

画像補修アルゴリズムの再定式化とパラメータ推定法

Reformulation of the Digital Image Inpainting Algorithm and How to Estimate its Model Parameters

安田 宗樹*
Muneki Yasuda

1 はじめに

画像補修 (Digital image inpainting) とは画像の欠損部を自動的に再構築する画像処理技術である。様々な手法が提案されているが、ほとんどの場合、画像における欠損部は人為的に手で決め、欠損部の色の情報はその境界の色の情報から推定するという方法で行われる。代表的なアルゴリズムは、欠損部の階調値の勾配が、外からの勾配とスムーズにつながるように欠損部の階調値を決めるという方針で設計されている [1, 2, 3, 4]。これらのアルゴリズムは直感的に理解しやすいものではあるが、やや複雑な手続きを必要とする。またそれらとは別に、テクスチャーに注目して欠損部に張り込む方法も検討されている [5]。一方、Yasuda らはマルコフ確率場に基づいて、全く異なる観点で画像補修の枠組みを提案した [6]。確率モデルによる画像処理については、ベイジアンネットなどの確率推論の理論との統合により多くの有用な結果が報告されている [8]。文献 [6] において提案された確率的な画像補修アルゴリズムには最適なモデルパラメータの値が存在することが指摘された [7]。したがって、その最適なモデルパラメータを探すアルゴリズムを開発することは重要である。

本稿は文献 [6] で導入された確率的画像補修アルゴリズムのモデルパラメータの推定方法について検討する。そのためにまず文献 [6] で導入された確率的画像補修アルゴリズムをより確率の観点から再定式化し、次にそれを用いたモデルパラメータの推定方法について検討する。

2 モデルの定式化とアルゴリズム

2.1 定式化

本節の目的は、欠損した画像を補修するためのモデルとそのアルゴリズムをマルコフ確率場の概念を基礎にして一般的な枠組みとして定式化することである。

画像のピクセルは正方格子的に並んでいるとし、最近接のピクセル対 (つまり、上下左右で隣り合っているピクセル同士) を隣り合った関係とみなす。補修対象としている画像の全てのピクセルそれぞれに異なるラベル i を割り当て、欠損領域に含まれるピクセルの集合を Ω とし、欠損していない領域のピクセルの集合を D とする。また領域 Ω の中で、領域 D に含まれるピクセルを最近接ピクセルに持つピクセルの集合を B_Ω とする。つまり、領域 B_Ω は領域 Ω 上の領域 D との境界線のピクセルの集合である。領域 D の中で、領域 Ω に含まれるピクセルを最近接ピクセルに持つピクセルの集合を B_D とする。領域 B_D は領域 D 上の領域 Ω との境界線のピクセルの集合である。更に集合 U を集合 Ω と集合 B_D との和集合 $U = \Omega \cup B_D$ とし、集合 B_{D2} を集合 D と B_D の差集合 $B_{D2} = D \setminus B_D$ とする。

$f = \{f_i | i \in \Omega\}$ を領域 Ω 上のピクセルの階調値とし、 $d = \{d_i | i \in B_D\}$ を領域 B_D 上のピクセルの階調値とする。また、 $g = \{g_i | i \in D \setminus B_D\}$ を領域 D 上で領域 B_D に含まれないピクセルの階調値とする。今、領域 U に対して次のようなコスト関数を導入する。

$$\begin{aligned} E_U(\mathbf{X} | \mathbf{g}, \boldsymbol{\alpha}) &\equiv \frac{1}{2} \sum_{(ij) \in N(U)} \alpha_{ij} (X_i - X_j)^2 \\ &\quad + \frac{1}{2} \sum_{i \in B_D} \sum_{j \in N_i(B_{D2})} \alpha_{ij} (X_i - g_j)^2 \\ &= \frac{1}{2} \sum_{(ij) \in N(U)} \alpha_{ij} (X_i - X_j)^2 \\ &\quad + \frac{1}{2} \sum_{i \in B_D} \gamma_i (X_i - \phi_i)^2 + \text{const} \quad (1) \end{aligned}$$

ここで、 $\mathbf{X} = \{X_i | i \in U\} = \{f_i | i \in \Omega\} \cup \{d_i | i \in B_D\} = \{\mathbf{f}, \mathbf{d}\}$ は領域 U 内のピクセルの階調値を表しており、 $\boldsymbol{\alpha} = \{\alpha_{ij}\}$ はインデックスに対して対称な正のモデルパラメータである ($\alpha_{ij} = \alpha_{ji}$)。また、 $N(Y)$ はある領域 Y 内で隣り合っているピクセル対の集合であり、 $N_i(Y)$ はある領域 Y 内でピクセル i と隣り合っているピクセルの

