

アクティブビジュアルサーボのための最適投影パターン

Optimal Projection Pattern for Active Visual Servoing

○正 荒井翔悟 (東北大) 学 宮本義大 (東北大)

正 小菅一弘 (東北大)

Shogo ARAI, Tohoku University, arai@tohoku.ac.jp

Yoshihiro MIYAMOTO, Tohoku University, miyamoto@irs.mech.tohoku.ac.jp

Kazuhiro KOSUGE, Tohoku University, kosuge@m.tohoku.ac.jp

Visual servo control uses images obtained by a camera to control robots. This paper focuses on a problem of positioning a object using a robotic manipulator with visual servo control. To perform the positioning task, image-based visual servoing requires visual features which can be extracted from appearance of a target object. Therefore, positioning error is remained for textureless objects, such as industrial parts. To solve this problem, this paper proposes a novel concept of visual servoing named "Active Visual Servoing". Active visual servoing projects pattern light on the target object using a projector. Thus designing of the projection pattern affects on the positioning error. We theoretically derive the optimal projection pattern which maximizes a cost function for image-based visual servoing. Experimental results show that the proposed active visual servoing achieves less positioning error compared to conventional image-based visual servoing.

Key Words: Visual servoing, Manipulator, Structured light

1 はじめに

カメラから撮像された画像を使って、物体の位置決めをする技術はビジュアルサーボと呼ばれ、広範に利用されている [5, 2]. ビジュアルサーボ法は大きく二種類に分類される。位置ベース法と画像ベース法である [1].

位置ベース法では、カメラで撮像された操作物体の画像から、物体の位置姿勢を推定し、物体の位置決めを行う。特に、エッジ、輪郭、コーナなどの画像特徴量を用いて物体の位置姿勢推定を行うことが多い [3]. 現時刻の画像特徴量だけでなく、時系列の画像特徴量が利用可能であるので、拡張カルマンフィルタなどのフィルタを利用して位置姿勢推定を行うことも可能である。工業部品などのテクスチャレスオブジェクトは、画像特徴量を安定して抽出することが困難であるので、位置ベース法をテクスチャレスオブジェクトの位置決め用いる場合には、何らかの工夫が必要となる。この問題を解決する一つのアプローチは、アクティブ三次元計測法を用いて、物体の三次元計測を行うことである。アクティブ三次元計測法では、プロジェクタから構造化光と呼ばれるパターン光を投影し、物体からの反射光をカメラで撮像する。撮像された画像を解析して三次元点群、または奥行き画像を取得し、物体の位置姿勢推定を行う。アクティブ三次元計測法を使った位置ベースビジュアルサーボでは、環境光の変化にロバストにテクスチャレスオブジェクトの位置決めが可能である一方で以下のような欠点がある。

1. カメラ-ロボット間のキャリブレーション誤差によって位置決め精度が左右される。つまり、高精度な位置決めのためには高精度なキャリブレーションが要求される。
2. 位置決め精度が、画像の量子化誤差とカメラのモデル化誤差の影響を受ける。
3. 一般に、三次元計測の精度をあげるためには、複数枚のパターン光を投影する必要がある。照射中に物体が移動することは望ましくないため、高速な位置決めを行うことができない。
4. 三次元計測、および三次元計測の結果から位置姿勢を推定するプロセスに要する計算に時間がかかり、この点も高速な位置決めへの妨げになる。

画像ベース法では、カメラで撮像された現時刻の操作物体画像から抽出された画像特徴量と予め与えられた目標画像から抽出される画像特徴量の偏差がゼロに収束する制御入力が計算される (Fig.1) [4]. 画像ベース法は、画像からの位置復元を必要としないので、位置ベース法で問題とされるカメラロボット間キャリブレーション誤差、画像量子化誤差、カメラのモデル化誤差に位置決め精度が影響されない利点がある。一方で、位置ベース法と同様、工業部品などのテクスチャレスオブジェクトに対しては、画像特徴量の抽出が困難であるという問題があった [1].

