

画像向きスパースコーディングの最適化基底に関する一考査 A Consideration on Sparse Coding Bases Optimum for Image

李 炎坤[†] 亀山 渉[‡] 関口 俊一^{†*}
Yankun Li Wataru Kameyama Shunichi Sekiguchi

1. はじめに

スパースコーディングの基底を求める一般的な方法では、原信号の再現性とスパース係数の L_0 ノルムの制約を利用する。これらの制約によって、原信号を画像とした場合、求められた基底は画像特徴の一端を表していると考えられるが、原信号が画像であることは一切考慮されていない。

文献[1]では、原信号の再現性に L_2 ノルムではなく、SSIM を適用する検討を行っているが、画像特徴そのものを制約項としている訳ではない。

そのため、本研究では、原信号を画像と限定した新たな制約項を導入することによって、画像向きの最適化基底を求める方法について検討した。新たな制約項として、画像の一般的な特徴である近傍画素類似性とエッジ特徴を追加し、スパースコーディングの基底がどのように変化するかについて検討を行った結果を報告する。

2. 研究手法

スパースコーディングの実装には K-SVD[2]を用いた。K-SVD の基底を求める計算式は、式(1)に示すように、原信号の再現性とスパース係数の L_0 ノルムの制約だけを考慮する。

$$\text{ksvd}(D, X) = \min_{D, X} \|Y - DX\|_2^2 \text{ s.t. } \forall i, \|x_i\|_0 \leq T \quad (1)$$

ここで、 D は辞書、 X はスパース係数、 Y は原信号、 T は非 0 の値を持つスパース係数の最大数である。また、画像から切り出したあるサイズの画像パッチ y_i を列要素としてならべたものが Y であり、 y_i に対応するスパース係数は x_i である。

2.1 近傍画素類似性制約の追加

近傍画素類似性は、画像復元等の処理で利用されているエネルギー関数[3]の平滑項を基に、 $2[\text{pixel}] \times 2[\text{pixel}]$ の計算窓 w を基本単位として計算する。計算窓 w 中の画素を P_i として、その平滑値 $\text{smooth}(w)$ を式(2)で計算する。

$$\text{smooth}(w) = |p_1 - p_2| + |p_1 - p_3| + |p_1 - p_4| + |p_2 - p_3| \\ + |p_2 - p_4| + |p_3 - p_4| \text{ s.t. } \forall p_i, p_i \in w \quad (2)$$

画像パッチ y_i 及び基底 d_i 中で、計算窓を縦横 $1[\text{pixel}]$ ずつずらして全体の平滑値を計算し、式(3)(4)に示すようにその値を並べ、それぞれ、平滑ベクトル $\text{smooth}(y_i)$ 、平滑ベクトル $\text{smooth}(d_i)$ とする。

$$\text{smooth}(y_i) = [\text{smooth}(w_1); \dots; \text{smooth}(w_n);] \text{ s.t. } \forall w_j, w_j \in y_i \quad (3)$$

$$\text{smooth}(d_i) = [\text{smooth}(w_1); \dots; \text{smooth}(w_n);] \text{ s.t. } \forall w_j, w_j \in d_i \quad (4)$$

ここで、 n は計算窓の数である。

原信号 Y に対応する全ての平滑ベクトルからなる行列を $\text{smooth}(Y)$ とし、辞書 D の基底 d_i に対応する平滑ベクトルからなる行列を $\text{smooth}(D)$ とし、式(1)に対して、式(5)のように近傍画素類似性制約を導入した。ここで σ は平滑制約調整のための重み値である。

$$\text{ksvd}2(D, X) = \min_{D, X} \|Y - DX\|_2^2 + \sigma \|\text{smooth}(Y) - \text{smooth}(D)X\|_2^2 \\ \text{ s.t. } \forall i, \|x_i\|_0 \leq T \quad (5)$$

