Using a smartphone camera to analyse rotating and vibrating systems: Feedback on the SURVISHNO 2019 contest

Hugo André, Q Leclère, D. Anastasio, Y. Benaïcha, K. Billon, M. Birem, F. Bonnardot, Z. Y. Chin, F. Combet, P.J. J Daems, et al.

Using a smartphone camera to analyse rotating and vibrating systems: Feedback on the SURVISHNO 2019 contest.

H. André[a,*], Q. Leclère[b,*], D. Anastasio[d], Y. Benaïcha[g], K. Billon[j], M. Birem[h], F. Bonnardot[a], Z.Y. Chin[l], F. Combet[f], P.J. Daems[e], A.P. Daga[d], R. De Geest[h], B. Elyousfi[a], J. Griffaton[c], K. Gryllias[i,k], Y. Hawwari[b], J. Helsen[e], F. Lacaze[j], L. Laroche[b,j], X. Li[j], C. Liu[i,k], A. Mauricio[i,k], A. Melot[g], A. Ompusunggu[h], G. Paillot[j], S. Passos[j], C. Peeters[e], M. Perez[j], J. QI[i,k], E.F. Sierra-Alonso[b,m], W.A. Smith[l], X. Thomas[f]

[a] *Université de Lyon, Université Jean Monnet de Saint-Etienne, LASPI EA3059, 20 Avenue de Paris, 42334 Roanne Cedex, France*
[b] *Laboratoire Vibrations Acoustique, Univ Lyon, INSA-Lyon, LVA EA677, F-69621 Villeurbanne, France*
[c] *Safran Aircraft Engines, Rond Point Rene Raveau - Reau, 77550 Moissy-Cramayel, France*
[d] *Dipartimento di Ingegneria Meccanica e Aerospaziale, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*
[e] *Vrije Universiteit Brussel, Department of Applied Mechanics, Pleinlaan 2, 1050 Brussels, Belgium*
[f] *DYNAE, rue Condorcet, 38090 Villefontaine, France*
[g] *Laboratoire de Tribologie et Dynamique des Systèmes, UMR CNRS 5513, Ecole Centrale de Lyon, 36 avenue Guy de Collongue, 69134 Ecully Cedex, France*
[h] *Flanders Make vzw, Corelab DecisonS, Gaston Geenslaan 8, 3001, Leuven, Belgium*
[i] *Division Mecha(tro)nic System Dynamics, Department of Mechanical Engineering, KU Leuven*
[j] *Univ Lyon, INSA-Lyon, LaMCoS, CNRS UMR5259, F-69621 Villeurbanne, France*
[k] *Flanders Make@KU Leuven – DMMS: Dynamics of Mechanical and Mechatronic Systems, Celestijnenlaan 300, BOX 2420, 3001 Leuven, Belgium*
[l] *School of Mechanical and Manufacturing Engineering, University of New South Wales, UNSW Sydney NSW 2052, Australia*
[m] *Universidad Nacional de Colombia - Sede Manizales, Signal Processing and Recognition Group, Km 9 way to magdalena, Manizales Caldas, Colombia.*

Abstract

A smartphone is a low-cost pocket wireless multichannel multiphysical data acquisition system: the use of such a device for noise and vibration analysis is a challenging task. To what extent is it possible to carry out relevant analysis from it? The Survishno conference, held in Lyon in July 2019, proposed a contest to participants based on this subject. Two challenges were proposed, wherein each a mute video showing an object moving/excited at different frequencies was provided. Due to the frequencies set and the video sampling characteristics, special effects occurred and are visible on both videos. From the first video, participants were asked to estimate the Instantaneous Angular Speed (IAS) of a rotating fan. From the second video, they were asked to perform the modal analysis of a cantilever beam. This paper gathers the interesting ideas proposed by the contestants and proposes a global method to solve these two problems. One major point of the paper might be the advantageous use of the rolling shutter effect, a well-known artefact of smartphone videos, to perform advanced mechanical analyses: the consideration of the unavoidable slight phase shift between the acquisition of each pixel opens up the possibility to perform a dynamic analysis at frequencies that are much higher than the video frame rate.

Keywords: IAS, Instantaneous Angular Speed, modal analysis, rolling shutter, wagon wheel, motion tracking, edge detection, SURVISHNO 2019 challenge

*Corresponding author
Email addresses: hugo.andre@univ-st-etienne.fr (H. André), quentin.leclere@insa-lyon.fr (Q. Leclère)

1. Introduction

Survishno is a joint organization of the conferences Surveillance, VISHNO (Vibration Shocks and Noise) and EVA (Experimental Vibration Analysis), which aims at gathering researchers, scientists and industry professionals, in the fields of vibrations, health monitoring and structural dynamics. In order to maintain a collegial atmosphere, each edition of these conferences featured a technical challenge open to all the attendees. Traditionally, the Surveillance challenges have focused on signal processing dedicated to rotating machine health monitoring, while the VISHNO challenge was oriented on modal analysis problems from the experimental point of view, modeling, data acquisition and identification. To satisfy the regulars of each community, the present challenge was divided into two parts, both relying on signals acquired using the same acquisition system: the camera and microphone of a regular smartphone, wherein lies the originality of the present work. This system is both universal in the sense that a smartphone can be found in every pocket, cheap when compared to the systems that are commonly used in our laboratories, and just needs to be tested to carry out advanced technical analysis.

Of course, the idea of using smartphones to ease life is not new. Medical scientists welcome the smartphone usage to acquire data form neurofeedback electrodes and other accessories (including a saliva chemical analysis) in order to detect fatigue and prevent car accidents [1]. A smartphone is indeed able to provide a mobile and low-cost electroencephalogram measurement system. A review was also recently dedicated to the description of the need for monitoring human stress using daily life signals rather than confining their subjects in a laboratory environment[2]. Other particularly attractive works offered the prospect of employing the smartphone solely, without connecting electrode cables, for recording the electrocardiogram to detect heart problems [3]. This is expected to enhance the feasibility of monitoring patients with heart diseases, irrespective of their possible frailties, and incapacitation. Watson et al recently proposed a transverse review evaluating medical apps' ability to perform clinical care through the use of classical smartphone sensors [4]. Smartphones may indeed pave the way to many applications, which may lead to a real transformation of the practice of medicine. There is an interesting parallel to draw with our fields of application, where scientists develop optimal and elegant but complex processing tools, validated within laboratory environments on simple machines through the use of efficient but expensive means, whereas industry needs to monitor complex machines with simple tools and cheap means. But mechanical engineers are also on the move. Recent works show that smartphone accelerometers could be used as instrumentation systems for modal identification of flexible structures such as pedestrian and automotive bridges [5]. Alavi et al [6] presented a comprehensive literature review of smartphone-centric research for the monitoring of civil infrastructure systems, while Grossi et al [7] recently listed out the mechanical signal acquisitions that could be done with smartphone sensors (high-resolution camera, microphone, accelerometer, gyroscope, magnetometer and GPS).

