

データのスパース性を利用した画像のノイズ除去とその計算量の削減

2014SC027 加藤慎也

指導教員：大石泰章

1 はじめに

画像には、センサ性能の限界や通信時のエラーなどによるノイズが含まれている。ノイズを除去することはさまざまな分野での利用に有用であるため、研究が進められている。その一環として、画像データのスパース性を仮定して多数のデータから必要な情報のみを抽出するスパースモデリングという方法が良いノイズ除去結果を与えるとして注目を集めている。しかしながら、スパースモデリングを用いた方法は往々にして計算量が大きく、高速化手法の提案は今もなお続いている [1]。

本研究では、まずスパースモデリングを用いた静止画像のノイズ除去を行い、その応用として動画のノイズ除去に挑戦する。動画のノイズ除去では隣接フレームの類似性を用いて計算量を小さくする工夫を行った。また、結果の良し悪しを客観的に判断するために、PSNR (peak-signal to noise ratio) 値という画像評価指標を用いる。

2 スパースモデリング

2.1 静止画像の取り扱い

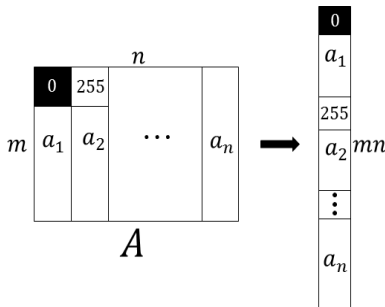


図1 静止画像の行列表現と列ベクトル表現

以下 5.1 節まで静止画像とそのノイズ除去について考え、動画のノイズ除去については 5.2 節で扱う。

図1のように、縦 m 個、横 n 個、全体で mn 個の画素からなる静止画像は、各画素の輝度を対応する位置の要素に持つ m 行 n 列の行列 A で表すことができる。ここでは輝度は 0 から 255 までの整数であり、0 が黒を、255 が白を表すとする。図1では、画像の左上の画素が黒、その右隣の画素が白なので、 $A(1,1) = 0$ 、 $A(1,2) = 255$ である。

画像行列 A を各列に分けて $(a_1 \ a_2 \ \dots \ a_n)$ と書くとき、すべての列を縦に積み重ねた mn 次元の列ベクトル $(a_1^T \ a_2^T \ \dots \ a_n^T)^T$ を考えれば、画像を列ベクトルで表現できる (図1)。

2.2 スパースモデリングを用いたノイズ除去方法

静止画像のノイズ除去問題をスパースモデリングを利用して解くために、問題を定式化する。ノイズの含まれていない原画像を列ベクトル表現して y とし、平均 0 のホワイト

トガウスノイズ n が一様に加わったものを観測画像 \tilde{y} とする。すなわち

$$\tilde{y} = y + n \quad (1)$$

である。この観測画像 \tilde{y} に基づいて原画像 y に近い画像を得ることが目標である。

画像のノイズ除去問題をスパースモデリングを利用して解くには、以下の最適化問題を解けばよい：

$$\underset{x}{\text{minimize}} \ \|x\|_0 \quad \text{subject to} \ \|\tilde{y} - Dx\|_2^2 \leq \epsilon. \quad (2)$$

ここで、 \tilde{y} はノイズの加えられた観測画像の列ベクトル表現で、次元は M であるとする。また、 x は N 次元の未知ベクトルであり、 N は M よりも大きいものとする。行列 D は辞書行列と呼ばれる所与の $M \times N$ 行列であり、 $D = (d_1 \ d_2 \ \dots \ d_N)$ の各列ベクトル d_i は画像を線形結合で表現するための基底画像である。しきい値 ϵ は所与の正の数であり、通常ノイズが大きいときは大きく選ぶ。また、0 ノルム $\|\cdot\|_0$ はベクトルの非零要素の個数であり、2 ノルム $\|\cdot\|_2$ はユークリッドノルムである。この最適化問題の解として得られた \hat{x} から、ノイズ除去の結果は以下によって得ることができる：

$$\hat{y} = D\hat{x}. \quad (3)$$

ここで \hat{x} の要素のほとんどは零であり、 \hat{y} は辞書行列 D の列をスパースに利用したベクトルであると言えることから、この方法をスパースモデリングと言う。ただし、0 ノルム最適化問題 (2) を直接解くことは困難である。