2.2 エッジ特徴制約の追加

Canny 線形検出器[4]を利用して、画像のエッジ特徴を制約項として導入する。画像パッチ y_i のエッジ画像 $\text{canny}(y_i)$ 及び基底 d_i のエッジ画像 $\text{canny}(d_i)$ を計算する。エッジ画像中で、エッジ部分を 1、他の部分を 0 としてエッジ情報を表す。原信号 Y に対応する全ての $\text{canny}(y_i)$ を並べた行列を $\text{canny}(Y)$ とし、辞書 D の基底 d_i に対応する $\text{canny}(d_i)$ を並べた行列を $\text{canny}(D)$ とし、式(1)に対して、式(6)のようにエッジ特徴制約を導入した。ここで ω はエッジ特徴制約調整のための重み値である。

$$\text{ksvd}3(D, X) = \min_{D, X} \|Y - DX\|_2^2 + \omega \|\text{canny}(Y) - \text{canny}(D)X\|_2^2 \\ \text{ s.t. } \forall i, \|x_i\|_0 \leq T \quad (6)$$

3. 実験

近傍画素類似性制約を導入したもの(方式 1)、エッジ特徴制約を導入したもの(方式 2)、両者を導入したもの(方式 3)を、K-SVD の基底と比較する。

3.1 実験手順

Caltech 101 データセット[5]から 100 枚の画像をランダムに選択し、それらから 10 万個の $8[\text{pixel}] \times 8[\text{pixel}]$ の画像パッチをランダムに選択して、輝度のみを学習データとした。画像パッチ中の各画素を $[0, 1]$ の値に変換後、パッチ中の画素の平均値を引いた値として正規化する。実験のパラメータを表 1 に示す。

表 1 実験のパラメータ

	辞書サイズ	σ	ω	T	反復計算数
実験 1	256	0.5	0.5	25	100
実験 2	256	0.1	0.02	25	100

基底を求めた後、同じ 100 枚の学習画像から $8[\text{pixel}] \times 8[\text{pixel}]$ の画像パッチ 10 万枚をランダムに選択し、求めた基底の評価を行った。

[†] 早稲田大学大学院基幹理工研究科

[‡] 早稲田大学基幹理工学部情報通信学科

* 三菱電機株式会社情報技術総合研究所

3.2 評価基準

評価基準として、RMSE (式(7)) と基底の近傍画素類似度 (式(8)) を求める。式(7)で P は画像中の画素の数を示す。式(8)で、 N は辞書のサイズ、 n は基底中の計算窓の数、 $\text{smooth}(w_j)$ は基底 d_i の j 番目の計算窓の平滑値である。なお、式(8)は、近傍画素類似性が良いほど値が小さくなる。

$$RMSE = \sqrt{\frac{1}{P} \sum_{i=1}^P (Y_i - (DX)_i)^2} \quad (7)$$

$$S = \sum_{i=1}^N \sum_{j=1}^n \text{smooth}(w_j) \text{ s.t. } \forall w_j, w_j \in d_i \quad (8)$$

3.3 実験結果

実験 1 の結果を表 2 と図 1 に示す。

表 2 実験 1 の結果

手法	RMSE	近傍画素類似度
K-SVD	0.0138	5.9647e+03
方式 1	0.0320	1.8531e+03
方式 2	0.0272	3.6520e+03
方式 3	0.0405	1.8399e+03

方式 1、2、3 は、K-SVD より基底画像の近傍画素類似性が高いが、信号の復元力は低下している。また図 1 から、それぞれの制約の導入によって、基底画像が変化していることが確認できる。しかしながら、再現性が元の K-SVD に比べて悪いいため、表 1 の実験 2 に示すパラメータ調整を行い、次に、RMSE がほぼ一定という条件の下で基底画像を比較する。

