

写真ベース認識法を用いたビジュアルサーボシステム

Visual Servoing System by Photo-model-based Matching Method

○学 椿伊吹 (岡山大)

正 田宏志 (岡山大)

正 見浪護 (岡山大)

非 舟久保龍希 (岡山大)

学 寇ギョウ郡 (岡山大)

非 松野隆幸 (岡山大)

Ibuki TSUBAKI, Okayama University, pn7k4tkl@s.okayama-u.ac.jp

Ryuki FUNAKUBO, Okayama University

Hongzhi TIAN, Okayama University

Yejun KOU, Okayama University

Mamoru MINAMI, Okayama University

Takayuki MATSUNO, Okayama University

Recently, robots have been used in industries for mass production with countless merits. However, there remain many challenges for robots in recognition, pose (position and orientation) detection operations, especially when the working object is deformable and every working object has unique shape and color. This paper proposes Photo-model-based Matching Method that has been adopted since the photo-model can be made at once by taking a photo of the target. In this paper, I report recognition accuracy and the results of this visual servoing system.

Key Words: Visual servoing, Photo-model-based, Dual-eye

1 緒言

1970 年に「高齢化社会」になって以降、日本は 2007 年に 65 歳以上の高齢化率が 21.5 % となり「超高齢社会」に入った。内閣府による「平成 27 年版高齢社会白書」[1]によると、2014 年 10 月の時点の高齢化率は 26 % で、今後も日本の高齢化率は上昇傾向が続くとみられている。さらに少子化も伴い、今後少子高齢化が進んでいく日本において、労働力人口の減少が大きな問題となっている。日本の労働力人口は、1998 年をピークに現在に至るまで減少し続けており、この先も減少傾向であると推測されている。その影響により、これまで人間が行ってきた作業を代替するロボットのニーズが高まってきている。

ハンドアイカメラロボットはカメラ視点を自由に選べる利便性から産業用ロボットとしてピンピッキングなどに広く用いられているが、単眼ハンドアイを用いている場合が多くカメラ視線奥行き方向の位置・姿勢計測精度が良くないという問題がある。このためカメラとレーザーレンジファインダの組み合わせた認識方法[2],[3]も研究されているが、カメラ画像内の認識対象物とレンジファインダで距離を検出した対象物が同一物体であるという前提「対象物体の同一性」があり、これが満たされていないとき誤認識しロボットの誤動作の原因となる。複眼カメラ画像情報から対象物の 3 次元位置・姿勢を計測する情報処理においても似た問題が発生する。すなわち 3 次元空間内の物体上のある点が複眼カメラ画像の中のどの点に対応しているかという「対応点の同一性」の問題である。対応点の同一性が満たされない場合、対象物の 3 次元位置・姿勢情報は誤差を含む[4],[5]ことになる。一方、ビジュアルサーボに基づくロボット制御[6]の研究では、物体の運動モデルと非線形オブザーバを利用した物体運動の予測手法が提案されているが、認識誤差がゼロ近くまで減少するまでに時間がかかるという問題がある。また、単眼ではカメラの奥行き方向の対象物の位置・姿勢計測が困難であることが指摘され、Luca らは既存のカメラの移動視に基づいてカメラと対象物までの距離を推定する方法を提案している[6]。

我々の研究グループでは、衣服の分類作業に対して Model-based Matching 法及び遺伝的アルゴリズムを適用し、指定する属性(対象物の形状と色)を持つ対象物の認識とハンドリングの方法を提案してきた[7]。我々の提案する写真ベース認識法は対象物の写真からモデルを作成するためロボットに対象を指示することが容易であり、対象の既知情報は写真のみで認識が可能である。本

研究では、水中生物の捕獲作業を想定して、蟹の模型に対するビジュアルサーボを行っている。本報では、対象物の位置姿勢の認識精度の確認を行い、ビジュアルサーボを行った結果について報告する。