3 0 ノルム最適化問題の近似解法

0 ノルム最適化問題 (2) を近似的に解くために、貪欲法を用いる。文献 [1] を参考にして、ここでは、貪欲法の中でも、良好な近似解を与えることが知られている OMP (orthogonal matching pursuit) 法を用いる。この手法の手続きは以下のとおりである。まず、残差ベクトルの初期値を $r^0 = \tilde{y}$ とし、残差 r^0 に 2 ノルムの意味で最も近い行列 D の列 d_i を選択し、選択された列の番号 i によって集合 S^0 を $S^0 = \{i\}$ のように定める。暫定解 x^0 を

$$\underset{x}{\text{minimize}} \ \|Dx - r^0\|_2^2 \quad \text{subject to} \ x_i = 0 \ (i \notin S^0) \quad (4)$$

の最適解として定め、暫定解 x^0 を用いて残差を $r^1 = r^0 - Dx^0$ のように更新する。残差 r^1 に最も近い行列 D の列 d_i を選び、列の番号 i をインデックス集合に追加して $S^1 = S^0 \cup \{i\}$ のように更新する。以下同様に繰り返し、終了条件 $\|r^k\|_2 \leq \epsilon$ が満たされたら終了する。終了したときの暫定解 x^k は 0 ノルム最適化問題 (2) の近似解と見なせる。

4 辞書学習

スパースモデリングを用いた画像のノイズ除去を行うために、辞書行列を準備する必要がある。文献 [1] を参考に

手順を記す。辞書を学習によって準備するには、訓練データベースとして ℓ 個の M 次元ベクトル $\{\mathbf{t}_i\}_{i=1}^{\ell}$ を与える。訓練データベースを行列 $\mathbf{T} = (\mathbf{t}_1 \mathbf{t}_2 \cdots \mathbf{t}_\ell)$ で表す。初期辞書として列を正規化したランダム行列を \mathbf{D}^0 として用いる。または、2次元離散コサイン変換行列などの列を正規化したものを \mathbf{D}^0 としてもよい。与えられた訓練データベースを用いて $i = 1, 2, \dots, \ell$ のそれぞれに対して以下の最適化問題

$$\underset{\mathbf{x}}{\text{minimize}} \|\mathbf{t}_i - \mathbf{D}^0 \mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_0 \leq k_0 \quad (5)$$

を解く。例えば上記の OMP 法を使った近似解でもよい。ここで、 k_0 は適当な正の整数である。得られた解を $\hat{\mathbf{x}}_i$ とし、これによって $N \times \ell$ 行列 $\mathbf{X}^0 = (\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2 \cdots \hat{\mathbf{x}}_\ell)$ を定める。次の最適化問題を考える：

$$\underset{\mathbf{D}}{\text{minimize}} \|\mathbf{T} - \mathbf{D}\mathbf{X}^0\|_F^2. \quad (6)$$

ここで、 $\|\cdot\|_F$ はフロベニウスノルムである。 \mathbf{X}^0 の一般化逆行列を \mathbf{X}^{0+} とするとき、最適解は $\mathbf{T}\mathbf{X}^{0+}$ であるが、これを暫定の辞書 \mathbf{D}^1 とする。以下同様に、辞書 \mathbf{D} を固定して \mathbf{X} を求め、 \mathbf{X} を固定して辞書 \mathbf{D} を求めるという作業を交互に繰り返し、収束した辞書 \mathbf{D} を結果とする。これが MOD (method of optimal direction) という辞書学習アルゴリズムである。

