Post-processing of PIV records to allow derivative computation.

J.M. FOUCAUT, J. CARLIER, M. STANISLAS
Ecole Centrale de Lille
LML URA 1441
BP 48
Cité Scientifique
F59651 Villeneuve d'Ascq Cedex
Tel: 33-(0)3-20-33-71-70
Fax: 33-(0)3-20-33-71-69
Email : jean-marc.foucaut@ec-lille.fr

Abstract : Particle Image Velocimetry provides raw data which needs post-processing to improve their quality and to allow derivative computation. This process includes: the spurious vector detection, the correction or removal of these vectors and the filling of the velocity maps in order to avoid holes. The recent developments of CCD have allowed to record a lot of images during a PIV test campaign. These images lead to a large number of velocity maps that have more or less spurious vectors depending on the quality of the records. A manual post-processing is then impossible. This paper proposes an automatic procedure to remove the spurious data based on a comparison between the raw vector and an "estimation". This estimator is a median filter applied to the surrounding velocity field. This procedure only requires an a priori estimation of the percentage of validated vectors in the map. After this procedure, each map is left with a certain number of holes. An adaptive and iterative interpolation method is then used, based on a 5 x 5 window interpolation filter. The maximum order of the filter is 7. The interpolated map obtained can then be used to compute the velocity gradients.

1. Data validation procedure.

As was extensively demonstrated in the Europiv program (Stanislas et al (2000)), raw PIV data coming out from the correlation procedure contains a certain amount of spurious vectors. This amount depends on several parameters of the recording procedure such as particle concentration, contrast, lens quality, in plane and out of plane loss of particles...

In order to use the PIV data for modelling purposes, it is important to remove these spurious vectors and, if velocity derivatives are needed, to replace them by the best possible estimation.

Due to the extensive use of CCD sensors in PIV, the number of records per test case has increased dramatically in the last few years. This imposes to develop automatic and reliable procedures for both outliers removal and interpolation.

The detection of spurious vectors has been addressed in the early development of the method by Keane & Adrian (1992) based on the quality and detectability of the peak in the correlogram and by Willert & Garib (1991) using a comparison between each vector and the nearby vectors. More recently Westerweel (1994) did suggest a method to estimate the velocity at a point by the median value of the neighbouring vectors. He showed that this estimation is better than the mean value which can be biased by the surrounding false vectors taken into account in the average. Raffel et al (1992) did propose an adaptive method, based on the same local detection which tries to take into account discontinuities in the velocity field by disconnecting the coherent domains. The coherence level is provided by a local variance computed on the nearby value. They used this approach on flows with strong gradients (like shock waves), in order to minimize the filtering effect on the discontinuity. The data validation is then made by a thresholding on the difference between the magnitude of the vector and its estimation. The threshold is a linear function of the local variance. Some other contributions like Raffel et al (1998) and Nogueira et al (1997) can be found in the literature concerning the detection of spurious vectors based on the same local concept.

In the case were a lot of spurious vectors are localized in the same region, a strong effect on the estimated value can be observed which can lead to a bias. If the spurious vectors are sparcely and randomly distributed in the field there are less problem to detect them. Thus, the first important step is to optimize the recording procedure in order to minimize the number of spurious vectors (Westerleel (1994)).

In the present contribution, an iterative method is proposed for data validation. This method is based on a comparison between the velocity at each point and its estimation using a median filter similar to the one proposed by Westerweel (1994). During the analysis of the correlograms, the three highest peaks are recorded and ordered in decreasing amplitude. A region of interest is used to search these peaks, based on the global histogram method (Raffel et al (1998)). The validation is then made in two steps:

- In a first step, the three peaks are ordered with respect to Δu, the modulus of the difference between the velocity given by the peak and that of the "estimator". This is done iteratively until no change is observed in this order on the whole map.
- In a second step, this modulus Δu, given by the best peak, is thresholded. For all values above a threshold ε, the vectors are removed and replaced by holes. This second step is also iterative in order to minimize the effect of the spurious vectors on the "estimator".

Both steps are executed until in a global iterative procedure, there is no change in the map.