2 対象物の位置姿勢計測法

本手法は Fig.1 のようなマニピュレータロボット (PA10-三菱重工業株式会社製) に 2 台のカメラを搭載したシステムを基準に構成されている。

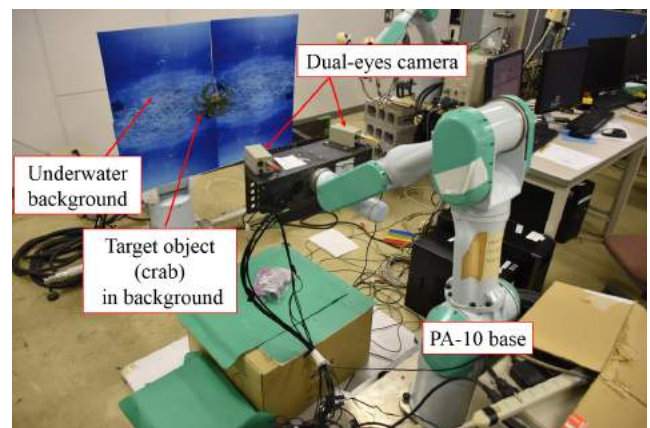


Fig.1 PA10

2.1 適合度関数の定義

左右ビデオカメラから入力される画像は、0 ~ 359 で表される色相値によって構成される。探索モデルは、Fig.2 に示すように対象物の評価とその周囲の色相変化を評価するために、内部領域 $S_{R,in}$, $S_{L,in}$ と帯領域 $S_{R,out}$, $S_{L,out}$ によって構成される。ここで、右画像領域 ${}^{IR}r_i^j(\phi_M^j)$ における色相値に基づく評価値を

$p_R(I^R r_i^j(\phi_M^j))$, 左画像領域に $I^L r_i^j(\phi_M^j)$ における色相値に基づく評価値を $p_L(I^L r_i^j(\phi_M^j))$ とする. 対象物を評価する内部領域の評価値を “+2”, 背景を評価する内部領域の評価値を “-0.005”, 背景を評価する帯領域の評価値を “-1” と設定する. また, 対象物を評価する内部領域において, モデルの色相値とカメラ画像の色相値が 50 以上離れていた場合は, 評価値を -1 と設定する. このような正負の評価値をもつ探索モデルと画像との相関を表す適合度関数を次式に示す.

$$F(\phi_M^j) = \left\{ \left(\sum_{I^R r_i^j \in S_{R,in}(\phi_M^j)} p_{R,in}(I^R r_i^j(\phi_M^j)) + \sum_{I^R r_i^j \in S_{R,out}(\phi_M^j)} p_{R,out}(I^R r_i^j(\phi_M^j)) \right) + \left(\sum_{I^L r_i^j \in S_{L,in}(\phi_M^j)} p_{L,in}(I^L r_i^j(\phi_M^j)) + \sum_{I^L r_i^j \in S_{L,out}(\phi_M^j)} p_{L,out}(I^L r_i^j(\phi_M^j)) \right) \right\} / 2$$

$$= \{F_R(\phi_M^j) + F_L(\phi_M^j)\} / 2 \quad (1)$$

$$p_{L,in}(I^L r_i^j(\phi_M^j)) = \begin{cases} 2, & \text{if } (|H_{IL}(I^L r_i^j(\phi_M^j)) - H_{ML}(I^L r_i^j(\phi_M^j))| \leq 30); \\ -1, & \text{if } (|H_{IL}(I^L r_i^j(\phi_M^j)) - H_{ML}(I^L r_i^j(\phi_M^j))| \geq 50); \\ -0.005, & \text{if } (|\bar{H}_B - H_{ML}(I^L r_i^j(\phi_M^j))| \leq 30); \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

$$p_{L,out}(I^L r_i^j(\phi_M^j)) = \begin{cases} 0.1, & \text{if } (|\bar{H}_B - H_{ML}(I^L r_i^j(\phi_M^j))| \leq 20); \\ -0.5, & \text{otherwise.} \end{cases} \quad (3)$$