In the SURVISHNO 2019 contest, the first challenge (C1) focuses on the determination of a fan instantaneous angular speed, with the fan installed on a hand-operated drill. This entertaining challenge gives prominence to a subject that currently is in the spotlight of the signal processing community: Instantaneous Angular Speed estimation in non stationary operating conditions. Indeed, the machines that happen to be operated under non stationary conditions need a phase reference, ie. a speed signal, to set up a cyclo-stationary approach based on angular (re)sampling. This is advocated by a complete special issue of MSSP published in 2014 and dedicated to angular approaches for non stationary machine monitoring [8]. However, industrial condition monitoring systems do not always have access to a tachometer signal. Therefore, the estimation of instantaneous angular speed with any kind of sensor except an angular encoder became of major importance in the field of rotating machine monitoring. It is no coincidence that the challenge organized during SURVEILLANCE 8, held in Roanne in 2015, was focused on the speed signal reconstruction of a jet turbine transmission using vibration sensors only [9]. This event was moreover an opportunity for the community to innovate by proposing new tools to tackle this issue: Multiple Order Probabilistic Approach [10] or cepstral based speed reconstruction [11], and others were recently compared in a dedicated review [12]. It is therefore not surprising that C1 was, on the whole, very well treated by the participants of SURVISHNO 2019. Several approaches, some complementary, and some based on different concepts, were proposed and are synthesized in this paper.

The second challenge (C2) aims at performing the first steps of the modal analysis of a cantilever beam. Through the identification of natural frequencies and mode shapes, vibration measurements are classically used to understand

Rank	C1	C2	Composition
1^{st} price	x	x	Julien Griffaton,
2^{nd} price	x	x	Alessandro Paolo Daga,Dario Anastasio
3^{rd} price	x	x	Cédric Peeters, Pieter-Jan Daems, Jan Helsen
	x	x	Frédéric Bonnardot
		x	Bilal Elyousfi
	x	x	Edgar Felipe Siera Alonso
	x	x	Francois Combet, Xavier Thomas
	x		Yasmine Hawwari, Luc Laroche
	x	x	Adrien Melot, Youness Benaïcha
	x	x	Agusmian Ompusunggu, Roeland De Geest, Merwan Birem
	x		Kevin Billon, Mathias Perrez,Sebastien Passos
	x		Florian Lacaze, Guillaume Paillot, Xiaowen Li
	x	x	Alexandre Mauricio, Konstantinos Gryllias , Chenyu Liu , Junyu Qi
	x		Zhan Yie Chin, Wade Smith

Table 1: Composition of the teams. Except the first 3 lines of this table, the teams are not sorted according to their results.

the dynamic behaviour of structures and to perform structural damage detection and health monitoring. The efficiency of the modal parameter estimation depends on the data collection system. The most traditional method relies on accelerometers. Although contact accelerometers are very accurate, the instrumentation of a complex structure can be tedious, and when installed on a small structure, the added mass cannot be neglected anymore. Non-contact methods, such as microwave interferometer [13], or Laser Doppler Vibrometer [14] are commonly used to avoid these drawbacks. Our interest is focused here on cameras measuring visible light, which once again provide a competitive method in term of price-quality ratio. They can range from accurate instruments for high-frequency and high-resolution video to inexpensive units such as those on smartphones. However, the video processing in order to extract quantitative information remains difficult, and there is still room for improvement, especially when applied to modal analysis. Just like Chen et al [15], who worked on motion magnification to transform a low resolution video in a way that small displacements can be visually appreciated and evaluated, papers about this topic are based on high speed cameras. Such high quality equipment does not suffer from the inherent limitations of smartphone cameras. This explains why contest participants found this second problem more difficult to deal with than the first one. Anyhow, this paper is an opportunity for the contest organisers to propose a formalized method of dealing with modal analysis using a smartphone camera.

The remainder of this paper is organized as follows. Section 2 aims at describing mobile phone limitations: image distortions, wagon-wheel effect and rolling shutter effect. Section 3 synthesizes the various approaches successfully used by the contestants regarding the speed reconstruction problem. Section 4 proposes an original method to tackle the modal analysis problem. A discussion is finally proposed in section 5 to describe the possibility to do dynamic analysis from videos at frequencies which are much higher than the classical Shannon's limit.

1.1. Competing teams

The present paper collects the results returned by fourteen competing teams in an attempt to solve the 2 challenges presented in the SURVISHNO Contests. The composition of the teams is detailed in Table 1 and their members are also all listed as co-authors of the paper.

1.2. After the contest

The data are provided as supplementary material of the present paper. The videos for the two challenges are shared with the corresponding frequency signals synchronized through the sound track at the following adress:`https://laspi.univ-st-etienne.fr/fr/equipe-tsi/liens-externes.html`. It is kindly asked that any publication that will result from subsequent use of the data references to the present paper and the *Survishno contest, Conference Survishno, July 08-10, 2019, Lyon, France*.

2. Smartphone video limitations

2.1. Image distorsions

The image captured by a camera is a 2D representation of a 3D scene. The 3D to 2D transformation is subject to image distortion: lines that are parallel in the real world are no longer parallel in the image because of perspective distortion, and straight lines of the real world can appear curved in the image because of lens distortion. The latter effect is intrinsic to the used optical device, and can be corrected by using an appropriate camera calibration operation [16, 17, 18]. The former one is trickier to correct, because it depends on the geometrical configuration of the scene [19]. However, once the geometrical configuration of a given plane in the captured 3D scene is retrieved by any a priori information or optimization process, it is possible to correct the perspective effects by geometric laws [20, 21]. This topic is particularly well documented in the field of computer vision, a domain in which the retrieval of physical information from digital picture is a key preprocessing step. Once image distortion effects have been corrected, a last issue is to correctly determine the scaling factor, that is generally estimated from the a-priori knowledge of any object dimension in the picture.
Note that the correction of image distortion is not the object of the present paper. A first reason is that these corrections do not constitute a real difficulty when processing the videos of the SURVISHNO 2019 contest, as we will see it in the following. A second reason is that the solutions to correct these effects are quite well documented in the field of image processing and computer vision, and some standard solutions exist to implement it (see for instance [22]). Built-in solutions are also eventually embedded in some smartphone image signal processors (at least for lens distortion), that sometimes deal with the combination of information obtained from several cameras with different focal lengths.

