射影ベース認識法を用いたビジュアルサーボによる対象物追従

the Performance of Object Tracking with Projection-Based Visual Servoing Method

○ 王 俊祥 (岡山大)　正 田 宏志（岡山大）
寇 ギョウ郡 (岡山大)　舟久保 龍希（岡山大）
李 想 (岡山大)　山本 太郎（岡山大）
正 松野 隆幸 (岡山大)　正 見浪 護（岡山大）

Junxiang WANG, Okayama University, p6386zth_shinj@s.okayama-u.ac.jp
Hongzhi TIAN, Okayama University
Yejun KOU, Okayama University
Ryuki FUNAKUBO, Okayama University
Xiang LI, Okayama University
Taro YAMAMOTO, Okayama University
Takayuki MATSUNO, Okayama University
Mamoru MINAMI, Okayama University

This paper proposes a visual servoing system based on a projection-based recognition method. The projection-based visual servoing system using only two calibrated cameras can identify the position and orientation of arbitrary-shaped unknown objects. For a visual servoing system, it needs to know the spatial dimension data of the observed object. However, in the detection of an unknown object, the visual servoing system generally directly obtains 2D images of the observed object, and doesn't know the three-dimensional size. Especially the detection of distance is important. Through an inverse projective transformation, we have realized the spatial transformation of 2D to 3D. And a genetic algorithm which is called "Real-Time Multi-Step Genetic Algorithm(RM-GA)" is used to solve on-line optimization problems. In this paper, the tracking ability of projection-based visual servoing method has been verified through arbitrary-shaped object tracking experiments. The results of experiments will be reported.

Key Words: Visual servoing, Projection-based, Inverse projection transformation, RM-GA

1 緒言

現在，日本は少子高齢化の影響で労働力が不足している問題がある．人口減少時代の経済成長のためには、人工知能やロボットによる労働の代替化は不可欠だと考えられる．近年では益々自律動作や未知の環境下での動作が必要されるようになった．災害地区や宇宙区間などで作業を行う極限ロボットや，病院や家庭で活躍するロボットにおいては，未知の環境下で，自律的に作業を実現するため，外界の情報を検出し制御をする機構が必要となる．そこで，ビジュアルサーボイングという手法が提案された．

ビジュアルサーボは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でロボットの動作を制御する方法であるので，常に変化する環境や未知の環境下でのロボットの活躍が期待されている．しかし，従来の研究では対象物の寸法，形または色などの情報がある程度必要であり，これによりシステムの柔軟性が低下する可能性がる．ビジュアルサーボの認識能力を向上させるために，ハンドアイ構成をとる方法 [1]，ステレオカメラを使用する方法 [2] など複数の研究方法が提案されている．これらの方法はカメラの数を増やすことで環境を認識する視点の数を増加させて，或いは深度センサーによって距離探索能力を向上する．

ターゲットの認識に対して事前に識別された対象物の寸法，形または色など情報を分かったら，そんな情報を先験知識と呼ばれる．ある研究 [3, 4] は先験知識を用いて対象物を認識して追跡する．

一方で，先験的な知識を使わずに未知対象物を認識する両眼立体視も応用が非常に幅広い研究方向である．その中に一般的な方法は左右画像の対応点を探し，左右画像の対応から 3 次元位置を計算する [5]．対応点を探すのはこの方法の難点で，間違いやすい．従って，我々は Model-based 方法を用いて 2 つカメラから構成する固定ハンドアイシステムに基づいて先験知識がいらない射影ベース認識法 (Projection-based recognition method) を提案した．

本文では，対象物のモデルの作り方と対象物の位置と姿勢の探索法を説明し，対象物追従実験を行い，実験結果により，射影ベース認識法を用いたビジュアルサーボシステムの性能を評価する．

2 対象物の位置・姿勢計測法

2.1 探索用モデル生成

まず，Fig.1(a) に示すような画像を左側のカメラから取得する．カメラ画像の中心から 240($Pixel$) 長さの正方形を抽出し，Fig.1(b) のように間隔が 13($Pixel$) で抽出した正方形画像の左上から $18 \times 18 = 324$(個) の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，左画像座標系 Σ_{IL} が Fig.1(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，他のは面部分 ($S_{L,in}$) と定義されている．

