

動的輪郭法のエネルギー解析による図地反転現象の考察 A Study of Perceptual Rivalry by Energy Analysis of Active Contour

佐藤 俊治[†]
Shunji Satoh

三宅 章吾[‡]
Shogo Miyake

1. まえがき

ルビンの壺に代表される図地反転現象は、視覚情報処理特性を示す非常に興味深い現象であり、これまでに多くの研究が行なわれてきた。これらの研究方法と目的は主に (a) 心理物理実験による視覚特性の測定、および得られた特性を説明する視覚アルゴリズム等の提案、(b) 神経生理学などで得られた知見を導入した神経回路モデルの提案、および数値実験によるモデルの妥当性の確認であった。

しかしながら、これまでに提案されてきた視覚アルゴリズムやモデルでは、入力画像にノイズがなく、図地検出に必要な情報があらかじめ与えられている等、理想的な状況を仮定して実験・検証が行なわれてきた。また、動作の理論解析が困難であるために定量評価が難しく、ノイズ・色・テクスチャなどの情報を含む自然画像に適用する際に問題が生じる。

従って本研究では、(i) 画像工学的立場から提案された物体検出のための画像処理方法(動的輪郭法)に着目し、これを実現する神経回路モデルを作成する。次に(ii) 作成したモデルによって図地反転が再現されることを数値実験により示す。また、(iii) 視覚心理実験で得られている知見との対応の調査を目的とする。

2. 動的輪郭法に基づく物体検出方法

2.1 動的輪郭法

動的輪郭法は、2次元画像の輝度 $I(\vec{x})$ ($\vec{x} = (x, y) \in \mathbb{R}^2$) を入力とし、物体の境界 $C(q) = (x_C(q), y_C(q)) \in C$ ($q \in [0, 1)$)、物体の外側(背景: 領域 B と記す)と内側(図: 領域 S)を求める方法である。また、物体の内外を表現するために2次元曲面 $\Phi(\vec{x})$ が導入され、各量が次式で表現される。

$$C = \{\vec{x} | \Phi(\vec{x}) = 0\}, \quad (1)$$

$$S = \{\vec{x} | \Phi(\vec{x}) > 0\}, \quad B = \{\vec{x} | \Phi(\vec{x}) < 0\}, \quad (2)$$

この表現方法は level set method と呼ばれている。

物体の境界 \hat{C} を求めるためにまず、 C のエネルギー $E_C[C(q, t)]$ を定義する (t は時刻)。次に適当な初期値 $C(q, t=0)$ を設定し、 E_C が小さくなるように C を動的に変化させる。境界 \hat{C} は $\hat{C} = \{C(q, t) | \partial C / \partial t = \vec{0}\}$ で与えられる。Caselles らは次のエネルギーを提案した [1]。

$$E_C[C] = \oint g(|\nabla I(C(q, t))|) \left| \frac{\partial}{\partial q} C(q, t) \right| dq. \quad (3)$$

関数 g は単調減少関数である。ここで次の問題点が生じる。定常解 \hat{C} を求める際、初期値によっては \hat{C} が E_C の局所最小解に陥り、期待する結果が得られない。

[†]東北福祉大学, IEICE, IEEE

[‡]東北大学, IEICE, JNNS, JPS

2.2 動的輪郭法の改良

我々は上記問題点を解決するために (3) の最小化で得られる C の定常解 C_0 を求め、

$$C_0 = \{\vec{x} | I_{\xi\xi}(\vec{x}) = 0\}, \quad (4)$$

で与えられることを示した [2]。この結果から、 $C(q, t=0) \in C_0$ とする方法を提案した。ただし I_{ξ} は $\hat{\xi}$ 方向の方向微分を示し、 $\hat{\xi} = \nabla I / |\nabla I|$ である。我々はこの解析結果と level set method を用い、 Φ のエネルギー $E_{\Phi}[\Phi]$ を次式で与えた。

$$E_{\Phi}[\Phi] = E_{\Phi}^1[\Phi] + E_{\Phi}^2[\Phi] + E_{\Phi}^3[\Phi], \quad (5)$$

$$E_{\Phi}^1[\Phi] = \frac{1}{2} \iint J(\vec{x}) |\nabla \Phi(\vec{x}, t)|^2 d\vec{x}, \quad (6)$$

$$J(\vec{x}) = \begin{cases} -1, & \text{if } \vec{x} \in C_0, \\ +1, & \text{otherwise.} \end{cases} \quad (7)$$

$$E_{\Phi}^2[\Phi] = g(|\nabla I(\vec{x})|) \frac{\lambda_{\text{NE}}}{2} \iint |\nabla \Phi(\vec{x}, t)|^2 d\vec{x}, \quad (8)$$

$$E_{\Phi}^3[\Phi] = \frac{\lambda}{2} \iint \{\Phi(\vec{x}, t)\}^2 d\vec{x}. \quad (9)$$

