触覚情報に基づく能動的物体認識のための探索行動設計
—相互情報量を規範とする最適制御によるアプローチ—

○田中大介　松原崇充　杉本謙二 (奈良先端大)

1. はじめに

人間との共存環境下で稼働するロボットにとって，環境認識の効率化は重要な課題である．本研究では，環境認識問題の一つとして，触覚情報に基づく物体認識問題 (図 1) を取り扱う．触覚情報はロボットが物体に触れることで初めて得られ，触り方に依存して得られる情報が大きく変わる．また，最適な触り方は物体に関して現在までに得られている情報にも依存する．このため，効率的に認識するためには状況に応じた探索行動の設計が重要となる．

これまで様々な探索行動の設計法が提案されているが，行動の計画問題とロボットの制御問題とが別々に取り扱われてきた（e.g. [1–4]）．この結果，ロボットにとって追従不可能な行動が計画される，もしくは不必要に大きなトルクが入力されることにより低エネルギー効率で危険な行動となる可能性があった．

我々は，この探索行動設計問題を最適制御問題として取り扱う．最適制御問題は，望む挙動をコスト関数で表し，そのコストが最小となるような制御則を求める問題であり，行動計画と制御問題を同時に解くことが可能である．探索行動設計問題で望む挙動は，対象に対する不確実性を下げることであり，加えてコスト関数の中にエネルギーやトルク制約を考慮することで，省エネルギーでかつ高コンプライアンスな最適探索行動が得られる．

2. 準備

2.1 逐次能動学習による能動的物体認識

我々は，物体認識問題をパラメータ推定問題 [3] として取り扱う．対象物体はそれぞれ固有のパラメータ (以降，物体パラメータと呼称する) が割り当てられており，探索行動により得られた触覚情報から物体パラメータを推定していく．

一般的な能動的物体認識の処理は以下のとおりである．

1. 物体パラメータの初期値をある確率分布として与える（この確率分布を信念（belief）と呼ぶ）
2. 現時点の信念に基づく最適な探索行動を設計する
3. 設計された探索行動を対象に実行し観測（触覚情報）を得る
4. 得られた観測から信念を更新する
5. 物体に関して確信が持てるまで（例えば信念の分散が十分小さくなるまで）2. から繰り返す
6. 信念の平均の最近傍の物体として認識結果を確定する

図 1 触覚情報に基づく物体認識問題

2.2 行動の最適性の定義

本稿では相互情報量を用いて探索行動の最適性を定義する．その時々で最適な探索行動は物体パラメータ $\boldsymbol{\theta}$ の確率分布で表される物体の信念 $p(\boldsymbol{\theta})$ に依存する．相互情報量 $\mathrm{I}[\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}]$ は，ある状態 $\mathbf{x}$ にて観測 $\mathbf{y}$ が与えられた時の $\boldsymbol{\theta}$ の不確実性が減った量，すなわち得られた情報の量を定量的に評価するもので，以下で与えられる．

$$\begin{aligned}\mathrm{I}\big[\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}\big] &\triangleq \mathrm{KL}(p(\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}) \| p(\boldsymbol{\theta})p(\mathbf{y}|\mathbf{x})) \\ &= \iint p(\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}) \log \frac{p(\boldsymbol{\theta}, \mathbf{y}|\mathbf{x})}{p(\boldsymbol{\theta})p(\mathbf{y}|\mathbf{x})} \mathrm{d}\mathbf{y}\mathrm{d}\boldsymbol{\theta} \quad (1)\end{aligned}$$

この量を最大化する状態（系列）にシステムを制御することで，パラメータ推定に有効な観測を得る．

3. 提案法: 最適制御に基づく逐次探索行動設計

3.1 対象システムと問題設定

物体認識に用いるロボットとセンサは，それぞれ以下の非線形状態方程式と観測方程式で表される離散時間システムモデルで表されるとする．

$$\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t) \quad (2)$$

$$\mathbf{y}_t = \mathbf{g}(\mathbf{x}_t, \boldsymbol{\theta}) + \boldsymbol{\epsilon}_t \quad (3)$$