To take into account eventual mean velocity gradients, before threholding, Δu is normalized by a reference velocity u. This reference velocity depends on the flow under study. If the flow is homogeneous in the two coordinate directions of the velocity map, u is just the mean velocity of the whole map. If the flow presents only one homogeneity direction, then u is averaged in that direction and is a function of the other coordinate. The most difficult case is the flow with a priori no direction of homogeneity, then either u can be defined from the a priori knowledge of the flow or the normalisation is useless.

The main difficulty of the second step is the choice of the threshold. In this study a method is proposed to determine ε. This value is directly deduced from the cumulative histogram of Δu / u. To do so, an estimation of the percentage Pv of valid vectors is done by hand in a few maps. Pv is easily obtained by looking at a few 10 by 10 windows in different regions of one map or a few and by counting the number of spurious vectors. This criterion is more objective than manual and iterative estimation of the threshold. Figure 1 gives an example of cumulative histogram in the case of a pure 2D sine wave velocity field. The histogram has been computed first without spurious vectors and second with 5 % of spurious vectors replacing randomly good vectors. In the first case, the histogram goes to 100 % very rapidly. Ideally it should be 100 % everywhere. The difference is due to the filtering effect of the estimator. In the second case, the effect of the spurious vectors is clearly evidenced by a slope change around P = 95 %. In this case, the best value of ε is 0.1.

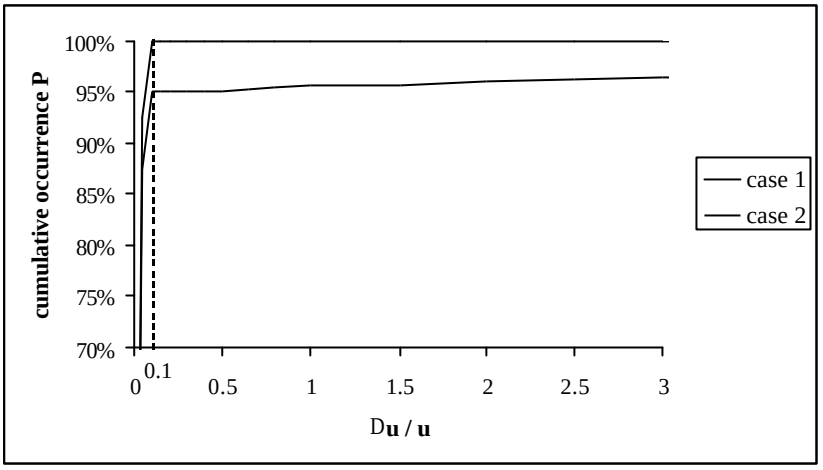

Figure 1: Cumulative histogram of Δu / u for a 2D sine wave.

Figure 2 gives an example of the cumulative histogram obtained from two analysis of the same field. This field, corresponding to a real turbulent flow, has been first analysed with 64 x 64 windows using a global shift and a gaussian peak fitting (Raffel et al (1998)). It has been also analysed with 24 x 24 window using a local shift deduced from the previous results. For the 64 x 64 window analysis Pv was estimated at 98 % by looking at five regions of 10 x 10 vectors. This represents about 10 % of the 5600 vectors of the full map. Using the cumulative histogram computed on the raw data (continuous line on figure 2) the threshold can be estimated at $\varepsilon = 14$. For the 24 x 24 window analysis 10 regions were observed (the total number of vectors was of the order of 40 000) giving Pv = 95 %. The histogram of the raw data is in dashed line in figure 2. The threshold is then $\varepsilon = 16$. As shown in figure 2, the shape of the histogram is smoother in the real case as compared to the model of figure 1. Nevertheless, the slope around the threshold shows a low sensitivity to the value of ε.

