Article:

Image-based tracking technique assessment and application to a fluid-structure interaction experiment

D.A. Mella[1], W. Brevis[2], J.E. Higham[3], V. Racic[4] and L.Susmel[1]

Abstract

This work analyses the accuracy and capabilities of two image-based tracking techniques related to Digital Image Correlation and the Lucas-Kanade optical flow method, with the subsequent quantification of body motion in a fluid-structure interaction experiment. A computer-controlled shaker was used as a benchmark case to create a one-dimensional oscillatory target motion. Three target frequencies were recorded. The measurements obtained with a low-cost digital camera were compared to a high precision motion tracking system. The comparison was performed under changes in image resolution, target motion, and sampling frequency. The results show that, with a correct selection of the processing parameters, both tracking techniques were able to track the main motion and frequency of the target even after a reduction of four and five times the sampling frequency and image resolution respectively. Within this good agreement, the Lucas-Kanade technique shows better accuracy under tested conditions, achieving up to 15.6% of lower tracking error. Nevertheless, the achievement of this higher accuracy is highly dependent on the position of the selected initial target point. These considerations are addressed to satisfactorily track the response of a wall-mounted cylinder subjected to a range of turbulent flows using a single camera as the measuring device.

Keywords

Accuracy assessment, Digital Image Correlation, Image-based tracking, Fluid-structure interaction, Lucas-Kanade technique

1 Introduction

The introduction of new technological advances associated to imaging devices has allowed the development of high precision and low-cost image-based techniques. In structural engineering, these techniques offer an alternative to traditional approaches such as accelerometers, strain gages, laser doppler vibrometer, among others. Their main advantages are related to the remote and non-contact nature, which can offer several benefits in cases of structures with difficult access or when simultaneous multi-point or full-field measurements are required. In general, image-based techniques use digital images to obtain displacements, geometry, or deformations of different targets [1]. In the context of engineering applications, these techniques have been used for the quantification of structural vibrations [2], damage detection [3], linear and non-linear deformations [4], crack growth during fatigue testing [5], and the characterisation of fluid-structure interactions [6, 7].

Two image-based techniques frequently used in engineering are based on Digital Image Correlation (DIC) and optical flow methods. Due to their extended application in engineering, this paper will focus on DIC calculated via standard Fast Fourier Transform (termed here as DIC-based technique), and the Lucas-Kanade approach (LK) [8], which is a linear approximation of the optical flow equations. Some applications of the DIC-based algorithm can be found in [9], where this technique was applied on a suspension bridge to estimate the modal frequencies and tension of hanger cables subjected

[1]Department of Civil and Structural Engineering, The University of Sheffield, Sheffield, UK
[2]Department of Mining Engineering and Department of Hydraulics and Environmental Engineering, Pontifical Catholic University of Chile, Santiago, Chile
[3]National Energy Technology Laboratories, Department of Energy, Morgantown, West Virginia, USA
[4]Department of Civil and Environmental Engineering, Politecnico di Milano, Milano, Italy

Corresponding author:
D.A. Mella, Department of Civil and Structural Engineering, The University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK.

Email: mvamellavivanco1@sheffield.ac.uk

to ambient vibration loads. The difference between their results and those obtained from simultaneous accelerometer measurements was within $\pm 0.5\%$. The response of a full-scale timber-framed structure subjected to synthetic seismic loads was successfully captured using DIC [10]. Their full-field measurements were validated with a Linear Variable Differential Transformer, obtaining differences of less than 5%. In the context of fluid-structure interaction, a stereo DIC-based algorithm was used to analyse differences in modal characteristics and vibrational responses of composite beams placed in different mediums [7]. Applications for the LK technique can be found in [11] for the estimation of the modal shapes and frequencies of a helicopter blade. The first two main modes of the structure were successfully extracted under an impulse excitation, showing good agreement with a finite element model and a classical experimental modal analysis. Amplitudes and frequencies of a targetless cable of a small pedestrian bridge has been measured using the LK technique [12]. The estimated natural frequencies were in close agreement with the spectral peak frequencies extracted from accelerometers. For the characterisation of a fluid-structure interaction problem, a computer vision technique based on the LK approach was used [13]. The seismic response of an underwater model subjected to three-dimensional synthetic seismic loads was successfully extracted.

Under controlled laboratory or field conditions, previous studies have validated the accuracy of the tracking techniques by comparing them to other external sensors. However, the selection of the image acquisition procedure and the parameters involved in the calculations often leads to an optimal local solution. In terms of the DIC-based technique, research on the quantification and optimisation of the variables involved in motion estimation (see, for instance, [14, 15]) indicate that there is an important dependence to the setup conditions and processing parameters. This consideration can be extended to the optical flow technique as well [16]. This work quantifies and compares the accuracy and robustness of both tracking techniques to the setup conditions and their corresponding processing parameters. Then, these results are used to measure the response of a body in a fluid-structure interaction experiment using a single recording device. For this, a high precision measurement technique and a computer controlled shaker are used to analyse the robustness and accuracy of the DIC-based and LK techniques. Their capabilities are evaluated by a comparison of the measured amplitude spectra and under changes in the tracking parameters, acquisition frequency, and image resolution. It is shown that, with proper

consideration of the experimental setup and processing parameters, both tracking techniques can satisfactorily track the target motion even after a reduction of four and five times the sampling frequency and image resolution respectively. The DIC-based technique is more robust under decrements in the image resolution. On the other hand, the LK technique is better in terms of accuracy, achieving up to 15.6% of lower tracking error in some cases. Nevertheless, this technique has an important dependency on the selection of the initial target point. This limitation is addressed through the implementation of the Forward-Backward (FB) tracking failure algorithm [17], where the response of a cylindrical obstacle subjected to different open-channel turbulent flows is successfully extracted using a single camera as the measuring device.

