

領域ベースの隠れ変数を用いた決定論的画像領域分割

Deterministic Image Segmentation by use of Region-Based Hidden Variables

三好 誠司*
Seiji Miyoshi

岡田 真人†
Masato Okada

Abstract: In image processing via Bayesian inference based on MRF model, introduction of hidden variables is effective to preserve edges in the image. We derive the image segmentation algorithm based on the Potts-spin-type region-based hidden variables and the variational method. The algorithm is applied to both the synthesized images contaminated by Gaussian noise and the natural image. Experimental results show its effectiveness and robustness.

Keywords: MRF, image segmentation, region-based, variational method

1 はじめに

多数の変数とその変数間の無向性相互作用からなる系はマルコフ確率場 (MRF) と呼ばれ, 画像の確率モデルとして広く利用されている [1, 2, 3]. MRF に基づく画像処理においては, 事後分布を用いるベイズ推定がよく用いられる [2, 3, 4]. この場合, ベイズの定理が事後分布の式を与えてくれるが, 実際の数値計算を行う段階で計算量的困難に直面することが多い. そのようなときに変分法に基づく推論, 確率伝搬法, モンテカルロ法など機械学習や統計力学の分野で近年開発された計算手法 [3] が威力を発揮することになる.

MRF とベイズ統計に基づいた画像処理を行う場合, 画像の事前分布を素朴なガウス分布とすると画像中のエッジの表現がどうしても難しくなる. エッジを表現するためには隠れ変数の導入が有効である. たとえば Kanemura らは Tipping と Bishop により提案されたベイズ超解像 [5] にエッジを表す隠れ変数を導入することによってすぐれた超解像処理を行うことに成功した [6].

画像処理においてエッジを表現するための隠れ変数に

は境界ベース [7] と領域ベース [8, 9] の二つの方法がある. 境界ベースは画素と画素の間に, そこがエッジであるかどうかを表す隠れ変数を置いてゆく考え方である. これに対して領域ベースは各画素がどの領域に属するかを示す隠れ変数を画素ごとに貼り付ける方法である. 境界ベースの隠れ変数の場合, 境界線がなるべく途切れず, 境界線が多くなりすぎず, 境界線がクロスすることが起こりにくくなるように多くの拘束条件を設ける必要がある [7]. これに対して領域ベースの隠れ変数の場合, 境界が自然に閉じたループになるなど好ましい性質を多く持つ [9]. ただし, 領域ベースの画像処理は局所解に陥りやすいという欠点があるため, あまり使われていない.

MRF に統計力学的なアプローチを試みる場合, 2 値をとるイジングスピンの変数としてよく用いられる. 領域ベースの隠れ変数を用いて画像処理を行う場合, これを D 値に拡張したポッツスピンを用いることにより表現の自由度があがると期待される.

ところで, 画像をある一定の特徴を持つ小領域ごとに分割する問題は領域分割 (セグメンテーション) と呼ばれる [2, 10]. 領域分割は画像に含まれる対象物を抽出する手法であるということも可能で, その後の画像の認識や理解のための第一次画像処理として重要である. また, 網膜という 2 次元センサーの信号から 3 次元の現実世界を再構成するための第一歩でもあることから視覚の計算論の基礎としても重要である. Geman ら [8] は MRF に領域ベースの隠れ変数を導入し, シミュレーテッドアニーリングを用いたモンテカルロ法により画像領域分割

*関西大学 システム理工学部, 564-8680 大阪府吹田市山手町 3 丁目 3 番 35 号, tel. 06-6368-1119, e-mail miyoshi@ipcku.kansai-u.ac.jp,

Faculty of Engineering Science, Kansai University, 3-3-35, Yamate-cho, Suita, Osaka, 564-8680 Japan