* 東北大学大学院情報科学研究科、学振特別研究員 DC2 e-mail: muneki@smapiip.is.tohoku.ac.jp

集合である。 γ_i, ϕ_i はそれぞれ

$$\gamma_i = \sum_{j \in Ni(B_{D2})} \alpha_{ij}, \quad \phi_i = \gamma_i^{-1} \sum_{j \in Ni(B_{D2})} \alpha_{ij} g_j$$

である。コスト関数 (1) を用いて、次を定義する。

$$P_U(\mathbf{X}|\mathbf{g}, \alpha) \equiv \frac{1}{Z_U} \exp \left(-E_U(\mathbf{X}|\mathbf{g}, \alpha) \right) \quad (2)$$

Z_U は規格化定数である。確率変数が $(-\infty, \infty)$ の間の任意の実数値を取る場合、この確率密度関数はガウシアングラフィカルモデルとなる。この確率密度関数は、「隣り合うピクセル同士は互いに近い階調値を取りやすい」という仮定を基礎に定義されており、マルコフ確率場における画像処理の基本理念である。

次に

$$P_{B_D}(\mathbf{d}|\mathbf{g}, \alpha) \equiv \frac{1}{Z_{B_D}} \exp \left(-E_{B_D}(\mathbf{d}|\mathbf{g}, \alpha) \right) \quad (3)$$

$$E_{B_D}(\mathbf{d}|\mathbf{g}, \alpha) \equiv \frac{1}{2} \sum_{(ij) \in N(B_D)} \alpha_{ij} (d_i - d_j)^2 + \frac{1}{2} \sum_{i \in B_D} \gamma_i (d_i - \phi_i)^2 + \text{const} \quad (4)$$

のように、 $\mathbf{g}$ を条件とした $\mathbf{d}$ の確率密度関数を定義すると、条件付き確率の性質より、領域 Ω に対する確率密度関数は

$$P_\Omega(\mathbf{f}|\mathbf{d}, \mathbf{g}, \alpha) = \frac{P_U(\mathbf{X}|\mathbf{g}, \alpha)}{P_{B_D}(\mathbf{d}|\mathbf{g}, \alpha)} = \frac{1}{Z_\Omega} \exp \left(-E_\Omega(\mathbf{f}|\mathbf{d}, \mathbf{g}, \alpha) \right) \quad (5)$$

となる。ただし、

$$E_\Omega(\mathbf{f}|\mathbf{d}, \mathbf{g}, \alpha) = \frac{1}{2} \sum_{(ij) \in N(\Omega)} \alpha_{ij} (f_i - f_j)^2 + \frac{1}{2} \sum_{i \in B_\Omega} \eta_i (f_i - \zeta_i)^2 + \text{const} \quad (6)$$

であり、 η_i, ζ_i はそれぞれ

$$\eta_i = \sum_{j \in Ni(B_D)} \alpha_{ij}, \quad \zeta_i = \eta_i^{-1} \sum_{j \in Ni(B_D)} \alpha_{ij} d_j$$

である。 $P_\Omega(\mathbf{f}|\mathbf{d}, \mathbf{g}, \alpha)$ が欠損領域 Ω に対する確率密度関数である。この定式化においては $\mathbf{f}$ は直接 $\mathbf{g}$ に寄らないので、 $P_\Omega(\mathbf{f}|\mathbf{d}, \mathbf{g}, \alpha) = P_\Omega(\mathbf{f}|\mathbf{d}, \alpha)$ である。

以上の定式化は確率変数の定義域には寄らず成り立つ一般的な枠組みである。一般的に確率密度関数 (5) を計算することは困難であるが、確率変数が $f_i \in (-\infty, \infty)$ の連続値である場合 (5) はガウシアングラフィカルモデル

ル (多次元ガウス分布) となり、容易に扱える。そのとき、確率密度関数 (5) は

$$P_\Omega(\mathbf{f}|\mathbf{d}, \alpha) = \sqrt{\frac{\det \mathbf{C}}{(2\pi)^{|\Omega|}}} \exp \left\{ -\frac{1}{2} (\mathbf{f} - \mathbf{C}^{-1} \boldsymbol{\xi})^t \mathbf{C} (\mathbf{f} - \mathbf{C}^{-1} \boldsymbol{\xi}) \right\} \quad (7)$$