本稿ではこの問題を解決するために、プロジェクタからパターン光を物体に照射し、反射光をカメラで撮像し、その画像を使って、物体の位置決めを行うアクティブビジュアルサーボを提案する (Fig.2). 提案するアクティブビジュアルサーボ法は、画像の輝度値を特徴量として採用し、特徴の抽出計算を必要としないので、対応点検索を必要としない。画像の輝度値を特徴量として採用し、テクスチャレスオブジェクトの位置決めを行う場合、目標画像と現時刻画像の画像偏差が小さくなり、収束誤差 (位置決め誤差) が残りやすく、位置決め速度が遅いという問題がある。この問題に対しては、画像偏差を評価関数として、その値を最大化する最適照射パターンを理論的に導出する。すなわち、本論文で提案するアクティブビジュアルサーボは、既存の構造化光を用いるアプローチではなく、全く新しい試みである。

本論文で提案するアクティブビジュアルサーボの特徴、および新規性を下に示す。

1. 提案法による位置決め精度は、カメラのキャリブレーション誤差、量子化誤差、カメラモデル化誤差に依存しないという画像ベース法の利点を保持している。
2. 提案法は、画像の輝度値を特徴量として採用し、画像からの特徴の抽出計算を必要としないので、従来法 [7] と比較して高速計算が可能
3. いくつかの仮定の下で、目標位置姿勢近傍で画像偏差を最大にする最適なパターン光を導出したまた、この仮定の下で、最適なパターンが物体形状によらないことを発見した。
4. 通常画像ベース法との比較実験を行い、提案法により高精度な位置決めが達成可能であることを示した

構成は以下のとおりである。2節で画像ベース法の基礎についての概略を説明する。3節でアクティブビジュアルサーボの提案とそれに使用する最適照射パタンの導出を行う。4節で検証実験について記し、5節で結論を述べる。

2 画像ベースビジュアルサーボの基礎

ビジュアルサーボ法の概要を示すために、Fig.3に示す状況を考える。ここでの目的は、マニピュレータによって把持された物体を目標位置姿勢に位置決めを行う。操作中の物体は環境に設置された一台のカメラによって撮像される。また、事前に目標位置姿勢でマニピュレータが物体を把持しているシーンをカメラで撮像し、目標画像とする。ビジュアルサーボでは、目標画像とフィードバックした現時刻の画像からロボットへの制御入力を計算する。制御入力の計算法には、位置ベース法と画像ベース法の二種類に大別されるが、1節に記述した通り本稿では画像ベース法に焦点を当てる。

画像ベース法では、画像から直接計算された特徴量をフィードバックしてロボットを制御する (Fig.1)。ここで、特徴量とは、注目物体のエッジ、重心座標など、ロボット-カメラ間キャリブレーションやカメラモデルを利用せずに計算できる特徴を表す多次元ベクトルである。最も基礎的な制御測は

$$\dot{\theta}_d = \lambda J^{\dagger} (s(I(t)) - s(I^*)) \quad (1)$$

によって与えられる。ここで、 $\theta \in \mathbb{R}^n$ はマニピュレータの関節角速度指令値、 λ はゲイン、 J^{\dagger} は画像ヤコビアン疑似逆行列、 $s(I)$ は画像 I から特徴量への写像である。画像ヤコビアンは、目標画像と現在画像の偏差からマニピュレータの関節速度空間への写像であり、厳密にはマニピュレータの関節角度に依存する。しかし、目標位置姿勢近傍では、画像ヤコビアンが一定とみなせると考え、時不変のヤコビアンを与えることも多い。この場合、画像ヤコビアンは、

$$J_{\text{img}}^{\dagger} = \Delta\Theta\Delta S^{\dagger} \quad (2)$$

で計算できる。ここで、

$$\Delta\Theta = [\Delta\theta_1 \ \Delta\theta_2 \ \dots \ \Delta\theta_n], \quad (3)$$

$$\Delta\theta_i = \theta^* - \theta_i \quad (4)$$

$$\Delta S = [\Delta s_1 \ \Delta s_2 \ \dots \ \Delta s_n] \quad (5)$$

$$\Delta s_i = s^* - s_i \quad (6)$$

である。式 (3)–(6) で、目標位置姿勢に操作物体があるときの画像特徴量と関節角度が、それぞれ s^* と θ^* で、操作物体を目標位置姿勢からわずかにずらしたときの画像特徴量と関節角度が、それぞれ s_i と θ_i で与えられている。つまり、式 (2) で画像ヤコビアンを計算する際には、操作物体を目標位置姿勢から n 回わずかにずらした画像を取得する必要がある。