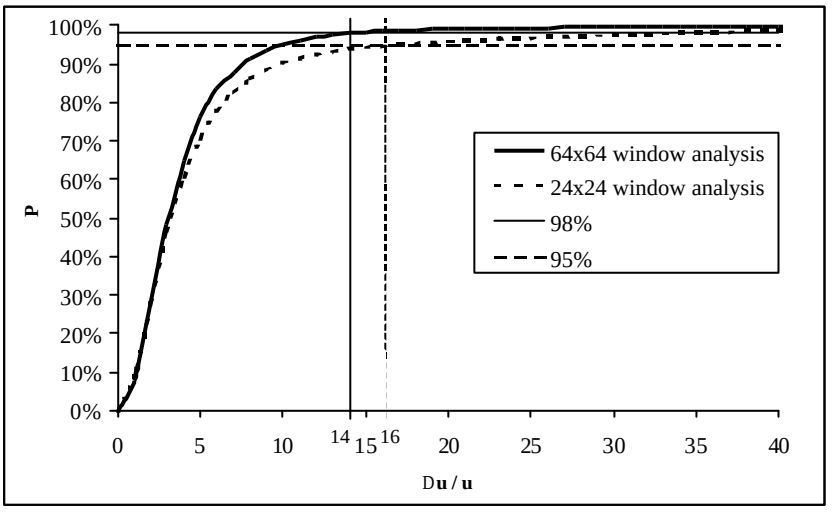

Figure 2: Cumulative histogram of Δu / u for a real turbulent flow.

This method allows to obtain a good detection of the spurious vectors with a minimization of the number of holes. Figure 3 gives an example of a raw velocity map (highest correlation peak) where spurious vectors can be easily identified. In figure 4 the same data are given after the application of the whole cleaning procedure. The threshold of 14 obtained by the cumulative histogram method was used (64 x 64 window analysis in figure 2). As can be seen, the flow is fairly different in the lower right corner, due to the use of higher order peaks. On the left part of the map, the outliers are replaced and on the right part several holes appear.

When the outliers have been removed, the velocity map contains some holes which size and position are more or less randomly distributed. These holes have to be interpolated in order to allow computation of the gradient fields.

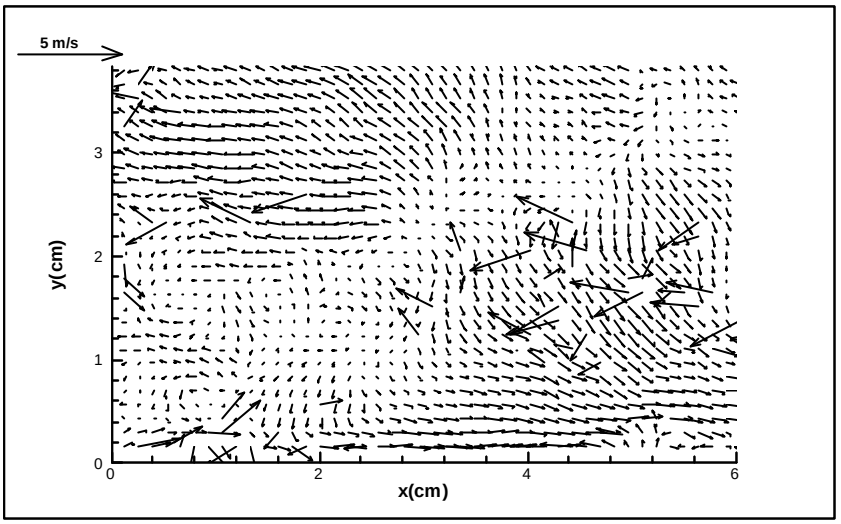

Figure 3: raw velocity map.

2. The interpolation method.

The aim of the present part is to propose a method to fill the holes by the use of an interpolation window. Nogueira et al (1997) have proposed an interpolation filtering by considering that the false vectors are high frequency signals. Their method uses low pass filters which modify each value of the field. In the present study, only the holes are interpolated. Given a velocity field $u_{i,j}$, the value corresponding to a hole is computed with eq. (1):

$$u_{i,j} = \sum_{\substack{l=-2,+2 \\ k=-2,+2}} F_{l,k}\, u_{i+l,j+k} \qquad (1)$$

where $F_{l,k}$ is a 5 by 5 interpolation matrix.

Table 1 gives the values of $F_{l,k}$ for the 7^{th} order interpolation filter.

