This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Exploiting Global Low-Rank Structure and Local Sparsity Nature for Tensor Completion

Yong Du, Guoqiang Han, Yuhui Quan, Zhiwen Yu, *Senior Member, IEEE*, Hau-San Wong, C. L. Philip Chen, *Fellow, IEEE*, and Jun Zhang, *Fellow, IEEE*

***Abstract*—In the era of data science, a huge amount of data has emerged in the form of tensors. In many applications, the collected tensor data are incomplete with missing entries, which affects the analysis process. In this paper, we investigate a new method for tensor completion, in which a low-rank tensor approximation is used to exploit the global structure of data, and sparse coding is used for elucidating the local patterns of data. Regarding the characterization of low-rank structures, a weighted nuclear norm for the tensor is introduced. Meanwhile, an orthogonal dictionary learning process is incorporated into sparse coding for more effective discovery of the local details of data. By simultaneously using the global patterns and local cues, the proposed method can effectively and efficiently recover the lost information of incomplete tensor data. The capability of the proposed method is demonstrated with several experiments on recovering MRI data and visual data, and the experimental results have shown the excellent performance of the proposed method in comparison with recent related methods.**

***Index Terms*—Orthogonal dictionary learning, sparse coding, tensor completion, weighted nuclear norm.**

Manuscript received February 7, 2018; revised June 15, 2018; accepted June 21, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61602184, Grant 61472145, Grant 61722205, Grant 61751205, Grant 61572199, Grant 61528204, and Grant U1611461616672241616021 84, in part by the Natural Science Foundation of Guangdong Province under Grant 2017A030313376 and Grant 2017A030312008, in part by the Science and Technology Program of Guangzhou under Grant 201707010147 and Grant 201604016040, in part by the Cultivation Project of Major Basic Research of NSF-Guangdong Province under Grant 2016A030308013, in part by the Applied Science and Technology Research Project of Guangdong Province under Grant 2016B010127003, in part by the Science and Technology Planning Project of Guangdong Province under Grant 2015A050502011, Grant 2016A050503015, and Grant 2016B090918042, in part by the Research Grants Council of the Hong Kong Special Administrative Region under Grant CityU-11300715, and in part by the City University of Hong Kong under Grant 7004884. This paper was recommended by Associate Editor S. Cruces. *(Corresponding author: Guoqiang Han.)*

Y. Du, G. Han, Y. Quan, Z. Yu, and J. Zhang are with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510640, China (e-mail: csyongdu@mail.scut.edu.cn; csgqhan@scut.edu.cn; csyhquan@scut.edu.cn; zhwyu@scut.edu.cn).

H.-S. Wong is with the Department of Computer Science, City University of Hong Kong, Hong Kong.

C. L. P. Chen is with the Department of Computer and Information Science, University of Macau, Macau, China (e-mail: philip.chen@ieee.org).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the author.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2018.2853122

I. Introduction

IN MANY data mining and machine learning scenarios, the collected data are in the form of multidimensional arrays, for example, multispectral images, image patch stacks, videos, and magnetic resonance imaging (MRI) data, which can be represented as tensors. The tensor data (i.e., data in the form of tensor) encode very rich structural information, and tensor analysis is widely used in gait recognition [1], [2]; object recognition [3]; dynamic texture recognition [4]; image recovery [5]–[7]; medical image processing [8], [9]; and many machine-learning fields [10]–[17].

The tensor data collected in the real world are often incomplete, which may be caused by occlusions, noise, partial damages, difficulties in collection, or data loss during transition. For example, people may wear sunglasses in surveillance videos. In recommendation systems, the preferences of users are often only available partially. Such incompleteness of tensor data may significantly decrease the quality of the data, making the analysis process very difficult.

This incompleteness problem can be remedied through tensor completion, which is to recover the complete tensor from the incomplete one. There have been some approaches proposed for this purpose, which can be roughly classified into two categories: 1) the matching-based approaches (e.g., [18] and [19]) that find optimal patch correspondences to fill the holes in data and 2) the low-rank approximation approaches (e.g., [20] and [21]), which exploit the low-rank structures of the original tensor data for the completion. The matching-based approaches are often designed for inpainting large regions, while the low-rank approximation approaches are applicable to both damaged regions and random missing entries.

A. Related Work

This paper focuses on the development of an effective low-rank approximation approach for tensor completion. In the following text, we first review the existing low-rank tensor completion methods.

The basic idea of the low-rank approximation approach is that high-dimensional data are likely to lie in some low-dimensional spaces, since they often exhibit large similarities. Such low dimensionality of data can be well characterized by low-rank constraints on the data and hence, the global structure of tensor data encoded in their similarities can be recovered via low-rank minimization. Nevertheless, unlike the rank of a