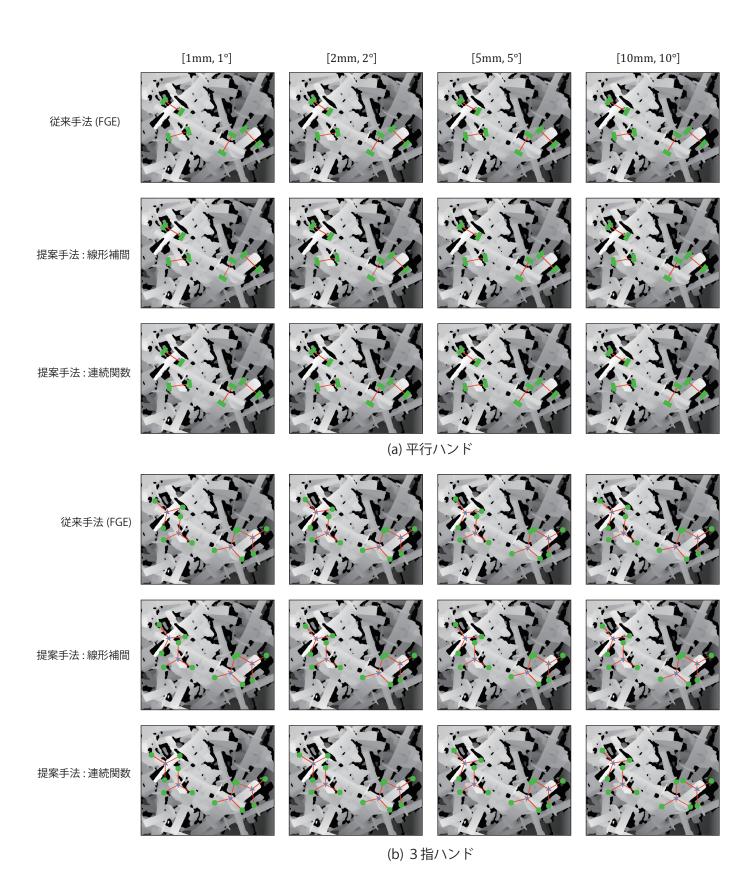


図 13 提案手法と FGE の把持位置検出の例