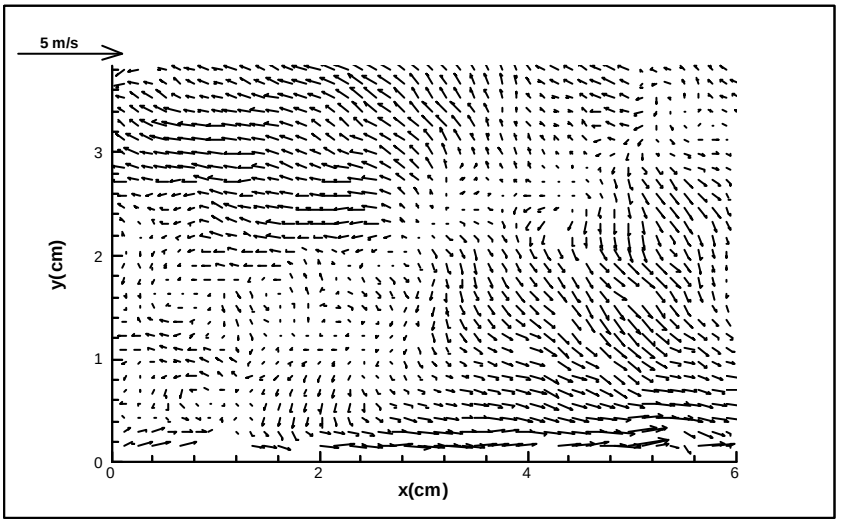

Figure 4: cleaned velocity field (raw data in figure 3).

l \ k	-2	-1	0	1	2
-2	-1/36	1/9	-1/6	1/9	-1/36
-1	1/9	-4/9	2/3	-4/9	1/9

0	-1/6	2/3	0	2/3	-1/6
1	1/9	-4/9	2/3	-4/9	1/9
2	-1/36	1/9	-1/6	1/9	-1/36

Table 1 : 7th order interpolation window.

A difficulty appears when there are other holes in the 24 values used for the interpolation. In this case, an adaptive and iterative interpolation is proposed. The interpolation matrix is deduced from a limited expansion around the value to interpolate which can be written:

$$u_{i,j} = u_{i+l,j+k} + \sum_{\substack{m=1,7 \\ n=1,7 \\ m+n<8}} \frac{dx^m dy^n}{m!\,n!} \frac{\partial^{m+n} u}{\partial x^m \partial y^n} \tag{2}$$

Equation (2) coupled with equation (1) gives:

$$u_{i,j} = \sum_{\substack{l=-2,2 \\ k=-2,2}} F_{l,k}\, u_{i+l,j+k} + \sum_{\substack{l=-2,2 \\ k=-2,2}} \sum_{\substack{m=1,7 \\ n=1,7 \\ m+n<8}} F_{l,k} \frac{\Delta x^m \Delta y^n l^m k^m}{m!\,n!} \frac{\partial^{m+n} u}{\partial x^m \partial y^n} \tag{3}$$

$F_{2,2}$	$F_{1,2}$	$F_{0,2}$	$F_{1,2}$	$F_{2,2}$
$F_{1,2}$	$F_{1,1}$	$F_{0,1}$	$F_{1,1}$	$F_{1,2}$
$F_{0,2}$	$F_{0,1}$	$F_{0,0}$	$F_{0,1}$	$F_{0,2}$
$F_{1,2}$	$F_{1,1}$	$F_{0,1}$	$F_{1,1}$	$F_{1,2}$
$F_{2,2}$	$F_{1,2}$	$F_{0,2}$	$F_{1,2}$	$F_{2,2}$

Table 2 : interpolation window taking into account symmetries.

To obtain the highest possible order the double sum in equation (3) has to be completely cancelled. For a 5 x 5 matrix, the maximum order which can be obtained is max(m+n) = 7. The complete expression of the interpolation equation, taking into account the symmetries appearing when m or n are odd (table 2) leads to the following system:

$$\begin{array}{r} 1 \\ 3 \\ \text{order}: 3 \\ 5 \\ 7 \end{array} \begin{bmatrix} 1 & 1 & 1 & 2 & 1 \\ 1 & 4 & 2 & 10 & 8 \\ 1 & 16 & 2 & 34 & 32 \\ 0 & 0 & 2 & 16 & 32 \\ 0 & 0 & 2 & 40 & 128 \end{bmatrix} \begin{pmatrix} F_{0,1} \\ F_{0,2} \\ F_{1,1} \\ F_{1,2} \\ F_{2,2} \end{pmatrix} = \begin{pmatrix} 1/4 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{4}$$