以上で説明した画像ベース法には、

1. 操作物体にテクスチャが乏しいと位置決め誤差が発生しやすい

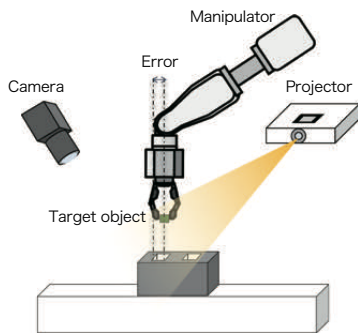


Fig.3 System configuration of active visual servoing.

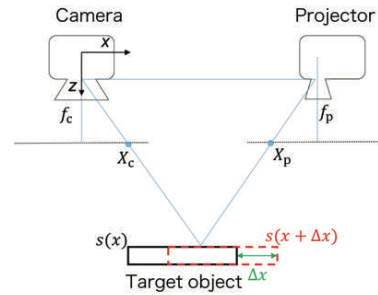


Fig.4 Arrangement of a camera, projector, and target object.

2. 目標位置近傍で一般に画像偏差が小さくなり、位置決めにかかる時間が増大する

という問題が指摘されている [1]。本論文では、こうした画像ベース法の問題を解決するためにプロジェクタを使った Active Visual Servo を次節以降で提案する。

3 アクティブビジュアルサーボの提案と最適照射パタンの導出

提案するアクティブビジュアルサーボ法では、プロジェクタを使ってパタン光を操作物体に照射し、その反射光をカメラによって撮像し、画像ベースビジュアルサーボを行う (Fig.2)。パタン光の投影によって、1. テクスチャが乏しい操作物体に対する位置決め誤差を低減すると同時に、2. 目標位置近傍での画像偏差が大きくなり、位置決めにかかる時間が短縮する、ことが期待される。アクティブビジュアルサーボの位置決め精度を決める大きな要因の一つに照射パタンがある。本節では、アクティブビジュアルサーボと最適な照射パタンを提案する。

Fig.4に示す二次元空間を考える。この空間には、操作対象物体、マニピュレータ、カメラ、プロジェクタが設置されており、座標系 Σ_c が固定されている。カメラは、その光学中心が座標系 Σ_c の原点に一致し、その光軸が z 軸方向に一致するように設置される。つまり、座標系 Σ_c はカメラ座標系である。カメラの撮像はピンホールカメラモデルに従うと仮定する。すなわち、位置 (x, z) にある点は、透視投影変換によって、カメラ画像平面上の

$$X_c = \frac{f_c}{z} x \quad (7)$$

に投影される。ここで、 f_c はカメラの焦点距離である。プロジェクタは、位置 (x_p, z_p) に、その光軸が z 軸に対して θ_p の角度をなすように設置される。このように設置されたプロジェクタの位置姿勢を $\xi_p := [x_p, y_p, \theta_p]^T$ と書く。また、プロジェクタの投影もピンホールカメラモデルに従うと仮定する。

いま、操作物体の位置姿勢を x と書き、二次元空間内で操作物体の表面形状を表す関数を $s(x)$ で表す。操作物体の目標位置姿勢を x^* とする。いま、マニピュレータによって物体を操作中に、少なくとも操作物体の一部がカメラによって撮像され、かつ撮像された箇所にはプロジェクタによってパタン光が照射されていると仮定する。以上の条件の下で、プロジェクタ画像平面上の X_p から照射された光が操作物体表面 $(x, s(x))$ で反射し、カメラ画像平面 X_c に到達すると思われることができる。いま、この関係を、プロジェクタ画像平面からカメラプロジェクタ平面への写像 g を用いて

$$X_c = g(x, s(x), X_p, \xi_p) \quad (8)$$

で表す。

次に、プロジェクタからの投影パタン光について考える。プロジェクタ画像平面 X_p から $I(X_p)$ の輝度で照射を行う。つまり、 $I(X_p)$ が投影パタンを表す関数である。

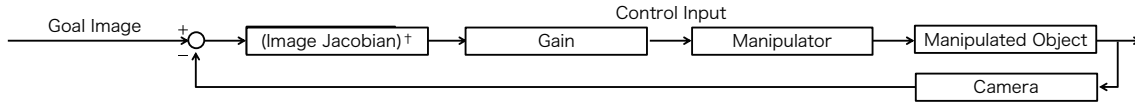


Fig.1 Block diagram of image-based visual servoing system.