ただし λ_{NE} と λ は正の定数である。エネルギー E_{Φ}^1 は (1) で与えられる C を画像中の物体境界と一致させる項である。エネルギー E_{Φ}^2 は C が画像中のノイズにトラップされることを防ぐ項であり、 E_{Φ}^3 は Φ の発散を防ぐ項である。曲面 Φ は E_{Φ} が減少するように動的に変化する。具体的にはエネルギー最小化の原理を用い、 Φ に関する反応拡散方程式を求めることで Φ のダイナミクスが得られる。なお (5) は (3) の解析から自然に導かれる。

初期値 $\Phi(\vec{x}, t=0)$ について考察する。まず (1) と (4) より、 $\Phi(\vec{x} \in C_0, t=0) = 0$ と設定すればよい。次に $\Phi(\vec{x} \notin C_0, t=0)$ について考察する。視覚心理実験で得られた知見によると、画像中のエッジが凸形をしている場合には、凸の内側が図として知覚されやすいことが知られている(凸性)。そこで凸の内側の位置に対して $\Phi(\vec{x}, t=0) > 0$ 、外側の位置に対して $\Phi(\vec{x}, t=0) < 0$ となるように初期値 $\Phi_0(\vec{x}) = \Phi(\vec{x}, t=0)$ を次式で与える。

$$\Phi_0(\vec{x}) = H(I_{\xi\xi} I_{\eta\eta} I_{\xi}^2) |\tilde{\kappa}_I(\vec{x})|. \quad (10)$$

ただし、関数 H は $H(0) = 0, H(y \neq 0) = \text{sgn}(y)$ ($y \in \mathbb{R}$) で定義され、方向 $\hat{\eta}$ は $\hat{\xi}$ を反時計回りに 90° 回転させたベクトルである。また $\tilde{\kappa}_I$ はエッジ強度で正規化した曲率関数であり、 $\tilde{\kappa}_I = I_{\xi}^2 I_{\eta\eta}$ で定義される。図 1(a) を入力 I とした時の、 $\Phi_0(\vec{x})$ を図 1(b) に示す。ただし、 Φ の値を輝度に置き換えて表示している。画像は全て 64×64 (pixel²) である。

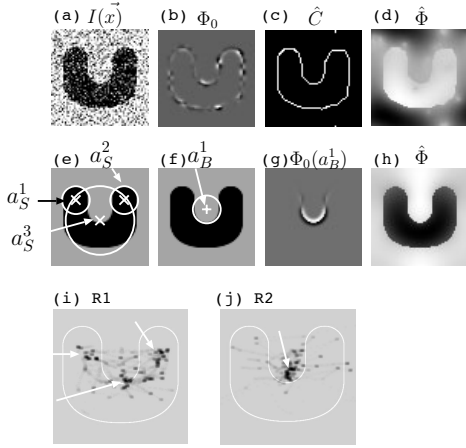


図 1: (a)~(h): 入力画像と諸量 (本文参照) . (i), (j): R1 または R2 を知覚しているときの注視位置 .

3. 物体検出モデル

3.1 物体検出モデル

値 Φ は図となる領域 (S) で正の値, 背景となる領域 (B) で負の値をとるように変化する. そこで値 Φ を, 図領域 (物体領域) に対して選択的に反応する細胞の活動度と読み換えて物体検出モデルを導出する. 具体的には (5) に対応する Φ の反応拡散方程式を求め, 得られた方程式に表われる位置 \vec{x} および時刻 t を離散化することで Φ のダイナミクス $\tau_{\Phi}(\Delta\Phi/\Delta t)$ が得られる. τ_{Φ} は時定数であり, $\tau_{\Phi} = 1$ (msec) とした.

