

統計力学的手法によるデジタル画像のノイズ除去

宇治野・高橋研究室 武藤 瑞生

1 はじめに

統計力学とは、粒子などの微視的な性質をもとに、それらの集団が全体としてどのような性質を示すかを研究する学問である [1]。

情報の分野においても、画像やニューラルネットワークといった、同じ性質をもつ要素の集合で構成されるものに対して統計力学を適用することができる。本研究では、デジタル画像のノイズ除去処理を統計力学的手法で行い、その有用性を考察する。

2 原理

2.1 2値画像のノイズ除去

はじめに、問題の簡略化のため扱う画像を2値画像に限定し、イジングモデルを用いてノイズ除去を行う [2] [3]。

イジングモデルは、格子状に配置された磁気モーメント μ を持つ N 個のスピンについて、スピンの向きをある軸に対して平行 (1) か反平行 (-1) かの2つに限定したものである。 N 個のスピンについて、 i 番目のスピンの向きを σ_i とする。各スピンは、1 または -1 のどちらかの値をとる。各スピンの値を画素値に対応させることで、このモデルを2値画像として考えることができる。

2値画像においては、 N 個の画素について、 i 番目の画素のもつ値（画素値）を σ_i とする。各画素値は、1 または -1 のどちらかの値をとる。これらの画素を N 個並べたものが2値画像となる。なお、今回は1を白、-1を黒とおくことにする。

イジングモデルにおいて、系のハミルトニアン（全体のエネルギー）は以下のように表される。

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - \mu H \sum_i \sigma_i \quad (1)$$

ここで、 J は正の定数、 H は外部磁場、 $\sum_{\langle ij \rangle}$ は隣り合うすべての画素の組についての和を表している。

このとき、カノニカル分布は以下のように表される。

$$P(\{\sigma_i\}) = \frac{e^{-\beta \mathcal{H}}}{\sum_{\sigma} e^{-\beta \mathcal{H}}} \quad (2)$$

β はボルツマン定数 k_B と温度 T の逆数 ($\beta = \frac{1}{k_B T}$) を示す。

$P(\{\sigma_i\})$ は、 $\sigma_1 \sim \sigma_N$ の画素値の組 $\{\sigma_i\}$ を持つ画像の存在する確率を示している。

式 (1) のハミルトニアンについて言及すると、第一項は隣り合うスピンどうしの向きが同じであれば小さく、異なれば大きくなることが分かる。また、第二項は各スピンの向きが外部磁場 H と同じであれば小さく、異なれば大きくなることが分かる。

これを2値画像のノイズ除去に適用すると、カノニカル分布に対して、第一項は「隣り合う画素どうしの相関を考慮するための項」となる。また、磁場の向きを受信画像の各画素値に置き換えることで、第二項は「通信の過程で乗るノイズ（画素値の反転）を考慮するための項」となる。

式 (1) における J や μ といったパラメータを適切に設定することで、 $P(\{\sigma_i\})$ を最大化する画素値の組 $\{\sigma_i\}$ が求まるべき画像となる。

式 (2) は、計算量が膨大になるため解くことはほぼ不可能である。

そこで、 $P(\{\sigma_i\})$ を計算する代わりに、平均場近似 [4] の考え方に沿って、隣接した画素にのみ着目して計算を行う。

平均場近似の元で、1 つのスピンに対するハミルトニアン \mathcal{H}_i は、

$$\mathcal{H}_i = -J\hat{z}\langle\sigma\rangle\sigma_i - \mu H\sigma_i \quad (3)$$

と書ける。なお、 \hat{z} は最近接格子点の数（ここでは二次元なので 4）であり、 $\langle\sigma\rangle$ はスピンの熱的平均値を示す。

更に、系のハミルトニアン \mathcal{H} を、 \mathcal{H}_i の和で、

$$\mathcal{H} \simeq \sum_i \mathcal{H}_i \quad (4)$$

と書けると仮定する。これにより、ハミルトニアンが独立な粒子からなる項の和で表されるので、状態和 Z は

$$\begin{aligned} Z &= \left(\sum_{\sigma_i=\pm 1} e^{-\beta\mathcal{H}_i} \right)^N \\ &= \left(\sum_{\sigma_i=\pm 1} e^{\beta(J\hat{z}\langle\sigma\rangle + \mu H)\sigma_i} \right)^N \\ &= (2 \cosh \beta(J\hat{z}\langle\sigma\rangle + \mu H))^N \end{aligned} \quad (5)$$

