

フェーズドアレイ気象レーダのデータ圧縮と 3階テンソル辞書学習を用いた高精度再構成

Data Compression and High-Quality Reconstruction for Phased Array Weather Radar Using Third-Order Tensor Dictionary Learning

川見 亮介[†] 北原 大地[†] 平林 晃[†] 吉川 栄一[‡] 菊池 博史* 牛尾 知雄*

[†]立命館大学 情報理工学部 [‡]JAXA 航空技術部門 * 首都大学東京 システムデザイン学部

R. Kawami[†] D. Kitahara[†] A. Hirabayashi[†] E. Yoshikawa[‡] H. Kikuchi* T. Ushio*

[†]Ritsumeikan University [‡]Japan Aerospace Exploration Agency *Tokyo Metropolitan University

アブストラクト フェーズドアレイ気象レーダ (PAWR) は、半径 60 キロメートルの全天気象状況を 30 秒間で観測できる高性能レーダである。PAWR の距離分解能は 100 メートルであり、30 秒ごとの取得データ量は約 491 Mbyte にも及ぶ。この大規模気象データを転送・保存するための圧縮方法として、ランダム間引きによりデータを削減し、凸最適化により再構成する手法が提案されてきた。最適化のための評価関数には全変動などが用いられてきたが、当該データに適したものは限らず、再構成精度に限界があった。そこで本研究では、過去の気象データの特徴を表現する 3 階テンソル辞書を事前に学習し、この辞書に基づいて現在の気象データを再構成する手法を提案する。計算機シミュレーションにより、提案法の有効性を示す。

1 はじめに

増加傾向にあるゲリラ豪雨などの局所的大雨を時間的・空間的に高分解能で観測するために、フェーズドアレイ気象レーダ (Phased Array Weather Radar: PAWR) が開発されてきた [1]。例えば、大阪大学で開発された PAWR は半径 60 キロメートルの全天気象状況を観測するが、その時間分解能と空間分解能はそれぞれ 30 秒と 100 メートルであり非常に高い。一方で、高分解能であるがゆえに取得するデータ量も膨大となってしまう。大阪大学の PAWR は、13 種類の気象パラメータを距離、方位角、仰角方向にそれぞれ 600, 300, 110 点ずつ観測する。これらのデータを 2 byte で記述するため、30 秒間で約 491 Mbyte のデータを取得する。PAWR をネットワークで運用していく上で、リアルタイムにデータを転送する必要がある。大阪大学の PAWR データをリアルタイムで転送するには約 137 Mbps の通信速度が必要となるが、通常のインターネット回線がこの速度を保証することは現実的ではない。通常の回線で PAWR を運用するためには、最大でも 35 Mbps 程度までの速度でデータを転送しなければならない、これはデータ量

を 25% に圧縮することに相当する。よって、PAWR データを最低でも 25%、あるいはより小さな値に圧縮し、そこから高精度に元データを再構成することが求められる。

この問題の解決のため、圧縮センシングに基づいた手法が開発されてきた。3 次元気象データから抽出された 2 次元データを圧縮し、その圧縮データから低ランク近似を用いて再構成する手法 [2] や、凸最適化手法の一つである SDMM アルゴリズムを用いて再構成する手法 [3] が提案されている。また、観測データを 3 次元データのまま圧縮し、SDMM を用いて再構成する手法も提案されている [4]。しかし、これらの手法では過去に観測された気象データの情報は用いられていない。本研究では、過去のデータを用いて辞書学習を行う再構成手法を提案する。ここで問題になるのは、高次元データに起因する膨大な計算量である。この問題に対して提案法では、辞書の各アトムに構造制約を持たせることで計算量を削減させる。これにより、3 階テンソル辞書学習を実現し、学習された気象データの特徴を利用して圧縮データからの再構成を行う。大阪大学の PAWR で実際に観測されたデータを使用した計算機シミュレーションにより、提案法の有効性を示す。

2 気象データの圧縮センシング

PAWR が観測する 13 種類のパラメータの内、本研究では反射因子 [dBZ] に着目し、この値を 3 階テンソル $\mathbf{X}_0 \in \mathbb{R}^{N_H \times N_V \times N_W}$ によって表現する。ここで、 $N_H := 600$, $N_V := 300$, $N_W := 110$ である。

