第13回コンピューテーショナル·インテリジェンス研究会
2018年6月16日-17日, 八王子

Projection-basedビジュアルサーボシステム

○寇　ギョウ郡　　田　宏志　　李　想　　見浪　護（岡山大学）

Projection-based Visual Sevroing System

*Y. Kou　H. Tian　X. Li　M. Minami (Okayama University)

Abstract– Nowadays, in the field of robot vision, the control method called visual servoing attracts attention. Visual servoing is one of the methods for controlling robots. By incorporating visual information obtained from the installed visual sensor into the feedback loop, it is expected to enable robots to work in the constantly changing environment or unknown environment. Now, in this field, there were some methods that have been proposed to realize visual servoing. However, in those methods, the pre-knowledge such as the size, color, shape of the target is needed to some extent, which may lead a low flexibility to the system. To this issue, we proposed the project-based method. In this method, we use a dual-eyes configuration to perform the visual servoing towards the arbitrary target object. First, the model was created by image shown in left camera in real time, by using homography in projection geometry, then the image in the left camera is projected into 3D space to make a search model, then project the 3D search model into the right camera again to estimate the position and rotation of target object. In this paper, the recognition accuracy of position and orientation of unknown target object was verified.

Key Words: Visual Servoing, Projection-based, RT-MS GA

1　緒言

近年，社会の発展に伴う，ロボットへの要求は，従来のように操縦者がロボットを介して作業を行うものから，変化する環境の中でロボット自身が自律行動をとるものへと変化している．ロボットは工場現場，深海探索などの常に変化する環境または未知の環境で作業することが期待されている．

その要求を対応するため，常に変化する環境中ロボットが正確的に距離を感知することが重要である．人間では，環境中の物体を二つ目で距離を測れる，それは視差により，大脳に三角測量と生活の経験で距離が分かる，精確的な結果と言えないですが，大体の距離感を掴むことができる．しかし，現段階では，ロボットが人間の知性というものを持つことができないから，障害物を避けるため，ロボットが一般的に距離センサを搭載し，周辺の環境を感知する．距離センサが主に光学式，超音波式，レーザー光線式に分かれている．距離センサでロボットが環境と距離を感知できるが，距離センサの検出距離が短い，影響を受けやすい，検出空間が狭いなどの欠点が持っている．さらに，複雑な環境をマッピングするなら時間をかかるの特徴がある．常に変化するの環境へ応用することが難しい．

その問題に対して，我々はビジュアルサーボが一つの解決策と考える．

ビジュアルサーボとは，ロボットを制御する方式のひとつである．搭載される視覚センサから得られる視覚情報をフィードバックループに組み込むことによりロボットを制御する．

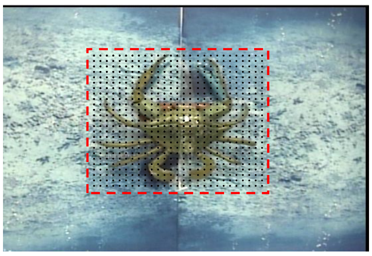
現在，ビジュアルサーボを実現するため，いくつかの方法が提案されているが，これらの方法は入力目標値により，位置ベース法（position-based method）, 特徴ベース法（image-based method）と 2-1/2-D 方 (2-1/2-2 method) 3 つに大別される [1),2)]．しかし，以上の方法は対象物の大きさ，色，或いは形状などの事前知識がある程度必要であり，これによりシステムに柔軟性が低下する可能性がある．この問題を解決するため，我々の研究グループは射影ベース認識法という方法を提案した．この方法では，任意対象物に対するビジュアルサーボを目的として，Dual-eyes カメラによる認識を行う．上記各認識手法の特徴を Fig.1 のテーブルに示す．

