

3次元再構成を伴わない観察による他者行動の模倣

Imitation by observation without 3-D reconstruction .

吉川 雄一郎 (阪大) 浅田 稔 (阪大)

Yuichiro YOSHIKAWA, Osaka University, 2-1, Yamadaoka, Suita, Osaka

Minoru ASADA, Osaka University

Abstract: This paper describes an imitation mechanism by observation, in a situation that an imitator's body structure is the same as demonstrator's one. The imitation mechanism consists of a desired trajectory generator and a follower. The desired trajectory generator applies the stereo epipolar constraint, by which 3-D reconstruction is not necessary, to generate a trajectory of endeffector in the stereo image pair, to imitate the observed behavior. The follower reproduces the observed behavior using the adaptive visual servoing to position own endeffector onto the desired trajectory.

Keywords: imitation learning, stereo epipolar constraint, no 3-D reconstruction

1 はじめに

ロボットに他者とのコミュニケーションの能力を持たせることは、教示者とのコミュニケーションによりロボットの学習を促進したり、協調行動を効果的に行なえるようになり、有用である。

最近サル¹⁾の F5 野にミラーニューロンと呼ばれるニューロンが発見されている。ミラーニューロンはサルが物をつかんだり、動かしたりするとき、および同じような行動をしている他人を観察したときの両方で発火するものである。このミラーニューロンのシステムと似たシステムが、人間のスピーチを司るという Broca 野に見つかっていることに基づき、Rizzolatti and Arbib¹⁾らはコミュニケーション能力の発達に行動の模倣、および模倣した行動の理解が関わると考えた。

Schaal²⁾は模倣行動を行動科学および認識科学、神経科学および認知神経学、ロボティクス および AI およびニューラルコンピューテーションの3つの観点から模倣について考察し、多自由度のヒューマノイドロボティクスの運動学習に通じる道として、プリミティブ形式の模倣学習を用いることの有用性を示唆した。模倣の本質的な必要条件は知覚と運動の結合であるとして、運動と知覚の表現形式、運動のプリミティブ、運動を発生するメカニズムでの運動認識、タスクのゴール理解についても考察し、それらが模倣に必要である事を指摘した。

ロボットの認知過程を理解し設計する認知ロボティクスの立場からも模倣行動の生成、およびその理解(自己の行動群とのマッチング)の能力は必要であると考えられる³⁾。呈示者の行動を模倣するためには、その呈示された行動を呈示者の座標系から自己の座標系に座標変換するシステムを持たなければならない。個々のロボットは、他者の内部状態については知り得ないとすると、相互の座標変換マトリクスは明示的には得られない。しかしその座標変換のマトリクスの代りとなるメカニズムが必要である。

従来の模倣行動の研究では観察した行動を自己の座標系に座標変換するのに3次元再構成のプロセスを含んでいる場合が多い。2次元画像からの3次元再構成

は計算のための事前知識(カメラパラメータやロボットの運動学パラメータ)が必要になる。ここでは3次元再構成を行わない手法を提案する。

ステレオ視覚を持つ学習ロボットと、学習ロボットと合同な身体構造(スケールファクタを除いて全く等価)を持つ呈示ロボットを考え、呈示ロボットの呈示する行動を視覚により捉え模倣するタスクを考える(Fig.1)。学習ロボットが、自己と呈示ロボットの身体構造の同一性を知っているという仮定に基づき、呈示された行動を観察し、ステレオエピポーラ拘束⁴⁾により、模倣すべき行動軌道を自己のセンサの座標系に写像し、適応型ビジュアルサーボ⁵⁾を用いて、観察による模倣を実現する。

2 手法

3次元再構成を伴わない、観察による他者行動の模倣メカニズムを次の2つの構成要素により実現する。

- ステレオエピポーラ幾何を利用して、観察した呈示ロボットの手先の画面上の軌道を学習ロボットの視点の画像平面に写像する。
- 適応型ビジュアルサーボによってその軌道を再現することにより模倣行動を実現する。