と表される。ヘルムホルツの自由エネルギー F は、

$$\begin{aligned} F &= -k_B T \log Z \\ &= -N k_B T \log (2 \cosh \beta(J\hat{z}\langle\sigma\rangle + \mu H)) \end{aligned} \quad (6)$$

と表される。 μ を単位とする磁化 M は、

$$\begin{aligned} M &= \left\langle \sum_i \sigma_i \right\rangle = \frac{\sum_{\sigma_i=\pm 1} \left\{ \left(\sum_i \sigma_i \right) e^{-\beta\mathcal{H}} \right\}}{\sum_{\sigma_i=\pm 1} e^{-\beta\mathcal{H}}} \\ &= -\frac{1}{\mu} \left(\frac{\partial F}{\partial H} \right)_T \\ &= N \tanh \beta(J\hat{z}\langle\sigma\rangle + \mu H) \end{aligned} \quad (7)$$

と表される。 $\frac{\langle \sum_i \sigma_i \rangle}{N} = \langle \sigma \rangle = x$ とおくと、以下の式が導かれる。

と書けると仮定する。これをもとに計算を進めると、以下の式が導かれる。

$$x = \tanh \beta(J\hat{z}x + \mu H) \quad (8)$$

式 (8) を自己整合方程式という。

式 (2) の折で述べたように、 $\beta = \frac{1}{k_B T}$ である。 $\frac{J}{k_B} = L$ 、 $\frac{\mu H}{k_B} = K$ とおいて、式 (8) をもとに各画素についての式を考える。ノイズの乗った画像の i 行 j 列目の画素値を $g_{i,j}$ とすると、数値計算において n ステップ目の修復の過程にある画像の画素値 $a_{i,j}^n$ を求める式は以下のようになる。

$$a_{i,j}^{n+1} = \tanh \left(\frac{L}{T} (a_{i+1,j}^n + a_{i,j+1}^n + a_{i-1,j}^n + a_{i,j-1}^n) + \frac{K}{T} g_{i,j} \right) \quad (9)$$

なお、 $a_{i,j}^0 = g_{i,j}$ とする。

2.2 グレースケール画像のノイズ除去

扱う画像をグレースケール画像とする。また、0 を黒、1 を白とし、各画素値は 0 から 1 までの実数値を取るとする。

グレースケールでは画素値が 2 値を超えるため、イジングモデルでの扱いや、2 値画像のような自己整合方程式を求めることが困難となる。よって、本研究ではグレースケール画像の修復の手法としてメトロポリス法を用いることにする。

計算量が膨大になる問題を、乱数を用いて統計的に解く手法にモンテカルロ法がある。特に、ある状態となる確率がその 1 ステップ前の状態のみで決まる、といった特徴を持つものをマルコフ連鎖モンテカルロ法 (MCMC 法) という。

メトロポリス法 [5] は、MCMC 法の 1 種である。これは、ある状態から遷移する別の状態を提案したとき、その状態への遷移を採用するかどうかを確率的に決定する手法である。

グレースケール画像の修復に当てはめた場合の考え方は、以下の通りである。

1. 今の状態の画像について、各画素のハミルトニアンを求め、式 (4) のように全体のハミルトニアン \mathcal{H} を求める。
2. 画素値をランダムに提案し、 i 番目の画素がその画素値だった場合の各画素のハミルトニアンを求め、全体のハミルトニアン \mathcal{H}' を求める。
3. 確率 $\min \left(1, e^{-\beta(\mathcal{H}' - \mathcal{H})} \right)$ で i 番目の画素の画素値を提案された画素値に変更する。
 β はボルツマン定数 k_B と温度 T の逆数 ($\beta = \frac{1}{k_B T}$) を示す。

上記の操作をすべての画素に対して複数回行うことで、 $P(\{\sigma_i\})$ が大きくなる画素値の組 $\{\sigma_i\}$ を近似的に求めることができる。このような処理を行うプログラムを実装し、グレースケール画像の修復を行う。