2.2. Wagon-wheel/stroboscopic effect

Videos are made of successive frames in which the exposure time is as short as possible to avoid blurred images. A video can therefore be seen as a series of brief episodes where the motion is absent, interrupted by longer episodes of invisibility, as if it was lit with a stroboscope. When it comes to switching from continuous-time signals to discrete-time signals, any MSSP reader knows that disrespecting the Nyquist–Shannon sampling theorem might produce some aliased oddity. As such, the wagon-wheel effect is a reference to stagecoaches or wagons in Western movies: the stroboscopic effect appears when the rotation speed, multiplied by the spoke number, is greater than half the image rate. This phenomenon is highlighted in the first challenges where the 6 holes of the fan appear to go in the opposite direction than the 10 fan blades, that rotates independently from the central hexagonal screw.

Although it does not present a cyclically symmetrical object, C2 is also subject to temporal aliasing. Of course, recordings of any regularly oscillating object, such as a singing guitar chord or a vibrating cantilever beam, will show an oscillating movement whose frequency is the remainder of the euclidian division of the actual frequency divided by half the sampling frequency. This second form of the effect is known as stroboscopic effect.

2.3. Rolling shutter effect

Rolling shutter is a method of image capture in which every video frame is captured not by taking a snapshot of the entire scene at a single instant in time, as it would be with a global shutter camera, but rather by scanning across the scene line by line. In signal processing language, the rolling shutter translates the fact that the acquisition of each pixel of a frame is multiplexed. Cameras embedded in smartphones are digital, and therefore rely on image sensors that convert the amplitude level of light waves into digital signals. More generally, two types of sensors, both developed in the 1960's, are used today to acquire digital images: Charge-coupled device (CCD) and Complementary-Metal-Oxide Semiconductors (CMOS). Broadly speaking, the first can be seen as powerful and expensive while the second dominates where overall cost is of concern. In other words, CCD are reserved for the most demanding applications (television broadcasting, professional, medical or scientific applications) while CMOS are embedded in most of modern smartphones. This distinction is of importance since many CMOS sensors present the rolling shutter effect. CMOS Sensors are a modern type of Active-Pixel Sensor, *ie* an image sensor where each pixel has a photodetector and an active amplifier. The scan proceeds from one photodetector to the next one either vertically or horizontally. Therefore, each pixel of the scene picture can be recorded at a slightly different instant. However, since the entire image is displayed at once during playback, as if it represents a single instant in time, fast-moving objects are subject to predictable distortions: wobble (vibrating camera induces wobbled pictures), skew (quickly moving

camera or object bends the picture diagonally), partial exposure (if a camera flash goes on for only part of the time of the exposure) and spatial aliasing (adjacent pixels are sampled in violation of the sampling theorem, when the camera or object motion is too rapid). In this paper, only one dimension of rolling shutter is considered: column pixels are supposed to be simultaneously acquired. However, determining the direction of rolling shutter (from left to right or from right to left) is a difficulty in itself that can only be overcome in the case of the hypnotic fan, since the real movement of the object can be reasonably assumed. This topic is discussed in more detail in the next section.

3. Hypnotic fan

3.1. Problem statement

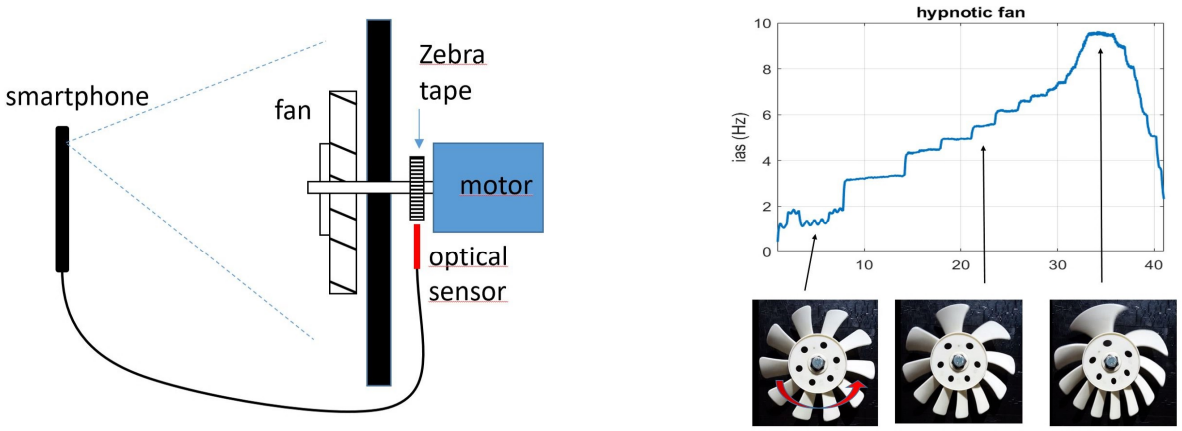

Figure 1: Left: Set up scheme for the hypnotic fan problem. Right: speed profile obtained from the zebra tape encoder with 3 miniature images presenting increasing levels of rolling shutter. The first images sequence obtained at slow regime indicates the direction of rotation.

The contestants were given a mute video file and were asked to provide the instantaneous speed of a fan mounted on a drill as a function of time. The video duration was a bit more than 40 seconds and the signal sampling frequency rendered by the participants should correspond to the video frame rate ($f_{fr} = 30Hz$). The soundtrack was removed since the audio input of the smartphone has been diverted from its intended use to monitor the TTL signal of an angular sensor: a 36 pulses per revolution zebra tape. The signal given by the optical probe was processed through a simple zero crossing technique [23], and used as the fan speed reference to evaluate the contributions of the contestants.

A symbolic diagram of the assembly is shown in figure 1. The top right plot shows the evolution of the speed signal measured by the zebra tape throughout the measurement. Three cropped excerpts of the video are presented below to illustrate the amplitude of the rolling shutter during the measurement. But except for the frame rate, which is intrinsically provided within the video file, no other parameters were given to the contestants. In particular, neither the existence, frequency nor the direction of the rolling shutter was communicated a-priori. And to make it even harder, or in order not to distort the amateur style of videos usually made with this kind of device, the smartphone autofocus system activated itself during the shooting.