と変形できる。ただし、行列 $\mathbf{C} \in R^{|\Omega| \times |\Omega|}$ とベクトル $\boldsymbol{\xi} \in R^{|\Omega|}$ はそれぞれ以下で定められる。

$$C_{ij} = \begin{cases} -\alpha_{ij} & (ij) \in N(\Omega) \\ 0 & (ij) \notin N(\Omega) \end{cases}, \quad C_{ii} = \sum_{j \in Ni(\Omega)} \alpha_{ij}$$

$$\xi_i = \begin{cases} \eta_i \zeta_i & i \in B_\Omega \\ 0 & i \notin B_\Omega \end{cases}$$

2.2 補修アルゴリズム

$\mathbf{g}$ や $\mathbf{d}$ は (欠損していないので) 既知の観測データであり、それを元に欠損領域のピクセルの階調値 $\mathbf{f}$ を推定する。確率密度関数 (5) を用いて、

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} P_\Omega(\mathbf{f}|\mathbf{d}, \alpha) \quad (8)$$

により得られる $\hat{\mathbf{f}}$ を推定値とする。

確率密度関数 (5) がガウシアングラフィカルモデルのとき、(8) により推定される $\hat{\mathbf{f}}$ は

$$\hat{f}_i = \frac{1}{C_{ii}} \left(\xi_i + \sum_{j \in Ni(\Omega)} \alpha_{ij} \hat{f}_j \right) \quad (9)$$

を反復法により解くことにより得ることができる。このアルゴリズムを用いることにより画像の欠損部を人為的に指定すること以外は自動的に修復を行うことができる。更にその処理は式 (9) を数値的に繰り返し計算するだけという非常にシンプルな形に帰着される。この補修アルゴリズムは文献 [6] のものと細部を除いてほぼ同等である。

ガウシアングラフィカルモデルによる補修アルゴリズム (9) を見てみると、ピクセル i の推定値は (境界上を除けば) 隣接ピクセルの推定値の重み付き平均になっていることが分かる。(もし $\{\alpha_{ij}\}$ が一定ならばピクセル i の推定値は単純に隣接ピクセルの推定値の平均として定まる。)

3 モデルパラメータの推定法

α の値を決めることができれば (9) で表される補修アルゴリズムを用いて画像を補修することができる。本節ではモデルパラメータ α の値を画像 (与えられたデータ) から推定する方法を考える。

最も自然な推定法の一つは、尤度関数

$$l(\mathbf{d}, \mathbf{g}, \alpha) \equiv -\ln \int d\mathbf{f} P_U(\mathbf{X}|\mathbf{g}, \alpha) \quad (10)$$

を最小にする α を推定値にする方法であろう。ここで、 $\int d\mathbf{f} = \prod_{i \in \Omega} \int df_i$ である。尤度関数 (10) を α_{ij} で微分すると、

$$\frac{\partial l(\mathbf{d}, \mathbf{g}, \alpha)}{\partial \alpha_{ij}} = \frac{1}{2} \left\{ \int d\mathbf{f} (V_i - V_j)^2 P_{\Omega}(\mathbf{f}|\mathbf{d}, \mathbf{g}, \alpha) - \int d\mathbf{X} (V_i - V_j)^2 P_U(\mathbf{X}|\mathbf{g}, \alpha) \right\} \quad (11)$$

となる。ただし、 $\int d\mathbf{X} = \prod_{i \in U} \int dX_i$ である。また、 $i \in \Omega$ のとき $V_i = f_i$ 、 $i \in B_D$ のとき $V_i = d_i$ 、 $i \in B_{D2}$ のとき $V_i = g_i$ である。(11) を用いて、パラメータを最急降下法により更新する。

$$\alpha_{ij}(t+1) = \alpha_{ij}(t) - \lambda \frac{\partial l(\mathbf{d}, \mathbf{g}, \alpha)}{\partial \alpha_{ij}} \quad (12)$$

ここで λ は小さな正の定数である。

しかし、この方針では一回の更新毎に統計量を計算する必要性がでてくるため、欠損領域が比較的大きい場合、この計算は多くの計算時間とメモリを必要としてしまう。(ガウシアングラフィカルモデルを仮定したとしても一回の計算毎に $R^{|U| \times |U|}$ のサイズの行列を扱うことになる。)

4 数値実験

図4の(a)が白地の文字により一部が欠損した256階調画像である。全体のおおよそ13%程度が欠損している。(b)は本稿で提案したアルゴリズムによりパラメータ推定し、補修した画像である。パラメータ推定の際の更新式(12)の計算には詳細は割愛するが、平均場近似の拡張であるTAP近似と呼ばれる近似的手法を用いている[10]。 $\lambda = 10^{-6}$ とし、20回程度更新式(12)を用いた結果である。計算時間は通常のPC(CPU 1GHz程度)で数秒程度である。

f_i^* を原画像のピクセル*i*の階調値、 $\hat{f}_i$ をピクセル*i*の階調値の推定値として、原画像との平均2乗誤差(MSE)

$$\text{MSE} = \frac{1}{|\Omega|} \sum_{i \in \Omega} (\hat{f}_i - f_i^*)^2 \quad (13)$$

