

単眼カメラのビデオ画像に基づく動作再現を目指したビジュアルサーボの応用に関する基礎検討

○李 堯希 藤本 博志 (東京大学)

1. 研究目的

イメージセンサの小型化, 高性能化や計算機の速度向上に伴い, イメージセンサはロボットが環境を検知する手法として広く用いられるようになった。

イメージセンサから取得した画像を用いてロボットを制御する研究は多く行われており, 例えば, ロボットの位置や経路を画像で与え, ビジュアルサーボのアプローチで移動ロボットの経路案内 [1] やロボットアームの動作を再現する研究 [2][3] 等が行われている。

画像を用いた経路案内では目標軌道を通った場合のビデオ画像の各フレームを目標として, それらの画像群に対して現在の画像を収束させることで結果的に軌道への追従を行うことができる [1] が, 単に目標画像に逐次的に収束させるだけでは動作時間のスケールを合わせられない問題点がある。

本稿においては, 図 1 に示すようにカメラの撮影した時系列の画像に対して現在の画像を同じスケールの時系列で合致させることを目標とする。この目標が達成された場合, ロボットは参照した動画像を取得した時と同一の動作を行うことになるため, 先ほどの経路のみを再現するものと区別するため, 以降, 動作再現と呼称する。

ビデオの動画像が指令値として与えられた場合, 動作再現を行うには, 目標値の定義の仕方と与え方, 目標値へと追従させる制御則の 2 つについて考える必要がある。

本稿では, 制御に用いる時系列の目標値をある特定のフレームに対する変位を用いて数値で定義する手法とその検出方法, 時不変のイメージャコビアンとそのリアルタイム推定を用いたフィードフォワード補償を用いたビジュアルサーボ則の 2 点について説明を行う。

2. 動作再現における目標値の定義方法

動作再現において, 経路を参照するための動画像は予め与えられているものとする。

カメラを用いてロボットを案内する際, 目標とする指令値をどのように与えるのかが 1 つの問題となる。

先行研究 [1] のように画像そのものを目標として用いた場合, 事前の特徴抽出などを行わない手軽さや, 特徴抽出等を行う際の情報の欠如がないことが利点として挙げられる。一方で, 画像から抽出したカメラの動作に関わる何らかの量, 例えば特徴点の座標などを予め画像群から抽出して用いることで経路をその値の軌道として扱うことができる [2]。この手法は直接保存しておくべき画像の量が減るので省メモリな手法であると言える。

画像からどのような量を抽出するのが良いかは対象の形状や想定する動きによって異なってくるが, 簡単

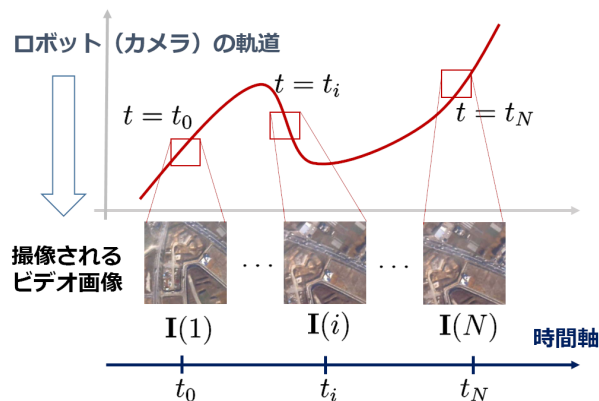


図 1 ロボットの時系列の動作が記録されたビデオ画像に対して時系列でカメラの取得画像を合わせることで動作再現を目指す

のため, [5] と同様な対象画像が平面かつカメラ平面に並行で, カメラの並進, 距離方向への移動とそのまわりの回転の 4 自由度である場合について記述する。

本稿で提案する画像の目標軌道の作り方を図 2 に示す。連続する動画像をフレームごとに切り出し, 動画の始まりのフレームを $\mathbf{I}(1)$ とし, 以降順番に各フレームの画像を $\mathbf{I}(n)$ とする。これら画像群は時系列にならんでおり, $t = t_i$ のタイミングで画像 $\mathbf{I}(i)$ をとるといったように定義される。

次に, 画像同士の変位を ξ と定義する。例えば, 今回の仮定において, ξ は画像の平行移動量 (ξ_x, ξ_y) , 回転 θ , 拡大縮小量 κ を要素にもつ 4 次元のベクトルである。この ξ を用いて目標画像群 $\mathbf{I}(i)$ ($i = 1, 2, 3, \dots, N$) を置き換えていく手法を次に述べる。

