

クオータニオンを用いたオンライン3Dポーズ遺伝的認識法

宋薇(福井大) 見浪護(福井大) 前泰志(福井大) 青柳誠司(関西大)

On-line Stable Evolutionary Recognition Based on Unit Quaternion

*Wei Song (Univ. of Fukui), Mamoru Minami (Univ. of Fukui),
Yasushi Mae (Univ. of Fukui), Seiji Aoyagi (Kansai Univ.)

Abstract— This paper presents a pose measurement method of a 3D object. The proposed method utilizes an evolutionary search technique of the genetic algorithm (GA) and a fitness evaluation based on a matching stereo model whose pose is expressed by unit quaternion. To improve the dynamics of recognition, a motion-feedforward compensation method is proposed for the hand-eye system. The effectiveness of the proposed method is confirmed by the simulation experiments.

Key Words: Pose measurement, Unit quaternion, GA, Motion-feedforward compensation

1. 緒言

ビジュアルサーボは、運動中の対象物体の認識及びロボットの運動制御という技術的課題を含み、運動中の対象物に対して何らかの作業を行うことが可能となるため、ロボットの応用範囲を広げる技術的課題である。本論文では実時間3次元物体の位置/姿勢計測法の提案を行う。

ビジュアルサーボ系はフィードバックループ内部に視覚センサーを含むため、ビジュアルサーボ系の問題点として視覚センサーの認識時間の遅れが問題となる。これは認識系の中に存在するダイナミクスが原因であり、この視覚センサーの遅れを補償するため、物体の運動を予測しながらハンドを動かすことが重要になる。物体の運動モデルと非線形オブザーバを導入することによって、物体の運動を予測する手法[1]が提案されが、非線形オブザーバを用いて正しい推定値が得られたあとは遅れが解消されているが、誤差の小さい推定値が得られるまでに時間がかかるという問題は残っている（誤差がない推定値は理論的に無限の時間がかかる）。初期の推定誤差より制御系が不安定になってしまふ場合も考えられる。また、オブザーバを用いると、高次の運動モデルに対して演算量が多くなるという欠点がある。Lucaらは非線形オブザーバを用いて、既知のカメラの運動からカメラから対象物までの距離“Z”値を推定する方法を提案した[2]。しかし、初期時刻“t=0”では正しく推定できず、ある時間の後“Z”値の推定誤差は次第に減少する。また、単眼であるため既知のカメラの運動が必要であり、ハンドが動かないカメラ固定の状態では推定できない。一方、一般的なCCDカメラのサンプリング周期は33[ms]、関節サーボ系のサンプリング周期は1[ms]程度とこれらの中には大きな差がある。サンプリング周期が1[ms]と非常に高速な画像処理を行うことができるビジョンチップを用いると予測や補償などを用いなくても、高速移動物体の追従が可能になる[3]。しかし、この高速ビジョンシステムの高価で広く実用化することは難しい。

本研究では、複眼を持つハンドアイマニピュレータが運動中に対象物の3次元位置・姿勢を認識するとき

のダイナミクスの改善のためのMotion-Feedforwardと呼ぶ補償方法を提案する。ここでは、省略するためにMFFと呼ぶ。ビジュアルサーボ時のロボットの動きを認識部にフィードフォワードすることによって、認識系はハンドアイの位置・姿勢の変化に影響を受けないという特徴を持つ。MFFを用いて、ロボットが物体の真の運動とハンドアイが動くことによる相対的な運動を区別することができるため、“t=0”からロボットの動きによる対象物の運動を予測できる。また、非線形オブザーバのように正しい推定まで“無限”的な時間は必要がない。

2. Motion-Feedforward (MFF) 法

我々人間はいつも自分の運動より対象物の運動を予測している。こういう能力をロボットが持つていればもっと知能化に近づけると考える。ロボットのビジュアルサーボは、マニピュレータの手先にCCDカメラが装着されるアイ・イン・ハンドシステムを用いることが多い。その場合、カメラから見た対象物の動きは二つの要素：物体の実際の運動とカメラを装着したマニピュレータの運動に影響される。本研究では、マニピュレータの運動（関節角速度）を観測して、マニピュレータの運動からカメラから見た対象物の位置/姿勢を予測することに着目する。