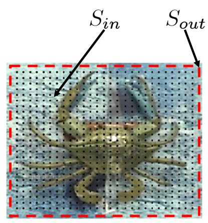
射影ベース法は，実時間にモデルを作り，変更することができるという特徴により，システムの柔軟性を向上しながら，任意物体を対象物として認識することが可能である．さらに，Dual-eyes カメラを搭載し，両カメラの視差により奥行き方向の距離を検出しやすい，可視領域が広いなどのメリットがある．本報では，射影ベースにより未知対象物の位置と姿勢の追従精度を検証する．

2　Projection-based 認識法

2.1　探索用モデル生成

まず，Fig.2(a) に示すような画像を左カメラから取得する．


(a)Background and target object



(b)Extraction of target object

Fig. 2: Model Generation Process

画像中心点から辺の長さが 240($Pixel$) の正方形をモデルとして切り出す．正方形内部が 13($Pixel$) ごとにサンプリング点を生成し，各サンプリング点の色情報が記録される．そして，画像座標系 $\Sigma_I L$ が Fig.2(a) に示すように画像の中心に設定されている．これにより，各サンプリング点が画像座標系 Σ_{IL} での位置情報が得られる．そこで，画像座標系での位置情報と色情報を持っ

	Position-based method	Image-based method	2-1/2-D Servoing	Projection-based method (本法)	Photo-model-based[3] (本研究室から提案)	Model-based [4] (本研究室から提案)
認識可能の対象物	三次元対象物	二次元対象物	三次元対象物	任意対象物	任意対象物	三次元対象物
モデルの定義・先験的知識	三次元モデルとして必要	特徴点として必要	画像と三次元情報[6]	実時間にモデルを作成，事前のモデル定義は不要	①対象物の写真 ②写真からカメラまでの距離	三次元モデルとして必要
対象物まで距離の認識	モデルにより認識可能	データベースにより認識可能[5]	認識可能	認識可能	認識可能[7]	認識可能[8]
正確のシステム校正(カメラのパラメータなど)	必要	不要	必要	必要	必要	必要
ロボットの形を確定	出来る	出来ない	出来る	出来る	出来る	出来る
利点	①既知モデルによりビジュアルサーボが簡単 ②デカルト座標系で直接に制御するため，軌道計画は簡単	システムの校正が不要のため，ロバスト性は高い	①三次元モデルは必要ない ②Position-basedとImage-based法の欠点を克服できる．	①先験的知識が不要のため，柔軟性は高い ②デカルト座標系で直接に制御するため，軌道計画は簡単 ③任意対象物までの距離を認識できる	①写真によって，速くモデルを作れ，速く対象物を変更できる ②対象物の接平面をモデルとして作るので，対象物の寸法，形と色などの定義が不要	①対応点の探す不要 ②既知モデルの三次元点群により画像認識精度高い
欠点	①モデルの定義は面倒，いろいろなターゲットに使用できない[4] ②カメラパなどのハードウェアからの影響を受けやすい	作業空間中のロボット制御は困難	事前情報が必要のため，柔軟性は不十分	現段階，対象物と背景の色は似てる場合が認識し難い	写真が撮る時，対象物とカメラの距離が必要	モデルの作りは時間がかかる

