

ヒトの視覚系の特徴を考慮したアモーダル補完に関する 視覚数理モデルの予備的検討

Preliminary Study on Computational Model of Amodal Completion Considering Features of Human Visual System

満 倉 英 一[†]

Eiichi MITSUKURA

要 旨

ヒトは物体の一部が他の物体によって遮蔽された場合に、被遮蔽領域の形状を補完する。この視覚機能はアモーダル補完と呼ばれる。遮蔽領域における解の候補は理論的には無数に存在する。そのため、視覚系はこのような領域に対して制約条件を設定し、不良設定問題を良設定問題に変換することで解いていると解釈されている（標準正則化理論）。多くの視覚数理モデルはこの制約条件に空間的な滑らかさを採用し、ヒトの知覚の再現に成功してきた。しかし、近年の研究において、制約条件は滑らかさではなく、“平坦性”の方がより適切である可能性が示唆されている。そこで、本研究では、制約条件として“平坦性”を用いた視覚情報補完モデルをアモーダル補完に適用し、数値シミュレーションによる検証を行った。その結果、“平坦性”を用いた視覚数理モデルによって、ヒトの知覚を再現できるパターンを見出した。さらに用いた数理モデルを神経生理学の観点から考察したところ、先行研究に矛盾しなかった。

キーワード：アモーダル補完、視覚数理モデル、標準正則化理論、曲率

1. 序論

ヒトは物体の一部が他の物体によって遮蔽された場合でも、被遮蔽領域の物体の形状を補完する（図1）。この視覚機能をアモーダル補完という。アモーダル補完は、心理物理学や神経生理学、さらにはこれらを数理的に記述する視覚数理モデル研究など、多くの分野で研究対象とされてきた（Gerbino, 2020）。

アモーダル補完における被遮蔽領域のように、対象とする空間領域の知覚が一意に定まらない、もしくはその候補が無数に存在する領域において知覚を一意に定める問題は、解が無数に存在する不良設定問題と捉えることができる（このような領域を本稿では不定領域と記す）。視覚系は不定領域に対して適当な制約条件を与え、不良設定問題を良設定問題（解が一意、または有限）に変換することで、最適化問題を解いていると解釈されてきた（Poggio, Torre & Koch, 1985）。このように、視覚系が最適化問題を解いているとされる解釈は標準正則化理論と呼ばれる。既存の多くの視覚数理モデル研究においては、この制約条件に空間的滑らかさを用いて、ヒトの視覚機能を統一的に定式化・説明できるとされてきた。この機能を数理的に捉えようと、不定領域における制約条件を対象とする物理量（明るさ、奥行きなど）の空間的滑らかさを用いて評価関数を定義し、評価関数が最小（もしくは最大）となる物理量を求める最適化問題と解釈できる。

[†] 愛知淑徳大学人間情報学部

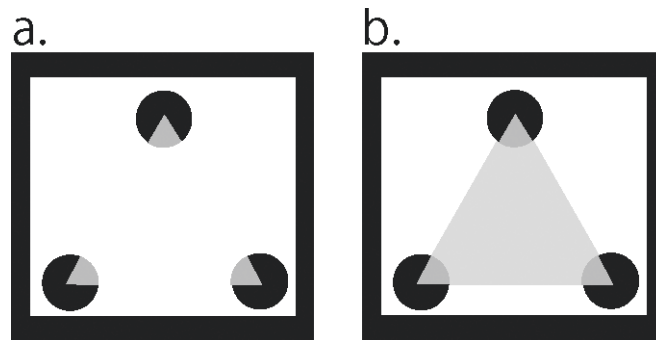


図1 アモーダル補完の例。灰色部は被遮蔽物体、白色部は遮蔽領域を示す。ヒトはaを観察した場合、遮蔽領域においてbのような形状を補完する。

アモーダル補完に関しては、心理物理学的には周囲の形状に応じて複数の補完，すなわち、大域的または局所的な補完がなされることが知られている (Boselie, 1988)。Ito, Onoda, Sakamoto, Kumada & Yano (2010) は、局所的補完と大域的補完 (図2) をともに再現可能なモデルを提案した。しかし、彼らがモデルに用いた制約条件は空間的滑らかさではなく対称性であった。一方、制約条件に空間的滑らかさを用いたアモーダル補完モデルも多数提案されている。例えば、Sarti, Malladi & Sethian (2000) は、形状検出アルゴリズム (Level Set Method) にヒトの形状補完特性を導入することでアモーダル補完モデルを提案した。また、物体と背景の境界領域における2交点間を最も滑らかな曲線で補完するアリゴリズムも提案されている (Mumford, 1994)。そして、これらのモデルに代表されるようにアモーダル補完に限らず、多くの視覚数理モデルの制約条件として、空間的“滑らかさ”が採用されてきた。