を計算してみると、図4の(b)でおよそ438.96であった。同様の欠損画像に対して文献[6, 7]の方法で補修したところ、最適なパラメータでMSEはおおよそ439.95であった。最適なパラメータはパラメータを振り、逐次計算して最適ところを見つけた。

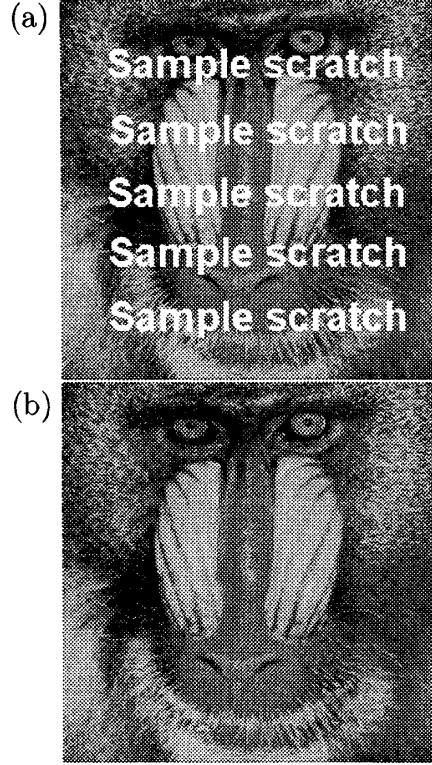


図1 提案アルゴリズムによる画像補修

(a) 欠損した画像。256階調でサイズは256×256である。

(b) 本提案アルゴリズムによりパラメータ推定し、補修された画像。

5 まとめ

本稿では文献[6]で導入されたマルコフ確率場の概念を用いた画像補修アルゴリズムをより一般的な枠組みとして再定式化し、それに対するモデルパラメータの推定方法を検討した。数値実験により、本稿で提案した手法は最適なパラメータを選んだ場合の文献[6]の手法の結果とほぼ同等の性能を与えていることを確認できた。

しかし、まだ本アルゴリズムのより一般的な性能や限界などまだ議論の余地はある。本アルゴリズムの数理的な構造やその処理の挙動などを今後詳しく調べ、パラメータの推定アルゴリズムやコスト関数の設計等を改良していきたいと考えている。

参考文献

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image Inpainting," in *Proc. SIGGRAPH 2000, Computer Graphics Proceedings*,

- Annual Conference Series*, edited by Kurt Akeley, pp 417–424, Reading, MA:Addison-Wesley, 2000.
- [2] M. Bertalmio, A. L. Bertozzi, and G. Sapiro, “Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting,” in *Proc. ICCV 2001*, pp. 1335 – 1362, IEEE CS Press 1. [CITY]:[PUB], 2001.
 - [3] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, “Filling-in by joint interpolation of vector fields and gray levels,” *IEEE Trans. on Image Processing*, vol. 10(8), pp. 1200 – 1211, August 2001.
 - [4] T. F. Chan and J. Shen, “Mathematical models of local non-texture inpaintings,” *SIAM J. Appl. Math.*, vol. 62(3), pp. 1019 – 1043, 2002.
 - [5] A. Criminisi, P. Perez and K. Toyama, “Region Filling and Object Removal by Exemplar-Based Image Inpainting”, *IEEE Trans. on Image Processing*, vol. 13(9), 2004.
 - [6] M. Yasuda, J. Ohkubo and K. Tanaka, “Digital image inpainting based on Markov random field,” Proceedings of International Conference on Computational Intelligence for Modelling, Control and Automation (November 28-30, 2005, Vienna, Austria), Vol.II, pp.747-752, IEEE Computer Society Press, March 2006.
 - [7] 安田宗樹, 大久保潤, 田中和之, “ガウシアングラフィカルモデルを用いた画像補修アルゴリズム”, FIT2006, LI-018, pp 225-228, 2006.
 - [8] K. Tanaka, “Statistical-mechanical approach to image processing,” *J. Phys. A: Math. Gen.*, vol 35, pp. R81 – R150, 2002.
 - [9] 田中和之, 確率モデルによる画像処理技術入門, 森北出版.
 - [10] K. Tanaka, H. Shouno, M. Okada and D. M. Titterton, “Accuracy of the Bethe Approximation for Hyperparameter Estimation in Probabilistic Image Processing,” *J. Phys. A: Math. Gen.*, vol 37, No. 36. pp.8675-8696, 2004.