ここで、ロボットの基準座標系 Σ_W 、物体の座標系 Σ_M 、右のカメラの座標系を Σ_{CR} とする。ハンドの手先座標系 Σ_H とし、右のカメラの座標系 Σ_{CR} と似通つて一致すると考えられる。

さて移動する座標系 Σ_{CR} 、 Σ_M 間の相対的移動速度について考える。図1に示すように Σ_W を基準にとり、 Σ_W から Σ_{CR} へのベクトルを Σ_W で表したもの ${}^W r_{CR}$ 、 Σ_W から Σ_M へのベクトルを Σ_W で表したもの ${}^W r_M$ とする。また Σ_{CR} から Σ_M へのベクトルを Σ_{CR} で表したもの ${}^{CR} r_{CR,M}$ とする。本研究ではカメラがロボットハンドの手先に装備されるので、カメラの速度とハンドの手先の速度は一致すると考える。そのため、 Σ_{CR} と Σ_W の回転行列 ${}^W R_{CR}$ は関節速度の関数になる。このとき次の関係が成立する。

$${}^{CR} r_{CR,M} = {}^{CR} R_W(q)({}^W r_M - {}^W r_{CR}(q)). \quad (1)$$

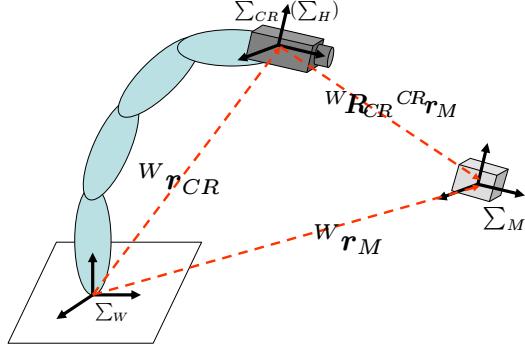


Fig.1 coordinate system

式(1)を時間について微分すると

$$^{CR}\dot{r}_{CR,M} = ^{CR}R_W(q)(^W\dot{r}_M - ^W\dot{r}_{CR}) + S(^{CR}\omega_W) \\ ^{CR}R_W(q)(^W r_M - ^W r_{CR}(q)). \quad (2)$$

ただし記号 $S(\cdot)$ は、任意の三次元ベクトル $a = [a_x, a_y, a_z]^T$ に対して、

$$S(a) = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}. \quad (3)$$

を意味する。

本稿では「4元数」クオータニオン(Quaternion)を用いて対象物の姿勢を表す。クオータニオンを使うとき、オイラ角、アングル/アクシス表現などで問題となる特異姿勢の問題を避けることができるので、3次元剛体の姿勢制御やロボットマニピュレータ制御などに適応する研究が盛んに行われている。ここで、 Σ_{CR} を基準とした Σ_M の回転角をクオータニオンで $^{CR}Q_M = \{^{CR}\eta_M, ^{CR}\epsilon_M\}$ とする。 $^{CR}\epsilon_M$ から $^{CR}\eta_M$ を導出できるため、三元数 $^{CR}\epsilon_M$ だけで対象物の姿勢を表す。 Σ_{CR} を基準とした Σ_M の座標原点の移動量を $^{CR}r_{CR,M}$ とすると、対象物の位置/姿勢は次の式で表される。

$$^{CR}\psi_M = \begin{bmatrix} ^{CR}r_{CR,M} \\ ^{CR}\epsilon_M \end{bmatrix}. \quad (4)$$