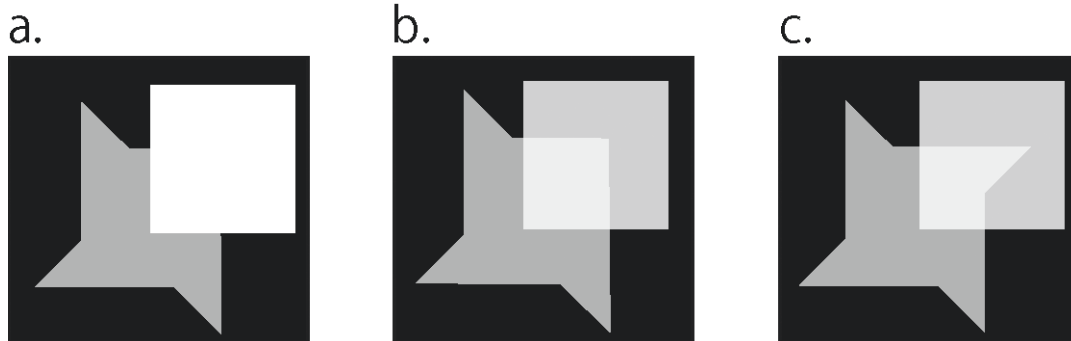


図2 局所的補完と大域的補完の例。灰色部は被遮蔽物体、白色部は遮蔽領域を示す。ヒトはaを観察した場合、遮蔽領域において、bまたは、cのような形状を補完する。b. 局所的補完。c. 大域的補完。

しかし、近年の研究において、多くの視覚数理モデルが制約条件に用いてきた空間的滑らかさは、最適ではない可能性が示唆されている。Ishikawa & Geiger (2006) は、心理実験によって、ヒトの視覚系は不定領域における奥行き情報の補完に用いる制約条件に3次元面曲率の1つであるガウス曲率=0を用いており、平坦な面を補完していることを見出した。また、Mitsukura & Satoh (2018) は、制約条件にレベルセットライン (図3b) とそれに直交するフローカーブラインの2種類の曲率と、ガウス曲率の関係を数学的に証明し、2種類の曲率によって評価される空間的“平坦性”を用いた奥行き情報補完モデルを提案した。さらに、このモデルを用いた数値シミュレーションの結果、不定領域においてガウス曲率=0となる面を補完し、上記の先行研究に矛盾しない結果が得られた。そして、この数理モデルは盲点領域における像の補完モデルとも数学的に等価であることから、像の補完においても、“平坦性”を用いていると考えられる。

以上のように、ヒトの視覚系が不定領域における像や奥行き情報を補完する際に用いる制約条件として、面の“滑らかさ”より“平坦性”の方が適している可能性がある。しかし、アモーダル補完の数理モデル研究