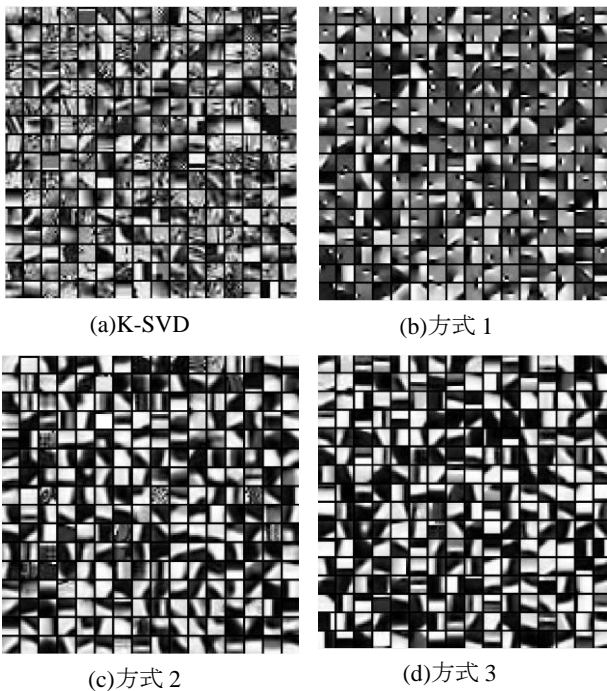


図 1 実験 1 の基底画像

実験 2 の結果を表 3 と図 2 に示す。

表 3 実験 2 の結果

手法	RMSE	近傍画素類似度
K-SVD	0.0138	5.9647e+03
方式 1	0.0144	5.6202e+03
方式 2	0.0139	5.6669e+03
方式 3	0.0149	5.6494e+03

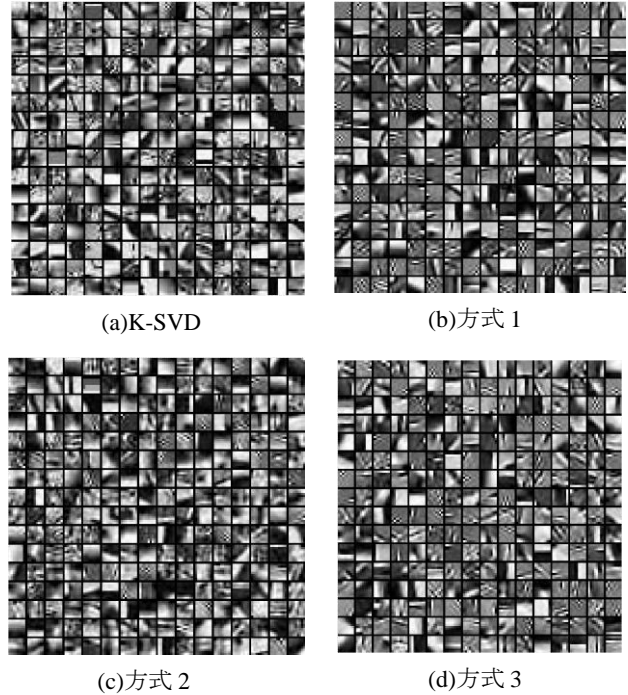


図 2 実験 2 の基底画像

表 3 から、新たな制約の導入によって、近傍画素類似度は少し改善されていることが分かる。また、図 2 から、基底画像はそれぞれの制約によって変化していることが確認できるが、影響はあまり大きくないことが分かる。

4. まとめと今後の課題

一般的な画像の特徴である近傍画素類似性とエッジ特徴を制約項としてスパースコーディングに導入し、基底画像の変化を確認した。今後の課題として、各種パラメータの最適値を求め、得られた基底を具体的な画像処理に適用し、新たな制約項導入の効果を評価する必要がある。

参考文献

- [1] A. Rehman, M. Rostami, Z. Wang, D. Brunet, and E. R. Vrscay, "SSIM-inspired image restoration using sparse representation", EURASIP J. Adv. Signal Process., Vol.2012, p.16, Jan. 2012
- [2] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation", IEEE Transactions on Signal Processing, Vol.54, pp.4311-4322, 2006
- [3] H. Ishikawa, "Higher-order clique reduction in binary graph cut", Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2009, pp.2993-3000, Jun. 2009
- [4] J. Canny, "A computational approach to edge detection", IEEE PAMI, Vol.8, No.6, pp.679-698, 1986
- [5] http://www.vision.caltech.edu/Image_Datasets/Caltech101/ (2016年6月1日 最終確認) .