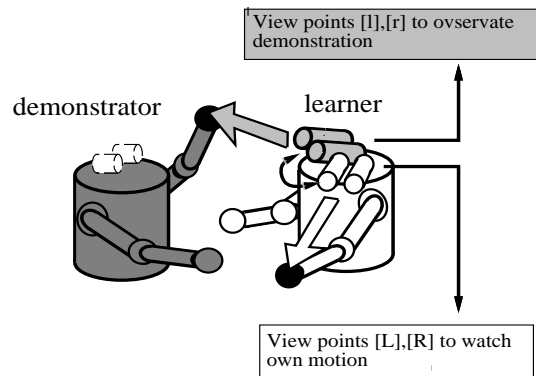


Fig.1 a learner and a demonstrator

2.1 エピポーラ幾何⁴⁾

3次元空間のある点 P をステレオ画像で捉えるとき、1つの視点の画像系列 $[l]$ への投影点 p_l が決まると、光幾何学より、もう片方の視点の画像系列 $[r]$ への投影点 p_r は直線(エピポーラ線) l_r 上に拘束される。この関係を式で表すと

$$p_r^T r F_l p_l = 0 \quad (1)$$

ここに

$$p_r = [X_r \ Y_r \ 1]^T \quad p_l = [X_l \ Y_l \ 1]^T$$

ただし $r F_l$ は 3×3 の行列で、画像系列 $[l], [r]$ のエピポーラ幾何を表す行列であり、ステレオカメラの相対的な位置関係、姿勢、焦点距離などのパラメータを含む。

2つの画面間の未知のエピポーラ拘束 $r F_l$ は $3 \times 3 = 9$ の未知数を持つ。このうち1つをスケールングファクターとして定めれば、(1)の方程式が8本あれば $r F_l$ を得ることができる。つまり3次元空間上の8つの点が両方の画像に投影されていれば、その8つの投影点の関係式(1)の連立方程式を解くことによってエピポーラ拘束を推定することができる。

ただし連立方程式に用いる3次元空間上の点が特別な配置にあるとき(視点を含む直線上にある、立方体を形成する、など)は連立方程式のランクが落ち、エピポーラ拘束を求ることができないことが示されている。このためエピポーラ拘束を推定する式(1)の数を増やし、つまり投影点を8点以上使い、連立方程式のランクが落ちないようにして最小自乗推定によりエピポーラ拘束を推定する。

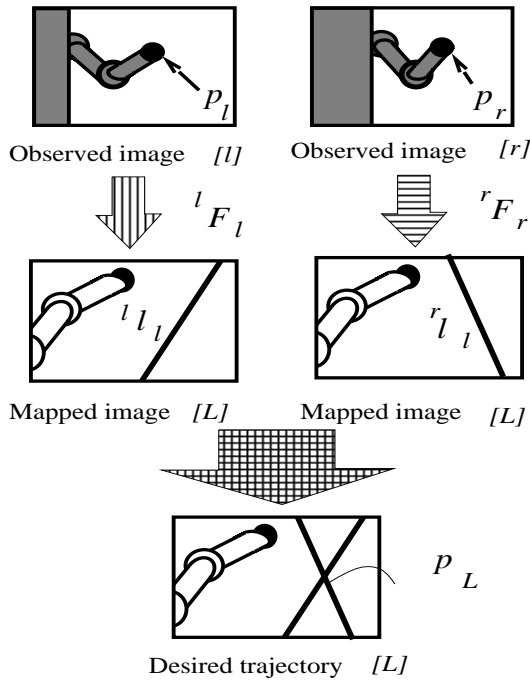


Fig.2 a translation method based on stereo epipolar constraint

2.1.1 エピポーラ幾何による点の写像

画像系列 $[l], [r], [L]$ に対して、 $l F_L, r F_L$ が既知であるとき、3次元空間の点 P の2つの画像系列 $[l], [r]$

への投影点 p_l, p_r はもう1つの視点の画像系列 $[L]$ 上のエピポーラ直線 l_L, r_L 上に限定される。したがってこのエピポーラ直線の交点として、画像系列 $[L]$ 上の対応点 p_L を見つけることができる (Fig.2)。

2.2 エピポーラ拘束を利用した目標軌道の写像

適応型ビジュアルサーボを用いて模倣行動を実現するために、学習者が呈示者の示す動作を観察する画像系列 $[l], [r]$ の軌道を、学習者自身を観察する画像系列 $[L], [R]$ に写像することを考える。