2.2 射影ベース認識法の概要

Fig.2 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

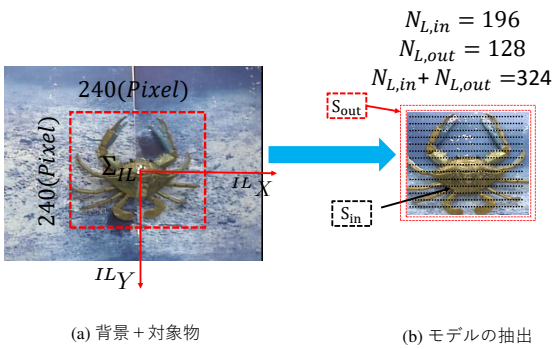

Fig.1 Model generation process

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.2 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 (第二章の第四節で説明される．) からの位置/姿勢情報 $\boldsymbol{\phi} = ({}^{CL}z_{MiC}, {}^{Hx}\theta_M, {}^{Hy}\theta_M)$ 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．そして，3 次元空間中に逆射影されたモデルの位置/姿勢が対象物の位置/姿勢と一致したとき，右カメラ画像に再度射影した二次元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，全ての可能性を計算する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．そして，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．

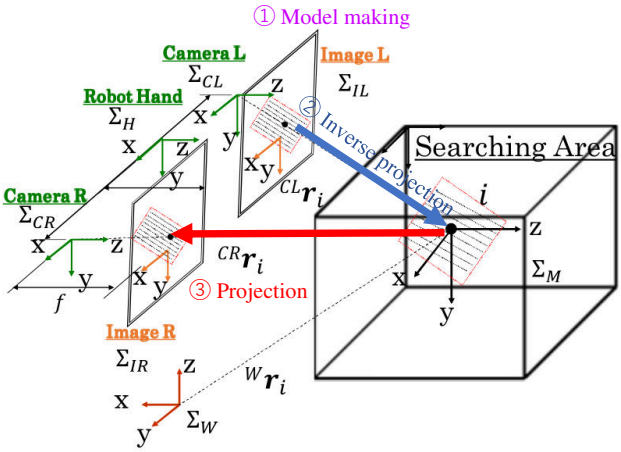

Fig.2 Overview of projection-based method

2.3 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.3 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 ${}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)$ 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の評価値を“ $+2$ ”，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 20 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．

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

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

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

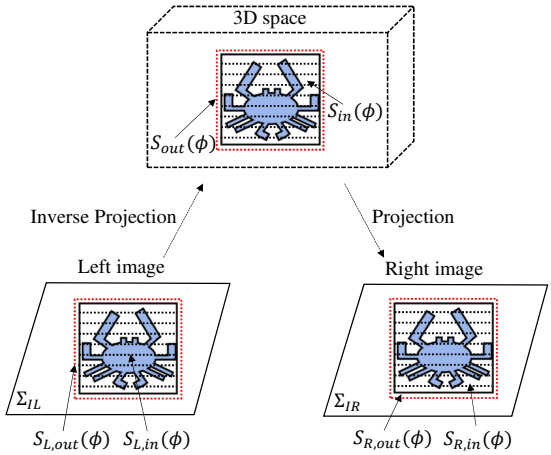

Fig.3 Searching model

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 $p_{R,in}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$ の合計値と $S_{R,out}(\boldsymbol{\phi}_M^j)$ に重なる入力画像の評価値 $p_{R,out}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$ を加えることで，右画像の適合度 $F(\boldsymbol{\phi}_M^j)$ 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，$F(\boldsymbol{\phi}_M^j) \leq 0$ の場合，$F(\boldsymbol{\phi}_M^j) = 0$ とする．

2.4 GA を用いた最適解探索法

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 $F(\boldsymbol{\phi}_M^j)$ 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．

$$\underbrace{01\cdots01}_{10bit}\underbrace{00\cdots01}_{7bit}\underbrace{11\cdots01}_{7bit}.$$