†東京大学大学院 新領域創成科学研究科, 277-8561 千葉県柏市柏の葉 5-1-5, 理化学研究所 脳科学総合研究センター, 351-0198 埼玉県和光市広沢 2-1, e-mail okada@k.u-tokyo.ac.jp, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa, Chiba, 277-8561 Japan, RIKEN Brain Science Institute, 2-1, Hirosawa, Wako, Saitama, 351-0198 Japan

の実験を行った。Bratsolis ら [11] はポツツスピンを領域ベースの隠れ変数に用いて平均場近似に基づく決定論的な手法で画像領域分割を行った。ただし、彼らは画素値をポツツスピンの初期値決定のためだけに用いた。

本稿では、領域ベースのポツツスピン型隠れ変数を用いて変分法に基づく推論により画像領域分割を行う決定論的アルゴリズムを導出する。また、雑音が重畳した人工画像と自然画像にそのアルゴリズムを適用した結果について報告する。

2 アルゴリズムの導出

画像を $x = \{x_i\}, i = 1, \dots, N$, 隠れ変数であるポツツスピンを $\mathbf{K} = \{\mathbf{k}_i\}, i = 1, \dots, N$ とする。ここで N は画素数である。本稿では各ポツツスピンは D 次元ベクトル $\mathbf{k}_i = \{k_{id}\}, d = 1, \dots, D$ であり、

$$\mathbf{k}_i \in \{(1, 0, \dots, 0)^T, (0, 1, \dots, 0)^T, \dots, (0, 0, \dots, 1)^T\} \quad (1)$$

とする。

画像領域分割の問題は、同時事前分布 $p(x, \mathbf{K})$ と画像 x が与えられたときに各画素の隠れ変数を推定することである。すなわち、事後分布 $p(\mathbf{K}|x)$ を求めたいわけであるが、これを直接求めるのは通常は困難である。そこで変分法に基づく近似的な推論を行う [3]。

任意の確率分布 $q(\mathbf{K})$ に関して、

$$\mathcal{L}(q(\mathbf{K})) = \sum_{\mathbf{K}} q(\mathbf{K}) \ln \frac{p(x, \mathbf{K})}{q(\mathbf{K})} \quad (2)$$

と定義し、これと $q(\mathbf{K})$ と $p(\mathbf{K}|x)$ のカルバック-ライブラー ダイバージェンス

$$\text{KL}(q(\mathbf{K})||p(\mathbf{K}|x)) = - \sum_{\mathbf{K}} q(\mathbf{K}) \ln \frac{p(\mathbf{K}|x)}{q(\mathbf{K})} \quad (3)$$

を用いると一般に

$$\ln p(x) = \mathcal{L}(q(\mathbf{K})) + \text{KL}(q(\mathbf{K})||p(\mathbf{K}|x)) \quad (4)$$

と書ける。式 (4) の左辺が \mathbf{K} に関して定数であることに注意すると、カルバック-ライブラー ダイバージェンスが最小であるという意味で事後分布 $p(\mathbf{K}|x)$ にもっとも近い試験分布 $q(\mathbf{K})$ を求めるためには、 $\mathcal{L}(q(\mathbf{K}))$ を最大化するような $q(\mathbf{K})$ を見つければよいことがわかる。

一般の $q(\mathbf{K})$ に関してこの最大化を行うことは困難なので、因子化仮定

$$q(\mathbf{K}) = \prod_i q_i(\mathbf{k}_i) \quad (5)$$

を満たす試験分布 $q(\mathbf{K})$ の中で $\mathcal{L}(q(\mathbf{K}))$ を最大化する分布を求めることにする。式 (5) を式 (2) に代入し因子のひとつ $q_i(\mathbf{k}_i)$ に関する依存性を取り出して計算を進めると、 $\mathcal{L}(q(\mathbf{K}))$ を最大にする試験分布 $q_i^*(\mathbf{k}_i)$ は

$$\ln q_i^*(\mathbf{k}_i) = \mathbb{E}_{j \neq i} [\ln p(x, \mathbf{K})] + \text{const.} \quad (6)$$

