

自由エネルギーを用いた操作主体感の数理モデリング

Mathematical modelling of sense of agency using free energy

○谷山建作（東京大学）*1

柳澤秀吉（東京大学）*2

*1 Kensaku Taniyama, The university of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan, 113-8656,
taniyama2032@mail.design.t.u-tokyo.ac.jp

*2 Hideyoshi Yanagisawa, The university of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan, 113-8656,
hide@mech.t.u-tokyo.ac.jp

キーワード: sense of agency, information theory, interface design

1. 緒 言

主体感とは、観測された物体や身体の運動が自己によって引き起こされていると認識する感覚である⁽¹⁾。本研究では、機械の操作における主体感を操作主体感と定義する。つまり、操作主体感とは、機械の挙動が自己の操作によって引き起こされていると認識する感覚である。例えば、自動車を意のままに操縦している状態は、操作主体感が高い状態である。操作主体感は、操作の楽しさや結果の責任感に影響を与えることが知られている⁽²⁾⁽³⁾。また、操作主体感が高いほど、機械を意のままに操作している感覚を得るため、操作自体が楽しいと感じられる。一方、操作主体感が失われると、自身が操作している感覚が失われるため、操作の結果に対する責任感が失われる。このことから、操作主体感は人間と機械のインタフェースにおいて重要な設計対象である。しかし、現状、操作主体感の設計指針となるモデルは見当たらない。そこで、本研究では、操作主体感の設計指針を確立させることを目的とする。そのために、操作主体感の数理モデルを提案する。数理モデルを開発することにより、一般的な工学応用可能になると考える。

2. 操作主体感の数理モデリング

2.1. コンパレータモデル

主体感を説明するモデルとしてコンパレータモデルがある⁽⁴⁾。例えば、対象に手を伸ばす到達課題において、人間は到達状態を実現するために、運動野より運動指令を出力する。運動指令に従い手を伸ばす一方、運動指令の遠心性コピーをもとに順モデルを用いて手を伸ばした結果を予測する。そして、予測された手の状態と感覚フィードバックを用いて推定された実際の手の状態を比較する。コンパレータモデルによると、比較の結果、予測された状態と実際の状態の誤差が小さいほど主体感が高い。

一方、主体感の因子として、コンパレータモデルによって説明される予測と実際の誤差以外に意図やタスクパフォ

ーマンスの結果などの認知レベルの要因も存在することが知られている⁽⁵⁾。

元来、コンパレータモデルは身体の運動において予測された身体の状態と身体に対する感覚フィードバック（e.g. 自己受容感覚、触覚、身体に対する視覚）を用いて推定された実際の状態の誤差が主体感に影響を与えることを説明するモデルである。ここで、機械を操作する場合、身体に対する感覚フィードバックだけでなく、機械の挙動に対する感覚フィードバック（e.g. 視覚、聴覚、触覚）も存在する。そして、予測された機械の挙動と機械の挙動に対する感覚フィードバックを用いて推定された実際の状態の誤差が生じる。身体における予測と実際の誤差が影響する主体感と身体外における予測と実際の誤差が影響する主体感はそれぞれ、*body agency*、*external agency* として明確に区別されている⁽⁶⁾。*Body agency* と *external agency* は、前述のコンパレータメカニズムによって説明される予測と実際の誤差を因子として持つ。一般的に身体における順モデルは十分に形成されているため、正確な予測が可能である。身体における誤差は小さいことが前提であるため、*body agency* は誤差に敏感である。すなわち、*body agency* において身体における予測と実際の誤差は支配的な因子である。一方、機械の挙動における順モデルは状況によっては十分に形成されていない。この場合、予測が正確でないことが前提であり、*external agency* は誤差に寛容である。つまり、機械に対する順モデルが十分に形成されていない場合の *external agency* においては、予測と実際の誤差よりも認知レベルの因子が支配的である。

本研究では、インタフェース設計に応用可能な知見を得るために、*external agency* に寄与する機械の挙動を対象とする。コンパレータメカニズムにおける誤差が *external agency* に与える影響を調べるために、機械に対する順モデルが十分に形成されると考えられる簡易な実験装置を用いる。また、認知レベルの因子の影響を排除するために、タスクの達成が意味を持たない実験装置を扱う。