呈示者と学習者は構造が合同であるので、原理的に呈示者の示す行動を学習者が再現することが可能である。再現された行動を模倣行動と呼ぶ。

呈示者と学習者のスケールが一致している場合、呈示者の示す動作を学習者が観察する画像系列 $[l], [r]$ と、学習者が模倣行動を実現している様子を呈示者が観察する画像系列 $[L], [R]$ は等価である (Fig.3)。

画像系列 $[L], [R]$ と学習者自身を観察する視点の画像系列 $[l], [r]$ は同一の学習者を捉えているので、それぞれエピポーラ幾何が推定可能である。したがって2.1.1節の手法により呈示者から観察するステレオ画像系列 $[l], [r]$ の目標軌道を学習者自身を観察するステレオ画像系列 $[L], [R]$ への写像が可能である。

ここに画像系列 $[L], [R]$ と画像系列 $[l], [r]$ が等価であるため、実際に存在する学習者が観察する画像系列 $[l], [r]$ と学習者自身を観察する画像系列 $[L], [R]$ のエピポーラ幾何を推定し、 $[l], [r]$ の呈示者の示す動作軌道を $[L], [R]$ に写像して軌道を生成する。これは学習者が再現すべき模倣行動の目標軌道になる。

2.1.1節の手法による写像はスケールに依存しないため、呈示者と学習者のスケールが一致しない場合でも身体構造が合同であれば、同様に学習者自身を観察するステレオ画像系列 $[L], [R]$ への写像が可能である。

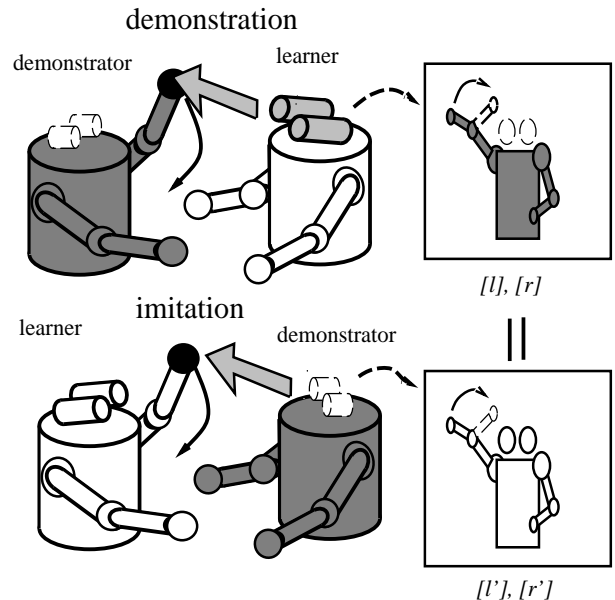


Fig.3 identifying learner's view with demonstrator's images

2.3 適応型ビジュアルサーボ⁵⁾に基づく目標軌道の実現

適応型ビジュアルサーボとはシステムの構造やパラメータに関する知識を必要としない、オンラインのヤコビ行列推定則によるヤコビ行列を用いてフィードフォワード項を含んだビジュアルサーボ系を構成する手法である。

ロボットシステムの運動を記述するシステム記述変数ベクトル $\theta \in \mathbb{R}^m$, 注視対象の画像上での特徴量を $x \in \mathbb{R}^m$ とすると運動のシステムは次式のように表せる。

$$x = x(\theta) \quad (2)$$

上式を微分すると

$$\dot{x} = J(\theta)\dot{\theta} \quad (3)$$

$J(\theta)$ は画像特徴量速度とシステム記述変数速度の関係を表すヤコビ行列である。

サンプリング時間 T が十分短く、その間ヤコビ行列 J が一定とみなせると仮定すると $u = T\dot{\theta}$ とし上式を離散化すると、

$$x(k+1) = x(k) + J(k)u(k) \quad (4)$$

ヤコビ行列 J の第 i 行ベクトルを j_i^T とすると

$$\{j_i^T(k+1) - j_i^T(k)\}u(k+1) = \{x(k+2) - x(k+1) - J(k)u(k+1)\}_i \quad (5)$$

これを満たす j_i^T で重みづけされた変化率が最小になる推定値 \hat{j}_i^T を外挿によって求めることとし、また外乱等による影響を考慮し、 $\|u\| = 0$ のときにも右边が発散しないように、推定則を次式のように定める。

$$\hat{j}_i^T(k+1) - \hat{j}_i^T(k) = \frac{\{x(k+1) - x(k) - \hat{J}(k)u(k)\}_i}{\rho_i + u(k)^T W_i(k)u(k)} W_i(k)u(k) \quad (6)$$