データを圧縮するために、 \mathbf{X}_0 から観測点をランダムに選択する。この処理を写像 \mathcal{A} で表す。 \mathcal{A} は選択されなかった観測点の反射因子値を 0 に置き換える写像である。これにより圧縮された観測信号 $\mathbf{Y} \in \mathbb{R}^{N_H \times N_V \times N_W}$ は

$$\mathbf{Y} = \mathcal{A}(\mathbf{X}_0) \quad (1)$$

と表される。また、このランダムパターンは PAWR が回転

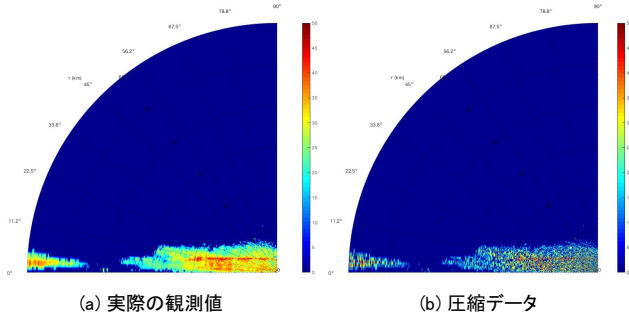


図 1: 実際の観測値と圧縮データ

するごとに生成される。実際の観測値 \mathbf{X}_0 と式 (1) の圧縮データ \mathbf{Y} を図 1 に示す。ここで、圧縮率は 25% とした。また、対流圏の外側では雲が発生しないことに基づいて、対流圏内のデータのみをランダムにサンプリングしている。

3 3階テンソル辞書学習

辞書学習とは、対象信号をよりスパースに表現する基底（あるいはフレーム）を生成する信号処理技術である。提案法では、計算量削減のために辞書の各アトムに構造制約を持たせる。すなわち、辞書を $\mathbf{D}_\Omega = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_K] \in \mathbb{R}^{n \times n \times n \times K}$ としたとき、各アトム $\mathbf{D}_k \in \mathbb{R}^{n \times n \times n}$ はある 3 つのベクトル $\mathbf{d}_{k,1}, \mathbf{d}_{k,2}, \mathbf{d}_{k,3} \in \mathbb{R}^n$ ($\|\mathbf{d}_{k,1}\|_2 = \|\mathbf{d}_{k,2}\|_2 = \|\mathbf{d}_{k,3}\|_2 = 1$) のテンソル積で表現されるものとする。この制約によって辞書生成の計算量を削減できる。学習データから $n \times n \times n$ のパッチを取り出したものを $\mathbf{Z}_p \in \mathbb{R}^{n \times n \times n}$ ($p = 1, 2, \dots, P$) とすると、3階テンソル辞書学習の問題は

$$\{\hat{\mathbf{D}}_\Omega, \{\hat{\mathbf{c}}_p\}\} = \underset{\mathbf{D}_\Omega, \{\mathbf{c}_p\}}{\operatorname{argmin}} \sum_{p=1}^P \|\mathbf{Z}_p - \mathbf{D}_\Omega \times_4 \mathbf{c}_p^T\|_F^2 \quad (2)$$

$$\text{s.t. } \forall k \mathbf{D}_k = \mathbf{d}_{k,1} \otimes \mathbf{d}_{k,2} \otimes \mathbf{d}_{k,3} \text{ and } \forall p \|\mathbf{c}_p\|_0 \leq T_0$$

と定式化できる。ここで、 $\mathbf{c}_p \in \mathbb{R}^K$ は表現係数、 $\|\mathbf{c}_p\|_0$ は \mathbf{c}_p の非ゼロ要素数、 T_0 はその上限値である。また、 \times_4 は第 4 モード積、 \otimes はテンソル積を意味する。式 (2) は一般に NP 困難な問題であるため、係数 $\{\mathbf{c}_p\}$ を更新するステップと辞書 \mathbf{D}_Ω を更新するステップに分解して、近似的に解決する。前者には MOMP アルゴリズム、後者には CP 分解を利用した K-CPD アルゴリズムを利用できる [5]。これにより、学習データから辞書を事前に生成する。