図 1(a) を入力画像 I としたときの, C の定常状態と, 対応する Φ を図 1(c) と (d) に示す. 時刻 $t = 190$ (msec) で $\partial C/\partial t = \{0\}$ となり「U 字の物体」が検出された. 以降 $\partial C/\partial t = \{0\}$ であるときの $\Phi(\vec{x}, t)$ を $\hat{\Phi}(\vec{x})$ と記す.

3.2 物体検出モデルの解析

エネルギー E_{Φ} の観点から物体検出モデルの動作解析を行ない, 物体検出モデルによって図地反転現象が再現される可能性を述べる. 図 2 は, エネルギー E_{Φ} , Φ の初期値, および初期値に対応する物体領域検出結果を模式的に示している. 例えば 図 2(P0) は Φ_0 を図示しており, エネルギー最小化の原理により (R1): 「U 字の物体が存在する」という結果を得る. (R1) を得ているときの Φ を $\hat{\Phi}^{R1}$, $E_{\Phi}[\hat{\Phi}^{R1}] = E_{\Phi}^{\min}$ と書くことにする.

ここで (5) で定義されるエネルギーが, Φ の偶関数であることに注意すると, $\hat{\Phi}^{R2} = -\hat{\Phi}^{R1}$ についても $E_{\Phi}[\hat{\Phi}^{R2}] = E_{\Phi}^{\min}$ が成立することがわかる (図 2(R2)). 従って Φ_0 とは異なる初期値を用いることで (R2): 「U 字型に穴があいた物体が存在する (喉の奥を見ているような知覚)」という図地反転結果が得られるはずである.

そこで, Φ_0 とは異なる初期値として $\Phi_0(\vec{x}; \vec{\mu}, \sigma)$ を次式で定義し, (R2) が得られることを確認する.

$$\Phi_0(\vec{x}; \vec{\mu}, \sigma) = \Phi_0(\vec{x}) \exp\left(\frac{-|\vec{x} - \vec{\mu}|^2}{2\sigma^2}\right). \quad (11)$$

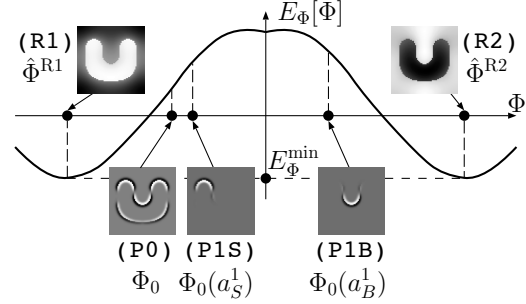


図 2: エネルギー E_{Φ} , Φ の初期値, および検出された物体領域の関係

図 1(f) に示すように U 字の窪みの位置を $\vec{\mu} = \vec{\mu}_B^1$, $\sigma = \sigma_B^1 = 8$ とする. このときの $\Phi_0(\vec{x}; \vec{\mu}_B^1, \sigma_B^1)$ を, 図 1(g) に示す (パラメータの組 $(\vec{\mu}_B^1, \sigma_B^1)$ を a_B^1 と記す). 初期値 $\Phi_0(\vec{x}; a_B^1)$ に対する物体領域検出結果を図 1(h) に示す. この結果から, 異なる初期値で (R1) と逆の (R2) が得られることがわかった.

パラメータ $a = (\vec{\mu}, \sigma)$ は空間的に不均一な強調・抑制を Φ に与えるが, これらパラメータを視覚的注意のメタファーの 1 つであるズームレンズモデルの中心位置 $\vec{\mu}$ と範囲 σ のモデル化であると解釈し, 以降パラメータ $a = (\vec{\mu}, \sigma)$ を視覚的注意と呼ぶことにする. また, $a^0 = (\cdot, \infty)$ とすることで注意は空間的に均一となり, $\Phi_0(\vec{x}; a^0) = \Phi_0(\vec{x})$ である. 以降 $a^0 = (\vec{\mu}, \sigma)$ を視覚的注意がない状態と呼ぶことにする.

4. 注意移動

4.1 特徴抽出

物体の詳細な情報を得るためには, 工学的な要請として, 検出された物体 (検出された図領域 S) を構成する様々な特徴を抽出・解析する必要がある. 例えば (R1) に対しては図 1(e) にあるように, U 字を構成する特徴の位置と固有スケール $a_S^n = (\vec{\mu}_S^n, \sigma_S^n)$, $n \in \{1, 2, \dots\}$ を計算し, 視覚的注意 a_S^n によって物体の特徴を抽出する必要がある. 実際, 視覚的注意は物体の特徴に向けられ易い傾向にあることが知られているため, 以降 a_S^n を物体の特徴とも呼ぶ.