This section is divided into three parts. The first describes the preprocess steps used by most of the competitors to prepare the ground before estimating speed. Then, two distinct approaches are successively presented, since two kinds of solutions were proposed by the candidates.

3.2. Preprocessing steps

The video is realized in 16:9 format with a 1080x1920 resolution in RGB24. Each colored image is thus a matrix of dimension 3 and size 1080x1920x3. Given that each element of this matrix is coded in 1 byte, an image has a size of little more than 6 megabytes. The objective being to label each image with one simple speed value, the problem already seems strongly oversized! It then seems relevant to pre-process each image to reduce the amount of data to be processed by the algorithms that will follow. To say it differently, the pre-processing is presented here as a way to reduce the number of variables to simplify the problem.

3.2.1. Grayscale conversion

The first way to simplify the image is to shrink the matrix to Dim-2, by converting the colored image into grayscale. Different formulae were used since there is no unique procedure to convert a colored image. Each color pixel is described by a triple (R,G,B) of intensities for red, green and blue. If the converted pixel value P must stay within $0-255$, the rendering obtained with the algorithms listed below is presented in figure 2. The first indicator $P_{Average}$ is the most intuitive conversion, and the third $P_{Luminosity}$ is most often used in image processing algorithms because it forms a weighted average accounting for human perception. However, although human appear to be more sensitive to green than other colors, the automatic thresholding algorithm that will process the image might not, and the visual appearance could be overlooked in favor of a non-linear approach that pushes aside the extreme values. As such, the information approach yielded good results.

$$\begin{aligned} P_{Average} &= (R+G+B)/3 \\ P_{Lightness} &= (min(R,G,B)+max(R,G,B))/2 \\ P_{Luminosity} &= 0.3\cdot R+0.6\cdot G+0.1\cdot B \\ P_{Information} &= (R\cdot G\cdot B)^{1/3} \end{aligned}$$

Figure 2: Gray-scale conversion using different formula. From left to right: Average, Lightness, Luminosity, Information

This paper does not dig any further into the subject, since in the specific case of this video, the grayscale conversion step does not pose any difficulty, given the binary character the images originally proposed. Even before gray-scale conversion, the fan is rather white while the background is rather black. Moreover, most of the contributions directly added a thresholding step, allowing to binarize the image. Each pixel is then coded with only two values. Although some participants chose to adapt the threshold locally, it was possible to find a single threshold to distinguish the rotor from the background. For more details, interested readers will find a comparison of Bradley's adaptive threshold [24] and of Otsu's Fisher-criterion-based threshold [25] in [26].

3.2.2. Perspective distortion correction

As explained before, the image distortion is due to two effects, namely lens (intrinsic) and perspective (extrinsic) distortion. Lens distortion correction requires a fine calibration of the device, using some pictures of a specific target as a chessboard [16, 18]. However, such a picture was not available for the camera used to make the contest videos. Moreover, lens distortion appeared to be very limited - at least of the second order as compared to perspective distortion.

Perspective distortion is also visibly limited for the hypnotic fan case: the fan plane seems to be almost perpendicular to the camera axis. But actually this is not completely true: some differences can be measured between the vertical and horizontal span of the fan: it does not follow a perfect circle as it should if the image wasn't distorted. The determination of the angles between the camera axis and the fan plane can thus be determined as an optional pre-processing, so as to rectify as much as possible the circularity of the fan.

3.2.3. Centering

The knowledge of the position of the rotation axis in the image is a necessary pre-processing step to any image-based rotation speed estimation problem. Moreover, this position can change as a function of time if the camera is not perfectly fixed to the stator of the rotating machine. The strategy that can be implemented to solve this issue strongly depends on the application case, on the shape and specificities of the rotating part of the machine that is captured by the video, and on the view angle (this specific problem is discussed in the previous section 3.2.2). A solution to the particular case of the surveillance 9 contest video is proposed hereafter. Even if this approach cannot be generalized

to any case, it gives an example of how this issue can be addressed from a practical point of view.
An interesting feature in the considered video is the white disc constituting the central part of the fan. A potential strategy is to find the circle of maximum radius completely belonging to that part. For a given candidate circle centre position (x_c, y_c), and a candidate radius value R, image pixels can be picked up along the corresponding circle. When increasing R, while the circle is completely inside the white part, the minimum of picked-up pixel values will be quite high (corresponding to the white color coding). When the circle will partly fall outside the white region, the minimum value will be low (black color coding). Repeating this analysis for each candidate center will lead to the center that maximizes the radius with minimum values above a given threshold.
The position of the center of the fan, extracted at each frame using this strategy, is drawn in Fig. 3. It can be seen that the fluctuation amplitude covers a range of several pixels along the x and y axes. Another observation is that the position exhibits a strong periodicity, related to the rotation speed, that could be exploited for any balancing purpose or orbital analysis. Some discontinuities appear on the curves, clearly visible on the x axis coordinate. These discontinuities correspond to the transitions between different regimes (different rotation speed set points). A last interesting observation is that the y axis position seems to slowly increase during the video, maybe due to a deformation of the smartphone support (a simple microphone stand) caused by the gravity.

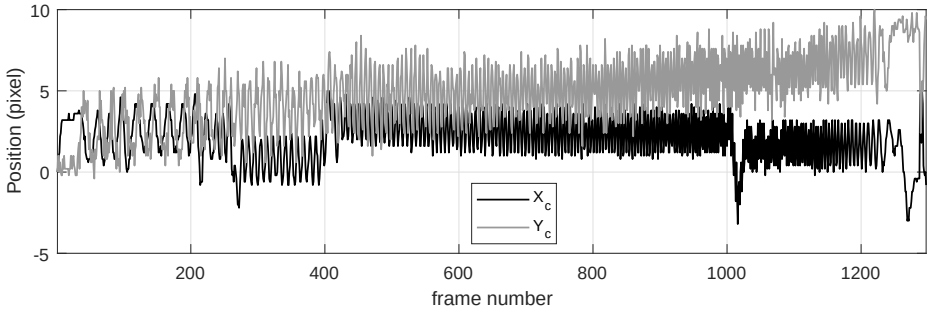

Figure 3: Evolution of the position of the rotation axis during the hypnotic fan video relatively to the position in the first frame.