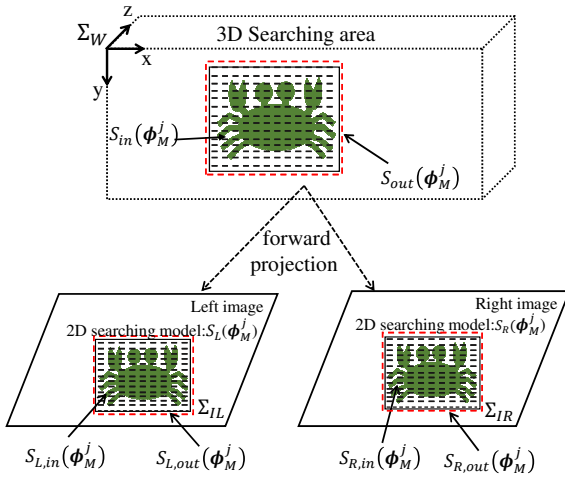


Fig.2 Searching model

この関数は, 右画像領域 $I^R r_i^j(\phi_M^j)$ の中で, $S_{R,in}(\phi_M^j)$ に重なる入力画像の評価値 $p_{R,in}(I^R r_i^j(\phi_M^j))$ の合計値と, $S_{R,out}(\phi_M^j)$ の評価値 $p_{R,out}(I^R r_i^j(\phi_M^j))$ の合計値を加えることで, 右画像の適合度 $F_R(\phi_M^j)$ を得る. 左画像においても同様に適合度 $F_L(\phi_M^j)$ を計算し, これらを加算し平均をとった適合度関数 $F(\phi_M^j)$ を用いて, 左右の画像を同時に評価する. この適合度関数 $F(\phi_M^j)$ は, 立体モデルの持つ位置/姿勢情報より一意に決定する関数であり, 立体モデルが対象物と一致すれば, 左右画像においても対象物と探索モデルが一致するはずである. ただし, $F_R(\phi_M^j) \leq 0$, または, $F_L(\phi_M^j) \leq 0$ の場合, $F_R(\phi_M^j) = 0$, または, $F_L(\phi_M^j) = 0$ とする.

2.2 GA を用いた最適探索法

前節で述べてきた適合度関数を用いることによって, 対象物の位置/姿勢を探索する問題は, 適合度関数 $F(\phi_M)$ の最大値を探索する問題に置き換えることができる. 本手法では, この最大値を探索する最適化問題に対して GA を適応することで探索時間の短縮を行う. また, 個体の位置/姿勢情報を表す遺伝子情報を以下に示す.

$$\begin{matrix} t_x & t_y & t_z & \epsilon_1 & \epsilon_2 & \epsilon_3 \\ 01 \cdots 01 & 00 \cdots 01 & 11 \cdots 01 & 01 \cdots 01 & 01 \cdots 11 & 01 \cdots 10 \\ \hline 12\text{bit} & 12\text{bit} & 12\text{bit} & 12\text{bit} & 12\text{bit} & 12\text{bit} \end{matrix}$$

個体の位置/姿勢は写真ベース認識法における立体モデルの位置/姿勢を表しており, 上位 36bit は 12bits ずつで, この遺伝子で与えられる立体モデルの位置座標を表わし, 残り 36bits は 12bit ずつで, 立体モデルの姿勢を表わしている. このとき使用する bit 数は探索時間短縮のために減らすことがある. 次に, 各個体は自分の位置/姿勢情報を用いて式 (10) に示した適合度関数 $F(\phi_M)$ より適合度値を得る. この値の優劣により進化処理が行われ, 世代交代による次世代の集合が構成される. このとき, 次世代の集合は前世代において適合度が高かった位置/姿勢, つまり対象物を表す適合度関数の最大値近辺に向かって近づく. この処理 (世代交代) を繰り返すことにより, GA は対象物の位置/姿勢を表す最大値を発見する.

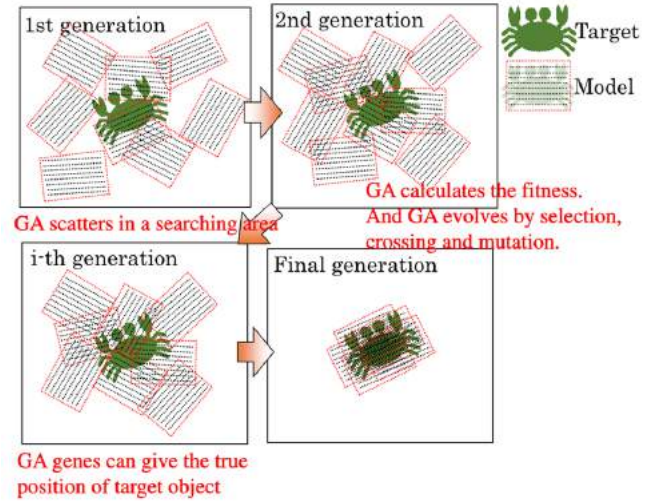


Fig.3 Fitness Distribution for xy

先ほど述べた GA では GA の収束を一定時間待つ必要があり, 適合度関数が十分に高い値を示すときに対象物の認識を完了すると, その時点では周囲の状況が大きく変化している可能性がある. そこで提案されたのが Real-Time Multi-Step GA であり, これは新しく入力された画像に対し, ビデオレート ([33ms]) で次の画像が入力されるまでの間だけ GA の進化を進め, その時点での最も高い適合度を与える個体をその時点における位置姿勢として出力する実時間遺伝的認識法である. RMGA のフローチャートを Fig.4 示す.