グレースケール画像について、平均場近似での各画素におけるハミルトニアン \mathcal{H}_i を以下のように定義する。

$$\mathcal{H}_i = -J \sum_j \text{com}_n(a_i, a_j) - K \text{com}_n(a_i, g_i) \quad (10)$$

ここで、 a_i は修復の候補となる画素値、 g_i は修復前の画素値を示す。また、 J, K は各項の重みを決定する係数であり、今回は推定した値を用いる。 j は i の示す画素に隣接する画素を示す。

式 10 における関数 $\text{com}_n()$ は、引数にある 2 つの画素値のエネルギー差を示す関数である。本研究で試行した関数は以下の 3 つである。

$$\text{com}_1(x, y) = 1 - |x - y| \quad (11)$$

$$\text{com}_2(x, y) = \exp\left(-\frac{\log 0.5}{0.2^2}(x - y)^2\right) \quad (12)$$

$$\text{com}_3(x, y) = \cos(\pi|x - y|) \quad (13)$$

2 値でのイジングモデルでは、2つのスピンの向きが同じか異なるかの2択でエネルギー差を決定していた。グレースケールでも同様に、2つの階調値が等しいか等しくないかで離散的にエネルギー差を決定することもできる。しかしこの方法では、階調数が増えるほど階調値の等しい画素の組は減り、モデルがうまく機能しなくなると考えられる。そこで採用したのが、上記の3種類の関数である。

$\text{com}_1()$ では、単純に画素値の差に比例してエネルギー差を決定している。

$\text{com}_2()$ では、ガウス分布を用いて、ステップ関数に近い形でエネルギー差を決定している。係数については、画素値の差が0.2のとき $\text{com}_2()$ の値が0.5となるように設定している。この係数のより良い設定方法については議論の余地がある。

$\text{com}_3()$ では磁性体のモデルのように各画素を0から π の範囲を向くスピンと考え、画素値を角度として2つのスピンの内積を取っている。

これらの関数は、どれも画素同士の画素値の差が小さいほど値が大きく、差が大きいほど値が小さくなる作用を持つ。よって、式(10)では、第一項によって、修復の候補となる画素値とその画素に隣接する画素値との差が小さいほどハミルトニアンは小さくなる。また、第二項によって、修復の候補となる画素値とその位置の元の画素値との差が小さいほどハミルトニアンは小さくなる。

カラー画像のノイズ除去 扱う画像をカラー画像とし、各画素値の r,g,b 要素がそれぞれ0から1までの実数値を取るとする。

グレースケール画像と同様にメトロポリス法を応用して、カラー画像のノイズ除去を行う。

カラー画像について、平均場近似での各画素におけるハミルトニアン \mathcal{H}_i を以下のように定義する。

$$\begin{aligned} \mathcal{H}_i = & -J \sum_j \text{com}_4(a_{r_i}, a_{g_i}, a_{b_i}, a_{r_j}, a_{g_j}, a_{b_j}) \\ & -K \text{com}_4(a_{r_i}, a_{g_i}, a_{b_i}, q_{r_i}, q_{g_i}, q_{b_i}) \end{aligned} \quad (14)$$

ここで、 $a_{r_i}, a_{g_i}, a_{b_i}$ は修復の候補となる画素値の rgb 要素、 $q_{r_i}, q_{g_i}, q_{b_i}$ は修復前の画素値の rgb 要素を示す。また、J,K は各項の重みを決定する係数であり、今回は推定した値を用いる。 j は i の示す画素に隣接する画素を示す。

式(14)における関数 $\text{com}_4()$ は以下に示すとおりである。

$$\begin{aligned} \text{com}_4(x_r, x_g, x_b, y_r, y_g, y_b) \\ = \frac{1}{3} \sqrt{2(x_r - y_r)^2 + 4(x_g - y_g)^2 + 3(x_b - y_b)^2} \end{aligned} \quad (15)$$

式(15)は、色の rgb 要素を基底としたときの単純な距離に対して、視覚への影響をもとに重みを乗じたものとなっている。2色の差を定義した式はいくつか考案されているが、本研究ではこの式を用いる。

なお、式(15)右辺のルート部分の最大値が3なので、 $\text{com}_4()$ の最大値を1とするために右辺に $\frac{1}{3}$ を乗じている。