2.2. 数理モデリング

予測と実際の誤差が大きいということは、起こりにくいことが起こったことを意味すると考える。例えば、ボタンを押すと音がフィードバックされる実験において、ボタン押しから感覚フィードバックまでの時間である応答遅れ時間を操作し、操作主体感を評価する方法がある⁽⁷⁾。応答遅れ時間が 0ms である条件を学習させた後応答遅れ時間が 0ms から 250ms までの条件で操作主体感を評価するグループは、応答遅れ時間が長くなるほど操作主体感は低下する。一方、応答遅れ時間が 250ms である条件を学習させた後応答遅れ時間が 0ms から 250ms までの条件で操作主体感を評価するグループは、応答遅れ時間が短くなるほど操作主体感は低下する。応答遅れ時間が 0ms の条件を学習したグループにとって応答遅れ時間が 250ms の条件は起こりにくい事象である。一方、応答遅れ時間が 250ms の条件を学習したグループにとって応答遅れ時間が 0ms の条件は起こりにくい事象である。そこで、操作主体感の尺度として、ある事象が起きた際それがどれほど起こりにくいかを表現する情報理論における情報量を用いることを提案する。操作主体感の尺度を式(1)のように表現する。

$$SoA \propto -I(s) = -(-\log p(s)) \quad (1)$$

s は感覚フィードバックであり、 $I(s) = -\log p(s)$ は感覚フィードバックを受け取った際に獲得する情報量である。情報量が小さいほど操作主体感が高いことを表す。

ここで、感覚フィードバックを用いてある外界の状態を推定する問題を考える。人間の認識のモデルとしてベイズモデルがある⁽⁸⁾。ベイズモデルによると、人間は感覚フィードバックをもとに状態を推定する際、事前の予測する状態と感覚フィードバックより得られる尤もらしい状態の重み付け和を認識する。ベイズモデルを式(2)に示す。

$$p(x|s) = \frac{p(s|x)p(x)}{p(s)}, p(s) = \int p(s|x)p(x)dx \quad (2)$$

x は状態変数であり、感覚フィードバックを用いて推定する対象である。 $p(s|x)$ は尤度関数であり、感覚フィードバックに対応する状態の尤もらしさを表す、状態を変数とする関数である。 $p(x)$ は事前分布であり、感覚フィードバックを受け取る前に予測する状態変数が従う分布である。 $p(x|s)$ は事後分布であり、感覚フィードバックを受け取った後に認識する状態変数が従う分布である。式(2)のように、尤度関数と事前分布の積が事後分布に比例する。

ベイズモデルを用いると、感覚フィードバックを受け取った際に獲得する情報量 $-\log p(s)$ は式(3)のように表される。

$$-\log p(s) = E_{p(x|s)}(-\log p(x, s)) - H(p(x|s)) \quad (3)$$

第一項は内部エネルギー、第二項はエントロピーの形式で表されることから、この情報量は熱力学のアナロジーより自由エネルギーと呼ばれる⁽⁹⁾。式(1)より、操作主体感の尺度は負の自由エネルギーによって表される。

2.3. モデル予測

ここで、2.2.で提案した操作主体感の数理モデルより、モデル予測を導出する。式(4)のように、尤度関数と事前分布を正規分布で近似する。

$$p(s|x) = N(\mu_l, \sigma_l^2), \quad p(x) = N(\mu_p, \sigma_p^2) \quad (4)$$

μ_l は感覚フィードバックに対応する尤もらしい状態変数の期待値、 σ_l^2 は感覚フィードバックの観測における分散、 μ_p は予測する状態変数の期待値、 σ_p^2 は予測する状態変数の分散である。ここで、 $|\mu_l - \mu_p|$ を予測誤差、 σ_l^2 を尤度の分散、 σ_p^2 を不確実性と呼ぶ。操作主体感の提案尺度は式(5)のように表現される。

$$-(-\log p(s)) = -\frac{1}{2} \left(\frac{1}{\sigma_l^2 + \sigma_p^2} |\mu_l - \mu_p|^2 + \log 2\pi(\sigma_l^2 + \sigma_p^2) \right) \quad (5)$$

式(5)をもとに予測誤差、尤度の分散、不確実性を変化させた時の操作主体感の変化を予測する。

予測誤差と尤度の分散が操作主体感に与える影響についてのモデル予測を図1に示す。予測誤差が小さい条件において、尤度の分散が小さいほど操作主体感が高く、予測誤差が大きい条件において、尤度の分散が大きいほど操作主体感が高い。

