POLITECNICO DI TORINO
Repository ISTITUZIONALE

An experimental study of the feasibility of phase-based video magnification for damage detection and localisation in operational deflection shapes

Original
An experimental study of the feasibility of phase-based video magnification for damage detection and localisation in operational deflection shapes / Civera, M., Zanotti Fragonara, L., Surace, C.. - In: STRAIN. - ISSN 0039-2103. - (2020), p. e12336. [10.1111/str.12336]

Availability:
This version is available at: 11583/2779308 since: 2020-01-21T17:07:13Z

Publisher:
Wiley

Published
DOI:10.1111/str.12336

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Wiley postprint/Author's Accepted Manuscript

This is the peer reviewed version of the above quoted article, which has been published in final form at http://dx.doi.org/10.1111/str.12336.This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

(Article begins on next page)

19 November 2026

strain
An International Journal for Experimental Mechanics

An Experimental Study of the Feasibility of Phase-Based Video Magnification for Damage Detection and Localisation in Operational Deflection Shapes

Journal:	*Strain*
Manuscript ID	STRAIN-1499.R1
Manuscript Type:	Full Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Civera, Marco; Politecnico di Torino, Department of Mechanical and Aerospace Engineering; Cranfield University, School of Aerospace, Transport and Manufacturing Zanotti Fragonara, Luca; Cranfield University, School of Aerospace, Transport and Manufacturing surace, cecilia; Politecnico di Torino, Department of Structural, Geotechnical and Building Engineering
Keywords:	damage detection, experimental modal analysis, motion magnification, operational deflection shapes, structural health monitoring, video processing

SCHOLARONE™
Manuscripts

An Experimental Study of the Feasibility of Phase-Based Video Magnification for Damage Detection and Localisation in Operational Deflection Shapes

Marco Civera[1*†], Luca Zanotti Fragonara[2‡] and Cecilia Surace[3§]

[1]Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, 10129, Turin, Italia. Tel.: 0110904911, Fax: 0110904911
[2]School of Aerospace, Transportation and Manufacturing, Cranfield University, College Rd, Cranfield, MK43 0AL, United Kingdom
[3]Department of Structural, Building and Geotechnical Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, 10129, Turin, Italia

For Peer Review

Abstract

Optical measurements from high-speed, high-definition video recordings can be used to define the full-field dynamics of a structure. By comparing the dynamic responses resulting from both damaged and undamaged elements, Structural Health Monitoring (SHM) can be carried out, similarly as with mounted transducers. Unlike the physical sensors, which provide point-wise measurements and a limited number of output channels, high-quality video recording allows very spatially dense information. Moreover, video acquisition is a non-contact technique. This guarantees that any anomaly in the dynamic behaviour can be more easily correlated to damage and not to added mass or stiffness due to the installed sensors.

However, in real-life scenarios, the vibrations due to environmental input are often so small that they are indistinguishable from measurement noise if conventional image-based techniques are applied. In order to improve the signal-to-noise ratio (SNR) in low-amplitude measurements, Phase-Based Motion Magnification (PBMM) has been recently proposed.

This study intends to show that model-based SHM can be performed on modal data and time histories processed with PBMM, whereas unamplified vibrations would be too small for being successfully exploited. All the experiments were performed on a multi-damaged box beam with different damage sizes and angles.

Keywords: damage detection, experimental modal analysis, motion magnification, structural health monitoring, video processing

*Corresponding author.
†E-mail: marco.civera@polito.it
‡E-mail: l.zanottifragonara@cranfield.ac.uk
§E-mail: cecilia.surace@polito.it

1 Introduction

Decades of continuous development allowed a consistent amount of vibration-based approaches to be proposed for the damage detection and assessment of structures (e.g. [1]). This is due to the global and local effects of the damage, which affects the overall dynamic characteristics of the structure but also works as a local source of added flexibility. Specifically, some of the most successful vibrational approaches exploit modal information for Structural Health Monitoring (SHM) purposes. Nowadays, Experimental and Operational Modal Analysis (EMA / OMA) are common practice for both academic or industrial applications [2]. However, some practical issues persist, despite their widespread use. For instance, mounted sensors, like, e.g., accelerometers, strain gauges, or Linear Variable Differential Transformers (LVDTs), inevitably add mass to the measured structure and they can also increase locally the bending stiffness, according to how they are attached to the structure. These pointwise local changes produce effects which can be easily confounded with damage [3]. Moreover, most of these techniques, while theoretically valid, requires a large number of output channels for the recording of the structural response. This is particularly true for damage localisation, which implicitly requires a high spatial resolution for both beam- and plate-like structures (e.g., [4–6] and [7,8]).

Considering the hierarchical scheme of damage identification proposed in [9], a second-level procedure technique focused on local changes can hardly be achieved by a single point-wise recording, even if some attempts exist [10]. In fact, the spatial resolution of the acquisition negatively affects most of the common damage localisation techniques which make use of mode shapes and their first and second derivatives. This becomes an issue in real-life scenarios, when affixed sensors are, for economic constraints or other practical reasons, deployed in a limited number. Hence, the information gathered cannot be but discretised and sparse and often suboptimal.

Optical measurements are a possible solution for some of these issues. Video cameras offer a non-contact measurement technology that can produce spatially dense information. Moreover, they are (relatively) low-cost and agile to deploy: if no physical obstacle subsists in the line of sight, spatial resolution is only a matter of focal distance between the camera lens and the investigated structure. They also allow different angles of observation, even resulting in a stereo acquisition if needed.

The pinhole camera model behind optical methods of measure can be defined with relative ease; once displacements are defined in terms of pixels, the conversion to SI length units (i.e., millimetres) is only a matter of calibration, by multiplying them by the known length of a reference object divided by the number of pixels it spans over [11–13]. Other non-contact alternatives exist: laser scanning is the main competitor of video-based techniques. A complete discussion will be too long to be reported here; an exhaustive comparison can be found in [14] and some related studies recently made by the Authors are reported in [15][16]. For the aim of this work, it is enough to say that high-speed cameras are much more common and easy to access than multi-point laser scanners, especially for untrained labourers.

Some other attempts to perform damage localisation from video recordings exist in literature. For instance, [17] used the Lukas-Kanade Optical Flow algorithm; whilst [18] developed a real-time digital image processing routine. Plenty of similar concepts exist; optical measurements extracted from video recordings are established and common in practice. Digital Image Correlation (DIC) [19], point tracking (PT) [20], and many other approaches (e.g., [21–24],

For Peer Review

to cite a few) are all currently used to extract time-dependent displacements. All these video acquisition techniques are known to be satisfactorily accurate; nevertheless, each one has its own caveats and technical requirements and limitations: some need high-contrast markers, speckle pattern or other arrangements before recording. Vision-based techniques are also often coupled with other more traditional techniques, especially for the monitoring of the dynamics of vast surfaces such as e.g. landslips or glaciers [25]. In particular, they have been applied extensively in the field of SHM (see, e.g., [26–29]). DIC techniques have been successfully use for inspection of bridges [30] and of other civil engineering structures [31] in both concrete and asphalt [32], as well as for mechanical components and machines, even by means of unmanned aerial vehicle video acquisitions [33]. Interested audience may find useful information, especially for the applications of DIC techniques for Aerospace and Civil Engineering instances, in [34].

The approach applied here, the Phase-Based Motion Magnification (PBMM) algorithm [35], differentiates from these other methods in resorting to a post-processed source rather than the actual video; hence, the useful information is magnified and the Signal-to-Noise Ratio (SNR) is artificially increased. The useful motion information, previously indiscernible, is so undisclosed. Specifically, it uses the Complex Steerable Pyramid [36] architecture for spatial and temporal filtering the video frame by frame, to amplify subtle motions undetectable by the naked eye. This can be used to extract both pixel-wise displacement time histories and Operative Deflection Shapes (ODSs) roughly correspondent to the structure's mode shapes [37].

The rest of the paper is organised as follow. In Section 2, the Phase-Based Motion Magnification procedure is outlined and discussed; the underlying theory and implementation steps are reported for completeness. A brief comparison with similar works is reported as well. In Section 3, the case study proposed is described in detail. Section 4 reports the obtained results. Finally, in Section 5 conclusions are drawn. Some hints regarding damage identification, assessment and modelling for multiple and/or angled damages are provided where needed.

2 Phase-Based Motion Magnification

Motion magnification techniques can be broadly classified between Lagrangian and Eulerian approaches, depending on if a given brightness-related feature is tracked frame-by-frame among the image for all pixels, or if instead the brightness intensity is measured at a fixed voxel grid. The PBMM algorithm belongs to the family of the Eulerian approaches; any subgroup of pixels can be used to obtain time series of brightness intensity and therefore of any related motion, and any arbitrarily large element of the grid can be used as a 'virtual' accelerometer [38]. The pros of this approach can be listed as follow: respect to a Lagrangian approach such as in [39], it allows to process the video decoupling the spatial and temporal problems; thus, the computational cost is greatly reduced, and selective amplification (that is to say, the magnification of only a certain range of temporal frequencies) may be performed. Time-consuming tasks like tracking, layer segmentation, and inpainting the manufactured video are avoided. With respect to the linear Eulerian technique proposed in [40], noise in the video is translated and not amplified with the content of interest; moreover, intensity-based Eulerian techniques support only smaller amplification factors and requires very high spatial frequency to perform optimally.

2.1 Mathematical formulation of the problem

Let I be a 2D image intensity profile under global translation over time. Let

For Peer Review

the frame being permanently w pixels wide and h pixels height. If $\boldsymbol{x}=(x,y), x\in[0,w], y\in[0,h]$ is the 2-dimensional pixel coordinates vector, t is the time expressed in terms of frames and $\delta(\boldsymbol{x},t)$ is the locally-defined, time-dependent displacement function, then $I(\boldsymbol{x}+\delta(\boldsymbol{x},t))$ defines frame-by-frame the spatial brightness profile of the whole captured image. For comparison, both Optical Flow (OF) and DIC revolve around *I* as well: OF methods approximate the motion field under the image brightness consistency constraint [41] between current and initial frames; DIC techniques estimate $\delta(\boldsymbol{x},t)$ through image correlation between $I(\boldsymbol{x},t)\ \forall t$ and a reference frame $I_0(\boldsymbol{x})=I(\boldsymbol{x},t_0)$.

The aim of motion magnification is to obtain a "manufactured" video montage where $I(\boldsymbol{x}+(1+\alpha)\delta(\boldsymbol{x},t))$ for an arbitrary magnification factor α. If it is legit to assume all motions (global and local ones) to be pixel-wise linear, $\delta(\boldsymbol{x},t)$ can be represented as a linear combination of sinusoids; this can be therefore achieved by considering a spatial 2-dimensional Fourier series decomposition, i.e.,

$$I(\boldsymbol{x}+\delta(\boldsymbol{x},t))=\sum_{\omega=-\infty}^{+\infty}\sum_{\theta=1}^{k}A_{\omega,\theta}e^{i2\pi\omega_0(\boldsymbol{x}+\delta(\boldsymbol{x},t))} \qquad (1)$$

along both x- and y-direction, for each complex sinusoid $S_{\omega,\theta}(\boldsymbol{x},t)=A_{\omega,\theta}e^{i2\pi\omega_0(\boldsymbol{x}+\delta(\boldsymbol{x},t))}$ corresponding to a given combination of spatial scale ω (at non-dimensional pulsation ω_0) and orientation θ, with the corresponding spatial amplitudes $A_{\omega,\theta}(\boldsymbol{x},t)$. The pivotal concept is, therefore, to amplify the related motion by modifying the spatial phase $\varphi_\theta(\boldsymbol{x},t)=2\pi\omega_0(\boldsymbol{x}+\delta(\boldsymbol{x},t))$, accordingly to the Fourier shift theorem. In fact, it is known [42] that phase provides a better estimation of the motion than the brightness amplitude, especially for very small motions and in conditions of changing illumination, thus outperforming both DIC and OF methods. Local displacement can be directly extracted from spatial phase information $\varphi_\theta(\boldsymbol{x},t)$. Focusing on φ_θ alone, regardless of the actual amplitude associated with it, one can observe that contours of constant spatial phase change from frame to frame, accordingly to the structure motion. For a fixed instant *t* and direction θ, it is thus possible to define $\varphi_\theta(\boldsymbol{x},t)=\varphi_\theta(x,y)\equiv c$ for any arbitrary value of c. By varying *t*, the phase contours become smooth space-time surfaces. Since the points move within the motion field, $\varphi_\theta(x(t),y(t),t)\equiv c$ must hold; thus, by deriving along time, according to the chain rule, it is possible to define the local velocities (in terms of pixels over the 2D frame space) from the gradient of spatial phase, as [42]

$$\nabla\varphi_\theta(\boldsymbol{x(t)},t)\cdot(\dot{\boldsymbol{u}},1)=\left(\frac{\partial\varphi_\theta(\boldsymbol{x(t)},t)}{\partial x},\frac{\partial\varphi_\theta(\boldsymbol{x(t)},t)}{\partial y},\frac{\partial\varphi_\theta(\boldsymbol{x(t)},t)}{\partial t}\right)\cdot(\dot{u},\dot{v},1)=0 \qquad (2)$$

Being $\frac{\partial u(\boldsymbol{x},t)}{\partial t}=\dot{u}(\boldsymbol{x},t), \frac{\partial v(\boldsymbol{x},t)}{\partial t}=\dot{v}(\boldsymbol{x},t)$ the velocities along the horizontal and vertical directions, respectively.

For Peer Review

By assuming that $\frac{\partial \varphi_{\theta=0}(x,y,t)}{\partial y} \approx 0$ and $\frac{\partial \varphi_{\theta=\pi/2}(x,y,t)}{\partial x} \approx 0$, one obtains that

$$\begin{aligned} \dot{u}(\boldsymbol{x},t) &= -\left(\frac{\partial \varphi_{\theta=0}(x,y,t)}{\partial x}\right)^{-1} \frac{\partial \varphi_{\theta=0}(x,y,t)}{\partial t} \\ \dot{v}(\boldsymbol{x},t) &= -\left(\frac{\partial \varphi_{\theta=\pi/2}(x,y,t)}{\partial y}\right)^{-1} \frac{\partial \varphi_{\theta=\pi/2}(x,y,t)}{\partial t} \end{aligned} \quad (3)$$

By comparing the velocity at any given frame respect to the initial value at $t=0$, displacement time histories can be defined for the whole duration of the video, theoretically for any given subregion of pixels [43]. However, homogeneous/textureless regions may be locally ambiguous. Hence, in [38] it is strongly suggested to restrain the use of this technique to edges, and only for the investigation of the motions happening in the direction perpendicular to them. This is what has been done here since the transverse motion was the major interest in this case study.

Magnification can be circumscribed to a specific range of frequency by bandpass filtering. This allowing also to remove the temporal mean $2\pi\omega_0 \boldsymbol{x} \; \forall \boldsymbol{x} = (x,y)$ and isolating the vibrational motion at any given location, i.e.,

$$B_\omega(\boldsymbol{x},t) = \delta'(\boldsymbol{x},t) = 2\pi\omega_0 \delta(\boldsymbol{x},t) \quad (4)$$

Where the bandpassed 2D matrix of local phases at time *t*, $B_\omega(\boldsymbol{x},t)$, can be multiplied by α (prime symbol in " δ' " only indicates a factorisation by the constant $2\pi\omega_0$). This cause the selected range of frequencies to undergo a magnified phase shift, hence a proportionately amplified motion, that is to say,

$$\hat{S}_{\omega,\theta}(\boldsymbol{x},t) \coloneqq S_{\omega,\theta}(\boldsymbol{x},t) e^{i\alpha B_\omega} = A_{\omega,\theta} e^{i 2\pi\omega_0 (\boldsymbol{x} + (1+\alpha)\delta(\boldsymbol{x},t))} \quad (5)$$

What defined so far applies to global motions, as the 2-D Fourier transform decomposes the frame in sinusoids along the whole image along the x- and/or y-directions. To perform it frame-by-frame in a local region of interest, a spatially multi-scale, spatially localised set of transform filters are needed. The Complex Steerable Pyramid (CSP) is a viable option in this sense.