ここで $W_i(k)$ は $\hat{j}_i^T(k)$ の変化に対する重み行列であり、共分散行列をとる。 ρ_i は右边が発散しないための正の数であり、 ρ_i の範囲を $0 < \rho_i \leq 1$ と限定すると、この推定則は最小二乗法となる。つまりこの推定則は最小二乗法を内包していることになる。

ヤコビ行列の推定値 \hat{J} を用いることにより、画像特徴量 $x(k)$ を与えられた目標値 $x_d(k)$ に追従させるビジュアルサーボ系を構成する。式 (4) より、サーボ系として、

$$u(k) = \hat{J}^+(k)\{x_d(k+1) - x_d(k)\} + \{I_n - \hat{J}^+(k)\hat{J}(k)\}k_r + K\hat{J}^T(k)\{x_d(k+1) - x(k)\} \quad (7)$$

を得る。ここで $\hat{J}^+(k)$ は行列 $\hat{J}(k)$ の疑似逆行列、 I_n は $n \times n$ の単位行列、 K は正定ゲイン行列である。

この制御則の右边第 1, 2 項は、ヤコビ行列が十分に推定されているときには、注視対象の画像特徴量を、与えられた目標値に追従させることのできるフィードフォワード項である。右边第 2 項はシステムの冗長性を表す項であり、 k_r はその冗長性を記述する項である。これに対して右边第 3 項は、フィードバック項であり、閉ループシステムの安定性を保証する項である。

この手法を用いて写像した目標軌道を実現する。したがって 3 次元再構成を伴わないので学習者の身体に関する知識は全く必要ない。

3 実験

提案する手法によって模倣行動が実現できるのか、実験を行なう。学習ロボット (*learner*) が視点 $[l], [r]$ より呈示ロボット (*demonstrator*) の行動を観察し (step 1)、ついで視点 $[L], [R]$ より学習ロボット自身の体を見ながら模倣行動を実現する (step 2)(Fig.4)。

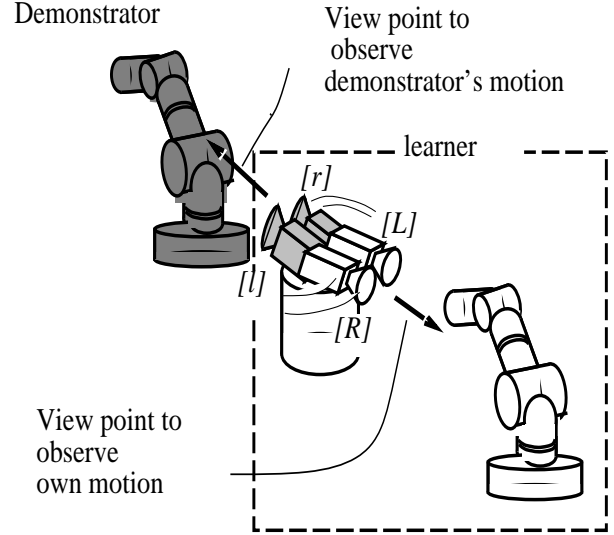


Fig.4 step1 and step2

3.1 実験システム

step 1 と step 2 は同時に起さず一台の 7 自由度 マニピュレータ PA10 に学習者と呈示者の両方の役割を負わせる。そのため学習者が呈示された行動を観察する視点 $[l], [r]$ のステレオカメラ、学習者が自分の行動を観察する視点 $[L], [R]$ のステレオカメラの計 4 台を Fig.4 の位置関係に相当するように配置する。