手順 1 ビデオの最初のフレーム $\mathbf{I}(1)$ を参照画像とし, k 番目のフレーム $\mathbf{I}(k)$ の $\mathbf{I}(1)$ からの変位を $\xi_1(k)$ とする。下付きの添字は何番目の画像を参照にしているかということを表す。

手順 2 $i+1$ 番目のフレームにて現在の参照画像からの変化が十分大きいと判断された場合, 参照画像を i 番目の画像に切り替え, 表記は $\xi_i(i+1)$ というように下付きの添字を変更する。これを全てのフレームの画像を変換するまで続ける。

注意として, 新たな参照画像は必ず既存の参照画像に対して対応付けが取れるようでは必要ではない。このことは後に指令値の微分を見るときや値の推定において連続性を保つために必要である。

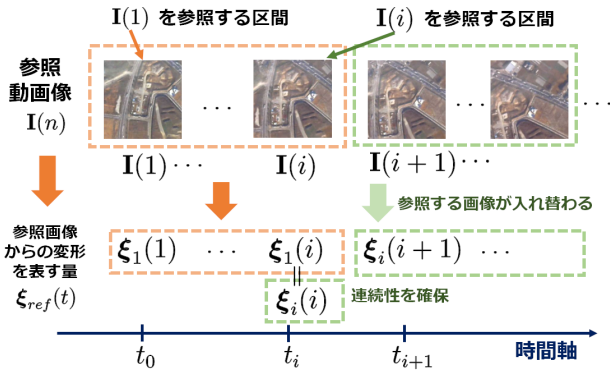


図2 参照画像群からパラメータへと変換する様子

3. 動作再現のための制御則

3.1 イメージベースビジュアルサーボ [4]

本稿ではイメージベースのビジュアルサーボの手法に倣って制御則を決定する。

ビジュアルサーボは3次元を再構成する位置ベースの手法と2次元画像を指令値に用いるイメージベースの手法とに大きく分けられるが、画像内の特徴量そのものを制御の偏差としてフィードバックするイメージベースの手法は3次元再構成の誤差にとらわれない。

3次元におけるカメラ位置を \mathbf{X} とし、それに応じてカメラから取得できる画像内特徴量を ξ と定義する。例えば、イメージベースのビジュアルサーボでは画像内特徴量として特徴点の座標 (x_i, y_i) を用いて $\xi = [x_1, y_1, \dots, x_n, y_n]^T$ などと取る手法が広く知られている。その場合、物体が静止している場合において3次元のカメラの速度 $\mathbf{V} = [v_x, v_y, v_z, \omega_p, \omega_r, \omega_y]^T$ と画像内特徴量の時間変化分 $\dot{\xi}$ とは一般に次のように関係づけられる。

$$\mathbf{J}\mathbf{V} = \dot{\xi} \quad (1)$$

ここで \mathbf{J} は画像ヤコビ行列、またはイメージヤコビアンと呼ばれる行列であり、一般に \mathbf{X} と ξ の関数で表される。例えば、先の例において i 番目の特徴点に対応する \mathbf{J}_i は

$$\mathbf{J}_i = \begin{bmatrix} -\frac{1}{Z_i} & 0 & \frac{x_i}{Z_i} & x_i y_i & -(1+x_i^2) & y_i \\ 0 & -\frac{1}{Z_i} & \frac{y_i}{Z_i} & 1+y_i^2 & -x_i y_i & -x_i \end{bmatrix} \quad (2)$$

のように表せる。 Z_i は各特徴点へのカメラからの距離を表し、時々刻々と変化するこの値を通常は単眼のカメラから得るのは困難である。

そのため、イメージベースの手法では、 Z_i を適当に近似したイメージヤコビアン \mathbf{J}_0 を用いて

$$\mathbf{V}_{ref} = -\lambda \mathbf{J}_0^+ (\xi_{ref} - \xi_c) \quad (3)$$

とすることで、誤差を逐次的に収束させる。 λ は収束の早さを決定するゲインである。

3.2 イメージヤコビアンの時不変化 [5]

本稿においては画像全体の平行移動量、拡大縮小量、回転量の4つのパラメータを画像内特徴量とする。ここで、対象が平面かつカメラ平面に平行に配置されて

いる系において、並進3自由度とz軸回転の合わせて4自由度の動きしか無い場合において、画像ヤコビ行列は参照画像周りで時不変の4行4列の行列にすることができる。詳細な導出は参考文献 [5] に譲り、以下では簡潔に説明する。

