

## 初期辞書設定による特異値分解を用いた辞書学習アルゴリズムの検討 Consideration of Dictionary Learning with Singular Value Decomposition by Initial Dictionary Setting

釋迦郡 充<sup>†</sup>      前田 道治<sup>†</sup>  
Ataru Shakagori      Michiharu Maeda

### 1. はじめに

本研究では、特異値分解を用いた辞書学習アルゴリズムで使用する高速かつ効果的な初期辞書を検討する。

### 2. 辞書学習

辞書学習で用いる K-SVD(K-Singular Value Decomposition) アルゴリズムとは、スパースコーディングの際に、最適化されたスパースな表現をもつ列ベクトル  $X_i$  から構成されているスパース表現行列  $X$  に基づいて辞書  $D$  における辞書の列(アトム)に対して 1 つずつ更新を行う際に、特異値分解を用いてアトムとそのアトムに対応するスパース表現ベクトルの最適化を行うアルゴリズムである。また、以下の最適化問題として定式化される。

$$\arg\min_{X,D} \|Y - DX\|_F^2 \quad \text{subject to} \quad \forall_i, \|X_i\|_0 \leq T_0 \quad (1)$$

ここで、 $T_0$  はスパース表現ベクトル  $X_i$  における最大非零要素数を表す。

以下に、K-SVD アルゴリズムについて示す。

#### Step 1. 初期化

初期辞書として用意した各要素に対して、ランダムに生成した値を与え、初期化を行いアトムの正規化を行う。

#### Step 2. スパースコーディング

辞書  $D$  は固定し、直行マッチング追跡を用いて式(2)の解と近似を行い、各  $y_i$  に基づいてスパース表現ベクトル  $X_i$  の計算を行い、そしてスパース表現行列  $X$  の作成を行う。

$$\arg\min_{x_i} \|y_i - Dx_i\|_F^2 \quad \text{subject to} \quad \forall_i, \|X_i\|_0 \leq T_0$$

$$\text{for } i = 1, 2, \dots, N \quad (2)$$

#### Step 3. コードブック更新

辞書  $D$  における各アトム  $d_j, j = 1, 2, \dots, K$  に対して以下の通りを繰り返す。

3.1 式(1)を以下のように分解を行い、 $K$  番目のアトム以外の残差をまとめた行列を残差行列  $E_k$  と置き、残差行列  $E_k$  との再構成誤差を最小化する  $d_k, x^k$  について求める。

$$\|Y - DX\|_F^2 = \left\| Y - \sum_{j=1}^K d_j x^j \right\|_F^2$$

$$= \left\| \left( Y - \sum_{j \neq k} d_j x^j \right) - d_k x^k \right\|_F^2$$

$$= \|E_k - d_k x^k\|_F^2 \quad (3)$$

ここで  $d_k$  は、辞書  $D$  における  $k$  列目を表し、 $x^k$  は、スパース表現行列  $X$  における  $k$  行目を表している。

3.2 アトム  $d_k$  を用いるかつ  $x^k$  内における非零要素のみの更新を行うための集合  $\omega_k$  を以下のように定義する。

$$\omega_k = \{i | 1 \leq i \leq K, x^k(i) \neq 0\} \quad (4)$$

3.3  $(\omega_k(i), i)$  成分には 1, それ以外には 0 をとる、大きさ  $N \times |\omega_k|$  の行列  $\Omega_k$  を定義する。  $x^k, Y, E_k$  に対して行列  $\Omega_k$  を用いて積をとり、 $x_R^k, Y_R^k, E_k^R$  を作成し、式(3)を用いて式(5)について定義する。

$$\|E_k \Omega_k - d_k x^k \Omega_k\|_F^2 = \|E_k^R - d_k x_R^k\|_F^2 \quad (5)$$

3.4  $E_k^R$  に対して特異値分解を用いて、 $E_k^R = U \Delta V^T$  と分解を行う。ここで、式(5)における最小化を行うため、 $d_k, x_R^k$  には、最大特異値に基づく特異ベクトルを用いる。アトム  $d_k$  の解を  $U$  における最初の列  $u_1$ 、 $x_R^k$  には  $V$  における最初の列と  $\Delta(1,1)$  の積をとった  $v_1 \times \Delta(1,1)$  として更新を行う。