Fig.5 はその実験のシステムを示したものである。4 台の CCD カメラの画像情報は横、および縦方向に 1/2 に圧縮され、一枚の画像 (640 × 480) として富士通製のトラッキングモジュールに送られる。このトラッキングモジュールは前もって与えられた画像テンプレートと、現在の画像の相関を取ることによって、そのテンプレートと相関がもっとも高い画像が現在の画像平面上のどこにあるかを検出できる。このトラッキングモジュールによって得られた座標はターゲットコンピュータ (OS:VxWorks) に送られ、計算が行なわれる。ターゲットコンピュータからリアルタイムネットワーク (arcnet, 2.5Mbps) を介してロボットコントローラに制御量が送られ PA10 を駆動する。ロボットコントローラはターゲットコンピュータにマニピュレータの関節角度を返す。制御プログラムはホストコンピュータ (IBM Think Pad) 上の Tornado によって開発しイーサネット経由でターゲットコンピュータ上で実行される。本実験では手先をひねる動きの自由度 1 つと手先の姿勢の 2 つの自由度を固定して 4 自由度をマニピュレータが持つ自由度としている。

3.2 目標軌道の写像

学習者自身を観察する視点 $[L], [R]$ の画像上に視点 $[l], [r]$ で観察した呈示者の手先の行動軌道を写像する。2.2 節より学習者自身を観察する視点 $[L], [R]$ の画像

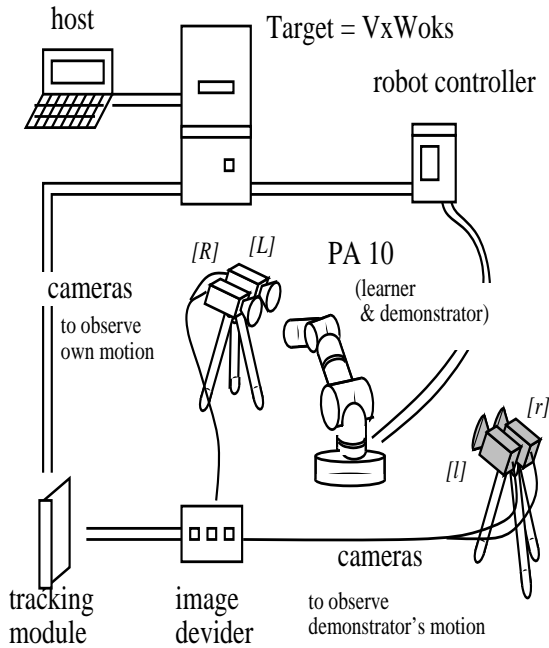


Fig.5 a setup of experiment

と呈示された行動を観察する視点 $[l]$ $[r]$ の画像の座標を用いてエピポラ拘束を推定する。

3.2.1 エピポラ拘束の推定

PA10 の手先にターゲットをつけ $[L]$ $[R]$ $[l]$ $[r]$ の 4 つの視点からそのターゲットをトラッキングする。PA10 に 56 種類の適当な関節角度を次々指定して 4 つの視点に投影された同一点の座標を 56 組 (p_L, p_R, p_l, p_r) 得る。これをそれぞれ式 (1) に代入し ${}^lF_L, {}^rF_L, {}^lF_R, {}^rF_R$ の行列の成分を変数とするそれぞれの連立方程式を最小自乗近似により解き iF_j を得る。

iF_j の最小自乗推定の残差 (ほぼ 1 桁) の自乗の平均値と分散は Table.1 のようになった。

Table 1 residuals of estimating F matrix

	平均値	分散
lF_L	0.000148	0.000000
rF_L	0.000155	0.000000
lF_R	0.000942	0.000003
rF_R	0.007990	0.015484

3.2.2 エピポラ拘束を利用した写像による目標軌道の生成

画像上に S 字の軌道を描くような関節角の組を PA10 に与え、PTP 制御を行って、PA10 に行動を呈示させる。このときに手先のターゲットの軌道を呈示する行動とする。

呈示行動の観察視点の画像系列 $[l]$ $[r]$ のターゲットの軌道の座標を 132 msec ごとに保存する。模倣行動の正確さを確かめるため、模倣行動の再現視点の画像系列 $[L]$ $[R]$ の呈示行動のターゲットの軌道 (真の模倣行動) も記憶しておく。2.2 節の手法より既に得たエ

ピポラ幾何を用いて、呈示行動の軌道を、画像系列 $[L]$ $[R]$ に写像し模倣の目標軌道を得る。

Fig.6 は、呈示された PA10 の S 字の軌道と 2.2 節の手法より写像された軌道を示している。4 つの画面のうち上の 2 つは呈示を観察する学習者の視点の画像であり、黒いフレームは呈示の軌道を表している。下の 2 つは学習者自身を観察する視点の画像であり、黒いフレームは真の模倣行動の軌道、黒い円は写像して生成された模倣行動の目標軌道である。黒いフレームと円はそれぞれ 132 msec ごとに点が表示されており、そのシーケンスにそれぞれの軌道を見ることができる。ただし Fig.6 の S 字の軌道は以下の実験に用いられている軌道とは、同様であるが異なる軌道である。