Fig. 1: The features of visual servoing method

ているモデルが生成される．モデルの外側から二層が帯部分 ($S_{L,out}$) であり，正方形の部分は面部分 ($S_{L,in}$) と定義されている．

2.2　射影ベース認識法の概要

Fig.3 に示すように基準点を手先座標系 Σ_H とする．また基準座標系 Σ_W とする．左のカメラ座標系を Σ_{CL}，画像座標系を Σ_{IL}，右のカメラ座標系を Σ_{CR}，画像座標系を Σ_{IR} とする．前節で説明したモデル生成の方法で左カメラ画像に映った対象物をモデルとして抽出し，Fig.3 に示すようにモデルを左カメラ画像から GA(遺伝的アルゴリズム) の遺伝子からの位置/姿勢 $\boldsymbol{\phi} = (x, y, z, \epsilon_1, \epsilon_2, \epsilon_3)$（$\epsilon$ はクォータニオンの姿勢変数）情報を用いて，3 次元空間へ逆射影する．逆射影されたモデルを右カメラ画像に再度射影することで，2 次元の平面モデルを得る．この平面モデルと右カメラ画像との照合により適合度関数を計算し，対象物の位置と姿勢を得ることができる．従って，平面モデルの位置/姿勢 $\boldsymbol{\phi}$（6 変数）が対象物の位置/姿勢（6 変数）と一致したとき，適合度関数の値が最大値を示す．

そのため，対象物の認識問題は，適合度関数の最大値探索問題に変換できる．適合度関数の最大値を探索・発見する場合，様々な方法があるが，一番単純かつ簡単な方法は全探索法である．これは，探索範囲内すべての可能性を探すことで最大値を発見するものであり，必ず最大値を発見できるが効率が悪い．つまり計算時間を多く費やすといった欠点を持っている．目標としているビジュアルサーボは動画像認識であり，動画像認識は計算処理が短いことが重要であるため，本研究では最適解探索に GA を適用することで，効率良く短時間で最大値探索処理を行う．従って，GA がある程度収束した時，GA の遺伝子により決定される立体モデルの位置/姿勢 $\boldsymbol{\phi}$ を対象物の位置/姿勢と考えることで，対象物の位置/姿勢が計測できる．

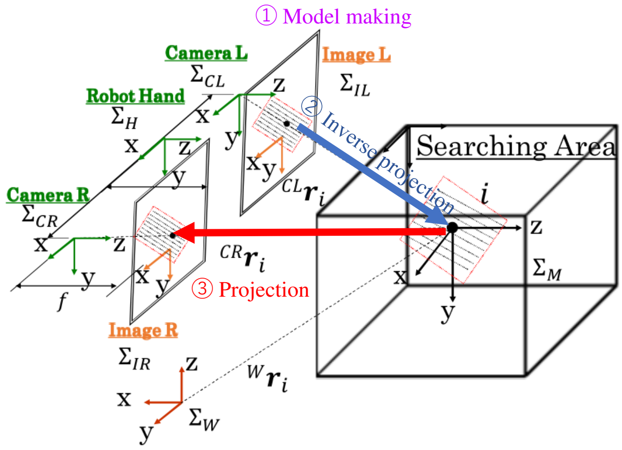


Fig. 3: Overview of Projection-based Method

3　認識結果評価手法

提案するシステムでは，両眼カメラから取得した画像を PC に転送し，Real-Time Multi-Step GA と Model-based Matching 法 [7] を組み合わせて対象物との相対的位置・姿勢を認識している．一般的に，GA（遺伝的アルゴリズム）を用いた最適解探索では，GA が最適解を決定する際ある程度の時間を要するため，最適解が得られた時には周囲の環境が大きく変化している可能性が高い．そこで本研究では，Real-Time Multi-Step GA を採用した．Real-Time Multi-Step GA とは，入力された画像に対し，ビデオレート（33[ms]）で次の画像が入力されるまでの間だけ GA 遺伝子の進化を進め，その時点で最大適合度を与える遺伝子情報を最新の対象物の位置・姿勢として出力する方法で，これにより実時間での認識を可能としている．また，このとき評価に使用する適合度とは，Model-based Matching 法で定義される対象物とモデルとの一致度合いを示す適合度関数から得られる値である．これにより，提案

システムでは対象物の３次元位置・姿勢の実時間認識を可能としている．また，提案システムによる対象物の位置・姿勢の認識情報は，Fig.3 に示す GA 探索空間内で獲得される．

