

反応拡散ネットワークによるステレオ視差の検出

野村 厚志[†], 一川 誠[‡], 三池 秀敏[‡]

山口大学教育学部[†], 山口大学工学部[‡]

1 はじめに

ステレオ画像から視差を検出する多くのモデルでは、視差分布に関する空間的な連続条件と、各画素位置における視差の唯一条件を課す。物体の端を除けば、物体表面の隣接領域では、奥行き視差は連続的に変化する。また、半透明の物体でない限り、画像の各画素位置においては、1つの視差値しか持ち得ない。MarrとPoggioは、ステレオ視差検出の問題において、連続条件と唯一条件を課し、興奮・抑制メカニズムによる協調モデルを提案した [1]。

ここでは、反応拡散ネットワークを用いた視差検出モデルを提案する。著者らは、既に反応拡散ネットワークが視覚の群化機能のモデルとして有効であることを確認しており [2]、さらに視差検出においても適用可能であると考えている。そこでまず、MarrとPoggioの協調モデルを1変数の反応拡散ネットワークで実現する。連続的な視差の分布を、複数の視差レベルに離散化し、各視差レベル・各画素位置において、MarrとPoggioの提案したセルに対応する反応項を導入する。これらの反応項を拡散で空間的に結合することによって、視差の連続条件を満足する反応拡散ネットワークを構成する。さらに、各視差レベルを支配する反応拡散ネットワークを相互抑制的に結合し、唯一条件を課す。

本研究では、上で実現した反応拡散ネットワークによるモデルに、新たに第2の変数に関する反応拡散ネットワークを結合させる。第2の変数は第1の変数の反応拡散ネットワークに対して自己抑制的に機能する。この2変数のネットワークモデルは、FitzHugh-Nagumo方程式 [3, 4] に、抑制性の拡散項を加えて拡張したものである。1変数の反応拡散ネットワークモデルでは、視差の定まらない領域については、既に視差の定まった隣接領域を、反応拡散によって拡張させることによって埋める。連続条件による視差分布の埋め尽くし処理は、微細ではあるが重要な特徴を誤って埋め尽くす。第2の変数の反応拡散ネットワークによって、視差分布の埋め尽くし処理を自己抑制し、微細な構造を残す。提案モデルを実画像に対して適用し、その有効性を確認する。

2 MarrとPoggioの協調モデル

MarrとPoggioの協調モデル [1] では、離散的な視差レベル $d = 0, 1, 2, \dots, N_d - 1$ を持つ視差分布地図の各画素位置 (x, y) にセルを考え、セルの状態: $S^t(x, y, d)$ を次式によって更新する。

$$S^{t+1}(x, y, d) = \sigma \left(\sum_{\Omega} S^t - \varepsilon \sum_{\Theta} S^t + C \right) \quad (1)$$

ここで、 Ω は視差の連続条件のための空間局所領域を、 Θ は唯一条件のための視差方向の領域を表す。 $\sigma(\cdot)$ は0または1の値をとる閾値関数、 ε は唯一条件のための抑制定数である。 $C(x, y, d)$ は、ステレオ画像を x 軸 (水平軸) に平行に視差 d だけずらして重ね合わせた各点 (x, y) を囲む局所領域において計算される類似度 (例えば相互相関係数) である。 (x, y) に位置するセルを囲む領域 Ω 内のセルが興奮状態となると、そのセルも興奮状態へと変化する。一方、同一画素位置における他の視差値を持つ Θ 内のセルが興奮状態となると、そのセルの興奮を抑制する。

