

曖昧性解消における視覚的注意へのトップダウン介入

A Top-down Penetration to the Visual Attention via Elimination of the Ambiguity

内海佑麻 *1
Yuma Uchiumi

福地庸介 *1
Yosuke Fukuchi

木本充彦 *1
Mitsuhiko Kimoto

今井倫太 *1
Michita Imai

*1慶應義塾大学 理工学研究科

Graduate School of Science and Technology, Keio University

In order to elucidate human perceptual functions, it is necessary to consider both bottom-up information processing, in which stimulus information received from the sensory organs is encoded into symbolic information, and top-down information processing, which is objective-oriented and based on memory, beliefs, and context. In this paper, we take the ResNet50 image classification problem as an example task, and conduct a basic study on the information processing when humans make judgments about visual information with ambiguities, and discuss the computation by which beliefs during task execution penetrates the discrimination results of the model in a top-down manner.

1. はじめに

人・生物・ロボットといった主体 (Agent) がおこなっている情報処理を考える際、知能システムと外部世界との間のインターフェースとして知覚 (Perception) と行為 (Act, Action) は重要な意味をもつ。本研究のモチベーションは、主体が知覚するときに作動している適応的な情報処理の解明である。

人がもつ知覚機能を正しく捉えるためには、感覚器官から受け取る刺激が意識表象へと符号化されるボトムアップ型の情報処理と、記憶・信念・文脈に基づいて刺激を解釈する目的志向のトップダウン型の情報処理を同時に考える必要がある。

本研究では、ResNet50 による画像識別タスクを例にとり、多義性・曖昧性をもつ視覚情報に対して人が判断を行う際の情報処理について基礎検討を行う。タスク実行時の信念分布がモデルの識別結果へトップダウンに浸入するプロセスについて考察し、Duck-Rabbit illusion における判断の反転現象の説明を試みる。

2. 関連研究

2.1 Duck-Rabbit illusion

視覚情報の曖昧性・多義性を表す有名な例として、図 1 に示す Duck-Rabbit illusion (原題: Kaninchen und Ente) というイラストがある。Duck-Rabbit illusion にはアヒルとウサギの特徴が同時に含まれており、観察者は「この画像がアヒルとウサギのどちらに見えるか」を問われると判断に迷う。このイラストは、歴史上、人のもつ認識や知覚のメカニズムに言及するとき、しばしば例示されてきた。たとえば、物的現象と心的現象の差に着目し「意識」について考察した F.Brentano や E.Husserl の志向性 (Intentionality)、言語のもつ「意味」に着目した後期 Wittgenstein の Aspect といった抽象概念を解説する際に Duck-Rabbit illusion は曖昧性の例として参照されてきた。

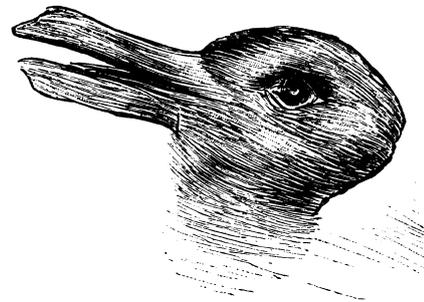


図 1: "Kaninchen und Ente" ("Rabbit and Duck"): ウサギとアヒルのどちらにも見える曖昧な画像。1892 年に Fliegende Blätter 紙上に掲載された。Source: <https://digi.ub.uni-heidelberg.de/diglit/fb97/0147>

2.2 知覚の理論負荷性

N.R.Hanson は、F.Brentano の「意識は常に何者かについての意識である」という志向性 (Intentionality) の前提に基づき、「すべての〈見る〉には〈……〉として見る〉が含まれている」という知覚論を提唱した。これを知覚の理論負荷性 (Perceptual theory-ladenness)[1] と呼ぶ。志向性や理論負荷性をもつ知覚処理系では、与えられた観測データに対して状況に応じて異なる処理をする必要があり、記憶・信念・文脈・欲求・感情に基づく脳の高次機能が特徴抽出・注意といった脳の低次機能へと介入するプロセス [2][3][4] が想定される。特に、認知科学では、脳の高次機能が脳の低次機能を支配するという考え方は認知的浸入 (Cognitive penetration) [5][6][7][8] という言葉で論じられてきた。