5 ノイズ除去と動画像への応用

5.1 ノイズ除去実験

本章では、2.2 節の問題を平均 0 で標準偏差 $\sigma = 30$ のホワイトガウスノイズを印加した 256×256 画素の静止画像に対して考え、実験を行う。実験には CPU i5-3340m, RAM 4GB の PC を利用した。画像の各画素を中心とした 8×8 画素の部分画像を作り、それぞれに対してスパースモデリングを用いたノイズ除去を行う。各ノイズを除去した部分画像の重複を単純平均により統合することにより、画像全体のノイズ除去を実現する。辞書行列を作るために、初期辞書を 2次元離散コサイン変換行列とし、訓練データベースには観測画像の各画素を中心とした 8×8 画素の部分画像のうち $1/5$ をランダムに選んで用い、 $k_0 = 4$ として MOD を 15 回反復させる。得られた辞書の基底画像を図 2 に示す。各部分画像におけるノイズ除去はしきい値を $\epsilon = 1$ として行った。図 3 左にノイズ除去前の画像を、図 3 中にノイズ除去後の画像を示す。ノイズ除去が成功していることがわかる。

5.2 動画像への応用

動画像は複数の静止画像（これをフレームという）を時間の順に並べたものである。したがって、動画像のノイズ除去をするには、上記の静止画像に対するノイズ除去の操作を各フレームに対して行えばよい [1]。しかしながら、静止画像のノイズ除去に比べ、さらに計算量が大きくなってしまいますので、動画像の隣接フレームの類似性を用いて計算量を小さくする方法を考える。すなわち、2 フレーム以降の各部分画像において前フレームの 0 ノルム最適化問題 (2) の近似解が OMP 法の終了条件を満たしていればその

ままその近似解を利用することで計算量を削減する。実験には 16 フレームからなる動画像を利用した。図 3 右に、ノイズ除去後の 16 フレーム目の画像を示す。16 フレーム目においてもノイズ除去に成功している。また、この方法による計算量削減の結果、計算時間は 34 秒程度短縮することができた。

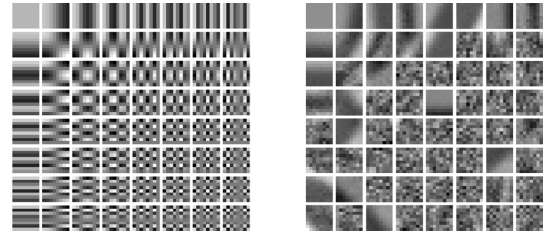


図 2 左：初期辞書の基底画像；右：学習後の基底画像



図 3 左：観測画像（1 フレーム目）；中：ノイズ除去後の画像（1 フレーム目）；右：ノイズ除去後の画像（16 フレーム目）

5.3 画像評価

画像の評価には PSNR 値がよく利用される。PSNR 値は以下の式により算出する：

$$\text{PSNR} = 10 \log_{10} \frac{\text{MAX}^2}{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \|\mathbf{A}(i, j) - \mathbf{A}'(i, j)\|^2}. \quad (7)$$

ここで MAX は対象画像の最大の輝度で \mathbf{A} が原画像、 \mathbf{A}' が劣化画像を示す。PSNR 値が高い方が原画像に近いことになる。

図 3 左での PSNR 値は 19.14[dB]、図 3 中では 31.84[dB] であるので、ノイズ除去により 12.7[dB] の向上が見られた。また、動画像の 16 フレーム目のノイズ除去（図 3 右）における PSNR 値の向上は 12.89[dB] であり、動画像に対してもノイズ除去ができていたことが客観的な数値でも確認された。

6 おわりに

本研究では、観測画像を訓練データとする学習で得た辞書を用いて、スパースモデリングによる画像のノイズ除去を行った。また、動画像への応用を行い、計算時間を短縮する工夫を行った。今後は、より良い結果を与える訓練データを得る方法を考える必要があると思われる。

参考文献

- [1] M. Elad (玉木 徹 訳)：『スパースモデリング： ℓ_1/ℓ_0 ノルム最小化の基礎理論と画像処理への応用』。共立出版、東京、2016。