

ビジュアルサーボ中の実時間ポーズトラッキング性能の解析

Analysis of Real-Time Pose Tracking Performance during Visual Servoing

西村 健太 (岡山大) 候 森 (岡山大)
前田 耕市 (岡山大) 正 見浪 護 (岡山大)
正 矢納 陽 (岡山大)

Kenta NISHIMURA, Okayama University, nishimurak@suri.sys.okayama-u.ac.jp
Sen HOU, Okayama University
Koichi MAEDA, Okayama University
Mamoru MINAMI, Okayama University
Akira YANOU, Okayama University

In this research, Genetic algorithm (GA) is used as pose-tracking method of an target object for 3-D visual servoing. Real-time Object tracking time has been shortened for real-time pose estimation by using 1-step GA and Eye-vergence function, and pose-tracking accuracy has been verified through fitness function distribution which is a correlation function between the target object projected in camera frame and model defined in the control computer. In this research, it verifies about the error of object recognition by showing the action of GA in time-domain during visual servoing.

Key Words: Visual feedback control, GA, Visual Servoing

1. 緒 言

ビジュアルサーボは、視覚情報を元にフィードバック制御を行うことによりロボットの動作を制御している。対象物の位置/姿勢を探索する問題は、適合度の最大値探索問題に置きかえることができる。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 を適用することで探索時間の短縮を行っている。周囲の状況変化に対応するために、実時間遺伝認識法として 1-step GA を用いる^[1]。

しかし、実際の物体の位置/姿勢を座標系 Σ_M の位置/姿勢で表現し、GA により推定された物体を $\Sigma_{\hat{M}}$ と表すと、 Σ_M と $\Sigma_{\hat{M}}$ の間には同次変換行列で表した誤差 ${}^M T_{\hat{M}}$ が存在することが一般的である。この認識誤差 ${}^M T_{\hat{M}}$ を減少させる必要がある。そこで、ビジュアルサーボ中における適合度の全探索における対象物の位置/姿勢と、1-step GA において GA が発見した対象物の位置/姿勢を比較することで、ビジュアルサーボ制御中の実時間認識誤差を把握することができ、認識追従誤差の関係を把握することが出来る。

本報では、対象物を周期的に動かす周波数応答実験において、ビジュアルサーボを行った (図 1)。ビジュアルサーボ中における GA 遺伝子の挙動を示すことにより、GA と対象物との関係性を明らかにする。

2. ビジュアルサーボで用いる認識手法

本研究で用いられているビジュアルサーボシステムは、動画像認識を必要とする。動画像は静止画像を時系列で並べたものであり、動画像から対象物を認識することは静止画像からの対象物を認識をビデオレート (33[ms]) 以内で連続して行うことで実現される。よって、ここでは、1 枚の静止画像に対する認識手法の概要を説明する。

3 次元の立体モデルの位置/姿勢 $\phi = (x, y, z, \epsilon_1, \epsilon_2, \epsilon_3)$ (ϵ はクォータニオンの姿勢変数) が GA の遺伝子により決定され、左右の画像に射影することで、2 次元の平面モデルを得る。この平面モデルを画像との照合により適合度関数を計算し、評価を行う。そして、立体モデルは、立体モデルの位置/姿勢 ϕ (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ように定義する。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^[1]。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、

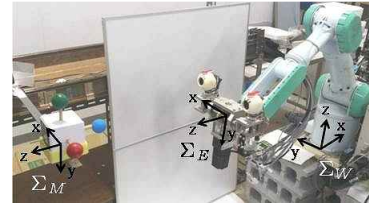


Fig. 1 Object and the visual-servoing system

一番単純かつ簡単な方法は全探索法である。これは、全ての関数値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く、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。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 ϕ 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。