#### Step 4. 停止条件

$\|Y - DX\|_F^2$  における再構成誤差の値に変化が十分に小さくなれば反復を終了し、変化が大きい場合は Step 2. に戻り繰り返し行う。

## 3. 初期辞書

### 3.1 離散コサイン変換を用いた辞書

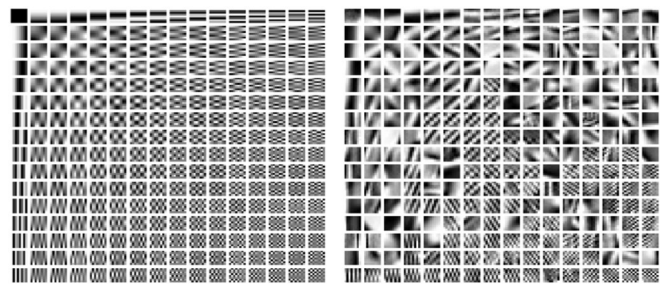
離散コサイン変換を用いて作成された辞書を DCT 辞書と呼び、従来手法として一般的に使われてきた辞書である。DCT 辞書の作成にあたって式(6)を用いて行った。

$$X_{\text{DCT}}(k) = \sqrt{\frac{2}{N}} C_k \sum_{n=0}^{N-1} x(n) \cos \frac{(2n+1)k\pi}{2N},$$

$$\text{for } k = 0, 1, \dots, N-1 \quad (6)$$

$$C_k = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0 \\ 1, & k \neq 0 \end{cases} \quad (7)$$

DCT 辞書と K-SVD アルゴリズムによって学習を行った K-SVD\_DCT 辞書の画像を図 1 に示す。



(a)DCT 辞書      (b)K-SVD\_DCT 辞書

図 1 従来手法 : DCT による辞書

<sup>†</sup> 福岡工業大学大学院工学研究科 Graduate School of Engineering, Fukuoka Institute of Technology

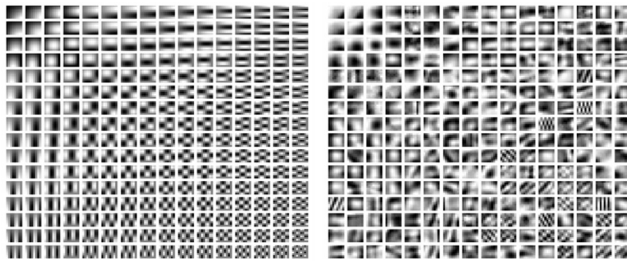
図 1 より, DCT 辞書の特徴として, (1,1)成分での値は直流成分だけから構成される DC 基底となっており, それ以外の基底は交流成分のみから構成されている為, AC 基底になっている. また, 左上から右下にかけて, 各基底の周波数が高くなるといった特徴をもっている. 学習後の辞書の特徴として, DC 基底が直流成分ではなくなり, 高周波成分における基底が低周波成分になり, 同じパターンの基底が幾つか作成されていることも分かる.

### 3.2 離散サイン変換を用いた辞書

提案手法 1 つ目の離散サイン変換を用いた辞書を DST(Discrete Sine Transform)辞書と呼ぶ. 辞書作成にあたって式(8)を用いて行った.

$$X_{\text{DST}}(u, v) = \frac{2}{N+1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) \sin\left(\frac{(m+1)(u+1)\pi}{2N}\right) \cdot \sin\left(\frac{(n+1)(v+1)\pi}{2N}\right) \quad \text{for } u, v = 1, 2, \dots, N-1 \quad (8)$$

DST 辞書と K-SVD アルゴリズムによって学習を行った K-SVD\_DST 辞書の画像を図 2 に示す.



(a)DST 辞書 (b)K-SVD\_DST 辞書

図 2 提案手法 1 : DST による辞書

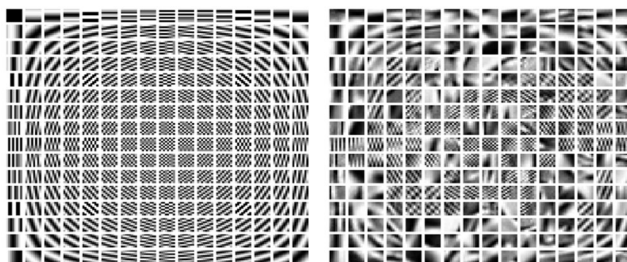
図 2 より, DST 辞書の特徴として, 正弦波からなる基底関数を用いるため, 基底における端点の値が 0 になるような滑らかな基底となっている. DCT 辞書同様に周波数が左上から右下にかけて, 高周波成分になっているといった特徴がある. 学習後の辞書の特徴は, 複雑的で局所的な変動をもつ基底が表現されており, 入力画像からの学習の為, 同じパターンの基底の作成がされていることが分かる.