Nomenclature

d pixel displacement between frames. $d = [d_x, d_y]$

d_{ij} Components of the affine transformation. $i = [x, y]$, and $j = [x, y]$

f_s shaker frequency

f_s' acquisition frequency of the resampled shaker image series

f_{nat} cylinder frequency measured in air

r cross-correlation function

s_k signal $k = [1, 2]$ obtained from the tracking techniques

u initial target pixel coordinates. $u = [x_o, y_o]$

w interrogation window size. $w = [w_x, w_y]$

x pixel image coordinates. $x = [x, y]$

A affine transformation matrix

D cylinder diameter

I first image pixel intensity values

J second image pixel intensity values

Q flow rate

N x-direction image pixel size

M y-direction image pixel size

P_1 initial tracking point position starting from the first image series

$P_1^{'}$ final tracking point position starting from the last image series

P_2 final tracking point position starting from the first image series

$\mathbf{R}$ correlation map

R_e Reynolds number

RMS_n root mean square between the normalised tracking signal and CODA signals

T transpose

U bulk flow velocity

W interrogation window region

W_1 interrogation window of the first image

W_2 interrogation window of the second image

ΔRMS_n root mean square between the difference of two signals

ϵ Lucas-Kanade technique minimisation function

$\mathcal{F}$ Fourier transform

$\mathcal{F}^*$ complex conjugate of the Fourier transform

$\mathcal{F}^{-1}$ inverse Fourier transform

ν kinematic viscosity

2 Computational implementation of the tracking techniques

2.1 Digital Image Correlation

Digital image correlation has been widely used for image processing. Particularly, this work uses the cross-correlation function to track the location of a single target. Consider two consecutive frames of equal size $[N, M]$. The intensity value of the first and second frame are $\boldsymbol{I}(\boldsymbol{x})$ and $\boldsymbol{J}(\boldsymbol{x})$ respectively, where $\boldsymbol{x} = [x, y]^T$ corresponds to the pixel coordinate of the frame ($x = [1, 2...N]$, $y = [1, 2...M]$). Here, the tracking technique is applied on a region of interest called interrogation window. If an initial target pixel location $\boldsymbol{u} = [x_o, y_o]^T$ and an interrogation window W of size $\boldsymbol{w} = [w_x, w_y]$ around $\boldsymbol{u}$ are selected, the discrete cross-correlation function is given by

$$r(\boldsymbol{d}) = \sum_{\boldsymbol{x}=\boldsymbol{u}-\boldsymbol{w}/2}^{\boldsymbol{u}+\boldsymbol{w}/2} \Big(\boldsymbol{I}(\boldsymbol{x}) \cdot (\boldsymbol{J}(\boldsymbol{x}+\boldsymbol{d})) \Big) \qquad (1)$$

where $\boldsymbol{d} = [d_x, d_y]^T$ is a vector of pixel displacement between frames (d_x and d_y are the pixel displacements in the x and y direction respectively). Eq. 1 calculates the cross-correlation at a given $\boldsymbol{d}$. If Eq. 1 is applied at every $\boldsymbol{d}$ within W, a correlation map $\mathbf{R}(\boldsymbol{d})$ is obtained. To increase the computational efficiency, the cross-correlation is determined in the frequency domain via a Fast Fourier Transform, where the Fourier transform of $\boldsymbol{I}$ and $\boldsymbol{J}$ are defined as $\mathcal{F}(\boldsymbol{I})$ and $\mathcal{F}(\boldsymbol{J})$ respectively. The cross-correlation is expressed as

$$\mathbf{R}(\boldsymbol{d}) = \mathcal{F}^{-1}\Big(\mathcal{F}\big(\boldsymbol{I}(\boldsymbol{x})\big) \cdot \mathcal{F}^*\big(\boldsymbol{J}(\boldsymbol{x}+\boldsymbol{d})\big) \Big) \qquad (2)$$

where $\mathcal{F}^*$ is the complex conjugate of its Fourier transform, and $\mathcal{F}^{-1}$ is the inverse Fourier transform. The $\boldsymbol{d}$ associated to the maximum value (highest peak) of $\mathbf{R}(\boldsymbol{d})$ corresponds to the most probable displacement of the target. An initial interrogation window W_1 around the target is chosen in the first frame. A second interrogation window W_2, with same size and location as W_1, is placed in the next frame. Their mean intensity values are extracted to normalise image regions under uniform changes in brightness. Then, fixing the position of W_1, Eq. 2 is applied between both interrogation windows changing $\boldsymbol{d}$ to cover all pixel positions within W_1. This procedure implies that the target displacement between frames is smaller than half the interrogation window size. In addition, the target displacement can be estimated within a pixel, reaching subpixel accuracy, by approximating the spatial distribution of correlations around the peak with a continuous function.

The open source library OpenPIV [18] provides the functions to obtain the position of the correlation peak with subpixel accuracy.

2.2 Optical Flow: Lucas-Kanade

The details of the LK implementation used in this study can be found in [19]. Consider two sequential frames with intensity values as $\boldsymbol{I}(\boldsymbol{x})$ and $\boldsymbol{J}(\boldsymbol{x})$ respectively. Given an initial target pixel location $\boldsymbol{u} = [x_o, y_o]^T$ in the first frame, the goal of the LK technique is to estimate its location in the second frame, $[\boldsymbol{u} + \boldsymbol{d}]^T$, such as its corresponding intensity values around the point of interest reaches a minimum difference. As in the previous section, the vector $\boldsymbol{d} = [d_x, d_y]^T$ corresponds to a pixel displacement between frames. This implementation considers a square-shaped interrogation window W of size $\boldsymbol{w} = [2w_x + 1, 2w_y + 1]$

around u, and assumes that the intensity values are subjected to the following affine transformation:

$$\mathbf{A} = \begin{pmatrix} 1 + d_{xx} & d_{xy} \\ d_{yx} & 1 + d_{yy} \end{pmatrix} \qquad (3)$$

where d_{xx}, d_{xy}, d_{yx}, and d_{yy} characterise the affine deformation of the interrogation window. Approximating the intensity values of the second frame by a first order Taylor expansion series of the first frame (valid for small displacements), the objective is to find the vector of pixel displacement d and the affine matrix $\mathbf{A}$ within W that minimise the following function:

$$\epsilon(\boldsymbol{d}, \mathbf{A}) = \sum_{\boldsymbol{x} = \boldsymbol{u} - \boldsymbol{w}/2}^{\boldsymbol{u} + \boldsymbol{w}/2} \left(\boldsymbol{I}(\boldsymbol{x}) - \boldsymbol{J}(\mathbf{A}\boldsymbol{x} + \boldsymbol{d}) \right)^2 \qquad (4)$$
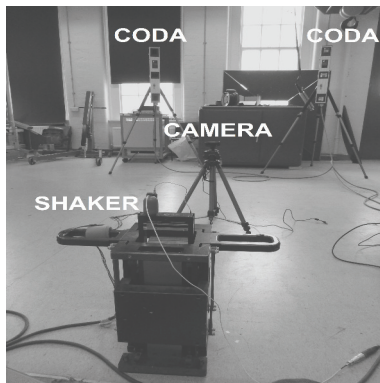
Eq. 4 is minimised following a Newton - Rapson iteration approach. In addition, subpixel accuracy is achieved through bilinear interpolation. In case of larger displacements, [19] considers a pyramidal scheme representation.