2.3 本研究の位置付け

本研究では特徴抽出や注意などの知覚の低次機能には何らかの理論負荷性が含まれると仮定し、Duck-Rabbit illusion の反転現象と曖昧性解消を例にして、従来研究に欠けていた計算論と情報処理に関する基礎検討を行う。

連絡先: 内海佑麻, 慶應義塾大学理工学研究科,
〒 223-8522 神奈川県横浜市港北区日吉 3-14-1,
E-mail: uchiumi@ailab.ics.keio.ac.jp

3. 提案

Duck-Rabbit illusion に対する判断の反転現象を、曖昧性解消を目的とした初期視覚（特徴抽出）機能の適応的変化によって説明し、信念分布に基づく注意計算を提案する。

3.1 曖昧性解消の計算論

一般に、2つの確率変数 X, Y に対して、 X の観測値 \mathbf{x} が与えられた下での Y の条件つき確率分布 $P(Y|X = \mathbf{x})$ を計算する関数 f を確率モデルという。このとき観測値 \mathbf{x}^* と確率モデル f によって得られる条件つき確率 $P(Y|X = \mathbf{x}^*)$ を $f(Y|\mathbf{x}^*)$ とかく。以下では、確率モデル f として識別モデルを考える。すなわち、識別モデル f は入力データ $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ から推論結果 $(f(Y|\mathbf{x}_1), f(Y|\mathbf{x}_2), \dots, f(Y|\mathbf{x}_n))$ を得る。

3.1.1 曖昧性 (Ambiguity)

入力変数 X のあるデータ \mathbf{x}^* に対して識別モデル f の推論結果がもつ曖昧性 (Ambiguity) は次のように定義できる。

$$A(Y; f, \mathbf{x}^*) = - \sum_{\mathbf{y}} f(\mathbf{y}|\mathbf{x}^*) \log f(\mathbf{y}|\mathbf{x}^*) \quad (1)$$

このとき、出力変数 Y の曖昧性 (Ambiguity) の大きさは、観測値 \mathbf{x}^* が与えられた下では、識別モデル f のみによって決定されることに注意する。

3.1.2 逸脱度 (Surprise)

次に、 Y についての事前知識を経験的に知っている場合、すなわち、 Y の生起確率を表すベイズ事後分布 $Y \sim p(\cdot)$ が既知の場合を考える。このとき、信念分布 p とモデル f による判断結果との逸脱度 (Surprise) は、識別モデル f が推定した Y の条件付き分布 $f(Y|\mathbf{x}^*)$ と Y の信念分布 $p(Y)$ との Kullback-Leibler Divergence として計算される。

$$S(Y; f, \mathbf{x}^*, p) = - \sum_{\mathbf{y}} p(\mathbf{y}) \log \frac{p(\mathbf{y})}{f(\mathbf{y}|\mathbf{x}^*)} \quad (2)$$

このとき、推論結果 Y の逸脱度 (Bayes surprise) の大きさは、観測データ \mathbf{x}^* と信念分布 p が与えられた下では、識別モデル f のみによって決定されることに注意する。

3.1.3 曖昧性解消プロセス

以上より、ある観測値 \mathbf{x}_i が与えられた下での識別モデル f の曖昧性解消プロセスは、

$$\min_f A(Y; f, \mathbf{x}^*) \quad (3)$$

となり、信念分布 p が与えられた場合の、識別モデル f の曖昧性解消プロセスは、 $\beta \in [0, 1]$ を用いて

$$\min_f A(Y; f, \mathbf{x}^*) + \beta \cdot S(Y; f, \mathbf{x}^*, p) \quad (4)$$

となる。

3.1.4 ベイズ事後確率との関係

出力変数 Y に関するベイズ事後分布 p が、モデル f と事前分布 π 、独立同分布を満たす n 個のデータサンプル $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ によって、

$$p(Y) := p(Y; \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \propto \sum_{i=1}^n f(Y|\mathbf{x}^{(i)})\pi(Y) \quad (5)$$

と導出されている場合、 Y の生起確率に関する信念分布 p は f によって推定されることになる。

3.2 曖昧性解消による初期視覚へのトップダウン介入

多義・曖昧な視覚情報に対する解釈を行う際、人間は視覚的注意によってこれを解決している [9][10]。実際、注視点を変化させることで、Duck-Rabbit illusion に対する解釈結果が異なるという実験結果 [11] もある。ここでは、曖昧性解消プロセスを初期視覚（特徴抽出）に作用させることで、入力画像に対する注意が適応的に変化するプロセスを考える。

