

回転不変特性を持つステレオカメラを用いた観察による模倣

辻義樹(大阪大) 吉川雄一郎(大阪大) 浅田稔(大阪大) 細田耕(大阪大)

View-based Imitation with Rotation Invariant Pan-Tilt Stereo Cameras

*Yoshiki TSUJI(Osaka University), Yuichiro YOSHIKAWA(Osaka University),
Minoru ASADA(Osaka University), and Koh HOSODA(Osaka University)

Abstract— In the previous work, the method for visual imitation by recovering the demonstrator’s view based on the stereo epipolar constraint has been developed. The method is applied to the stationary pair of the stereo cameras, therefore, the visual field is limited. This paper presents a method to extend the previous work by adopting a pair of rotation invariant stereo cameras that has pan and tilt motions without changing the optical center, therefore, the stereo epipolar equation does not change. The spherical projection is used to represent the constraint. The experimental results are shown.

Key Words: imitation learning, stereo epipolar constraint, rotation invariant

1. はじめに

ロボットにおける模倣学習の研究は、多自由度、多感覚を有するロボットの行動学習を加速する手段として¹⁾、また人間の知能の構成論的なモデル化の研究題材として²⁾、近年注目を集めている。従来研究では、運動の呈示者視点の情報(関節角や手先の軌道)が得られると仮定し、それをいかに習熟あるいは内在化するか、という問題に焦点を当てたものが多い。それらでは呈示者視点の情報を得るために、高精度な3次元観察³⁾や明示的な座標変換⁴⁾などを行っており、それらには設計者によるキャリブレーションが必要となる。しかし、自律的なヒューマノイドロボットを考える場合には、ロボット自身がセンサ情報に基づいて呈示者視点の情報を獲得するしくみを考える必要がある。

吉川ら⁵⁾は、呈示者視野復元のための視野変換と復元された呈示者の軌道を再現するための適応型ビジュアルサーボ⁶⁾からなる観察による模倣の手法を提案した。しかし、この手法ではエピポーラ幾何と呼ばれる観察画像間の幾何学的な拘束を用いているため、呈示者の運動を観察するときに学習者はカメラを動かすことができず、呈示者や学習者の運動が制限されてしまう。そこで本論文では、吉川ら⁵⁾の手法を拡張し、呈示者や学習者の運動を捉えるためにカメラを動かす必要がある場合にも適用可能な手法を提案する。光学中心が固定であればカメラが回転してもエピポーラ幾何は不変であることを利用するため、光学中心を変化させずにパン、チルト方向の回転が可能なステレオカメラを用い、回転しても不変となるエピポーラ幾何の表現方法として球面射影を用いて呈示者視野復元を行う。

2. 球面画像間の視野変換

2.1 エピポーラ幾何

光学中心を中心とする球の表面上の画像を球面画像と呼び、3次元空間上の注視点を、それと光学中心とを結ぶ直線と球面画像との交点へ射影することを球面射影と呼ぶ。ここで、球面射影カメラのカメラ座標系は3次元座標系である。

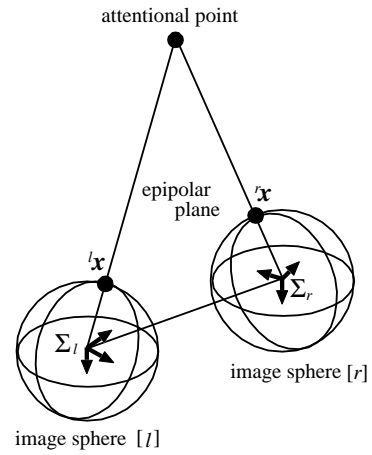


Fig.1 Epipolar geometry between two image spheres.

