

操作主体感のデザインのための自由エネルギーモデル

遅延を伴う操作系を用いた検証

Free Energy Model for Designing Sense of Agency

Validation Using Control Systems with Delay

○谷山建作（東京大学）^{*1} 澁江秀明（(株) 本田技術研究所）^{*2} 柳澤秀吉（東京大学）^{*3}

^{*1} Kensaku Taniyama, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku Tokyo-to, 113-8654, taniyama-kensaku688@g.ecc.u-tokyo.ac.jp

^{*2} Hideaki Shibue, Honda R&D Co., Ltd., 4630 Shimotakanezawa Haga-machi Haga-gun Tochigi-ken, 321-3321, hideaki_shibue@jp.honda

^{*3} Hideyoshi Yanagisawa, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku Tokyo-to, 113-8654, hide@mech.t.u-tokyo.ac.jp

キーワード: 操作主体感, 自由エネルギー原理, 遅延

1. 緒言

操作主体感とは、機械の挙動を引き起こしているのは自分である感覚を意味する⁽¹⁾。操作主体感は、運動主体感から派生した概念であり、運動主体感とは、行動（身体の運動）を引き起こしているのは自分である感覚を意味する⁽²⁾。運動主体感の喪失は無力感（helplessness）を引き起こすことや、運動主体感は自身の行動に対する責任感と強く結びついていることが知られている⁽³⁾。つまり、運動主体感は、喜びや責任感の要因であると考えられる。同様に、操作主体感も機械を操作する喜びや機械の挙動に対する責任感の要因であると考えられる。操作主体感の因子を明らかにすることは、人機械系のユーザ体験や安全性を向上させる上で重要である。また、様々な条件下における操作主体感を事前に定量的に予測することができれば、設計の段階で操作主体感を考慮することができる。そのため、操作主体感をデザインするためのアプローチとして、操作主体感の数理モデリングが有用であると考えられる。

筆者らは、脳が処理する情報量が感覚値に対応するという仮説の下、操作主体感の数理モデリングを行ってきた^{(4),(5)}。本研究の目的は、脳の情報理論である自由エネルギー原理をもとに構築した操作主体感の数理モデルの変数を自動車の運転条件に適用し解釈すること、そして、その条件を操作した実験によって数理モデル予測を検証することである。

2. 研究方法

2.1. 操作主体感の数理モデル

運動主体感の定性的なモデルとしてはコンパレータモデルがある。コンパレータモデルは、元来、運動制御を説明する内部モデルであり、望まれる状態をもとに運動指令が出力され、運動指令をもとに状態の予測と運動が行われる過程を説明する⁽⁶⁾。コンパレータモデルによると、運動制

御の内部モデルを構成する順モデルや逆モデルを学習するために、望まれる状態・予測される状態・実際の状態の3者間の比較が行われる。予測される状態と実際の状態の差は予測誤差と呼ばれ、運動主体感を減少させると考えられている。

予測誤差を定式化するために、知覚・学習・行動を説明する情報理論である自由エネルギー原理を用いる。自由エネルギー原理によると、知覚・学習・行動は自由エネルギーというコスト関数を最小化するように決まる⁽⁷⁾。自由エネルギーは予測誤差と解釈される情報量であり、式(1)のように表される。

$$F = \int q(x) \log \left(\frac{q(x)}{p(y, x)} \right) dx \quad (1)$$

x, y はそれぞれ隠れ状態、感覚信号を意味し、 $q(x)$ は認識分布、 $p(y, x)$ は生成モデル（内部モデル）を表す。自由エネルギー原理によると、知覚過程を通して、自由エネルギー F を最小化するように認識分布 $q(x)$ が決まる。認識分布 $q(x)$ が事後分布 $p(x|y)$ に一致する時、自由エネルギー F は最小化され、式(2)のように表される。

$$\min_{q(x)} F = -\log p(y) \quad (2)$$

知覚過程を通して最小化される自由エネルギーは、サブライズと呼ばれ、コンパレータモデルにおける予測誤差に対応すると考えられる。従って、サブライズの符号を反転させたもの（負のサブライズ）が運動主体感を表し、 x が機械の隠れ状態の場合、操作主体感を表すと考えられる。

負のサブライズは、事前分布（予測）と尤度（観測）で表される。ここで、事前分布と尤度を正規分布で近似すると、負のサブライズは、式(3)のように表される。

$$\log p(y) = -\frac{1}{2} \left(\frac{1}{\sigma_p^2 + \sigma_l^2} |\mu_l - \mu_p|^2 + \log 2\pi(\sigma_p^2 + \sigma_l^2) \right) \quad (3)$$