3.1　適合度関数

左右カメラ映る画像と逆射影した結果を評価するため，本研究は適合度関数を利用して，認識結果を評価する．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.4 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，N は面部分と帯部分の評価点の総点数,$H_{IR}({}^{IR}r_i)$ は右カメラ画像の色相値，$H_{MR}({}^{IR}r_i)$ はモデル各サンプリング点の色相値と定義し，右画像領域 ${}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)$ における色相値に基づく評価値を $p_R({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$ とする．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 20 以上離れていた場合は，評価値を“ −1 ”と設定する．モデルの色相値と右カメラ画像の色相値の差が 20 以下の場合は，評価値を“ +2 ”と設定する．また，モデルの色相値と右カメラ画像の色相値の差が 20 以下の場合は，評価値を“ +0.1 ”と設定し，他の点の評価値を“ −0.5 ”とする．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．

$$F(\boldsymbol{\phi}_M^j) = \left\{ \sum_{\substack{{}^{IR}\boldsymbol{r}_i^j \in \\ S_{R,in}(\boldsymbol{\phi}_M^j)}} p_{R,in}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) + \sum_{\substack{{}^{IR}\boldsymbol{r}_i^j \in \\ S_{R,out}(\boldsymbol{\phi}_M^j)}} p_{R,out}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))) \right\} / N \tag{1}$$

$$p_{R,in}({}^{IL}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) = \begin{cases} 2, & \text{if}(|H_{IR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - H_{MR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))| \le 20); \\ -1, & \text{if}(|H_{IR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - H_{MR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))| > 20); \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

$$p_{R,out}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) = \begin{cases} 0.1, & \text{if}(|H_{IR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - H_{MR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))| \le 20); \\ -0.5, & \text{otherwise.} \end{cases} \tag{3}$$

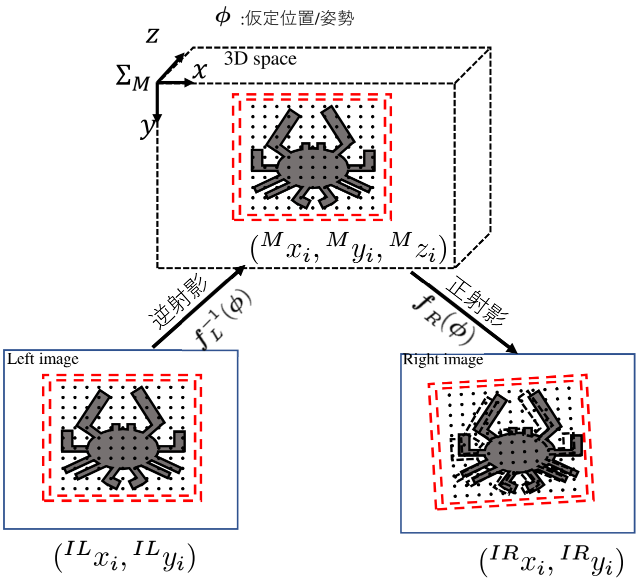


Fig. 4: Searching Model

この関数は，右画像領域 ${}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)$ の中で，$S_{R,in}(\boldsymbol{\phi}_M^j)$ に重なる入力画像の評価値 $p({}^{IR}\boldsymbol{r}_i)$ の合計値から，$S_{R,out}(\boldsymbol{\phi}_M^j)$ のそれを差し引くことで，右画像の適合度 $F_R(\boldsymbol{\phi}_M^j)$ を得る．これらを加算した適合度関数 $F(\boldsymbol{\phi}_M^j)$ を用いて，右の画像と再射影された平面モデルとを評価する．この適合度関数 $F(\boldsymbol{\phi}_M^j)$ は，空間中に逆射影されたモデルの持つ位置/姿勢情報より一意に決定する関数であり，逆射影されたモデルが対象物と一致すれば，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，$F_R(\boldsymbol{\phi}_M^j) \le 0$ の場合，$F_R(\boldsymbol{\phi}_M^j) = 0$ とする．