3次元空間上の注視点が2つの球面画像へ射影されたとき、3次元空間上の注視点と射影された点、および光学中心は一つの平面上に存在し、この平面はエピポーラ平面と呼ばれる(Fig.1参照)。また、この幾何学的な拘束はエピポーラ幾何と呼ばれ、式(1)のエピポーラ方程式を満たす。ここで、 ${}^l x$ 、 ${}^r x$ はそれぞれ球面画像 $[l]$ 、 $[r]$ に射影された点の座標であり、 ${}^l E \in \mathbb{R}^{3 \times 3}$ は基本行列と呼ばれ、二つのカメラ座標系の幾何学的な関係により決定される。

$${}^l x^T {}^l E {}^r x = 0 \quad (1)$$

対応する射影点が8点以上与えられれば、式(1)の連立方程式を最小自乗法に基づいて解くことにより基本行列を推定することができる⁷⁾。

2.2 エピポーラ幾何に基づく視野変換

球面画像 $[l]$ 、 $[r]$ で観察する点と同じ点を観察しているもう一つの球面画像 $[L_D]$ 、 $[R_D]$ を考える(Fig.2参照)。ここで、球面画像 $[l]$ 、 $[r]$ での点に対応する点を球

面画像 $[L_D]([R_D])$ でどのようにして見つけるかが問題となる。

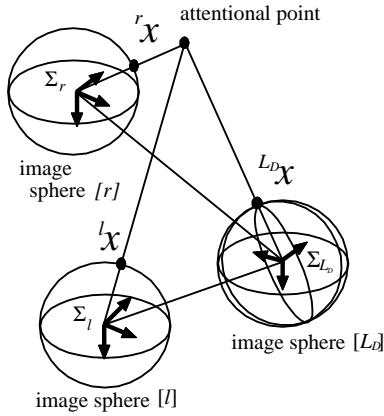


Fig.2 The mechanism of the view transformation based on epipolar geometry.

エピポーラ幾何により, l_x に対応する点 $L_D x$ ($R_D x$) はエピポーラ平面上, さらに球面画像 $[L_D]([R_D])$ に拘束される。つまり, $L_D x$ ($R_D x$) はエピポーラ平面と球面画像 $[L_D]([R_D])$ との交わる面上の大円上に拘束される。同様に, 球面画像 $[r], [L_D]([R_D])$ との関係から $L_D x$ ($R_D x$) は別のもう一つの大円上にも拘束され, この二つの大円の交点を求めることにより $L_D x$ ($R_D x$) が求まる。

実際には, 式 (2) ~ 式 (4) の連立方程式を解くことにより $L_D x$ ($R_D x$) を決定する。

$$l_x^T l L_D E L_D x = 0 \quad (2)$$

$$r_x^T l L_D E L_D x = 0 \quad (3)$$

$$L_D x^T L_D x = c^2 \quad (4)$$

式 (2), (3) は $L_D x$ がエピポーラ平面上に拘束されていることを, 式 (4) は $L_D x$ が球面画像上 $[L_D]$ にあることを表しており, 任意の正の定数 c は球面画像 $[L_D]$ の半径を表す。 $R_D x$ も同様にして決定することができる。

3. 呈示者視野の復元⁵⁾

視野変換を用いて呈示者の運動を模倣するための呈示者視野の復元方法について述べる。

呈示者と学習者の身体構造 (リンクパラメータとカメラパラメータ) が同じであると仮定すると, 呈示者の視野で捉えられる呈示者自身の運動の軌道は, 学習者が正確に呈示運動を再現したときに学習者自身を観察する視野で捉える運動の軌道と等しいとみなせる。したがって, 呈示者の視野を復元し, それを目標軌道として呈示運動を再現することで模倣が実現可能である。

Fig.3 に呈示者の視野と学習者の視野の関係を示す。 V_O^A は, エージェント A の運動を観察するエージェント O の視野を表す (A, O : 呈示者または学習者)。学習者と呈示者の身体構造が同じであるので, 学習者は $V_D^D([L_D], [R_D])$ における呈示運動の軌道を得ること

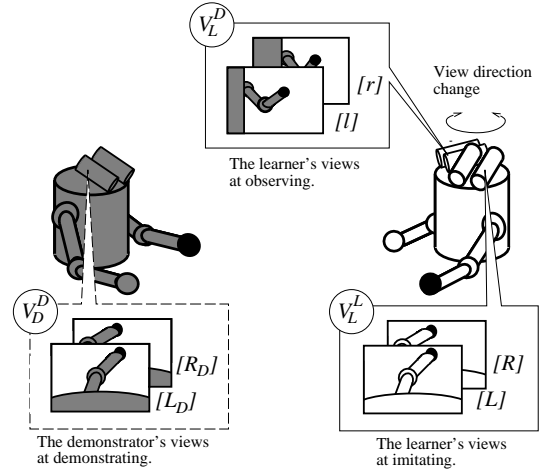


Fig.3 Demonstrator's view recovery.