2.2 Complex Steerable Pyramid

Basically, the pyramid architecture is an overcomplete, steerable, and complex-valued band-pass filter bank, able to decompose a given input (here, a frame) according to its orientation, position and spatial scale. In the signal processing ambit, the overcompletness of a filter refers to the degree of redundancy of its frame. That is to say, the frames of a Gabor system (i.e. a time-frequency shift made up by functions constructed from modulating, dilating and translating a generating function) not being a Riesz basis of a Hilbert space, can produce more robust and stable decompositions than by resorting to a nonredundant basis, but at the cost of more computational effort and less compactness [44]. This property is here exploited as it provides a larger amplification capability. In the same context, the steerability of the same indicates an orientation-selective convolution kernel, which may be expressed via a linear

For Peer Review

combination of a set of basis filters, i.e., rotated versions of itself [45].

The Complex Steerable Pyramid has a wavelet-like impulse response, in the sense of finite spatial support, necessary to capture the variations of the locally-defined motion in a 2D frame. It is important to remark that the wavelets themselves are not suited for motion estimation; in fact, the well-known Mallat algorithm for Discrete Wavelet Transform (DWT) is affected by aliasing problems (i.e. low-frequency contaminations), due to the implicit limitations of orthogonal separable wavelet decompositions. Even Laplacian Pyramid, another widespread alternative for multi-scale representation, may suffer the same issues and, moreover, it is not steerable and more difficult to invert for signal/video reconstruction (i.e., it is not self-inverting; because of that, the error introduced by sub-bands quantisation also produces low-frequency distortions in the reconstructed image). Eventually, the aim is to maintain the useful multi-scale sub-band decomposition typical of wavelets, often used for time or space localisation in the frequency domain, but also having them (almost) aliasing-free in the space domain, thus equivariant with respect to translation. This is obtained thanks to the sub-sampling, overcomplete scheme of the CSP. Its filters resemble complex Gabor filters, that is to say, globally-defined complex sinusoids windowed by a Gaussian envelope. It is this modulation that gives them finite spatial support. Resorting to complex numbers is essential to deal efficiently with local amplitudes and local phases simultaneously. Again, similar to wavelets, all the transfer functions in the oriented bands of the pyramid are scaled and *rotated* 'children' of a basic filter, defined by their scale and orientation, that is, $\Psi_{\omega,\theta}$, so that the discrete Fourier transform of a single frame can be defined (in the frequency domain) as

$$\tilde{I} = \sum_{\omega=-\infty}^{+\infty} \sum_{\theta=1}^{k} \tilde{S}_{\omega,\theta} \Psi_{\omega,\theta} \quad (6)$$

where $\tilde{S}_{\omega,\theta}$ are the DFTs of the several spatial frequency bands $S_{\omega}(\boldsymbol{x},t)$.

It must be noted that a CSP's transfer functions only contain the positive frequencies of the respective real pyramid's ones. Since the decomposition in the frequency domain is polar-separable, the independence of scale and orientation representation is guaranteed. For any number of orientation bands k, the basis functions of the pyramid are k^{th} order directional derivative operators, spanning over a rotation-invariant subspace. The major issues are overcompleteness, which equals $4k/3$ (by comparison, Laplacian Pyramid's overcompleteness is 4/3, and orthogonal wavelets are not overcomplete at all. Moreover, complex pyramids are twice as overcomplete than real-valued ones), and ringing artefacts, which are mainly due to the not perfect reconstruction in the space domain. Another limit, which will be discussed in more detail later, regards the size of the local support.

Collapsing the CSP produce a manufactured, motion-amplified frame for the resulting video montage. The steerability of the pyramid allows performing this amplification in every direction, even if for the case study here reported, transversal displacements are of major interest (given that the axial displacements are orders of magnitude smaller and of much less interest). Thus, confounding them with the vertical component of motion (with almost no error in practice), the direction of greatest interest is for $\theta = \pi/2$.

2.3 Motion Amplification Procedure

By applying the filtering procedure described above to every spatial scale and every orientation of the CSP, a collection of amplified and non-amplified motions is defined

For Peer Review

(the overall scheme is sketched in Fig 1). The results are ad hoc displacement THs and ODSs, magnified of an arbitrary factor in the frequency range of interest.

The procedural steps are as follow. Firstly, the 2D Fourier Transform is computed (step 'a' in Fig 1); this way, each point represents a frequency contained in the image in the spatial domain. Being the image size the same in both the spatial and the frequency domain, the range of inspected spatial frequencies is directly linked to the image definition, i.e., to the spatial sampling; the higher the number of pixels $N = w \times h$, the broader the range becomes. On the other hand, for frames, the data become a 3D $w \times h \times M$ tensor of complex numbers, with a very large size and very sparse useful information. This decomposition along spatial sinusoids is well-suited for Modal Analysis, as the expected mode shapes to be extracted are sinusoidal as well. The zero-frequency component – that is to say, the spatial average brightness of the frame – is shifted to the centre of the image in the amplitude spectrum, and higher frequencies are confined to the edges. The bright spot in the middle of Fig 1 image in the top right corner is evidence of the prevalence of content at the lower spatial frequencies.

At this step, scaled and steered real-valued filters are applied to the complex matrices (step 'b' in Fig 1). Filters can be thought of as 'masks' applied in the Fourier domain of the single frames; the results are the components at the several scales and directions. Moreover, being the filter bank pyramidal, the number of pixels is halved at any decomposition step, not too different from what happens with the aforementioned Mallat's Algorithm for DWT. This convolution process produces coefficients that, being complex numbers, define both the local phase and the local amplitude of the windowed basis functions at any point and along any direction. According to the user's desires, a frequency band of interest is defined between the two bound values f_{low} and f_{high}; the so-obtained coefficients are then filtered, bandpassing in the time domain, to isolate the temporal frequencies required while cutting off lower and higher frequency content which shall not be magnified (step 'c' in Fig 1).

The spatial amplification itself is user-defined in terms of the magnification factor α. Thus, multiplying scalar-wise only their phase component, the motion is 'amplified' (in the sense, over-displaced) without any manipulation of brightness' amplitude (step 'd' in Fig 1). Moreover, an optional amplitude-weighted spatial smoothing was performed over the σ adjacent pixels, to avoid unrealistic discontinuities due to noisy measurements and/or excessive manipulation. The unmodified/untouched low-pass and high-pass residual information are then finally added to reach complete reconstruction (step 'e' in Fig 1). The procedure, up to this point, can be performed with no major modifications to the software codes provided by the authors of [35,40] and [43] in the MatLab toolboxes available at [46] and [47].

It is then possible to extract the ODSs from the manufactured video output by means of edge detection (step 'f' in Fig 1). In the study here reported, a Canny edge detection routine [48] was found to produce a more robust and reliable results respect to similar alternatives (Sobel, Prewitt, or Roberts algorithms). The routine is simply required to recognise frame-by-frame the edges of the beam, to isolate them from the surroundings, and to extract them. To do so, a lower and a higher threshold of the Canny edge detection algorithm are empirically set, according to the illumination and the background, to discard as much unneeded information as possible.

The preserved pixelwise edges (i.e., the ones with edge strength above the lower limit and below the upper one) are then used as an approximation of the actual mode shape which natural frequency falls into the prescribed range of temporal frequencies.

For Peer Review

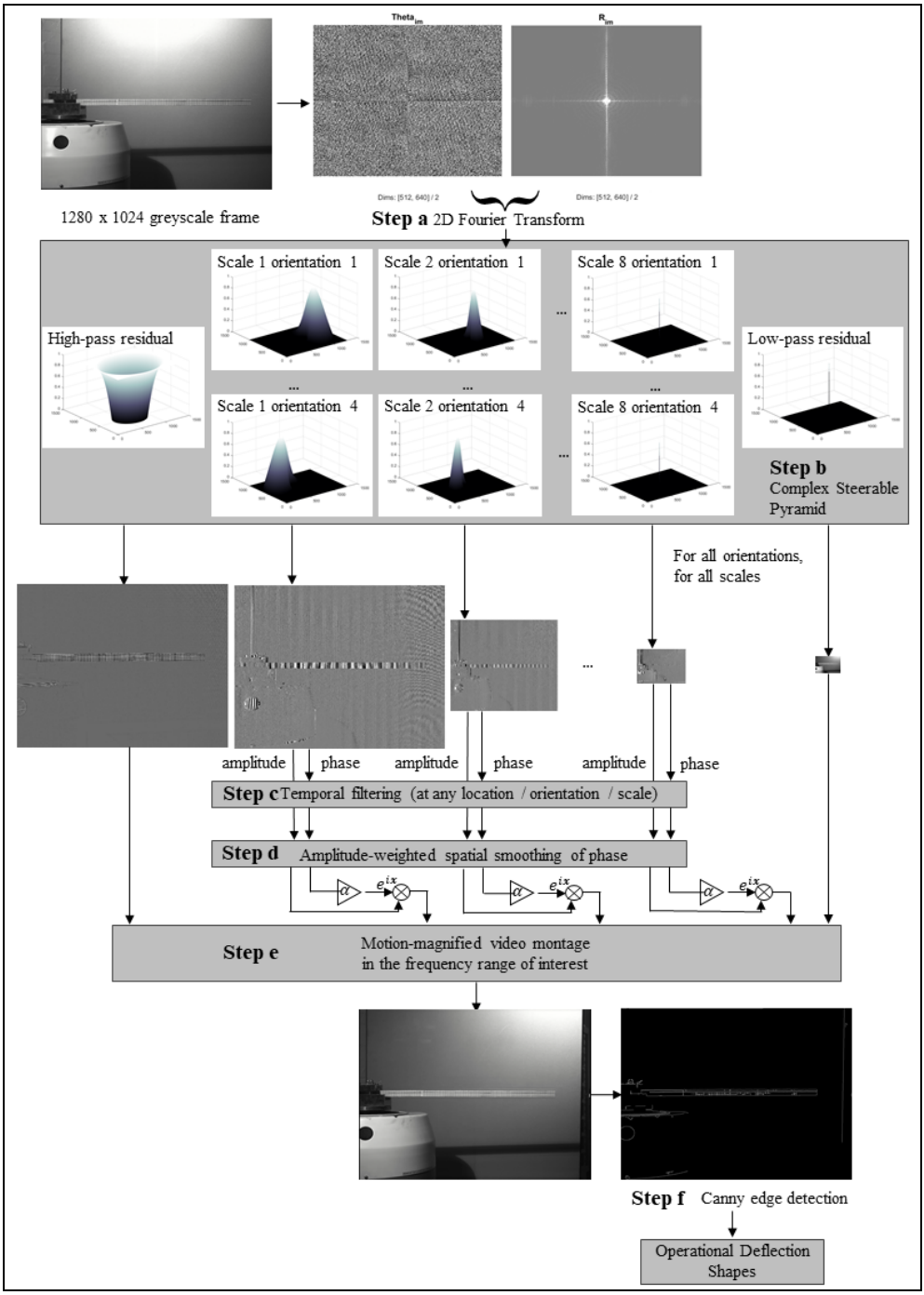


Fig 1. Scheme of the Motion-Magnification and ODSs extraction procedure.

2.4 Limitations of the method

For reasons that will emerge later on, it is important to recall the theoretical and practical limits of this spatial amplification procedure. The length of the windowing Gaussian function defines their spatial support, in terms of pixels. An optimal compromise is to consider octave-bandwidth spatial filters. However, the local phase can only be shifted within the filter window support, as any further translation would (a) have strongly

reduced or null intensity and (b) produce artefacts in unwanted neighbouring locations. This point is better discussed in [35]; for simplicity sake, it is sufficient to say that to sustain larger phase shifts, wider windowing function would be needed. However, resorting to sub-octave bandwidth means to increase the overcompleteness of the problem, as adjacent supports would overlap. Moreover, overlapping windows make the estimation of local phase more difficult and are thus less apt for the direct extraction of displacement signals.

Frequency-wise, satisfactory results were obtained by applying octave-width pyramids with relatively limited computational costs, maximum compactness and minimal over-completeness. For ODSs, larger magnification factors were needed, in order to better highlight local discontinuities; thus, sub-octave CSPs (specifically, with four filters per band) were used there.

2.5 Distinctions with similar works

PBMM is a novel technique, as mentioned previously, firstly proposed in [35] for several applications, encompassing (but not limited to) the vibrational monitoring of structures. The approach was also recently conjugated with a Blind-Source Separation technique for the operational modal analysis of simple structures [49], extended to stereo camera vision for the investigation of 2-dimensional plates [50], and combined with Machine Learning [51]. During the writing of this paper, another very recent work was brought to the Authors' attention, which merged PBMM and Particle Tracking Velocimetry to perform 3-dimensional modal identification [52]. The same approach has been thereafter applied to the structural dynamic characterisation of wind turbine blades [53]. Differently from what performed there, here a direct comparison between the pristine and the damaged ODSs has been used to pinpoint the location of reduced stiffness; several kinds and extension of damage have been also considered.

The work presented here is a methodological and schematic experimental investigation of the PBMM algorithm possibilities as a video processing tool, applied in the ambit of enhanced visual inspection for SHM purposes. More specifically, the three issues of damage detection, localisation and severity assessment are all addressed, plus the direct visualisation of the amplified motion of the damage's edges. The efficiency of the procedures for damage detection and localisation is also studied for different depths and angles, with single and multiple saw-cuts. This experimental validation is intended to fall in the broader effort of the whole SHM community to make optical methods in general, and the video-magnification algorithm specifically, a robust and reliable tool for practitioners in civil and mechanical engineering.

3 Experimental Setup

The experimental setup for this study is pictured in Fig 2; the characteristics of the aluminium box beam portrayed are reported in Table 1. The tests were performed at Cranfield with a Data Physics® Signal Force™ modal shaker; DP760 close-loop control system was used to induce 11 μs-long triangular impulses at the clamped base, with a peak value of v_g = 5 g. Measurements were performed with an Olympus ® I-speed 3 camera, setting always a frame rate of $f_s = 1000$ frames per second (fps). Due to the camera internal memory capacity, recordings were limited to 4897 frames, 1280x1024 pixels wide (width x height). The high pixel density was useful to achieve a high spatial frequency resolution. All videos were recorded and then stored uncompressed in *.hsv* format. An open-faced lighting fixture with a tungsten light source was used for illumination. The camera was mounted on a levelled tripod and it is assumed to be perfectly static.

A video stabilisation routine (available on the MatLab ® Computer Vision System Toolbox) was nevertheless applied before the phase-based magnification, to neglect any imperceptible and random extraneous movement due to possible environmental excitations. Focal length was set to about 160 cm, chosen in order to be not too close (to avoid lens distortion at the frame edges) neither too far, having the aluminium box beam occupying most of the horizontal width of the image. The camera and the light source were both slightly moved during the experiment for practical reasons. This turned out to be also useful to test the method robustness with respect to prospective and illumination changes, which proved to be good.

Theoretically, the Nyquist's limit of the acquisitions is $fs/2 = 500$ Hz. However, for practical reasons, many sample points per period are necessary for accurate time-frequency estimation. This is particularly important for impact tests, where instantaneous variations in frequency and amplitude in the free decay response matter and high temporal frequency resolution is needed. Thus, only frequency content up to 1/20th of f_s (50 Hz) was inspected. This eventually limited the study to only the first eigenfrequency of the beam, as the second mode is above 250 Hz.