ここで，$\mathbf{x} \in \mathbb{R}^{d_\mathbf{x}}$ は観測可能なロボットの状態，$\mathbf{u} \in \mathbb{R}^{d_\mathbf{u}}$ はロボットへの入力，$\mathbf{y} \in \mathbb{R}^{d_\mathbf{y}}$ はロボットのセンサからの観測，$\boldsymbol{\epsilon} \in \mathbb{R}^{d_\mathbf{y}}$ は平均 $\mathbf{0}$，共分散行列 $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \mathrm{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_{d_\mathbf{y}}^2\}$ の $d_\mathbf{y}$ 次元ガウス分布に従う観測ノイズである．ロボットへの入力は，$\mathbf{u}_{\min} \leq \mathbf{u} \leq \mathbf{u}_{\max}$ に制限されるとする．また，$\boldsymbol{\theta} \in \mathbb{R}^{d_\boldsymbol{\theta}}$ は物体パラメータである．

このシステムにより前節で説明した能動的物体認識の処理を行う．処理 2. はロボットのダイナミクス (2) の

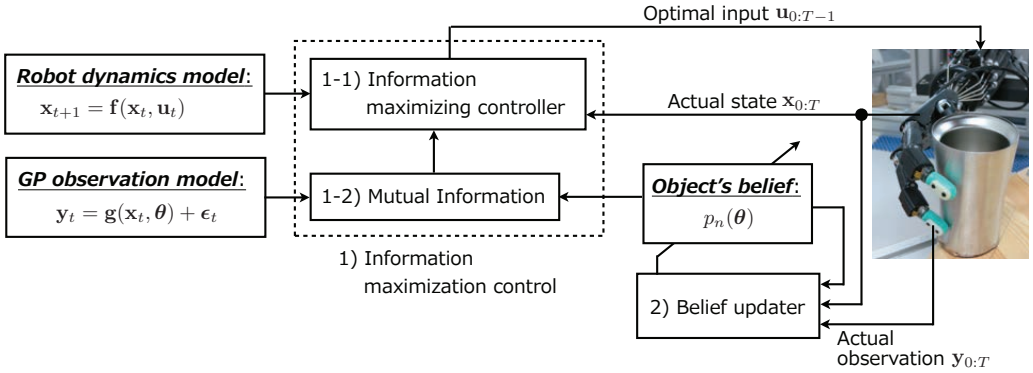

図 2 提案法の概要図．1-1) 得られる情報量を最大にする制御器，1-2) ガウス過程回帰による観測モデルから計算される相互情報量から構成される 1) 情報量最大化制御，2) 得られた観測からの信念の更新の 2 要素で構成される．

制約を考慮した上で，コストの総和が最小となる状態系列に対応する行動系列を探索する問題，つまり相互情報量を規範とする最適制御問題に帰着させる（**3.2.1**）．このコストには相互情報量 (1) を用いるが，2 重積分を簡単に計算する工夫として，ガウス過程回帰 (Gaussian Process Regression) を用いた観測モデルを利用する（**3.2.2**）．処理 4. では得られた状態系列から信念を更新する（**3.3**）．

3.2 情報量最大化制御

3.2.1 局所化に基づく最適制御問題の近似解法

まず，最適制御問題を定式化する．これは，式 (2) のダイナミクスに従うシステムに対して，次式で与えられる評価関数を最小化するような制御則 $\mathbf{u}_t = \boldsymbol{\pi}(t, \mathbf{x}_t)$ を定める問題である．

$$J_{0:T} = h(\mathbf{x}_T) + \sum_{t=0}^{T-1} \ell(t, \mathbf{x}, \boldsymbol{\pi}(t, \mathbf{x}))$$