ができれば, その軌道を再現することで模倣が実現可能である。しかし, 学習者は $V_D^D([L_D], [R_D])$ を観察することができないので, 学習者が観察可能な $V_L^L([l], [r])$ から $V_D^D([L_D], [R_D])$ を復元する必要がある。もし, V_L^L, V_D^D 間のエピポーラ幾何の基本行列が既知であるならば, 学習者は前節で述べた視野変換により V_D^D における呈示運動の軌道を獲得でき, 適応型ビジュアルサーボ⁶⁾ によってその軌道を追従することで模倣が実現可能である。

3.1 基本行列の推定

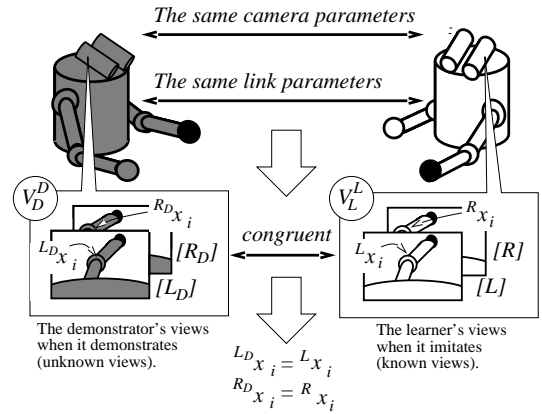


Fig.4 The key idea of the method of the parameter estimation.

呈示者視野復元のため, 学習者は V_L^L, V_D^D 間のエピポーラ幾何の基本行列, ${}^l L_D E, {}^r L_D E, {}^l R_D E, {}^r R_D E$ を推定する必要がある。前節で述べたように, 基本行列の推定には 8 点以上の対応点が必要である⁷⁾ が, 学習者は V_D^D における対応点を直接知ることはできない。しかし, 呈示者の初期姿勢 (関節角) と学習者の初期姿勢が同じであると仮定すると, 学習者と呈示者の身体構造は同じであるので, V_L^L 上に射影された学習者自身の身体と V_D^D 上に射影された呈示者自身の身体は全く同じになる (Fig.4 参照)。つまり, 学習者が観察できな

い視野 V_D^D における点を使う代わりに，学習者が観察可能な視野 V_L^L における対応点を使うことで基本行列を推定できる．

V_L^L における学習者の身体の i 番目の射影点を Lx_i , Rx_i , それに対応する V_D^D における呈示者の身体上の点を ${}^{LD}x_i$, ${}^{RD}x_i$ とすると，それらは以下の式 (5) を満たす．これは，学習者は基本行列を推定するのに ${}^{LD}x_i$, ${}^{RD}x_i$ の代わりに Lx_i , Rx_i が使えることを意味する．

$$\begin{aligned} {}^{LD}x_i &= {}^Lx_i \\ {}^{RD}x_i &= {}^Rx_i \end{aligned} \quad (5)$$

初期姿勢が同じであるという仮定をなくすために，そのとき推定した E 行列を使って視野変換を行った結果の誤差を最小にするように関節角を制御することでより正確に近い基本行列を推定する手法⁸⁾が提案されている．

4. 実験

提案手法により呈示者の視野復元が可能であるかを検証するために実ロボットを使って実験を行った結果を示す．本実験では，2 台の同じマニピュレータをそれぞれ学習者および呈示者の身体と仮定している (Fig.5) ．

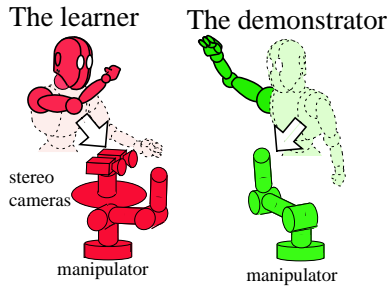


Fig.5 The learner and the demonstrator have identical manipulators as bodies.