The first line of the system corresponds to the first sum of equation (3) and the other lines to the double sums of this equation. The solution of this system gives the 7th order interpolation filter presented in table 1.
Each line of equation (4) corresponds to an order of the filter. When there is a hole around $u_{i,j}$ the corresponding $F_{l,k}$ is cancelled and the last line of (4) is replaced by four 0 and a 1 for the corresponding column. The order is thus decreased.
For example, if the hole to interpolate is in the second column of the map, the values of $F_{l,2}$ is zero for all l. System (4) leads to the following:

$$\begin{array}{r} 1 \\ 3 \\ \text{order}: \\ \\ \end{array} \begin{bmatrix} 1 & 1 & 1 & 2 & 1 \\ 1 & 4 & 2 & 10 & 8 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} F_{0,1} \\ F_{0,2} \\ F_{1,1} \\ F_{1,2} \\ F_{2,2} \end{pmatrix} = \begin{pmatrix} 1/4 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{5}$$

which corresponds to a 3rd order interpolation matrix (table 3).

0	0	0	0	0
0	-1/4	1/2	-1/4	0
0	1/2	0	1/2	0
0	-1/4	1/2	-1/4	0
0	0	0	0	0

Table 3 : 3rd order interpolation matrix close to the border.

In the proposed procedure, when there is a hole in the velocity map, other holes are looked for in the 5 x 5 interpolation window. For each hole found a line from the bottom of the matrix is modified. The system is then solved to obtain the adapted interpolation filter. If the central hole is close to too many others (all $F_{l,k}$ have to be cancelled), it is not interpolated. The minimum interpolation order is thus 1. On the border of the map or of regions with too many holes, an extrapolation, based on the same procedure, is also used with a 4 x 5 matrix which maximum order is 5. After one pass of interpolation on the map, some holes are eliminated. A new iteration is then computed to improve the order of interpolation (and thus the accuracy) and to decrease the number of holes.
It can be noted from tables 1 and 3 that the coefficient $F_{0,0}$ of $u_{i,j}$ is zero. In order to minimize the random error due to the measurement noise, it is possible to add an equation taking into account the value of $u_{i,j}$. This can be done by taking $F_{0,0} = a$ and by replacing $F_{l,k}$ of equation (1) by $(1-a)F_{l,k}$. a is a coefficient which depends on the order of the interpolation filter. The measurement noise on the interpolated value is then given by $\sigma^2 = \left(a^2 + (1-a)^2 \sum_{\substack{l=-2,2 \\ k=-2,2}} F_{l,k}^2 \right) \sigma_u^2$ where σ_u is the measurement noise on the input data. The value of a which minimizes the noise amplification is $a = \sum_{\substack{l=-2,2 \\ k=-2,2}} F_{l,k}^2 \Big/ \left(\sum_{\substack{l=-2,2 \\ k=-2,2}} F_{l,k}^2 + 1 \right)$.

For the 7th order filter, the interpolation matrix is then the same as the one proposed by Nogueira et al (1997) (table 4). That correction can be applied whatever the order of interpolation is. Iterations are applied until the order of the filter does not increase, which takes about 10 iterations. The iteration procedure is illustrated in figure 5 for two consecutive holes, in the case of a sinusoidal distribution of velocity in both directions. Figure 6 presents the velocity map of figure 4 after the interpolation procedure has been applied. The contours represent the interpolation order. As can be seen, the order of the last filter is 7 in most of the field and goes down to 3 close to the border where interpolation is more difficult. After this procedure the number of holes starting from 5 % comes down to less than 0.5 %.

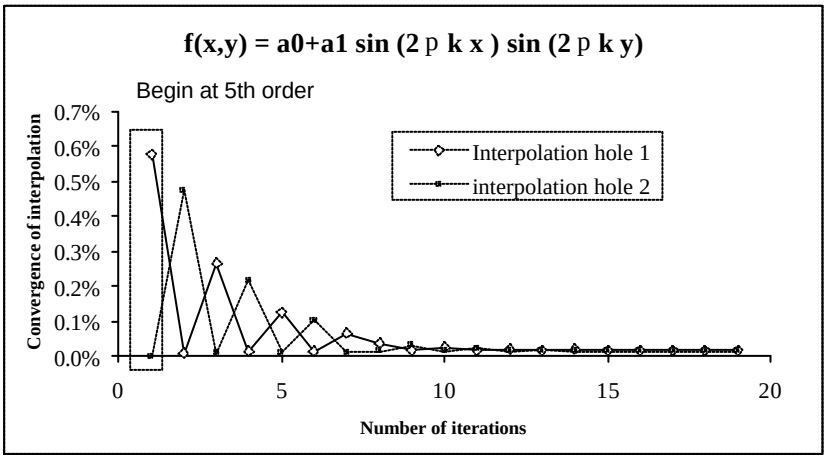

Figure 5: convergence of 7th order interpolation for 2 neighbouring holes.