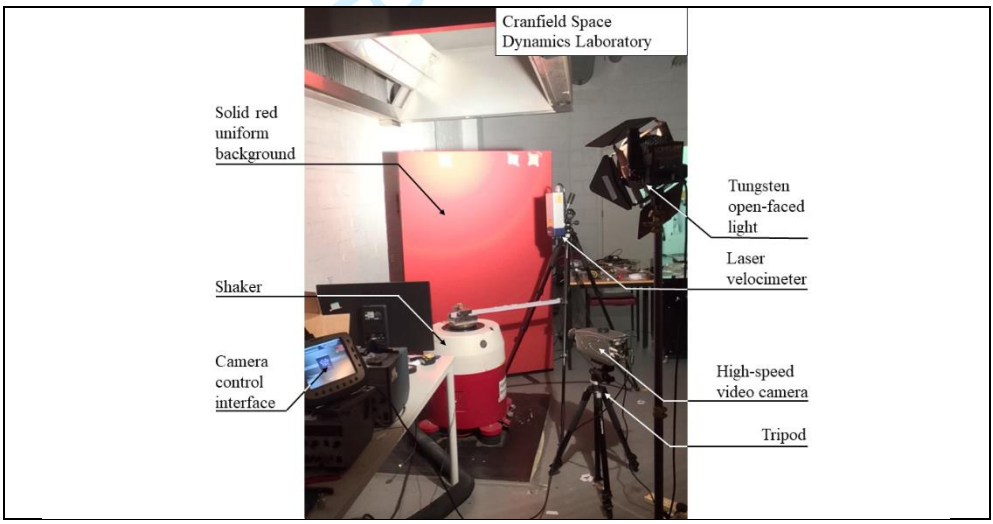


Fig 2. Experimental Setup.

Table 1: Material and geometrical properties of the aluminium bar

Young Modulus E	$69\ GPa$
Mass per unit of length μ	$0.322\ kg/m$
	(Density ρ $2170\ kg/m^3$)
Free Length L	$716\ mm$
Side (W=H)	$25\ mm$
Thickness t	$2\ mm$
Cross-section A	$184\ mm^2$
Moment of inertia I	$1.63 \cdot 10^4\ \ mm^4$

The target "damage" has been emulated by means of two u-shaped slots, regular and equal throughout the whole beam width, artificially inserted by sawing the box beam at two separate locations (cut thickness 1 mm; Fig 3.a). The first notch was positioned at $x'_1 = 205\ mm$ from the clamped section. The second notch was introduced in $x'_2 = 319\ mm$, closer to the box beam mid-length. The two notches are axially distanced apart enough ($\Delta x = 114\ mm$) to not influence the stress fields of one another. The crack growth was simulated by further sawing the specimen: damage conditions are reported in Fig 4.a - 4.g; the order of increasing damage is reported in Table 2. The second slot was cut with a 5:1 mm inclination (vertical rise to horizontal run; thus $\beta \cong 90° - 78.7° = 11.3°$) to investigate both the effects of multiple damages and the influence of slanted cuts. In this case, the damage location is intended as the intersection between the notch and the outer beam edge.

Emulating the damage by means of saw-cuts with known geometrical properties allows the repeatability of the experiment and enable to better discern the components (angle of slope, length of the cut, etc) that influence the frequency shift and the increase in the transverse displacements. This is a convenient expedient used since decades (e.g. [54]) and still pretty common in recent works to study damages within a fixed location and increasing depth, even if obviously only approximating real-life cracks. For a deeper discussion on the comparison of naturally occurring fissures and man-made crack-like notches, particularly in the context of experimental validation for vibration-based techniques, the reader is referred to [55]. To reduce any possible environmental-induced uncertainty, all the experimental tests were performed on one day; the beam was not dismounted from the clamp until the end.

For Peer Review

Table 2: Damage Levels

Damage severity case (reference #)	damage depth (*)	
	d_1	d_2
no damage	-	-
#1	1 mm	no damage
#2	5 mm	no damage
#3	10 mm	no damage
#4	15 mm	no damage
#5	15 mm	5 mm
#6	15 mm	10 mm
#7	15 mm	15 mm

(*) intended as the vertical component of length; corresponding to damage extension for $l_1 = d_1$ but not for $l_2 = d_2 / \cos(\beta) > d_2$.

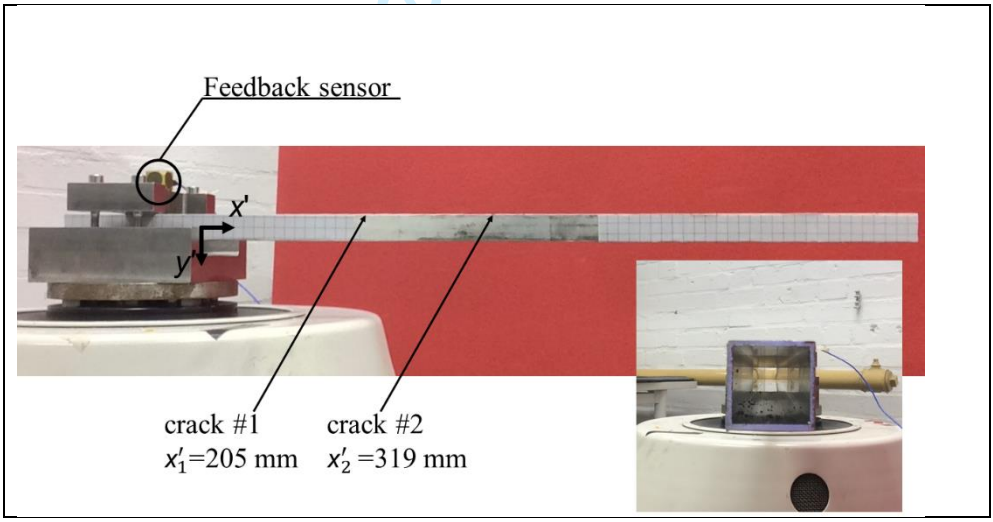


Fig 3. Box beam lateral view and hollow cross-section. The beam reference frame {O,*x*',*y*'} is also shown; black arrows point to the intersections between the edge slots and the edge itself.

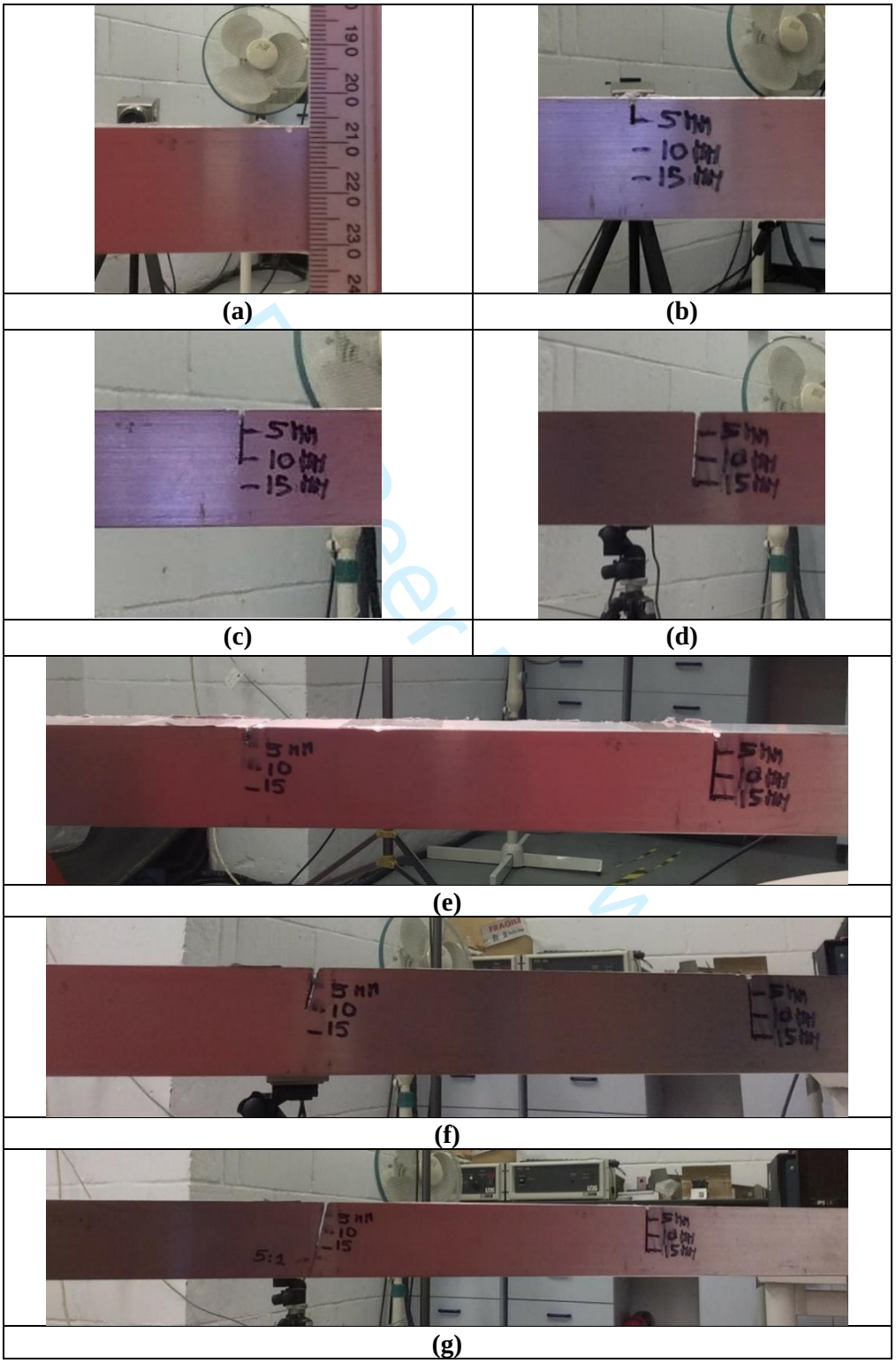


Fig 4. (a – g) increasing damage conditions: in the same order, case #1, #2, #3, #4, #5, #6, and #7 (see Table 2 for better description).

4 Results of the Damage Investigation

Damage detection and damage localisation were performed on different video recordings for the same damage level conditions, to establish separately that the frequencies extracted from the first video acquisition were actually confirmed in the second. For both damage detection and damage localisation, short sections were extracted from videos, starting right after the impulse application. This way, the captured transverse motion was assumed as pure free oscillations and there was no need to depurate it from the input motion v_g imposed at the base.

The dynamic behaviour of a multi-cracked box beam depends on the number of damaged cross-sections present, N_c, and on their parameters – namely, for the i^{th} damage, the crack depth d_i, the slant angle β_i, and their respective location x'_i. All these parameters eventually affected the damage assessment process.

The selected range of frequencies for motion magnification was defined between 30 – 50 Hz. That range surrounds the undamaged structure's first natural frequency, $f_n \cong 46$ Hz, while being slightly skewed to the lower frequencies to account for its expected decrease when damage is introduced. In order to have comparable results, it was maintained equal for all the damage scenarios.

4.1 Damage Detection and Severity Estimation

The frequency shift of the first fundamental mode is expected to vary appreciably when damage is inserted [1]. Indeed, the frequency shift in the structural response from impact tests is a classic methodology for damage detection, well-known in Literature and standard in common practice since decades. The comparison is made with the displacement time history from the pristine structure under a similar input: with no additional sensors' masses attached, all changes can be directly linked to a (local) reduction in structural bending stiffness. Tracking the decreasing trend of the first eigenfrequency as the fissure length is artificially incremented can be considered as an index for damage severity assessment. The exact position of the crack is also known to influence this frequency shift [56]; for instance, for any combination of boundary conditions and vibration mode, the inflexion points of the corresponding eigenshape will be unreactive to the presence of damage, irrespective to its depth and slope. However, there are caveats in applying these differences in frequency shift as an approach for damage localisation. In fact, these methodologies are arguably more subject to confounding influences than eigenshapes-based techniques like the one here utilised and described in the following Section.

By applying the PBMM algorithm, a time history (TH) of spatial phase changes (and thus, of related magnified displacement) was extracted in the range of frequencies analysed, by comparing the n^{th} frame with the first one. The extrados upward edge at mid-length cross-section was used as a 'virtual' sensor for output recording. Only the vertical component was considered, assuming it equal to the whole transverse motion (with a negligible error for very small oscillations). Fourier Transform was then performed on them; the procedure is pictorially described in Fig. 5 for better comprehension. 2-second-long (2000 frames) fractions of the several videos were considered, to reduce the computational cost of the whole procedure. This resulted in a reduction of resolution in the frequency domain ($\Delta f = f_{s,Video} / N_{samples} = 0.5$ Hz), which is still considered a good resolution for the specific case here investigated. Results are enlisted in Table 3; all frequency values are rounded up to the first decimal

For Peer Review

unit. As it can be observed, there is a divergence between the values reached and the ones acquired from laser velocimeter on random noise (RN) tests. This may also be due to the different sampling frequency (being for the laser velocimeter $f_{s,LDV} = 2048$ Hz). The frequency shifts respect to the pristine structure's fundamental frequencies are also reported.

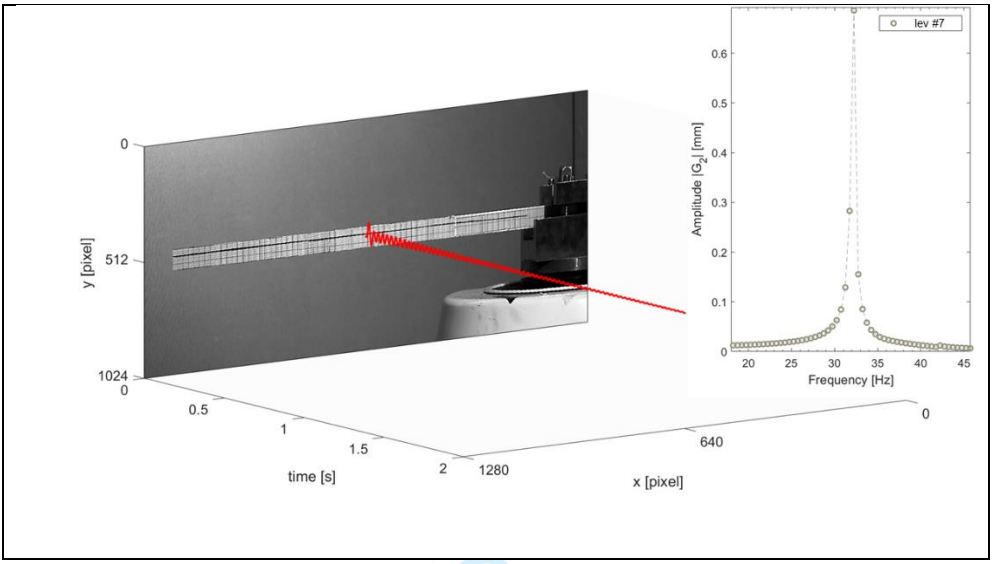


Fig 5. A conceptual sketch of the TH extraction and Fourier Transform process.

Table 3: First natural frequency, as extracted by the laser tachometer from RN tests and through video PBMM from impulse tests.

damage level	laser velocimeter (*) (random noise)	Frequency shift	video PBMM (*) (impulse)	Frequency shift
	[Hz]	[%]	[Hz]	[%]
no damage	46.3	-	46.3	-
#1	45.8	-1.1	46.3	~0
#2	44.7	-3.5	45.7	-1.3
#3	42.2	-8.9	40.7	-12.1
#4	37.0	-20.1	34.8	-24.8
#5	36.7	-20.7	34.5	-25.5
#6	36.1	-22.0	33.8	-27.0
#7	34.4	-25.7	32.2	-30.5

(*) data recorded at $x' = L/2$

The frequency shift between no damage and damage level #3 (corresponding to a 10-mm-deep indentation of the upper surface) exceeds 5% of the initial undamaged value, which in real-life operative conditions can be considered as a safe threshold between actual damage and statistical variations due to changing environmental conditions e.g. temperature. Further changes can be linked to the increased damage severity level. By

inserting the second slot, the first natural frequency is additionally reduced. However, multiple damages with different depths and/or angles make the damage investigation problem ill-posed, as slots with different lengths and orientations can produce identical frequency shifts of the first resonance [57]. Thus, damage sizing is only viable for the first four cases analysed (damage levels #1 to #4).

