

プロジェクションのメカニズムを探る： 重ね描きとプレディクション

Towards the mechanism for projection: overlaying and prediction

横山裕樹 *1 岡田浩之 *2 鈴木宏昭 *3
Hiroki Yokoyama Hiroyuki Okada Hiroaki Suzuki

*1*2 玉川大学脳科学研究所

Tamagawa University Brain Science Institute

*3 青山学院大学教育人間科学部

College of Education, Psychology and Human Studies, Aoyama Gakuin University

1. 問題の背景

目の前に自分の好物のリンゴがある。するとそれは美味しそうだと感じる。そうした感覚、感情が生じると同時に、私たちに「美味しそうなるリンゴがある」という認識が生じる。つまり推論の結果生まれた「美味しそう」という情報は、単に感じられる、そう考えたという、心の中の出来事だけでなく、「ある」という形で世界に実在するものとなっている。

この（感じた、考えた）→（ある、存在する）という関係はいかにして確立されるのだろうか。残念ながら、認知科学を含めた心の科学は、この問題に対する知見をほとんど有していない。この問題の困難さに気づいた哲学者の一部には、これをもって内的表象の不在、認知科学の不要を説くものもいる [Searle 84]。

2. プロジェクションとは何か

この困難を乗り越えるためには、内的（脳内、情報処理システム内）に構成された表象を世界と結びつける心の働きが必要となる。我々はこの心の働きをプロジェクションと名付け、その研究の必要性を訴えてきた [鈴木 16, 鈴木 19]。物理世界の発する情報を元に内的に作り出された表象＝世界の意味は、プロジェクションの働きにより世界へと返されることにより、世界はその人なりの固有の意味に彩られた経験世界となる。そしてこの経験世界の中で人はまた認識と行為を繰り返す。

プロジェクションは、いくつかのタイプに分けられる。内的に構成された表象がそれを生み出した物理世界の事象とある程度まで一致した形で行われる、いわゆる投射が考えられる。目の前のものを特別な意味なしに知覚する、音の発生源の定位を行うなどは、通常環境下では投射の元となる情報の発生源＝ソースと、投射の先＝ターゲットが一致している。

一方、物理世界のソースとは異なる対象へのプロジェクションである異投射も考えられる。例えば、自分の手に対する刺激を目の前のゴムの手に感じるラバーハンド錯覚 [Botvinick 98]、音源とは異なる場所を音源としてしまう腹話術効果などは、典型的な異投射である。またフェティシズムに見られるように、長年使用してきたモノ、特別な経験の時に用いられたモノ、有名人が使用したモノには、そのモノ自体が持つ性質とは異なる性質が付与される [Bloom 11]。これも異投射の1つと考えることができる。また PC 上のエージェントに特別な感情を抱いたり [小野 16]、自分の身体の一部と同期するアバターなど

の CG に自己との同一感を感じるなどの現象も異投射の現れと考えられる [渡邊 16]。

またソースが存在しない、あるいは曖昧であるにも関わらず投射が起こることもある。これを虚投射という。盛んに研究がなされている色字共感覚では、黒い文字で印字されているものに色がついたと見え、それが数字によって異なることが明らかにされている [Cytowic 09]。統合失調における幻聴、東日本大震災後に被災地に現れる幽霊 [工藤 16]、極限状況下において現れる他者＝サードマンなども [Geiger 09]、明確なソースがないにも関わらず投射が起こる虚投射の例となるだろう。また脳への直接的な刺激を与えて脳の状態を変化させることで、虚投射が生じることも報告されている。[Arzy 06] は、てんかん治療のために埋め込まれた電極を通して、左側頭頭頂結合部に刺激を与えることで、現実には存在しない人の気配を生み出せるという報告もある [Arzy 06]。

以上のことの中には、共感覚のような厳密な実験を通して得られたデータもあれば、ひとりの体験者の事後報告に基づくものもある。さらに、その体験の中にはサードマン、幽霊などのような一見オカルト風のものも含まれている。そうしたことから、これらの現象は科学の対象にならないと考える人もいるかもしれない。ただここで主張したいことは、これらの認識対象（幽霊やサードマン自体）の研究をすべきということではなく、そうした認識を生み出す心の仕組みを解明することは重要だ、ということである。つまり色字共感覚者が見る色が実際に印字されたもののどこかに存在するという、実際に幽霊がどこかにいることの研究を勤めているのではなく、色や幽霊が見えてしまうプロジェクションの仕組みを明らかにすべきということである。