4.1 回転不変特性を持つカメラヘッド

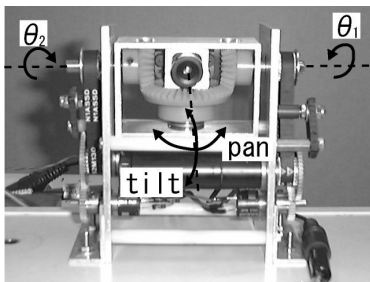


Fig.6 The pan-tilt camera.

カメラの回転により光学中心の位置が変化しないという特性を回転不変特性と呼ぶ．本実験では，回転不変特性を持つ 3 つのはずば歯車からなるカメラヘッドを用いた (Fig.6) ．まず，2 つのモータのトルクはベルトを通して上部のはずば歯車に伝達される．それから，

CCD カメラの台になっている下部のはずば歯車へ伝達され CCD カメラが動く．カメラ台の回転中心は 3 つのはずば歯車の回転軸の交点であり，その回転中心と CCD カメラの光学中心が一致するように設計されている．回転角 ($pan, tilt$) は以下の式 (6), (7) によって求められる．ここで， θ_1, θ_2 は上部のはずば歯車の回転角度， r, R はそれぞれ上部のはずば歯車，下部のはずば歯車の半径である．

$$pan = r/R \cdot (\theta_1 - \theta_2) \quad (6)$$

$$tilt = \theta_1 + \theta_2 \quad (7)$$

このカメラヘッドはカメラ台の回転中心と CCD カメラの光学中心が一致するように設計されているが，现阶段ではカメラを手で台に取り付けるため回転が不変である保証はない．よって今後，カメラの位置を調節できるようにする必要がある．

4.2 実験装置

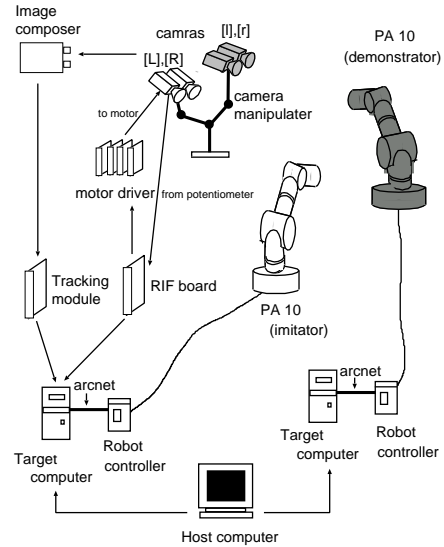


Fig.7 Experimental setup

実験装置の詳細を Fig.7 に示す．学習者と呈示者の身体として 2 台の同じな 7 自由度マニピュレータ (PA10, MHI) を用い，学習者の視覚として上で述べた 2 台のパン・チルトカメラヘッドを用いた．今回の実験では，マニピュレータの 3 自由度のみを利用し，残りは固定した．

2 台の CCD カメラから入力された画像は，それぞれ多重変換装置により縦方向に半分に圧縮され (画像サイズ: 640[pixel] × 240[pixel])，1 つの画像 (画像サイズ: 640[pixel] × 480[pixel]) に結合された後に，富士通製のトラッキングモジュールに送られる．マニピュレータにはトラッキングのためのターゲットを付けた．CPU (VxWorks, WindRiver) は計算を行い，マニピュレータとカメラヘッドのコントローラに制御信号を出力する．

4.3 球面画像の射影

球面画像を利用するために、学習者は平面画像、球面画像間のマッピングを行う必要がある。光学中心が回転中心と一致しているため、カメラ台の角度 (*pan, tilt*) は極座標系の角度と等しい。よって、平面画像上の注視点をカメラ台を回転させることで平面画像の中心に持っていくという経験を通してマッピングを獲得することができる。

3次元空間上の特徴点 100 点を学習者に注視させ、得られたデータをもとに、3つのユニットを持つ1つの隠れ層からなるフィードフォワードニューラルネットワークの逆誤差伝播法による学習を行うことで、平面画像上に射影された特徴点の座標からカメラ台の角度へのマッピングを学習した。マッピングの自乗誤差が十分に小さくなっていることから (Fig.8), 学習者はマッピングを学習により獲得できたといえる。

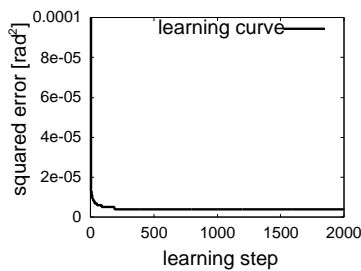


Fig.8 Learning curve : squared error of the mapping from image plane to image sphere.