3.2.1 特徴抽出・情報量・注意

モデル f の入力変数 X が出力変数 Y に対してもつ曖昧性や逸脱度を計算するためには、入力変数 X の出力変数 Y に対する条件つき確率 $p(Y|X)$ を計算する必要がある。このとき、モデル f には入力変数 X から出力変数 Y に対してより有用な情報を取り出す操作、特徴抽出が必要となる。すなわち、モデルが入力データ \mathbf{x} から特徴抽出を行い、特徴量 $\phi(\mathbf{x})$ から入力データ \mathbf{x} の（出力変数 Y に対する条件つき）情報量を計算し、入力データ \mathbf{x} のもつ（出力変数 Y に対する条件つき）平均情報量 H に対して入力データ \mathbf{x} の各要素 \mathbf{x}_i の寄与度（注意、顕著性） $w(\mathbf{x}_i)$ が計算される。

$$\mathbf{x} \xrightarrow{\text{特徴抽出}} \phi(\mathbf{x}) \xrightarrow{\text{情報量の推定}} H(Y|\phi(\mathbf{x})) \xrightarrow{\text{注意の計算}} w(\mathbf{x}_i)$$

実際、情報量最大化による顕著性の計算 [12] では、画像特徴量の基底関数 (basis function) を経験的に学習することで、画像の部分領域がもつ生起確率を学習することで、顕著性マップの導出を行なっている。

3.2.2 畳み込みカーネルへの実装

畳み込みニューラルネット (Convolutional Neural Nets, CNN) [13] では入力層に近い中間層の畳み込み処理が特徴抽出を担っており、その処理結果は畳み込みカーネルのパラメータに依存している。よって、曖昧性の最小化を目的として、中間層の畳み込みカーネルを更新することで、多義・曖昧な視覚情報に対する適応的な特徴抽出と注意計算が可能になる。

与えられた入力画像 \mathbf{x}^* に対して、初期視覚（特徴抽出）が適応的に変化することで、モデル f の曖昧性が解消されるプロセスを以下にまとめる。ただし、モデル f の中間層の畳み込みカーネルがもつパラメータを θ とおく。

(1) 曖昧性の計算 ある入力画像 \mathbf{x}^* に対して、モデル f の判断結果がもつ曖昧性は次のように計算される。

$$\ell(f_\theta, \mathbf{x}^*, p) := A(Y; f_\theta, \mathbf{x}^*) + \beta \cdot S(Y; f_\theta, \mathbf{x}^*, p) \quad (6)$$

(2) 特徴抽出プロセス (低次視覚) への介入 モデル f による推論結果の曖昧性を最小化するために、 θ は勾配情報に基づく値の更新を繰り返す。 θ が変わると、モデル f によって抽出される \mathbf{x}^* の特徴量が変わる。

$$\theta \leftarrow \theta + \eta \cdot \frac{\partial \ell(f_\theta, \mathbf{x}^*, p)}{\partial \theta} \quad (7)$$

(3) 入力画像に対する注意計算 入力画像 \mathbf{x}^* の i 番目の要素 \mathbf{x}_i^* がもつ、モデル f の判断結果の曖昧性に対する寄与度 $w(\mathbf{x}_i^*)$ は勾配情報として次のように計算される。

$$w(\mathbf{x}_i^*) := \frac{\partial \ell(f_\theta, \mathbf{x}^*, p)}{\partial \mathbf{x}_i^*} \quad (8)$$

$$= \frac{\partial A(f_\theta, \mathbf{x}^*)}{\partial \mathbf{x}_i^*} + \beta \cdot \frac{\partial S(f_\theta, \mathbf{x}^*, p)}{\partial \mathbf{x}_i^*} \quad (9)$$

