未知領域を含むオブジェクト同定による窓問題知覚の説明
A computational account for an aperture problem by object identification with uncertainty

日髙　昇平†，高橋　康介‡
Shohei Hidaka, Kohske Takahashi
†北陸先端科学技術大学院大学，‡中京大学
Japan Advanced Institute of Science and Technology, Chukyo University
shhidaka@jaist.ac.jp

Abstract

The idea of "representation" has been pervasive among empirical sciences and philosophy of mind including cognitive science. It implicitly assumes the existence of a correspondence between a construct in the cognitive system and an object in the outer world. Beyond this implicit assumption, can we formulate a potential mechanism of recognition without "representation"? This study explores an answer to this question by considering with a visual illusion, called Barberpole illusion, which has intrinsic ambiguity in its interpretation. Our approach exploits a sort of structural consistency as a basis to infer the "object" which is not directly accessible to the observer, and gives an account for the visual illusion. This theoretical account may be a step toward a potentially novel mechanism replacing "representation".

Keywords — Representation, object identification, Barberpole illusion, aperture problem

1. はじめに

我々が何かを認識するとき、それはその何かの鋳型のような“表象”を我々が持っている、あるいはそれに対応づける、という考え方は、認知科学にとどまらずデカルト以来の心の哲学においても主流の考え方である。もしそうだとしたら、その鋳型あるいは表象はどこから来たのだろうか。一見すると、この考え方は何かを説明しているようにも思えるが、“脳内小人”ホムンクルスが外の世界を見ている、という無限後退に陥る。

本研究では、こうしたなんらかの鋳型を前提としない認識過程の説明を模索する。そのような説明原理の一つとして、“オブジェクト同定”による視覚的物体の説明[1, 2, 3]に着目し、具体的な視覚現象の説明を試みる。具体的に、Barberpole illusion [4]と呼ばれる曖昧図形の知覚を一つの事例として取り上げ、その数理的な説明を提案する。

2. オブジェクト同定に基づく運動知覚

ある2つの時点(多くの場合、その2時点の物理的時間の間隔は十分短い)の視覚的パタン(以下,「視覚データ」)が与えられると仮定する。このとき、1つの線形変換(場合によってアフィン変換まで拡張)で対応がつく2時点の視覚データのある部分(ベクトル空間上の点の集合)を「オブジェクト」と呼ぶ。また、2時点の視覚データが与えられたとき、それらの視覚データから可能なオブジェクトを探索し、オブジェクトを特定することを、「オブジェクト同定」と呼ぶ。

ある2時点の視覚データが与えられたとき、それらの間の“最大マッチング”により、オブジェクト同定を行う。これは、視覚データとオブジェクトモデルの間の予測誤差を最小にすることと同値である。オブジェクト同定で考えられるケースとして、大きく以下の3つのケースが考えられる。：

(1) 1つのオブジェクトモデルで視覚データが過不足なく説明できる場合(well-posed)
(2) 無数のオブジェクトモデルが視覚データを説明できる場合(ill-posed)
(3) どの1つのオブジェクトモデルでも視覚データが説明できない場合 (over-posed)

直感的に言えば、(1)は、データの複雑さとモデルの複雑さが本質的に同じ場合、(2)はデータがモデルを制約するのに十分でなく、説明できるモデルが1つに決まらない場合、(3)はデータの制約が強すぎて、1つのモデルでは説明できない(独立な2つのオブジェクトを同時に仮定する必要がある)場合である。それぞれ、線形代数で言うところの、(1)モデルの未知変数と方程式の数が一致する well-posed (良設定)問題，(2) モデルの未知数が方程式の数より多い(ill-posed)不良問題，(3) モデルの未知数より方程式

の数が多い(over-posed)過剰設定問題に相当する。

3. 窓問題におけるオブジェクト同定

具体的に、窓問題(Aperture problem)[5,6]あるいはBarberpole [4]と呼ばれる直線運動を例として、オブジェクト同定による曖昧図形（前節の(2)の場合）の知覚を説明する。