ここで、 $^{CR}r_{CR,M} = [t_x, t_y, t_z]^T$, $^{CR}\epsilon_M = [\epsilon_1, \epsilon_2, \epsilon_3]^T$ 。対象物の位置/姿勢速度を次の式で定義する。

$$^{CR}\dot{\psi}_M = \begin{bmatrix} ^{CR}\dot{r}_{CR,M} \\ ^{CR}\dot{\epsilon}_M \end{bmatrix}. \quad (5)$$

回転速度に関して、カメラ座標系 Σ_{CR} の角速度を Σ_W で表したベクトルを $^W\omega_{CR}$ 、物体の座標系 Σ_M の角速度を Σ_W で表したベクトルを $^W\omega_M$ とするとき、 Σ_{CR} から見た対象物の回転角速度 $^{CR}\omega_{CR,M}$ は

$$^{CR}\omega_{CR,M} = ^{CR}R_W(q)(^W\omega_M - ^W\omega_{CR}). \quad (6)$$

Σ_{CR} から見た対象物の位置速度 $^{CR}\dot{r}_{CR,M}$ を式(2)で表し、姿勢速度 $^{CR}\dot{\epsilon}_M$ と回転角速度 $^{CR}\omega_{CR,M}$ との関

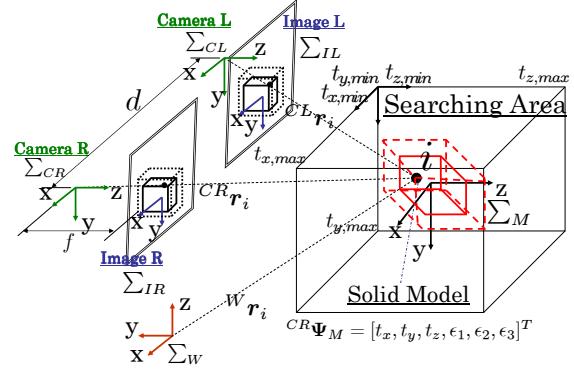


Fig.2 Coordinate systems

係は

$$^{CR}\dot{\epsilon}_M = \frac{1}{2} (^{CR}\eta_M I - S(^{CR}\epsilon_M)) ^{CR}\omega_{CR,M} \quad (7)$$

となる。

カメラの速度はハンドの手先の速度と等しくヤコビ行列 $J(q) = [J_p^T(q), J_o^T(q)]^T$ を用いて表現できる。ここで、手先の並進速度と関節速度との関係を表したヤコビ行列の上3行 $J_p(q)$ 、手先の回転角速度と関節速度との関係を表したヤコビ行列の下3行は $J_o(q)$ と表現される。

$$^W\dot{r}_{CR} = J_p(q)\dot{q}, \quad (8)$$

$$^W\omega_{CR} = J_o(q)\dot{q}, \quad (9)$$

$$S(^{CR}\omega_W) = -^{CR}R_W(q)S(^W\omega_{CR})^W R_{CR}(q) \\ = -^{CR}R_W(q)S(J_o(q)\dot{q})^W R_{CR}(q). \quad (10)$$

式(8), (9), (10)を用いて式(2), (7)を書き直すと、カメラから見た対象物の運動速度

$$\begin{aligned} ^{CR}\dot{\psi}_{CR,M} &= \begin{bmatrix} ^{CR}\dot{r}_{CR,M} \\ ^{CR}\dot{\epsilon}_M \end{bmatrix} \\ &= \begin{bmatrix} -^{CR}R_W(q)J_p(q) + ^{CR}R_W(q) \\ S(^W R_{CR}(q)^{CR}r_{CR,M})J_o(q) \\ -\frac{1}{2} (^{CR}\eta_M I - S(^{CR}\epsilon_M)) ^{CR}R_W(q)J_o(q) \end{bmatrix} \dot{q} \\ &\quad + \begin{bmatrix} ^{CR}R_W(q) & 0 \\ 0 & ^{CR}R_W(q) \end{bmatrix} \begin{bmatrix} ^W\dot{r}_M \\ ^W\dot{\epsilon}_M \end{bmatrix} \\ &= J_m(q)\dot{q} + J_n(q)^W \dot{\psi}_M \end{aligned} \quad (11)$$