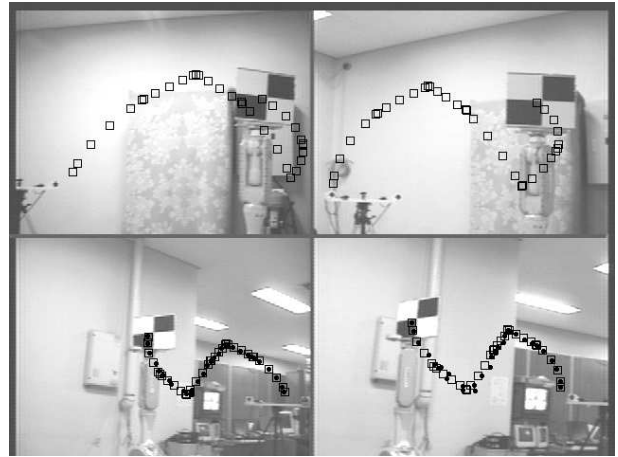


Fig.6 observed trajectory in $[l]$ and $[r]$ (upper), true trajectory(lower), and desired trajectory(lower ●) of given demonstration

3.2.3 目標軌道の実現

画像系列 $[L]$ $[R]$ のターゲットの座標を画像特徴量、PA10 の 4 つジョイントの関節角 $(\theta_i, i = 1, \dots, 4)$ をシステム記述変数、目標軌道を画像特徴量目標値として適応型ビジュアルサーボによる制御を行なう。

(7) 式の制御系の右辺第二項の冗長項に関する評価は行わず、 $k_r = 0$ とし、またコントローラからマニピュレータへの指令値の出力の遅れ時間を考慮して、フィードフォワードのゲインを下げ、制御則を以下のように改める。

$$u(k) = 0.4\hat{J}(k)^+ \{x_d(k+1) - x_d(k)\} + K\hat{J}(k)^T \{x_d(k+1) - x(k)\} \quad (8)$$

ヤコビ行列の初期値を以下のように適当に定める。

$$\hat{J}(0) = \begin{bmatrix} 10.0 & 0.0001 & 0.0001 & 10.0 \\ 0.0001 & -10.0 & 10.0 & 0.0001 \\ 10.0 & 0.0001 & 0.0001 & 10.0 \\ 0.0001 & -10.0 & 10.0 & 0.0001 \end{bmatrix}$$

ここで与えたヤコビ行列は真の値と異なるため、目標軌道を与えても始めうまく追従しない。このため模倣の目標軌道を与える前に、PA 10 に微振動を与えヤコビ行列の推定を進めておく。画像特徴量の目標軌道

を 132 msec ごとに更新して、(8) 式の制御則で制御する。

このとき得た 132 msec ごとの、学習者自身を眺める視点の画像系列 $[L]$, $[R]$ 上の手先の軌道を横軸を x 座標、縦軸を y 座標として示す (Fig.7, Fig.8)。さらに画像系列 $[L]$, $[R]$ へ写像された目標軌道、および 3.2.2 節で呈示行動を観察するときに保存しておいた画像系列 $[L]$, $[R]$ における真の模倣行動の軌道も加えて示す。

これら 3 種類の軌道は良く一致している。したがって上記の手法による、学習者自身を観察する視点の画像系列への模倣の目標軌道の写像、および適応型ビジュアルサーボによる目標軌道の実現が 3 次元空間でうまく行なわれていると言える。

