

A Distributed and Tensor-based Approach for Multidimensional Dictionary Learning

Thiago P de B Vieira

Department of Electrical Engineering
University of Brasilia
Brasília-DF, Brazil
Email: tpbvieira@gmail.com

André L. F. de Almeida

Department of Teleinformatics Engineering
Federal University of Ceará
Fortaleza, Brazil
Email: andre.almeida@ieee.org

Florian Roemer

and João Paulo C. L. da Costa
Institute for Information Technology
Ilmenau University of Technology
Ilmenau, Germany

Abstract—The abstract goes here.

I. INTRODUCTION

Dictionary Learning is a signal processing technique for sparse representation of signals as basis vectors, learning the representations from training data, as dictionaries. The sparse representation in terms of such dictionaries has attracted increased interest for compressive sensing and for solving problems such as denoising, compression, classification, data decomposition, feature extraction and image processing [1], [2], [3], [4].

In some applications the data and its dictionary are multidimensional, e.g., when estimating jointly behavior of users in social networks. Computing tensor decompositions of multiway datasets is particularly useful to extract hidden patterns and structure in multidimensional data analytics problems [5]. Tensor-based algorithms for dictionary learning can improve the performance for cases of multidimensional and separable data, regarding the dictionary identification rating, the required number of training samples and iterations for the optimization problem [6].

Multidimensional parameter estimation and learning multidimensional separable dictionaries are growing research problems. Roemer *et al.* [6] show that the multidimensional dictionary estimation problem can be efficiently formulated in terms of tensors, and that their results outperform existing schemes by exploiting the multilinear structure of the problem.

Existing dictionary learning schemes can be applied to multidimensional analysis and obtain valuable results. However, the performance of tensor-based algorithms outperform existing schemes when dealing with growing multidimensional datasets, as can be seen in Figures 1 and 2, which show that the detection rating of traditional algorithms decreases over the dataset increasing.

Figures 1 and 2 show the evaluation of tensor-based and traditional dictionary learning algorithms for dictionary reconstruction of a multidimensional and separable data. Figure 1 presents the results for dictionary reconstruction from a matrix of 1000 elements, while Figures 2 presents the results from a matrix with 24000 entries. It is possible to observe better detection rating for tensor-based algorithms. However, tensor-based algorithms face challenges to achieve reasonable

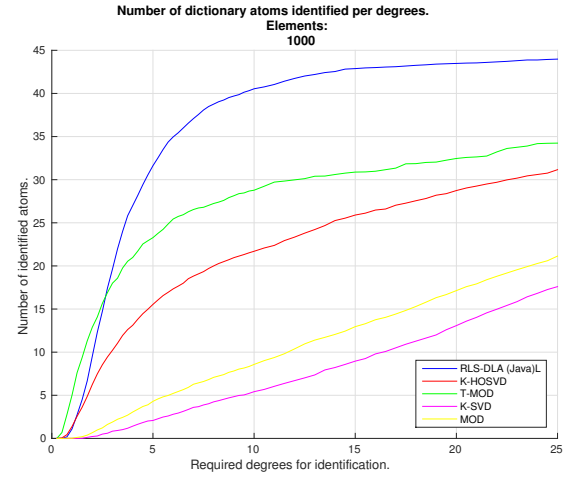


Fig. 1. Matrix of 1.000 elements, 2000 sample training, 100 iterations

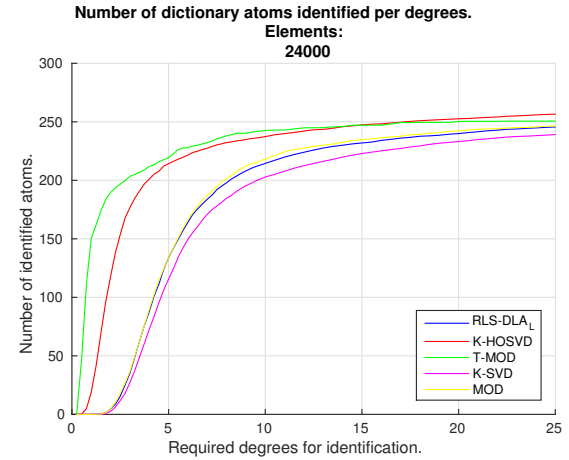


Fig. 2. Matrix of 24.000 elements, 2000 sample training, 100 iterations

processing time to handle large-scale tensor factorizations [7], demanding efforts in order to explore distributed processing techniques for tensor-based analytics of large datasets.

We propose a distributed and tensor-based approach for multidimensional dictionary learning in order to obtain better processing time for larger datasets. We focus on a distributed

implementation of T-MOD [6] algorithm, based on a modification of Almeida and Kibangou's [7] approach to use Tucker-2 decomposition. We perform experiments to test different methods for learning dictionaries, evaluating the processing time and the reconstruction of a dictionary of multidimensional and separable data.

II. DATA MODEL

Consider a generic sparse recovery problem of the following form

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S} + \mathbf{W}, \quad (1)$$

where $\mathbf{X} \in \mathbb{C}^{M \times T}$ represents T consecutive observations from M features, $\mathbf{A} \in \mathbb{C}^{M \times N}$ is the overcomplete dictionary, $\mathbf{S} \in \mathbb{C}^{N \times T}$ represents the sparse coefficient matrix, $\mathbf{W} \in \mathbb{C}^{M \times T}$ is the additive noise, and $M < N < T$.