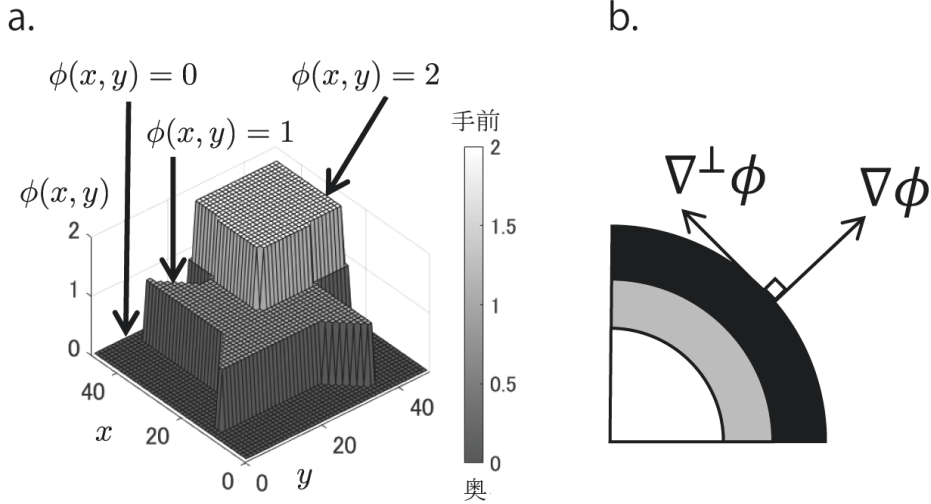


図3 本研究で用いた数理表現の概略図。a. 図2aを $\phi(x, y, 0)$ で記述した例。 $\phi(x, y) = 2$ は遮蔽領域、 $\phi(x, y) = 1$ は被遮蔽物体、 $\phi(x, y) = 0$ は背景領域を表す。b. 式(2)、式(3)の曲率に関する概略図。実線部は、3次元面 ϕ のレベルセットライン（等高線）を表す。

において、制約条件として“平坦性”を用いたモデルは提案されていない。もしヒトのアモーダル補完結果に矛盾しない数理モデルを、空間的平坦性を用いて構築できれば、ヒトの視覚系は視覚情報補完において空間的平坦性を用いているという予測に対する新たな科学的根拠になると期待される。

そこで、本研究では、ヒトの視覚系のメカニズムを統一的に説明可能な数理モデル構築のための予備的検討として、標準正則化理論の枠組みで神経生理学的に無矛盾なアモーダル補完モデルを提案することを目的とする。具体的には、先述の Mitsukura & Satoh (2018) が構築した奥行き情報の補完モデルをアモーダル補完に適用し、ヒトの知覚結果と比較する。さらに、神経生理学的観点からアモーダル補完モデルとしての妥当性も考察する。

2. 数理モデル

アモーダル補完において、遮蔽領域と被遮蔽領域の間に奥行き関係が存在することに着目し、これを $\phi(x, y)$ で記述する（図3a）。さらに、平坦性の評価関数として、Mitsukura & Satoh (2018) が提案した評価関数：式(1)を用いる。

$$E[\phi] = \iint_B (\bar{\kappa}^2 + \bar{\mu}^2) \cdot \|\nabla\phi(x, y)\| dx dy. \quad (1)$$

ここで、 $\bar{\kappa}(x, y)$ 、 $\bar{\mu}(x, y)$ は等高線、すなわちレベルセットライン $\kappa(x, y)$ （図3b）とそれに直交するフローカーブラインの曲率 $\mu(x, y)$ を用いて、それぞれ次式で定義される（Florack, Romeny, Koenderink & Viergever, 1992）。

$$\bar{\kappa}(x, y) = \kappa(x, y) \cdot \sqrt{\phi_x^2 + \phi_y^2} = \frac{\phi_y^2 \phi_{xx} - 2\phi_x \phi_y \phi_{xy} + \phi_x^2 \phi_{yy}}{\phi_x^2 + \phi_y^2} \quad (2)$$

$$\bar{\mu}(x, y) = \mu(x, y) \cdot \sqrt{\phi_x^2 + \phi_y^2} = \frac{(\phi_x^2 - \phi_y^2) \phi_{xy} + \phi_x \phi_y (\phi_{yy} - \phi_{xx})}{\phi_x^2 + \phi_y^2} \quad (3)$$

ただし、 ϕ の添え字は、 $\phi_x \triangleq \partial\phi/\partial x$ 、 $\phi_{xy} \triangleq \partial^2\phi/\partial x \partial y$ のように偏微分を表す。また、記載を平易にするため、右辺の座標 (x, y) を省略している。遮蔽領域 B において式(1)が最小となる ϕ が得られる更新測を導出するために、時間変数 t を導入し、再急降下法を式(1)に適用すると、以下のダイナミクスが得られる（Mitsukura

& Satoh, 2018)。

$$\frac{\partial}{\partial t} \phi = \nabla(\Delta\phi) \cdot \nabla^\perp \phi + \bar{\kappa}. \quad (4)$$