3.2　RT-MS GA

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 $F_R(\boldsymbol{\phi}_M^j)$ 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．

$$\overbrace{\underbrace{01\cdots01}_{10bit}}^{{}^{CL}z_{MiC}}\overbrace{\underbrace{00\cdots01}_{7bit}}^{{}^{Hx}\theta_M}\overbrace{\underbrace{11\cdots01}_{7bit}}^{{}^{Hy}\theta_M}.$$

個体の位置/姿勢は Projection-based Matching 法における立体モデルの位置/姿勢を表しており，上位 10bit がこの遺伝子で与えられる空間中に逆射影したモデルのカメラからの距離を表わし，残り 14bits は 7bit ずつ空間中に逆射影した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．

探索時に，Fig.5 に示すようにまず，探索区間中に GA の初期個体を生成する．次に，式 (1) に示した適合度関数 $F(\boldsymbol{\phi}_M^j)$ より各個体の適合度値を得る．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．

しかし，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，新しく入力された画像に対し，ビデオレート (33[ms]) で次の画像が入力されるまでの間だけ GA の進化を進め，その時点で最も高い適合度を与える個体をその時刻における位置・姿勢として出力する実時間遺伝的認識法である．

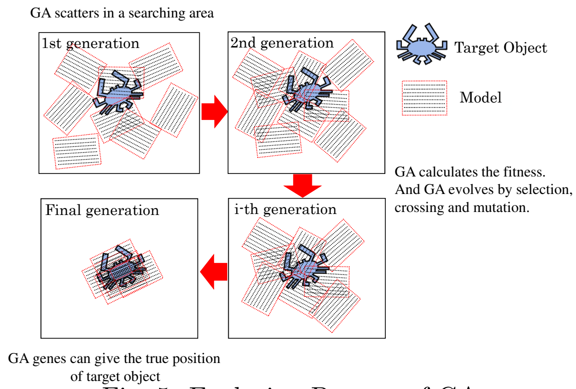


Fig. 5: Evolution Process of GA

本研究では，まず GA-PC1 で左画像から切り出した部分をモデルとして左画像中 RT-MS GA で $({}^{IL}x_{ij}, {}^{IL}y_{ij})$ を探索する．求められた $({}^{IL}x_{ij}, {}^{IL}y_{ij})$ をルーターで GA-PC2 に転送し，逆射影を行う．逆射影の計算により空間中探索用モデルの位置 $({}^Mx_{ij}, {}^My_{ij}, {}^Mz_{ij})$ を見つける．そして RT-MS GA で各探索用モデルを右カメラ比較し，適合度が高い結果を求める．ここで，現段階射影ベース法を用いたビジュアルサーボシステムが 33[ms] の中に，左カメラの進化世代数は 20 代，右カメラの進化世代数は 15 代，その特徴により実時間で対象物の位置と姿勢を検出することができる．

4　移動中対象物に向けた認識実験

4.1　システム構成

本システム構成は対象物把持ロボット (Target Object-Robot) とビジュアルサーボ認識ロボット (V.S.-Robot) 二つ部分で構成されている．ロボットは PA10 三菱重工業株式会社製であり，認識部のカメラセンサーが FCR-IX11A(Sony 製) である．各部分の座標関係は Fig.6 で表す．