ただし，$h(\mathbf{x}_T) \geq 0$ は終端コスト，$\ell(t, \mathbf{x}, \mathbf{u}) \geq 0$ はランニングコストを表す．探索行動設計に対しては，これらコスト関数はエネルギー効率と行動の最適性を用いて与える．実際の与え方については後述する．この問題は，システムが線形でかつ 2 次コストであれば Riccati 方程式に基づく方法で最適制御則 $\boldsymbol{\pi}^*$ を導出可能であるが，ここではシステムを初期入力列 $\bar{\mathbf{u}}_{0:T-1}$ を入力した時の初期状態列 $\bar{\mathbf{x}}_{0:T}$ 周りで局所的に線形システムを構築し近似的に解く iterative Linear Quadratic Regulator (iLQR, [5]) を用いる．

この手法を用いる利点は，局所的なフィードバック制御を構築可能なことである．iLQR によりフィードバックゲイン $\mathbf{L}_t$ が与えられ，実際に与える入力 $\mathbf{u}_t$ は $\mathbf{u}_t = \bar{\mathbf{u}}_t + \mathbf{L}_t(\mathbf{x}_t - \bar{\mathbf{x}}_t)$ で得られる [5]．$\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t$ はそれぞれ iLQR で得られた最適な状態と入力であり，$\mathbf{x}_t$ はロボットの実際の状態である．このフィードバックにより，ダイナミクスに物体との相互作用が考慮されていない，もしくはモデル化誤差が存在する場合にも，その状態での最適な入力を生成可能である．この場合にダイナミクスを正確にモデル化することは困難であるため，このようなフィードバックが有効に機能することが期待される．

3.2.2 相互情報量による逐次コストの定義

観測モデル (3) は，ガウス過程回帰 [6] により $\mathbf{y}$ の各次元 $y_a = g_a(\mathbf{x}, \boldsymbol{\theta}) + \epsilon_a, a \in \{1, 2, \ldots, d_{\mathbf{y}}\}$ がモデル化されているとする．訓練データ $\mathcal{D} = \{\mathbf{x}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{y}^{(i)}\}_{i=1}^N$ が与えられた時，予測分布は $\mathbf{x}, \boldsymbol{\theta}, y_a$ のそれぞれの訓練データ集合 $\mathbf{X}, \boldsymbol{\Theta}, \mathbf{y}_a$ を用いて以下のガウス分布で与えられる [6]．

$$\begin{aligned} &p(y_a|\mathbf{x}, \boldsymbol{\theta}, \mathbf{X}, \boldsymbol{\Theta}, \mathbf{y}_a) \\ &= \mathcal{N}(\mu_a(\mathbf{x}, \boldsymbol{\theta}; \mathbf{X}, \boldsymbol{\Theta}, \mathbf{y}_a), s_a^2(\mathbf{x}, \boldsymbol{\theta}; \mathbf{X}, \boldsymbol{\Theta}, \mathbf{y}_a)) \end{aligned}$$

本稿では μ_a, s_a^2 の計算に用いられるカーネル関数は以下で表される 2 乗指数カーネル関数 (squared exponential kernel function) を仮定する．

$$k_a(\mathbf{z}, \mathbf{z}') = \alpha_a^2 \exp\left(-\tfrac{1}{2}(\mathbf{z} - \mathbf{z}')^{\mathrm{T}} \left(\mathbf{H}_a^{\mathbf{z}}\right)^{-1} (\mathbf{z} - \mathbf{z}')\right)$$