次に a_S^n の計算方法について述べる. 筆者らは先に, スケールスペース解析と視覚特性を反映した注意移動のモデルを構築した. 詳細は文献 [3] を参照していただきたいが, 物体の特徴 $a_S^n = (\vec{\mu}_S^n, \sigma_S^n)$ は基本的に次式によって計算することができる.

$$\{(\vec{\mu}_S^n, \sigma_S^n)\} = \arg\text{LocalMax}_{\{\vec{\mu}_g, \sigma_g\}} [L(\vec{\mu}_g; \sigma_g)], \quad (12)$$

$$L(\vec{\mu}_g; \sigma_g) = -\sigma_g^2 \nabla^2 g(\vec{\mu}_g; \sigma_g) * \hat{\Phi}(\vec{\mu}_g) \quad (13)$$

ただし関数 g は平均 $\vec{0}$, 分散 σ_g^2 の 2 次元ガウス関数, 演算子 $*$ は畳み込み演算である. また, 関数 $\arg\text{LocalMax}$ は L が局所最大値を取る複数の $\{(\vec{\mu}_S^m, \sigma_S^m)\}$, $m \in \{1, 2, \dots\}$ を計算し, $L(\vec{\mu}_S^n; \sigma_S^n) > L(\vec{\mu}_S^{n+1}; \sigma_S^{n+1})$ となるように $\{(\vec{\mu}_S^m, \sigma_S^m)\}$ を並べ変える関数である. 例えば,

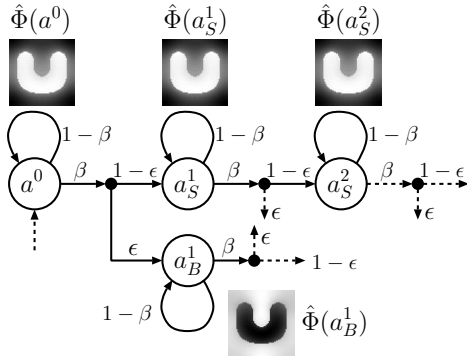


図 3: 注意移動の状態遷移図

(12) を図 1(d) に示す $\hat{\Phi}$ に適用すると、図 1(e) が得られる。正の整数 n は特徴の重要度を表現しており、特徴 a_S^n は a_S^{n+1} よりも重要度が高いことを意味する。

背景の特徴 a_B^n は (12) における L の符号を反転させることで得られる。なお、 $a_B^1 = (\bar{\mu}_B^1, \sigma_B^1)$ を図 1(f) に示すが、これは 3.2 で述べた U 字の窪みの位置と固有スケールである。

4.2 注意移動

注意の移動機構を物体検出モデルに導入する。ある特徴に注意が向けられている時間 T が t 以下である確率 $\Pr(T < t; \gamma)$ を次式で与える。

$$\Pr(T < t; \gamma) = 1 - \exp(-t/\gamma). \quad (14)$$

ただし γ は正の定数であり、時間の次元を持つ。式 (14) は、平均で γ 単位時間後に注意の移動が発生し、逆に、単位時間内に $\beta = \gamma^{-1}$ 回の確率で注意の移動が発生することを意味する。値 β は入力 I に含まれる特徴によって異なる値を取ると予測されるが、本研究では簡単のために $\beta = 2.0$ (sec^{-1}) の定数とした。特徴 $a = (\bar{\mu}, \sigma)$ に注意が移動すると、新たな初期値が (11) で設定され、再び Φ の計算が行なわれる。

次に、物体の特徴 $\{a_S^n\}$ への注意と背景の特徴 $\{a_B^n\}$ への注意が、 $1 - \epsilon: \epsilon$ の割合で発生すると仮定する。ただし $0 < \epsilon < 0.5$ の定数である。これは物体の特徴に選択的に注意が向けられ易いという視覚特性を反映するものである。また、重要度が高い特徴へ優先的に注意が向けられるようにする。加えて復帰抑制効果を取り入れ、一度注意を向けた特徴には再び注意が向けられないようにする。全ての特徴に注意が向けられた後は状態 a_0 に戻る。