用いるパラメータは、現在の画像 \mathbf{I} の座標 $\mathbf{x} = (u, v)^T$ のテンプレートの画像 \mathbf{I}^* の座標 $\mathbf{x}^* = (u^*, v^*)^T$ に対する、平行移動量 $(\xi_x, \xi_y)^T$ 、回転 θ 、拡大縮小量の逆数 κ とする。拡大縮小量はイメージヤコビアン時不変化のために、通常のスケーリングの逆数をとる。すなわち、 $scaling^{-1} = \kappa$ である。重要なのは平行移動量の定義であり、平行移動成分 $\mathbf{t} = (\xi_x, \xi_y)^T$ と置いた時の定義を式 (4) に示す。

$$\mathbf{x} = \mathbf{A}(\mathbf{x}^* + \mathbf{t}) \quad (4)$$

ここで、角度 θ の2次元回転行列 $\mathbf{R}(\theta)$ にスケーリングをかけた行列を $\mathbf{A} = scaling \mathbf{R}(\theta)$ としている。

これはすなわち、平行移動量を目標画像の座標系 $\mathbf{x}^* = (u^*, v^*)^T$ の上で定義していることになり、Z方向の変化や現在の回転の変化に対して完全に独立して平行移動量を扱うことが出来るということを示している。

最終的に、式 (4) で定義された平行移動量、回転、拡大縮小量を用いて画像内の変形パラメータ ξ を式 (5) のように定義することでイメージヤコビアンを固定値の関数にすることができる。

$$\xi = [\xi_x, \xi_y, 1/\kappa, \theta]^T \quad (5)$$

この時のイメージヤコビアン \mathbf{J}_{const} は次のように表せる。

$$\mathbf{J}_{const} = \begin{bmatrix} \frac{f}{Z_0} \cos \Theta_0 & \frac{f}{Z_0} \sin \Theta_0 & 0 & 0 \\ -\frac{f}{Z_0} \sin \Theta_0 & \frac{f}{Z_0} \cos \Theta_0 & 0 & 0 \\ 0 & 0 & \frac{1}{Z_0} & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (6)$$

ここで、 Θ_0 と Z_0 はそれぞれ目標画像を取得した際のカメラの角度とカメラと対象の距離を表し、前者は検出した角度変位と現在の角度から計算可能であるが、後者はロボットの動きと画像内特徴量の変化から推定することができる。その手法は後の章で述べる。

3.3 動作再現における制御則

動作再現を行うにあたって問題設定を、時々刻々と変化する画像内特徴量指令値 $\xi_{ref}(t)$ に対して現在の画像内特徴量 $\xi_c(t)$ を一致させるように制御すること定義する。

その際用いられる関係式はイメージベースビジュアルサーボの時と同様に式 (7) のようになる。

$$\mathbf{J}(\mathbf{X}, \xi) \mathbf{X}(t) = \dot{\xi}(t) \quad (7)$$

先ほど導出した関係式や式 (6) の \mathbf{J}_{const} を用いて考えると、最も単純な制御則は式 (3) を見習って式 (8) のようになる。

$$\mathbf{V}_{ref}(t) = -\lambda \mathbf{J}_{const}^{-1} (\xi_{ref}(t) - \xi_c(t)) \quad (8)$$

式 (8) の右辺は指令値の変化に対してフィードバック的に動作するため、固定位置に対しては収束可能だが、時系列で変化する軌道に対して完全には収束しない。このことは後のシミュレーションにおいて実際に示す。

次に、本稿における提案手法を式 (9) に示す。

$$V_{ref}(t) = -\lambda J_{const}^{-1} (\xi_{ref}(t) - \xi_c(t)) + J_{const}^{-1} \frac{d}{dt} \xi_{ref}(t) \quad (9)$$

式 (9) の第 2 項は軌道の情報を先見的に与えるフィードフォワード項であり、軌道の時間変化を補償する一方、第 1 項はフィードバック的に位置の偏差やフィードフォワード項の誤差を校正するように動く。

この式に関連して、近似的に固定したイメージャコピアンを用いて、画像の軌道の偏差の情報をフィードフォワードしようとする手法 [2] やそれを最適制御の問題に帰着した研究 [3] がなされてきた。

本稿における独自性は近似ではなく、特定の条件下でイメージャコピアンを時不変を用いたことにある。

この手法を用いる前提として、第 2 項として加えられたフィードフォワード項におけるイメージャコピアン J_{const} を正確に推定する必要がある。なお、第 1 項目のイメージャコピアン J_{const} はそれと異なり正確に推定しなくとも良いため、式 (9) では敢えて式を 2 項に分けている。

4. 目標位置における距離 Z_0 の推定手法

2 章で述べた指令値の連続性から、目標画像が変化した場合も、前回の距離の推定値と画像から得たスケールリングを用いて滑らかに距離 Z_0 をつなぐことができるため、ここでは式 (6) にて用いられる、最初の目標画像 $I(1)$ における距離 Z_0 を推定する。