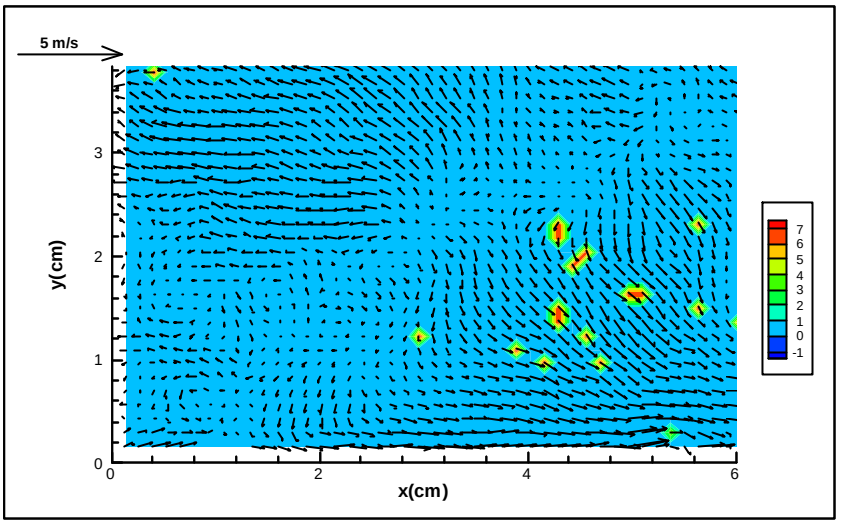

Figure 6: interpolated velocity field (data from figure 3), contours correspond to the order of interpolation.

-1/136	1/34	-3/68	1/34	-1/136
1/34	-2/17	3/17	-2/17	1/34
-3/68	3/17	100/136	3/17	-3/68
1/34	-2/17	3/17	-2/17	1/34
-1/136	1/34	-3/68	1/34	-1/136

Table 4 : 7th order interpolation window taking into account the noise.

When a value is interpolated, an uncertainty appears which can be estimated by the truncation error and the measurement noise more or less amplified by the filter. Figure 7 give a comparison of the 3rd and 7th order interpolation filter error for an isotropic 2D sine wave, k is the wave number in both direction. The truncation errors are respectively:

$$\Delta u = \frac{1}{36}\left(\Delta x^4 \Delta y^4 \frac{\partial^8 u}{\partial x^4 \partial y^4}\right) \tag{6}$$

and

$$\Delta u = \frac{1}{12}\left(\Delta x^4 \frac{\partial^4 u}{\partial x^4} + \Delta y^4 \frac{\partial^4 u}{\partial y^4}\right) \tag{7}$$

for the 7th and 3rd order interpolation filter. As can be seen the 7th order interpolation presents a better frequency response. For the low frequency, only noise errors of the order of 2% for both filter are detectable. When k Δx increases beyond the value of 0.1 for the 3rd order and 0.15 for the 7th order the truncation errors become greater. The obtained errors appear to grow rapidly with k. They can thus introduce bias on the high frequency part of the PIV data. Thus, it is important to limit the number of holes to a minimum to assess the turbulent statistics.