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.2 に示すようにまず，逆射影によって三次元の探索区間中に GA の初期個体を生成する．次に，右画像へ再投影し，式 (1) に示した適合度関数 $F(\phi_M^j)$ より各個体の適合度値を得る．後は Fig.4 に示すように適合度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．

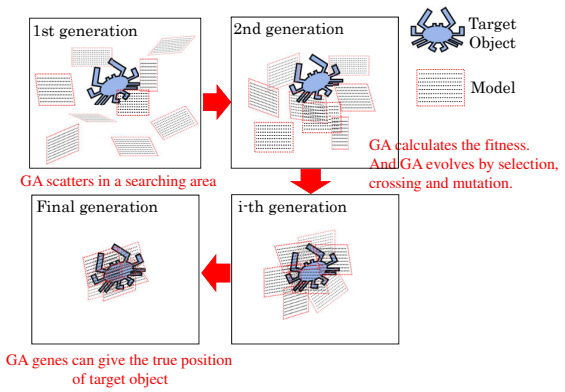

Fig.4 Evolution process of GA

3 姿勢追従実験

3.1 実験環境

Fig.5 に示すような実験環境で提案したビジュアルサーボシステムの有効性を実機によって確認する． システムの構成として

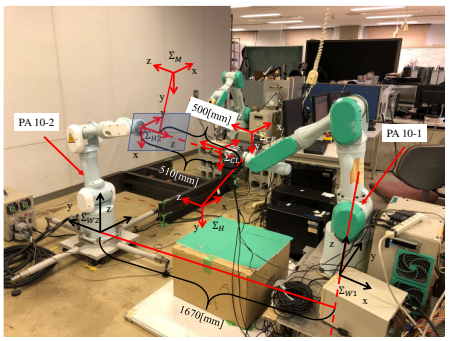

Fig.5 the Coordinates of target object and robots

は，Fig.6 に示すように，Target Object-Robot と V.S.Robot に分けられ，V.S.-Robot の部分はまた，制御部と認識部に分けられている．Target Object-Robot と制御部のロボットは PA10-

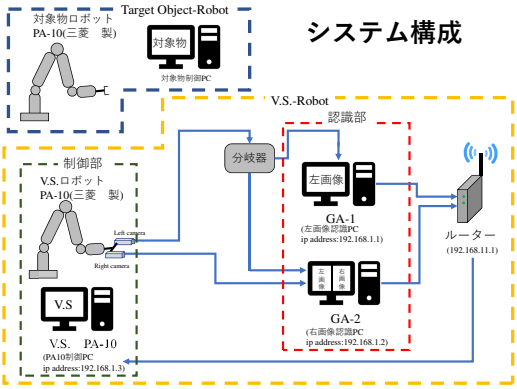

Fig.6 Configuration of System

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．各部分の座標系は Fig.5 に示すように設定されている．

3.2 実験内容

本実験では，まず，カニを対象物としてビジュアルサーボをする．途中に海星を持って一瞬にカニを隠して対象物としてビジュアルサーボをする．そして，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，ビジュアルサーボをする．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．

$$^{H}\phi_M = [0, -80[mm], 500[mm], 0, 0, 0]. \tag{4}$$

対象物の軌道については，一つ目の実験では，対象物がカニの場合では，対象物を ^{M}x 軸回り (ε_1) に周期が 30[s] で振幅が 10[deg] で正弦関数の周期運動をさせ，海星に変更した後に実験者が海星を持って，^{M}x 軸回り (ε_1) に回転する．

二つ目の実験では，対象物がカニの場合では，対象物を ^{M}y 軸回り (ε_2) に周期が 30[s] で振幅が 10[deg] で正弦関数の周期運動をさせ，海星に変更した後に実験者が海星を持って，^{M}y 軸回り (ε_2) に回転する．