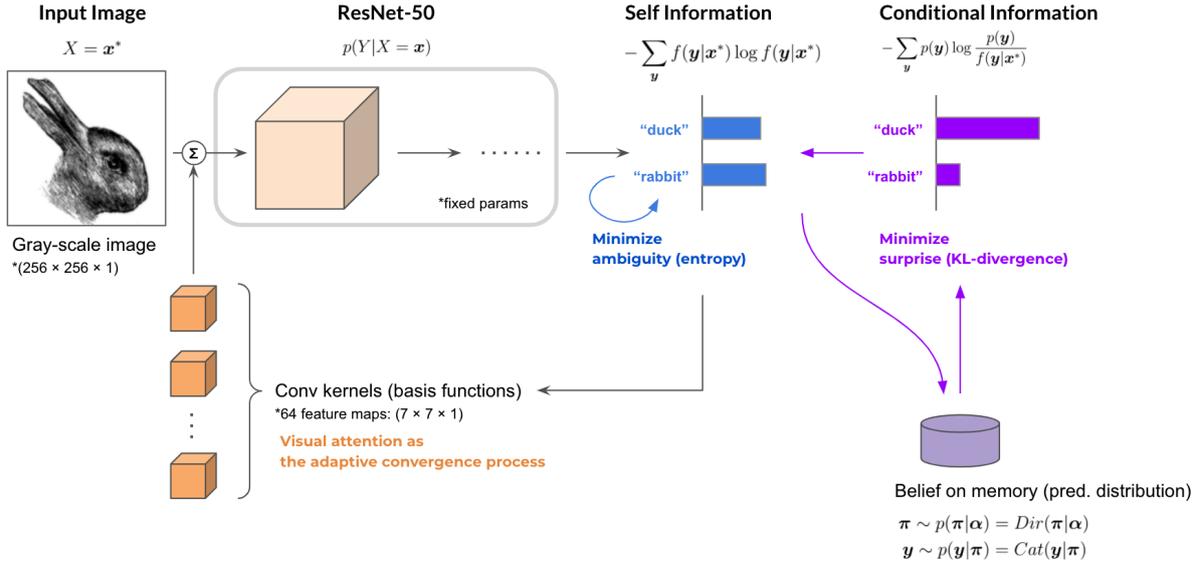


図 2: システム構成図: モデルの出力変数 Y に対する入力画像 \mathbf{x}^* の平均情報量 (曖昧性, ambiguity) と, 出力変数 Y に対する信念分布 $p(\mathbf{y}|\pi)$ に対する KL-Divergence (逸脱度, surprise) を同時に最小化するように畳み込みカーネルのパラメータを更新.

4. 実験と考察

4.1 データとモデル

Google Open Images Dataset[14] から, Object Detection タスクの Duck クラスと Rabbit クラスの画像を 150 枚ずつ用意した. また, 自由度調節のために同じ枚数のランダムなダミー画像を用意した. 画像は 256×256 の大きさへリサイズした. 識別モデルには ImageNet[15] で事前訓練した ResNet-50[16] を使い, 用意したデータに対して再学習を行なった.

4.2 判断に対する信念分布

画像識別タスクに対する信念を Dirichlet-Categorical 分布によってモデリングする. 入力画像を \mathbf{x} に対して ResNet-50 が予測するカテゴリカル変数の One-Hot ベクトルを $\mathbf{y} \in [0, 1]^K$ とおく. \mathbf{y} は, 各クラス k の生起確率 π_k をもつカテゴリカル分布 $\text{Cat}(\mathbf{y}|\pi)$ に従い, 各クラス k の生起確率ベクトル π は, 集中度 α をもつディリクレ分布 $\text{Dir}(\pi|\alpha)$ に従うとする.

$$\begin{cases} \mathbf{y} \sim p(\mathbf{y}|\pi) = \text{Cat}(\mathbf{y}|\pi) = \prod_{k=1}^K \pi_k^{y_k} \\ \pi \sim p(\pi|\alpha) = \text{Dir}(\pi|\alpha) = C(\alpha) \cdot \prod_{k=1}^K \pi_k^{\alpha_k - 1} \end{cases} \quad (10)$$

ただし $C(\alpha) = \frac{\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K \Gamma(\alpha_k)}$ とする. Dirichlet-Categorical 分布に対して, パラメータ α の初期値 $\alpha^{(0)}$ と, n コの観測値 $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}$ を与えると, π の事後分布は,

$$p(\pi|\alpha^{(n)}) \propto p(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}|\pi) p(\pi|\alpha^{(0)}) \quad (11)$$

$$\propto \left(\prod_{i=1}^n p(\mathbf{y}^{(i)}|\pi) \right) p(\pi|\alpha^{(0)}) \quad (12)$$

$$= C(\alpha^{(0)}) \cdot \prod_{k=1}^K \pi_k^{\sum_{i=1}^n y_k^{(i)} + \alpha_k^{(0)} - 1} \quad (13)$$

$$\propto \text{Dir}(\pi|\alpha^{(n)}) \quad (14)$$

となる. ただし, $\alpha_k^{(n)} := \sum_{i=1}^n y_k^{(i)} + \alpha_k^{(0)}$ とする.