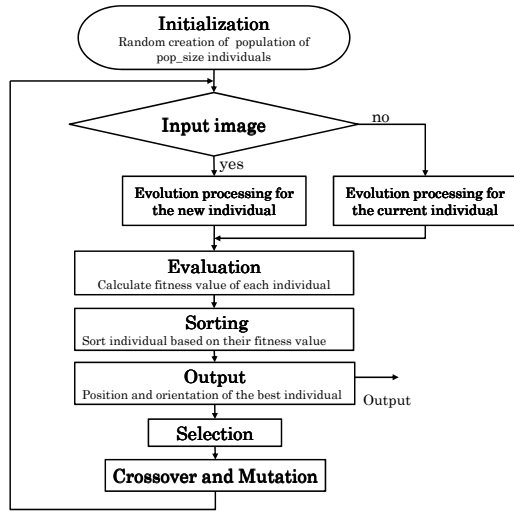


Fig.4 Real-time Multi-Step GA flowchart

3 実験

3.1 実験環境

実験で用いるロボットと対象物の座標系をそれぞれ Fig.5 に示す. 対象物座標系 Σ_M をハンド座標系 Σ_H から見た ($x=0\text{mm}$, $y=0\text{mm}$, $z=500\text{mm}$) の位置に設定する.

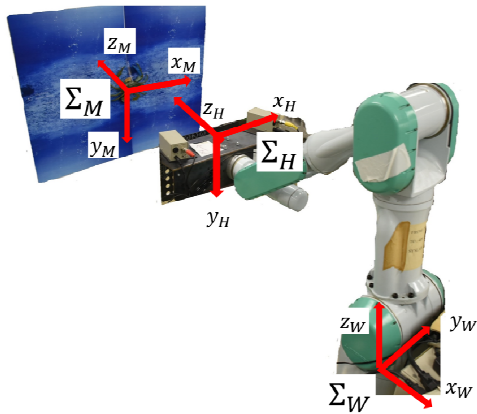


Fig.5 Coordinate System

3.2 対象物のモデル生成

まず, Fig.6(a) に示すような単一色の背景をカメラから取得する. カメラ画像にある点 (i, j) の色相値 $(H(i, j))$ を計算し, 式 (4) を用いて平均値 $H_{average}$ を求め, その値に基づいた背景の色相値の範囲 $H_{background}$ を式 (5) のように設定する.

$$\sum_{i=0}^{640} \sum_{j=0}^{480} H(i, j) / (640 \times 480) = H_{average} \quad (4)$$

$$H_{average} - 20 \leq H_{background} \leq H_{average} + 20 \quad (5)$$

次にモデルとなる対象物を Fig.6(b) のように設置し, 背景を含めた画像から色相値をとり, 先ほど設定した背景の色相値範囲と比較する. この時, 範囲内に収まっている点は背景, 範囲内に収まっていない点をモデルとし, モデルのフレームを決定する.

次にモデル枠内の各点の色相値をサンプリングし, Fig.6(c) に示されるような黒い点群で対象物の部分を示す面モデル (S_{in}) を

生成する. 最後に, Fig.6(c) に示されるような赤い点群で背景と重なる部分を示す帯モデル (S_{out}) を生成する.

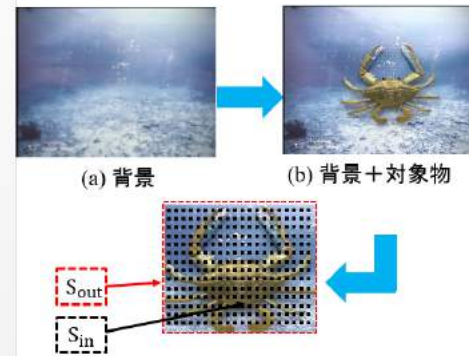


Fig.6 Model Generation Process

3.3 実験内容

本報では, 写真ベース認識法を用いたビジュアルサーボにおいて GA による対象物の認識が可能であるかどうかの検証を行うため, 適合度分布の調査を行った. 今回は Σ_H における $x, y, \epsilon_1, \epsilon_2, \epsilon_3$ について実験を行った. 本システムは適合度が 0.3 を上回り, かつ適合度分布において単峰性のピークが現れていれば認識可能であると判断することができる.