The OpenCV library [20] contains a well-documented version of this algorithm.

3 Experimental Setup
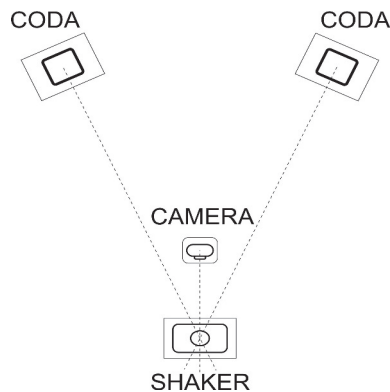
(a) Top view sketch.

(b) Photograph of the experiment.

Figure 1. Shaker recorded by a PS3 camera and a CODA system.

The experiments were performed at The University of Sheffield, United Kingdom. The APS 400 Electro-Seis shaker was used to produce a one-dimensional oscillatory, rigid motion. The amplitude and velocity of the shaker changed simultaneously to reach a particular frequency. As a consequence, a normalisation process is necessary to accurately compare across frequencies (explained later in this section).

Two systems are used to register the movement of the shaker, a three-dimensional (3D) motion tracking system called CODA CX1, used as a baseline, and a PS3 Eye recording camera (see Figure 1). The CODA system is composed of two self-calibrated 3D scanners, which track high-intensity infra-red LED markers at up to 800 Hz. The accuracy of this system in the target plane, indicated by the manufacturer, is approximately 0.05 mm when the distance between the CODA scanners and the marker is three meters. Thus, the CODA system is placed at two meters from the shaker position. The PS3 Eye camera is capable of recording 8-bit uncompressed images with a resolution of 640x480 pixels and a acquisition frequency of 75 Hz. The camera is mounted at 0.5 m from the shaker position with a field-of-view of 75 and a focal length of 2.1. These parameters produce less than 1% of distortion according to the manufacturer.

The movement of the shaker was tracked at three different shaking frequencies: 1 Hz, 3 Hz, and 4 Hz. The amplitude of the shaker for these frequencies correspond to 120.19 mm, 45.55 mm, and 25.42 mm respectively. The CODA system was set to track two-dimensional motions at 200 Hz. Simultaneously but unsynchronised, the PS3 Eye Camera recorded for 40 seconds with an acquisition frequency of 75 Hz. The discrete signals obtained by the CODA system and the ones used for the tracking techniques are subjected to the following normalisation process. Firstly, the CODA signal is sub-sampled using a spline cubic interpolation to match the sampling frequency of the PS3 Eye camera. Secondly, an equal temporal shift is applied to align the starting recording time of both systems. The shifted position is selected when the root mean square (RMS) between the CODA data and the displacements estimated by the two tracking systems reached a minimum value. Finally, both signals are normalised by their corresponding mean amplitude, followed by the subtraction of their mean positions to obtain independent coordinate signals. The mean displacement between consecutive data points of the normalised coda signal is equal to 0.056, 0.158, and 0.211 for the 1 Hz, 3 Hz, and 4 Hz respectively. The RMS between the normalised

tracking and CODA signals, termed here as RMS_n, is used as an error measurement throughout this study.

Even though the experiments were carefully conducted to reduce perspective and lens distortions, a calibration technique was applied to further minimise these elements. In the context of small displacements (e.g. shaker amplitude is 25.42 mm at 4 Hz), these sources of error can affect the accuracy of both tracking techniques and increment their RMS_n when they are compared to the high-precision CODA system. A calibration plate, composed by a regular grid of black circles with 3 mm diameter and 10 mm of separation, was temporally fixed on the shaker plane of motion. Then, an image of this plate was taken and used for the calibration process. A multi-quadric radial interpolation function was used to determine a transformation between the known marker distances and their corresponding pixel coordinates. Thus, undistorted real-world displacements are obtained from the original image pixel coordinates. A similar approach can be found for shallow-flow visualisations [21].
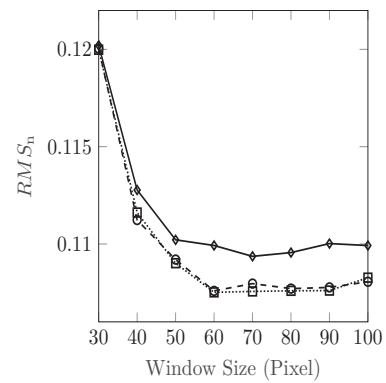
4 Sensitivity analysis

4.1 DIC-based technique: Interrogation window size and subpixel estimation method

Square-shape interrogation windows with sides from 30 to 100 pixels and 10-pixel increments are investigated. The influence of the Centroid, Parabolic and Gaussian subpixel methods on the overall accuracy of the DIC-based technique are also analysed.

Fig. 2a shows the RMS_n calculated for the 4 Hz series as a function of interrogation window sizes and subpixel estimation methods. For all interrogation window sizes, the Centroid estimation method gives the highest RMS_n compared with the Gaussian and Parabolic subpixel functions. Between these last two methods, there are no significant differences in their RMS_n. The RMS_n divided by the maximum RMS_n of a given shaker frequency is shown in Fig. 2b. As [14] indicates, there is a range of interrogation window sizes that maximise the technique accuracy, which is dependent on the target displacement between frames.

4.2 Lucas-Kanade technique: Integration window and initial target point selection

An evenly spaced squared grid of 121 initial target points with one pixel separation was placed in the shaker zone of motion. The LK technique was applied to individual
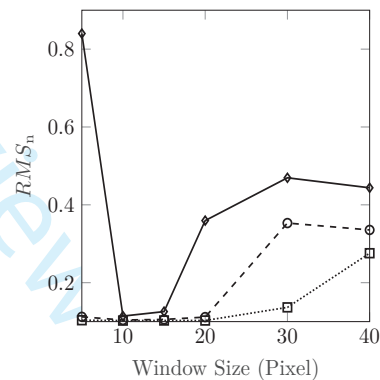

(a) ◇ : Centroid, -○- : Parabolic, ...□...: Gaussian.
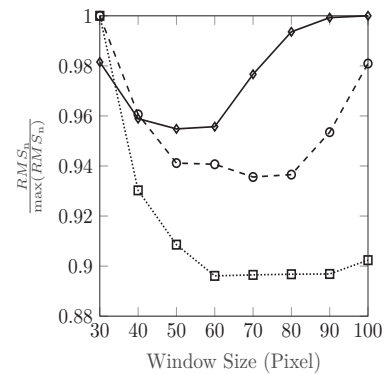
(b) ◇ : 1 Hz, -○- : 3 Hz, ...□...: 4Hz.

