



ベイズ因果推論を用いたハンドリダイレクションにおける 多感覚統合のモデル化

村山皓平¹⁾, 松本啓吾^{2,3)}, 鳴海拓志¹⁾, 葛岡英明¹⁾

1) 東京大学 (〒 113-0033 東京都文京区本郷 7-3-1, {murayama, narumi, kuzuoka}@cyber.t.u-tokyo.ac.jp)

2) 筑波大学 (〒 305-8577 茨城県つくば市天王台 1-1-1, matsumoto@iit.tsukuba.ac.jp)

3) 産業技術総合研究所 (〒 305-8560 茨城県つくば市梅園 1-1-1)

概要: ベイズ因果推論 (BCI) モデルは, 従来の感覚統合モデルとは異なり, 感覚情報を統合すべきか否かを判断した上で推定を行うため, 多感覚知覚のメカニズムやその個人差をより正確に説明できると期待されている. 本研究では, 視覚と固有感覚の統合を利用したハンドリダイレクション (HR) という手法に着目し, HR 適用時の手の位置推定データに対してモデルフィッティングを行うことで, BCI モデルの有効性を検証した.

キーワード: ベイズ因果推論, ハンドリダイレクション, 多感覚統合, バーチャルリアリティ

1. はじめに

ベイズ因果推論モデル [1] (BCI モデル) は感覚情報間の因果関係を考慮した感覚統合モデルである. 従来のモデルは多感覚統合を前提としているが, BCI モデルは統合を仮定していない. さらに, BCI モデルは感覚尤度の標準偏差などの感覚入力に基づく要因と, 感覚情報の提示位置や因果構造に対する信念などの認知的要因を併せ持つ. これにより, BCI モデルは従来のモデルと比較して, 多感覚統合を精度よく予測できると考えられる. 視覚や固有感覚に関連する錯覚現象であるラバーハンド錯覚については, Samad ら [2] による BCI モデルを用いた研究が既に行われている. しかし, ハンドリダイレクション (HR) のように運動を伴う視覚や固有感覚に関する錯覚においても, このモデルが適用可能であるかは明らかではない.

さらに, HR における知覚閾値は感覚入力の精度といったボトムアップの感覚的要因と物体の位置や姿勢の予測といったトップダウンの認知的要因の双方の影響を受ける可能性が示唆されている. そのため, 感覚的要因と認知的要因をパラメータとして持つ BCI モデルによる HR のモデル化が有効であると考え, その適用可能性を検証する.

2. 関連研究

2.1 最尤推定モデル

Ernst と Banks [3] は多感覚統合を説明するための最尤推定 (MLE) モデルを提案した. このモデルでは, 各モダリティの感覚推定値をその推定分散の逆数で重み付けし統合する. 各モダリティにおける推定値を $\hat{X}_i$, 実環境の状態を S_i とし, 神経系による推定演算を f とし, この演算で分散 σ_i^2 を持つ独立なガウスノイズが $\hat{X}_i$ に付与されると仮定する. このとき統合推定値 $\hat{X}$ は式 1 のように表される.

$$\hat{X} = \sum_i w_i \cdot f(S_i) \quad (1)$$

ここで, 重み w_i は推定の分散を最小化するように決定され, 式 2 で示される.

$$w_i = \frac{\frac{1}{\sigma_i^2}}{\sum_j \frac{1}{\sigma_j^2}} \quad (2)$$

したがって, 統合推定値は各モダリティの推定値とその分散に依存して決定される.

2.2 ベイズ統合モデル

最尤推定モデルは感覚の信頼度に基づいて統合を決定する. しかし知覚には経験に基づく事前信念など認知的要因も関与すると考えられる. Ernst [4] はこの点を踏まえベイズ統合モデルを提案した. 真の環境状態を S , 感覚入力を X とすると事後確率は式 3 で与えられ, 事後分布 $P(S | X)$ を最大化することで環境状態が推定される.

$$P(S | X) = \frac{P(X | S)P(S)}{P(X)} \propto P(X | S)P(S) \quad (3)$$

2.3 ベイズ因果推論モデル

従来のモデルはいずれも感覚入力の統合を前提とする. しかし腹話術錯覚やゴム手袋錯覚などで, 感覚信号源の距離が大きく乖離すると統合が生じない現象が報告されている. Kording ら [1] は感覚信号間の因果構造の推定を伴うベイズ因果推論モデルを提案した. 図 1 のように, 因果構造を C で表し, $C = 1$ が共通原因, $C = 2$ が独立原因を表し, 感覚入力を X すると, 因果構造の事後確率は式 4 で求められ, この事後確率を用いることで, 感覚位置の推定値が計算される.