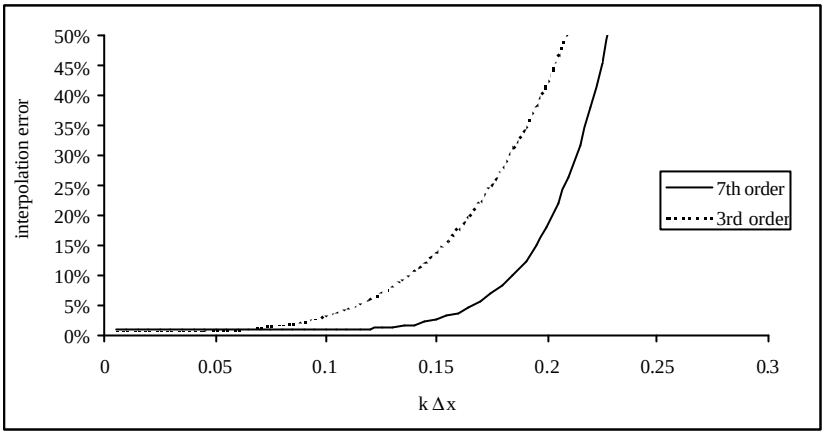

Figure 7: error due to the interpolation on a 2D sine wave comparison between 3rd and 7th order interpolation.

Figure 8 presents a comparison of the cumulative histogram of Δu / u for raw, cleaned and interpolated data. These results correspond to the 64 x 64 window analysis of figure 2. In that case, the valid vectors were estimated a priori at Pv = 98 %, giving a threshold of ε = 14. The dashed line gives the same raw data histogram as in figure 2. The thin continuous line presents the same histogram after removal of spurious vectors. As can be seen, above Δu / u =14, the histogram saturates close to Pv. This is due to the holes in the map. Before Δu / u =14, the histogram is nearly unchanged. The thick continuous line is the histogram after the interpolation procedure. The maximum level is very near to 100 % showing that most of the holes have been interpolated.
Finally, figure 9 shows the interpolation error if all the vectors of a good quality map (no spurious vector) would be replaced by their interpolated values. This is done on a real velocity field and only for the u component. The flow is turbulent with a turbulence level slowly increasing from 5 % to 15 % with x. As can be seen the error increases with the turbulent level from less than 1 % to about 3 %. The errors close to the border increase also, but due to the use of a 3rd order filter. It is there around 5 %.

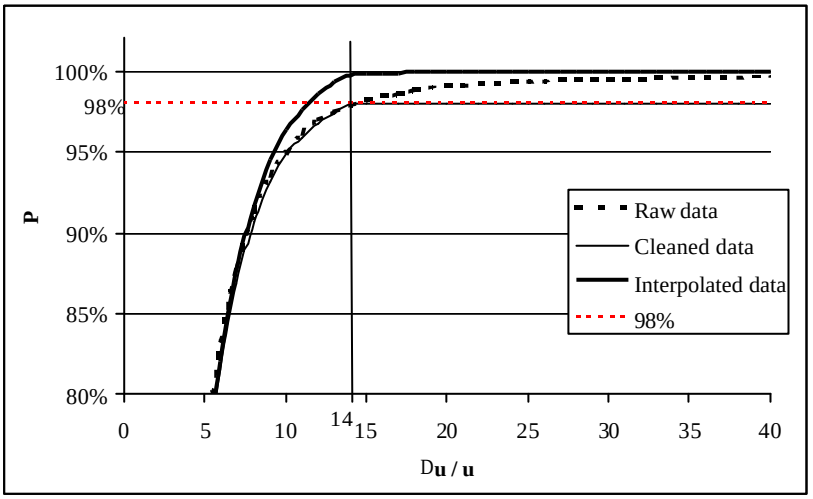

Figure 8: Cumulative histogram of Δu / u, comparison of raw, cleaned and interpolated data.