ここで $\mathbf{z} = \left[\mathbf{x}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{d_{\mathbf{z}}}, d_{\mathbf{z}} = d_{\mathbf{x}} + d_{\boldsymbol{\theta}}$ を用いた．また α_a^2 は g_a の分散で，モデルの簡単化のために $\mathbf{x}$ と $\boldsymbol{\theta}$ は独立であると仮定し，$\mathbf{H}_a^{\mathbf{z}} = \mathrm{block\ diag}\{\mathbf{H}_a^{\mathbf{x}}, \mathbf{H}_a^{\boldsymbol{\theta}}\}$ と定めた．$\mathbf{H}_a^{\mathbf{x}}, \mathbf{H}_a^{\boldsymbol{\theta}}$ は共に正の要素を対角に持つ対角行列である．ハイパーパラメータは $\gamma_a = \{\alpha_a^2, \sigma_a^2, \mathbf{H}_a^{\mathbf{x}}, \mathbf{H}_a^{\boldsymbol{\theta}}\}$ であり，訓練データに対する周辺対数尤度関数最大化により得る [6]．

ここからは，$\boldsymbol{\theta}$ の信念をガウス分布 $p(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ で表す．ある 1 点の $\mathbf{x}$ が与えられた時，相互情報量 $\mathrm{I}\left[\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}\right]$ は，GP 観測モデルの訓練データとハイパーパラメータを用いて，以下のように表される．

$$\mathrm{I}\left[\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}\right] = -\frac{1}{2}\log\left(\frac{\det \tilde{\boldsymbol{\Sigma}}(\mathbf{x})}{\det \boldsymbol{\Phi}(\mathbf{x}, \mathbf{x}) \det \boldsymbol{\Sigma}}\right)$$

$$\tilde{\boldsymbol{\Sigma}}(\mathbf{x}) = \begin{bmatrix} \boldsymbol{\Phi}(\mathbf{x}, \mathbf{x}) & \mathbf{C}^{\mathrm{T}}(\mathbf{x}) \\ \mathbf{C}(\mathbf{x}) & \boldsymbol{\Sigma} \end{bmatrix}$$

ただし，$\boldsymbol{\Phi}(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^{d_{\mathbf{y}} \times d_{\mathbf{y}}}$ の (a, b) 要素 Φ_{ab} と $\mathbf{C}(\mathbf{x}) \in \mathbb{R}^{d_{\boldsymbol{\theta}} \times d_{\mathbf{y}}}$ は以下で与えられる．

$$\begin{aligned} \Phi_{ab} &= \boldsymbol{\beta}_a^{\mathrm{T}} \boldsymbol{\Lambda}_{ab}(\mathbf{x}, \mathbf{x}') \boldsymbol{\beta}_b - m_a(\mathbf{x}) m_b(\mathbf{x}') \\ &\quad + \delta_{\mathbf{x}\mathbf{x}'} \delta_{ab} \left(\alpha_a^2 - \mathrm{Tr}\left(\mathbf{K}_a^{-1} \boldsymbol{\Lambda}_{aa}(\mathbf{x}, \mathbf{x}')\right)\right) \\ \mathbf{C} &= \boldsymbol{\Psi}(\mathbf{x}) - \boldsymbol{\mu}\mathbf{m}(\mathbf{x})^{\mathrm{T}} \end{aligned}$$

ここで行列 $\mathbf{K}_a$ はカーネル行列であり，その (i,j) 要素 $K_{a,ij}$ は $K_{a,ij} = k_a(\mathbf{z}^{(i)}, \mathbf{z}^{(j)}) + \delta_{ij}\sigma_a^2$ で与えられる．また，$\mathbf{m}(\mathbf{x}) \in \mathbb{R}^{d_\mathbf{y}}$ の a 番目の要素は $m_a = \boldsymbol{\beta}_a^{\mathrm{T}} \boldsymbol{\lambda}_a(\mathbf{x})$ であり，$\boldsymbol{\beta}_a = \mathbf{K}_a^{-1}\mathbf{y}_a \in \mathbb{R}^N$ とした．$\boldsymbol{\Lambda}_{ab}(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^{N\times N}$ については付録に示すが，$\boldsymbol{\lambda}_a(\mathbf{x}) \in \mathbb{R}^N$，$\boldsymbol{\Psi}(\mathbf{x}) \in \mathbb{R}^{d_{\boldsymbol{\theta}}\times d_\mathbf{y}}$ については文献 [3] を参照されたい．