$|\mu_l - \mu_p|$ は、事前分布と尤度の期待値の差、 σ_p^2 は、事前分布の分散、 σ_l^2 は尤度の分散であり、それぞれ、予測誤差、予

測不確実性、観測不確実性と解釈される。予測誤差が増加すると負のサプライズは減少し、不確実性が増加すると負のサプライズは、予測誤差が十分小さい場合は減少し、予測誤差が十分大きい場合は増加することが知られている⁽⁶⁾。つまり、この数理モデルは、予測誤差と不確実性が交互作用を持って操作主体感に影響することを予測する。

2.2. 数理モデル予測の検証実験

2.2.1. 実験目的 筆者らは、離散操作系を用いて、予測誤差と予測不確実性の交互作用の検証⁽⁴⁾と予測誤差と観測不確実性の交互作用の検証⁽⁵⁾を行ってきた。一方、操作主体感は自動車の運転などの連続操作系において重要な概念であると考えられるため、連続操作系を用いた検証が必要である。そこで、本実験の目的は、ドライビングシミュレータを用いた予測誤差と予測不確実性の交互作用の検証とする。

2.2.2. モデル変数と運転条件の対応 先行研究では、予測と実際の不一致を引き起こすためによく遅延が用いられる^{(4),(5),(8)-(10)}。自動車における遅延には、例えば、ハンドル操作から車両挙動までの遅延などがある。この場合、予測誤差は、車両挙動の遅延に対する予測と観測の期待値の差、予測不確実性は、車両挙動の遅延に対する予測の分散である。実際の運転条件に適用すると、例えば、予測誤差は、運転者が学習している車両挙動の遅延と路面状態の変化や自動車の乗り換えなどによる変化後の車両挙動の遅延の差、予測不確実性は、学習している車両挙動の遅延のばらつきなどに対応する。

2.2.3. 実験方法 本実験では、ハンドル操作するとディスプレイ上の視界が変化するドライビングシミュレータを用いる。コースは、直線・クロソイド曲線・円弧によって構成されるものとする。図1に1試行で走行するコースの鳥瞰図を示す。

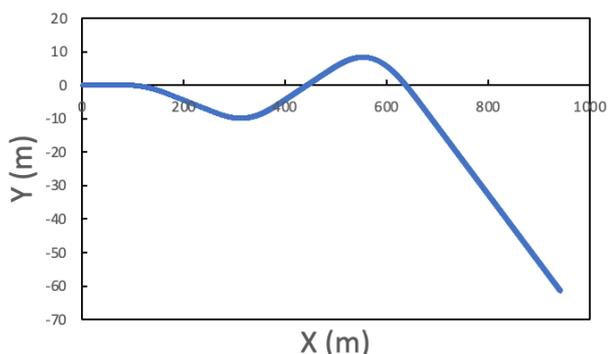


Fig.1. Overview of the Course

遅延に対する予測を形成するために操作主体感を測定する試行の前に遅延を学習させる試行を設定する。学習試行における遅延の期待値は200msとし、測定試行における遅延の期待値が200msの条件を予測誤差が小さい条件、測定試行における遅延の期待値が400msの条件を予測誤差が大きい条件とする。また、学習試行における遅延のばらつきがない条件（標準偏差が0msの条件）を予測不確実性が小さい条件、学習試行における遅延のばらつきがある条件（標準偏差が100ms程度の条件）を予測不確実性が大きい条件とする。ばらつきがある遅延の1試行にお

ける時系列データを図2に示す。

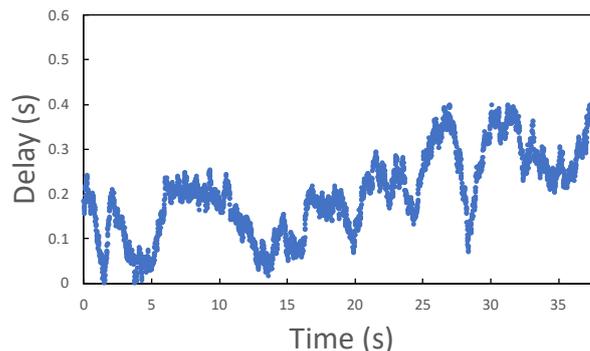


Fig.2. Time-series data with variable delays

操作主体感は質問紙を用いた主観評価で行う。質問は「車両の運動を引き起こしているのはどの程度自分だと感じましたか？」とし、0~100のビジュアルアナログスケールで回答する。

実験参加者は24名とし、東京大学大学院工学系研究科研究倫理審査委員会の承認（承認番号：KE23-9）、および、Hondaの研究開発における生命倫理委員会の承認（承認番号：99HM-067H）を得て実施した。実験手順を下に示す。