Barberpole 知覚(以下 BP)とは、2 次元平面上の、“窓”から移動する直線を覗いたときに生じる直線の運動知覚を指す(図 1)。図 1 において、青枠で囲われた領域の内側の領域のみ観測でき、ある時点 0 で直線 0($L_1 = \{(x_1, x_2) \mid ax_1 + bx_2 = c_0\}$)が観測され、また次の時点 1 において直線 1($L_1 = \{(x_1, x_2) \mid ax_1 + bx_2 = c_1\}$)が観測されたとする。また、直線 0 と直線 1 はいずれもある法線ベクトル$(a, b)^\top$に直交するとする。

異なる 2 つの時点の 2 つの直線を対応付ける平行移動は、直線がその方向への運動の知覚と解釈できる。このケースでは、運動知覚と解釈可能な平行移動、すなわち、2 つの直線 0 と 1 を対応づけるアフィン変換(この場合では平行移動のみを考える)は、無数に存在する。つまり、任意の$\gamma \in \mathbb{R}$に対し、以下のどの平行移動も直線 0 上の任意の点を直線 1 のある点に対応付ける。:

$$(\alpha, \beta)^\top = \frac{c_1 - c_2}{2(a^2 + b^2)}(a, b)^\top + \gamma(b, -a)^\top.$$

つまり、この 2 つの時点の直線を観測することでは、直線がどのような移動をしたのか一つに定めることはできない。原理的には、(線分ではなく、無限遠にまで伸びる)直線であれば、窓の内側だけを観測しない場合(窓がない、あるいは窓が無限に大きい場合)にでも、この対応付けの不定性は生じる。逆に、端点のある線分の場合は、その端点の存在から、線分上のすべての点を過不足なく対応付けるアフィン変換は一つに定まる。この現象が往々にして“窓問題”と呼ばれる理由は、仮に端点の存在する線分であっても、その端点が窓枠の外にあり、観測できない場合には、2 直線の対応付けの不定性と同質の不定性が生じるからである。

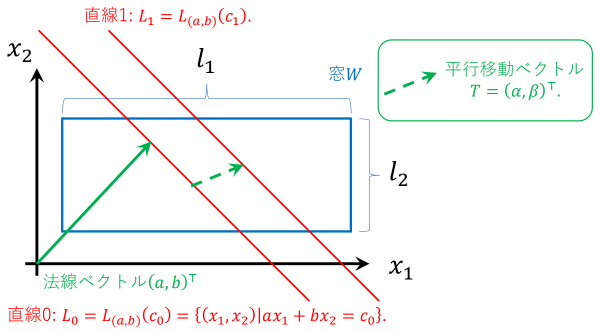

図 1: Barberpole 知覚の定式化.

以上のように、原理的には 2 つの直線を対応付ける変換は無数に存在するに関わらず、典型的には、特定の窓から直線の移動を観測すると、ある一定の方向に運動するように知覚される。これが BP 知覚である。BP 知覚の典型例は、その名称の指す通り、床屋のポール(Barberpole)が、水平方向に回転しているにもかかわらず、鉛直方向に縞模様(直線の束)が平行移動しているように見える経験は誰しもあるだろう。

4. 観測されない視覚データの最小化

原理的に不定であるはずの運動方向が、しかし経験的には一定の方向への運動を生じる、という BP 知覚は、いかに説明されるだろうか。本研究では、窓の存在によって生じる対応づけの不確実性を着目して BP 知覚の説明を与える。図 2 に、ある平行移動(破線矢印)を想定した場合に生じる「未知領域」(灰色の領域)を図示する。「未知領域」とは、仮にある平行移動によってある時点 0 から 1 へと平面全体が変換されたと考えたときに、窓 W の内に時点 1 で観測されるが、時点 0 では窓の外にあるような領域を指す。この領域にある点は、平面のある平行移動で対応づける場合、時点 0 で観測されないが、“対応づけられる位置に時点 1 と同じ視覚的パタン(直線や空白など)があった”と想定しなければ、その平行移動による平面の「解釈」と整合性が保てない。つまり、未知領域にある点に関して、特定のパタンを“でっちあげる”ことで、平面全体を平行移動したという解釈が整合的に成立する。