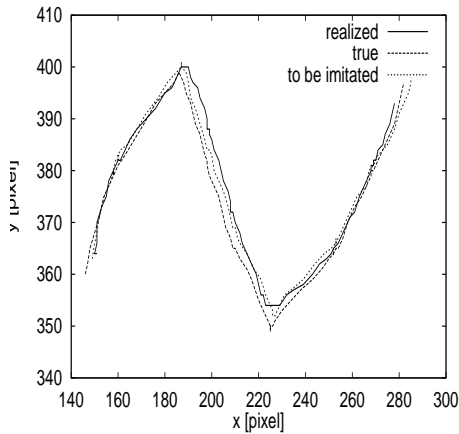


Fig.7 a trajectory in $[L]$

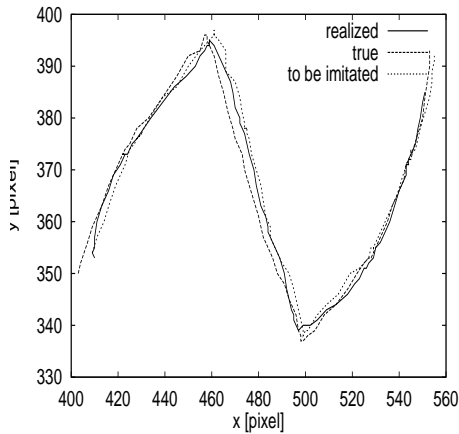


Fig.8 a trajectory in $[R]$

視点 $[L]$, $[R]$ の画面上の呈示された行動の軌道、目標軌道および適応型ビジュアルサーボにより実現した軌道上の点を 132 msec の同じタイミングで比較することにより、3 次元空間の位置に加え、速度、タイミングに関するも呈示行動を模倣できているかを見ることができる。Fig.9, Fig.10 はそれぞれの画像系列の同じ時刻での、実現した模倣行動の軌道と真の模倣行動の軌道の差 (total), および実現した模倣行動の軌道と生

成した目標軌道の差 (avs), 生成した目標軌道と真の模倣行動の軌道の差 (translation) を表したものであり、横軸は時間、縦軸は画像上の絶対距離 [pixel] である。

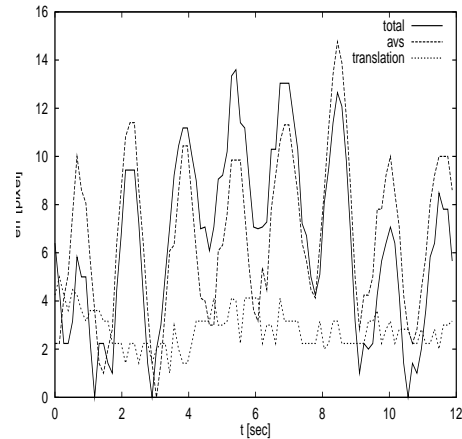


Fig.9 trajectory errors of the imitation in $[L]$

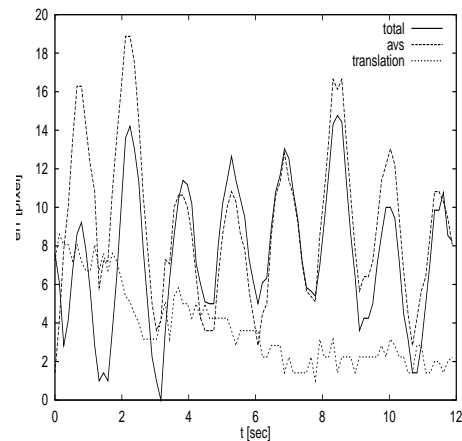


Fig.10 trajectory errors of the imitation in $[R]$

Fig.9, Fig.10 に見られる、8 個のピークは 8 組の目標関節角の PTP 制御によって呈示の軌道が与えられていることに対応する。目標軌道が進んでいくのに対し、フィードフォワード項のゲインを下げているのですぐには追従せず目標軌道に対して手先が離れていくにつれて、フィードバック項の出力がだんだん大きくなり追いつくということを繰り返していると考えられる。また PTP 制御において PA10 には比較的大きな移動量を与えられている。大きな移動ではヤコビ行列は変動するので、推定が不十分なときに追従の失敗が起っている可能性も考えられる。従って速度、タイミングという点でみれば、この手法によって呈示された行動を完全には模倣していないと言える。

それぞれのサンプリングのタイミングでの模倣行動の目標軌道の、真の模倣行動に対する誤差 (写像誤差, trans), 適応型ビジュアルサーボで実現した軌道の、目標軌道に対する誤差 (制御誤差, avs), 適応型ビジュ