### 3.3 DHT(Discrete Hartley Transform)辞書

提案手法 2 つ目の離散ハートレー変換を用いた辞書を DHT(Discrete Hartley Transform)辞書と呼ぶ. 辞書作成にあたって式(9)を用いて行った.

$$X_{\text{DHT}}(u, v) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{cas}\left(\frac{2\pi}{N}(mu + nv)\right) \quad \text{for } u, v = 1, 2, \dots, N-1. \quad (9)$$

DST 辞書と K-SVD アルゴリズムによって学習を行った K-SVD\_DST 辞書の画像を図 3 に示す.



(a)DHT 辞書 (b)K-SVD\_DHT 辞書

図 3 提案手法 2 : DHT による辞書

図 3 より, DHT 辞書の特徴として,  $\text{cas}\theta = \sin\theta + \cos\theta$  を表した関数を用いて辞書の作成を行っている為, 周期的かつ規則的な基底となっており, 基底配列における四隅が低周波成分となり, 中心に向かうにつれ高周波成分になっている. また, 実数のみの計算となる為, 高速で構築することが可能である. 学習後の特徴として, 中心成分に向けて基底が低周波成分に変化が起きていることが分かる. また, 他 2 つの辞書同様, 同じパターンの基底が作成されていることが分かる.

## 4. 評価実験

本評価実験では, K-SVD における辞書学習の時間の比較と初期辞書と学習辞書を用いて画像のノイズ除去を行い辞書性能の評価について行った. 学習に用いる初期辞書として, DCT 辞書, DST 辞書, DHT 辞書を使用し辞書学習の時間の計測を行った. K-SVD 辞書学習を行うにあたって, 学習繰り返し回数を 15 回とし, 試行回数を 30 回行った. 以下に, 各辞書の学習にかかった時間について表 1 に示す.

表 1 : K-SVD 辞書学習にかかる時間

	DCT	DST	DHT
Time(sec)	350.63	408.48	336.09

次に, 標準偏差  $\sigma = 5.0, 10.0, 15.0, 20.0$  のガウシアンノイズを付与した画像に対して初期辞書と学習辞書を用いてノイズ除去を行う. ノイズ除去における評価方法として, PSNR[dB]を用いて行った. 各辞書を用いてノイズ除去を行った結果を表 2 に示す.

表 2 : ノイズ除去における PSNR[dB]

辞書 \ $\sigma$	5.0	10.0	15.0	20.0
DCT	34.035	31.775	29.561	27.687
KSVD_DCT	36.382	32.672	29.829	27.171
DST	30.397	29.109	27.590	26.112
KSVD_DST	35.141	32.264	29.568	27.505
DHT	35.416	32.366	29.750	27.672
KSVD_DHT	36.363	32.665	29.840	27.723

表 1 より, 辞書学習にかかる時間について, 着目すると従来手法である DCT 辞書の学習時間より DHT 辞書の方が高速であることが分かった. 表 2 より, PSNR についても標準偏差が 15.0, 20.0 でも精度が高いことが分かる.

## 5. まとめ

本研究では, K-SVD による辞書学習において初期辞書の設定では, 従来手法である DCT 辞書より DHT 辞書を用いた方が有用であることが分かった.

### 参考文献

- [1] Michal Aharon, Michael Elad, and Alfred Bruckstein, "K-SVD : An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," IEEE Transactions on Signal Processing, Vol. 54, No. 11, November 2006, pp. 4311-4322.
- [2] R. N. Bracewell, O. Buneman, H. Hao, and J. Villasenor, "Fast Two-Dimensional Hartley Transform," Proceedings of the IEEE, Vol. 74, No. 9, September 1986, 1282-1283.
- [3] Vladimir Britanak and K. R. Rao, "Two-Dimensional DCT/DST Universal Computational Structure for  $2^m \times 2^n$  Block Sizes," IEEE Transactions on Signal Processing, Vol. 48, No. 11, November 2000, pp. 3250-3255.
- [4] M. Elad, 玉木 徹, スパースモデリング, 共立出版, 2016.