以上の注意移動機構を状態遷移図で示したのが図 3 である。状態はどの特徴に注意を向けているのかを示す。状態遷移が微小時間 $\Delta t = 1$ (msec) で発生すると仮定すると、 $\beta = 0.002$ (msec^{-1}) である。

5. 数値実験・考察

入力画像 I を図 1(a) とし、(R1) または (R2) が得られている時間幅 T_{R1} および T_{R2} を測定し、視覚心理実験によって得られた時間幅に関する知見との対応を調査・

考察する。パラメータは $\lambda = 2.0$, $\lambda_{NE} = 0.5$, $\epsilon = 0.05$ とし、図地反転が繰り返し発生することを確認した。

注意 a が与えられているときの $\hat{\Phi}$ を $\hat{\Phi}(a)$ と記し、図 3 に示す。この結果から特徴 a_S^n に注意が向けられているときは物体領域検出結果として (R1)、特徴 a_B^n に注意が向けられているときは (R2) が得られることがわかる。また、時間幅 T_{R1} と T_{R2} はそれぞれ、Erlang 分布 $\text{Er}(T; 4, \gamma)$ と $\text{Er}(T; 1, \gamma)$ で近似できることがわかった (χ^2 適合度検定, $p < 0.10$)。ただし $\text{Er}(T; \alpha, \gamma) = \frac{\gamma^{-\alpha}}{(\alpha-1)!} T^{\alpha-1} e^{-T/\gamma}$ である。同様の結果は、入力画像 I をルビンの壺とした場合にも確認することができた。なお、 $\epsilon \rightarrow 0$ とすると、時間幅は厳密に Erlang 分布に従うことが簡単な計算から導くことができる。

視覚心理実験により実際に得られる時間幅は Γ 分布 $\Gamma(T; \alpha, \gamma) = \frac{\gamma^{-\alpha}}{\Gamma(\alpha)} T^{\alpha-1} e^{-T/\gamma}$ で近似できることが知られており、さらに Murata らは、 α が自然数であることを見出した [4]。Erlang 分布は、 Γ 分布のパラメータ α を自然数にすることで得られる分布であるため、提案モデルの図地反転に関する時間幅と、視覚心理実験で得られた時間幅に関する知見が一致しているといえる。以上の結果から、 α は注意を与えられた物体領域中の平均特徴数、 γ はある特徴に注意を与えている平均時間と解釈することができる。

次に $\bar{\mu}$ と注視位置の相関を、眼球運動測定装置 ISCAN を用いて調査した。図 1(i) と (j) はそれぞれ、(R1) または (R2) が得られているときの注視位置を測定した結果である (被験者 TS は 20 歳男性。提示パターンまでの距離 145cm。パターンの視野角 $14.9^\circ \times 17.5^\circ$ 。測定時間 60sec。) 図 1(i) と (j) の結果から、注視位置は (R1) と (R2) で明らかに異なることがわかる。また、(12) で得られる $\bar{\mu}_S^n$ (または $\bar{\mu}_B^n$) は、ヒトの注視位置と相関が高いことがわかる。しかし、注意位置と注視位置は異なる場合があるため、測定した注視位置の違いを注意位置の違いと結論付けることはできない。また、Murata らの実験では注視位置を固定している。したがって、図地反転現象と注視・注意の関係を明確にするにはより詳細な実験と検証が必要となる。この問題については今後の課題とする。

6. まとめ

物体領域検出モデルをエネルギーの観点から解析し、注意移動機構を取り入れることで、図地反転現象が再現された。今後はより詳細な視覚心理実験を行ない、提案モデルの妥当性を調査する予定である。

謝辞 本研究の一部は文部科学省科学研究費補助金 (No. 14780254) の援助を受けて行なわれた。

参考文献

- [1] V. Caselles, R. Kimmel and G. Sapiro, Int. J. Computer Vision, vol. 22, no. 1, pp. 61–79, 1997.
- [2] 佐藤 俊治, 三宅 章吾, “動的輪郭法の計算原理に基づく物体検出モデル,” 信学論, 2004 (投稿中).
- [3] 佐藤 俊治, 三宅 章吾, 信学論, vol. J86-D-II, no. 10, pp. 1490–1501, 2003.
- [4] T. Murata *et al.*, NeuroReport, vol. 14, no 10, pp. 1347–1352, 2003.