これら様々なタイプのプロジェクションの解明を通して、人も含めた生体にとって意味に彩られた経験世界がいかに作り出されるのか、そしてその中で生体はどう活動するのかを描き出すことが可能になる。

3. 認識と生成

物理学者のファイマン (Richard Feynman) は “What I cannot create, I do not understand. (私は自分に作れないものは、理解できない)” という言葉で認識と生成を一對のこととして捉える重要性を示唆している。本稿で議論しているプロジェクションによる認知のプロセスは、世界に対して予測モデルを生成し、認識と生成のプロジェクションループを回すという点で同様の問題意識を共有している。このように認識と生成

連絡先: hiroki.yokoyama@okadanet.org

を投射というプロセスを介して一対として考えることがプロジェクション科学の大きな特徴であると考えられる。

ここではプロジェクション科学における認識と生成の問題を統計的予測問題における生成モデルと識別モデルの対比で考える。あるデータがあったときに、そのデータを入力とした何らかの予測器を構成することが統計的予測問題である。

リンゴを見たときに「大きい」とか「小さい」とか判断したり、「美味しそう」とか「酸っぱそう」とか味を推測すること、即ちデータ (\mathbf{x}) に対するクラス (C_k) の条件付き確率 (事後確率 $p(C_k)$) を直接データ (\mathbf{x}) だけからモデル化することを識別モデルと言う。大量の教師データ (\mathbf{x} を入力としたときに対応するクラス (C_k) が得られる場合には識別モデルのような仕組みでリンゴの見た時の各人様々な認識を獲得することが可能であるが、実世界で変化に富んだ環境で暮らす我々にはこのような状況は稀なことであると言わざるを得ない。

一方、生成モデルでも同様にデータに対するクラスの条件付き確率 (事後確率 $p(C_k)$) をモデル化する点では識別モデルと同様だが、生成モデルではこの事後確率をベイズの定理により以下のように求める。

$$p(C_k|\mathbf{x}) = \frac{p(\mathbf{x}|C_k)p(C_k)}{p(\mathbf{x})} \quad (1)$$

この式の右辺において、 $p(\mathbf{x}|C_k)$ はクラスに対するデータの条件付き分布であり、これを経験的に求めるのは難しくない。つまり、単にリンゴの画像からリンゴの形状 (丸いとか小さいとか) や位置、光源などのパラメータを推定する (認識) だけでなく、生成過程をシミュレーションすることによりリンゴの画像の予測を可能にしているのである。

このような生成モデルの考え方はプロジェクションによる認識と非常に親和性があると考えられる。即ち生成と認識は一対のものとして考える、その認知プロセスを説明するのがプロジェクション科学であると考えられる。

4. 重ね描きによるプロジェクション

プロジェクションの問題を難しくするのは、その経路に物理的基盤が存在しないという事実である。外界からの刺激の受容から処理までは、物理的、生理的な基盤が存在する。視覚で言えば、受容されるのは電磁波の一種である光であり、その後の処理は脳内の各部位が担当する。一方、プロジェクションについて、それを字義通りに受け取れば、それは我々から世界へ向けてのプロセスである。このプロセスの一部は、明らかに行為である。しかし、行為が及ぶのは身体の到達可能な範囲までのことであり、視覚や聴覚のような遠感覚には及ばない。言うま

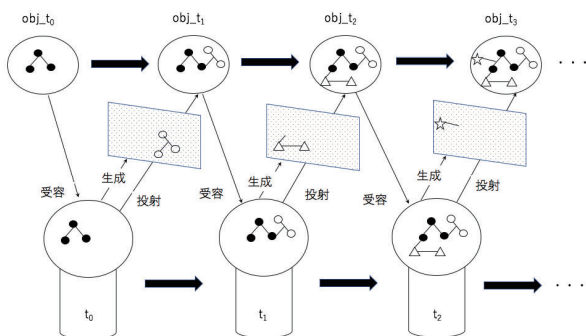


図 1: フィルタの漸次的更新による認識の変容

でもなく、人はプロジェクタではないので、目から光線を発するわけではない。だとすると、内的に作り出された意味をいかにして世界の中に実在させられるのだろうか。

この難問解決の候補を考える際に、大森が述べる「重ね描き」は大きな意味を持つ [大森 82]。重ね描きについて、大森は概略以下のように論じる。赤いメガネをかければ、世界は赤く見えるのと同様に、脳に何らかの異常が生じれば、世界はその異常を反映した形で知覚される、と*1。つまり大森は、赤メガネ、脳も含めた身体を前景として、世界という遠景を透視するという図式を提案している。

