Matching pursuit with POD modes dictionaries in the analysis of 2D turbulence signals and images

Ch.H. Bruneau, P. Fischer, Z. Peter*, A. Yger

Institut de Mathématiques de Bordeaux, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence, France

Abstract

This paper is devoted to decomposition-reconstruction methods in the analysis of turbulence data, issued from a Direct Numerical Simulation of the two-dimensional Navier-Stokes equations. The Proper Orthogonal Decomposition (POD) method, based on the Hilbert-Schmidt theory, will be combined with the matching pursuit algorithm in order to localize the turbulent flow patterns that are coherents with a particular dictionary.

Keywords: 2D turbulence; Matching pursuit algorithm; Proper orthogonal decomposition

1. Introduction

The matching pursuit (MP) algorithm, introduced by Mallat [1], allows a clever decomposition of a given signal s into a linear combination of functions (also called atoms), which are selected from a redundant dictionary of signals with normalized energy equal to 1. These atoms are selected in order to best fit the structure of the signal. The first selected atom d_1 is chosen so that the modulus $|<s, d_1>|$ of its correlation coefficient with s is maximal; then d_2 is chosen within the dictionary so that

$$|\langle s - \langle s, d_1 \rangle \, d_1, \ d_2 \rangle| \tag{1}$$

is maximal, and so on. The sequence of coefficients is decreasing and indicates the order in which the corresponding atoms are selected. In the case of orthogonal bases, the number of iterations equals the number of atoms needed for a required ratio of reconstruction as a new atom is selected at each iteration.

2. The MP algorithm with POD modes

The POD, also called Karhunen-Loeve decomposition [2], is a classical method developed in statistics. Given a random process U, the overall algorithm can be summarized as follows:

1. Compute the autocorrelation matrix A of a set of realizations (also called 'snapshots') of U, $U_1, \ldots, U_q$.
2. Perform the Singular Value Decomposition of A, and thus organize the eigenvalues of A (the singular values) in decreasing order: $\lambda_1 \geq \lambda_2 \geq \ldots$.
3. Take $m \leq q$ and select an orthonormal system (in the L^2 sense) of vectors (α_{ij}, $1 \leq i \leq q$, $j = 1, \ldots, m$, such that $(\alpha_{ij})_i$ is an eigenvector with respect to the eigenvalue λ_j.
4. Compute the POD modes

$$V_j := \sum_{i=1}^{q} \alpha_{ij} U_i, \quad j = 1, \ldots, m$$

When $m = q$ this four-step process provides the best (orthogonal) basis for the set of realizations $\{U_1, \ldots, U_q\}$ with respect to the L^2 norm.

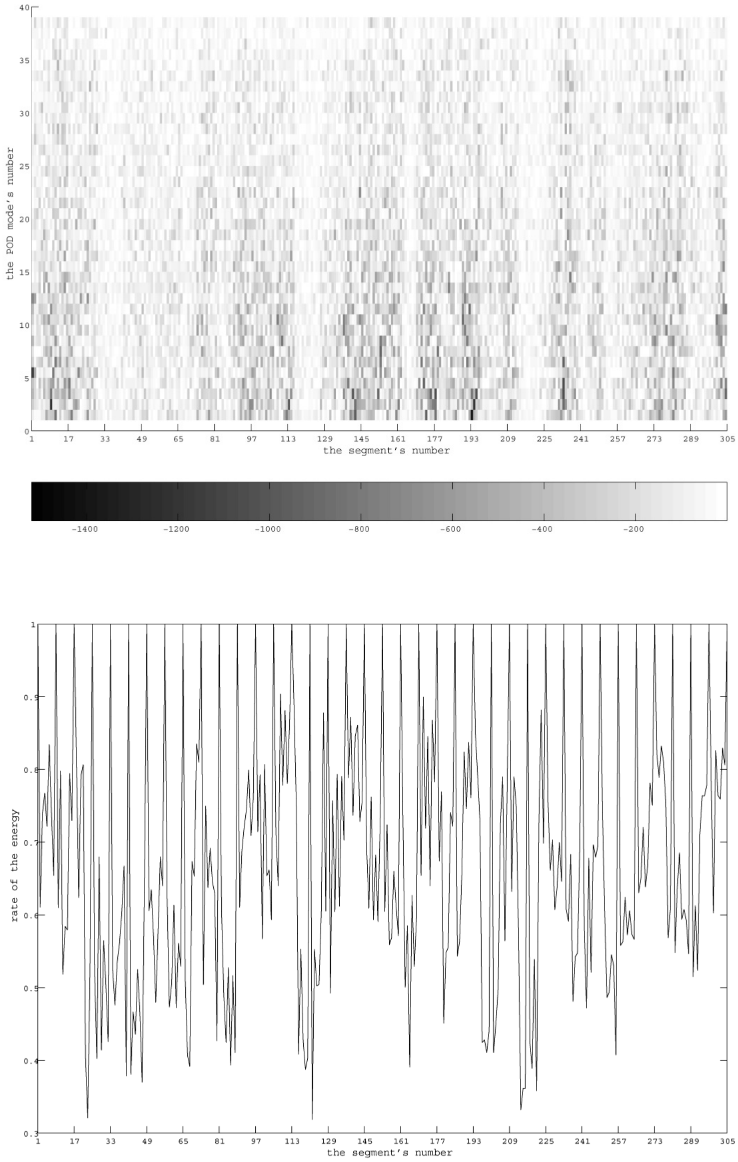
2.1. Application to one-dimensional signals