It can be inferred from Table 3 that the effects of increasing the size of the second notch are much less evident than for the first one. This is coherent from what known from Literature. Indeed, as a consequence of oblique crack propagation, mixed-mode stress conditions are induced at the crack tip – more precisely, a combination of mode I (tensile) and mode II (in-plane shear) loading. Thus, the apparent Stress Intensity Factor (SIF) at the crack tip should be partitioned into its distinct components [58]. Moreover, as long as the two damages have different depths, the larger one has majorly significant effects on the first natural frequency, especially if it is closer to the clamped end. Also, the farther the two damages are located, the smaller is the frequency reduction caused by the superposition of their effects (for the same damage severity and if both are located far from any antinodal point of the corresponding mode shape).

4.2 Operational Deflection Shapes & Damage Localisation

Any damaged mode shape is known to deviate from its corresponding undamaged, original version, as increased transverse displacements happen due to the local effects of stiffness reduction. This is a much more convenient parameter for damage localisation than natural frequencies, as their behaviour to varying crack position and length is very complicated and subject of studies since decades [9]. By running a Canny edge detection routine frame-by-frame on the processed video montage, the ODSs of the amplified motion corresponding to the first mode's natural frequency were isolated and stored. The assumption made is that, being the modes well-separated, the actual motion very small and the selected bandwidth of frequencies relatively narrow, the PBMM is able to cleanse the motion from the other superimposed modes. Hence, the amplified ODSs can be considered as a good approximation of the first mode, for both the pristine structure and all the damage scenarios.

For this aim, 0.2 seconds-long tracts of the video captured were selected for video processing, taken straight after the end of the input force application. Thus, 200 frames out of 4897 recorded, i.e., a tenth of the amount used for damage detection, were employed. That was intended to speed up the procedure and to reduce the computational burden. Moreover, it was observed that after relatively a few oscillations, even amplified motion decay to a point where the damage-induced local increase in flexibility becomes indistinguishable. Instead, for any recording considered, the difference in amplitude among the few periods so selected was found to be always negligible.

Being the undamaged first natural frequency roughly equal to 46 Hz, up to nine periods could be obtained from the test on the pristine structure; for damaged conditions with lower eigenfrequencies and longer periods, slightly fewer instantaneous pictures of the damaged ODSs at their largest deflection were captured. In turn, any period of oscillation for a damaged beam can be further divided into a compressive half-cycle and a tensile half cycle, considering the direction of the forces acting on the crack's edges and tip. For the specific case tested here (i.e. edge saw-cuts located on the extrados), the compressive semi-period corresponds to the upward transverse deflection of the beam, while tensile stresses develop during downward bending. It is crucial to remind how the operative shapes, bending in the two directions, should be theoretically identical as the system remains linear (no evident signs of nonlinearity where found in the recorded output, and the surfaces of the two slots were visibly seen to never come in touch; this point will be better addressed later on in Subsection

For Peer Review

4.3). However, the effects of gravity cause these two configurations to differ evidently in practice.

A comparison between a downward and the following upward deflection is depicted in Fig 6.a, for the maximum level of damage (case #7, magnified 16 times). Respect to the undamaged case, the effects of the local reduction in stiffness are immediately noticeable. Indeed, for the same level of external excitation and for the same amplification factor, transverse displacements at the beam free tip are much more pronounced in the damaged scenario respect to the pristine structure, especially for downward deflections. This is qualitatively represented in Fig 6.b and 6.c.

In all cases, amplitude factors up to $\alpha = 16$ were used, for better evidence the otherwise invisible vibrations; quarter-octave CSPs were utilised to support larger amplification. These made the problem 56x over-complete, increasing the number of spatial filters from 34 to 258. Unit spatial smoothing was applied ($\sigma = 1$).

Aiming at a centimetre-wise localisation, 66 points, included between $x'_0 = 55$ mm and $x'_{TIP} = 716$ mm and equally spaced of 10 mm from each other (apart from the last point, which is only 9 mm from the second to last), were considered along the reference line (Fig. 7). In-between points were interpolated linearly, with sub-pixel accuracy. For simplicity, relative positions are therefore expressed respect to the normalised length $\Delta\xi' = (x' - x'_0)/(x'_{TIP} - x'_0)$ (the two slots positions being $\Delta\xi'_1 = 0.23$ and $\Delta\xi'_2 = 0.40$ according to this non-dimensionalised length).

As expected from the dynamics of a slant-cracked beam [59], the vibrations are larger in the case of the transverse cut; oblique slot angles produces a lower decrease in bending stiffness and thus weaker, less evident effects, even if their length $l_i = d_i / \cos(\beta)$ is actually slightly longer than for their transverse cut of equivalent depth (where $l_i = d_i$). The resulting ODSs are reported in Fig 8. Moving mean was used to further smooth out the measurement noise; ODSs are normalised over their maximum displacement for comparability. The Modal Assurance Criterion (MAC) was also applied as a global damage index, by comparing the damaged cases and the baseline eigenshape of the pristine structure; the values are reported in Table 4 (1 meaning complete correspondence with the undamaged beam and 0 no correspondence).

Damage-induced effects can be noticed as an apparent stiffening of the beam tract included between their position and the beam left (clamped) end. This is caused by the normalisation respect to the maximum transverse displacement; as aforementioned, it is the complementary tract between the damaged cross-section and the beam right (free) end that undergoes larger transverse displacements. Results are consistent with what expected, even if some noise and other effects due to subpixel interpolation and approximation persist.

For Peer Review

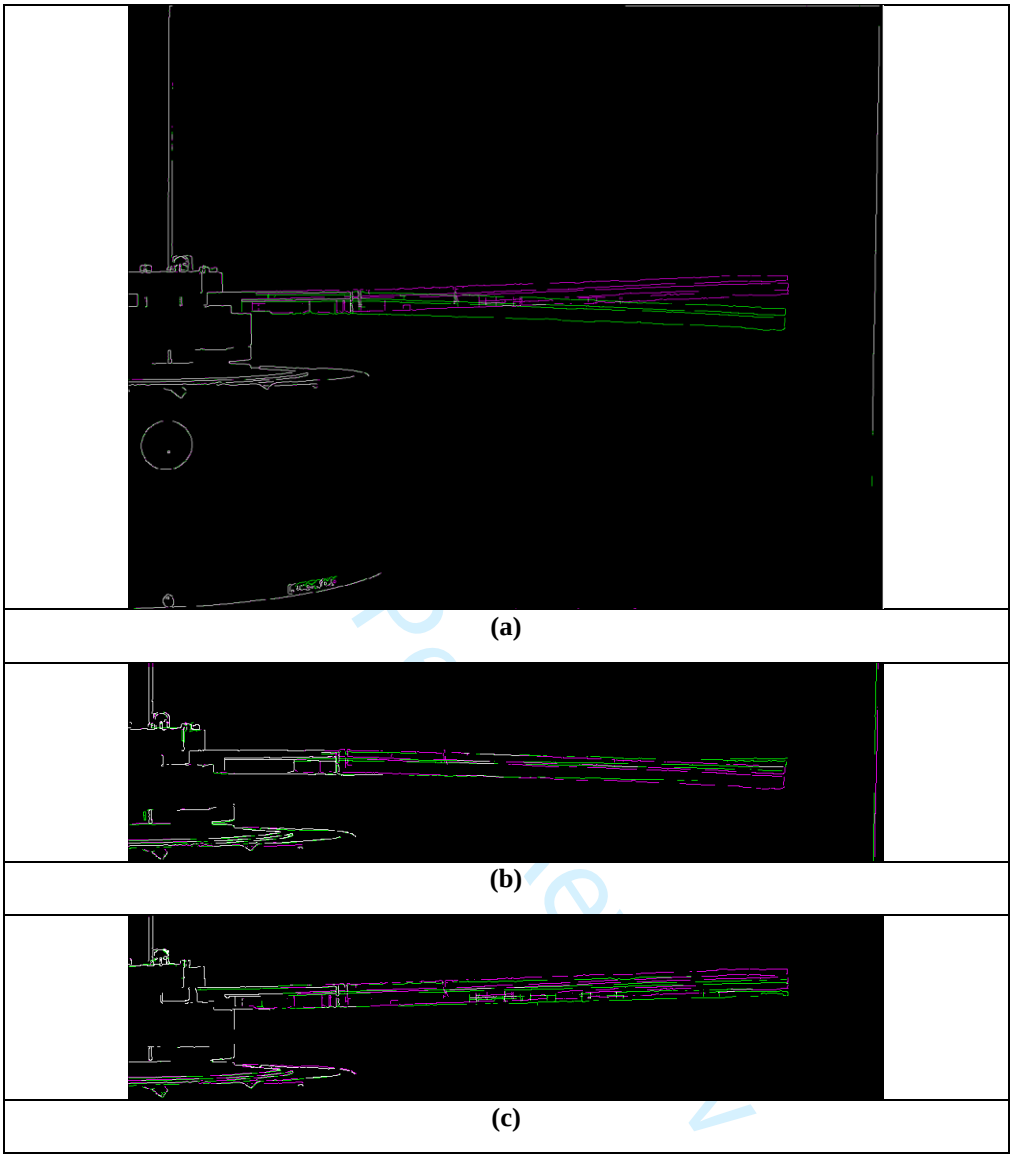


Fig 6. (a) Comparison of the maximum downward (green) and the maximum upward (purple) deflection for the most damaged case (#7). (c). For both (b) and (c), undamaged beam's edges are shown in green, damaged one's in purple. In all cases, parameters $\alpha = 16$ and $\sigma = 1$ were used.

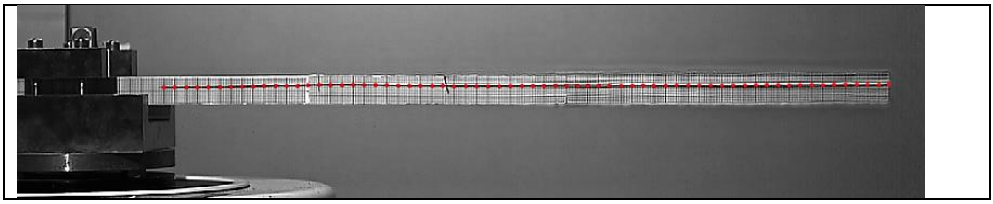


Fig 7. The 66 points considered for damage localisation, superimposed to a frame from the post-processed video ($\alpha = 4$ and $\sigma = 1$).

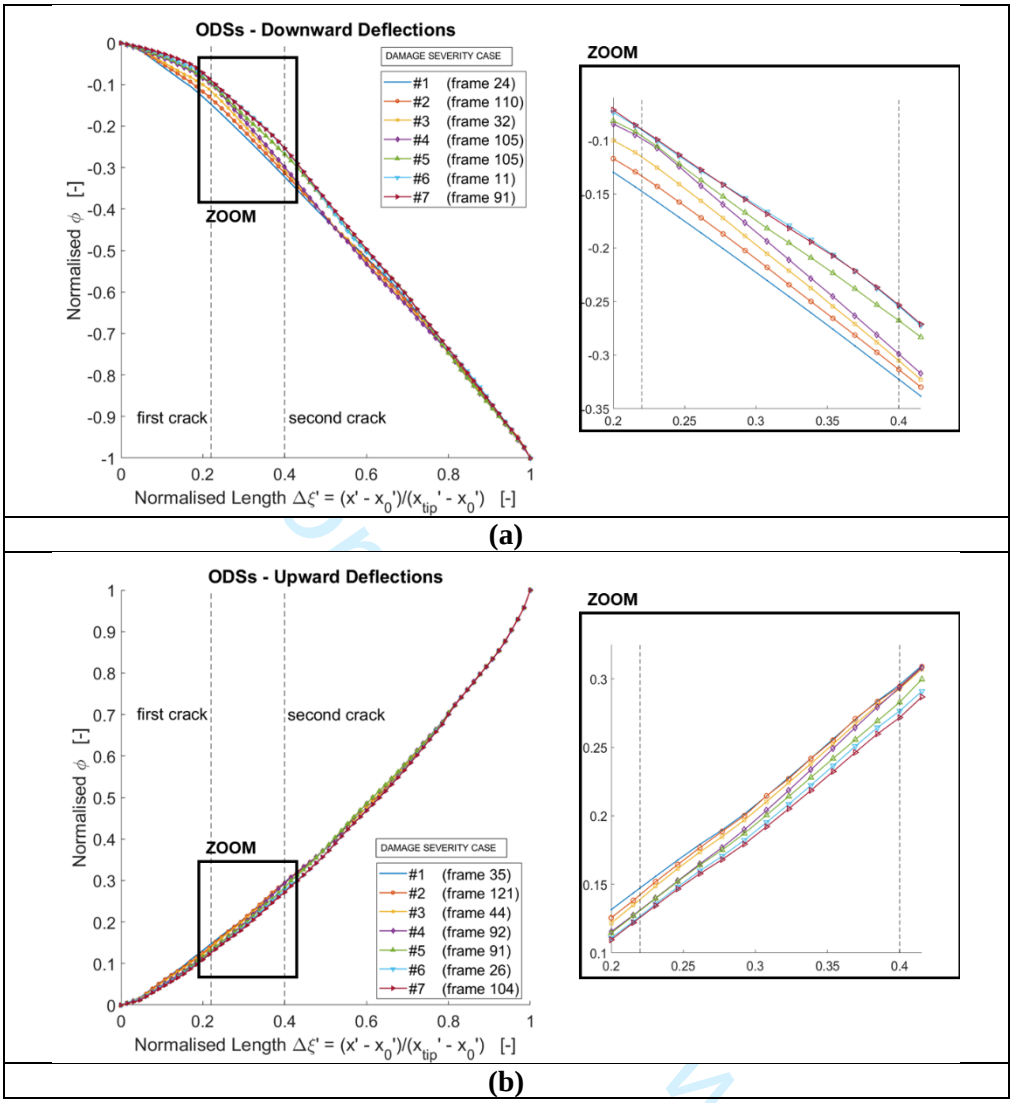


Fig 8. (a) downward and (b) upward deflections of the ODSs considered, with a zoom on the portion of most interest (respectively, (b) and (d)). The beam tract length shown is normalised between 0 ($x'_0 = 55$ mm) and 1 ($x'_{TIP} = 716\ mm$).

Table 4: MAC values for the damaged ODSs (normalised over the maximum displacement)

Damage level	MAC value (*1) (downward deflection)	MAC value (*2) (upward deflection)
	[-]	[-]
#1	> 0.99999	> 0.99999
#2	0.99983	> 0.99999
#3	0.99923	0.99994
#4	0.99824	0.99975
#5	0.99671	0.99963
#6	0.99564	0.99949
#7	0.99531	0.99933

(*1)- (*2) calculated respect to the undamaged downward and upward ODS, respectively.

It seems from the extracted ODSs that the local effects of the slanted cut are more evident than the ones of its transverse equivalent. Again, this can be a side-effect of the ODS normalisation, but it may also be due to the two slots respective positioning. Slot angle seems to influence the results nevertheless, as its prominence is more evident in downward deflections than in the upward ones; in general, upward effects seem to be strongly attenuated respect to the ones induced by downward deflections, probably due to gravity affecting the results, emphasising displacements to the downside. The two edge slots' location at the extrados may also be a cause for this double-sided behaviour.

It is crucial to remark that the two slots introduced here do *not* follow the breathing crack model, as their edges never come in contact and the structural response was linear. This aspect will be better analysed in the next subsection.