$$p(C = 1 | X) = \frac{p(X | C = 1)p(C = 1)}{p(X | C = 1)p(C = 1) + p(X | C = 2)p(C = 2)} \quad (4)$$

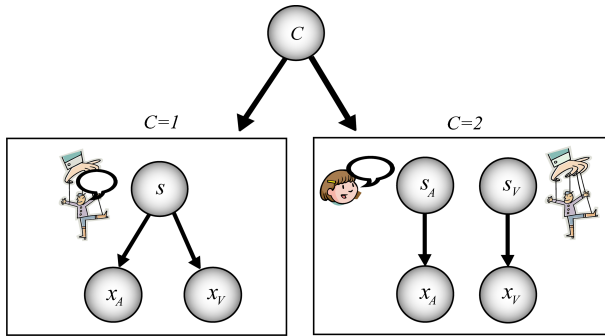


図 1: ベイズ因果推論のイメージ [1]. $C=1$ が共通原因, $C=2$ が独立原因を表す.

2.4 ハンドリダイレクション

ハンドリダイレクション [5] は VR 環境においてユーザに気付かれないようにバーチャルハンドの位置を実際の手の位置からずらす手法である. この手法は, 人間の知覚における視覚優位性を利用することで実現されており, 触覚提示やリハビリテーションなどへの応用が期待されている.

ハンドリダイレクションの適用において重要なのは, ユーザに悟られないように操作を適用し, ユーザの没入感や快適さを損なわないことである. ユーザが気づかない最大の操作量は知覚閾値, 操作に気付いた後でも許容できると感じる最大の操作量は許容閾値と呼ばれる.

3. 実験

3.1 実験目的

BCI モデルを用いたフィッティングを行うには, ユーザの手とバーチャルハンドの真の位置と, ユーザが報告した両者の位置が必要である. そこで, 本実験では, モデルフィッティングに必要なデータを空間定位課題で取得するとともに, 知覚閾値と許容閾値を取得するための閾値測定課題も実施した.

本実験は, 東京大学大学院情報理工学系研究科倫理審査委員会の承認を得て行われた.

3.2 実験参加者および実験環境

実験参加者は視覚と腕部の動作に障害を持たない男女 12 名 (男性 8 名, 女性 4 名; 平均年齢 22.0 歳, ($SD = 1.0$)) であった. 全ての実験参加者は実験に関する事前知識を持たなかった.

実験では, 被験者の手を所定の位置へ移動するため, 力覚提示デバイス 3D Systems® Touch¹ を使用した. 本装置は 6 自由度 (6-DoF) の位置・姿勢計測と 3 自由度 (3-DoF) の力覚フィードバックを提示できる.

VR 環境の提示には, HTC VIVE Pro Eye と Meta Quest 3 およびコントローラを使用し, 実験環境の開発には Unity² (ver.2019.3.6f1) を用いた.



図 2: リダイレクション操作のイメージ.

黒色の球体はリダイレクションが適用された指先の位置 ($\vec{p}_v$), 黄色の球体は実際の人差し指の位置 ($\vec{p}_r$) を示し, $\vec{p}_r$ は実験中には表示されなかった.

3.3 ハンドリダイレクションの実装

HR の実装手法には Zenner と Krüger の Gain Warp [5] を採用した. デバイスを用いたリーチング中にハンドトラッキングを行うと手がデバイスに隠れてしまい正確な描画が行われない可能性がある. これを避けるため, デバイスの座標系から VR 空間上に座標変換を行い, 人差し指の先端のみを小さな球体として表示した. 小球の位置 $\vec{p}_v$ は式 5 のように計算される.

$$\vec{p}_v = \vec{o} + g(\vec{p}_r - \vec{o}), \quad (5)$$

ここで, $\vec{p}_r$ は実際の人差し指の位置, $\vec{o}$ はワープの原点をそれぞれ表す. 実験中, $\vec{p}_r$ や $\vec{o}$ は表示せず, VR 環境にはリダイレクション適用後の球 $\vec{p}_v$ のみを呈示した (図 2).