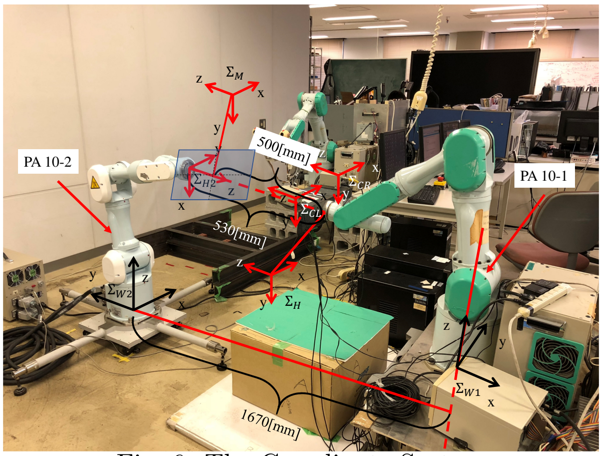


Fig. 6: The Coordinate System

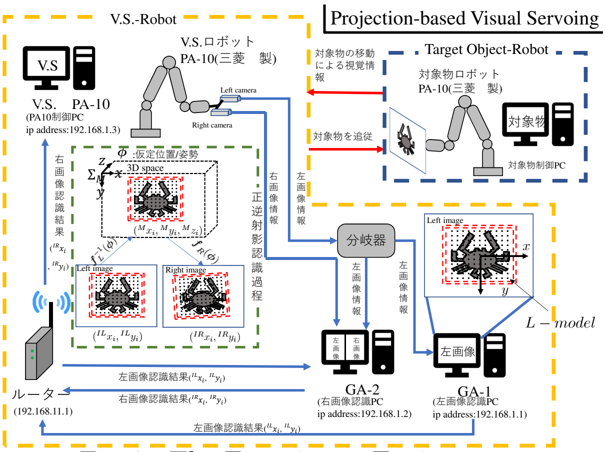


Fig. 7: The Experiment Environment

4.2　実験内容

実験の構成図は Fig.7 で表されている．本実験では，対象物とハンドの間の位置と姿勢を一定の値を保つことを目標として，20 秒の周期で奥行方向 (Σ_x 方向) のビジュアルサーボをする．ここで，対象物とハンドの位置と姿勢の関係は式 (4) のように表されている．また，対象物の移動は式 (5) で表されている．まず GA-PC1 で左画像から切り出した部分をモデルとして左画像中 RT-MS GA で $({}^{IL}x_{ij}, {}^{IL}y_{ij})$ を探索する．求められた $({}^{IL}x_{ij}, {}^{IL}y_{ij})$ をルーターで GA-PC2 に転送し，逆射影を行う．逆射影の計算により空間中探索用モデルの位置 $({}^{M}x_{ij}, {}^{M}y_{ij}, {}^{M}z_{ij})$ を見つける．そして RT-MS GA で各探索用モデルを右カメラ比較し，適合度が高い結果を求める．そして，対象物を ${}^{M}X$ に軸回るの回転角度を ϵ_1, ${}^{M}Y$ に軸回るの回転角度を ϵ_2 として，30 秒周期で $-5°\sim 5°$ の往復回転についてビジュアルサーボをする．

$$ {}^{H_d}\boldsymbol{\phi}_M = [0[mm], -80[mm], 500[mm], 0, 0, 0]. \tag{4}$$

$$ L = -100 + 100\cos(\omega t)[mm]. \tag{5}$$

そこで，式 (4) の関係を用いて，ハンドの目標位置と姿勢が認識された対象物の位置と姿勢から得られる．ハンドの目標位置と姿勢を用いて，ビジュアルサーボをすることができる．

4.3　実験結果

実験の結果は Fig.8 で示す．図中赤い線は対象物位置の真値，黒点線は追従結果である．結果から見ると，射影ベース法を用いたビジュアルサーボシステムは対象物の移動軌道を沿って，見失うことなしで連続に対象物を認識した．最初の 4 秒間で対象物を探索し，4 秒以後認識結果は対象物の軌道に収束している．奥行方向 (Σ_x) の認識結果は約最大 50[mm] の誤差が持っている．上下，左右及び姿勢の認識結果が真値に収束していることが分かる．