を得る。

式(11)を用いれば、 Σ_{CR} から見た対象物の位置速度 $^{CR}\dot{r}_{CR,M}$ 及び姿勢速度 $^{CR}\dot{\epsilon}_M$ とマニピュレータの関節速度 \dot{q} の関係が $J_m(q)$ で、対象物の実際の位置/姿勢速度 $[^W\dot{r}_M, ^W\dot{\epsilon}_M]^T$ との関係が $J_n(q)$ で簡潔に表現できる。

本稿で提案した motion-feedforward (MFF) 方法は、対象物の実際の運動速度を考えずに、式(11)の関係行列 $J_m(q)$ を用いてマニピュレータの関節速度より Σ_{CR}

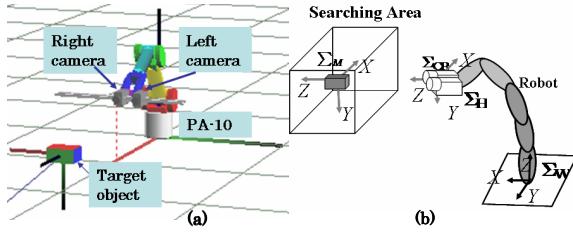


Fig.3 (a) Simulation experiment system created by OpenGL. (b) coordinate systems of simulation experiment

から見た対象物の位置/姿勢速度を予測することに着目し、カメラの動きから認識への影響を抑える。そのため、 ${}^W\dot{\psi}_M$ を無視し式(11)を書き直して

$${}^{CR}\dot{\psi}_{CR,M} = J_m(q)\dot{q}. \quad (12)$$

となる。

式(12)を用いれば、 $t + \Delta t$ 時の対象物の位置/姿勢を予測することができる。予測した $t + \Delta t$ 時の対象物の位置/姿勢は ${}^{CR}\hat{\psi}_M = [{}^{CR}\hat{r}_{CR,M}, {}^{CR}\hat{\epsilon}_M]^T$ と定義し、

$${}^{CR}\hat{\psi}_M(t + \Delta t) = {}^{CR}\psi_M(t) + {}^{CR}\dot{\psi}_M \Delta t. \quad (13)$$

を得る。 ${}^{CR}\dot{\psi}_M \Delta t$ は時刻 t の位置/姿勢より将来の時刻までどのくらい変化があるかを示す。式(13)を使って Δt 後の将来の対象物の位置/姿勢を予測できると考える。

3. 三次元位置/姿勢計測方法

探索方法としては、まず3D空間内に対象物と同じ形状かつ同じ大きさで、位置と姿勢情報を持つ仮想的な立体モデルを配置する。本研究では、特異姿勢がないクオータニオンを用いて対象物や立体モデルの姿勢を表す。その立体モデルから射影幾何学に基づく射影変換によって左右画像における2次元探索モデルを求め、このモデルと画像との間で相関関数値の分布が決まる。探索する対象物と探索領域に配置した立体モデルが一致すると入力画像中においても、認識対象物と探索モデルとが一致し、相関関数は最大値を示す。以下では、相関関数値 F の最大値を与える ψ を解く最適化問題をGAで求解するため、 F を適合度関数と呼ぶ。

ここで、最大値を与える変数の値の求解方法として、最適解探索に遺伝的アルゴリズム(GA)を用いる。本研究における個体の位置/姿勢はGAの遺伝子情報 ${}^{CR}\psi_{GA}$ で表す。GAのリアルタイム性の問題に対し、ここでは目標物体の探索操作の終了を待たずに、探索中の位置/姿勢の遺伝子の中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を適用する。実時間遺伝的認識法は“1-Step GA”と呼ばれ、泳ぐ魚をロボットにより捕獲する研究事例より実時間認識が可能であることを示した[4]。