3.4 実験条件

3.4.1 空間定位課題

この課題は, 視覚単条件, 固有感覚単条件, 2 感覚条件の 3 つの条件で構成される. 全ての課題において, 条件に応じて実際の手, バーチャルハンド, またはその両方の手が, 5 段階の目標距離 (230, 260, 290, 320, 350 mm) のいずれかへ動かされた. 各距離は 8 回ずつテストされた.

視覚単条件: 参加者は右手をテーブル上に置いた状態で, VR 内には仮想的な指先を表す球体のみが表示された. 球体は 5 つの目標位置のいずれかへリーチング運動を行った. その後, 参加者は左手のコントローラを使い, 記憶していた球体の到達位置に球体を移動させ, トリガーを押してその位置を記録した.

固有感覚単条件: この条件では, 指先を表す球体は一切表示されなかった. Touch が参加者の手を目標位置まで動かした後, 参加者は Touch のスタイラスを自身が感じた到達位置まで動かし, トリガーを押すことで位置を記録した.

2 感覚条件: Touch が参加者の手を動かすと同時に, 3.3 節で述べた手法でリダイレクトされた球体が表示された. その後, 参加者は両方の定位課題を遂行した. 2 つの目標距離は単条件と同様の 5 段階の距離からそれぞれ独立に抽出さ

¹<https://www.3dsystems.com/haptics-devices/touch>

²<https://unity.com/>

れ、計 25 通りの条件がランダムに提示された。

3.4.2 閾値測定課題

参加者はリーチング運動を行い、バーチャルハンドには、操作ゲイン（0.50 から 1.50 まで 0.1 刻みの 11 段階）に応じたリダイレクションが適用された。各リーチングの後、参加者は「動きは一致したか（知覚閾値）」と「その動きは許容できるか（許容閾値）」という 2 つの質問に、一肢強制選択（1AFC）形式で回答することが求められた。

この一連の試行（各ゲイン 10 回、計 110 試行）を、2 つの異なる運動条件下で実施した。まず受動リーチング条件では、デバイスが参加者の手を動かし、VR 内では指先を示す球体のみが表示された。さらに、デバイスを用いた定位データから構築したモデルが、より一般的な HR 適用時の閾値についても検証するために、能動リーチング条件を実施した。この条件では、参加者が自発的に手を動かし、その動きがハンドトラッキングによって捉えられた。

4. 結果と考察

4.1 モデルフィッティング

本研究で用いた BCI モデルは、以下の 7 つの自由パラメータを持つ：

- P_{common} : 視覚と固有感覚入力が共通の原因から生じる事前確率
- σ_v : 視覚の標準偏差
- σ_p : 固有感覚の標準偏差
- γ_v : 視覚尤度の平均に追加される定数
- γ_p : 固有感覚尤度の平均に追加される定数
- σ_s : 感覚事前分布の標準偏差
- μ_s : 感覚事前分布の平均

BCI モデルフィッティングには BCIToolbox [7] を用いた。全被験者のデータに対してモデルを適合させたところ、高い適合度（ $R^2 = 0.8977$ ）を示したが、個人ごとのデータに対する適合度は中程度（ $R^2 = 0.6565 \pm 0.1440$ ）であった。この結果は、モデルが参加者の平均的な定位傾向を適切に捉えている一方で、各参加者の試行数が不十分であり、個人差を完全に説明するには至らなかったことを示唆している。一方で、実験参加者の疲労を考慮すると試行回数の増加は定位精度の低下につながる恐れがあり、少ないデータに対する高精度のフィッティング手法が求められる。

4.2 閾値測定課題

本実験では視覚の標準偏差（ σ_v ）と固有感覚の標準偏差（ σ_p ）の大きさがほぼ同程度であり、両感覚の精度差は小さいという前提が成り立つ。このとき σ_p と σ_v の統合後の標準偏差 σ_c は式 6 のように単調に $\sigma_v^2 + \sigma_p^2$ （すなわち分散の和）に依存して増減する。

$$\sigma_c = \sqrt{\frac{\sigma_v^2 \sigma_p^2}{\sigma_v^2 + \sigma_p^2}} \approx \sqrt{\frac{\sigma^4}{2\sigma^2}} = \frac{\sigma}{\sqrt{2}} \quad (6)$$