3.3 状態系列に基づく信念の更新

現在の信念 $p_n(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$ を基にして先に述べた制御法により探索行動が計画後実行され，時刻 $t=0$ から $t=T$ までの状態列 $\mathbf{x}_{0:T}^n$ とその時の観測 $\mathbf{y}_{0:T}^n$ が得られたとする．これを用いて信念の更新則は，文献 [3] で提案されている状態と観測が 1 点ずつ与えられた時の更新則を拡張し，以下のように与えられる．

$$\boldsymbol{\mu}_{n+1} = \boldsymbol{\mu}_n + \mathcal{C}_n \mathcal{S}_n^{-1}(\mathcal{Y}_n - \mathcal{M}_n)$$
$$\boldsymbol{\Sigma}_{n+1} = \boldsymbol{\Sigma}_n - \mathcal{C}_n \mathcal{S}_n^{-1} \mathcal{C}_n^{\mathrm{T}}$$

ただし，$\mathcal{C}_n \in \mathbb{R}^{d_{\boldsymbol{\theta}}\times(T+1)d_\mathbf{y}}$，$\mathcal{S}_n \in \mathbb{R}^{(T+1)d_\mathbf{y}\times(T+1)d_\mathbf{y}}$，$\mathcal{Y}_n \in \mathbb{R}^{(T+1)d_\mathbf{y}}$，$\mathcal{M}_n \in \mathbb{R}^{(T+1)d_\mathbf{y}}$ は以下のように定義した．

$$\mathcal{C}_n = \begin{bmatrix} \mathbf{C}(\mathbf{x}_0^n) & \cdots & \mathbf{C}(\mathbf{x}_T^n) \end{bmatrix}$$
$$\mathcal{S}_n = \begin{bmatrix} \boldsymbol{\Phi}(\mathbf{x}_0^n, \mathbf{x}_0^n) & \cdots & \boldsymbol{\Phi}(\mathbf{x}_0^n, \mathbf{x}_T^n) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Phi}(\mathbf{x}_T^n, \mathbf{x}_0^n) & \cdots & \boldsymbol{\Phi}(\mathbf{x}_T^n, \mathbf{x}_T^n) \end{bmatrix}$$
$$\mathcal{Y}_n = \begin{bmatrix} (\mathbf{y}_0^n)^{\mathrm{T}} & \cdots & (\mathbf{y}_T^n)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$$
$$\mathcal{M}_n = \begin{bmatrix} \Big(\mathbf{m}\big(\mathbf{x}_0^n\big)\Big)^{\mathrm{T}} & \cdots & \Big(\mathbf{m}\big(\mathbf{x}_T^n\big)\Big)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$$

物体認識は前節で示したように，行動計画と信念の更新とを繰り返すことで達成される．例えば終了条件として，適切な閾値 ϵ を用いて $\|\boldsymbol{\mu}_{n+1} - \boldsymbol{\mu}_n\|^2 < \epsilon$，$\|\boldsymbol{\Sigma}_n\|_{\mathrm{F}}^2 < \epsilon$，もしくは $n_{\max}$ 回繰返すまでと設定することで，最終的に最近傍の $\boldsymbol{\theta}$ に対応する物体との識別結果を得る．

4. 数値シミュレーション

4.1 シミュレーション設定

図 3 に示す 1 リンクアームで，提案する手法の有効性を検証する．ただし，このアームの可動域は $-\pi/2 \le q \le \pi/2$ とし，運動方程式は $\Delta t = 0.01$[s] でオイラー離散化した．物体はバネ κ・ダンパ D を並行に接続した系に基づく押しこみ反力モデルと動摩擦モデルを仮定した．手先に仮想的に圧力センサが設置されていることとし，以下の物体押しこみ方向の力 f_1 と擦る方向の力 f_2 を得た．

$$f_1 = -(\kappa\xi_x + D\dot{\xi}_x),\ f_2 = -\mathrm{sign}(\dot{\xi}_y)\mu' f_1$$