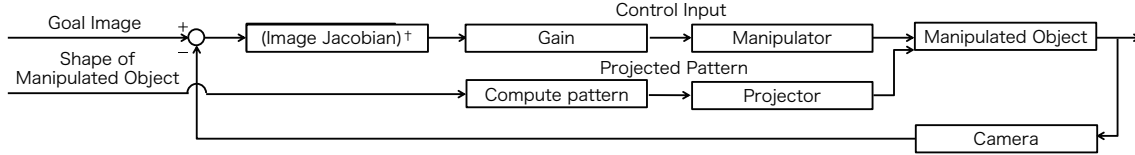


Fig.2 Block diagram of active visual servoing system proposed in this paper.

いま、カメラ画像平面 X_c に入射する光線の輝度を考える。この光線は、プロジェクタ画素

$$X_p = g^{-1}(x, s(x), X_c, \xi_p) \quad (9)$$

から照射された光である。ここで、 g^{-1} は g の逆関数である。このプロジェクタ画素から照射された光線の輝度は $I(X_p)$ であるので、式 (9) を使うと、 X_c に入射する光の強度 $p(X_c, x)$ は

$$p(X_c, x) = I(g^{-1}(x, s(x), X_c, \xi_p)) \quad (10)$$

とかける。ここでは、プロジェクタから照射された輝度がカメラで観測される輝度に等しいと仮定した。

いま、物体が目標位置 ξ^* に移動したとき、同じカメラ画像平面 X_c に入射する光線は、プロジェクタ画素

$$X_p^* = g^{-1}(X_c, x^*, s(x^*)) \quad (11)$$

から照射される。したがって、カメラ画素 X_c で観測される輝度は

$$p(X_c, x^*) = I(g^{-1}(X_c, x^*, s(x^*))) \quad (12)$$

となる。

ここで、提案するアクティブビジュアルサーボの制御則を

$$\theta_d = \lambda J^{\dagger}(I(t) - I^*) \quad (13)$$

で与える。(13) 式は、(1) 式の特徴量 s を画像 I に変更することによって得られる。画像 I は各要素に輝度値が格納されている行列である。したがって、画像から特徴量を抽出する計算を必要せず、(1) 式で表される従来の画像ベース法と比較して、高速な計算が可能である。

位置決め誤差を低減し、位置決めにかかる時間を短縮するために、目標位置近傍で (13) 式の画像偏差を最大化する照射パターン

$$I^* = \arg \max_I |J(I)|_2 \quad (14)$$

を求める。ここで、

$$J(I) = I(g^{-1}(X_c, x, s(x))) - I(g^{-1}(X_c, x^*, s(x^*))) \quad (15)$$

であり、 $|\xi|_2$ はベクトル ξ のユークリッドノルムを表す。位置姿勢 x が目標位置姿勢近傍であることを考慮して、式 (14) を x^*

近傍でテーラー展開すると、

$$\begin{aligned} J(I) &= (I(g^{-1}(X_c, s(x^* + \Delta x))) - I(g^{-1}(X_c, x^*, s(x^*))))^2 \\ &= \left[I \left(g(s(x^*), X_c) + \frac{\partial g}{\partial s} \Big|_{s=s(x^*)} \frac{\partial s}{\partial g} \Big|_{x=x^*} \Delta x \right) \right. \\ &\quad \left. - I(g^{-1}(X_c, x^*, s(x^*))) + O(\Delta x^2) \right]^2 \quad (16) \\ &= \underbrace{\frac{\partial I}{\partial g} \Big|_{g=g(s(x^*), X_c)}}_A \underbrace{\frac{\partial g}{\partial s} \Big|_{s=s(x^*)}}_B \underbrace{\frac{\partial s}{\partial x} \Big|_{x=x^*}}_C \Delta x \\ &\quad + O(\Delta x^2) \quad (17) \end{aligned}$$