と求まる。ここで $\mathbb{E}_{j \neq i}[\cdot]$ は $j \neq i$ であるすべての \mathbf{k}_j による分布 q での期待値を表す。

いま、式 (1) を考慮してエネルギー関数を

$$E(x, \mathbf{K}) = \sum_{l \sim m} (\mathbf{k}_l \cdot \mathbf{k}_m (x_l - x_m)^2 + (1 - \mathbf{k}_l \cdot \mathbf{k}_m) \lambda) \quad (7)$$

とおく。ここで $\sum_{l \sim m}$ は隣接画素対すべてに関する和を表す。すなわち、隣接する画素 l と m が同じポツツスピンを持つ場合は $(x_l - x_m)^2$ がエネルギー関数に加えられ、異なるポツツスピンを持つ場合は定数 λ が加えられるようにしておく。同時事前分布 $p(x, \mathbf{K})$ にボルツマン分布

$$p(x, \mathbf{K}) = \frac{1}{Z} \exp\left(-\frac{\rho}{2} E(x, \mathbf{K})\right) \quad (8)$$

を仮定すると、

$$p(x, \mathbf{K}) = \text{Ber}(\{\mathbf{k}_l \cdot \mathbf{k}_m\}|\nu) N(x|\mathbf{0}, (\rho \mathbf{A}_K)^{-1}) \quad (9)$$

となる。ここで、 $\text{Ber}(\cdot|\nu)$ は平均 ν のベルヌーイ分布であり、

$$\nu = \left(1 + \exp\left(-\frac{\lambda\rho}{2}\right)\right)^{-1} \quad (10)$$

である。また、 $N(x|\mathbf{0}, (\rho \mathbf{A}_K)^{-1})$ は平均 $\mathbf{0}$ 、共分散行列 $(\rho \mathbf{A}_K)^{-1}$ のガウス分布であり、 \mathbf{A}_K の要素は

$$A_{ij} = \begin{cases} \sum_{k \in \mathcal{N}(i)} \mathbf{k}_i \cdot \mathbf{k}_k, & (i = j) \\ -(1 - \mathbf{k}_i \cdot \mathbf{k}_j), & (i \sim j) \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

である。ここで、 $\mathcal{N}(i)$ は画素 i に隣接する画素の集合を表す。

式 (9) を用いて式 (6) の計算を実行すると最終的にポツツスピンの平均値に関する方程式が以下のように得られる。

$$\langle k_{id} \rangle = \frac{\exp\left(\sum_{j \in \mathcal{N}(i)} \frac{\rho}{2} \langle k_{jd} \rangle (\lambda - (x_i - x_j)^2)\right)}{\sum_{d=1}^D \exp\left(\sum_{j \in \mathcal{N}(i)} \frac{\rho}{2} \langle k_{jd} \rangle (\lambda - (x_i - x_j)^2)\right)} \quad (12)$$

式 (12) を反復法で解くことにより、ポツツスピンの平均値が得られる。

3 実験

図1は画素値が一様な3つの領域からなる画像にPSNRが24dBであるようなガウス雑音を重畳することにより作成した人工画像である。この人工画像に対して3値ポツツピンで領域分割を行った結果を図2~4に示す。ハイパーパラメータは $\rho = 130, \lambda = 0.2$ とした。これらの図は式(12)を反復法で解くことにより得られたポツツピン各要素の平均値を表している。白,黒はそれぞれ1,0に対応する。図2~4から,今回提案する方法でほぼ完全な領域分割が行えていることがわかる。

図5にこの画像の画素値のヒストグラムを示す。雑音の影響で暗い二つの領域は画素値はオーバーラップしており,閾値だけで領域分割を行う場合は,閾値を注意深くチューニングしたとしても良好な分割を行うことは無理であることがわかる。このことから今回提案するアルゴリズムが有効に機能していることがわかる。