Consider a sparse recovery problem for a separable 2-D manifold, i.e., a manifold that can be written as the product of two 1-D manifolds. If we choose a separable 2-D sampling grid, we can write the 2D sparse recovery problem as

$$\mathbf{X} = (\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)}) \cdot \mathbf{S} + \mathbf{W}, \quad (2)$$

where $\mathbf{A} = (\mathbf{A}^{(1)} \otimes \mathbf{A}^{(2)})$, $\mathbf{A}^{(1)} \in \mathbb{C}^{M_1 \times N_1}$, $\mathbf{A}^{(2)} \in \mathbb{C}^{M_2 \times N_2}$, $M = M_1 \times M_2$, and $N = N_1 \times N_2$.

The Kronecker model (2) can be rewritten in an equivalent tensor form. Applying the algebraic rules for unfoldings of n-mode products [6] we can rewrite (2) into a "Tucker-2" decomposition, as

$$\mathcal{X} = \mathcal{S} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} + \mathcal{W}, \quad (3)$$

where $\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \times T}$, $\mathcal{S} \in \mathbb{C}^{N_1 \times N_2 \times T}$, and $\mathcal{W} \in \mathbb{C}^{M_1 \times M_2 \times T}$ are rearranged versions of the matrices \mathbf{X} , \mathbf{S} , and \mathbf{W} such that $\mathbf{X} = [\mathcal{X}]_{(3)}^T$, $\mathbf{S} = [\mathcal{S}]_{(3)}^T$, and $\mathbf{W} = [\mathcal{W}]_{(3)}^T$, respectively.

We consider a randomly generated data in order to evaluate the proposed approach, where initially $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ are gaussian distributed and randomly generated, then the dictionary \mathbf{A} is obtained. In order to be able to evaluate the dictionary reconstruction, the sparse data \mathcal{X} is generated from the dictionary \mathbf{A} , which is used for further comparisons. We adopt 20 as the signal-to-noise ratio (snr) and 5 as the sparseness factor of the data \mathcal{X} .

The dictionary estimation is evaluated by applying the proposed approach to estimate the dictionary \mathbf{A} from the sparse data \mathcal{X} , according the selected number of sample training and iterations for the optimization problem.

III. PROPOSED APPROACH

...

IV. EXPERIMENTS

...

V. CONCLUSION

The conclusion goes here.

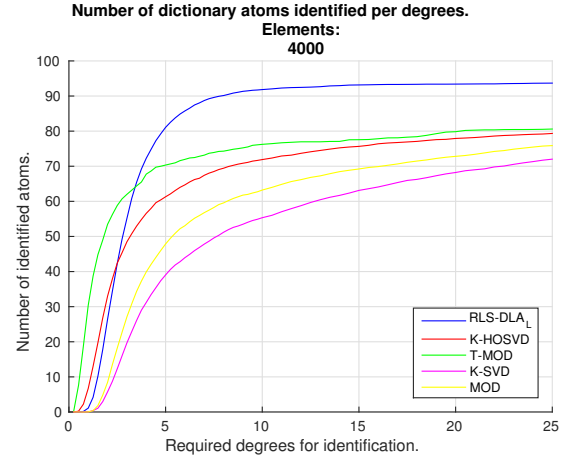


Fig. 3. Matrix of 4,000 elements, 2000 sample training, 100 iterations

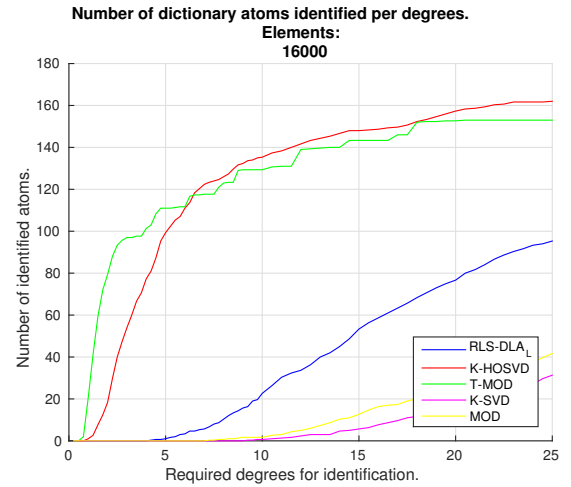


Fig. 4. Matrix of 16,000 elements, 500 sample training, 100 iterations

ACKNOWLEDGMENT

This research was supported by ...

REFERENCES

- [1] I. Tosić and P. Frossard, "Dictionary learning," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 27–38, 2011.
- [2] Q. Zhang and B. Li, "Discriminative k-svd for dictionary learning in face recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2691–2698.
- [3] P. Zhu, Q. Hu, C. Zhang, and W. Zuo, "Coupled dictionary learning for unsupervised feature selection," in *AAAI*, 2016, pp. 2422–2428.
- [4] S. Ravishanker and Y. Bresler, "Mr image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE transactions on medical imaging*, vol. 30, no. 5, pp. 1028–1041, 2011.
- [5] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [6] F. Roemer, G. Del Galdo, and M. Haardt, "Tensor-based algorithms for learning multidimensional separable dictionaries," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3963–3967.
- [7] A. L. De Almeida and A. Y. Kibangou, "Distributed large-scale tensor decomposition," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 26–30.