を得る。ここで、 $O(\Delta x^2)$ は $\Delta x := x - x^*$ の 2 次以降の剰余項を表す。

式 (17) において、B と C はそれぞれ、操作物体の形状、カメラとプロジェクタの位置姿勢に依存するので、制御することはできない。項 A は、照射パターンに依存する項なので、これを最大化する I を求めればよい。いま、 $X_p = g(s(x^*), X_c)$ であるので、項 A は

$$K_{con} = \frac{\partial I}{\partial X_p} \Big|_{X_p=g(s(x^*), X_c)} \quad (18)$$

とかける。この式は、プロジェクタ画像平面上で照射輝度を画素座標で一階微分した際の大きさを表している。いま、プロジェクタ画像はある大きさをもつ画素から構成されることを考慮すれば、(17) 式、すなわち (14) 式を最大化する照射パターンは、格子パターンとなる。

4 検証実験

提案したアクティブビジュアルサーボ法の位置決め精度と収束速度を確認するため、通常の画像ベースビジュアルサーボとの比較実験を実施した。

実験環境を Fig.5 に示す。カメラはフォトロン製のハイスピードカメラ IDP-Express R2000、プロジェクタはエプソン製の EB-W420 を使用した。カメラとプロジェクタの解像度はそれぞれ、 512×512 (pixel)、 $1,280 \times 800$ (pixel) で、カメラの動作周波数は 50 [fps] である。マニピュレータとしてデンソーウェーブ社製 VS068 を使用した。平行グリッパをロボットハンドとして採用した。位置決め精度を測定するために、レーザーセンサ LJ-V7300 (KEYENCE 社製) を Fig.5 に示す位置に設置した。

実験では、Fig.5 に示す操作物体を平行グリッパに把持させ、把持誤差 (画像ヤコビアン生成時の把持位置からのズレ) を与えて比較実験を行った。アクティブビジュアルサーボと通常のビジュアルサーボの制御測は、式 (13) を採用した。ゲインは 2.0

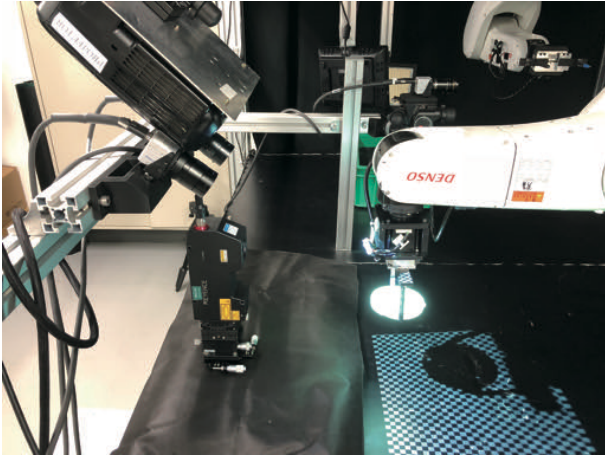


Fig.5 Experimental environment.

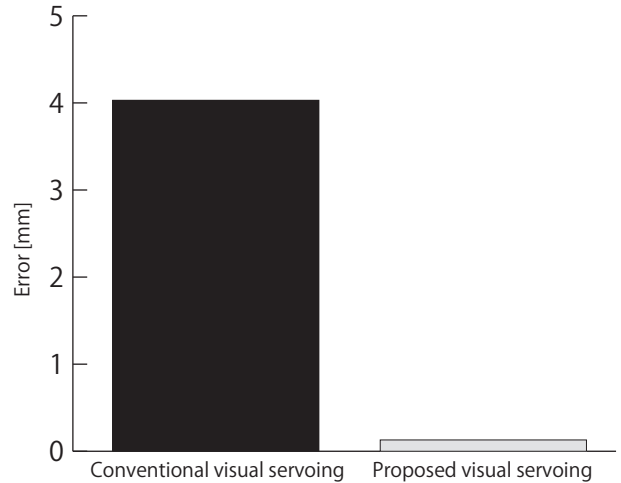


Fig.7 Positioning error.