3.3. Frame to frame correction

Once the preprocessing steps have been carried out, we have an image with simplified color levels (grayscale or binarized), that are framed on the rotor. Only its rotation speed determination remains. The "frame to frame" solution, preferably chosen by the competitors, measures the angle travelled by the rotor from one frame to the next. Once this angle is known, the rotation speed is trivially derived by considering the frequency of image capture (ie. video frame rate). The main difficulty with this approach lies in taking into account the wagon wheel effect to determine the actual angle change rather than the apparent angle change. Various solutions have been implemented by the candidates to overcome this difficulty and are summarised hereafter.

3.3.1. Wagon wheel effect

As a logical continuation of the centering step (presented § 3.2.3), a polar coordinate system (CS) can be attached to the rotor pole. Hence, let $M(\rho, \theta)$ be a specific point on the rotor defined by its distance ρ from the pole, and the angle θ made between the line segment from the pole to M and the positive horizontal x-axis. Contributions varied widely on the choice of the M point, as shown in figure 4. Some participants attached it to an object that does not suffer from the wagon wheel effect: the hardness index of the screw (M_1) or a forgotten stain on the inner disc (M_2). Indeed, since the fan speed did not exceed $f_{fr}/2 = 15Hz$ (see fig 1), the angle actually swept by one of these points between two consecutive frames corresponds to the measured angle on the video. Others preferred to jump in feet first into the wagon wheel effect, and tracked an object suffering from cyclic repetition: one of the six holes of the inner disc (M_3) or one of the 10 blades (M_4).

The solutions proposed by candidates tracking one of the objects suffering from cyclic symmetry are described with the help of Fig 5. This figure illustrates the problem through the case of the 6-holes-tracking during the beginning

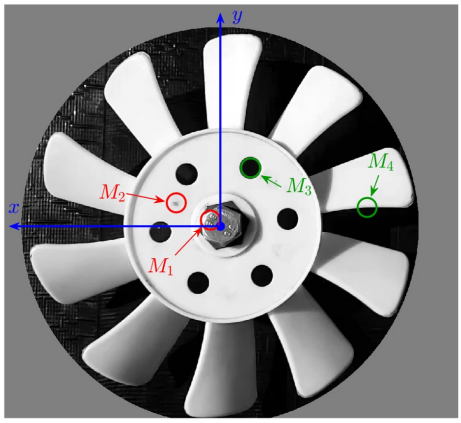

Figure 4: Location of the points tracked in the different strategies, once the image is cropped around the fan center. In green, 2 examples of point subject to the wagon wheel effect.

of the measurement, when the wagon wheel effect does not cloud the issue yet. Each color shows the evolution of θ_i, with $i \in [1:6]$ where θ_i here stands for the i^{th} hole found in the trigonometric direction starting from the fixed $\vec{x}$ axis. Therefore, each color does not always represent the same physical hole, except during during the first elapsed second of the measurement, when the fan starts up. Some competitors took advantage of this brief peculiarity to set up a recursive approach. Indeed, estimating the speed of the first frames is trivial since the machine is still. The absence of rolling shutter effect is the best evidence of this necessary base case assumption. The induction step uses the object position and speed known for step n, to estimate the position of the object at step $n+1$, and once the closest object is matched, deduce the speed at step $n+1$. A very similar approach, if not the same in other words, was to select the object of step $n+2$ minimizing the angular acceleration of the tracked object whose position is known for the two previous steps n and $n+1$. These simple recursive methods, both based on the derivatives of the position, have shown excellent results. The other method, favoured by most of competitors, was to recognize the object among its clones using a singular pattern (M_1 or M_2).

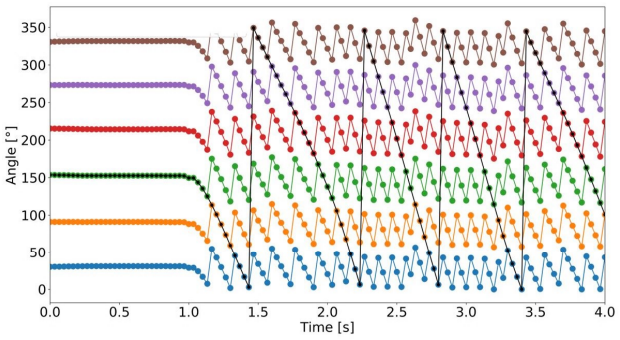

Figure 5: Angles of the six holes during the beginning of the measurement. The colours respect the order of the holes regarding the fixed $\vec{x}$-axis . The black line presents the angle of one specific hole.

In the end, it turned out that the tracking of a singular object only, despite the smaller number of pixels involved in the operation (and thus risking to amplify the influence of noise), despite the impossibility to average the speed

estimations obtained with every elements of the cyclic symmetry, was sufficient to obtain an excellent estimate of the fan speed. Even in the case of M_1, although its close position from the rotation pole limits to the extreme the number of pixels carrying information, an approach described by Daga in [26] managed to properly reconstruct the fan speed using a tracking system based on GA-adaptive template matching. In this approach, image processing only concerns the fan fixing screw (its hexagonal shape and its hardness index). The focus being limited to a restricted area of the image, the rolling shutter effect was rightly and advantageously ignored. For those tracking objects moving on a wider zone, the rolling shutter effect had to be taken into account...

3.3.2. Rolling shutter effect

In the frame-to-frame methods, the rolling shutter is considered as a nuisance rather than a source of information. Amongst the various solutions proposed to solve the problem, one aims at processing the entire image to simulate a simultaneous acquisition of the entire frame. This procedure is detailed since it provides a general analysis of the rolling shutter effect applied on a stable but rotating object, and also because it provides the opportunity to explain how to determine rolling shutter direction, in concrete terms.

Let's first focus on any point M attached to the rotating fan, using the polar CS. $M_0(\rho, \theta(t_0) = \theta_0)$ stands for the position of this point when the first column of the frame is acquired and $M_i(\rho, \theta_i)$ stands for its position at time t_i, date of the i^{th} column acquisition. Parallax and other perspective distortions are assumed to have been corrected beforehand (see section 3.2.2), so ρ does not depend on the time the point M is acquired. Let f_{sh} be the constant velocity of the rolling shutter expressing the delay separating M_0 and M_i : $t_i - t_0 = i/f_{sh}$. in a first approach, $f_{sh} \approx \pm f_{fr} \times N_C = \pm 32400Hz$, where f_{fr} is the frame rate of the video, and N_C the number of columns to be successively captured. The sign of f_{sh} depends on the rolling shutter direction and will be determined in the next paragraph. Finally, let Ω be the rotating speed of the fan. If Ω is assumed constant during the frame acquisition, then again $\theta_i = \Omega \cdot (t_i - t_0) + \theta_0$. The formula combination of the previous expressions yields a velocity composition relation valid for any point attached to the rotor fan :