MFF認識法を用いると、カメラから見た対象物の動きをロボットの運動から予測できるので、一旦GAが対象物の付近に収束したら、個体をカメラの動きと共に

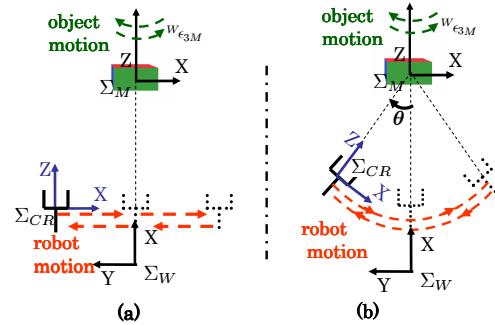


Fig.4 (a) motion "A": given position changing of end-effector. (b) motion "B": given orientation changing of end-effector

に動かすことで、ロボットの運動が速くてもGAの個体が対象物の付近に収束することを維持できる。

式(12)の $J_m(q)$ を用いて対象物の運動速度 ${}^{CR}\dot{\psi}_{CR,M}$ を求め、それを用いて次の $(i+1)$ 世代の個体の位置/姿勢を予測することができる。

$${}^{CR}\hat{\psi}_{GA}^{i+1} = {}^{CR}\psi_{GA}^i + {}^{CR}\dot{\psi}_{CR,M} \Delta t. \quad (14)$$

式(14)を用いて予測したカメラ座標内の対象物の運動速度でGAの個体を動かすことで、カメラの動きによる認識への影響を抑えることができる。そこで、カメラを動かしても、固定したカメラのように対象物を認識することができると考えられる。

4. シミュレーション

提案した3D位置/姿勢計測法と実時間認識法を用いて、直方体形状の対象物の認識シミュレーションを行う。対象物は、各面に赤、緑、青の色を持ち、大きさは $(x, y, z) = (250, 100, 150)[mm]$ である。Fig. 3(a)に作成したシミュレータの画面を示す。Fig. 3(b)にシミュレータシステムの各座標関係を示す。

対象物の目標の姿勢変化は、

$${}^W\epsilon_{3M} = \sin(15\sin(\omega_o t)/2) \quad (15)$$

と与え、他のパラメーターは $({}^Wt_{xM}, {}^Wt_{yM}, {}^Wt_{zM}) = (-75, 150, 800)[mm]$, $({}^W\epsilon_{1M}, {}^W\epsilon_{2M}) = (0, 0)$ に設定する。対象物の運動の周波数は ω_o 、マニピュレータの周波数は ω と定義する。ここでは、 $\omega_o = 0.125[rad/s]$ と設定する。マニピュレータが動かない場合は $\omega = 0[rad/s]$ となる。

(1) 手先を左右往復並進する状態の認識 (Fig. 4(a))

ここでは、マニピュレータは Σ_W のy軸方向を左右往復並進という目標軌道(目標軌道の説明が省略)を設定した。マニピュレータの初期の状態をFig. 3(b)に示す。ここでは、初期探索対象物位置/姿勢が既知であると仮定する。つまり、 $t = 0$ のとき探索立体モデルは探索対象物の真の位置/姿勢情報を持つので、探索された位置/姿勢は探索対象物の真の位置と一致する。