次に、予測誤差と不確実性が操作主体感に与える影響についてのモデル予測を図2に示す。予測誤差が小さい条件において、不確実性が小さいほど操作主体感が高く、予測誤差が大きい条件において、不確実性が大きいほど操作主体感が高い。予測誤差と不確実性が操作主体感に与える影響についてのモデル予測は番場、柳澤⁽¹⁰⁾が行なった研究の結果と一致する。

本研究では、未検証である、予測誤差と尤度の分散が操作主体感に与える影響についてのモデル予測を実験検証する。すなわち、予測誤差が小さい条件において、尤度の分散が小さいほど操作主体感が高く、予測誤差が大きい条件において、尤度の分散が大きいほど操作主体感が高いことを実験検証する。

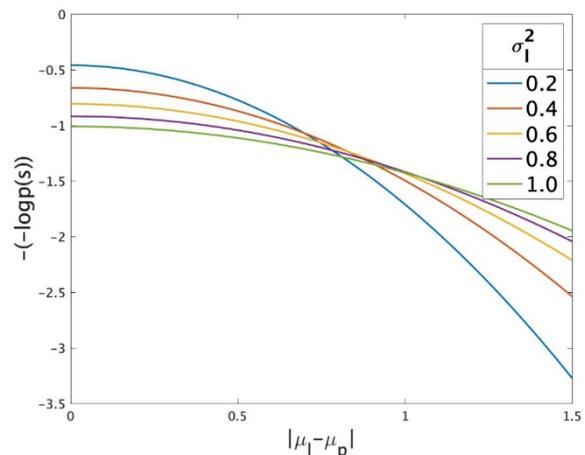


Fig.1 Sense of agency scale as a function of prediction error for different likelihood variance

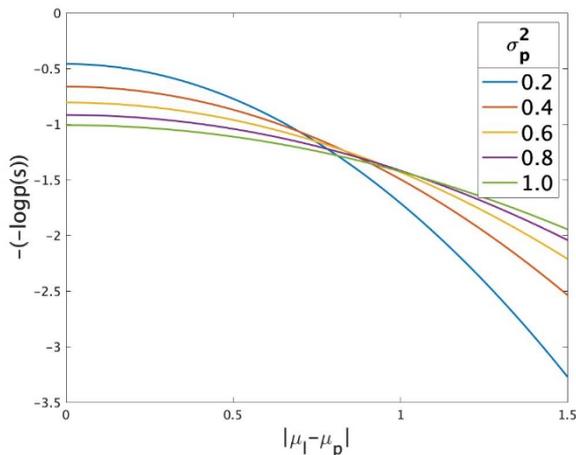


Fig.2 Sense of agency scale as a function of prediction error for different uncertainty

2.4. モデルパラメータの操作方法

2.4.1. 予測誤差の操作方法 (応答遅れ時間)

操作主体感の実験として、ボタン押しから感覚フィードバックまでの時間である応答遅れ時間を操作し、操作主体感を評価する実験がある⁽⁷⁾。この実験において、状態変数 x は応答遅れ時間である。予測する状態変数の期待値 μ_p は、ある応答遅れ時間を学習させることにより操作可能である。感覚フィードバックに対応する尤もらしい状態変数の期待値 μ_1 は、評価段階における感覚フィードバックの応答遅れ時間により操作可能である。本実験では、ボタンを押すと感覚がフィードバックされる装置を用いて、応答遅れ時間 0ms の条件を学習させた後、応答遅れ時間 0ms から 500ms の条件において評価させることにより、予測誤差を操作する。

2.4.2. 尤度の分散の操作方法 (感覚モダリティ数)

複数モダリティの感覚統合のモデルとして、cue integration model がある⁽¹¹⁾。Cue integration model によると、異なるモダリティの感覚が同じ特徴を参照する時、人間は複数モダリティの感覚を統合することにより、感覚による推定の分散を減少させ感覚の信頼度を高める。以上より、感覚フィードバックのモダリティ数を用いて尤度の分散を操作する。

