Proper orthogonal decomposition based outlier correction for PIV data

H. Wang¹, Q. Gao¹, J. Wang¹

1: Key Laboratory of Fluid Mechanics, Ministry of Education, Beijing University of Aeronautics and Astronautics, Beijing, China
*Correspondent author: qgao@buaa.edu.cn

Keywords: PIV, post-processing, outlier correction, POD

Introduction

It is quite obvious that a fast, reliable and fully automatic post-processing of PIV outlier correction (OC) is essential. Two procedures of outlier correction are often performed: detecting and replacing of spurious velocity vectors. Most methods (Shinseeb et al. 2004, Westerweel and Scaran 2005) detect outliers based on a local statistical model, and then fill the blanks with a mathematical interpolation method. None of them is capable to do both works simultaneously. This paper presents a novel technique based on Proper Orthogonal Decomposition (POD) to detect and correct velocity field simultaneously, named as POD-OC method. It offers an efficient and robust algorithm of picking the optimal number of POD modes for reconstructing a reference velocity field for both outlier detection and correction. It has been easily extended to one instantaneous velocity field and 3D velocity field. It is proved by facts that this method is feasible and stable.

POD-based outlier detection and correction

POD is an approach for data analysis aiming to approximately describe a high/infinite-dimensional dynamical problem using a low-dimensional model. It can extract large-scale coherent structure with energy-contained modes. Thus, POD can approximately reconstruct the original pure flow field, this is the fundamental idea of POD-OC. This method uses dynamic modes denoted by \( n \), to reconstruct reference flow field for identifying and replacing outliers. This key parameter is calculated according to correlation analysis. After detecting spurious vectors through absolute error between original and reconstructed field; a process for outlier correction is needed to fill the blanks. In current work, both outlier detection and correction are performed in each iteration, and the corrected velocity directly calculated from the optimal reconstructed velocity and a residual which is the median value of eight neighboring points at the \( i \)th grid if a 3×3 local scanning window is applied.

Compared with normalized median test method (Westerweel and Scaran 2005), POD-OC is more efficient for clustered outliers. It is because that the normalized median test can’t find correct references from neighboring vectors. Differently from the local statistical algorithm as normalized median test, POD-OC method is a global iterative approach that reconstructs field from the entire dataset based on POD-modes, which can avoid the local influence of clustered outliers in each instantaneous field.

Block POD-OC: an extended method for a single flow field

POD-OC method has a notable shortcoming that the mode decomposition statistically requires a large amount of velocity fields. To solve this issue, a novel way called block POD-OC is devised to extend the POD-OC application to be suitable for a single flow field. The idea of block POD-OC is to divide the single field into multiple blocks of sub-fields for POD. To balance both sides, the number and size of sub-fields, the ratio of the sub-field size to the original full-field size defined as \( R \) was tested on instantaneous fields of the wake flow. When \( R \) is in the range of [0.2, 0.5], the relative error of u-component is stable and minimal. Therefore, the optimized value of \( R \) is better be in the range of 0.2–0.5. An experimental instantaneous field with real PIV-processed outliers was tested and compared with all-in-one technique (Garcia 2011) to reveal that the block POD-OC is feasible and reliable.

Conclusions

A new POD-based outlier correction for PIV data has been presented in this paper. Two types of spurious vector, scattered vectors and clustered vectors, have been analyzed on their statistical characteristics, which has been found to satisfy a Gaussian distribution on probability. Artificial outliers have been generated in a wake flow field behind a cylinder to test the POD-OC method. It is found that outliers could contaminate the POD modes, but only on high-order modes, which suggests that low-order high-energy modes still can approximate the original field. This is the fundamental idea of POD-OC method. Using certain number of POD modes to reconstruct a reference flow field can be used not only for outlier detection, but also for outlier repairing. The number of POD modes \( n_{p} \) for the reference field is a critical parameter, which is dynamically calculated from the latest mode decomposition from the corrected field during the iteration. The self-adaptive parameter \( n_{p} \) is also applicable for gappy POD. Against normal local OC algorithms, POD-OC approach is more reliable for detecting clustered outliers, which can also provide more accurate interpolation of missing vectors comparing other common mathematical tools.

In addition, a block POD-OC algorithm is proposed for post-processing on a single instantaneous velocity field, since regular POD decomposition needs large sample of data to extract the energy-contained modes. In the block POD-OC method, sub-fields divided from the full field are used for mode decomposition. The size ratio \( R \) is found insensitive to the performance of OC in range of 0.2 to 0.5. The block POD-OC can keep most information from the original field, while only fix the outliers. It is better than all-in-one method on OC, which will introduce non-physical noise into the flow by smoothing.

References