$$\theta_0 = \theta_i - \rho \cdot \Omega \cdot \frac{i}{f_{sh}} \tag{1}$$

The ambiguity of the sign of f_{sh} is removed by determining the rolling shutter direction using the beginning sequence of the movie, when the fan starts up (see Fig. 1, left miniature). Indeed, the apparent stillness cannot result from a rotating speed equal to the frame rate, since the absence of rolling shutter effect indicates that the fan freezes during the frame acquisition time span. Then, Fig 5 indicates a negative speed around $t = 1s$, and the fan will not show such still behaviour before the end of the video. This proves that the rotating speed stays negative during the whole capture, and the rolling shutter direction can now be deduced from a skewed frame, as illustrated on fig 6, left plot. The frame observed on this plot is obviously subject to rolling shutter: blades that appear wider than in reality go in the direction of the rolling shutter, while those appearing thinner than in reality go in against it. The rolling shutter direction in the video is therefore from the right to the left of the picture: $f_{sh} > 0$.

Once this is set, all pixels of each frame are expressed in the polar CS and can be spatially transported at time t_0 using eq. 1. Figure 6 middle plot shows the effect of a rolling shutter correcting process once applied on all pixel of a cropped frame. Beyond the apparent quality obtained on the central image, the limits of the process cannot be ignored:

- The pixels showing the background are also transported though they do not rotate with the fan. Nevertheless, this phenomenon is not deleterious if the background is plain, which is the case here.
- After the correction, the density of pixels is not uniform anymore. Some zones are empty of information and therefore appear darker than they should. This common noise problem is easily dealt by *closing* the image. the closing of a set A by a structuring element B is a mathematical morphology defined such as $A \bullet B = (A \oplus B) \ominus B$. A is the central image of fig 6 and $A \bullet B$ is the right image if B is a simple all-ones 3-3 matrix. This classical operation actually is the combination of two basic operations, dilation and erosion, removing respectively small dark and small light spots and defined such as

$$(A \oplus B)(a) = \sup_{b \in B}(A(a \cdot b)) \qquad A \ominus B(a) = \inf_{b \in B}(A(a \cdot b))$$

- Last but not least, eq. 1 relies on two unknowns (Ω, f_{sh}), one of which happens to be the goal of C1 while the other cannot be precisely determined without an optimization approach (an estimation method is provided in §4.2). This method was however used by the winning competitor, who started at tracking a singular spot (M_2 in Fig. 4, for instance) to have a first rough estimation of Ω, corrected the rolling shutter effect, and finally tracked the blades for a more precise speed estimation.

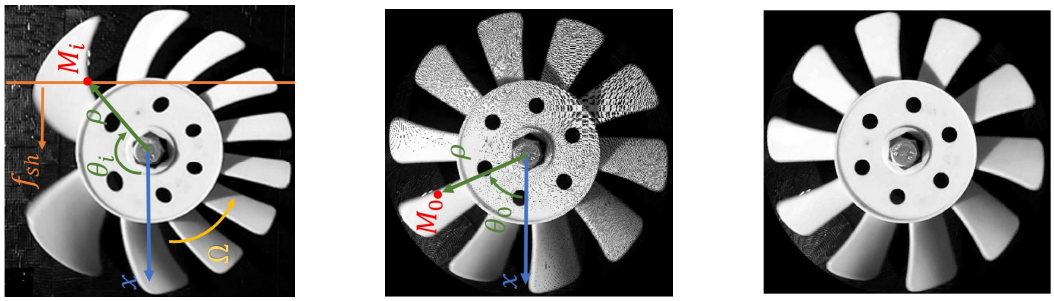

Figure 6: Left plot shows a frame where the fan appears heavily distorted by the rolling shutter effect. Distortion is corrected using Eq. 1 on middle plot, and is further processed through a $3x3$ closing operation on right plot.

This being said, it was not necessary to correct the image to deal with the problem. Contributions that only monitored the appearance of cyclically repeating objects on a single column of the image, for example, were prevented from doing so. Note, however, that the speed estimation suffered from a slight phase shift depending on the selected column. Moreover, such approach limited the number of points used to estimate the speed, and could therefore be expected to lower the signal to noise ratio.

3.4. Single frame IAS estimation

A generic procedure is given in the following, allowing the estimation of the rotation speed based on the analysis of the image distortion caused by the rolling shutter. A specific radius ρ is chosen on the fan, exhibiting a known number of periods B as a function of angle θ (rotation around the fan center, expressed in revolution). On the studied fan, if the radius coincides to the blade radial position, the period is $1/10$ ($B = 10$ blades). Image pixels are extracted along a circle of radius ρ, using a regular sampling of the angle θ (sampling step $d\theta = (2\pi\rho)^{-1}$), by picking the value of the closest pixel to the position (ρ, θ) (see Fig. 7, left). Note that the circle coordinates eventually take into account the perspective correction discussed in section 3.2.2. The blade edge angular positions, noted $\theta_b (b \in [0, ...(B-1)])$ are detected using a simple thresholding procedure (see Fig. 7, center).
The rolling shutter, presented in section 2.3 and put into equations in section 3.3.2, is then taken into account to estimate the time instants at which blade positions θ_b are detected depending on the x axis position :

$$t_b = x_b / f_{sh} = \rho \cos(2\pi\theta_b) / f_{sh}$$

Note that the rolling shutter direction in the video is from the right to the left of the picture (this can be inferred from the very beginning of the video, as explained in section 3.3.2), and that the x axis is oriented from right to left.
The position of the first blade at each time instant t_b is recovered from the position of each blade :

$$\theta_0^{(b)} = \theta_b - b/B$$

Assuming a constant rotation speed Ω (in Hz) during the frame acquisition, the position of the first blade follows:

$$\theta_0^{(b)} = \Omega t_b + \Theta$$

where Θ (in revolution) is the unknown position of the first blade at the beginning of the frame capture ($t = 0$). Writing this equation for all blades gives a linear system of n equations, Ω can be estimated using the classical least squares fitting (see Fig. 7, right):

$$\Omega = \frac{\sum_b \left(\theta_0^{(b)} - \overline{\theta_0}\right).(t_b - \bar{t})}{\sum_b (t_b - \bar{t})^2}$$

where $\overline{\theta_0}$ and $\bar{t}$ are average values of $\theta_0^{(b)}$ and t_b over b.