2.4.3. 実験仮説

モデル予測は、予測誤差が小さい条件において、尤度の分散が小さいほど操作主体感が高く、予測誤差が大きい条件において、尤度の分散が大きいほど操作主体感が高いことである。そして、予測誤差を応答遅れ時間を用いて操作し、尤度の分散を、感覚モダリティ数を用いて操作することを考えた。実験仮説は、応答遅れ時間 0ms の条件を学習させた時、応答遅れ時間が小さい条件において、感覚モダリティ数が多いほど操作主体感が高く、応答遅れ時間が大きい条件において、感覚モダリティ数が少ないほど操作主体感が高いことである。

3. モデル予測の検証実験

3.1. 実験目的

実験目的は、モデル予測である、予測誤差が小さい条件において、尤度の分散が小さいほど操作主体感が高く、予

測誤差が大きい条件において、尤度の分散が大きいほど操作主体感が高いことの検証である。

3.2. 実験方法

感覚モダリティ数を用いて尤度の分散を操作し、応答遅れ時間を用いて予測誤差を操作する。操作主体感は、質問紙を用いて主観評価した。質問紙は、「自分のボタン操作によってフィードバックが来たと思いますか。」という質問に対して 0%から 100%で回答するものである。実験装置として、ボタンを押すと応答遅れ時間後に音や振動がフィードバックされるものを用いる。応答遅れ時間は 0, 50, 100, 150, 200, 250, 300, 400, 500ms の 9 条件、感覚モダリティは音 (S)、振動 (V)、音と振動 (SV) の 3 条件、合計 27 条件とした。実験は、S ブロック、V ブロック、SV ブロックにより構成される。そして、それぞれのブロックは、学習段階、評価段階により構成される。学習段階で、応答遅れ時間 0ms の条件において 2 分間自由にボタンを押し感覚フィードバックを受け取る。評価段階で、応答遅れ時間 9 条件それぞれにおいて 2 回試行、質問紙回答を 5 セット行う。つまり、1 条件あたり 5 回の操作主体感の評価、合計 27 条件において 135 回の操作主体感の評価を行う。被験者は 18 名であった。感覚モダリティの条件 (ブロック) の順番はカウンターバランスが取られ、評価段階における応答遅れ時間の条件の順番はランダム化された。

3.3. 実験結果

応答遅れ時間と感覚モダリティのそれぞれの条件における操作主体感の評価を図 3 に示す。横軸は評価段階における応答遅れ時間、縦軸は操作主体感の主観評価、凡例は感覚モダリティの条件である。条件ごとに操作主体感の評価の平均を計算した。表 1 に感覚モダリティと応答遅れ時間を要因とする二元配置分散分析の結果を示す。応答遅れ時間が大きくなるほど操作主体感の評価は低下している (応答遅れ時間の効果 $P < 0.001$)。次に、応答遅れ時間 0ms の条件において、感覚モダリティが S や V より SV の条件の時の方が操作主体感の評価は高く、応答遅れ時間 500ms の条件において、感覚モダリティが SV より S や V の条件の時の方が操作主体感の評価は高い (感覚モダリティと応答遅れ時間の交互作用 $P < 0.01$)。

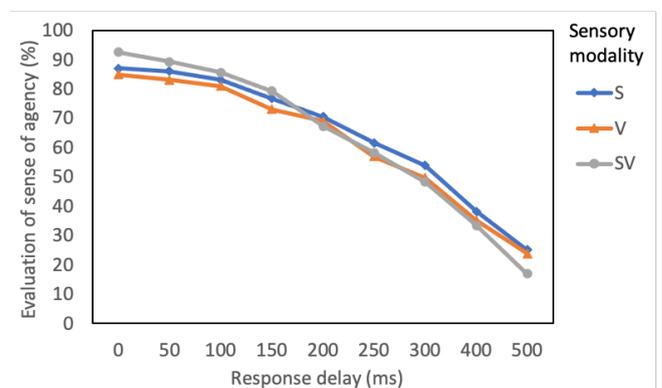


Fig.3 Evaluation of sense of agency for different response delay and sensory modality (S, V and SV indicate sound, vibration and sound-vibration respectively)

Table.1 Effect of sensory modality and response delay and interaction between sensory modality and response delay on evaluation of sense of agency