2.1 Model-based Matching 法

左右ビデオカメラから入力される画像は、0 ~ 255 で表される色相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 ${}^{IR}r_i$ における色相値を $p({}^{IR}r_i)$ 、左画像領域に ${}^{IL}r_i$ における色相値を $p({}^{IL}r_i)$ とする。対象物を評価する内部領域の色相値評価分布を“+1”，周囲背景を評価する帯領域をのそれを“-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。

$$F_{site}(\phi) = \left\{ \left(\sum_{{}^{IR}r_i \in S_{R,in}(\phi)} p({}^{IR}r_i) - \sum_{{}^{IR}r_i \in S_{R,out}(\phi)} p({}^{IR}r_i) \right) + \left(\sum_{{}^{IL}r_i \in S_{L,in}(\phi)} p({}^{IL}r_i) - \sum_{{}^{IL}r_i \in S_{L,out}(\phi)} p({}^{IL}r_i) \right) \right\} / 2 \\ = \{F_{R,site}(\phi) + F_{L,site}(\phi)\} / 2 \quad (1)$$

この関数は、右画像領域 ${}^{IR}r_i$ の中で、 $S_{R,in}(\phi)$ に重なる入力画像の色相値 $p({}^{IR}r_i)$ の合計値から、 $S_{R,out}(\phi)$ のそれを差し引

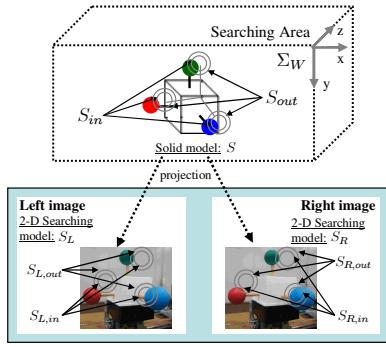


Fig. 2 Searching model

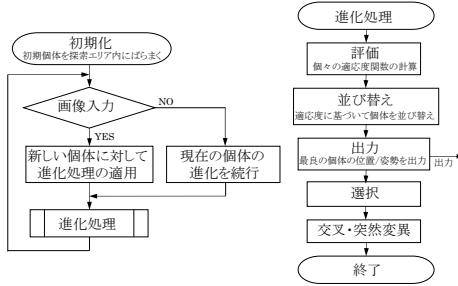
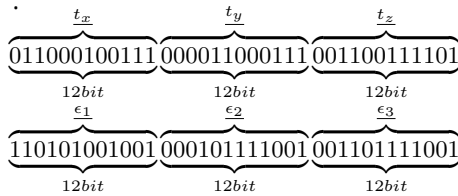


Fig. 3 1-step GA

くことで、右画像の適合度 $F_{R,site}(\phi)$ を得る。左画像においても同様に適合度 $F_{L,site}(\phi)$ を計算し、これらを加算し平均をとった適合度関数 $F_{site}(\phi)$ を用いて、左右の画像を同時に評価する。この適合度関数 $F_{site}(\phi)$ 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、 $F_{R,site}(\phi) \leq 0$ 、または、 $F_{L,site}(\phi) \leq 0$ の場合、 $F_{R,site}(\phi) = 0$ 、または、 $F_{L,site}(\phi) = 0$ とする。

2.2 GA を用いた最適解探索法

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 $F(\phi)$ 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、研究における個体の位置/姿勢情報を表す遺伝子情報を以下に示す。



個体の位置/姿勢は Model-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 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 (1) に示した適合度関数 $F(\phi)$ 、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。しかし、GA は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