ここではまず F の最大値を F_M とし、時刻を t_1, t_2, \dots, t_n と定める。このとき

$$\bar{F} = \frac{1}{n} \sum_{i=1}^m F_M(\psi_{t_i}) \quad (16)$$

$\bar{F}, \tilde{\Delta\psi}$		$\omega=0$	$\omega=0.125$	$\omega=0.25$	$\omega=0.5$
1-step GA	\bar{F}	0.9625	0.7047	0.5794	0.4378
	$\tilde{\Delta x}$ [mm]	1.238529	15.43175	16.53167	30.21539
	$\tilde{\Delta y}$ [mm]	3.963243	18.58162	26.21253	32.96989
	$\tilde{\Delta z}$ [mm]	5.544477	25.84152	36.72746	44.76831
	$\tilde{\Delta\epsilon}_1$	0.011146	0.018499	0.021081	0.039064
	$\tilde{\Delta\epsilon}_2$	0.015599	0.041138	0.054877	0.123591
FF. + 1-step GA	\bar{F}	0.9473	0.9556	0.9590	
	$\tilde{\Delta x}$ [mm]	2.128153	1.316153	1.923763	
	$\tilde{\Delta y}$ [mm]	6.43577	2.329286	3.173726	
	$\tilde{\Delta z}$ [mm]	6.873411	3.128897	4.313114	
	$\tilde{\Delta\epsilon}_1$	0.000485	0.001283	0.00084	
	$\tilde{\Delta\epsilon}_2$	0.00952	0.013151	0.013249	
	$\tilde{\Delta\epsilon}_3$	0.014976	0.015944	0.013401	

Fig.5 Conclusion of simulation under motion “A”
左右

を適合度の時間平均値と定義する。また、探索された位置/姿勢 $\psi_M(t_i)$ と探索対象物の真の位置/姿勢との差を認識の位置/姿勢誤差 $\tilde{\Delta\psi}$ と定義し、 $\tilde{\Delta\psi}(t_i) = [\tilde{\Delta x}, \tilde{\Delta y}, \tilde{\Delta z}, \tilde{\Delta\epsilon}_1, \tilde{\Delta\epsilon}_2, \tilde{\Delta\epsilon}_3]^T$ と表す。位置/姿勢の誤差の rms 値は、

$$\tilde{\Delta p} = [\tilde{\Delta x}, \tilde{\Delta y}, \tilde{\Delta z}, \tilde{\Delta\epsilon}_1, \tilde{\Delta\epsilon}_2, \tilde{\Delta\epsilon}_3]^T$$

$$\tilde{\Delta p} = \sqrt{\frac{1}{n} \left(\sum_{i=1}^m \tilde{\Delta p}^2(t_i) \right)}$$

$$\Delta p = \tilde{\Delta p} = \Delta x, \Delta y, \Delta z, \Delta\epsilon_1, \Delta\epsilon_2, \Delta\epsilon_3. \quad (17)$$

と定義する。

本シミュレーションでは 0[s] から 72[s] までの各世代の最大適合度の時間平均値 \bar{F} と位置/姿勢の誤差の rms 値 $\tilde{\Delta\psi}$ を用いて認識性能を評価する。クオータニオンで姿勢を表すとき、変数 $\epsilon_1, \epsilon_2, \epsilon_3$ は単位が明確でない。対象物が x 軸より角度 1[deg] で回転した姿勢はクオータニオンで表すと $\epsilon_1 = 0.008, \epsilon_2 = 0, \epsilon_3 = 0$ になる。つまり、0.008 のクオータニオンの値は角度約 1[deg] を表すことが分かる。

Fig. 5 は異なるハンド周期運動の角周波数 ω でのシミュレーションの \bar{F} と $\tilde{\Delta\psi}$ を計算してまとめたである。MFF を用いない “1-step GA” だけで認識する場合はマニピュレータの周波数 ω は 0[rad/s] から 0.5[rad/s] まで増加するとともに、各世代の最大適合度の平均値 \bar{F} は小さくなつて、位置/姿勢の誤差の rms 値 $\tilde{\Delta\psi}$ は大きくなる。最大位置誤差は約 35[mm]、最大姿勢誤差は約 12[deg] になつてしまつ。つまり、認識の性能はだんだん低くなる。“1-step GA + MFF” を用いるとき、マニピュレータの運動速度が速くなつても、 \bar{F} と $\tilde{\Delta\psi}$ はあまり変化しない。位置誤差は約 3[mm]、姿勢誤差は約 1[deg] 程度と維持している。MFF 認識法、マニピュレータの運動から認識への影響を抑えて、固定したカメラと同じように認識できることを Fig. 5 は示している。