ただし、 $\nabla^\perp \phi$ は、 $\nabla\phi$ に対して直交するベクトルである (図 3b)。また、右辺の座標 (x, y) は省略している。遮蔽領域 B における補完結果は、式 (1) の評価関数: $E[\phi]$ が最小となる $\phi(x, y)$ であり、式 (4) の安定状態である。

このモデルを用いた数値シミュレーションには、Kanizaの三角形、Itoら (2010) 及び、Oliver, Haro, Dimiccoli, Mazin & Ballester (2016) を元に作成した画像を用いた。数値シミュレーション結果を図 4b, d, f に示す。数値シミュレーション図 4b のように、形状の補完結果が直線である場合、ヒトの知覚と相同な結果が得られた。しかし、図 2b のように予想される補完結果の形状が直線であっても、ヒトの知覚とは異なる形状が補完された。また図 4e のように円の一部分を遮蔽したパターンに対しても、直線状の形状が補完された。

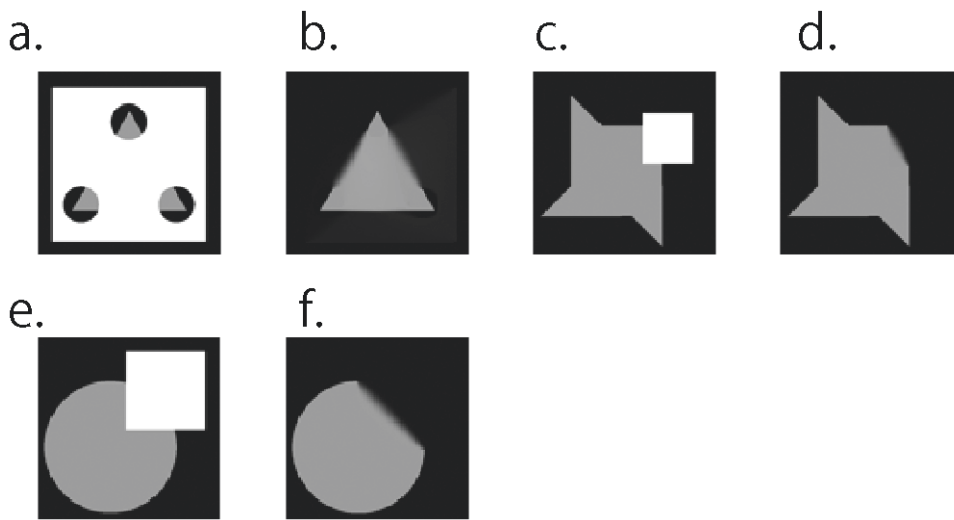


図 4 数理モデルを用いたシミュレーション結果。a, c, e. 数値シミュレーションの初期状態： $\phi(x, y, 0)$ 。白色部はモデルを適用した遮蔽領域 B 、灰色部は被遮蔽物体。b, d, f. 面の平坦性を最小化するモデル式 (4) を遮蔽領域に適用した結果。

3. 考察

本研究では 3 次元空間における面の平坦性が最小となる面を補完する奥行き情報補完モデルをアモーダル補完モデルとして用いて、アモーダル補完モデルの制約条件に平坦性を用いることの妥当性を検討した。その結果、図 4a のようなパターンの場合、ヒトの知覚を再現できることが示された。しかし、図 4c に対する局所的及び大域的補完 (図 2b, d) や、図 4e に対する補完を再現できなかった。これらの結果について、神経生理学及び、数理的観点から考察する。

Ban, Yamamoto, Hanakawa, Urayama, Aso, Fukuyama & Ejima (2013) は、fMRI を用いてアモーダル補完が V1 野や V2 野において表象されていることを示した。ゆえに、少なくとも本研究で用いたモデル式 (4) の各項が、V1 野及び V2 野において符号化可能な量で記述されなければならない。式 (4) は、V1 野及び V2 野における神経細胞で符号化可能な量を用いて構築された盲点における像の補完モデルと数学的に等価である (Mitsukura & Satoh, 2018)。すなわち、奥行きや像のみならず、V1 野及び V2 野の神経細胞で表象されている量を用いて、本研究で用いたモデル式 (4) を記述できると考えられる。したがって、本研究で用いた補完モデルは、ヒトのアモーダル補完結果の一部を再現可能であり、神経生理学的知見にも矛盾しないことが分かった。このことは、像、奥行き、形状の視覚情報補完は、同様のモデルで記述できる可能性を示唆している。