表 1: 各参加者へのモデルフィッティングの結果

パラメータ	平均 $\pm$ SD	単位
P_{common}	0.358 $\pm$ 0.355	
σ_v	45.89 $\pm$ 34.63	mm
σ_p	61.27 $\pm$ 29.25	mm
σ_s	1451.75 $\pm$ 1341.01	mm
μ_s	255.24 $\pm$ 22.34	mm
γ_v	26.82 $\pm$ 4.56	mm
σ_p	3.25 $\pm$ 0.89	mm

閾値幅と二つの推定パラメータ（共通原因の事前確率 P_{common} と統合感覚の標準偏差 σ_c ）との相関関係を表 2 に示す。

相関解析の検出力は十分ではなかったが、受動リーチング条件において二つの注目すべき傾向が得られた。

第一に、統合感覚の標準偏差（ σ_c ）が大きい参加者ほど、知覚閾値幅が広がる傾向が確認された。この結果は、式??のように表される BCI モデルの独立原因に対する共通原因のオッズ近似式 [6]

$$\begin{aligned} \log \frac{P(C=1 | X_v, X_p)}{P(C=2 | X_v, X_p)} &= \log \frac{P(C=1 | X_v, X_p)}{1 - P(C=1 | X_v, X_p)} \\ &= \log \frac{P_{\text{common}}}{1 - P_{\text{common}}} - \frac{(X_v - X_p)^2}{2(\sigma_v^2 + \sigma_p^2)} \end{aligned} \quad (7)$$

$$(8)$$

に照らすと、分散和（ $\sigma_v^2 + \sigma_p^2$ ）が大きいほど第 2 項のペナルティが弱まり、相対的に $\log P(C=1 | X_v, X_p)$ が高く維持されるため、視覚手掛かりと固有感覚手掛かりを「共通原因に由来する」と判断しやすくなる。結果として、固有感覚を通じて感じられる手の位置と目で見える手の位置が離れていてもそのずれに気づきづらくなり、知覚閾値幅が広がると解釈できる。

第二に、許容閾値幅は共通原因事前確率 P_{common} と中程度の正の相関を示した（受動: $r = 0.401$, 能動: $r = 0.304$ ）。 P_{common} が高い被験者は手の位置のずれをリーチングを行う前から想定しない傾向があり、大きな操作であっても許容する可能性が示された。

一方、能動リーチングでは σ_c と知覚閾値幅の相関は弱く（ $r = 0.164$, $p = .629$, $N = 12$ ）、図 3, 4 の通り知覚・許容閾値幅のいずれも受動リーチング条件より有意に狭かった（Wilcoxon, $p < .05$, $r \geq 0.840$ ）。能動リーチングと受動リーチングとでは感覚情報の信頼度や事前分布が異なるため、受動リーチング条件のデータで構築したモデルでは、能動リーチング条件下の知覚をうまく説明できなかったと考えられる。

視覚の標準偏差は通常約 0.36° であり、本実験では 1, 2mm に相当する。固有感覚の標準偏差は通常約 10mm であり、本研究で推定された視覚および固有感覚の標準偏差は、これらを大幅に上回る結果であった。主な要因として、手全体ではなく指先のみを描画したことによる視覚情報量の低下、指先

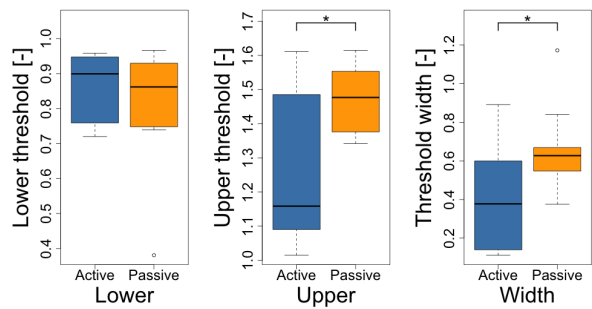


図 3: 能動リーチング条件と受動リーチング条件での知覚閾値の比較

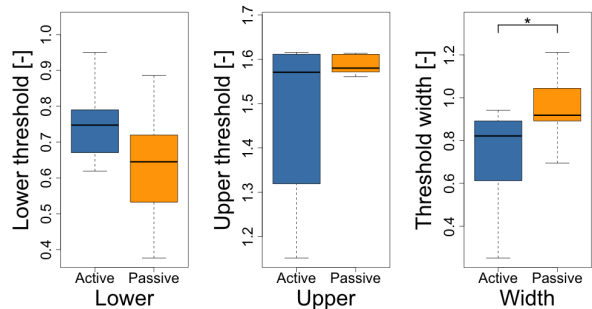


図 4: 能動リーチング条件と受動リーチング条件での許容閾値の比較

の位置を VR 空間上に描画する際の誤差,そして受動的な運動による固有感覚の信頼性低下が挙げられる。これらの影響が,特に受動リーチング条件において観測された感覚精度の低下と,それに伴う広い知覚閾値をもたらしたと推測される。