ここで，$\boldsymbol{\xi} = [\xi_x, \xi_y]^{\mathrm{T}}$ は手先のデカルト座標である．

この実験では物体認識問題をダンパ係数 D の推定問題とする．物体パラメータは $\theta = D$ であり，探索行動により図 3 の物体の D を推定する．ここからまず観測モデル学習の設定について述べる．バネ定数を

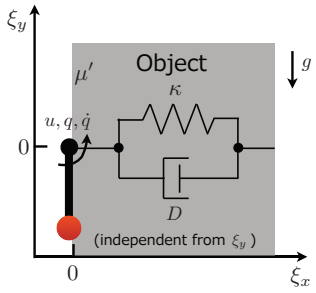

図 3 シミュレーション実験設定．物体はバネダンパ並列モデル（バネ定数 κ，ダンパ定数 D）で反発力をモデル化し，表面には動摩擦係数 μ' の摩擦モデルを仮定した．これらは，物体表面の高さ方向の位置に依存しないと仮定する．アームの先は赤丸で表しており，アームが ξ_x 軸方向と平行のとき $q=0$ である．また，図中のアームの姿勢は $\mathbf{x}_0 = [-\pi/2, 0]^{\mathrm{T}}$ に対応している．入力 u や $q, \dot{q}$ の正方向は反時計回りに定めた．

$\kappa = 1$，動摩擦係数 $\mu' = 0.5$ と固定し，ダンパ係数 D を $D \in \{1, 3, 5\}$ と変化させた．観測を $\mathbf{y} = [f_1, f_2]^{\mathrm{T}}$，状態を $\mathbf{x} = \begin{bmatrix} q, \dot{q} \end{bmatrix}^{\mathrm{T}}$ として，それぞれの D に対して，$-\pi/2 \le q \le \pi/2, -15 \le \dot{q} \le 15$ について観測 $\mathbf{y}$ を得て，訓練データ $\mathcal{D}$ を構成した．データ数は各 D に対して 225 個，よって総データ数は $N = 675$ である．信念の更新回数は $n_{\max} = 10$ とした．

次に最適制御則の導出に関する設定を述べる．初期状態は $\mathbf{x}_0 = [-\pi/2, 0]^{\mathrm{T}}$ とし，1 行動の長さ $T = 100$ として初期入力列は $u_t = 3, t = 0, 1, \ldots, T-1$ とした．また $p_n(\theta)$ に基づく探索行動についての逐次コスト $\ell_n(t, \mathbf{x}, u)$ は以下のように設定した．

$$\ell_n(t, \mathbf{x}, u) = \exp\Big(-\rho^{(n)} \mathrm{I}\big[\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}\big]\Big) + 1000\exp(-(15 - \dot{q})) + 1000\exp(-(15 + \dot{q}))$$

ただし，$\rho^{(n)} > 0$ は最大化問題を最小化問題に変更するために用いた定数であり，実験的に $\rho^{(n)} = 10\exp(-0.2n^2)$ と選んだ．また終端コストは $h(\mathbf{x}) = \ell_n(T, \mathbf{x}, 0)$ とした．このコスト関数の第 2 項と第 3 項は，アームの角速度が訓練データの範囲を逸脱しないように入れた罰則項である．さらに，初期信念として，$p_0(\theta) = \mathcal{N}(3, 5^2)$ を与えた．この平均は，訓練データの θ の平均に等しい．

4.2 シミュレーション結果

観測モデルから計算される初期信念 $p_0(\theta)$ に対する相互情報量の分布を図 4 に示す．この分布について考察する．ダンパの力は押しこみ方向の速度 $\dot{\xi}_y$ に依存して変化する．$q=0$ のときには $\dot{\xi}_y$ は 0 になるため，ダンパ係数に関する情報は得られない．対して，$|q|$ が $\pi/2$ に近い部分では，大きな $\dot{\xi}_y$ が得られるため，ダンパ係数に関する情報は得られやすい．よって，この相互情報量の分布は妥当だと考える．