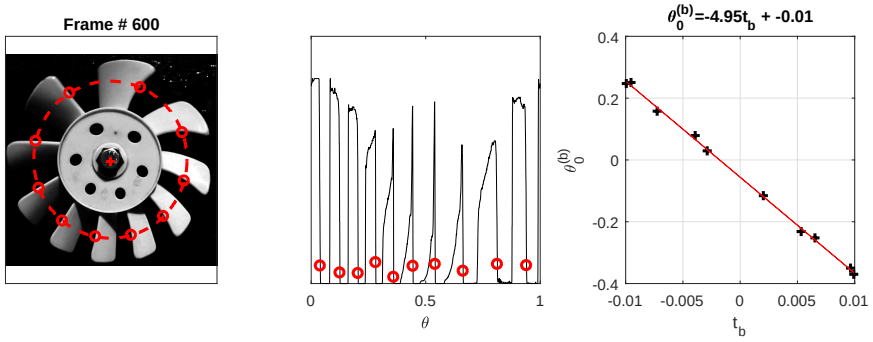

Figure 7: Single frame based IAS estimation. Left: detection of blade leading edges in the picture. Center: pixel value as a function of θ for a given ρ, red circles corresponding to white to black edge detections. Right: least squares fitting of the $\theta_0^{(b)} = \Omega t_b + \Theta$ curve.

Note that the least squares fitting estimates an average value Ω of the rotation speed during the frame acquisition, but a second output is the position of the first blade at the beginning of the frame acquisition Θ. This second information can be exploited to estimate a rotation speed using a frame to frame approach as described in section 3.3.
Another interesting point of the single frame approach is that it is based on B estimations of the first blade position during the frame acquisition, at B different -non equally spaced- time instants. Moreover, different values of radius ρ would also lead to B different time instants, exploiting either the blades or the holes on the inner part of the fan. This approach could hence be used to analyse the instantaneous speed fluctuations with a much higher sampling frequency than the frame rate. This is discussed in section 5.
An important parameter when implementing the single frame approach is the speed of the rolling shutter frequency, noted f_{sh}, whose actual value is not known exactly. It can be assumed that its order of magnitude should be at least equal to the frame rate f_{fr} multiplied by the number of columns N_C scanned by the rolling shutter during the acquisition of a frame, because the beginning of a new frame cannot precede the end of the previous one. However, taking $f_{sh} = f_{fr} \times N_C$ leads to underestimate the rotation speed Ω. A comparison between a frame to frame approach and the single frame approach using $f_{sh} = f_{fr} \times N_C$ shows that the underestimation is by a factor of about 0.923, which means that the actual rolling shutter speed is equal to $f_{sh} = f_{fr} \times N_C / 0.923$.

3.5. Analysis of the estimated IAS

As already mentioned, a reference IAS of the fan was estimated using an optical encoder. The encoder output signal was plugged into the audio input of the smartphone, so that the encoder signal was recorded as the soundtrack of the video. Of course, the sound was removed from the file sent to contestants. The optical device is made of an optical probe facing a zebra tape pasted on the rotor of the fan, with 36 marks. This system is easy to mount, however a difficulty appears because the perimeter of the rotor is never a multiple of the marks width on the tape. One of the marks is thus smaller or larger than the others, this has to be corrected a posteriori (see [27] for instance) when estimating the IAS.
Contestants where asked to provide an IAS signal at a sampling frequency equal to the frame rate. The resolution of the IAS measured with the optical encoder was 36 samples by rotation (36 marks), an angle to time resampling operation was required so as to compare it to contestant feedbacks. The encoder signal has been used to build a time to angle law, in order to get the angular position at constant time intervals (1/30 s) using a linear interpolation, and to

finally build the reference $30Hz$ IAS signal. An estimation error has then been calculated for each contribution, using the following equation :

$$E_{cand} = \frac{\int \left(IAS_{cand}(t) - IAS_{ref}(t)\right)^2 dt}{\int \left(IAS_{ref}(t)\right)^2 dt}$$

Note that in the video file provided as supplementary material to the present paper, the video and audio tracks are not perfectly synchronized, a small delay (shift of the audio track by 127ms backward) has to be applied to the reference IAS so as to minimise the estimation error.
The error score distribution obtained by the different teams is presented in table 2. While 4 contributions where out of range ($E > 50\%$), 9 teams reached honourable values among which 4 contributions obtained excellent scores lower than 1.5%.

total number of contributions	13
contributions with $E > 50\%$	4
contributions with $10\% < E < 50\%$	3
contributions with $2\% < E < 10\%$	2
contributions with $E < 2\%$	4

Table 2: Distribution of error scores reached by the contestants

The eight best IAS curves are drawn in Fig. 8, together with the zebra tape speed.

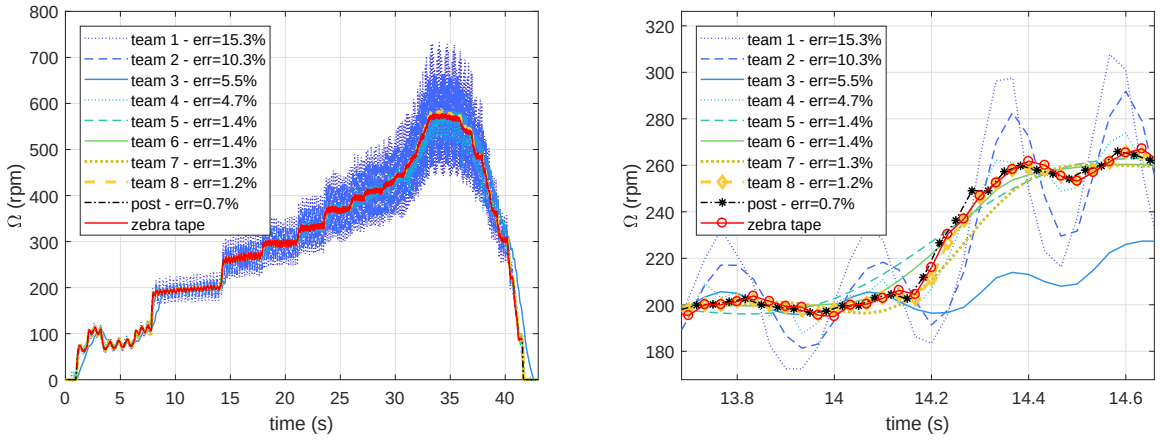

Figure 8: IAS curves with the lowest error scores, together with the IAS estimated with the reference (zebra tape), and an optimal IAS estimated a posteriori (post), for the whole video duration (left) and a close up view over 1s (right).