Figure 2. DIC sensitivity analysis. a) RMS_n vs window size using different subpixel techniques b) Percentage of the maximum RMS_n vs window size.

(a) ◇ : 90th percentile, -○- : 60th percentile, ...□...: 30th percentile.
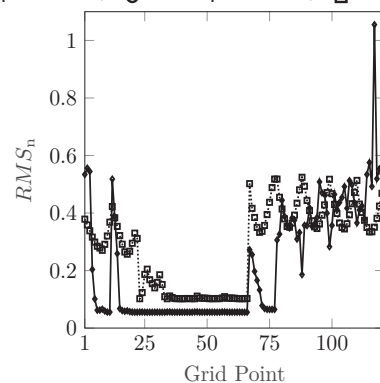
(b) ◇ : 1 Hz, ...□...: 4 Hz

Figure 3. LK sensitivity analysis. a) RMS_n vs interrogation window size (4 Hz shaker frequency). b) RMS_n at different initial target points.

grid points using a range of square-shape interrogation windows of sides from 5 to 40 pixels. Fig. 3a shows their corresponding RMS_n in terms of three percentiles: 30th, 60th, and 90th. For interrogation window sizes less than 20 pixels, at least 30% of the initial target points reached the global minimum RMS_n, which corresponds to the smallest RMS_n value considering all target points. Furthermore, at 90th percentile, the RMS_n decreases toward an optimum of 10 pixels. This optimum value is where the dependency of the LK technique on its initial target point is minimised. Outside this optimum of 10 pixels, smaller interrogation window sizes rely on a small amount of information to accurately track the movement of the shaker, which makes the tracking technique more sensitive. On the other hand, larger interrogation windows sizes produce excessive smoothing in the motion estimation and the error increases, as explained in [19]. Analysing the dependency of the LK technique on its initial target point, Fig. 3b shows all RMS_n from the 121 initial grid points at two different shaker frequencies. A suboptimal interrogation window size of 30 pixels, as shown in Fig. 3a, was used to illustrate the effects of different target displacements. Between grid points 25 to 65, a low RMS_n region is observed for both shaker frequencies. This particular region within the target corresponds to a high intensity gradient zone where the tracking technique achieves its highest accuracy. Except in a few grid points, the RMS_n decreases with the shaker frequency. Moreover, there is a higher number of target points reaching RMS_n closer to the global minimum at lower shaker frequencies. As the only experimental change was the shaker frequency, this result could be explained by the relationship between the camera exposure time and the target displacement between frames. Considering a constant exposure time, higher target displacements reduce the image intensity gradients and, as a consequence, lowers the tracking technique accuracy. The ratio between the 5th and 95th RMS_n percentiles obtained from Fig. 3b is used to quantify the LK accuracy regarding the initial target point selection. For the 1 Hz and 4 Hz shaker frequency, this ratio corresponds to 9.7 and 5 respectively. Furthermore, there is a high variability of the RMS_n between consecutive nearby target points. These results indicate that the LK technique is highly dependent on the initial target point selection.

5 Changes in the operational parameters

In this section, changes in the shaker frequency, image resolution, and acquisition frequency are used to test the accuracy and capabilities of both tracking techniques. In addition, their tracking motion estimation is analysed in the frequency domain.

5.1 Spectral analysis

Fig. 4 shows the normalised power spectral density (PSD) of the displacement obtained with the tracking techniques and CODA signals. The frequency of the shaker f_s was used to normalise the spectra. The Welch method with a window size of 225 data points, and a Hamming window was used for the power spectral density calculations. In general, there is a good agreement between the tracking techniques and the CODA system. In all cases, the normalised main peak frequency is equal to 0.975, 1 and 0.995 for the 1 Hz, 3 Hz, and 4 Hz respectively. However, there are some discrepancies in the estimation of the energy contributions in the high-frequency range. Within this range, the LK technique decreases the spectral energy contribution. On the other hand, the DIC-based technique amplifies this energy contribution, producing a flat spectrum at the highest frequencies. This flatness is normally associated with the lack of performance of the measuring apparatus and tracking technique to estimate small amplitude, high-frequency displacements [22].

5.2 Acquisition frequency

Changes in the camera acquisition frequency were obtained by resampling the original time series of images. Different set of images with sampling frequencies $f_s^{'}$ equal to the original acquisition frequency, 75 Hz, divided by an integer factor were generated. A lower limit for this reduction is given by the sampling theorem (Nyquist criteria) to avoid aliasing effects in the signal. This limit is also constrained by the existence of enough similarities between consecutive frames, in which both tracking are able to track the target with acceptable accuracy. After a series of tests above the Nyquist frequency, it was found that the lower for the shaker motion at 1 Hz, 3 Hz, and 4 Hz, are 12.5 Hz, 15 Hz, and 18.75 Hz respectively. The acquisition frequency of the CODA signal was also resampled to match the tested sampling frequencies.