しかしながら、図 4c, e のようなパターンの補完を再現できなかった。この原因は本稿で用いたモデル式(4)の各項が、空間的隣接領域の量を用いて算出される微分量であるためであると考えられる。特に図 2c のように大域的な形状の特徴を用いた補完の再現は、原理的に不可能である。これらの補完を再現するには、V1 野、V2 野より高次の視覚野の情報を用いる必要があると考えられる。遮蔽領域を含む画像に対して選択的に反応を示す神経細胞が V4 野に存在する (Bushnell, Harding, Kosai & Pasupathy, 2011)。また、単一の神経細胞が情報を符号化する領域である受容野のサイズは、視覚野の階層が高いほど広がる。ゆえに、V4 野において大域的な形状の特徴が符号化されている可能性が高い。もし V4 野の特性を数理的に記述し、それを本モデルに導入できれば、大域的補完も再現可能なモデルを構築できると期待される。

したがって、V4 野の視覚機能を定式化し、それを標準正則化理論の枠組みで本稿のモデル式に導入することでヒトの大域的補完も再現可能なアモーダル補完モデルの構築を今後の課題とする。

謝 辞

本研究は愛知淑徳大学研究助成 (22TT01), JSPS 科学研究費 (JP21K18034) の助成を受けたものです。

参考文献

- Ban, H., Yamamoto, H., Hanakawa, T., Urayama, S.-I., Aso, T., Fukuyama, H., & Ejima, Y. (2013). Topographic representation of an occluded object and the effects of spatiotemporal context in human early visual areas. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 33(43), 16992–17007. <https://doi.org/10.1523/JNEUROSCI.1455-12.2013>
- Boselie, F. (1988). Local versus global minima in visual pattern completion. *Perception & Psychophysics*, 43(5), 431–445. <https://doi.org/10.3758/BF03207879>
- Bushnell, B. N., Harding, P. J., Kosai, Y., & Pasupathy, A. (2011). Partial Occlusion Modulates Contour-Based Shape Encoding in Primate Area V4. *Journal of Neuroscience*, 31(11), 4012–4024. <https://doi.org/10.1523/jneurosci.4766-10.2011>
- Florack, L. M. J., Romeny, B. M. ter H., Koenderink, J. J., & Viergever, and M. A. (1992). *Scale and the differential structure of images. 1*, 376–388.
- Gerbino, W. (2020). Amodal Completion Revisited. *I-Perception*, 11(4). <https://doi.org/10.1177/2041669520937323>
- Ishikawa, H., & Geiger, D. (2006). Illusory volumes in human stereo perception. *Vision Research*, 46(1–2), 171–178. <https://doi.org/10.1016/j.visres.2005.06.028>
- Ito, T., Onoda, T., Sakamoto, K., Kumada, T., & Yano, M. (2010). A Computational Model That Enables Global Amodal Completion Based on V4 Neurons. In K. W. Wong, B. S. U. Mendis, & A. Bouzerdoum (Eds.), *Neural Information Processing. Theory and Algorithms* (Vol. 5864, Issue PART 2, pp. 9–16). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-10684-2>
- Mitsukura, E., & Satoh, S. (2018). Computational study of depth completion consistent with human bi-stable perception for ambiguous figures. *Neural Networks*, 99, 42–55. <https://doi.org/10.1016/j.neunet.2017.11.016>
- Mumford, D. (1994). *Elastica and Computer Vision BT-Algebraic Geometry and its Applications: Collections of Papers from Shreeram S. Abhyankar's 60th Birthday Conference* (C. L. Bajaj (ed.); pp. 491–506). Springer New York. https://doi.org/10.1007/978-1-4612-2628-4_31
- Oliver, M., Haro, G., Dimiccoli, M., Mazin, B., & Ballester, C. (2016). A Computational Model for Amodal Completion. *Journal of Mathematical Imaging and Vision*, 56(3), 511–534. <https://doi.org/10.1007/s10851-016-0652-x>
- Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature*, 317(6035), 314–319. <http://dx.doi.org/10.1038/317314a0>
- Sarti, A., Malladi, R., & Sethian, J. A. (2000). Subjective surfaces: A method for completing missing boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), 6258–6263. <https://doi.org/10.1073/pnas.110135797>