3 実装

3.1 2値画像のノイズ除去

式 (9) について、 $a_{i,j}^{n+1}$ と $a_{i,j}^n$ の差が無視できるほど小さくなるまで数値計算を行う。その結果として得られる画素値の組 $\{a_i\}$ が、修復画像となる。

上記の計算をプログラムとして実装し、これを実行した結果を図 1 に示す。



図 1: 2 値画像のノイズ除去

図 1(a) は元画像、(b) は 10% の確率で画素を反転させた（ノイズを乗せた）画像、(c) は (b) の画像のみをもとにノイズ除去を行った画像である。

各画像を比較すると、ノイズ除去の効果が表れていることが確認できる。

3.2 グレースケール画像のノイズ除去

式 (10) の各関数のハミルトニアンを用いて、メトロポリス法でノイズ除去処理を行った結果を図 2 に示す。

なお、本研究では計算量の短縮のため 256 階調である画素値を 8 階調に下げる処理を行ったうえでノイズ除去処理を行っている。



$$(\beta = 8.0, J = 1.0, K = 2.0)$$

図 2: グレースケール画像のノイズ除去

図 2 より、どの関数のハミルトニアンを用いた処理でもある程度ノイズが除去できていることが確認できる。

さらに、どの程度元画像に近づいているかを数値化して確かめる。

元画像の画素値と修復画像の画素値の、1 画素当たりの平均の差を画素値差と呼ぶことにする。画素値差が 0 ならば、2 つの画像は全く同じということになる。

各関数のハミルトニアンを用いてノイズ除去処理を行い、得られた修復画像と元画像の画素値差を求めた。この処理を 5 回ずつ行った平均の画素値差を表 1 に示す。

表 1: 各関数のハミルトニアンごとの画素値差（グレースケール画像）

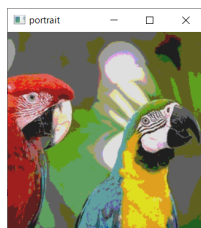
関数	画素値差
ノイズ除去処理なし	0.034899139
$(\beta = 8.0, J = 1.0, K = 2.0)$ com ₁ ()	0.027587509
com ₂ ()	0.029083633
com ₃ ()	0.033935547

表 1 から分かるように、どの関数を用いた場合でも、ノイズ除去処理を行った場合のほうがノイズの乗った画像より元画像に近づいていることが確認できる。また、関数 com₁() のハミルトニアンを用いたときのノイズ除去処理が最も精度が良いことが分かる。

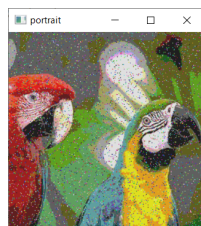
3.3 カラー画像のノイズ除去

グレースケール画像と同様に、式 (14) のハミルトニアンを用いてメトロポリス法でノイズ除去処理を行った結果を図 3 に示す。

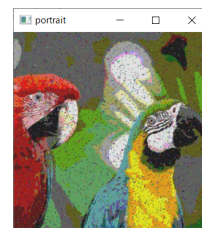
なお、本研究では計算量の短縮のため 256 階調である画素値の rgb 要素をそれぞれ 4 階調に下げる処理を行ったうえでノイズ除去処理を行っている。



(a) 元画像



(b) ノイズの乗った画像



(c) 修復画像

$(\beta = 8.0, J = 1.0, K = 2.0)$

図 3: カラー画像のノイズ除去

図 3 より、明るい画素値のノイズはある程度除去できていることが確認できる。一方で、暗い画素値のノイズは除去できていないことが分かる。

3.4 他の手法との比較

8 階調のグレースケール画像において、メトロポリス法による手法とその他の手法によるノイズ除去処理の精度を比較した。

比較対象の手法として、以下の 2 つを行った。

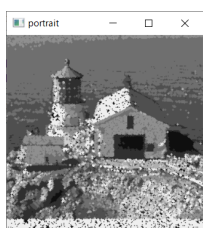
1. 膨張処理の応用

2 値画像において、白の画素の 4 近傍の画素をすべて白で置き換える処理を膨張処理という。その後、黒の画素の 4 近傍の画素をすべて黒で置き換えることで、孤立した黒の画素を取り除くことが出来る。同様に、ある階調値の画素の、4 近傍の画素の画素値をすべてその階調値に置き換えるという処理を、小さい階調値から順にすべての階調値で行う。