これを本論文の主旨から解釈すれば、前景はフィルタであり、遠景は世界ということになるだろう。つまり、人は推論も含めた知覚システムによってフィルタを生成しながら、世界を知覚したり、行為を通して世界に働きかけたりするというサイクルを回しながら認知を行っている。このサイクルの中で、ただのリンゴは赤くて艶やかなリンゴとなり、おいそなリンゴとなり、(たとえば) アップルパイの材料にしたいものとして知覚されることになる (図 1 参照)。こうした重ね描きという働きによって、心の世界と物理世界を結びつけること、すなわちプロジェクションが可能になるのではないだろうか。

5. 予測符号化理論による説明

こうした図式は予測符号化理論との適合性が高い。予測符号化理論は、Rao らが大脳皮質にみられる求心性、遠心性の双方向結合の役割を説明するために提案した計算モデルである [Rao 99]。初期視覚野のニューロンは、線分方位や局所的な動きといった情報を表現し、これらを組み合わせるとどのような網膜像が得られるかを予測する。この予測は、初めは正確ではないため、実際の網膜像との間に予測誤差が生じる。この予測誤差を減少させるためにニューロン活動を修正する。同様に、高次の領野は初期領野からの入力を良く予測するように、より抽象的な概念を表現する。下位概念を x 、上位概念を y と表すと、各領野は上位概念が与えられた時に下位概念がどのように生成され得るかを $p(x|y)$ として表現していることになる。感覚入力の生成則を表現しているため、予測符号化モデルは生成モデルの一種である。 $p(y|x)$ を表現し、感覚入力から抽象概念へと一方向かつボトムアップに計算して行く識別モデルとは、考え方が逆となる。

例えば、高次領野で、目の前の物体がりんごであることを y_2 に表現しているとすると、この領野は、これまでの経験から、りんごが円形で赤いことが $p(y_1|y_2)$ に表現されており、これに基づいてりんごの画像を構成する曲線や色の情報が y_1 に表現されていることを予測する。そして初期領野は曲線や色の情報に基づいて網膜像を予測する。もし実際の網膜像が予測よりも黄色を帯びていて凹凸があれば、初期領野は x の予測誤差を減少させるために y_1 を修正し、高次領野はそれに伴って生じる y_1 の予測誤差を減少させるために y_2 を修正する。この場合は、物体がりんごではなく、みかんであるということが表現される。さらに、味覚との関係が $p(y_3|y_1, y_2)$ として表現されていれば、 y_3 が予測され、(赤くて艶のあるりんごなので) 「美味しそうなりんごがある」と感じることができる。このよ

*1 大森は表象を認めないどころか、主観-客観の区別を否定するという立場から論を進めており、本論文のような物理世界と心理的世界 (表象) を分けて考える立場の対極に位置する。ただし、2つの世界を一体化するという意味では、大森と本論文は問題意識を共有している。大森は主客が一体化していることを世界、人の在り方として存在論的に描き出した。一方、本論文はそれを認識論的に描き出すことを目指している。

うな予測と誤差のやりとりを繰り返し行うことによって、感覚入力として得られた x を最も良く説明する ($p(y_1, y_2, \dots | x)$ を最大にする) y_1, y_2, \dots を求めようとする過程で、りんごに様々な意味が重ね描きされる。

色や形、単語といった静的な情報だけでなく、それに対する行動やその結果といった時系列の情報も重ね合わせの対象となり得る。Rawlik らは、エージェントの行動と報酬を折り込んだ生成モデルを提案し、強化学習を教師なし学習の一種として記述している [Rawlik 12]。また、深層ニューラルネットワークによる生成モデルの学習を実現した GAN (Generative Adversarial Networks, [Goodfellow 14]) は、データセットを再現しながら、そこには含まれていない新たなパターンを生成することができる。この GAN のように、複雑でありながら高い汎化性能を持った生成モデルは、自己の行動やそれに伴う報酬を含む経験を学習することで、新たな行動を生成し、その結果を予測することも可能であろう。そのため、目の前のりんごに対して想起した、りんごを剥くという動作、りんごを食べている恋人といった情報から、「このりんごを彼氏のために剥いてあげたい」などという考えが浮かび、一つの物体に対して様々な意味が重ね描きされていくと考えられる。