Fig. 5a, Fig. 5b, and Fig. 5c show the accuracy of the tracking techniques at different sampling frequencies. Even though both tracking techniques were able to track the shaker motion across all tested $f_s^{'}$, there is an overall increment on the RMS_n as the sampling frequency decreases. As previously discussed, this could be explained by a reduction in the image intensity gradients as the target displacement between frames increases. The maximum increment in
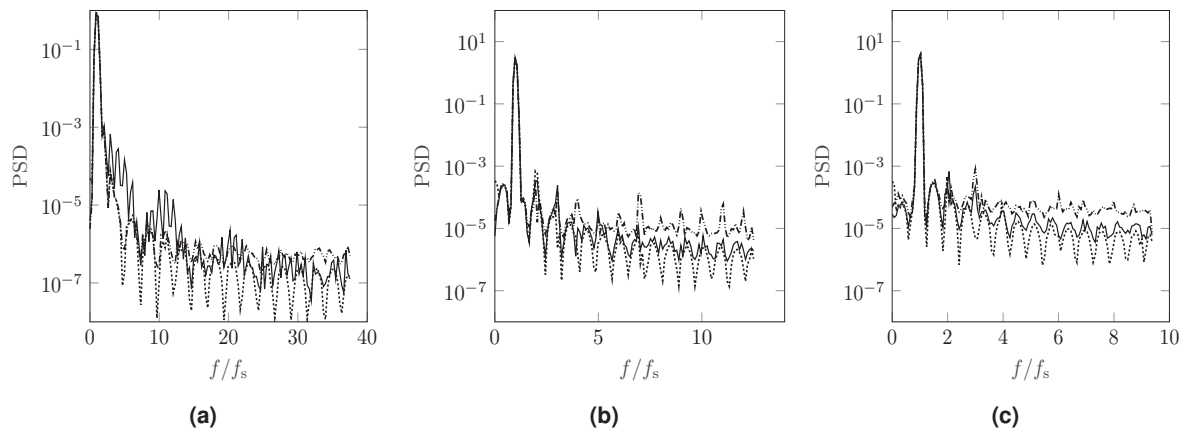
Figure 4. Power spectral density. ——: CODA. ┄┄┄: LK. ——: DIC-based. a) 1 Hz. b) 3 Hz. c) 4 Hz.
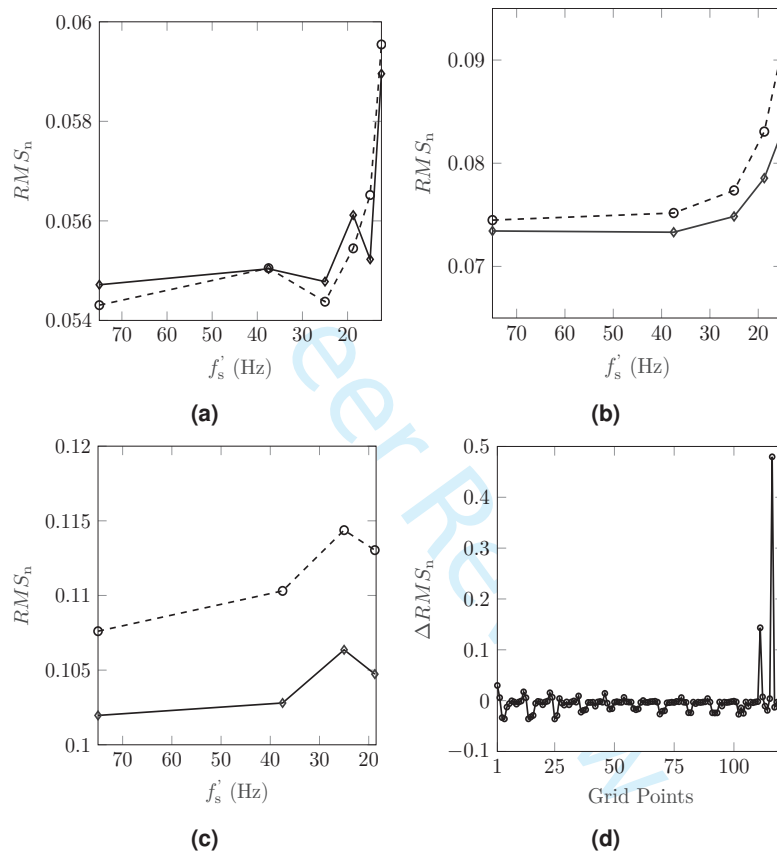

Figure 5. RMS_n vs $f_s^{'}$. ◇: LK, ⊖: DIC-based. a) 1 Hz. b) 3 Hz. c) 4 Hz. d) ΔRMS_n, where $s_1 = 75$ Hz and $s_2 = 18.75$ Hz, considering a shaker frequency of 4 Hz.

RMS_n across all $f_s^{'}$ is 7.9% and 12.3% for the LK and DIC-based technique respectively. On average, the LK technique shows a higher accuracy of 5.6% over the DIC-based algorithm at the highest shaker frequency. It is important to determine the influence of $f_s^{'}$ in the LK technique dependency to the initial target point selection. Changes of a specific parameter are assessed by introducing ΔRMS_n, defined as the difference between the RMS_n of two resulting

signals, s_1 and s_2, obtained from the same initial target point, i.e:

$$\Delta RMS_n = RMS_n^{s_1} - RMS_n^{s_2} \qquad (5)$$

Considering the grid described on the sensitivity analysis for the LK technique applied to the 4 Hz shaker frequency case, Fig. 5d shows the ΔRMS_n of two signals sampled at $s_1 = 75$ Hz and $s_2 = 18.75$ Hz. It is observed that decrements in $f_s^{'}$ do not significantly increase the LK dependency to the initial target point selection.

5.3 Image resolution

The dimensions of the original images were reduced in size up to a factor of 10 using four different interpolation techniques: nearest (the nearest pixel from the original image is used directly), bilinear (2x2 pixel window), bicubic (4x4 pixel window), and Lanczos. This last technique is a high-quality downsampling filter available in the Python Pillow library (https://pillow.readthedocs.io). Then, these low-resolution images were scaled back up to the original image dimensions using the nearest interpolation technique, which preserves their image intensity gradients. Poor tracking accuracy is achieved for the LK and DIC-based techniques when the original image resolution was reduced by a factor of six and eight respectively.
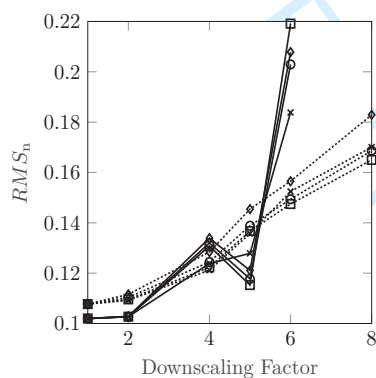
(a) ——: LK. ········: DIC-based. ◇: Bilinear, ○: Bicubic, □: Lanczos, ✕: Nearest

(b)
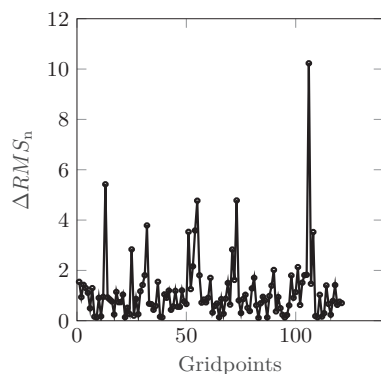
Figure 6. Downscaling effect. 4 Hz shaker frequency. a) RMS_n vs downscaling factor using different interpolation techniques. b) ΔRMS_n using s_1: original set, and s_2: five times downscaled.