Let us consider a signal s (of length 40 000), corresponding to the recording of the vorticity at one point along time. We divide $s_1 = s(1 : 39936)$ into 39 consecutive non-overlapping segments of length 1024. Each segment plays the role of a snapshot, thus leading to a dictionary of 39 snapshots and to an autocorrelation matrix of size 39×39. The POD algorithm then provides 39 POD modes of length 1024 [3].

The MP algorithm implemented as in [4] is applied to a family of 305 segments of length 1024, which are obtained by translation of 128 points from s_1. In Fig.

* Corresponding author. Tel.: +33 (5) 40 00 2175; Fax: +33 (5) 40 00 2626; E-mail: peter@math.u-bordeaux1.fr

Computational Fluid and Solid Mechanics 2005
K.J. Bathe (Editor)



(a)

(b)

Fig. 1. Influence of the position of the segments on the MP decomposition with the POD modes dictionary. (a) MP coefficient intensity; (b) reconstruction rate.

1(a) are plotted the absolute values of the MP coefficients (from dark gray for the biggest ones to light gray for the smallest ones) obtained when decomposing each of the 305 segments. The clear vertical strips show the intermittencies of the signal. In addition, Fig. 1(b) shows that in these strips, taking away the 39 snapshots, the reconstruction rate (always using the whole set of POD modes) is much lower.

2.2. Application to two-dimensional images

Starting from 200 snapshots, obtained for successive time steps, one can calculate the 200 POD modes, by using the same method as above. We represent in Fig. 3 the reconstructions of the original field (Fig. 2), with 20 POD modes as follows: using the first 20 (most energetic) modes, provided by the decreasing sequence of the singular values (Fig. 3(a)) and with the most significant 20 modes, provided by the first 20 iterations of the matching pursuit (Fig. 3(b)). One can see that the

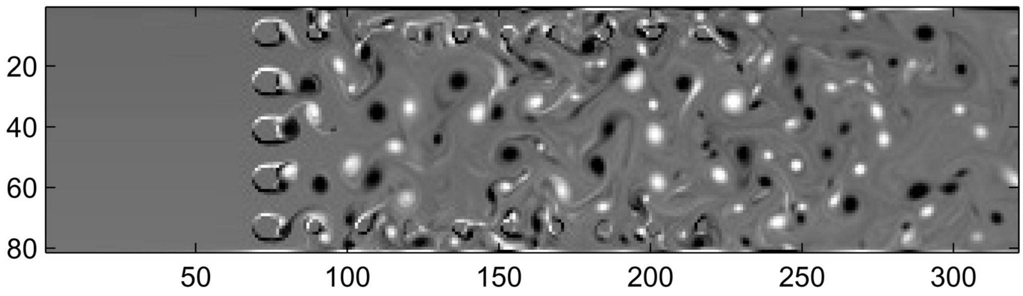


Fig. 2. The vorticity image (also the first snapshot).