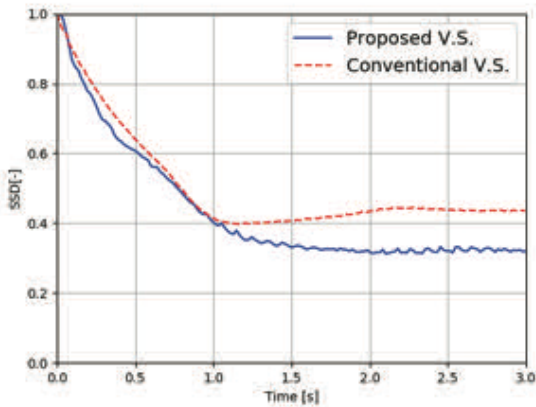


Fig.6 Time series data of SSD.

で固定した。提案するアクティブビジュアルサーボ法では、前節で示した通り、最適照射パターンである格子画像を照射した。

画像偏差 (式 (13) の $I(t) - I^*$) の SSD に関する時系列データを Fig.6 に示す。画像 L に関する SSD は、

$$SSD = \sum_{i=1}^N \sum_{j=1}^M L_{i,j} \quad (19)$$

で与えられる。ここで、 $L_{i,j}$ は画像 L の (i, j) pixel の輝度値、画像 L の解像度は $N \times M$ (pixel) である。Fig.6 において、初期時刻の SSD で画像偏差を正規化した。従来法と提案法において、約 1.5[s] で収束が見られたことがわかる。また、このグラフから提案法のほうが SSD 上では収束誤差が小さいことがわかる。ただし、SSD は画像偏差であるので、このグラフから三次元の位置決め精度について論じることは難しい。

次に、Fig.7 にレーザーセンサで計測した、終端時刻における位置決め誤差を示す。このグラフが示す通り、通常の画像ベース法と比較して、提案法によって 97% 以上の位置決め誤差低減が可能であることを示した。

5 結論

本稿では、従来の画像ベース法では困難であったテクスチャレス物体の高精度位置決めを目的としたアクティブビジュアルサーボを提案した。提案法では、プロジェクタからパターン光を操作対象物体に照射し、人工的に、目標画像との画像偏差を作り出す。この方法では、どのようなパターン光を照射するかが位置決め精度

に大きな影響を及ぼす。そこで、本稿では、画像偏差を最大化するパターン光を解析的に導出した。導出された最適なパターン光を使った検証実験の結果、通常の画像ベース法と比較して、97% 以上の位置決め誤差低減が可能であることを示した。

参考文献

- [1] F. Chaumette and S. Hutchinson, "Visual servo control. I. Basic approaches," IEEE Robotics & Automation Magazine, Vol. 13, No. 4, pp. 82-90, 2006.
- [2] C. Kingkan, S. Ito, S. Arai, T. Nammoto, and K. Hashimoto, "Model-based virtual visual servoing with point cloud data," IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 5549-5555, 2016.
- [3] 橋本, "ビジュアルサーボ - IV - 位置ベースビジュアルサーボ," システム/制御/情報, Vol.54, No.3, pp.117-123, 2010.
- [4] 橋本, "ビジュアルサーボ - V - 特徴ベースビジュアルサーボ," システム/制御/情報, Vol.54, No.5, pp.206-213, 2010.
- [5] M. Morita, S. Kagami, K. Hashimoto and K. Kosuge, "Markerless 6-DoF pose estimation of industrial wire-harness components from shapes in 2-D images", International Symposium on Robotics, pp.928-933, 2012.
- [6] J. Pagès, C. Collewet, F. Chaumette, and J. Salvi, "Plane-to-plane positioning from image-based visual servoing and structured light," IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1004-1009, 2014.
- [7] J. Pagès, C. Collewet, F. Chaumette, and J. Salvi, "A camera-projector system for robot positioning by visual servoing," IEEE Int. Workshop on Projector-Camera Systems, pp. 2-9, 2006.