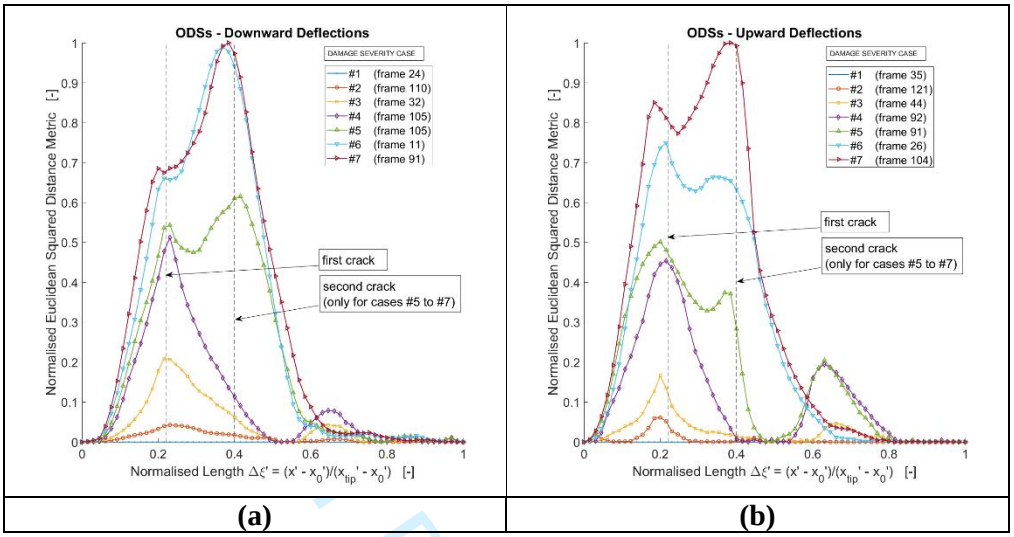
Euclidean Square Distance (ESD) was used as a baseline-based technique for damage localisation, by comparing the difference in the transverse displacements of the normalised damaged and undamaged shapes at any point; results are depicted in Fig 9. This approach was applied as a global feature on a previous work [60], where its efficacy was proved to be dependent on the damage position; thus, it has been used here to locally determine the position of the damage. The concept can be seen as a complication of the well-known Mode Shape Difference (MSD) method and it can be analytically defined as

$$\mathrm{ESD}_{\mathrm{p}} = \left|\varphi_{\mathrm{p}}^{\mathrm{d}} - \varphi_{\mathrm{p}}^{\mathrm{u}}\right|^2, \forall \mathrm{p} = 1, 2, ..., L \quad (7)$$

Where $\boldsymbol{\varphi}^d = (\varphi_1^d, \varphi_2^d, ..., \varphi_L^d)$ is the ODSs of the damaged beam, spanning over L measurement points, and φ_p^d is its p^{th} term; same applies for the undamaged baseline ($\boldsymbol{\varphi}^u$ and φ_p^u respectively). All values have been normalised to the maximum value of the most damaged case (#7) for easiness of graphical comparison. The normalised ESD proved to be functional, except for some unavoidable side effects which return false peaks close to the beam free end; this is due to the ODSs normalisation at the maximum displacement, as already noted in [60]. Nevertheless, the highest peak was in all cases corresponding to the actual slot position. These two methodologies are here applied as a basic example for merely didactic purposes; any viable mode shape-based approach

For Peer Review

would fit the aim.



(a) **(b)**

Fig 9. Normalised Euclidean Square Distance (ESD,as a metric of difference between damaged and undamaged ODSs, considering downward (a) and upward (b) deflections. The two metrics peak in correspondence of the cut cross-sections.

4.3 Visualisation of the open slot's surfaces motion

So far, the steerability of the CSPs has been basically unused, as the major interest was on transverse displacements only. However, when applied along other directions than the vertical, motion magnification allows visualising the otherwise imperceptible motions of the damage surfaces.

Two main categories of crack models exist: the already-cited breathing crack (which behaves nonlinearly) and the fully open crack. By assuming the man-made slots as representative (in vibrational terms) of a crack, the case here investigated belongs to the second group, as the slots' faces were found to never come in contact, and no perceptible nonlinearity was found in the recorded data. On the other hand, while not formally 'breathing', both notches (especially the transverse one at x_1) enlarge and reduce their surfaces' distance under the influence of both gravity and unbalance forces acting on their respective cross-sections. This mechanical behaviour, unnoticeable by the naked eye or in the pre-amplification video, is visible in the zoomed picture in Fig 10. A slight yet evident enlargement happens at every period during downward deflections. For this configuration, the camera was moved closer to the box beam, with a resulting focal length of circa 75 cm; for the given pixel resolution and amplification factor, the slot's surface actual horizontal motions are estimated to be sub millimetric.

.

For Peer Review

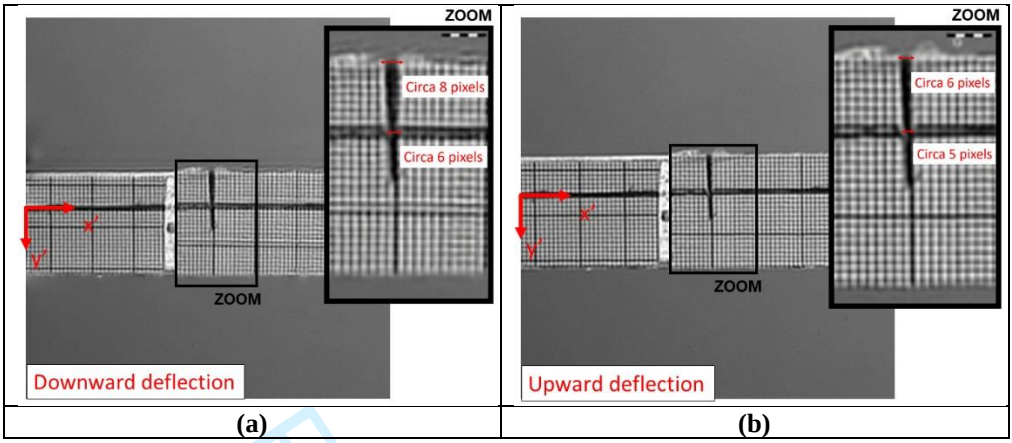


Fig 10. The magnified motion of the transverse slot's surfaces: (a) enlarging during downward deflection and (b) shrinking during the upward tracts ($\alpha = 8, \sigma = 1$).

5 Discussion and Conclusions

By looking at the discrete Fourier transform of a single frame extracted from a recorded video sequence, the amplitude component contains all the spatial frequencies of the image in the spatial domain, while the phase part brings the information needed to reorder the basis functions correctly for inverse 2-dimensional DFT. It comes obvious that altering the phase along time, the translation of the related pixels in the space domain – that is to say, the captured motion – can be manipulated. Thus, a Phase-Based approach is a well-suited candidate for motion magnification. The study presented here is intended as a systematic proof of concept for its application to Computer Vision-based Structural Health Monitoring. Several compelling and demanding configurations have been considered for this experimental investigation, such as angled, multiple damages, applied to a simple case study (a cantilevered box beam), which is well-representative of a multitude of similar, relatively slender structures.

The concept is that subtle changes in the vibrational behaviour, imperceptible by eye, become evident when the interesting temporal frequency range is amplified dozens times. By benchmarking the results from the experimental modal analysis (eigenfrequencies and ODSs) against the pristine data, damage detection, sizing and localisation have been proved feasible (with due care). Displacements THs were extracted from a chosen point and used for damage detection and sizing, resorting to a comparison in the frequency domain with the pristine baseline. ODSs, extrapolated via quarter-octave Complex Steerable Pyramids and Canny edge detection, were used again for damage severity assessment (by comparing the MAC value between any of them with the undamaged case) and for damage localisation. The amplified videos allowed even to perceive the movements of the edge slots' faces.

Specifically, it was found that fissures as small as a 1 mm-long indentation on the 25x25 mm hollow cross-section of the box beam can produce noticeable effects on its frequency response; a 5 mm-long saw-cut generated appreciable deviation in the damaged ODS respect to the pristine baseline. These results were obtained in a controlled laboratory environment and assuming notches as a good approximation of the vibrational behaviour of actual cracks. Less deep saw-cuts went unnoticed to the techniques applied to the motion magnified frames. Damage sizing was also proved feasible, at least for the single damage case and for a transverse-cracked beam.

The benefits of the technique are evident. With no added masses or stiffening due to screwing/attaching sensors, video recording is non-contact and thus completely non-invasive. Therefore, in the absence of mounted transducers, the pure dynamic response of the damaged or pristine structure can be investigated 'as is'. No wiring is needed, and no

For Peer Review

power signal transmission issues may arise.

If proper care is used in the acquisition, the technique may be implemented with relative simplicity. To the current state, experimental results may be not as accurate as with other more traditional techniques such as laser vibrometer and attached accelerometers. Nevertheless, other advantages make video processing competitive with the established techniques to rapidly acquire a large amount of data with high spatial and temporal resolution, even for weakly-excited structures. For instance, unlike single-point LDV, video cameras allow capturing the whole structure of interest, or at least a large, continuous part of it, along any direction belonging to the focal plane. Multipoint LDV would produce more accurate ODSs, but not necessarily denser, and at cost of longer experimental procedures, more post-processing, and costlier instrumentations. Using better video equipment with superior specifics, the accuracy of video acquired data can be increased significantly.

Results showed a good agreement between the frequencies extracted via laser and video acquisition, with discrepancies between the measurements never larger than 7% even in the worst cases and less than 1% in the best ones. PBMM specifically overcomes the issues of other similar visual enhancement strategies based on motion amplification, both Eulerian and Lagrangian.

The remaining main issue is to efficiently mine all the data captured by the video camera. The volume of the acquired information becomes easily very high when increasing the frame-per-second rate and/or the pixel density. Cropping the image in order to isolate the interesting part of the structure can improve the efficiency of the procedure; nevertheless, one should always remember that reducing the number of pixels also decrease the range of spatial frequencies which can be studied. State-of-the-art model order reduction tools for image compression can be used to reduce the dimensionality of the problem, speeding up the computational process, with an acceptably small decrease in accuracy; big data analysis techniques might turn out to be useful. This may eventually lead to an application downloadable from the internet and applicable on commercial mobile devices with consumer-grade cameras. That leaves large room for improvement and further researches.

For Peer Review

Acknowledgements

The authors would like to thank Dr Ivan Petrunin for his help with the experimental setup, Dr Mudassir Lone for kindly providing the High-Speed camera, and doctoral student Alessandro Pontillo for his technical support. The First Author would like to thank Cranfield University and Politecnico di Torino for their support and for allowing this joint research project.

Bibliography

1. Farrar, C.R., and Worden, K. (2013) *Structural Health Monitoring: A Machine Learning Perspective*.

2. Stanbrifge, A.B., and Ewins, D.J. (1999) Modal Testing Using a Scanning Laser Doppler Vibrometer. *Mech. Syst. Signal Process.*, **13** (2), 255–270.

3. Salawu, O.S. (1997) Detection of structural damage through changes in frequency: a review. *Eng. Struct.*, **19** (9), 718–723.

4. Civera, M., Surace, C., and Worden, K. (2017) Detection of Cracks in Beams Using Treed Gaussian Processes, Springer, Cham, pp. 85–97.

5. Corrado, N., Durrande, N., Gherlone, M., Hensman, J., Mattone, M., and Surace, C. (2018) Single and multiple crack localization in beam-like structures using a Gaussian process regression approach. *J. Vib. Control*, **24** (18), 4160–4175.

6. Surace, C., Archibald, R., and Saxena, R. (2013) On the use of the polynomial annihilation edge detection for locating cracks in beam-like structures. *Comput. Struct.*, **114**–**115**, 72–83.

7. Corrado, N., Gherlone, M., Surace, C., Hensman, J., and Durrande, N. (2015) Damage localisation in delaminated composite plates using a Gaussian process approach. *Meccanica*, **50** (10), 2537–2546.

8. Surace, C., Saxena, R., Gherlone, M., and Darwich, H. (2014) Damage localisation in plate like-structures using the two-dimensional polynomial annihilation edge detection method. *J. Sound Vib.*, **333** (21), 5412–5426.

9. Rytter, A. (1993) Vibrational Based Inspection of Civil Engineering Structures.

10. Civera, M., Zanotti Fragonara, L., and Surace, C. (2017) A novel approach to damage localisation based on bispectral analysis and neural network. *Smart Struct. Syst.*, **20** (6), 669–682.

11. Schmidt, T., Tyson, J., and Galanulis, K. (2003) Full-Field Dynamic Displacement and Strain Meaurement-Specific Examples Using Advanced 3D Image Correlation Photogrammetry: Part II. *Exp. Tech.*, **27** (4), 22–26.

12. Schmidt, T., Tyson, J., and Galanulis, K. (2003) Full-Field Dynamic Displacement and Strain Meaurement-Specific Examples Using Advanced 3D Image Correlation Photogrammetry: Part 1. *Exp. Tech.*, **27** (3), 47–50.

13. Schreier, H., Orteu, J.-J., and Sutton, M.A. (2009) *Image Correlation for Shape, Motion and Deformation Measurements*, Springer US, Boston, MA.

14. Warren, C., Niezrecki, C., Avitabile, P., and Pingle, P. (2011) Comparison of FRF measurements and mode shapes determined using optically image based, laser, and accelerometer measurements. *Mech. Syst. Signal Process.*, **25** (6), 2191–2202.

15. Civera, M., Zanotti Fragonara, L., and Surace, C. Video Processing Techniques for the Contactless Investigation of Large Oscillations. *Proc. AIVELA XXVI Meet. (in Publ.*

16. Civera, M., Zanotti Fragonara, L., Surace, C., Civera, M., Zanotti Fragonara, L., and Surace, C. (2019) Using Video Processing for the Full-Field Identification of Backbone Curves in Case of Large Vibrations. *Sensors*, **19** (10), 2345.

17. Morlier, J., Salom, P., and Bos, F. (2007) New Image Processing Tools for Structural Dynamic Monitoring. *Key Eng. Mater.*, **347**, 239–244.

18. Lee, J.J., Fukuda, Y., Shinozuka, M., Cho, S., and Yun, C.-B. (2007) Development and application of a vision-based displacement measurement system for structural health monitoring of civil structures. *Smart Struct. Syst.*, **3** (3), 373–384.

19. Helfrick, M.N., Niezrecki, C., Avitabile, P., and Schmidt, T. (2011) 3D digital image correlation methods for full-field vibration measurement. *Mech. Syst. Signal Process.*, **25** (3), 917–927.

20. Baqersad, J., Niezrecki, C., and Avitabile, P. (2015) Extracting full-field dynamic strain on a wind turbine rotor subjected to arbitrary excitations using 3D point tracking and a modal expansion technique. *J. Sound Vib.*, **352**, 16–29.

21. Olaszek, P. (1999) Investigation of the dynamic characteristic of bridge structures using a computer vision method. *Measurement*, **25** (3), 227–236.

22. Wahbeh, A.M., Caffrey, J.P., and Masri, S.F. (2003) A vision-based approach for the direct measurement of displacements in vibrating systems. *Smart Mater. Struct.*, **12** (5), 785–794.

23. Caetano, E.F., Silva, S., and Bateira, J. (2007) Application of a vision system to

For Peer Review

the monitoring of cable structures.

24. Feng, M.Q., Asce, F., Fukuda, Y., Feng, D., Asce, S.M., and Mizuta, M. (2015) Nontarget Vision Sensor for Remote Measurement of Bridge Dynamic Response.

25. Dematteis, N., Giordan, D., Zucca, F., Luzi, G., and Allasia, P. (2018) 4D surface kinematics monitoring through terrestrial radar interferometry and image cross-correlation coupling. *ISPRS J. Photogramm. Remote Sens.*, **142**, 38–50.