1. 奇数番号の実験参加者の場合は、遅延の期待値が200msかつ遅延のばらつきがない条件を、偶数番号の実験参加者の場合は、遅延の期待値が200msかつ遅延のばらつきがある条件を8試行学習する。
2. 遅延の期待値が200msかつ遅延のばらつきがない条件・遅延の期待値が200msかつ遅延のばらつきがある条件・遅延の期待値が400msかつ遅延のばらつきがない条件・遅延の期待値が400msかつ遅延のばらつきがある条件の4条件を1試行ずつ行い、それぞれにおいて操作主体感を測定する。
3. 1,2を行う。
4. 奇数番号の実験参加者の場合は、遅延の期待値が200msかつ遅延のばらつきがある条件を、偶数番号の実験参加者の場合は、遅延の平均が200msかつ遅延のばらつきがある条件を8試行学習する。
5. 2を行う。
6. 4,5を行う。

2.2.4. 解析方法 操作主体感は、実験参加者のパフォーマンスや自分の操作とその結果の因果関係に対する信念によってばらつくと考えられる^{(11),(12)}。実験参加者間の操作主体感の測定値のばらつきを抑えるため、ビジュアルアナログスケールの値を実験参加者毎に正規化したものを操作主体感の測定値とする。そして、操作主体感の測定値に対して、測定試行における遅延の期待値（予測誤差）と学習試行における遅延のばらつき（予測不確実性）を因子とする二元配置分散分析を行う。

3. 結果

実験参加者毎に正規化した操作主体感の測定結果を図3に示す。左図の測定試行における遅延のばらつきがない場合、測定試行における遅延の期待値が増加すると操作主体

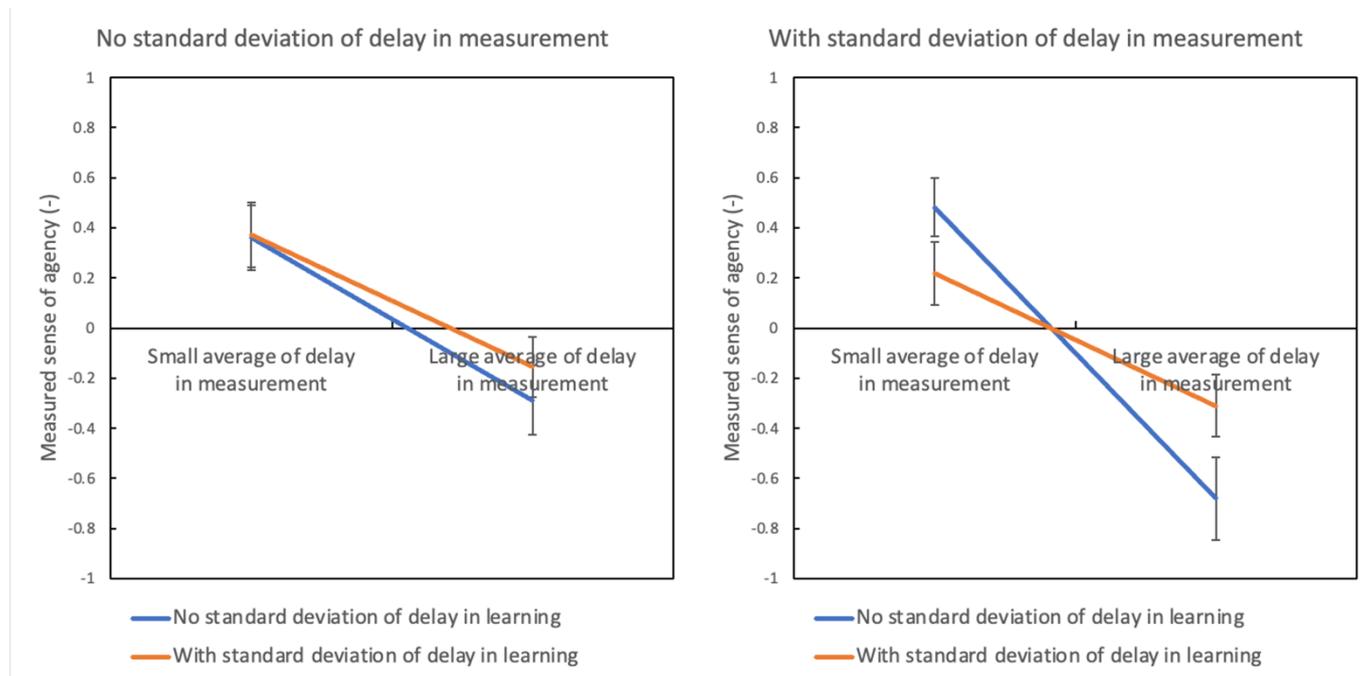


Fig.3. Normalized measurements of sense of agency for each experimental participant