二つの実験とも，海星を対象物として追従する場合では，実験者が海星を持って，運動させるため，目標値の真値が取られない．

3.3 実験結果

世界座標系 Σ_W，手先座標系 Σ_H，対象物座標系 Σ_M が Fig.5 に示されている．実験の結果が Fig.7 と Fig.8 のように表されている．グラフの横軸が時間を表し，縦軸がそれぞれに位置と姿勢を表す．ブラックの破線 $(^{W}x_{Hd}, ^{W}y_{Hd}, ^{W}z_{Hd}, \varepsilon_{1d}, \varepsilon_{2d}, \varepsilon_{3d})$ が手先位置と姿勢の目標値であり，赤い実線 $(^{W}x_{Hr}, ^{W}y_{Hr}, ^{W}z_{Hr}, \varepsilon_{1r}, \varepsilon_{2r}, \varepsilon_{3r})$ が認識値であり，青い長鎖線 $(^{W}x_{H}, ^{W}y_{H}, ^{W}z_{H}, \varepsilon_1, \varepsilon_2, \varepsilon_3)$ が手先の実際の位置と姿勢である．

二つの実験結果から，Fig.7 と Fig.8 の示すように，射影ベース認識法を用いるビジュアルサーボシステムはカニと海星を対象物として，位置と姿勢を認識して追従できたことがわかる．特に，追従途中に追従対象物を変更して，安定に追従することもできたことが分かる．

4 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追従実験結果を報告した．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と姿勢における追従ができた．特に，追従途中に突然に入る新しい対象物とも追従できた．これは我々の射影ベース認識法の特徴点である．今後は更に対象物の位置と姿勢追従精度を向上させることを目指して，任意の未知形状の物体の認識が安定にできるシステムを目標として，研究を進めていく予定である．