26. Yoon, H., Elanwar, H., Choi, H., Golparvar-Fard, M., and Spencer, B.F. (2016) Target-free approach for vision-based structural system identification using consumer-grade cameras. *Struct. Control Heal. Monit.*, **23** (12), 1405–1416.

27. Khuc, T., and Catbas, F.N. (2017) Completely contactless structural health monitoring of real-life structures using cameras and computer vision. *Struct. Control Heal. Monit.*, **24** (1), e1852.

28. Feng, D., and Feng, M.Q. (2018) Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review. *Eng. Struct.*, **156**, 105–117.

29. Spencer, B.F., Hoskere, V., and Narazaki, Y. (2019) Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering*, **5** (2), 199–222.

30. Nonis, C., Niezrecki, C., Yu, T.-Y., Ahmed, S., Su, C.-F., and Schmidt, T. (2013) Structural health monitoring of bridges using digital image correlation. 869507.

31. Malesa, M., Szczepanek, D., Kujawińska, M., Świercz, A., and Kołakowski, P. (2010) Monitoring of civil engineering structures using Digital Image Correlation technique. *EPJ Web Conf.*, **6**, 31014.

32. Koch, C., Georgieva, K., Kasireddy, V., Akinci, B., and Fieguth, P. (2015) A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv. Eng. Informatics*, **29** (2), 196–210.

33. Tzitzilonis, V., Malandrakis, K., Zanotti Fragonara, L., Gonzalez Domingo, J.A., Avdelidis, N.P., Tsourdos, A., and Forster, K. (2019) Inspection of Aircraft Wing Panels Using Unmanned Aerial Vehicles. *Sensors*, **19** (8), 1824.

34. Niezrecki, C., Baqersad, J., and Sabato, A. (2019) Digital Image Correlation Techniques for NDE and SHM, in *Handbook of Advanced Nondestructive Evaluation*, Springer International Publishing, Cham, pp. 1545–1590.

35. Wadhwa, N., Rubinstein, M., Durand, F., and Freeman, W.T. (2013) Phase-based video motion processing. *ACM Trans. Graph.*, **32** (4), 1.

36. Simoncelli, E.P., and Freeman, W.T. (1995) The Steerable Pyramid: A Flexible Architecture For Multi-Scale Derivative Computation. **III**.

37. Chen, J.G., Wadhwa, N., Durand, F., Freeman, W.T., and Buyukozturk, O. (2015) Developments with Motion Magnification for Structural Modal Identification Through Camera Video, Springer, Cham, pp. 49–57.

38. Chen, J.G., Wadhwa, N., Cha, Y.-J., Durand, F., and Freeman, W.T. (2015) Modal identification of simple structures with high-speed video using motion magnification.

39. Liu, C., Torralba, A., Freeman, W.T., Durand, F., and Adelson, E.H. (2005) Motion Magnification. *ACM Trans. Graph.*, **24** (3), 519–526.

40. Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., and Freeman, W. Eulerian Video Magnification for Revealing Subtle Changes in the World.

For Peer Review

41. Horn, B.K.P., and Schunck, B.G. (1981) Determining optical flow. *Artif. Intell.*, **17** (1–3), 185–203.

42. Fleet, D.J., and Jepson, A.D. (1990) Computation of component image velocity from local phase information. *Int. J. Comput. Vis.*, **5** (1), 77–104.

43. Wadhwa, N., Chen, J.G., Sellonc, J.B., Wei, D., Rubinstein, M., Ghaffari, R., Freeman, D.M., Buyukozturk, O., Wang, P., Sun, S., Kang, S.H., Bertoldi, K., Durand, F., and Freeman, W.T. (2017) Motion microscopy for visualizing and quantifying small motions. *Proc. Natl. Acad. Sci.*, **114** (44), 11639–11644.

44. Christensen, O. A short introduction to frames, Gabor systems, and wavelet systems.

45. Freeman, W.T., and Adelson, E.H. (1991) The Design and Use of Steerable Filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, **13** (9), 891–906.

46. http://people.csail.mit.edu/nwadhwa/phase-video/. *Accessed 4 April 2019*.

47. http://people.csail.mit.edu/nwadhwa/motion-microscope/. *Accessed 4 April 2019*.

48. Canny, J. (1986) A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-8** (6), 679–698.

49. Yang, Y., Dorn, C., Mancini, T., Talken, Z., Kenyon, G., Farrar, C., and Mascareñas, D. (2017) Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. *Mech. Syst. Signal Process.*, **85**, 567–590.

50. Poozesh, P., Sarrafi, A., Mao, Z., Avitabile, P., and Niezrecki, C. (2017) Feasibility of extracting operating shapes using phase-based motion magnification technique and stereo-photogrammetry. *J. Sound Vib.*, **407**, 350–366.

51. Oh, T.-H., Jaroensri, R., Kim, C., Elgharib, M., Durand, F., Freeman, W.T., and Matusik, W. Learning-based Video Motion Magnification.

52. Yunus, E.H., Gulan, U., Holzner, M., and Chatzi, E. (2019) A Novel Approach for 3D-Structural Identification through Video Recording: Magnified Tracking. *Sensors*, **19** (5).

53. Sarrafi, A., Poozesh, P., Niezrecki, C., and Mao, Z. (2017) Mode extraction on wind turbine blades via phase-based video motion estimation. *Proc. SPIE 10171, Smart Mater. Nondestruct. Eval. Energy Syst. 2017*.

54. Ju, F.D., and Mimovich, M.E. (1988) Experimental diagnosis of fracture damage in structures by the modal frequency method. *J. Vib. Acoust. Stress. Reliab. Des.*, **110** (4), 456–463.

55. Rytter, A., Krawczuk, M., and Kirkegaard, P.H. (2000) Experimental and Numerical Study of Damaged Cantilever. */ J. Eng. Mech.*, **126** (1), 60–65.

56. Praisach, Z.-I., Gillich, G.-R., and Minda, A.-A. *Relative Frequency Shift Curves Fitting Using FEM Modal Analyses*.

57. Morassi, A., and Rollo, M. (2001) Identification of Two Cracks in a Simply Supported Beam from Minimal Frequency Measurements. *J. Vib. Control*, **7** (5), 729–739.

58. Pook, L.P. (1971) The effect of crack angle on fracture toughness. *Eng. Fract. Mech.*, **3** (3), 205–218.

59. Behera, R.K., Pandey, A., and Parhi, D.R. (2014) Numerical and Experimental Verification of a Method for Prognosis of Inclined Edge Crack in Cantilever Beam based on Synthesis of Mode Shapes. *Procedia Technol.*, **14**, 67–74.

For Peer Review

60. Martucci, D., Civera, M., Surace, C., and Worden, K. (2018) Novelty Detection in a Cantilever Beam using Extreme Function Theory. *J. Phys. Conf. Ser.*, **1106** (1), 012027.

For Peer Review

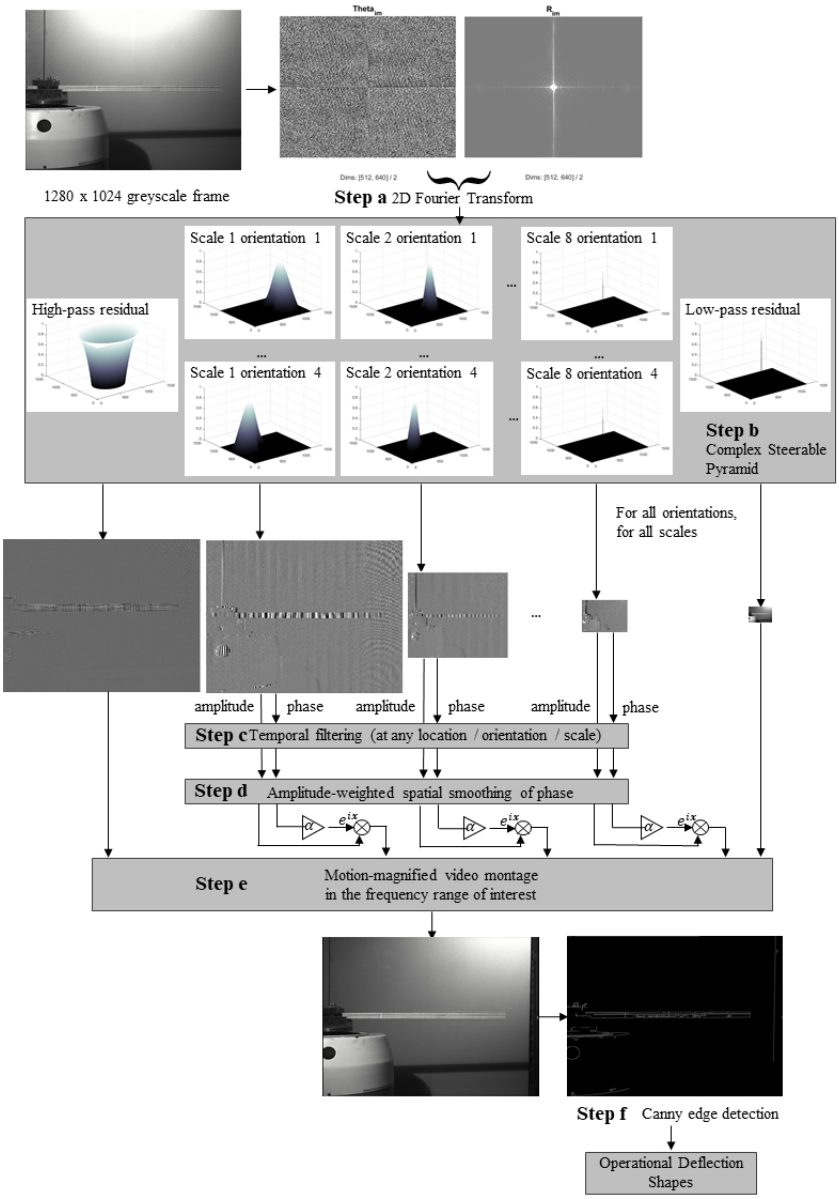


Fig 1. Scheme of the Motion-Magnification and ODSs extraction procedure.

210x297mm (96 x 96 DPI)

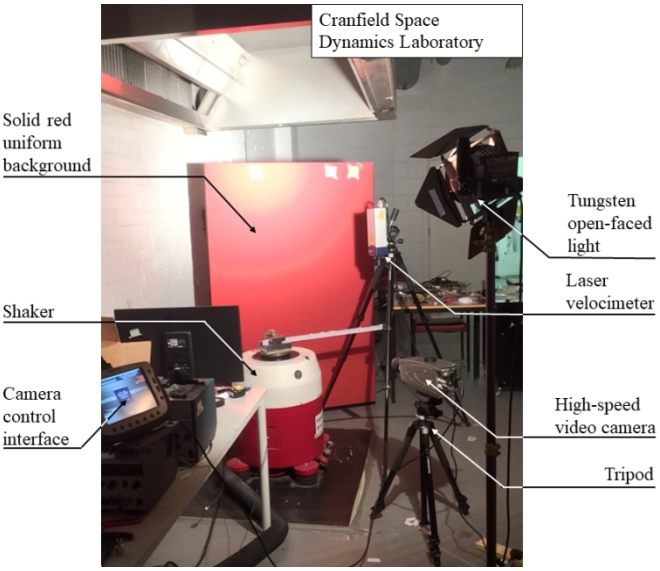


Fig 2. Experimental Setup.

338x190mm (96 x 96 DPI)

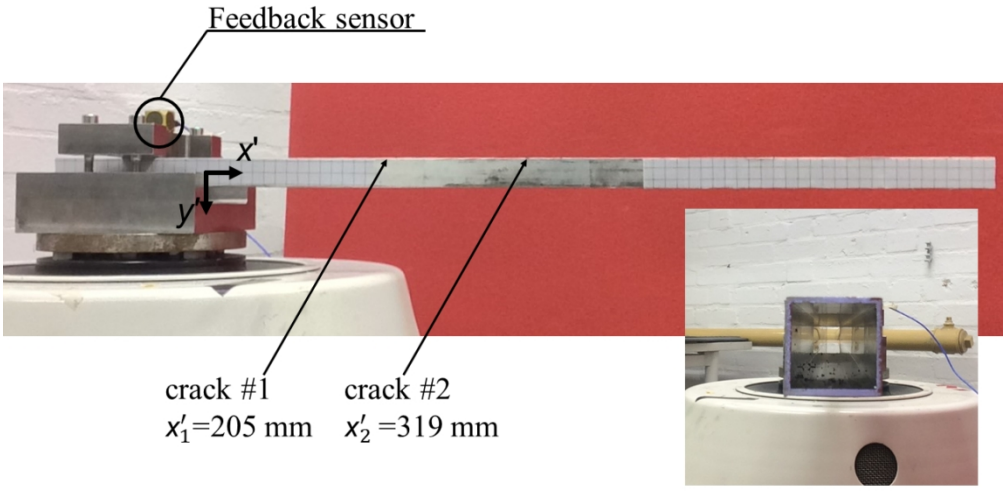


Fig 3. Box beam lateral view and hollow cross-section. The beam reference frame {O,x',y'} is also shown; black arrows point to the intersections between the edge slots and the edge itself.

338x190mm (96 x 96 DPI)