感の測定値は有意に減少した ($p < 0.001$)。右図の測定試行における遅延のばらつきがある場合も、測定試行における遅延の期待値が増加すると操作主体感の測定値は有意に減少した ($p < 0.001$)。また、右図においては、学習における遅延のばらつきが増加すると、測定における遅延の期待値の増加による操作主体感の測定値の減少量は有意に減少した ($p < 0.05$)。

4. 考察

測定試行における遅延のばらつきがある場合の操作主体感の測定値に対して、測定試行における遅延の期待値と学習試行における遅延のばらつきの交互作用が有意に見られたが、これは数理モデルが予測する予測誤差と予測不確実性の交互作用を支持する。一方、測定試行における遅延のばらつきがない場合の操作主体感の測定値に対しては、測定試行における遅延の期待値と学習試行における遅延のばらつきの交互作用は有意には見られなかったが、これは、測定試行において遅延を学習したことが原因であると考えられる。遅延のばらつきがない場合は遅延のばらつきがある場合より遅延を学習しやすいと考えられる。つまり、予測を形成するための学習効果が、遅延のばらつきがない測定試行を通して薄れたと考えられる。

解析では、実験参加者間のばらつきを抑えるため、実験参加者毎に正規化した操作主体感の測定値を用いた。実験参加者間のばらつきは、実験参加者のパフォーマンスや自分の操作とその結果の因果関係に対する信念によるものと考えられる。本研究方法の限界は、実験参加者のパフォーマンスや因果関係に対する信念を、数理モデル上で記述できていないこと、実験でコントロールできていないことである。今後は、因果推定のフレームワーク⁽¹³⁾を用いた数理モデルの拡張による因果関係に対する信念の記述や実験参加者のスキルレベルの測定方法とそれに合わせて実験課題の挑戦レベルを設定する実験方法の開発などの研究

が望まれる。

5. 結 言

数理モデル予測の実験検証を通して、遅延を操作した連続操作系において、予測誤差が小さい場合は予測不確実性が小さいと操作主体感が高く、予測誤差が大きい場合は予測不確実性が大きいと操作主体感は維持されることが分かった。この知見から、操作主体感を維持するためには、予測誤差が小さい場合は予測不確実性を小さく、予測誤差が大きい場合は予測不確実性を大きくすれば良いという設計指針が得られる。本研究のオリジナリティは、連続操作における遅延とそのばらつきが操作主体感に与える影響を数理モデルで予測している点である。

文 献

- (1) 柳澤秀吉, 眞鍋美祈, 上田一貴: 操作フィードバックの感覚様相と期待一致性が操作主体感に与える影響, 設計工学・システム部門講演会講演論文集, Vol. 28, 1201, 2018.
- (2) Gallagher, S.: Philosophical conceptions of the self: implications for cognitive science, Trends in cognitive sciences, Vol. 4, No. 1, 14-21, 2000.
- (3) Haggard, P. and Chambon, V.: Sense of agency, Current biology, Vol. 22, No. 10, 390-392, 2012.
- (4) 番場雅典, 柳澤秀吉: 情報利得を用いた操作主体感の数理モデリング (応答遅れ時間を用いた検証), 日本機械学会論文集, Vol. 87, No. 893, 20-35, 2021.
- (5) Taniyama, K., Maki, T. and Yanagisawa, H.: Modeling Sense of Agency using Free Energy, International Symposium on Affective Science and Engineering, 1-4, 2021.
- (6) Frith, C. D., Blakemore, S. J. and Wolpert, D. M.:

- Abnormalities in the awareness and control of action, *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, Vol. 355, No. 1404, 1771-1788, 2000.
- (7) Friston, K., Kilner, J. and Harrison, L.: A free energy principle for the brain, *Journal of physiology-Paris*, Vol. 100, No. 1-3, 70-87, 2006.
- (8) Blakemore, S. J., Frith, C. D. and Wolpert, D. M.: Spatio-temporal prediction modulates the perception of self-produced stimuli, *Journal of cognitive neuroscience*, Vol. 11, No. 5, 551-559, 1999.
- (9) Sato, A. and Yasuda, A.: Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership, *Cognition*, Vol. 94, No. 3, 241-255, 2005.
- (10) Haering, C. and Kiesel, A.: Was it me when it happened too early? Experience of delayed effects shapes sense of agency, *Cognition*, Vol. 136, 38-42, 2015.
- (11) Wen, W., Yamashita, A. and Asama, H.: The sense of agency during continuous action: performance is more important than action-feedback association, *PloS one*, Vol. 10, No. 4, e0125226, 2015.
- (12) Wegner, D. M., Fuller, V. A. and Sparrow, B.: Clever hands: Uncontrolled intelligence in facilitated communication, *Journal of personality and social psychology*, Vol. 85, No. 1, 5, 2003.
- (13) Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B. and Shams, L.: Causal inference in multisensory perception, *PLoS one*, Vol. 2, No. 9, e943, 2007.