Fig. 6a shows the RMS_n of both tracking techniques at different image resolutions. In general, the interpolation technique does not have a significant impact on the RMS_n variations at different image resolutions. There is an inverse relationship between resolution and accuracy for both tracking techniques. As the image resolution decreases, the same amount of information is contained in a lower number of pixels. As a consequence, the image intensity gradients are reduced along with the accuracy of the tracking techniques. The DIC-based technique shows an almost linear relationship across the analysed range of image resolutions. In contrast, the LK technique does not show a clear pattern up to a downscaling of five. Then, there is a sudden increment in RMS_n for a downscaling factor of six, indicating a significant reduction in its tracking motion estimation. Considering a downsampling factor of five and the images generated using the Lanczos algorithm (lowest overall RMS_n), the LK technique has a 15.6% lower RMS_n compared with the DIC-technique. Moreover, considering the same conditions as before, there is an increment in RMS_n of 13.4% and 20.7% for the LK and DIC-based technique respectively. From these observations, it seems that a decrement in the image resolution (i.e. loss of information) affects the tracking techniques differently. The LK technique relies on the image intensity gradients, where a lower image resolution introduces larger restrictions compared with those involved in the DIC-based technique, which relies on the target intensity patterns. Analogous to the previous section, ΔRMS_n was calculated to determine if changes in the image resolution increase the dependency of the LK technique on the initial target point selection. Selecting s_1 and s_2, where s_2 was obtained from the original dataset and s_2 from the images downscaled five times in resolution, Fig. 6b shows that the image resolution has an important influence on the location of the best initial target point selection.

6 Fluid-Structure Interaction experiment

In this section, the estimation of the spatio-temporal displacement of a cylinder subjected to a range of incoming turbulent flows is analysed. The comparative analysis from previous sections has shown that the LK technique is moderately better in terms of tracking accuracy, and it will be used in this section.

6.1 Experimental setup

An experiment performed at The University of Sheffield, United Kingdom, was designed to analyse the response of a slender, lightweight, wall-mounted, emerged cylinder subjected to three flow conditions. The Reynolds number is defined here as $R_e = UD/\nu$, where U is the bulk flow velocity, D is the diameter of the cylinder, and ν is the kinematic viscosity of the fluid. The tested flow conditions correspond to R_e 430, 660, and 950 at a discharge rate of $Q_1 = 0.0151$ m^3 s^{-1}, $Q_2 = 0.0231$ m^3 s^{-1}, and $Q_3 = 0.0355$ m^3 s^{-1} respectively. Q_3 is also the highest possible

discharge rate in the facility. For all tested flow conditions, the incoming turbulent intensity was 5%. The experiment was composed by a recirculating water channel covered by clear cast acrylic sheets, a cylindrical obstacle, and a recording device (see Fig. 7). The flume has a longitudinal fixed slope of 0.001 $\mathrm{m\,m}^{-1}$ and a rectangular cross-sectional area of 486 mm width. A water depth of 347 mm was maintained using a computer-controlled system. The cylindrical obstacle was made of clear cast acrylic of $D = 5$ mm diameter and a density of 1.19 $\mathrm{g\,cm}^{-3}$. The obstacle was tightly embedded in an acrylic sheet by fixing one end of the cylinder and leaving the other end unsupported. This arrangement allowed the cylinder to oscillate freely in both longitudinal (x-axis) and transverse (y-axis) directions. The cylinder length measured from its tip to the acrylic base was 491 mm with a natural frequency f_{nat} of 6.74 Hz. This value was calculated imposing an initial one-dimensional displacement on its free end. The response was recorded and tracked, obtaining a f_{nat} closer to its theoretical value considering an Euler-Bernoulli cantilever beam behaviour.
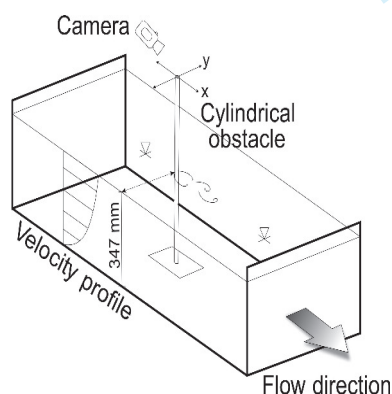

Figure 7. Experimental setup of a cylinder subjected to an open-channel turbulent flow. Cylinder oscillations in transverse (y-axis) and longitudinal (x-axis) directions.

6.2 Tracking estimation

The maximum observable cylinder displacement for the tested flow conditions was approximately one diameter in the transverse direction. This imposes a requirement for the experimental setup, in which the transverse displacement needs to be contained within the imaged region as well as to capture the smallest longitudinal displacements with a high enough resolution. Previously, it has been shown that a lower image resolution or, likewise, an increment in the camera-object distance reduce the accuracy of both tracking techniques. Furthermore, the experimental setup presents length constraints in terms of camera mounting, available space, and object distance. As a consequence, a high-resolution MX 4M camera with a spatial resolution of

2048x2048 pixels was employed. The camera focused on a black marked circle at the free end of the cylinder. This mark generates an area with high-intensity gradients and is used as the initial tracking target. Each experiment was recorded at 8-bit and 70 Hz for two minutes. Pictures of a calibration plate, LaVision model 058-5, positioned at the cylinders free end were taken to correct for perspective and optical distortions. The calibration was performed using the software DAVIS 8.3, following a similar procedure as described in the shaker experimental setup.

As shown previously, the LK technique has a strong dependency on the selection of the initial target point. This drawback is addressed using the Forward-Backwards (FB) tracking failure technique [17]. This algorithm is based on the fact that the optical flow is independent on the direction of time, i.e. the tracking of an object from the spatial position P_1 to P_2 in time is the same as between point P_2 to P_1 moving backwards in time. In practice, P_1 and the final position tracked from P_2 backwards in time, defined as P_1', will not be the same. Their Euclidean difference is called FB-error and is a measure of the LK technique accuracy. This process is applied to a number of initial target points to identify the ones that satisfactorily represent the true motion of the target. If a cloud of initial points is considered, it is possible to select the signals whose FB-error is in the lowest 5th percentile. Once these signals are identified, their corresponding mean displacement is subtracted to make them coordinate independent. At each time step, the displacement is calculated as the median value of the selected signals. Using the shaker data, Fig. 8 shows the comparison between the tracking motion estimation using the FB technique and the minimum RMS_{n} obtained on the LK sensitivity analysis section for a range of interrogation window sizes. The maximum difference across all shaker frequencies is 7.5%, showing that the FB technique effectively finds the initial target points that maximise the accuracy of the LK technique without using additional external sensors. Since the LK technique dependency is also a function of the target displacement, interrogation window sizes of 5 up to 40 pixels were tested on the recorded cylinder data. The minimum FB-error was achieved at interrogation window sizes of 10, 15, and 20 pixels. Within these options, the lower interrogation window size was selected (10 pixels) since smaller interrogation window sizes reduce the smoothing of the tracking estimation.