3 反応拡散ネットワーク

拡散方程式に反応項を加えたものを反応拡散方程式という。変数 $u(x, y, t)$ に関して空間的に一様な拡散係数 D_u を持つ拡散方程式は、

$$\partial_t u = D_u \nabla^2 u \quad (2)$$

ここで、 $\partial_t = \partial/\partial t$ 、 $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ 。一方、 u の時間変化に関する常微分方程式: $u' = f(u)$ を考える [$f(u)$ を反応項という]。 $f(u) = u(1-u)(u-a)$ とすると、 $0 < a < 1$ のとき u の初期値に対して閾値 a で分ける閾値方程式となる。MarrとPoggioの提案した興奮・抑制の考え方を反応拡散方程式に導入すると、視差 d の変数 u_d について次式となる。

$$\partial_t u_d = D_u \nabla^2 u_d + \frac{1}{\epsilon} u_d (1 - u_d) (u_d - a) + rC \quad (3)$$

但し、 ϵ, r は定数で $0 < \epsilon \ll 0$ 、 a は次式により決める。

$$a = \frac{1}{4} [1 + \tanh(u_{\max} - a_0)] \times \frac{1}{2} [1 + \tanh(d_a)] \quad (4)$$

*Stereo disparity detection with reaction-diffusion networks
Atsushi Nomura (Faculty of Education, Yamaguchi University),
Makoto Ichikawa and Hidetoshi Miike (Faculty of Engineering,
Yamaguchi University)

$$u_{\max} = \max_{d' \in \Theta} u_{d'}, \quad d_a = |d - \arg_d(u_{\max})| \quad (5)$$

ここで a_0 は定数。 d_a は u_{\max} を有するセルに対応付けられた視差レベル d と、 u_d の視差レベル d との差：すなわち距離を表す。式 (3) における拡散項は隣接領域に興奮性を伝播させる (連続条件)。また、式 (4),(5) を通じて異なる視差レベル間で相互に抑制する (唯一条件)。

本研究では、もう一つの変数: v_d を加え、各視差レベル内で自己抑制的な効果を持つ次の 2 変数反応拡散方程式を導入する [2]。

$$\partial_t u_d = D_u \nabla^2 u_d + \frac{1}{\epsilon} [u_d(1-u_d)(u_d-a) - v_d] + rC \quad (6)$$

$$\partial_t v_d = D_v \nabla^2 v_d + u_d - b v_d \quad (7)$$

ここで、 D_u は拡散係数、 b は定数である。 $D_v = 0$ のとき、式 (6),(7) は FitzHugh-Nagumo 方程式として知られている [3, 4]。 a が一定の場合、 u_d が増加すると式 (7) によって v_d が増加し、さらに v_d が増加すると式 (6) によって u_d が減少する。 $D_u < D_v$ の場合、 v_d の拡散が u_d の拡散よりも大きい場合自己抑制的となる。提案式 (6),(7) を条件 $D_u < D_v$ でステレオ画像に対して適用する [5]。

提案式は空間刻み: $\delta x, \delta y$ 、時間刻み: δt で差分化する。ラプラシアン演算子 ∇^2 は 5 点近似及び Crank-Nicolson 法により差分化する。初期値は空間領域の至る所で全ての視差レベルに対して $u_d = v_d = 0$ とし、領域の境界はノイマン条件を課す。得られる線形方程式を Gauss-Seidel 法を用いて解き、解 (u_d, v_d) を得る。視差分布 $D(x, y)$ は、各画素点において $u_d > 0.5$ を満たす時間長で重み付け平均して求める。すべての d に対して $u_d > 0.5$ を満たさない画素位置では視差は未定とする。提案モデルの式 (6),(7) は、相互に結合された $2 \times N_d$ 層の反応拡散ネットワークを形成している。