次に事前に学習した辞書 $\hat{\mathbf{D}}_\Omega$ を用いて、圧縮データ \mathbf{Y} から元データ \mathbf{X}_0 を復元する。このために以下の式を解く。

$$\{\hat{\mathbf{X}}, \{\hat{\mathbf{c}}_s\}\} = \underset{\mathbf{X}, \{\mathbf{c}_s\}}{\operatorname{argmin}} \sum_{s=1}^S \|R_s(\mathbf{X}) - \hat{\mathbf{D}}_\Omega \times_4 \mathbf{c}_s^T\|_F^2 \quad (3)$$

$$\text{s.t. } \mathbf{Y} = \mathcal{A}(\mathbf{X}) \text{ and } \forall s \|\mathbf{c}_s\|_0 \leq T_0.$$

ここで、 R_s はパッチを抽出する作用素、 s は対象データのパッチインデックス、 S はパッチの総数、 $\mathbf{c}_s \in \mathbb{R}^K$ は各

表 1: シミュレーション結果

	従来法 [4]	提案法
方位 no. 100–104	11.82%	11.51%
方位 no. 200–204	10.78%	10.54%

パッチの表現係数である。この最適化問題を解くことによって、圧縮データから再構成結果 $\hat{\mathbf{X}}$ を得る。

4 シミュレーション

大阪大学吹田キャンパスに設置された PAWR で観測されたデータを用いてシミュレーションを行なった。2014 年 3 月 30 日に観測されたデータにおいて、第 1 から 4 方位のデータを学習データとした。再構成対象は、同日に観測されたデータの中で、第 100 から 104 方位のデータと第 200 から 204 方位のデータである。圧縮率は 25%、 $T_0 = 8$ 、パッチサイズは $n = 4$ 、ストライドは 2、辞書のアトム数は $K = 256$ とした。シミュレーションの結果を表 1 に示す。表の値は正規化誤差であり、 $100 \|\hat{\mathbf{X}} - \mathbf{X}_0\|_F / \|\mathbf{X}_0\|_F$ で計算される。表 1 が示す通り、提案法は従来法と比べて正規化誤差を第 100 から 104 方位のデータに対しては 0.31%、第 200 から 204 方位のデータに対しては 0.24% 低減した。

5 おわりに

PAWR の圧縮データに対して、辞書学習を用いた 3 次元再構成手法を提案した。シミュレーション結果より、提案法が従来法よりも高精度に再構成が行えることを示した。今後は、更なる高精度化を目指す。

参考文献

- [1] F. Mizutani, T. Ushio, E. Yoshikawa, S. Shimamura, H. Kikuchi, M. Wada, S. Satoh, and T. Iguchi, “Fast-scanning phased-array weather radar with angular imaging technique,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2664–2673, 2018.
- [2] K. V. Mishra, A. Kruger, and W. F. Krajewski, “Compressed sensing applied to weather radar,” in *Proceedings of IEEE Geoscience and Remote Sensing Symposium*, pp. 1832–1835, 2014.
- [3] R. Kawami, A. Hirabayashi, N. Tanaka, M. Shibata, T. Ijiri, S. Shimamura, H. Kikuchi, G. Kim, and T. Ushio, “2-Dimensional high-quality reconstruction of compressive measurements of phased array weather radar,” in *Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 7 pages, 2016.
- [4] R. Kawami, H. Kataoka, D. Kitahara, A. Hirabayashi, T. Ijiri, S. Shimamura, H. Kikuchi, and T. Ushio, “Fast high-quality three-dimensional reconstruction from compressive observation of phased array weather radar,” in *Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 44–49, 2017.
- [5] G. Duan, H. Wang, Z. Liu, J. Deng, and Y. W. Chen, “K-CPD: Learning of overcomplete dictionaries for tensor sparse coding,” in *Proceedings of International Conference on Pattern Recognition (ICPR)*, pp. 493–496, 2012.