An example highlighting the benefits of the FB implementation for the analysis of the cylinder displacements are presented Fig. 9 and Fig. 10. The values were normalised using the parameters D and f_{nat}. A window size of 350
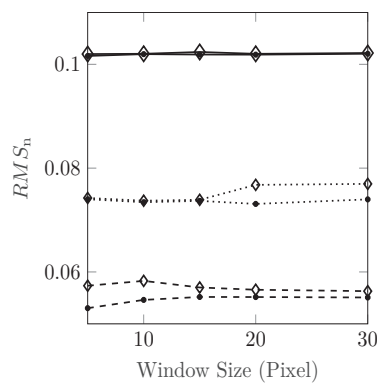
Figure 8. FB with LK using the shaker images. - - -: 1 Hz, ·······: 3 Hz, and ——:4 Hz. At each shaker frequency, ●: $\min(RMS_n)$, and ◇ : FB.

data points, and a Hamming window was considered for the Welch power spectra density calculations. In these figures, a comparison between the tracking displacements obtained implementing the FB algorithm and the tracking results that correspond to the 99th FB-error percentile (worst initial target point) were performed for different R_e. Significant differences in displacements and frequencies are observed. These differences are larger in the longitudinal direction, specifically in the low and high-frequency spectral range. The longitudinal response of the cylinder is up to 20 times smaller than its transverse motion. As a consequence, it is more difficult to track its longitudinal motion at a given image resolution.

7 Conclusions

This paper analyses the accuracy and robustness of the cross-correlation via Fast Fourier Transform (DIC-based) and Lucas-Kanade (LK) techniques under changes in acquisition frequency, image resolution, and internal tracking parameters. These results are then used for the quantification of body motion in a fluid-structure interaction experiment. The selection of the processing parameters of both tracking techniques plays a critical role in the estimation of the true target motion. The LK technique shows a strong dependency on the selection of the initial target point, which varies according to the target displacements and image intensity gradients. There is a range of interrogation window sizes in which the LK technique reaches its maximum accuracy. Within this range, a correct selection of the interrogation window size has the additional benefit of substantially decrease the dependency of the initial target point. On the other hand, the Parabolic and Gaussian subpixel methods of the DIC-based technique show a superior accuracy compared to the Centroid method. There is an optimum range of interrogation window sizes that

maximise the technique accuracy. With the correct selection of the tracking technique parameters, both techniques were able to track the main motion and frequency of the target even after a reduction of four and five times the sampling frequency and image resolution respectively. Considering a reduction of five times the original images resolution, the LK technique has a 15.6% lower RMS_n compared with the DIC-technique. In addition, the LK technique has, on average, a lower 5.6% RMS_n over the DIC-based algorithm at the highest tested shaker frequency and at different acquisition frequencies. Overall, the LK technique shows a superior accuracy when it is able to track the target. On the other hand, the DIC-based technique shows a higher robustness under systematic reductions in image resolution. Despite its moderately higher accuracy, the LK technique shows an important dependency on its initial target point, which difficult the achievement of its better accuracy. This dependency can be minimised through the implementation of the Forward-Backward tracking failure technique without depending on additional external sensors. These considerations were addressed in a fluid-structure interaction experiment, were the body motion of a cylindrical obstacle subjected to a range of turbulent flow was successfully extracted using a single recording device.

8 Acknowledgements

References

[1] Baqersad J, Poozesh P, Niezrecki C et al. Photogram-metry and optical methods in structural dynamics A review. *Mechanical Systems and Signal Processing* 2017; 86: 17–34.

[2] Helfrick MN, Niezrecki C, Avitabile P et al. 3D digital image correlation methods for full-field vibration measurement. *Mechanical Systems and Signal Processing* 2011; 25: 917–927.

[3] Poudel UP, Fu G and Ye J. Structural damage detection using digital video imaging technique and wavelet transformation. *Journal of Sound and Vibration* 2005; 22: 869–895.
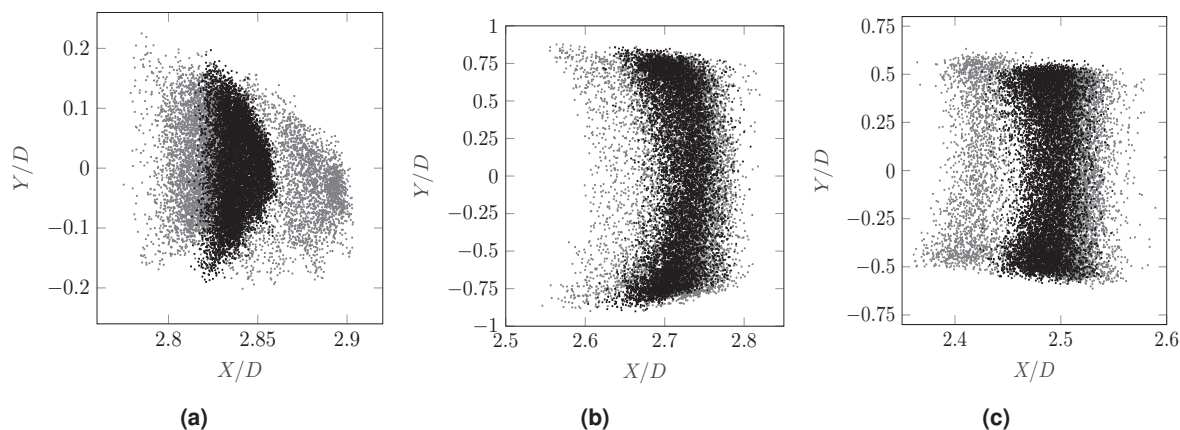
(a) **(b)** **(c)**

Figure 9. Cylinder displacement at different flow rates. Comparison between LK with the FB implementation and a LK result with a 99th FB-error percentile. ●: LK with FB. ●: 99th FB-error percentile. a) Q_1. b) Q_2. c) Q_3