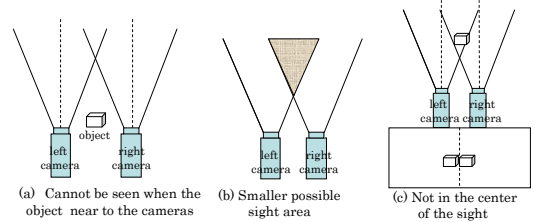


Fig. 4 Disadvantage of fixed camera system

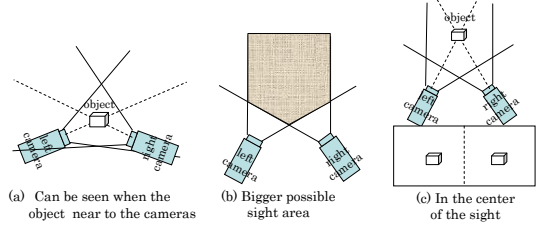


Fig. 5 Advantage of Eye-vergence system

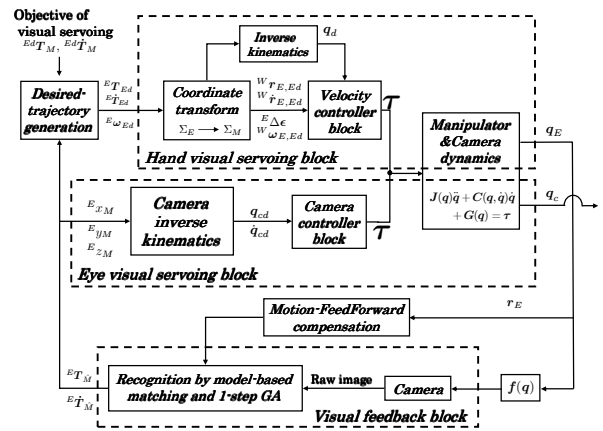


Fig. 6 Hand & Eye-Vergence Visual Servo System

が大きく変化している可能性がある。そこで、1-step GA を用いる (図 3)。これは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対象物の位置/姿勢として出力する実時間遺伝的認識法⁽¹⁾である。

3. ハンドアイビジュアルサーボシステム

3.1 Eye-vergence system

ハンドアイ構成は固定カメラ構成に比べて追従動作の際のハンドの振動によってサーボ動作が不安定になりやすい欠点を持つが、適応的に視点を選択できるといった長所がある。本報では二つのカメラを持つハンドアイ構成のビジュアルサーボについて考える。対象物形状が既知であることを仮定すると、単眼でも位置/姿勢の 6 変数を計測することは可能であるが、カメラ対象物間の距離の計測精度に問題があることは、良く知られており、ここでは複眼構成を採用する。

ハンドアイシステムであっても、カメラがハンドに固定されている固定ハンドアイシステムは、下記のいくつかの欠点を持つ。例えば、対象物がカメラに近いとき対象物を認識できないこと (図 4 (a))、また二つのカメラの可視可能領域が狭い (図 4 (b))、さらに可視可能領域内の対象物であっても図 4 (c) に示すようにカメラ画像視野の中心に写像されることはない。以上より固定ハンドアイシステムは位置/姿勢計測が不正確になる場合や、不可能になる場合があるという本質的な問題を持っていることがわかる。そこで上述の問題を解決するため、本研究では

Eye-vergence という、画像の中心で目標物を捉えるためにカメラ自身に姿勢を変化させる自由度を与えるシステムを用いる。カメラの姿勢を変えることが可能になるので、対象物をよりよく観測することができる。図 5 の (a)-(c) は図 4 の問題点 (a)-(c) に対応した Eye-Vergence システムにメリットを示す。(a), (b) は両眼撮像可能領域が拡大すること、(c) はレンズの中心で目標物を観測することを示している。

3.2 ハンドアイビジュアルサーボコントローラ

ハンド Eye-Vergence 二重ビジュアルサーボのブロック線図を図 6 に示す。ハンドビジュアルサーボに対して使用されるコントローラは B.Siciliano^[2] によって提案されたものである。マニピュレータの各関節角度を $q_E = [q_1, \dots, q_7]$, カメラのパンチルト角を $q_c = [q_8, q_9, q_{10}]$ とし、各リンクの目標角度を q_d とすると、ハンドビジュアルサーボコントローラは

$$\tau = K_{SP}(q_d - q) + K_{SD}(\dot{q}_d - \dot{q}) \quad (2)$$

と示される。 K_{SP} はバネ定数、 K_{SD} は粘性抵抗を表す行列である。また、カメラの目標角度を q_{cd} とおき、角度の誤差を

$$\Delta q_c = q_{cd} - q_c \quad (3)$$

と定義すると、Eye-Vergence ビジュアルサーボコントローラは以下のように計算できる。

$$\dot{q}_{cd} = K_{Pc} \Delta q_c \quad (4)$$

K_{Pc} は P ゲインとして与えられる行列である。これらのコントローラを用いてビジュアルサーボを行っている (詳細は文献^[3] 参照)。

4. 周波数応答特性測定実験を用いた GA の実時間

4.1 実験条件

ポーズトラッキング性能の解析

ハンドの初期姿勢を Σ_{E_0} とする。一方、物体の初期位置は Σ_{M_0} として定義する。 $\Sigma_W, \Sigma_{E_0}, \Sigma_{M_0}$ の関係を図 7 に示す。 Σ_W から Σ_{M_0} への同次変換行列、 Σ_W から Σ_{E_0} への同次変換行列は以下のように与えられる。