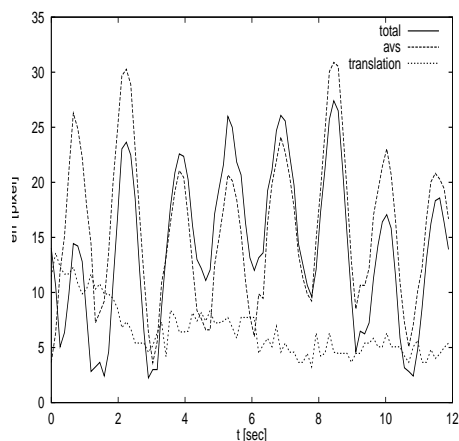


Fig.11 sum of trajectory errors of the imitation in [L] and [R]

アルサーボで実現した軌道の、真の模倣行動に対する誤差 (total) の平均値と標準偏差を示す (Table.2) . また画像系列 [L], [R] の軌道の各誤差を単純に足し合わせたものを Fig.11 に示す . 横軸が時間 [sec] , 縦軸が誤差 [pixel] を表す .

これらによると、実現した軌道の真の模倣行動に対する誤差が制御誤差よりも小さくなっている、あるいは写像誤差と制御誤差の単純な和よりも小さくなっていることが分かる . 写像誤差を含む軌道は実現する際に矛盾した目標値になっているため、適応型ビジュアルサーボによって 3 次元空間上で完全に追従することができない . これは制御誤差に含まれる . 実現できる範囲で目標軌道に追従した結果、写像誤差の矛盾を打ち消すような軌道が生成され、目標軌道の矛盾のため、実現した軌道の制御誤差は大きくなるが、真の軌道に対する誤差が小さくなるということが起きていると考えられる . 2.2 節の目標軌道の生成と適応型ビジュアルサーボによる目標軌道の実現が補い合う形で現れていると言える .

Table 2 an average and a standard of trajectory errors in [L] and [R]

	[L]	[R]
写像誤差	2.80 (0.80)	3.71 (2.01)
制御誤差	6.69 (3.24)	9.04 (3.95)
真の模倣行動と実現された 模倣行動との誤差	6.45 (3.62)	7.49 (3.54)

4 おわりに

本論文では、身体構造の同様なロボットの呈示した行動を視覚により捉え、エピソード拘束を利用したメカニズムにより知覚可能なセンサ空間上に目標軌道を生成し、適応型ビジュアルサーボを用いてその目標軌

道を実現することにより、呈示された行動の模倣を実現する手法を提案し、この手法により、3次元空間の位置に関して、真の模倣の軌道と良く一致する軌道となる模倣行動が実現可能であることを示した . しかし、生成した模倣の軌道と真の模倣の軌道を時系列で見ると、その速度変化は同一ではなかった . 行動をする際に、速度には修飾的な内容 (急いで、力強く等) などが含まれるので、これを理解するために時系列での軌道の一致が今後の課題となる .

またエピソード拘束を利用したメカニズムにより生成した目標軌道は、3次元再構成を含まない、光幾何学に基づいた単なる座標変換であり、行動の理解が単に画像上の軌道として行なわれるともみなせる . しかし動き発生メカニズムを用い、自分で再現してみることによって、視覚によって軌道という形でしか理解できなかった行動を、モータコマンド (関節速度) という形の表現に変えることができたといえる . 今後は、行動を観察し理解するために、モータコマンド (あるいは他の形で) 自分の行動レパートリを作成すること、レパートリと観察した行動のマッチングによる観察行動の理解、が課題になると考えられる .

参考文献

- [1] Giacomo Rizzolatti and Micheal A Arbib. Language within our grasp. *TINS*, Vol. 21, No. 5, pp. 188-194, 1998.
- [2] Stefan Schaal. Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, 1999.
- [3] 浅田稔, 石黒浩, 國吉康夫. 認知ロボティクスの目指すもの. 日本ロボット学会誌, Vol. 17, No. 1, pp. 2-6, 1999.
- [4] H. C. Longuet-Higgins. A computational algorithm for reconstructing a scene from two projections. *Nature*, Vol. 293, pp. 133-135, 1981.
- [5] 細田耕, 浅田稔. 構造やパラメータに関する先験的な知識を必要としないフィードフォワード補償器を持つ適応型ビジュアルサーボ系の構成. 日本ロボット学会誌, Vol. 14, No. 2, pp. 313-319, 1996.