もし未知領域にある特定の視覚的パタンが存在する事前確率があるとすれば、ある平行移動を想定したときの未知領域の面積が大きいほど、その平行移動の事後確率が小さくなる。あるいは、“でっちあげる”領域が大きいほど、そのようなことが生じる確率が小さくなるとも言える。したがって、未知領域に関する一様な事前分布を考え、ベイズ推定における事後確率をより大きくする対応付けを知覚する確率が高いと考えれば、未知領域の最小化する平行移動が知覚される確率が最大となる。

この考えに従って計算すれば、矩形の窓である場合(図 1,2 の場合)、縦横比に応じて、より長い辺に平行移動するように知覚されるという最大事後確率解が得られる。具体的に、図 2 のように矩形窓の横の長さをl_1、縦の長さをl_2とした場合、未知領域の面積

A は平行移動パラメタγの関数として

$$A(\gamma) = (l_1, l_2)^{\top} \left| \frac{c_2 - c_1}{2(a^2 + b^2)} \binom{a}{b} + \gamma \binom{b}{-a} \right|$$

とかける。これを最小にする平行移動パラメタγを解けば、矩形の窓枠長と生じる知覚の関係を定性的に説明する平行移動が得られる。

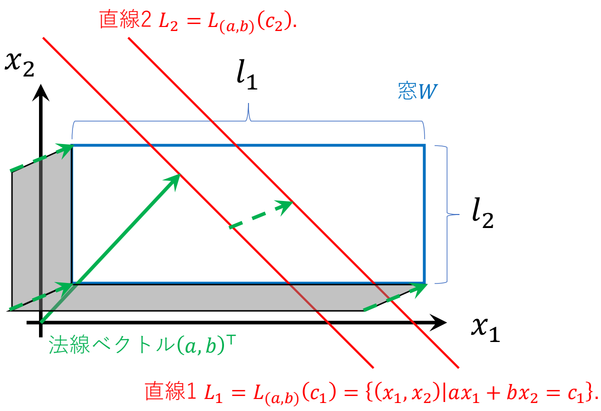

図 2: ある平行移動(破線矢印)を想定した場合に対応づける点が観察されない「未知領域」(灰色の領域).

5. 今後の課題

本稿では、曖昧図形の全体性をもつ知覚の一種である窓問題・Barberpole 知覚のオブジェクト同定による説明を与えた。今後の課題として、これと同様の説明により、BP 刺激の 2 重の重ね合わせによる Plaid motion [7] など、他の運動知覚や、あるいは運動を伴わない静的な曖昧図形の知覚の説明へとスコープを広げていくことが挙げられる。また、本稿で提示した説明による予測を実験的に検討していく予定である。

参考文献

[1] Hidaka, S. (2018). From Machine Learning to Machine Understanding. Japanese-German Frontiers of Sciences Symposium, 2018 年 9 月 7 日, 京都ブライトンホテル.

[2] 日高昇平 (2018). 記号接地問題における地とは何か：視覚的物体の同一性の分析. 日本認知科学会第 35 回大会論文集. (OS10-2).

[3] 高橋康介・日高昇平 (2019). 恒常性（constancy）の構造と認知的錯覚への適用. 錯覚現象のモデリングとその応用・第 13 回 錯覚ワークショップ.

[4] Guilford, J.P. (1929) "Illusory Movement from a Rotating Barberpole." American Journal of Psychology 41: 686–687.

[5] McDermott, J., Weiss, Y., Adelson, E.H. (2001). Beyond junctions: Nonlocal form contraints on motion interpretation. Perception, 30: 905-923.

[6] Adelson, E. & Movshon, A. (1982) Phenomenal coherence of moving visual patterns. Nature, 300, 523-525.

[7] Kim, J. & Wilson, H. (1993) Dependence of plaid motion coherence on component grating directions. Vision Research, 33, 2479-2489.