Factor	Type III Sum of Squares	df	Mean Square	F	P-value
Sensory modality	3306.7490	2	1653.3745	4.6374	0.0098
Response delay	1187298.6008	8	148412.3251	416.2674	$P < 0.001$
Sensory modality * Response delay	11508.8066	16	719.3004	2.0175	0.0095
Error	856744.4444	2403	356.5312		
Total	2058858.6008	2429			

4. 考察

応答遅れ時間が大きくなるほど操作主体感の評価は低下した。この結果は、先行研究の結果と一致する⁽⁷⁾⁽¹⁰⁾。この結果より、学習段階における2分間の応答遅れ時間0msの条件の学習が十分であったと考える。つまり、予測する応答遅れ時間の期待値 μ_p と不確実性 σ_p^2 は十分固定することができた。そして、評価段階における応答遅れ時間の期待値 μ_l を操作することにより、予測誤差 $|\mu_l - \mu_p|$ を操作することができた。

次に、応答遅れ時間0msの条件において、感覚モダリティ数が多い方が操作主体感の評価が高く、応答遅れ時間500msの条件において、感覚モダリティ数が少ない方が操作主体感の評価が高いことが分かった。この結果は、提案した数理モデルから導出されるモデル予測(図1)と一致する。すなわち、この結果は、提案した操作主体感の数理モデルを支持する。

また、感覚モダリティがSとVの条件と応答遅れ時間の条件の交互作用は確認されなかった(図3)。この結果から、音と振動は、応答遅れ時間の推定において感覚の信頼度(尤度の分散)は同程度であると考えられる。

5. 結 言

本研究では、情報論的自由エネルギーを用いた操作主体感の数理モデルを提案した。そして、数理モデルより、予測誤差と尤度の分散が操作主体感に与える影響についてのモデル予測を導出した。予測誤差を応答遅れ時間を用いて操作し、尤度の分散を感覚モダリティ数を用いて操作し、操作主体感を主観評価することにより、モデル予測を検証した。実験結果はモデル予測と一致した。すなわち、応答遅れ時間が小さい条件において、感覚モダリティ数が多いほど操作主体感は高く、応答遅れ時間が大きい条件において、感覚モダリティ数が少ないほど操作主体感が高いことが分かった。

本研究の設計応用として、操作主体感を高めるためには、応答遅れ時間が小さい機械においては感覚モダリティ数を多くする、応答遅れ時間が大きい機械においては感覚モダリティ数を少なくするという設計指針が得られる。

本研究では、予測誤差の操作方法として応答遅れ時間、尤度の分散の操作方法としてcue integration modelをもとに感覚モダリティ数を用いた。今後の課題は、予測誤差の操作方法として空間的な誤差や質的な誤差を用いること、尤度の分散の操作方法としてcue integration modelに依らない感覚フィードバックに対するノイズなどを用いることによる、提案した数理モデルの一般性の検証である。また、本

研究は、ボタン操作という離散操作を対象としたものであったが、ハンドル操作などの連続操作も対象とした数理モデルへの発展が必要である。

文 献

- (1) Haggard, P.: Sense of agency in the human brain, *Nature Reviews Neuroscience*, Vol. 18, No. 4, 196, 2017.
- (2) Karsh, N. and Eitam, B.: I control therefore I do: judgements of agency influence action selection, *Cognition*, Vol. 138, 122-131, 2015.
- (3) Moretto, G., Walsh, E. and Haggard, P.: Experience of agency and sense of responsibility, *Consciousness and Cognition*, Vol. 20, No. 4, 1847-1854, 2011.
- (4) 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.
- (5) Synofzik, M., Vosgerau, G. and Newen, A.: Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and Cognition*, Vol. 17, No. 1, 219-239, 2008.
- (6) Wen, W.: Does delay in feedback diminish sense of agency? A review. *Consciousness and Cognition*, Vol. 73, 102759, 2018.
- (7) 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.
- (8) Yanagisawa, H., Kawamata, O. and Ueda, K.: Modeling emotions associated with novelty at variable uncertainty levels: a bayesian approach, *Frontiers in Computational Neuroscience*, Vol. 13, 2, 2019.
- (9) Friston, K.: The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, Vol. 11, No. 2, 127-138, 2010.
- (10) 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, 20-00035, 2021.
- (11) Moore, J. W. and Fletcher, P. C.: Sense of agency in health and disease: a review of cue integration approaches, *Consciousness and Cognition*, Vol. 21, No. 1, 59-68, 2012.