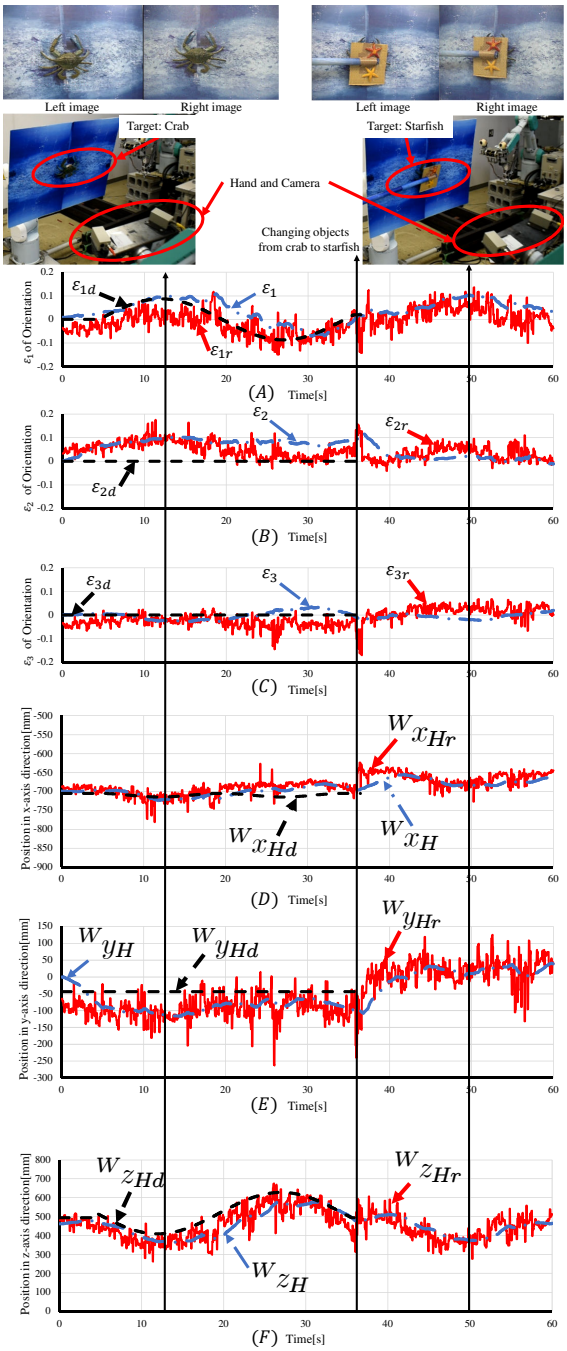

Fig.7 Position and orientation tracking results rotating with ε_1

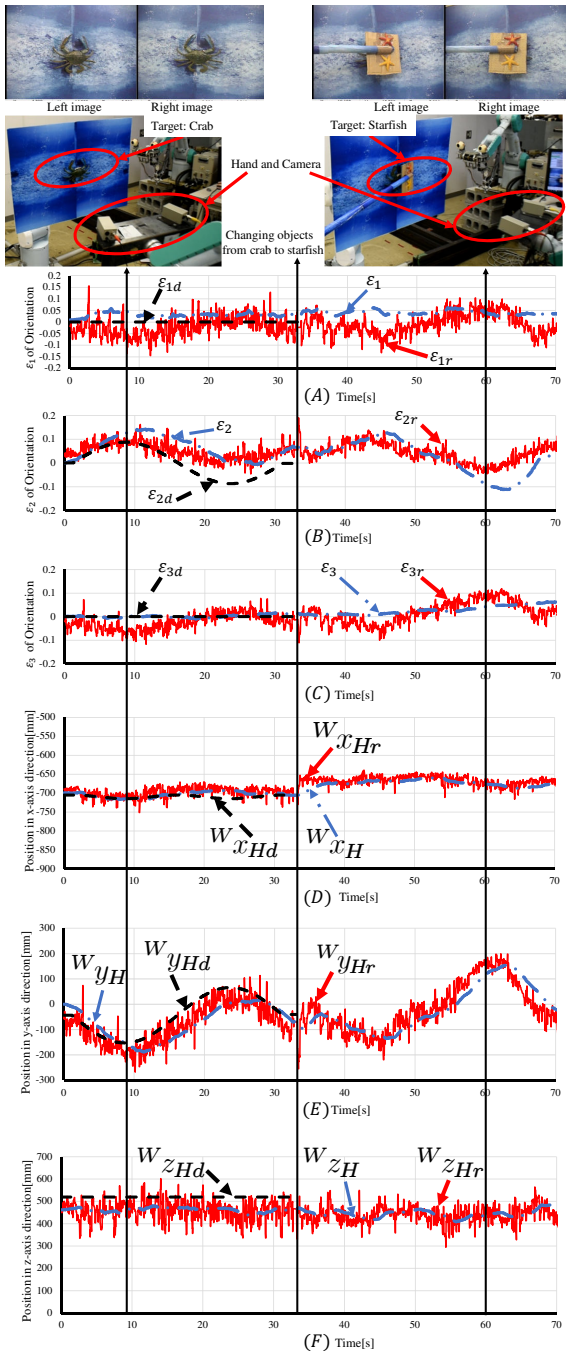

Fig.8 Position and orientation tracking results rotating with ε_2

参考文献

[1] 松浦精太郎, 丸典明 : “視空間誤差を用いた仮想バネダンパ仮説に基づくダイナミックビジュアルサーボによる Eye-In-Hand 型ロボットの位置姿勢制御”, 日本機械学会論文集 (C 編). 77-776, pp.186-195, 2011.

[2] W. Song, M. Minami, Y. Mae and S. Aoyagi: “On-line Evolutionary Head Pose Measurement by Feedforward Stereo Model Matching,” IEEE Int. Conf. on Robotics and Automation (ICRA), 4394/4400, 2007)

[3] Petsch, Susanne, and Darius Burschka. ”Estimation of spatio-temporal object properties for manipulation tasks from observation of humans.” Robotics and Automation (ICRA), 2010 IEEE International Conference on. IEEE, 2010.

[4] Cui, Yu, et al. ”Analyses about Trackability of Hand-eye-vergence Visual Servoing in Lateral Direction.” Applied Mechanics and Materials. Vol. 772. Trans Tech Publications, 2015.

[5] Woodfill, John Iselin, Gaile Gordon, and Ron Buck. ”Tyzx deepsea high speed stereo vision system.” Computer Vision and Pattern Recognition Workshop, 2004. CVPRW’04. Conference on. IEEE, 2004.