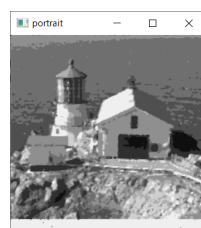
2. メディアンフィルタ

ある画素の画素値を、その画素と 8 近傍の画素の画素値のうちの中央値で置き換えるという処理をすべての画素に対して行う。

これらの手法で 8 階調のグレースケール画像のノイズ除去処理を行った結果を図 4 に示す。



(a) 修復画像：膨張処理の応用



(b) 修復画像：メディアンフィルタ

図 4: 各手法によるグレースケール画像のノイズ除去

図 4 より、膨張処理の応用では、画像がノイズの乗った画像よりも乱れていることが分かる。また、メディアンフィルタでは、画像がのっぺりした印象を受けるが、ノイズは除去できていることが分かる。これらについても画素値差を求め、どの程度元画像に近づいているかを評価する。

各手法を用いてノイズ除去処理を行い、得られた修復画像と元画像の画素値差を求めた。この処理を 5 回ずつ行った平均の画素値差を表 2 に示す。

なお、メトロポリス法ではもっとも精度の良かった関数 $\text{com}_1()$ のハミルトニアンを用いた。ノイズ除去処理をしない場合及びメトロポリス法のデータは表 1 と同じものを用いている。

表 2: 各手法での画素値差（グレースケール画像）

手法	画素値差
ノイズ除去処理なし	0.034899139
メトロポリス法（関数 $\text{com}_1()$ ）	0.027587509
膨張処理の応用	0.05366478
メディアンフィルタ	0.032133865

表 2 から、膨張処理はノイズの乗った画像よりも元画像から遠ざかっていることが分かる。また、メディアンフィルタは、ノイズの乗った画像よりも元画像に近づいていることが分かる。

表 1 および 2 より、本研究で試行した手法の中では、関数 $\text{com}_1()$ のハミルトニアンを用いたメトロポリス法によるノイズ除去処理が最も良い精度でノイズ除去処理を行ったことが分かる。

一方で、メディアンフィルタは、グレースケール画像においては階調数に関わらず画素数に比例した時間で処理が行えるので、本研究で行ったメトロポリス法よりも計算が速いという利点がある。

4 まとめと今後の展望

本研究では、イジングモデルについて学習し、これをもとに 2 値画像のノイズ除去処理を行うプログラムを作成した。また、プログラムを実行し、実際にノイズ除去が行えることを確認できた。

加えて、メトロポリス法を用いてグレースケール画像およびカラー画像のノイズ除去処理を行うプログラムを作成した。グレースケール画像のノイズ除去についてはいくつかの手法を比較し、統計力学的手法の優位性を確かめることができた。

本研究ではイジングモデルと MCMC 法的一种であるメトロポリス法を用いて画像修復を行ったが、他にもノイズ除去処理に活用できる統計力学的手法は存在する。例えば、条件付きボルツマンマシンというモデルを用いて、類似の画像から画素値の分布を学習し、それをもとに対象の画像を修復するという手法も考えられる。今後はこういった手法についても学習し、研究していきたい。

また、カラー画像については試行の余地が多分にあるため、さらに研究を進めていきたい。

例えば、今回はカラー画像のハミルトニアンには式 (14) のみを用いたが、色の差を数値化する方法にはいくつか種類がある。他の方法を用いてハミルトニアンを設定した際のノイズ処理の精度の変化について確かめたい。また、カラー画像を rgb 空間から hsv 空間に変換した際のノイズ除去の精度の変化についても確かめたい。

参考文献

- [1] 北原和夫・杉山忠男『統計力学』,(講談社,2010)
- [2] 田中和之『統計力学的手法をもとにした画像修復』日本物理学会誌 54 巻 1 号 p.25-33,(1999)
- [3] 稲葉文香『統計力学的手法による 2 値画像の修復』お茶の水女子大学卒業論文予稿 p.15-16,(2015)
- [4] 岡部豊『統計力学』,(裳華房,2007)
- [5] 花田政範・松浦壮『ゼロからできる MCMC』,(講談社,2022)