5. むすびに

本研究は,ベイズ因果推論 (BCI) モデルを用いて Hand Redirection (HR) 下の多感覚知覚を解析し,その妥当性と

表 2: 各条件での閾値幅とパラメータ間のピアソン相関係数, 95 %信頼区間, p 値.

条件	パラメータ	r	95 % CI	p
受動リーチング				
知覚閾値	σ_c	0.610	[-0.17, 0.92]	.108
	P_{common}	0.250	[-0.55, 0.81]	.550
許容閾値	σ_c	0.051	[-0.57, 0.63]	.881
	P_{common}	0.401	[-0.26, 0.81]	.222
能動リーチング				
知覚閾値	σ_c	0.164	[-0.48, 0.70]	.629
	P_{common}	0.304	[-0.36, 0.77]	.363
許容閾値	σ_c	0.153	[-0.49, 0.69]	.653
	P_{common}	0.304	[-0.36, 0.77]	.363

今後の課題を検討した。本実験で得られた空間定位データに BCI モデルを適用したところ,参加者全体の行動は高い精度 ($R^2 = 0.8977$) で説明できた。一方,個人レベルの適合度は中程度 ($R^2 = 0.6565 \pm 0.1440$) にとどまり,1 条件あたりの試行数が少ないため,個人差を捉えるにはデータ量が不足していたことが示唆された。さらに,推定されたパラメータと知覚閾値および許容閾値の相関を分析したところ,受動的なリーチングでは一部の相関が確認されたが,一般的なリーチング運動への汎化性能は限定的であった。

今後の課題として,モデルフィッティング手法と実験系の改良が必要だろう。フィッティング手法については,HR の実験で許容される少数試行からでも精度を保てる階層ベイズ推定や変分推論に加え,行動と知覚を同時に説明できる能動的推論を用いたモデル化が有効であると考ええる。また,実験系は受動刺激を再現させるのではなく,実験参加者が能動的に手を動かしながら視覚・固有感覚刺激の定位を行えるような実験系を構築することが必要であるだろう。

本研究の知見が,将来的には個人の感覚特性に合わせて表示を最適化する,適応的な VR/AR システムの設計に貢献することを期待する。

謝辞 本研究は, JST さきがけ (JPMJPR22S9), JST BOOST (JPMJBY24H1), JST ACT-X (JPMJAX22AG), JSPS 科研費 (25K03162, 23H00076) の支援を受けた。

参考文献

[1] Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., Shams, L.: Causal Inference in Multisensory Perception, PLOS ONE, Vol. 2, No. 9, pp. 1–10, 2007.

[2] Samad, M., Chung, A. J., Shams, L.: Perception of Body Ownership Is Driven by Bayesian Sensory Inference, PLOS ONE, Vol. 10, No. 2, pp. 1–23, 2015.

[3] Ernst, M., Banks, M.: Humans integrate visual and haptic information in a statistically optimal fashion, Nature, Vol. 415, pp. 429–433, 2002.

[4] Ernst, M.: A Bayesian view on multimodal cue integration, Human Body Perception from the Inside Out, pp. 105–131, 2006.

[5] Zenner, A., Krüger, A.: Estimating Detection Thresholds for Desktop-Scale Hand Redirection in Virtual Reality, Proc. of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 47–55, 2019.

[6] Ma, W. J., Rahmati, M.: Towards a neural implementation of causal inference in cue combination, Multisensory Research, Vol. 26, Nos. 1–2, pp. 159–176, 2013.

[7] Zhu, H., Beierholm, U., Shams, L.: BCI Toolbox: An open-source python package for the Bayesian causal inference model, PLOS Computational Biology, Vol. 20, No. 7, pp. 1–12, 2024.