(2) 手先を左右往復回転する状態の認識 (Fig. 4(b))

ここでは、マニピュレータは左右往復回転という目標軌道 (目標軌道の説明が省略) を設定する。左右往復回転運動というのはカメラから見た対象物の位置を変

$\bar{F}, \tilde{\Delta\psi}$		$\omega=0$	$\omega=0.125$	$\omega=0.25$	$\omega=0.5$
1-step GA	\bar{F}	0.9625	0.7445	0.5089	0.3987
	$\tilde{\Delta x}$ [mm]	1.238529	7.084036	12.50464	15.94141
	$\tilde{\Delta y}$ [mm]	3.963243	15.10137	19.43937	24.79264
	$\tilde{\Delta z}$ [mm]	5.544477	21.11374	26.54229	33.70501
	$\tilde{\Delta\epsilon}_1$	0.011146	0.023746	0.030825	0.069354
	$\tilde{\Delta\epsilon}_2$	0.015599	0.073936	0.106551	0.118325
FF. + 1-step GA	\bar{F}	0.9556	0.9682	0.9443	
	$\tilde{\Delta x}$ [mm]	1.212835	2.475826	2.069171	
	$\tilde{\Delta y}$ [mm]	5.841326	3.708801	5.811863	
	$\tilde{\Delta z}$ [mm]	3.082948	4.88349	4.00924	
	$\tilde{\Delta\epsilon}_1$	0.006454	0.012079	0.008623	
	$\tilde{\Delta\epsilon}_2$	0.016637	0.02126	0.024202	
	$\tilde{\Delta\epsilon}_3$	0.017566	0.012186	0.013883	

Fig.6 Conclusion of simulation under motion “B”
回転

化させずに、姿勢だけで左右回転するためのマニピュレータの運動である。ここでも、初期のときの探索対象物の位置/姿勢が分かると仮定する。

また、Fig. 6 に Fig. 5 と同じような三つの角周波数でのシミュレーションの \bar{F} と $\tilde{\Delta\psi}$ をまとめる。Fig. 6 より、“1-step GA + MFF” を用いるとき、マニピュレータの運動速度が速くなつても、ハンド静止時に得られる高い \bar{F} と低い $\tilde{\Delta\psi}$ を維持することができる。提案した “1-step GA + MFF” 認識方法の有効性を再び確認した。

5. 結論

本研究では、動画像中から対象物の位置 / 姿勢六変数を同時に認識することを目的として、Model-based Matching 法と GA を用いて 3 次元位置 / 姿勢計測法を提案した。また、マニピュレータの手先に CCD カメラが装着されるアイ・イン・ハンドシステムを用いて認識する場合は、マニピュレータの運動が対象物認識に与える影響を抑えるために “1-step GA + MFF” 認識方法が有効であることを示した。

将来的には、実際にロボットマニピュレータによる対象物（六変数）へのビジュアルサーボを行いたい。“1-step GA + MFF” を用いてロボットが物体の真の運動と自分の動きからの偽の相対的な運動を区別でき、マニピュレータの運動から認識への影響を抑えることでより安定的なビジュアルサーボを期待する。

- [1] 橋本, 木村：“ビジュアルサーボイング—非線形オプザーバアプローチー”，日本ロボット学会誌，Vol.13, No.7, pp986-993, 1995
- [2] A. De Luca, G. Oriolo and P. R. Giordano ”Online Estimation of Feature Depth for Image-Based Visual Servoing Schemes”, IEEE Int. Conf. on Robotics and Automation (ICRA2007).
- [3] 中坊, 石川：“1ms 高速ビジョンを用いたビジュアルサーボイング”，計測と制御，Vol.40, No.9, pp636-640, 2001
- [4] M. Minami, H. Suzuki, J. Agbanhan, T. Asakura: ”Visual Servoing to Fish and Catching Using Global/Local GA Search”, 2001 IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics Proc., pp.183-188, 2001.