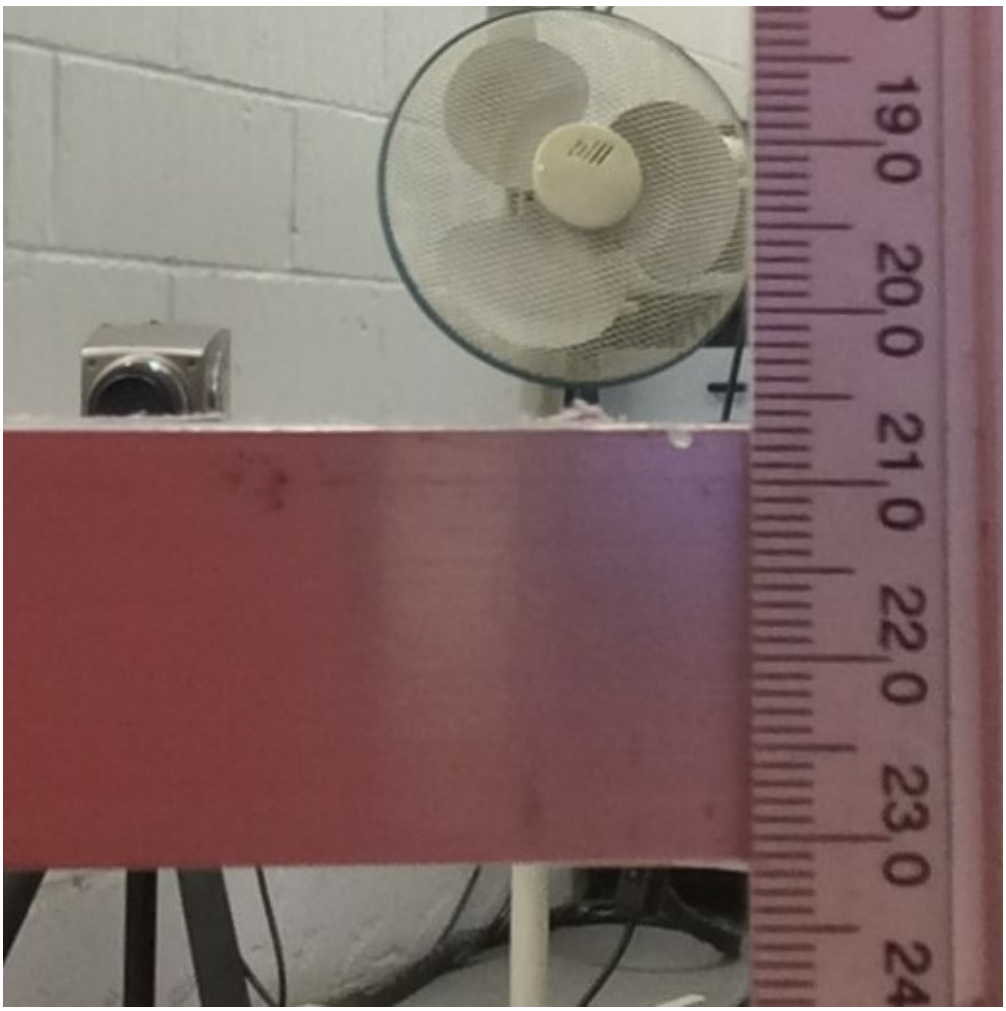
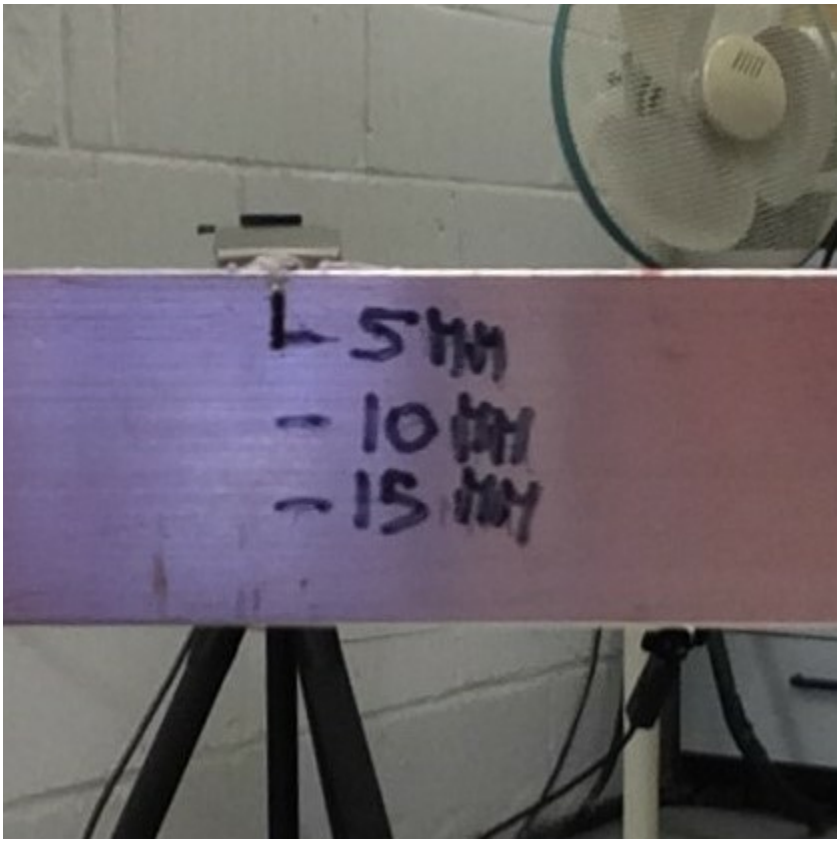
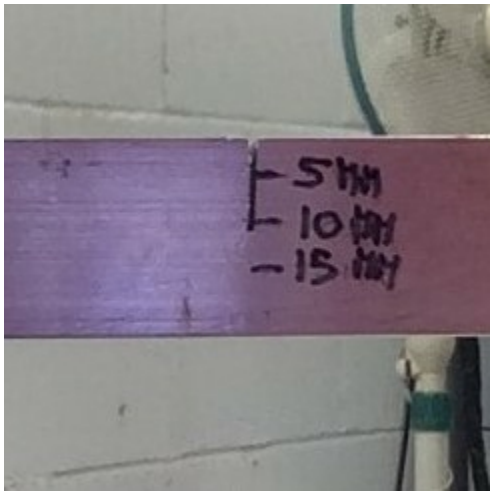
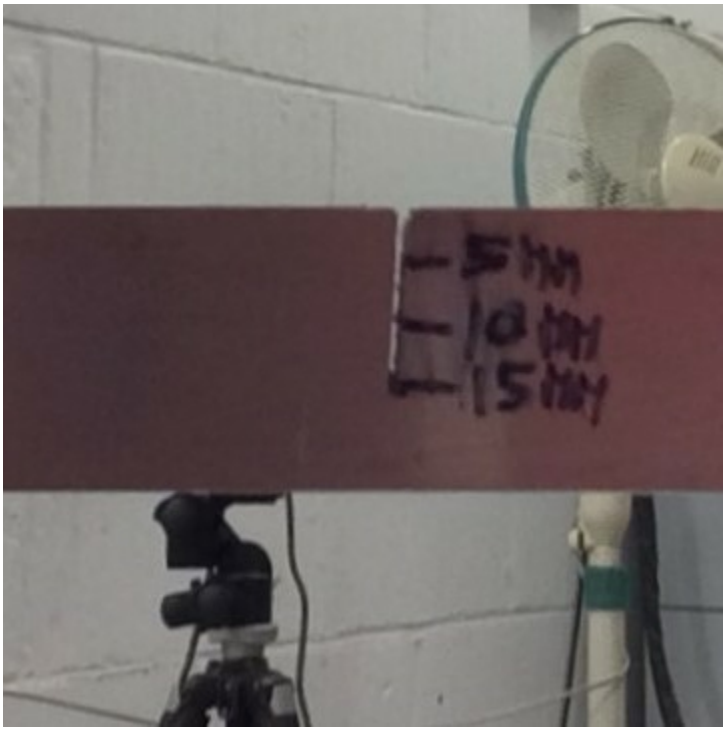
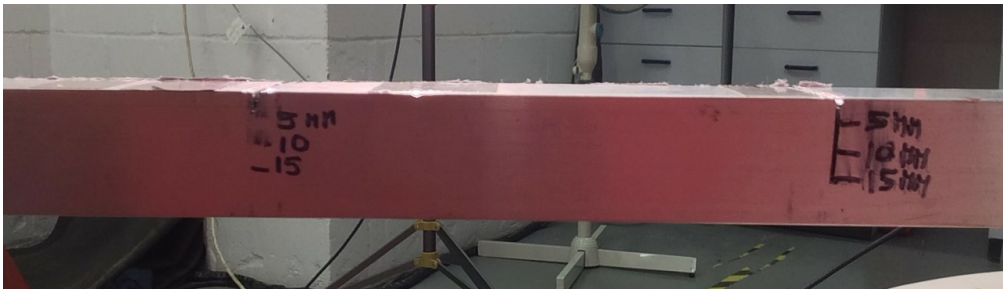


Fig 4. (a – g) increasing damage conditions: in the same order, case #1, #2, #3, #4, #5, #6, and #7 (see Table 2 for better description).


L 5mm
- 10mm
- 15mm


5mm
10mm
15mm

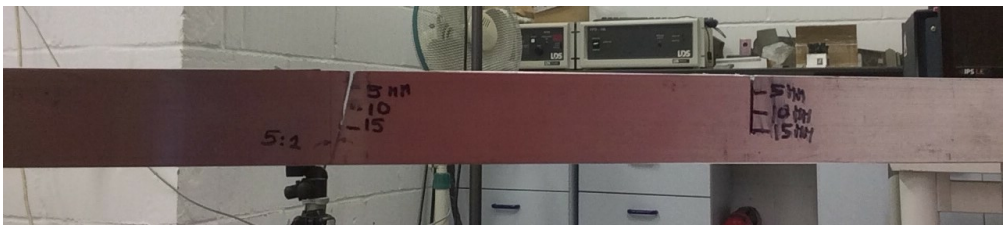

5 MM
10 MM
15 MM



162x46mm (200 x 200 DPI)



155x38mm (200 x 200 DPI)



162x35mm (200 x 200 DPI)

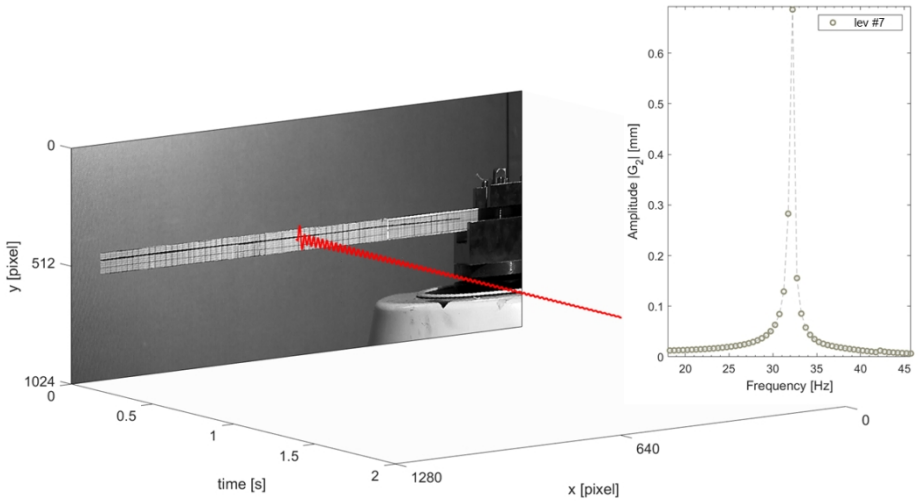


Fig 5. A conceptual sketch of the TH extraction and Fourier Transform process.

338x190mm (96 x 96 DPI)

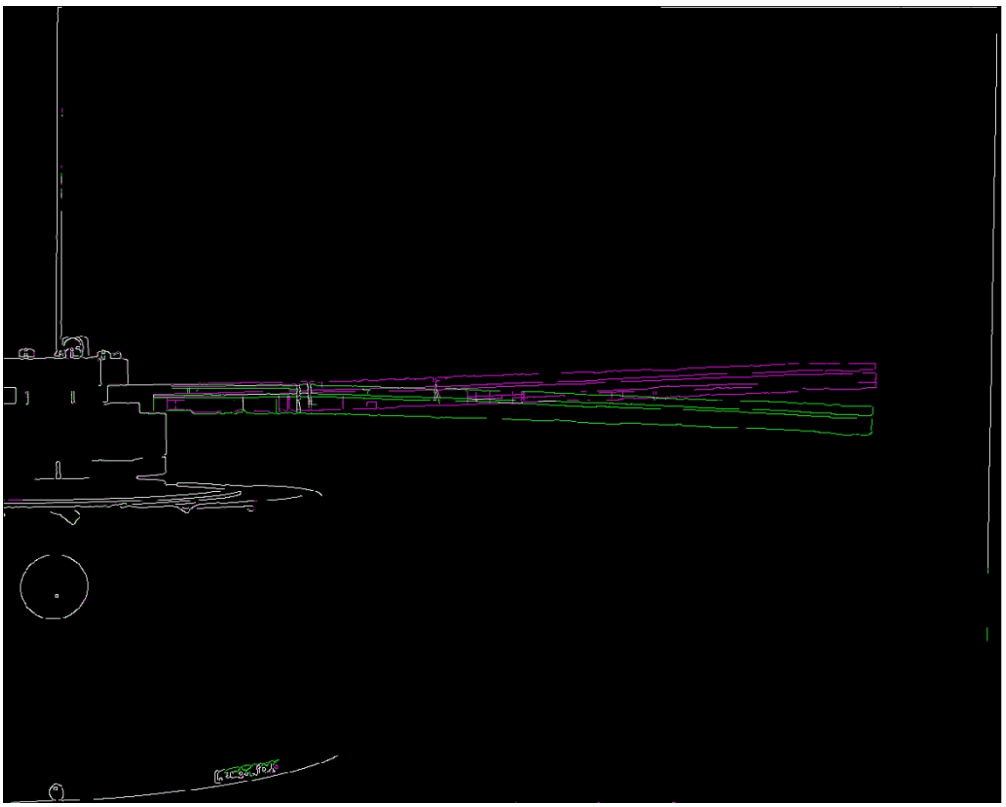
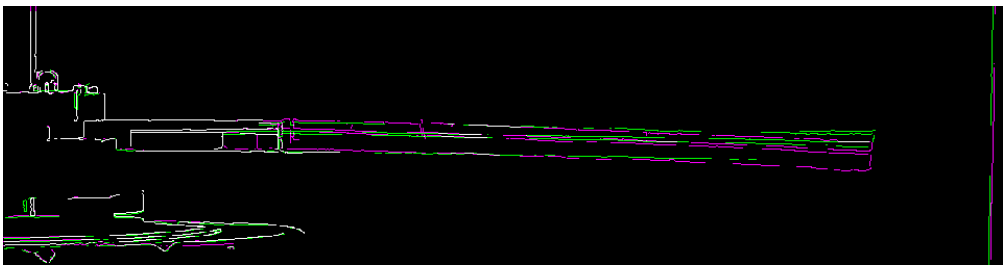
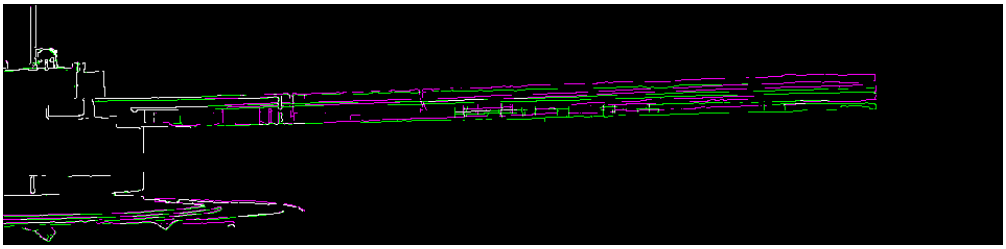


Fig 6. (a) Comparison of the maximum downward (green) and the maximum upward (purple) deflection for the most damaged case (#7). (c). For both (b) and (c), undamaged beam's edges are shown in green, damaged one's in purple. In all cases, parameters α=16 and σ=1 were used.





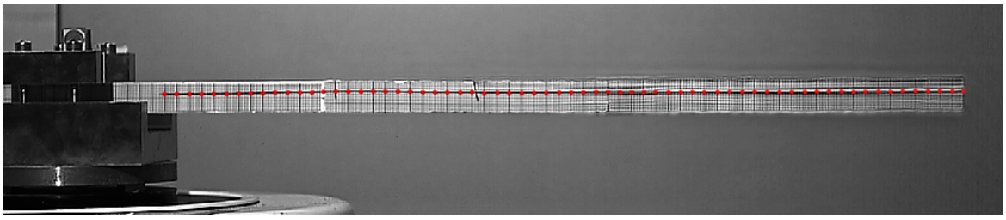


Fig 7. The 66 points considered for damage localisation, superimposed to a frame from the postprocessed video (α = 4 and σ = 1).