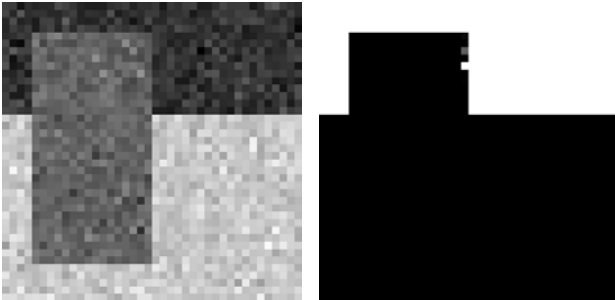


図1: ノイズを重畳して生成した人工画像。PSNRは24dB。

図2: $\{k_{i1}\}$



図3: $\{k_{i2}\}$

図4: $\{k_{i3}\}$

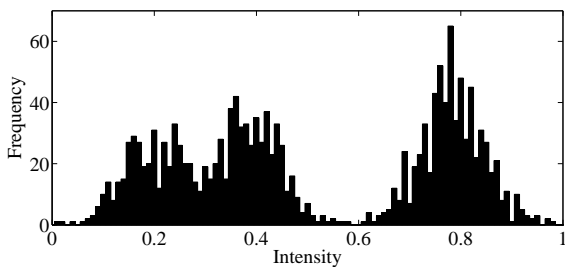


図5: 図1のヒストグラム

図6はPSNRを18dBにした人工画像である。この人工画像に対して3値ポツツピンで領域分割を行った結果を図7~9に示す。ハイパーパラメータは図1の場合と同じ $\rho = 130, \lambda = 0.2$ とした。

図10にこの画像の画素値のヒストグラムを示す。図1~5の場合よりも雑音を6dB増大したために三つの領域の画素値は大きくオーバーラップしており,閾値だけで領域分割を行うことは無理であることがわかる。この場合でも図7~9から比較的良好な分割が行えており,今回提案するアルゴリズムが頑健な性能を有していることがわかる。しかし,図7~9をよく見ると,雑音の影響で正しい分割が得られていない部分も存在する。また,一部に細かい“市松模様”がみられるが,これは雑音による誤分割というよりは,今回のアルゴリズムが局所解にトラップされていることを表すものと考えられる。

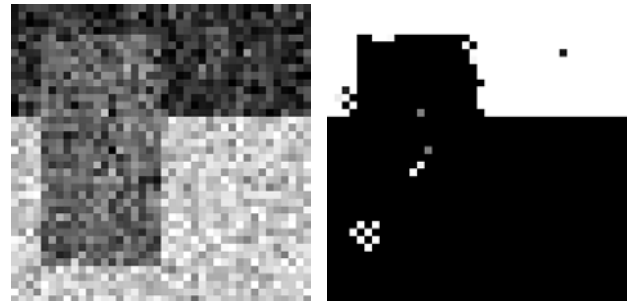


図6: ノイズを重畳して生成した人工画像。PSNRは18dB。

図7: $\{k_{i1}\}$



図8: $\{k_{i2}\}$

図9: $\{k_{i3}\}$

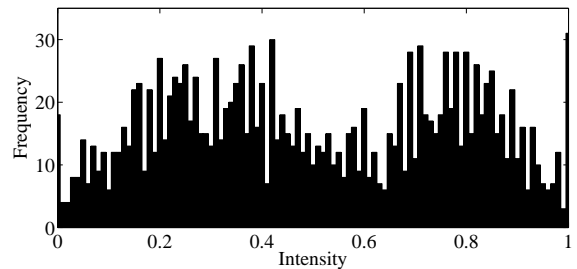


図10: 図6のヒストグラム

次に図11に示す自然画像を対象に領域分割を行った。

4 値ポッツスピンの場合の結果を図 12~15 に示す．ハイパーパラメータは $\rho = 130, \lambda = 0.015$ とした．これらの図も式 (12) を反復法で解くことにより得られたポッツスピンの各要素を表している．白，黒はそれぞれ 1,0 に対応するが，事後分布 $p(K|x)$ による平均値 $\langle k_{id} \rangle$ を計算してソフトな推定を行っているので図 12~15 の各画素は 1 (白) か 0 (黒) かの 2 値ではなく 0~1 のアナログ値 (グレー) をとっていることがよくわかる．これらの図から，今回提案する方法で比較的良好な領域分割が行えていることがわかる．