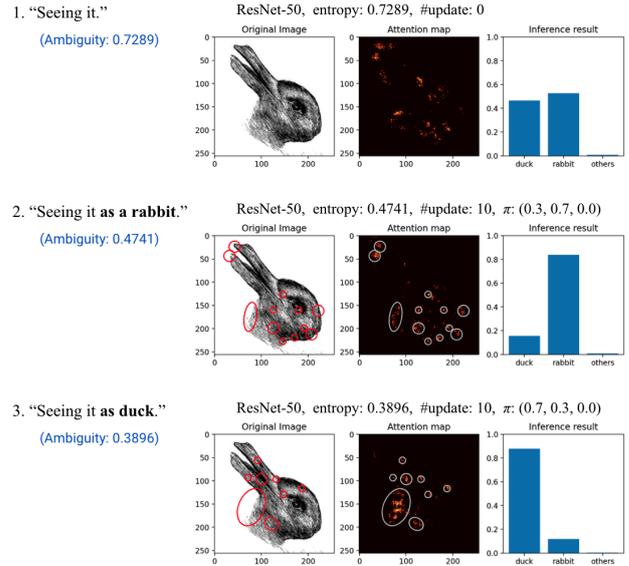


図 3: Duck-Rabbit-illusion における曖昧性と注意計算: 1.) 元の ResNet-50. 2.) Rabbit 優位の信念分布により更新された特徴抽出カーネルをもつ ResNet-50. 3.) Duck 優位の信念分布により更新された特徴抽出カーネルをもつ ResNet-50.

4.3 Duck-Rabbit illusion に対する注意マップ

ResNet-50 と信念分布を用いて, Duck-Rabbit illusion に対する推論を行い, 推論結果のもつ曖昧性を解消するために特徴抽出パラメータを更新した結果を図 3 に示す. 実験では, 予め設定された偏った生起確率 π をもつ Duck · Rabbit · Others クラスのカテゴリカル分布からサンプルされた 30 枚の画像を使って, Duck-Rabbit illusion に対する推論を行う前に異なる信念分布を形成させ, 曖昧な入力画像に対するモデルの判断結果と注意マップの違いを見た. 注意マップは, モデル f の曖昧性

$\ell(f_{\theta}, \mathbf{x}^*, p)$ に対して入力画像がもつ勾配情報の、Integrated Gradient[17]を計算することで導出した。元の ResNet-50 を使った推論では、入力画像の特徴から偏りなく判断しており、推論結果のもつ曖昧性が大きい。一方、推論結果のもつ曖昧性 (ambiguity) と信念分布に対する逸脱度 (surprise) を解消するように更新された特徴抽出カーネルをもつ ResNet-50 を使った推論では、入力画像の部分的な特徴が判断結果へ強く寄与しており、推論結果のもつ曖昧性が小さい。

4.4 ResNet-50 の各中間層の特徴抽出機能

ニューラルネットワークモデルの推論処理は、入力データから有用な情報を取り出す特徴抽出と、取り出した特徴量と出力変数との対応関係の学習を同時に行なっている。すなわち、ニューラルネットワークモデルの性質として、入力層に近い中間層は入力データのもつ情報から出力変数の予測や推定に有用な特徴のみを抽出し、出力層に近い中間層は特徴量のもつ情報を維持したまま低次元の出力空間へ情報を縮約している。

実験に使った学習済み ResNet-50 のそれぞれの中間層が、入力層と出力層に対してもつ独立性を測るため訓練データを用いて HSIC(Hilbert-Schmidt Independence Criterion)[18]を計算した。結果を図4に示す。ResNet-50 の構成は、最初の畳み込み層、4つの Residual block、プーリングと全結合層の6つからなるため、それぞれの中間層の出力ベクトルを layer-1~6 と表記した。今回の実験では、特徴抽出機能として layer 1 のパラメータのみを更新したが、図4をみると、layer-1~4 は入力層との依存性が高く、layer-5~6 は出力層との依存性が高い。