$${}^W T_{M_0} = \begin{bmatrix} 0 & 0 & -1 & -1435[mm] \\ 1 & 0 & 0 & 0[mm] \\ 0 & -1 & 0 & 499[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$${}^W T_{E_0} = \begin{bmatrix} 0 & 0 & -1 & -890[mm] \\ 1 & 0 & 0 & 0[mm] \\ 0 & -1 & 0 & 499[mm] \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

対象物は Σ_{M_0} の z 軸に沿って運動し、次式で与える。

$${}^{M_0} z_M(t) = -150 + 150 \cos(\omega t)[mm] \quad (7)$$

物体とハンドの相対的位置・姿勢の目標関係は時不変とし、

$${}^{Ed} \psi_M = [0, -90[mm], 545[mm], 0, 0, 0] \quad (8)$$

と定める。GA の探索範囲は $-200 \leq t_x \leq 200[mm]$, $-195 \leq$

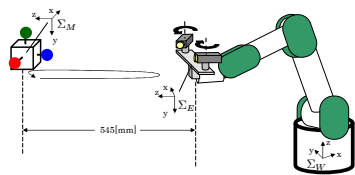


Fig. 7 Coordinate system

$t_y \leq 5[mm]$, $350 \leq t_z \leq 750[mm]$, $-0.3 \leq \epsilon_1, \epsilon_2, \epsilon_3 \leq 0.3$ と設定した。また、適合度の値が低い場合、対象物を誤認識していると考えられるので、実験中の安全性を考慮し、適合度の値が 0.1 以下の場合、PA10 ハンド部分は停止するように設定した。

4.2 実験結果

実験条件にて、位置/姿勢の 6 変数のうち、位置 3 変数を未知とし、姿勢 3 変数に真値を与えた場合において、対象物を周期的に動かす周波数応答実験を行った。物体とハンドの相対的位置/姿勢の関係は式 (8) とする。本報では対象物の周期 z 軸方向に $\omega = 0.209$, $\omega = 1.256$ の二種類の場合においてビジュアルサーボを行い、 z - x 方向における適合度の全探索を行った。全探索で得られた結果に GA の位置/姿勢を書き込むことにより GA と対象物との関係を示す。

図 8, 図 9 は Σ_E での z - x 方向において $\omega = 0.209$, 図 10, 図 11 は Σ_E での z - x 方向において $\omega = 1.256$ で 4 秒ごとに適合度を全探索することによって得られた結果と GA 遺伝子を白丸で表示し、その時の Σ_W におけるハンドの奥行き方向の位置と GA による認識結果、実際の対象物の奥行き方向の位置の関係、GA による認識結果を示したものである。色が濃いほど適合度が高く、色が薄いと適合度が低い。

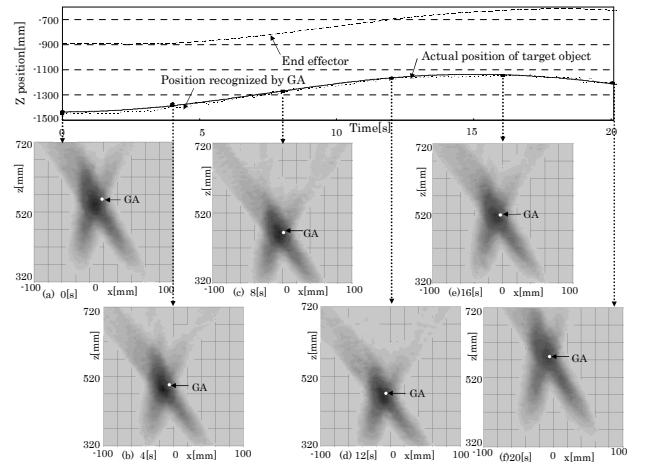


Fig. 8 Relation between z-x fitness value, end effector, actual position of target object and position recognized by GA in T=30 (a)