(a) From top to bottom: Q_1, Q_2, and Q_3

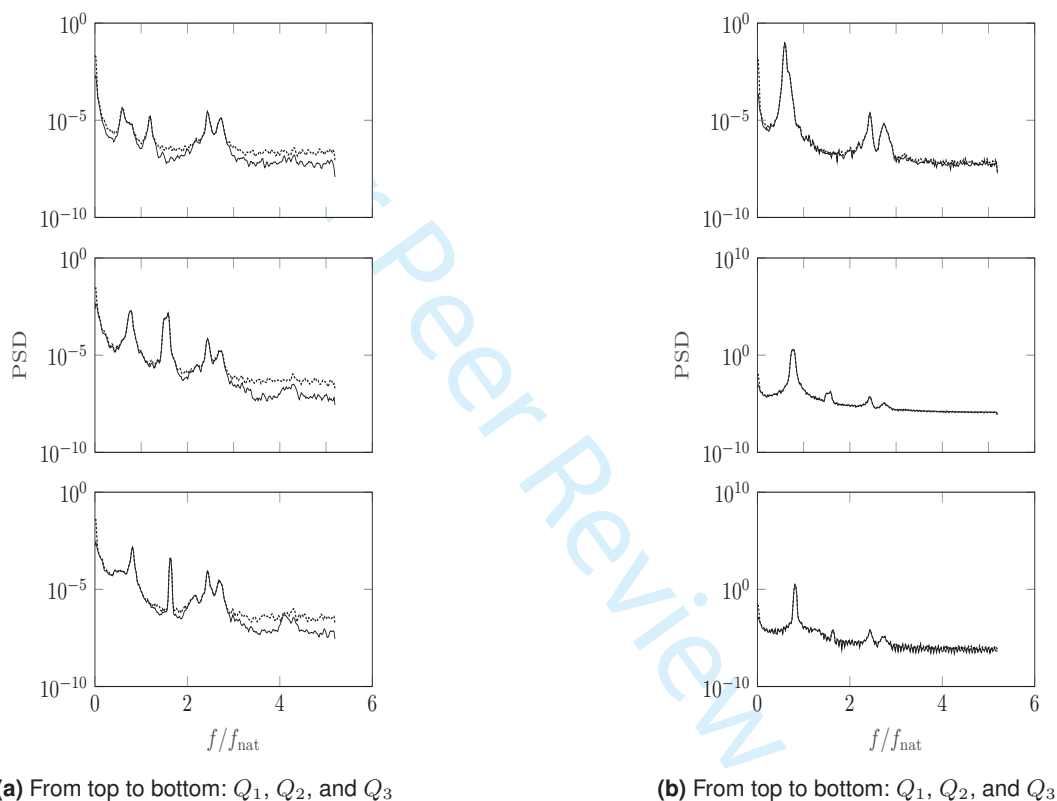**(b)** From top to bottom: Q_1, Q_2, and Q_3

Figure 10. ——: LK with FB, ⋯⋯: 99th FB-error percentile. a) Longitudinal spectra displacement. b) Transverse spectra displacement.

[4] Ehrhardt DA, Allen MS, Yang S et al. Full-field linear and nonlinear measurements using Continuous-Scan Laser Doppler Vibrometry and high speed Three-Dimensional Digital Image Correlation. *Mechanical Systems and Signal Processing* 2017; 86: 82–97.

[5] Vanlanduit S, Vanherzeele J, Longo R et al. A digital image correlation method for fatigue test experiments. *Optics and Lasers in Engineering* 2009; 47: 371–378.

[6] Huera-Huarte FJ. An optical instrument based on defocusing for dynamic response model testing in water or wind tunnels. *Ocean Engineering* 2014; 79: 92–100.

[7] Kwon YW, Priest EM and Gordis JH. Investigation of vibrational characteristics of composite beams with fluid-structure interaction. *Composite Structures* 2013; 105: 269–278.

[8] Lucas BD and Kanade T. An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proceedings of the DARPA Image Understanding Workshop 1981*. pp. 121–130.

[9] Kim SW and Kim NS. Dynamic characteristics of suspension bridge hanger cables using digital image processing. *NDT & E International* 2013; 59: 25–33.

[10] Sieffert Y, Vieux-Champagne F, Grange S et al. Full-field measurement with a digital image correlation analysis of a shake table test on a timber-framed structure filled with stones and earth. *Engineering Structures* 2016; 123: 451–472.

[11] Morlier J and Michon G. Virtual Vibration Measurement Using KLT Motion Tracking Algorithm. *Journal of Dynamic Systems, Measurement, and Control* 2010; 132: 11003–11011.

[12] Ji YF and Chang CC. Nontarget Image-Based Technique for Small Cable Vibration Measurement. *Journal of Bridge Engineering* 2008; 13: 34–42.

[13] Rao GVR, Sreekala R, Kumar KS et al. Seismic Response Measurement of an Under-Water Model Through High Speed Camera and Feature Tracking. *Experimental Techniques* 2016; 40: 83–90.

[14] Bornert M, Brémand F, Doumalin P et al. Assessment of digital image correlation measurement errors: Methodology and results. *Experimental Mechanics* 2009; 49: 353–370.

[15] Zappa E, Mazzoleni P and Matinmanesh A. Uncertainty assessment of digital image correlation method in dynamic applications. *Optics and Lasers in Engineering* 2014; 56: 140–151.

[16] Atcheson B, Heidrich W and Ihrke I. An evaluation of optical flow algorithms for background oriented schlieren imaging. *Experiments in Fluids* 2009; 46: 467–476.

[17] Kalal Z, Mikolajczyk K and Matas J. Forward-backward error: Automatic detection of tracking failures. In *Proceedings - International Conference on Pattern Recognition 2010*. pp. 2756–2759.

[18] Taylor ZJ, Gurka R, Kopp GA et al. Long-duration time-resolved PIV to study unsteady aerodynamics. *IEEE Transactions on Instrumentation and Measurement* 2010; 59: 3262–3269.

[19] Bouguet JY. Pyramidal implementation of the affine lucas kanade feature trackerdescription of the algorithm. *Intel Corporation* 2001; 5: 1–10.

[20] Bradski G. The OpenCV Library. *Dr Dobbs Journal of Software Tools* 2000; 25: 120–125.

[21] Brevis W and García-Villalba M. Shallow-flow visualization analysis by proper orthogonal decomposition. *Journal of Hydraulic Research* 2011; 49: 586–594.

[22] Bendat JS and Piersol AG. *Random data: Analysis and Measurement Procedures*. 4th ed. New York: John Wiley & Sons, 2011.