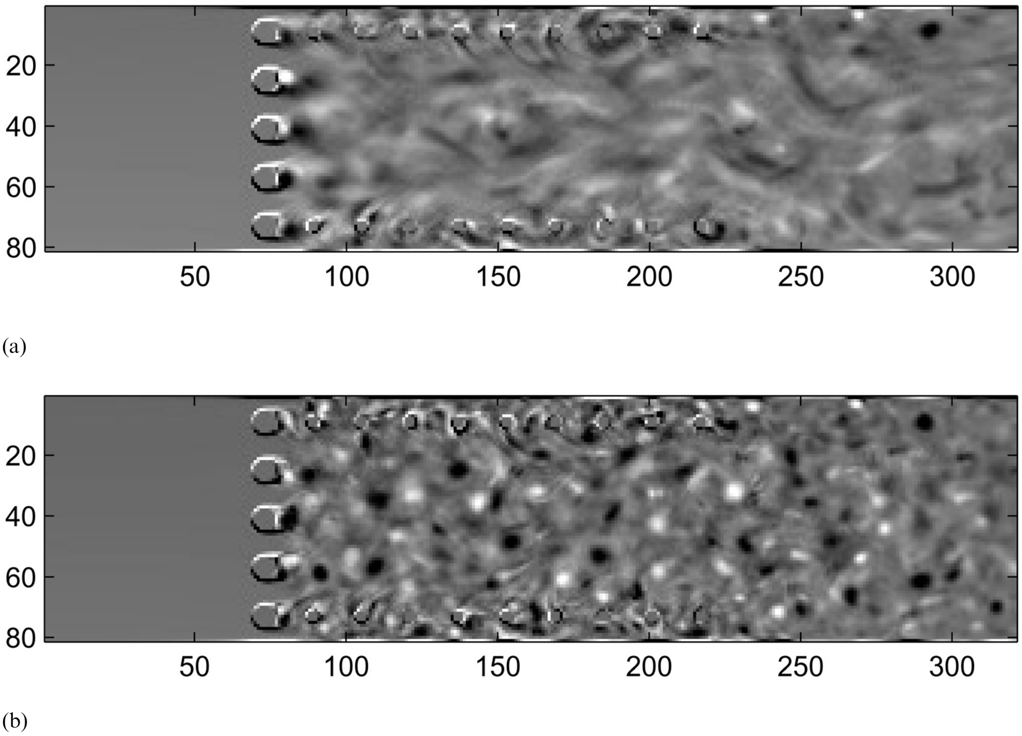


Fig. 3. The first snapshot of the vorticity rebuilt with 10% of the total number of POD modes: (a) with the first 20 POD modes; (b) with the 20 POD modes chosen by MP.

reconstruction is closer to the original image in the second figure (which keeps 65.67% of the Frobenius norm) than in the first figure (rebuilding only 27.76% of the energy) as the coherent structures are better captured. The field of Fig. 3(b), obtained by the MP algorithm, has more vortex cores located in the correct place than the field of the Fig. 3(a), where only a few coherent structures are visible. Using more modes it is possible in both cases to reconstruct completely the original snapshot.

One can wonder, when a POD mode basis is built using a subset of the snapshots, if it is possible to reconstruct one of the snapshots that are not used to construct the basis. The answer is no, as the rate of the reconstructed energy is less than 70%, even with all the available POD modes.

3. Conclusions

The implicit order of the POD modes is given by the decreasing sequence of the eigenvalues, that are related to the energy content of these modes. Thus the POD method orders the modes after their average importance for each snapshot. Therefore, it is possible that for a particular snapshot not the most energetic POD modes are the best adapted to obtain a predefined reconstruction rate. The MP algorithm is able to find these most significant modes.

The combination of the above methods appears to be efficient for analyzing one of the snapshots, but turns out to be less adapted for a segment randomly chosen in the signal or for a different image.

References

[1] Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Process 1993; 41(12): 3397–3415.

[2] Loeve M. Probability Theory. Princeton, NJ: Van Nostrand, 1955.

[3] Bruneau CH, Fischer P, Peter Z, Yger A. Comparison of numerical methods for the computation of energy spectra in 2D turbulence. Submitted to Sampling Theory in Signal and Image Processing.

[4] Yger A. Théorie et analyse du signal. Cours et initiation pratique via MATLAB et SCILAB. Paris: Editions Ellipses, 2000.