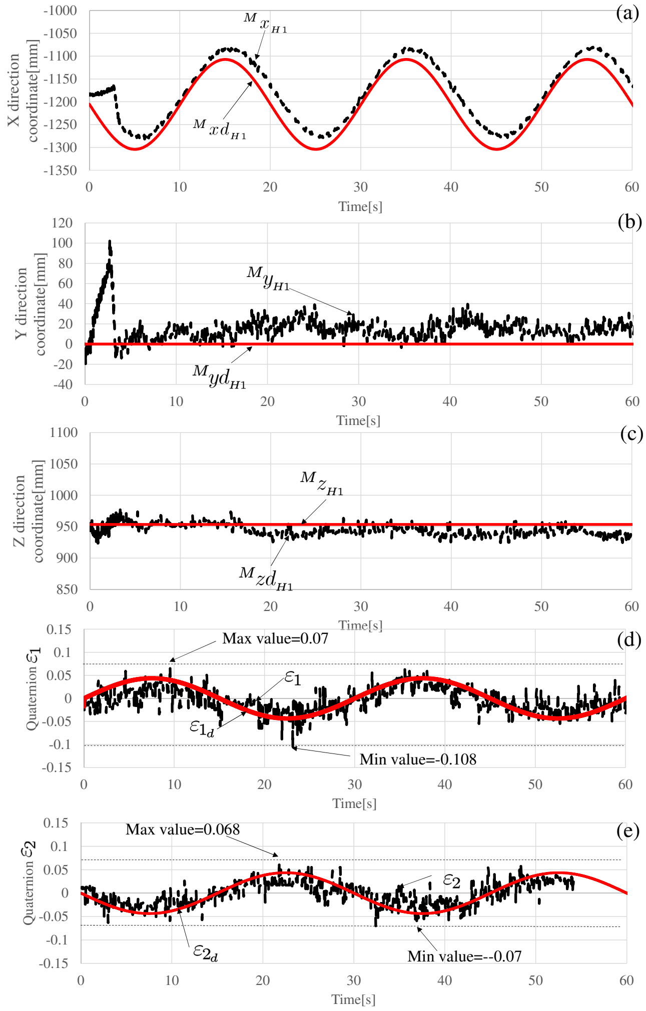


Fig. 8: The Experiment Result

5　結言

本文では，射影ベース認識法を用いたビジュアルサーボシステムで対象物の認識結果を報告した．今回行った実験結果より，我々が提案したビジュアルサーボシステムが位置及び姿勢における認識できることが分かる．今後は対象物の位置/姿勢認識精度を向上させることを目指して，本システムロバスト性を検証し研究を進めていく予定である．

Refences

1) Kragic, Danica and Henrik I. Christensen. “ Survey on Visual Servoing for Manipulation. ”

2) Hutchinson, Seth, Gregory D. Hager, and Peter I. Corke. "A tutorial on visual servo control." IEEE transactions on robotics and automation 12.5 (1996): 651-670.

3) Khaing Win Phyu, Ryuki Funakubo, Ryota Hagiwara, Hongzhi Tian, Mamoru Minami, "Verification of photo-model-based pose estimation and handling of unique clothes under illumination varieties", Journal of Advanced Mechanical Design, Systems, and Manufacturing, Vol.12, No.2.

4) 田 宏志, 寇 ギョウ郡, 見浪 護 "横軸方向周波数応答実験による両眼転導ビジュアルサーボの評価", 日本機械学会論文集 Vol.87, No.857, 2017

5) Song, Jinseon and Yongwan. “ Image Features Comparison-Based Position Estimation Method Using a Camera Sensor. ” (2016).

6) Chaumette, Fran?ois, and Ezio Malis. "2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoings." Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on. Vol. 1. IEEE, 2000.

7) 舟久保 龍希, Khaing Win Phyu, 田 宏志, 寇 ギョウ郡, 見浪 護, 矢納 陽 "複眼ハンドアイロボットを用いた不定形単品衣服ハンドリングとビジュアルサーボ", 第 27 回インテリジェント・システム・シンポジウム

8) Yejun Kou, Hongzhi Tian, Mamoru Minami, Takayuki Matsuno, "Improved eye?vergence visual servoing system in longitudinal direction with RM?GA", Artif Life Robotics, Vol.23, Issue 1