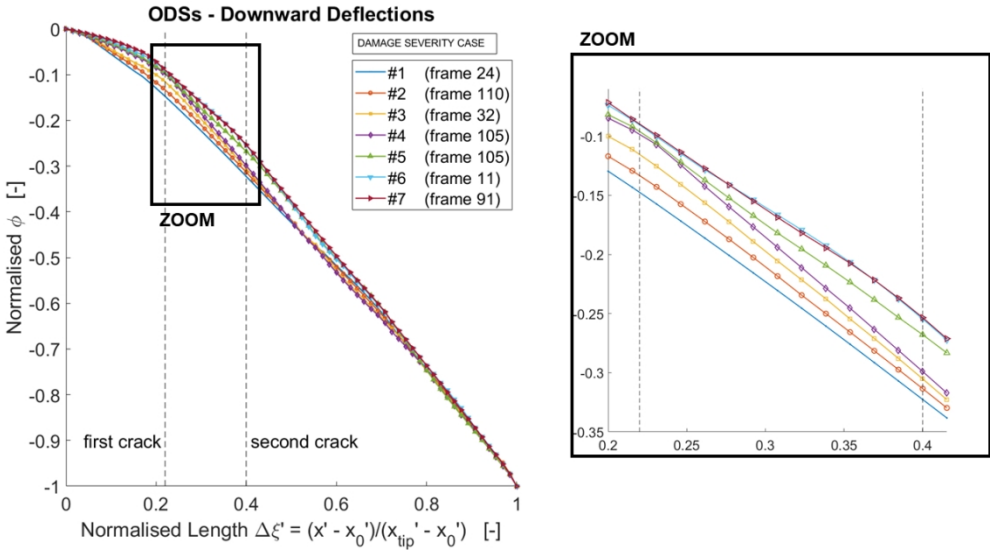
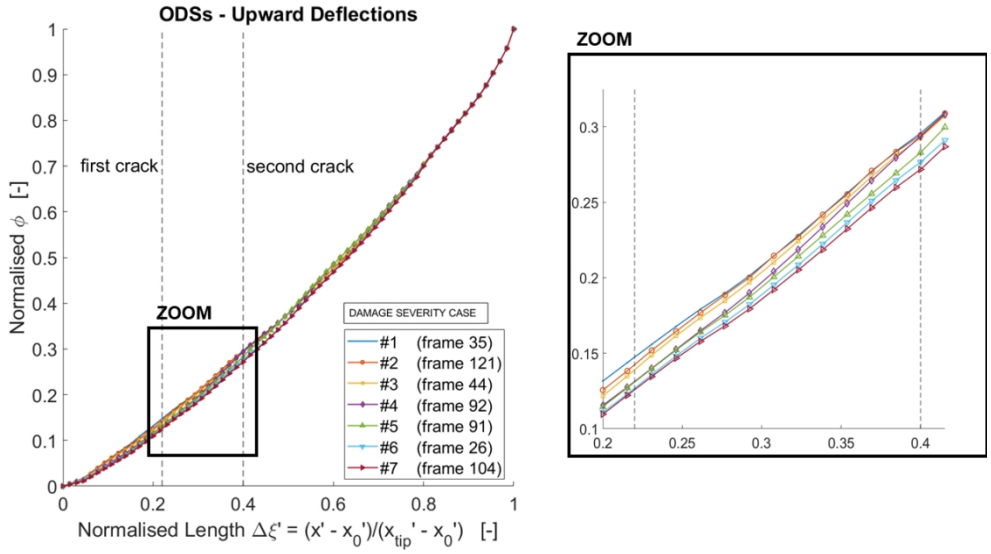


Fig 8. (a) downward and (b) upward deflections of the ODSs considered, with a zoom on the portion of most interest. The beam tract length shown is normalised between 0 (x'_0 = 55 mm) and 1 (x'_{TIP} = 716 mm).

338x190mm (96 x 96 DPI)



338x190mm (96 x 96 DPI)

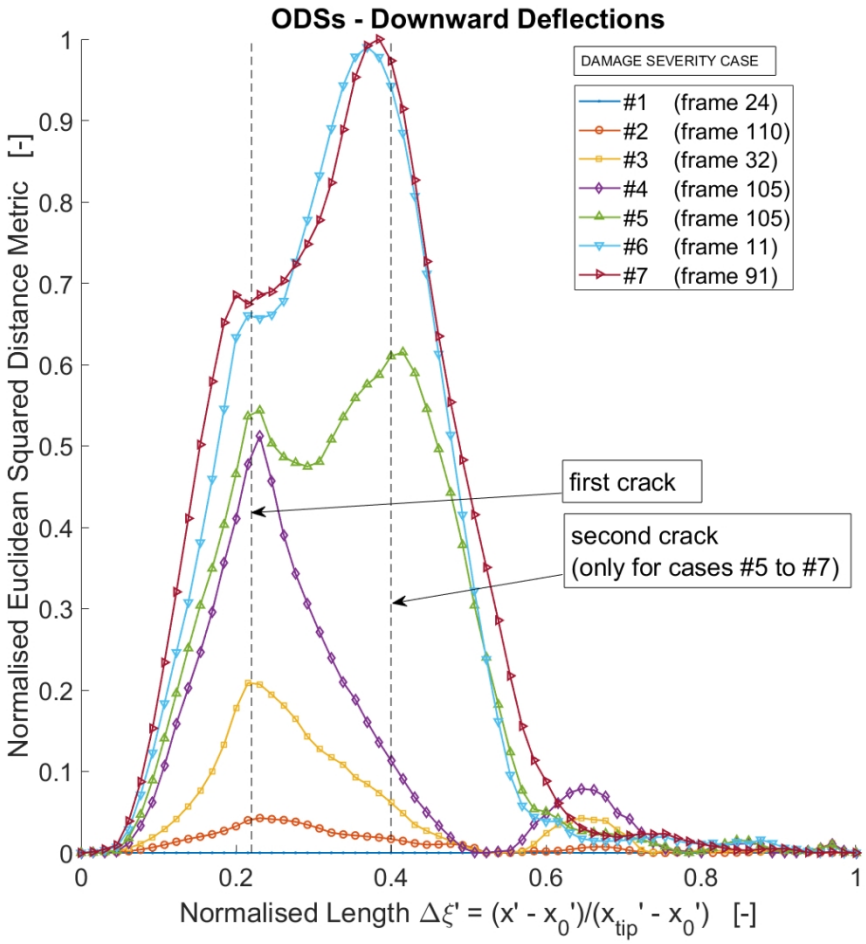
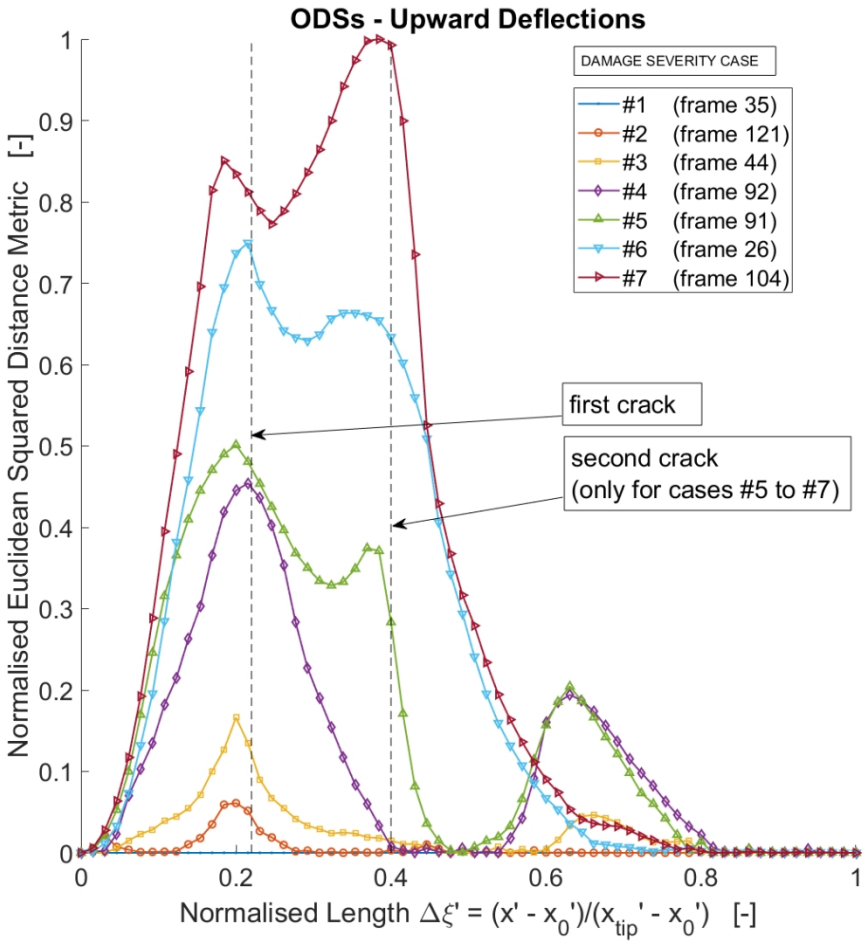


Fig 9. Normalised Euclidean Square Distance (ESD) as a metric of difference between damaged and undamaged ODSs, considering downward (a) and upward (b) deflections. The two metrics peak in correspondence of the cut cross-sections.

317x316mm (96 x 96 DPI)



317x316mm (96 x 96 DPI)

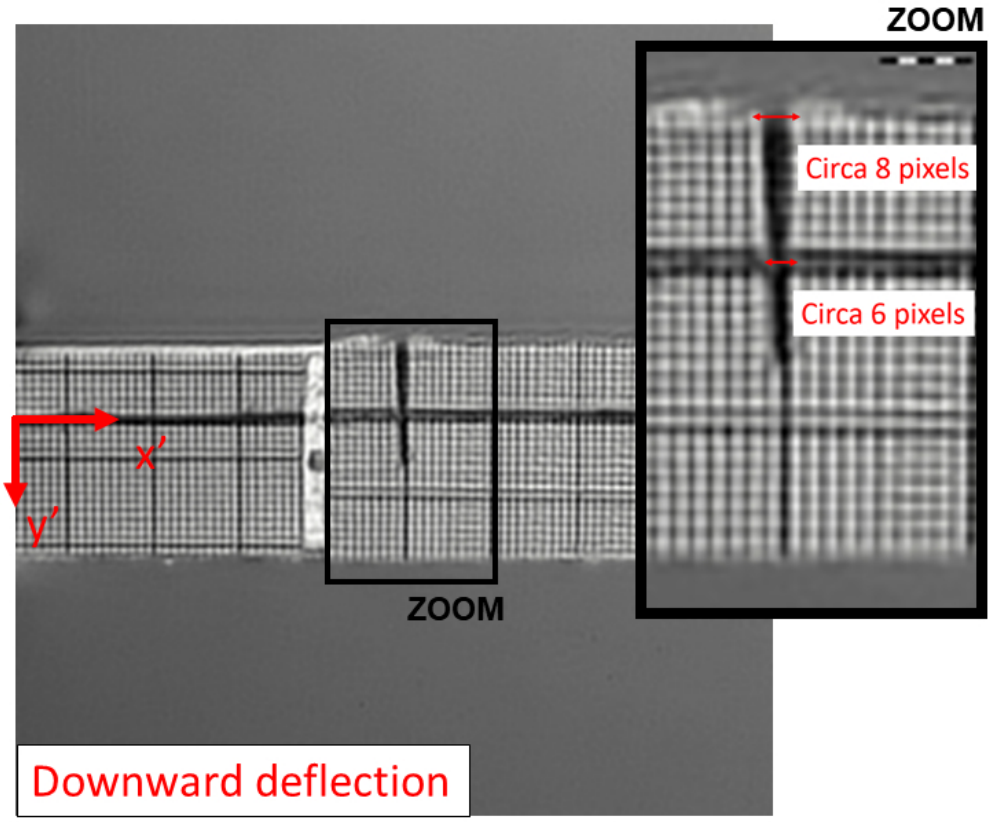
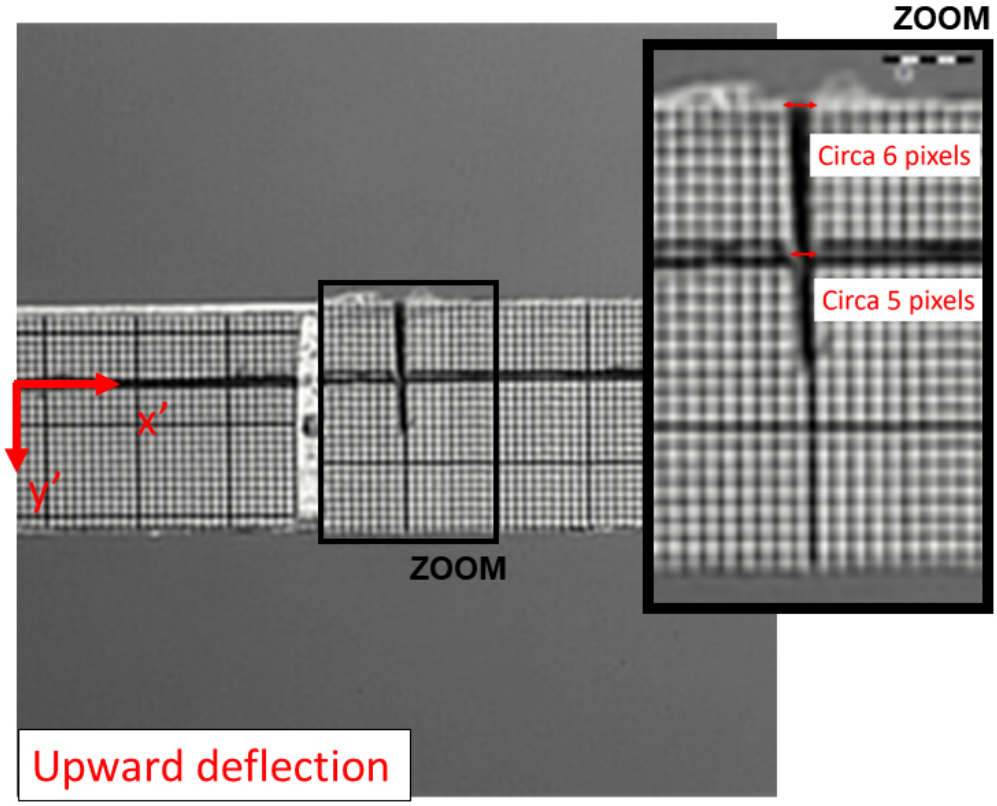


Fig 10. The magnified motion of the transverse slot's surfaces: (a) enlarging during downward deflection and (b) shrinking during the upward tracts (α = 8, σ = 1).

192x160mm (96 x 96 DPI)



193x157mm (96 x 96 DPI)

Table 1: Material and geometrical properties of the aluminium bar

Young Modulus E	$69\ GPa$
Mass per unit of length μ	$0.322\ kg/m$ (Density ρ $2170\ kg/m^3$)
Free Length L	$716\ mm$
Side (W=H)	$25\ mm$
Thickness t	$2\ mm$
Cross-section A	$184\ mm^2$
Moment of inertia I	$1.63 \cdot 10^4\ \ mm^4$

For Peer Review

Table 2: Damage Levels

Damage severity case (reference #)	damage depth (*)	
	d_1	d_2
no damage	-	-
#1	1 mm	no damage
#2	5 mm	no damage
#3	10 mm	no damage
#4	15 mm	no damage
#5	15 mm	5 mm
#6	15 mm	10 mm
#7	15 mm	15 mm

(*) intended as the vertical component of length; corresponding to damage extension for $l_1 = d_1$ but not for $l_2 = d_2 / \cos(\beta) > d_2$.

For Peer Review

Table 3: First natural frequency, as extracted by the laser tachometer from RN tests and through video PBMM from impulse tests.

damage level	laser velocimeter (*) (random noise)	Frequency shift	video PBMM (*) (impulse)	Frequency shift
	[Hz]	[%]	[Hz]	[%]
no damage	46.3	-	46.3	-
#1	45.8	-1.1	46.3	~0
#2	44.7	-3.5	45.7	-1.3
#3	42.2	-8.9	40.7	-12.1
#4	37.0	-20.1	34.8	-24.8
#5	36.7	-20.7	34.5	-25.5
#6	36.1	-22.0	33.8	-27.0
#7	34.4	-25.7	32.2	-30.5

(*) data recorded at $x' = L/2$

For Peer Review

Table 4: MAC values for the damaged ODSs (normalised over the maximum displacement)

Damage level	MAC value (*[1]) (downward deflection)	MAC value (*[2]) (upward deflection)
	[-]	[-]
#1	> 0.99999	> 0.99999
#2	0.99983	> 0.99999
#3	0.99923	0.99994
#4	0.99824	0.99975
#5	0.99671	0.99963
#6	0.99564	0.99949
#7	0.99531	0.99933

(*[1])- (*[2]) calculated respect to the undamaged downward and upward ODS, respectively.

For Peer Review