6. おわりに

本報告では、プロジェクションがなぜ必要なのかを述べ、プロジェクションのタイプ分けをまず行った。次に、プロジェクションがいかなる計算過程を経てなされるのかを、生成モデル、大森の「重ね描き」、予測符号化理論をベースにして考察した。これらはまだ提案の段階であり、その実装、シミュレーションは今後の課題となる。

プロジェクション科学は、情報の受容から表象の構成の段階に留まっていた心の科学を、世界へと結びつけ、人（あるいは種）に固有な意味世界を描き出す新しいパラダイムである。このため、これまでの知見の再検討、再解釈に加えて、心の科学とは疎遠であった様々な分野（社会関連諸科学、経営学、臨床諸科学）との共同が必要となる。またこの展開の過程では、意識、クオリア、志向性、自己などの哲学上の難問への挑戦も必要となるだろう。

謝辞

本研究は、青山学院大学総合研究所「投射の科学」ユニットからの助成を受けて行われた。

参考文献

- [Arzy 06] Arzy, S., Seeck, M., Spinelli, L., Ortigue, S., and Blanke, O.: Induction of an illusory shadow person, *Philosophical Transaction of the Royal Society B*, Vol. 443, p. 287 (2006)
- [Bloom 11] Bloom, P.: *How Pleasure Works: The New science of Why We Like What We Like*, Vintage Books (2011), (小松淳子 (2012). 『喜びはどれほど深い：心の根源にあるもの』インターシフト)
- [Botvinick 98] Botvinick, M. and Cohen, J.: Rubber hands feel touch that eyes see, *Nature*, Vol. 391, p. 756 (1998)
- [Cytowic 09] Cytowic, R. E. and Eagleman, D.: *Wednesday is Indigo Blue: Discovering the Brain of Synesthesia*, MIT Press, Cambridge, MA (2009), (山下篤子 (訳)

(2010). 『脳の中の万華鏡：共感覚のめくるめく世界』河出書房新社.)

[Geiger 09] Geiger, J.: *The Third Man Factor: The Secret to Survival in Extreme Environments*, Penguin (2009), (伊豆原弓 (訳) (2010). 『サードマン：奇跡の生還を導く人』新潮社.)

[Goodfellow 14] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: *Generative Adversarial Networks* (2014)

[工藤 16] 工藤 優花：死者たちが通う街：タクシードライバーの幽霊現象、金菱 清 (ゼミナール) (編), 呼び覚まされる霊性の震災学：3・11 生と死の狭間で, 新曜社 (2016)

[小野 16] 小野 哲雄：「プロジェクション・サイエンス」の視点からの認知的メカニズムのモデル論的理解, 日本認知科学会第 33 回大会発表論文集, pp. 26 – 30 (2016)

[大森 82] 大森 莊蔵：新視覚新論, 東京大学出版会 (1982)

[Rao 99] Rao, R. P. N. and Ballard, D. H.: Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects, *Nature Neuroscience*, Vol. 2, No. 1, pp. 79–87 (1999)

[Rawlik 12] Rawlik, K., Toussaint, M., and Vijayakumar, S.: On Stochastic Optimal Control and Reinforcement Learning by Approximate Inference, in *Robotics: Science and Systems* (2012)

[Searle 84] Searle, J. R.: *Minds, Brains, and Science*, BBC (1984), (土屋俊 (訳) (1993). 『心・脳・科学』岩波書店)

[鈴木 16] 鈴木 宏昭：プロジェクション科学の展望, 日本認知科学会第 33 回大会発表論文集, pp. 20 – 25 (2016)

[鈴木 19] 鈴木 宏昭：プロジェクション科学が目指すもの, 認知科学, Vol. 26, pp. 1 – 22 (2019)

[渡邊 16] 渡邊 翔太, 川合 伸幸：自己身体と操作対象の身体図式の一致はより強い運動主体感・身体所有感を喚起する, 日本認知科学会第 33 回大会発表論文集, pp. 31 – 40 (2016)