4 実験結果とまとめ

提案モデルの有効性を確認するため、試験画像としてよく利用される "University of Tsukuba" のステレオ画像を解析した [6, 7]。図 1(a), 1(b) はその左右のステレオ画像を表し、図 1(c) は真の視差分布図である。ステレオ画像を視差 $d = 0, 1, 2, \dots, 19 (N_d = 20)$ まで変化させながら重ね合わせ、各画素位置を中心とした上下左右の 5 点を用いて相互相関係数 $C(x, y, d)$ を得た。提案モデルを $t = 50$ まで計算し、得られた視差分布地図 $D(x, y, t = 50)$ を図 1(d) に示す。文献 [6] の図 17 に他の多くのアルゴリズムによる解析結果が示されている。これらと本解析結果を比較しても、概ね良好な視差分布を得ている。但し、特に背景右上角の領域においてはテクスチャが少ないため、誤った視差を得ている。

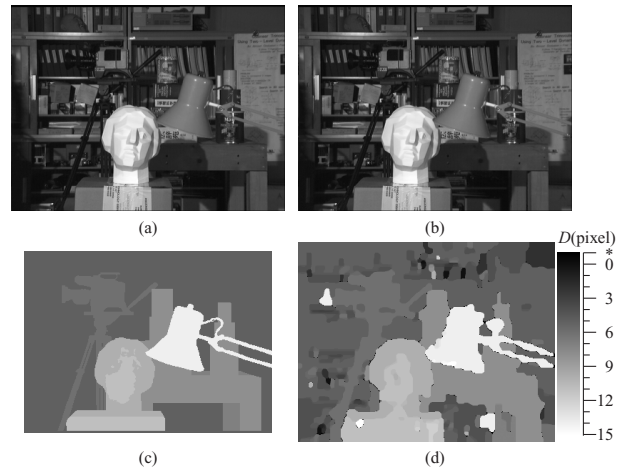


図 1: 試験画像: "University of Tsukuba" [6, 7] の解析結果。(a) 左画像、(b) 右画像。画像のサイズ: 384×288 (pixel)。但し、画像の端を切り取り、 360×280 (pixel) のサイズで解析した。(c) 真の視差分布図。(d) 提案法によって得られた視差分布図。* 印で示されたレベルの領域は視差が未定となったことを表す。モデルパラメータは $D_u = 0.5, D_v = 1.5, a_0 = 0.15, b = 10, \epsilon = 1/100, r = 3.0$ 。離散化のための時間・空間間隔は $\delta t = 1/100, \delta x = \delta y = 1/5$ 。

提案モデルでは、2 つの拡散係数を $D_u < D_v$ とした。それに対して、 $D_u > D_v$ と設定した場合や、式 (3) で記述される 1 変数のみの反応拡散ネットワークを用いた場合、試験画像中の細かい構造 (デスクライトの腕の部分) が、図 1(d) ほど正しく検出されなかった。従って、2 つめの式 (7) による自己抑制効果が有効であることが示された。従来の手法では、各画素位置における異なる視差レベル間での相互抑制性が導入されてきた。本提案モデルは、微細な構造を保つために空間についても自己抑制性を導入することが有効であると指摘している。

参考文献

- [1] Marr, D. & Poggio, T. (1976) "Cooperative computation of stereo disparity," *Science* **194**, 283–287.
- [2] 野村厚志, 一川誠, 三池秀敏 (2004) "反応・拡散モデルによる群化過程の実現," 情報処理学会論文誌 (コンピュータビジョンとイメージメディア) **45**, 26–39.
- [3] FitzHugh, R. (1961) "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical J.* **1**, 445–466.
- [4] Nagumo, J., Arimoto, S. & Yoshizawa, S. (1962) "An active pulse transmission line simulating nerve axon," *Proc. IRE* **50**, 2061–2070.
- [5] Nomura, A., Ichikawa, M. & Miike, H. (2005) "Stereo vision system with the grouping process of multiple reaction-diffusion models," *Lecture Notes in Computer Science 3522, Part I*, 137–144.
- [6] Scharstein, D. & Szeliski, R. (2002) "A taxonomy and evaluation of dense two-frame stereo correspondence algorithm," *Int. J. Comp. Vis.* **47**, 7–42.
- [7] <http://www.middlebury.edu/stereo>