図 11: 自然画像

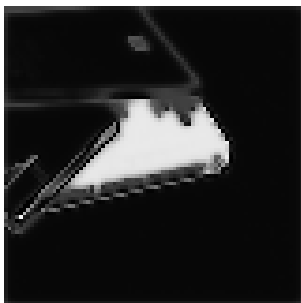


図 12: $\{\langle k_{i1} \rangle\}$

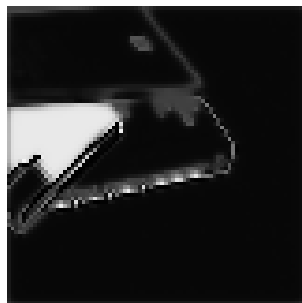


図 13: $\{\langle k_{i2} \rangle\}$

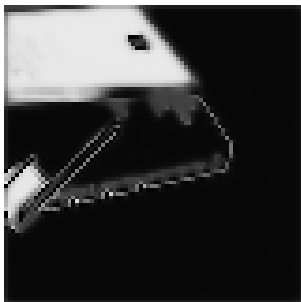


図 14: $\{\langle k_{i3} \rangle\}$



図 15: $\{\langle k_{i4} \rangle\}$

4 まとめと今後の課題

領域ベースのポッツスピン型隠れ変数を用いて変分法に基づく推論で画像領域分割を行う決定論的アルゴリズムを導出した．また，雑音が重畳した人工画像と自然画像にそのアルゴリズムを適用した．その結果，比較的良好な領域分割が行える可能性を示すことができた．しかし，ポッツスピンの平均値を計算するソフトな推定を

行っているにもかかわらず，局所解にトラップされやすいという欠点が現れた．この点の解決は今後の課題である．また，ポッツスピンの次元 D やハイパーパラメータ ρ, λ の自動設定も今後検討すべき課題である．

参考文献

- [1] Stan Z. Li, Markov Random Field Modeling in Image Analysis (Third Edition), Springer, 2009.
- [2] 田中和之, 確率モデルによる画像処理技術入門, 森北出版, 2006.
- [3] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
- [4] K. Tanaka, "Statistical-mechanical approach to image processing," J. Phys.A: Math. Gen., **35**, R81-R150, 2002.
- [5] M. E. Tipping and C. M. Bishop, "Bayesian Image Super-resolution," Advances in Neural Information Processing Systems, **15**, 1279-1286, 2003.
- [6] A. Kanemura, S. Maeda and S. Ishii, "Superresolution with compound Markov random fields via the variational EM algorithm," Neural Networks, (in press)
- [7] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," IEEE trans. PAMI, **6**(6), 721-741, 1984.
- [8] D. Geman, S. Geman, C. Graffigne and P. Dong, "Boundary detection by constrained optimization," IEEE trans. PAMI, **12**(7), 609-628, 1990.
- [9] 岡田真人, 銅谷賢治, 吉岡利福, 川人光男, "位相を隠れ変数として持つ領域ベース結合 MRF モデル," 信学技報, NC98-184, 1999.
- [10] N. R. Pal and S. K. Pal, "A review of image segmentation techniques," Pattern Recognition, **26**(9), 1277-1994, 1993.
- [11] E. Bratsolis and M Sigelle, "Image relaxation by use of the Potts model with a fast deterministic method," J. Opt. Soc. Am. A, **14**(5), 1033-1043, 1997.