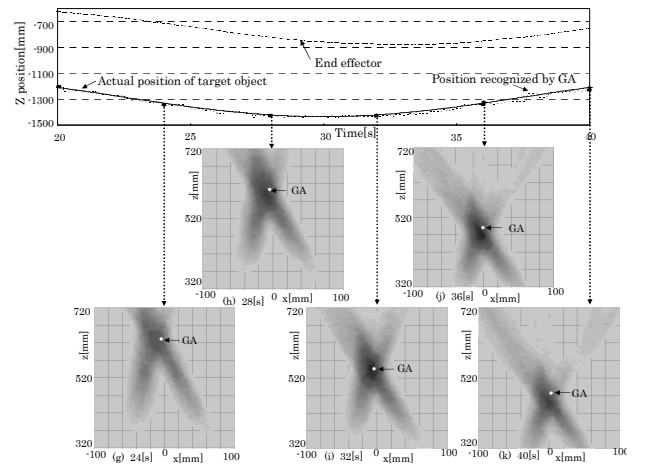


Fig. 9 Relation between z-x fitness value, end effector, actual position of target object and position recognized by GA in T=30 (b)

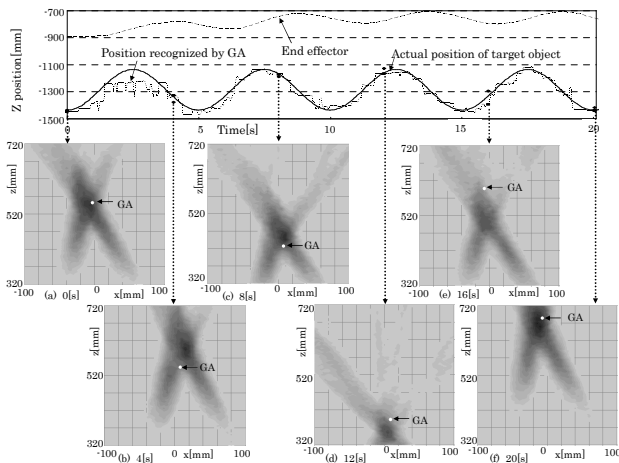


Fig. 10 Relation between z-x fitness value, end effector, actual position of target object and position recognized by GA in T=5 (a)

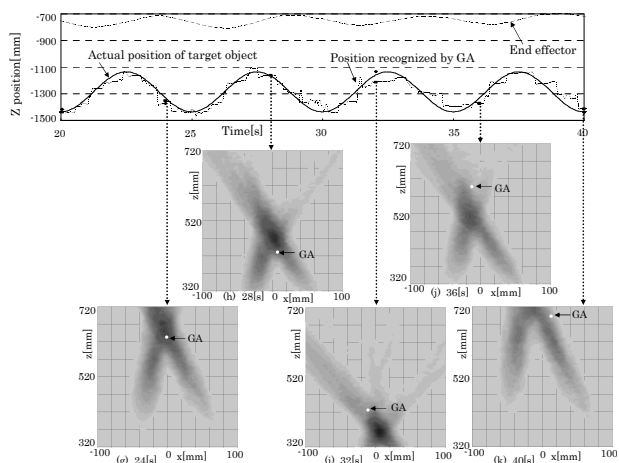


Fig. 11 Relation between z-x fitness value, end effector, actual position of target object and position recognized by GA in T=5 (b)

4.3 考察

適合度の全探索結果より、適合度は色の濃い、適合度の高い地点を中心にx状に分布していることが挙げられる。これは、本ビジュアルサーボは複眼構成であり、左右カメラから得られた画像を同時に評価し、適合度を求めているからである。(図12)

図8, 図9は対象物の速度が $\omega = 0.209$ と遅いので、GAは適合度の高い、色の濃い地点、つまり対象物を発見し、遅れることなく追従している。図10, 図11は対象物の速度が $\omega = 1.256$ と速いので、GAは適合度の低い、色の薄い地点に多く、対象物に対して遅れて追従している。また、適合度の全探索結果における対象物の位置・姿勢、特にz方向の位置とハンドとの関係は式(8)に示した設定値である545[mm]の間隔を保てていない。図10(d)や図11(i)は対象物とハンドのz方向の距離が設定値よりも近い関係であり(図13(a)), 図10(f)や図11(k)は対象物とハンドのz方向の距離が設定値よりも遠い関係である(図13(b))。これはハンドのトルクが大きく、対象物の速い動きに対して位相の遅れが生じているためである。