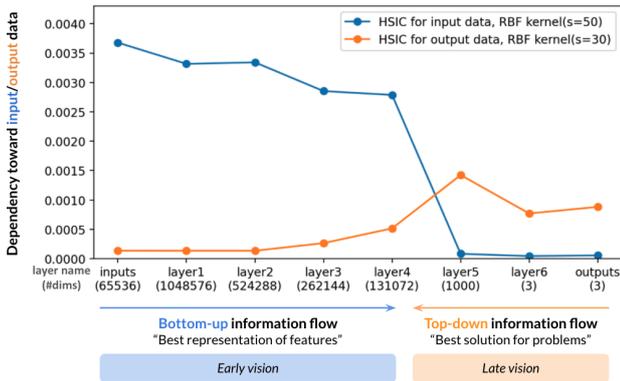


図4: ResNet-50 の中間層と入力層・出力層との独立性 (HSIC)

さらに、Duck・Rabbit クラスの訓練データを ResNet-50 に推論させた際の中間層の状態を t-SNE[19] で可視化した。結果を図5に示す。Duck・Rabbit いずれかのクラスに優れた特徴抽出をもつ ResNet-50 では、元の ResNet-50 よりも遅い layer-3 の段階で、各サンプルの分類領域が形成されている。Duck-Rabbit illusion は layer-3 でどちらかのクラスへ分類されており、特徴抽出は layer-1~3 が担っていることがわかる。

5. 結論

5.1 まとめ

本研究では、知覚の理論負荷性に着目して Duck-Rabbit illusion の反転現象の説明を試みた。すなわち、識別モデルの初期視覚 (特徴抽出) 機能を信念と曖昧性解消を目的として適応的に変化させることで、判断結果が変化するプロセスを実験

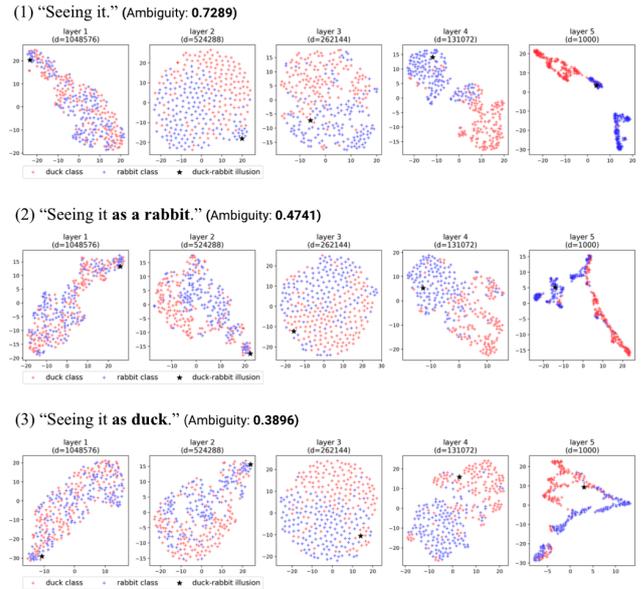


図5: ResNet-50 の中間層の t-SNE プロット

的に示した。また、初期視覚 (特徴抽出) 機能が変わる際に、識別モデルが入力データに対してもつ注意計算を可視化した。

5.2 将来研究

本稿では十分に検討できなかったが、本研究に関連する重要な研究トピックを挙げておく。

(1) **感覚: 信念とモダリティ** 本研究では、画像識別タスクにおいて識別クラスに対して形成される信念として、モデルによって予測されたカテゴリ変数のデータを用いた。しかし、実際の人間は複数の感覚機能を持っているため、ある信念が、視覚などの単一のモダリティに基づいて形成されるものなのか、他のモダリティを参照して形成されるものなのか、を考える必要がある。感覚統合は非常に難しいテーマだが、信念や欲求といった高次認知機能の発達には欠かせない条件である。

(2) **記憶: 注意と作業記憶** 視覚的作業記憶と視覚的注意の関係 [20][21][22] は、認知科学や心理学において近年注目されている研究トピックである。選択的注意と作業記憶は相互参照する関係にあるため、これらを同時に捉えるアプローチが求められている。すなわち、あるタスクを実行する時、どの情報を選択するか? という注意機能は、あるタスクを解く際に形成される事前の作業記憶を参照するし、あるタスクを解く際に形成される作業記憶は、どの情報を選択したか? という事前の注意機能を参照する。また、信念が生成され消滅する過程は、人間があるタスクに従事する際に、作業記憶を生成しやがて放棄する過程と似ていることから、信念と記憶の関係について考察する価値がある。

