Parallel POD Compression Accelerated by Binary-Swap Compositing

CHONGKE BI and KENJI ONO†

1. Introduction

Proper orthogonal decomposition (POD) can greatly help researchers to compress large-scale datasets for visualizing and analyzing them. Due to the limitation of memory size and computational cost, a parallel POD method has been proposed in [1]. However, this straightforward way brings the degradation of parallel efficiency because the half of processors becomes idle in recursive procedure. To address this issue, the binary-swap compositing [2] is employed.

Our approach mainly deals with time-varying fluid datasets ([x×y×z×*timesteps*]). Firstly, a large-scale dataset is divided into several subspace datasets. Then, these subspace datasets are compressed using POD method in parallel described in the next paragraph. Finally, the compressed dataset can be linearly restored.

In order to demonstrate parallel POD compression accelerated by the binary-swap compositing method, an example of compressing 8 time steps dataset with four processors is explained. Firstly, each processor compresses two neighbor time steps into one POD basement. Then the obtained 4 POD basements are subdivided into 8 parts to avoid the idling of processors. Here, the binary-swap compositing is employed for the data transfer between different processors. This process is recursively until the dataset cannot be compressed any more.

2. Results and Discussions

In this section, the proposed method is evaluated using the following three technique issues: *compression ratio* and *computational cost*.

Table 1: The computational cost and compression ratio of a compressed large-scale dataset with 8,000 (125*64) processors. The dataset is obtained from a time-varying flow turbulence simulation around a car.

Computational cost

Calculation Time	14.73s
Data Transfer Time	3.33s

Compression Ratio

Size of Original Dataset	100.00%
Size of Compressed Dataset	0.78%

Table 1 shows the result to compress a dataset (1000×400×250×128) obtained from a time-varying flow turbulence simulation around a car. 8,000 (125*64) processors are used to compress this dataset, which is divided into 125 subspace datasets (200×100×80×128) for compression. The proposed approach successfully compresses the dataset with the compression ratio of 0.78%, which means the size of the compressed data is 0.78% of the original dataset. Furthermore, the computational cost in Table 1 fully proved that the binary-swap compositing greatly accelerated the algorithm.

Acknowledgement: Part of the results is obtained by early access to the K computer at the RIKEN Advanced Institute for Computational Science.

References

[1] C. Bi and K. Ono, "POD-Based Parallel Compression for Visualizing Large-Scale Dataset," High Performance Computing Symposium 2013, no. 30, pp. 83, Tokyo, Japan, January 2013.

[2] K. Ma, J. S. Painter, C. D. Hansen, and M. F. Krogh, "Parallel Volume Rendering Using Binary-Swap Compositing," IEEE Computer Graphics and Applications, Vol. 14(4), pp. 59-68, July 1994.

† Advanced Visualization Research Team
Advanced Institute for Computational Science, RIKEN.