また実時間で対象物の認識が可能であるか検証を行うため, 対象物 x 方向に振幅 100[mm] 周期 20[s] を入力し, 入力した値に対して認識した値が同様な変化を示すかどうか認識実験を行った. またビジュアルサーボが可能であるかどうかの検証を行うため, 対象物の追従実験を行った. 対象物の x, y 方向ともに振幅 100[mm] 周期 30[s] の軌道を入力し, 認識した値をマニピュレータに出力し対象正面を維持しながら追従を行った. 手先が対象物を追従しているかどうかを検証した

3.4 適合度分布

Fig.7-9 に適合度分布をそれぞれ示す. 各適合度分布で最大適合度が 0.3 を上回り, 周辺に適合度の高い局所解がないことから単峰性のピークが現れていることがわかる.

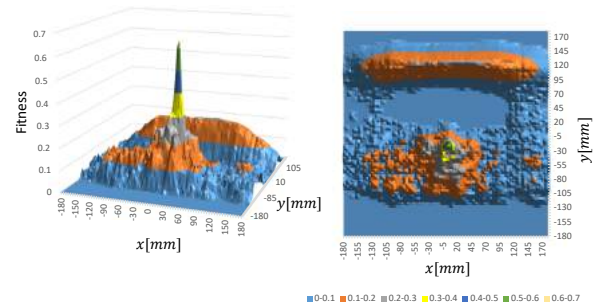


Fig.7 Fitness Distribution for xy

3.5 認識実験結果

Fig.10 に認識実験の結果を示す. 対象物に入力された軌道に対して認識した値が同様に変化をしており, 入力していない値は振動がみられるものの大きな変化がみられないため RMGA による実時間で対象の認識できていることがわかる.