(3) **運動: 注視点と眼球運動** 人間の視覚系がとらえる刺激強度は注視点を中心として周辺へと減衰する傾向があり、受け取る刺激要素の形状は、画像データのような矩形ではなく円形に近いことが知られている。加えて、注視点は眼球運動によって制御されており、視覚探索課題の実験を通じて、眼球運動、注視点、選択的注意に纏わる様々な刺激-反応の相関関係が明らかになっている。将来研究として、視覚系を、知覚推論と運動制御からなるダイナミックな情報処理系として捉えて、眼球運動、注視点、選択的注意といった現象を一貫する、計算の機序を考える必要がある。

謝辞

本研究は、JST、CREST、JPMJCR19A1 の支援を受けたものである。

参考文献

- [1] Norwood Russell. Hanson. *Patterns of discovery : an inquiry into the conceptual foundations of science / by Norwood Russell Hanson*. Cambridge University Press Cambridge, 1958.
- [2] S. Ullman. *High-level Vision*. The MIT Press, 1996.
- [3] 横澤一彦. 高次視覚の認知心理学と脳機能計測. 認知神経科学, Vol. 1, No. 2, pp. 152–156, 1999.
- [4] 横澤一彦. 注意と認知. *Technical Report on Attention and Cognition*, Vol. 1, , 2003.
- [5] Jack Lyons. Circularity, reliability, and the cognitive penetrability of perception. *Philosophical Issues*, Vol. 21, pp. 289 – 311, 10 2011.
- [6] Shaiyan Keshvari, Ronald van den Berg, and Wei Ji Ma. Probabilistic computation in human perception under variability in encoding precision. *PLOS ONE*, Vol. 7, No. 6, pp. 1–9, 06 2012.
- [7] 正秀西村. 運動知覚の認知的侵入可能性. 彦根論叢, No. 412, pp. 52–67, 2017.
- [8] Zenon Pylyshyn. Is vision continuous with cognition? the case for cognitive impenetrability of visual perception, 1998.
- [9] Noaki M. Kawabata N, Yamagami K. Visual fixation points and depth perception. *Vision Res.*, Vol. 18, No. 7, pp. 853–4, 1978.
- [10] Kawabata N. Attention and depth perception. *Perception*, Vol. 15, No. 5, pp. 563–72, 1986.
- [11] 岸本充史, 川端信男. 局所的・大域的情報選択モデルによる多義図形の非あいまい化. テレビジョン学会誌, Vol. 50, No. 5, pp. 594–598, 1996.
- [12] Neil D. B. Bruce and John K. Tsotsos. Saliency based on information maximization. In *Proceedings of the 18th International Conference on Neural Information Processing Systems, NIPS’05*, p. 155–162, Cambridge, MA, USA, 2005. MIT Press.
- [13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, Vol. 25. Curran Associates, Inc., 2012.
- [14] Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Mallocci, Alexander Kolesnikov, Tom Duerig, and Vittorio Ferrari. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. *IJCV*, 2020.
- [15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.
- [16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, Vol. abs/1512.03385, , 2015.
- [17] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70 of *Proceedings of Machine Learning Research*, pp. 3319–3328, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR.
- [18] Arthur Gretton, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf. Measuring statistical dependence with hilbert-schmidt norms. In Sanjay Jain, Hans Ulrich Simon, and Etsuji Tomita, editors, *Algorithmic Learning Theory*, pp. 63–77, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.
- [19] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, Vol. 9, pp. 2579–2605, 2008.
- [20] Paul Downing and Chris Dodds. Competition in visual working memory for control of search. *Visual Cognition*, Vol. 11, No. 6, pp. 689–703, 2004.
- [21] Theeuwes J. Olivers CN, Meijer F. Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *J Exp Psychol Hum Percept Perform*, Vol. 32, No. 5, pp. 1243–65, 2006.
- [22] Chunyue Teng and D. Kravitz. Visual working memory directly alters perception. *Nature Human Behaviour*, pp. 1–10, 2019.