図13(a)のように、ハンドが対象物に対して近い位置にある場合でも、図10(d), 図11(i)のように、全探索で得られる適合度の最大値は高い値を示す。これは、図5(a)で示した、両眼撮像可能領域が拡大するというEye-vergenceシステムの利点を示している。図10の0秒から4秒の間で、実際の対象物の位置

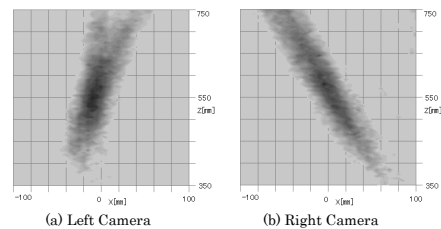


Fig. 12 Fitness value of the left camera and right camera

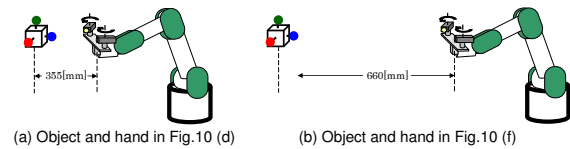


Fig. 13 Object and hand in Fig.10, Fig.11

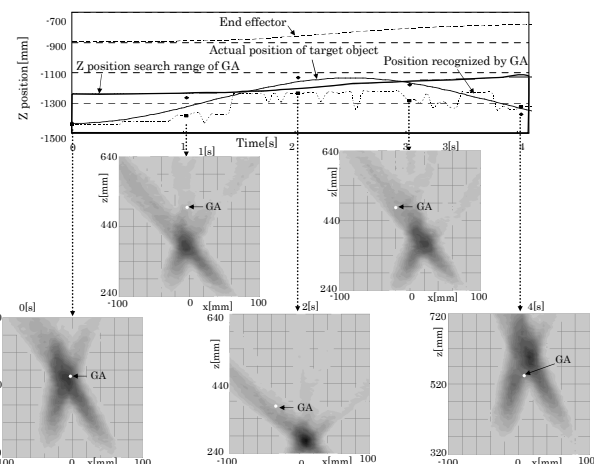


Fig. 14 Relation between z-x fitness value, end effector, actual position of target object and position recognized by GA in T=5 (c)

とGAによる認識結果の誤差が大きい。そこで、0秒から4秒の間を1秒単位で適合度の全探索を行った。結果を図14に示す。図10, 図11ではGAと対象物の認識の誤差が大きい場合は、適合度の分布が全体的に薄く、対象物を見つけやすい状態とはいえなかった。しかし、図14より、適合度の分布が濃く、対象物を見つけやすい状態であっても、GAは対象物を発見できていない場合があることがわかった。これは、ハンドの追従が遅く、GAの探索範囲外に対象物が位置しているためにおこると考えられる。

5. 結言

本報では、周波数応答実験においてビジュアルサーボを行い、ビジュアルサーボ中におけるGA遺伝子の挙動を示し、GAと対象物との関係性を明らかにすることにより、実時間認識追従誤差を把握することができた。GAの探索範囲外に対象物が位置している場合があったので、探索範囲の再考や、認識誤差の改善が課題である。

文献

- [1] H. Suzuki, M. Minami, "Visual Servoing to catch fish Using Global/local GA Search", IEEE/ASME Transactions on Mechatronics, Vol.10, Issue 3, 352-357 (2005.6).
- [2] B.Siciliano and L.Villani: *Robot Force Control*, ISBN 0-7923-7733-8.
- [3] 候 森, 侯 森, 見浪 護, 前田 耕市, 于福佳, 矢納 陽 "Eye-Vergenceを用いたビジュアルサーボの周波数応答特性", 日本ロボット学会第30回記念学術講演会, RSJ2012AC4I2-4, 2012