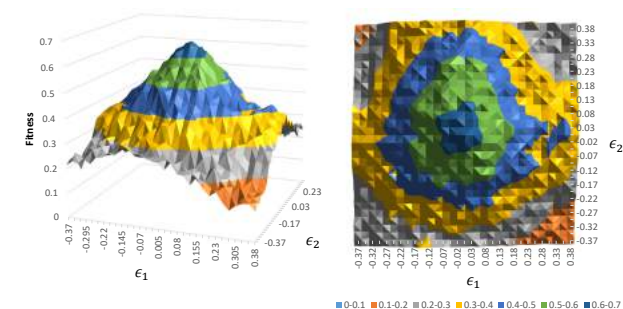


Fig.8 Fitness Distribution for ϵ_1, ϵ_2

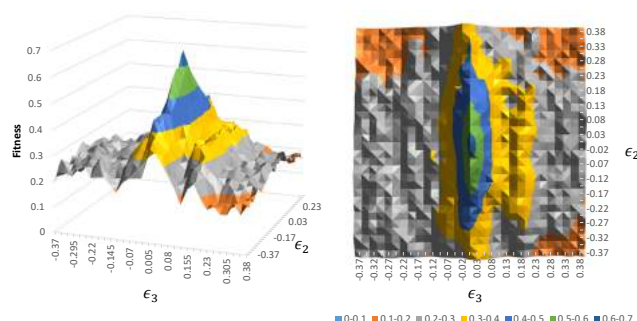


Fig.9 Fitness Distribution for ϵ_2, ϵ_3

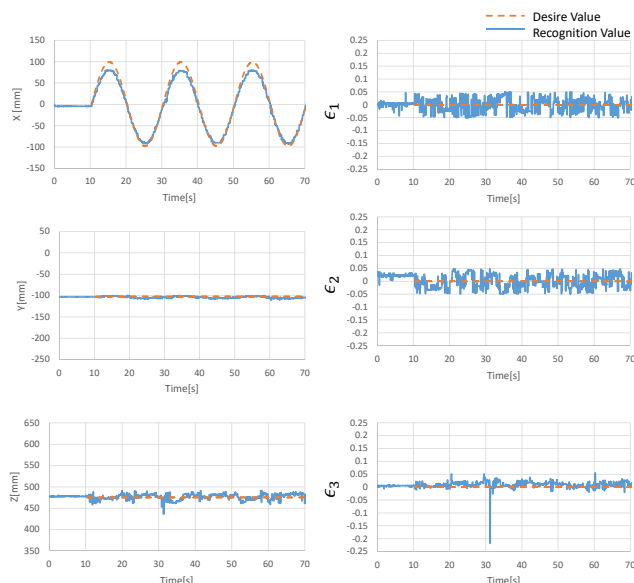


Fig.10 Ability of Tracking Target by RMGA

3.6 追従実験結果

Fig.11に追従実験の認識した値と実際の手先の軌道を示す。認識した値は振動がみられるものの入力した軌道に対して認識した値が同様に変化している様子が見られる。また認識した値に対して実際の手先も同様に変化している様子が見られることから対象物を追従していることがわかる。

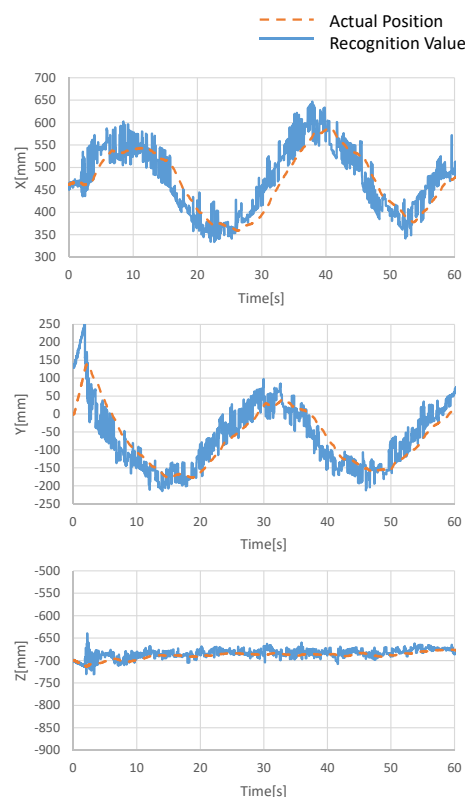


Fig.11 Ability of Tracking Target in Visual Servoing

4 結言

本報では、我々の提案する写真ベース認識法を用いたビジュアルサーボについて報告した。今回行った実験結果より、実際にビジュアルサーボを行うことができた。今後は更に認識精度向上を目指し研究を進めていく予定である。

参考文献

- [1] “平成 27 年版高齢社会白書（概要版）-内閣府”，http://www8.cao.go.jp/kourei/whitepaper/w-2015/html/gaiyou/s1_1.html
- [2] 栗井真志, 清水隆史, 山下淳, 金子透: “単眼カメラとレーザレンジファインダ搭載移動ロボットによる人物追跡と自律帰還”, 日本機械学会ロボティクス・メカトロニクス講演会 (2011), 2A1-H10.
- [3] 佐藤芳紀, 神酒勤, 岡本茂樹, 藤井祥平: “移動ロボットに実装したレーザレンジファインダと単眼カメラによる物体検出・追跡”, 電子情報通信学会, 信学技法, PRMU2011-137 (2011), pp.71-76.
- [4] 中原智治, 顧海松, 荒木秀和, 藤井裕之, 廣田雅之: “3次元認識によるピンピッキングシステムの実用化”, システム制御情報学会論文誌, Vol.14, No.4 (2001), pp.226-232.
- [5] 小野桂太郎, 小川卓哉, 前田雄介, 中谷茂樹, 永安剛, 清水領, 大内規高: “ステレオビジョンを用いた巻ばねの認識とピンピッキング”, 日本機械学会論文集 (C 編), Vol.79, No.804 (2013), pp.2769-2779.
- [6] De Luca, A., Oriolo, G. and Giordano, P. R.: “On-line Estimation of Feature Depth for Image-Based Visual Servoing Schemes”, Proceedings of 2007 IEEE Int. Conf. on Robotics and Automation (ICRA2007) (2007), pp.2823-2828.
- [7] Ryuki Funakubo, Khaing Win Phyu, Hongzhi Tian, Mamoru Minami: “Recognition and Handling of Clothes with Different Pattern by Dual Hand-eyes Robotic System”, IEEE/SICE International Symposium(2016), pp. 742-747.