次に，$D=1$ の物体に対する認識結果を図 5 に示す．横軸は信念の更新回数であり，各信念の更新には $\mathbf{x}_{0:T}$

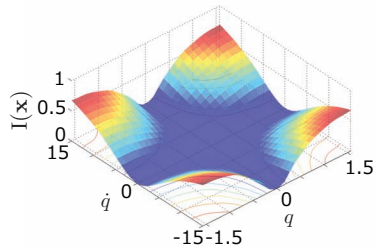

図 4 相互情報量の分布．初期信念 $p_0(\theta) = \mathcal{N}(3, 5^2)$ に対して得た．

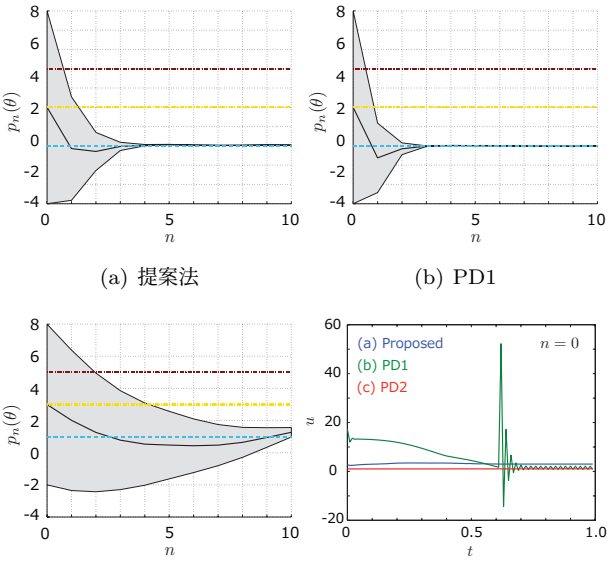

図 5 パラメータの推定結果．(a-c) 横軸は信念の更新回数で，$p_0(\theta) = \mathcal{N}(3, 5^2)$ である．真値 $D = 1$ は水色の点線で表した．黒線と灰色の領域は，$p_n(\theta)$ の平均と標準偏差を表す．これらの探索行動は，(a) 提案法，(b) 高いゲインの PD コントローラで，(c) 低いゲインの PD コントローラでそれぞれ得られた．(d) (a-c) のそれぞれの探索行動の入力トルクの関係を表す．

が用いられている．図 5(a) は提案法により得た探索行動により信念の更新を行った結果であり，$n = 4$ の時点でほぼ結果が収束していることが確認できる．この結果と比較するため，$\mathbf{x}_d = [\pi/2, 15]^{\mathrm{T}}$ に向かう PD コントローラを設計し，探索行動を行った．この目標値は，図 4 から設定した．図 5(b) は高いゲインに設定した時で，提案法よりも少ない更新回数で収束したように見えるが，図 5(d) に示すように高いトルクが入力されており危険である．これに対して低いゲインに設定した場合には入力トルクは小さいが図 5(c) に示すように収束は遅い．これらの比較から，提案法の有効性が確認できた．

5. おわりに

本稿では触覚情報に基づく能動的物体識別のための探索行動設計問題に対して，相互情報量を最大化する最適制御によるアプローチを提案した．今後，図 6 に示す Shadow Dexterous Hand を用いた実機実験システム [4] を用いて提案法の有効性を確認する．