4.4 呈示者視野において観察された呈示運動の復元

提案手法の正当性を検証するために、呈示者に呈示運動としてエンドエフェクタにより三角形を描かせ、学習者に呈示者視野の観察軌道を学習者視野における軌道へと復元させた。3次元空間上の 79 点を与え、基本行列の推定を行った結果の残差を Tab.1 に示す。基本行列推定の残差は、画像の量子化による誤差やマッピングの学習時の誤差、そしてカメラヘッドの回転不変特性が保証されていないことによる誤差などに起因していると考えられる。

Table 1 The residuals in the estimation of the essential matrices in the real robot experiment.

essential matrices	residual of estimation
$lL_D E$	1.1×10^{-3}
$rL_D E$	8.2×10^{-4}
$lR_D E$	1.0×10^{-3}
$rR_D E$	5.8×10^{-4}

また、呈示者のエンドエフェクタの軌道を復元した結果を Fig.9 に示す。ここで、Fig.9 は球面画像 $[R_D]$ 上での復元点 (dots) と真の点 (cubes) を表している。それらはほぼ同じであるので、視野変換により呈示者視野上の対応点が正しく推定されているといえる。

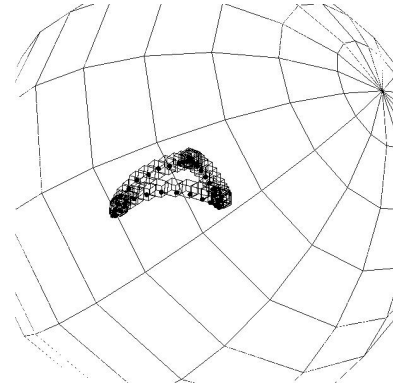


Fig.9 The trajectory recovered by the view transformation and the true trajectory measured by the designer in advance.

5. おわりに

本論文では、ロボットにおける模倣のためのエピポラ幾何による呈示者視野復元に基づく手法を拡張し、カメラが光学中心回りに回転した場合にも適用可能な手法を提案した。学習者のステレオカメラの光学中心が回転不変特性をもつと仮定することで、呈示者視野と学習者視野間のエピポラ幾何に基づく視野変換により呈示者視野復元が可能になる。これにより、これまでの手法とは異なり学習者は観察範囲を広げるためにカメラを動かすことができる。また、実ロボットによる実験により提案手法の正当性を確認した。回転不変特性を保証するため、カメラの取り付け位置を調節する機構をカメラヘッドに組み込むといったシステムの改良が必要である。

参考文献

- 1) S. Schaal: Is imitation learning the route to humanoid robot?, Trends in Cognitive Sciences, 3, 6, pp.233-242(1999).
- 2) 浅田稔, 石黒浩, 國吉康夫: 認知ロボティクスを目指すもの, 日本ロボット学会誌, 17, 1, pp.2-6(1999).
- 3) 戸嶋麻樹, 稲邑哲也, 中村仁彦: 力学的身体性を用いたミメシスの閉ループ化, ロボティクス・メカトロニクス講演会 '01 講演論文集.
- 4) 宮本弘之, 川人光男: 作業レベルのロボット学習のための見まねによる教示, 電子情報通信学会論文誌, J81-D-II, 10, 2401-2410(1998).
- 5) 吉川雄一郎, 浅田稔: 3次元再構成を伴わない観察による他者行動の模倣, 第5回ロボティクスシンポジウム予稿集, pp.158-163(2000).
- 6) 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, 14, 2, pp.313-319(1996).
- 7) H. C. Longuet-Higgins: A computer algorithm for reconstructing a scene from two projections, Nature, 293, pp.133-135(1981).
- 8) 吉川雄一郎, 浅田稔, 細田耕: 幾何学的な整合性の追及に基づく呈示者視野の復元と姿勢の模倣, 第19回日本ロボット学会学術講演会講演論文集, pp.401-402(2001).