4.1 手法 1：拡大縮小量を用いた距離推定

拡大縮小量の逆数 κ は距離 Z と 1 対 1 で対応するため、カメラの Z 軸方向の運動と拡大縮小量の時間変化分の関係を用いて目標画像における距離 Z_0 を推定することができる。式 (10) に示す関係から逐次最小二乗法を用いて推定を行った。

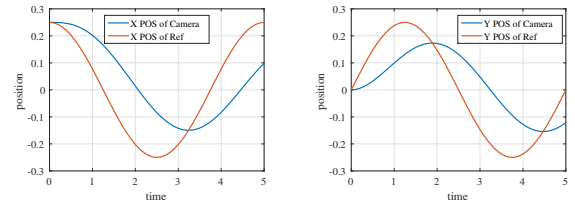
$$\Delta\kappa = \frac{1}{Z_0} \Delta Z \quad (10)$$

ここで Δ は該当する変数の 1 ステップでの変化分を示す。

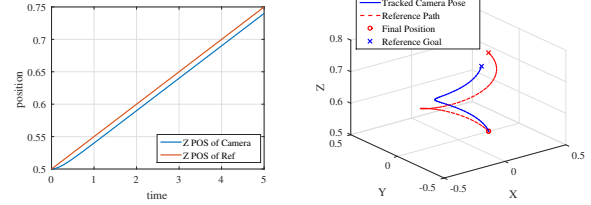
ただし、カメラの光軸が対象平面に対して垂直でない時、すなわち画像平面と対象が平行でない時、カメラの光軸周りの回転が誤差として非常に大きく寄与してくることに留意する必要がある。また、 κ は通常 1 前後の値であるため、距離方向の移動分が小さい場合もノイズに埋もれやすい傾向がある。

4.2 手法 2：並進量を用いた距離推定

先ほどの拡大縮小量から直接対象との距離を計測する手法は距離 Z が変化しない場合には用いることができない。そのため、x,y 方向の変位から距離を推定することを考える。代用として、画像平面上での移動距離と 3 次元空間における対象平面に対して平行移動した



(a) 螺旋軌道を描いた時の X 方向の位置指令とカメラの動き (b) 螺旋軌道を描いた時の Y 方向の位置指令とカメラの動き



(c) 螺旋軌道を描いた時の Z 方向の位置指令とカメラの動き (d) 目標軌道 (赤点線) とカメラの通った軌道 (青線)

図 3 FF 項を入れない場合の軌道と時系列の目標に対する応答

距離は、焦点距離と実際の距離の比として表せるため、次のような式になる。

$$\sqrt{\Delta\xi_x^2 + \Delta\xi_y^2} = \frac{f}{Z_0} \sqrt{\Delta X^2 + \Delta Y^2} \quad (11)$$

この式にも同様に逐次最小二乗法を当てはめることで、カメラと対象の距離と焦点距離の比を測ることができ。焦点距離を事前に求めておくことによってカメラと対象の距離をここから求めることが可能である。

5. シミュレーション

提案手法の有効性を確かめるためにシミュレーションを行った。目標とする画像内特徴量の指令値 $\xi_{ref}(t)$ は予め全て 1 つの目標画像に対して記述されているとし、現在の $\xi_c(t)$ を検出した。

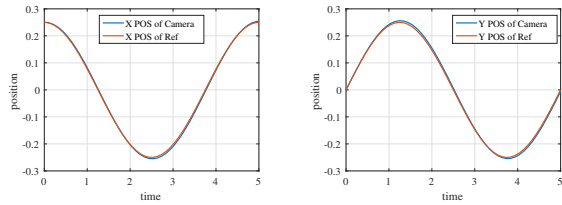
目標とするカメラの動きは X,Y 平面で円を描きつつ Z 方向に等速で直進する螺旋状の運動とし、ロボットは生成される速度指令値に対して遅れなしで追従するものとする。図 3 に示すフィードバックのみの場合と図 4 に示す提案手法との比較から複雑な軌道に対しては予見的なアプローチが有効かつ必要であることがわかる。

6. 予備実験

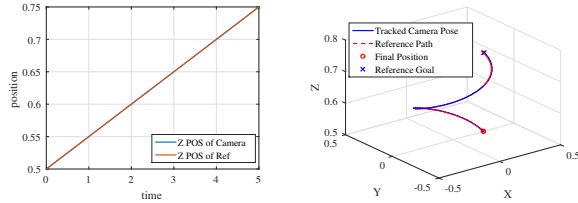
実験の前段階として、各々の章で提案した内容の検証実験を行った。RTLinux を用いて駆動する安川電機のロボットアーム開発環境 RTLab を使い、別の PC を用いてアーム上のカメラ画像を解析、画像処理の結果をソケット通信で送受信している。