図 6 実機実験用の設定

付録 A　$\boldsymbol{\Lambda}$ の定義

訓練データとハイパーパラメータを用いて，$\boldsymbol{\Lambda}_{ab}$ の (i, j) 要素 $\Lambda_{ab,ij}$ は次のように表される．

$$\begin{aligned}\Lambda_{ab,ij} = \alpha_a^2 \alpha_b^2 \det\left(\left(\left(\mathbf{H}_a^{\boldsymbol{\theta}}\right)^{-1} + \left(\mathbf{H}_b^{\boldsymbol{\theta}}\right)^{-1}\right)\boldsymbol{\Sigma} + \mathbf{I}\right)^{-\frac{1}{2}} \\ \times \exp\left(-\tfrac{1}{2}\left(\boldsymbol{\theta}^{(i)} - \boldsymbol{\theta}^{(j)}\right)^{\mathrm{T}}\left(\mathbf{H}_{ab}^{\boldsymbol{\theta}}\right)^{-1}\left(\boldsymbol{\theta}^{(i)} - \boldsymbol{\theta}^{(j)}\right)\right) \\ \times \exp\left(-\tfrac{1}{2}\left(\boldsymbol{\theta}_{ab}^{ij} - \boldsymbol{\mu}\right)^{\mathrm{T}}\mathbf{R}_{ab}^{-1}\left(\boldsymbol{\theta}_{ab}^{ij} - \boldsymbol{\mu}\right)\right) \\ \times \exp\left(-\tfrac{1}{2}\left(\mathbf{x} - \mathbf{x}^{(i)}\right)^{\mathrm{T}}\left(\mathbf{H}_a^{\mathbf{x}}\right)^{-1}\left(\mathbf{x} - \mathbf{x}^{(i)}\right)\right) \\ \times \exp\left(-\tfrac{1}{2}\left(\mathbf{x}' - \mathbf{x}^{(j)}\right)^{\mathrm{T}}\left(\mathbf{H}_b^{\mathbf{x}}\right)^{-1}\left(\mathbf{x}' - \mathbf{x}^{(j)}\right)\right)\end{aligned}$$

ただし，$\mathbf{H}_{ab}^{\boldsymbol{\theta}} = \mathbf{H}_a^{\boldsymbol{\theta}} + \mathbf{H}_b^{\boldsymbol{\theta}}$ であり，

$$\boldsymbol{\theta}_{ab}^{ij} = \mathbf{H}_b^{\boldsymbol{\theta}}\left(\mathbf{H}_{ab}^{\boldsymbol{\theta}}\right)^{-1}\boldsymbol{\theta}^{(i)} + \mathbf{H}_a^{\boldsymbol{\theta}}\left(\mathbf{H}_{ab}^{\boldsymbol{\theta}}\right)^{-1}\boldsymbol{\theta}^{(j)}$$

$$\mathbf{R}_{ab} = \left(\left(\mathbf{H}_a^{\boldsymbol{\theta}}\right)^{-1} + \left(\mathbf{H}_b^{\boldsymbol{\theta}}\right)^{-1}\right)^{-1} + \boldsymbol{\Sigma}$$

を用いた．

参考文献

[1] J. A. Fishel and G. E. Loeb: “Bayesian exploration for intelligent identification of textures,” Frontiers in Neurorobotics, vol.6, no.4, 2012.

[2] J. Sinapov, C. Schenck, K. Staley, V. Sukhoy, and A. Stoytchev: “Grounding semantic categories in behavioral interactions: Experiments with 100 objects,” Robotics and Autonomous Systems, vol.62, no.5, pp.632–645, 2014.

[3] H. Saal, J.-A. Ting, and S. Vijayakumar: “Active sequential learning with tactile feedback,” Proceedings of the 13th International Conference on Artificial Intelligence and Statistics, pp.677–684, 2010.

[4] D. Tanaka, T. Matsubara, K. Ichien, and K. Sugimoto: “Object manifold learning with action features for active tactile object recognition,” Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014. to appear.

[5] W. Li and E. Todorov: “Iterative linear quadratic regulator design for nonlinear biological movement systems,” Proceedings of International Conference on Informatics in Control, Automation and Robotics, pp.222–229, 2004.

[6] C. E. Rasmussen and C. K. I. Williams: Gaussian processes for machine learning. MIT Press, 2006.