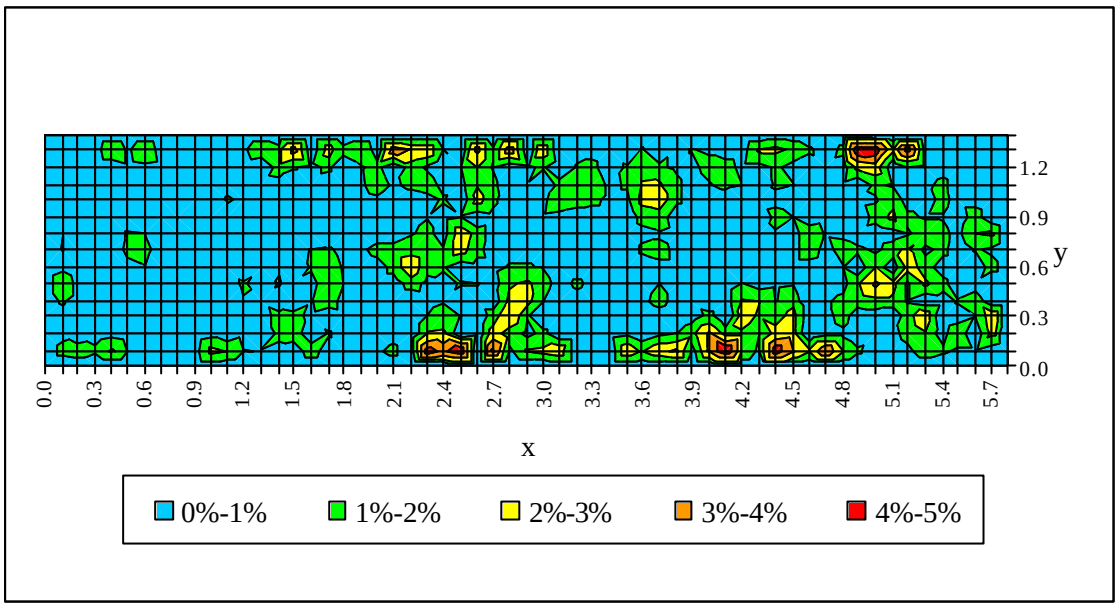

Figure 9: error due to the interpolation on a real u velocity field of a turbulent flow.

In this contribution, a method of data validation in PIV maps has been proposed. That method is based on a comparison between a vector and an estimator deduced from the nearby data (Westerweel (1994)). An objective criterion obtained from the cumulative histogram of the error and a manual estimation of the number of valid vectors in the maps allows to remove the spurious vectors and to replace them by holes. A look at the synthetic results of figure 1 indicates that it should be possible to define automatically the criterion.
A second part of this paper concern the interpolation of the data in the holes. An adaptive and iterative method is proposed which allows to get the highest order of interpolation whatever the distribution of holes is in the map. This filter uses a 5 x 5 interpolation window and is optimised with regards to the noise. Its maximum order is 7^{th}. A study of accuracy corresponding to the case of a single hole has been done. In the future, the case of two or more consecutive holes should be studied for synthetic and real flow.
The proposed method of post-proceeding brings an improvement on the smoothness of the velocity field with an optimization of the interpolation accuracy. It does not modify the valid vectors and gives an estimation of the accuracy of the interpolated values. After the complete procedure the derivative computation can be done.

This work has been performed under the EUROPIV2 project. EUROPIV2 (A joint program to improve PIV performance for industry and research) is a collaboration between LML URA CNRS 1441, DASSAULT AVIATION, DASA, ITAP, CIRA, DLR, ISL, NLR, ONERA and the universities of Delft, Madrid, Oldenburg, Rome, Rouen (CORIA URA CNRS 230), St Etienne (TSI URA CNRS 842), Zaragoza.

The project is managed by LML URA CNRS 1441 and is funded by the European Union within the 5^{th} frame work (Contract N°: G4RD-CT-2000-00190).

Reference.

Keane R D & Adrian R J (1992), Theory of cross-correlation analysis of PIV images, Applied Scientific Research, vol. 49, pp 191-215.

Nogueira J et al (1997), Data validation, false vectors correction and derived magnitudes calculation on PIV data, Meas. Sci. Technol., vol. 8, pp 1493-1501.

Raffel M et al (1992), Data validation for Particle Image Velocimetry, 6^{th} Int. Symposium on Application of Laser Techniques to Fluid Mechanics, Lisbon.

Raffel M et al (1998), Particle Image Velocimetry, A practical guide, Springer.

Raffel M & Kompenhans J (1996), Post processing: data validation, Von Karman Institute, Lecture Series, 1996-03.

Stanislas M et al (2000), Particle Image Velocimetry: progress toward industrial application, to appear March 2000, Kluwer Academic Press.

Westerwell J (1994), Efficient detection of spurious vectors in particle image velocimetry data, Experiments in Fluids, vol. 16, pp 236-247.

Willert C E & Garib M (1991), Digital particle image velocimetry, Experiments in Fluids, vol. 10, pp 181-193.