6.1 動画像からの目標値の生成

動画像からの目標値の生成において変位の抽出手法として文献 [5] でも用いられていた位相限定相関法 [6] を用いた。変位と同時に観測できる類似度の尺度であるピーク値を用いて切り替え条件を設定し、例として生成された目標軌道は図 5 のようになる。



(a) 螺旋軌道を描いた時の X 方向の位置指令とカメラの動き (b) 螺旋軌道を描いた時の Y 方向の位置指令とカメラの動き



(c) 螺旋軌道を描いた時の Z 方向の位置指令とカメラの動き (d) 目標軌道 (赤点線) とカメラの通った軌道 (青線)

図 4 FF 項を入れた場合の軌道と時系列の目標に対する応答

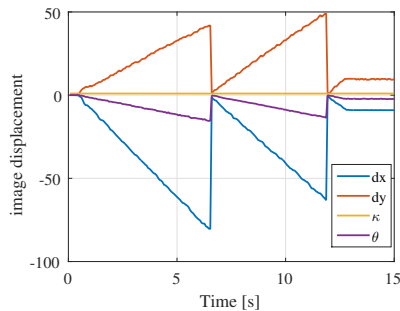


図 5 目標値 $\xi_{ref}(t)$ の表され方。この例では目標値が 3 回切り替わっている。

6.2 距離の推定実験

4. 章において示した距離推定則を実験において検証した。実験において、文献 [5] と同様の条件にもとづき、並進方向に関して位置制御を行うビジュアルサーボを行い、その際、同時に距離 Z_0 についてリアルタイムで推定を行った。どちらも逐次最小二乗法を用い、忘却計数は 0.95 から 1 に徐々に漸近させるようにしている。図 6 に示すように、拡大縮小量は値が 1 前後の小さな値で変動が小さく、ノイズの影響を受けやすいため結果が振動的になる問題がある。一方で、並進を使った推定は比較的安定するが、一方ではじめに計測した焦点距離 f の値に左右され、一定の誤差が生じると考えた方がよい。より良い距離の計測にはこれらの情報をカルマンフィルタ等の手法を用いて統合する必要があると思われる。

7. まとめ

カメラの動画をを用いて動作再現を行うための基礎的な検討を行った。動作再現の目標値として各フレームにおける画像群ではなく、ある特定のフレームの画像に対する変位を用いることでパラメータ化すること

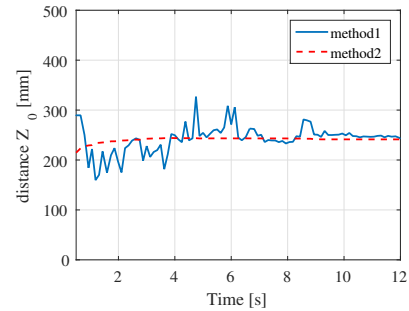


図 6 実験に基づくリアルタイム距離推定の結果。青線：拡大縮小量を用いた手法、赤点線：並進量を用いた手法

を提案した。また、従来のイメージベースのビジュアルサーボの手法に距離計測を伴うフィードフォワードを追加した式を提案し、時不変なイメージャコビアンを適用することで効率よく目的を果たすことができることを示した。予備実験に基づいて動作再現に必要な先の各要素を検証し、実用にいたるまでの道筋を示した。

参考文献

- [1] A. Dame and E. Marchand. “A new information theoretic approach for appearance-based navigation of non-holonomic vehicle.” *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011.
- [2] 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.159-165, 1996.
- [3] 遠藤 公誉, 田中 弘一, 荒川 賢一, 武川 直樹: “予見制御による視覚軌道追従制御法”, *日本ロボット学会誌* Vol. 15 No. 4, pp.565-572, 1997.
- [4] F. Chaumette and S. Hutchinson. : “Visual servo control. Part 1: Basic Approaches.” *IEEE Robot. Autom.Magazine* pp.82-90, 2006.
- [5] Ri, Y., and H. Fujimoto. :“Proposal of visual servoing using phase-only-correlation (POC).” *Industrial Electronics Society, IECON 2015-41st Annual Conference of the IEEE*. IEEE, 2015.
- [6] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi and K. Kobayashi. :“High-Accuracy Subpixel Image Registration Based on Phase-Only Correlation,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* Vol. E86-A, pp. 1925-1934. 2003.