Error scores of the contributions are provided in the legend. Some IAS curves (teams 1,2,4) exhibit large fluctuations around the reference, probably due to a difficulty to correctly handle the rolling shutter effect. Another curve (team 3) seems to be slightly delayed, because of the use of a causal low pass filtering. Remaining curves (teams 5 to 8) are in excellent agreement with the reference, with similar error values between 1.2 and 1.4 %.
A last IAS curve estimated from the contest video is drawn in Fig. 8, based on the methodology described in this paper. This result has been obtained combining best approaches proposed by the contestants, and with the knowledge of the reference. The error of this last result is 0.7 %, this is only a little better than the contest best contributions, which validates their excellent level.

4. Shake the snake: Modal analysis of a cantilever beam

4.1. Problem statement

This problem focuses on the modal analysis of a cantilever beam clamped on a shaker. The participants were expected to find the resonance frequencies corresponding to two bending modes of the beam under forced vibrations. Both modes are second bending modes oscillating on two orthogonal planes: their frequencies are distinct because the profile of the beam is not perfectly circular: a small flat plane was manually tooled on the beam. Although the modes are slightly coupled with their respective orthogonal planes, the camera was positioned in such a way that the direct observation clearly favoured one mode, while observation of the beam shadow favoured the other. This configuration is illustrated in figure 9, where green arrows stand for the oscillations visible on the direct observation while the red ones correspond to the oscillations magnified by the beam shadow. Fig. 9 also shows an image averaging every frame of the 50 seconds video. On this averaged image, the blurred beam presents one nodal point on each plan, both modes shows second mode shapes.

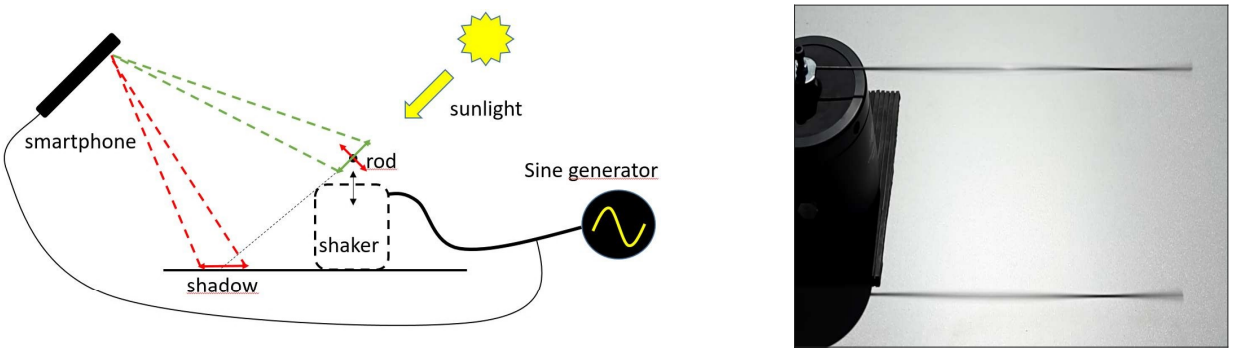

Figure 9: the set-up scheme illustrates the position of the camera, and explains the different behaviour of the oscillations observed on the beam directly or on its shadow. On the right, the image is the average of all the video frames.

The video was recorded using the same smartphone as in the previous experiment. Once again, the hand free kit analog input was diverted from its original use to acquire the signal from the function generator driving the shaker. This $48kHz$ video soundtrack was not provided initially: Participants had to identify the real resonance frequency through the aliased observation of the operating deflection shapes (ODS) of the $30Hz$ frame rate video. Fig. 10 top plot reads the real excitation frequency f (in blue) and the aliased one f_{alias} (in red): The apparent motionless of the beam at the beginning of the movie is only an illusion, the shaker is actually exciting the beam at a frequency corresponding to twice the sampling frequency of the camera ($2 \cdot f_s = 60Hz$). The frequency was then manually changed to scan both resonance modes. The 3-seconds averaged images, also shown in Fig. 10, enlighten respectively the apparent still position, the direct bending mode around $59Hz$ and the shadow bending mode around $55Hz$.

Last but not least, the rolling shutter effect clouds the issue even more by dephasing the oscillating waves, thus creating the odd illusion that the beam belly-dances as if a propagating wave was slowly traveling through it. Once again, the non-synchronous acquisition of the successive columns of each frame creates an oddity that can be used to overtake the Nyquist–Shannon sampling theorem. The image resulting from the average of the first frames (Fig 10 bottom left) explains the title of this problem. If the apparent immobility of the beam is due to aliasing, the snake shape is due to rolling-shutter.

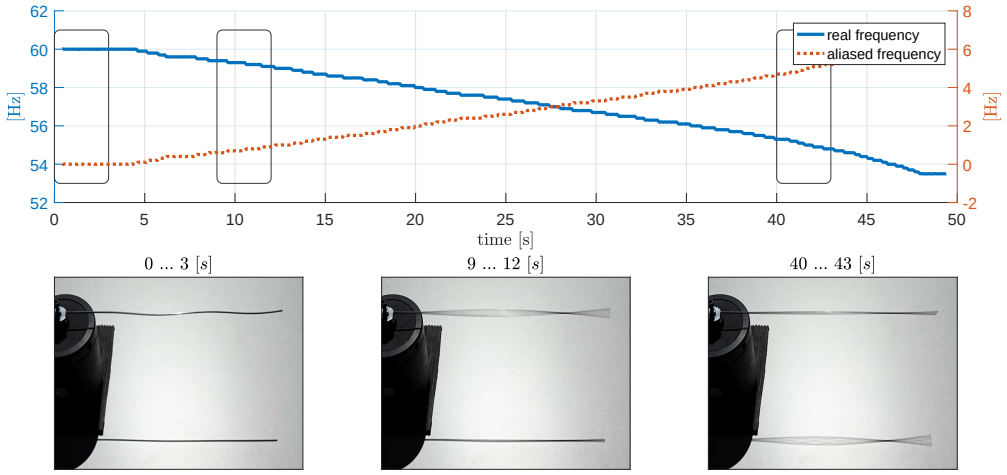

Figure 10: On top, evolution of the real (f) and aliased (f_{alias}) excitation frequencies during the whole run. Bottom: 3-seconds average image during three specific time zones. Middle and right images focus on the mode observed on the direct view and the shadow view, respectively.

4.2. Solution

The accumulation of these difficulties, no doubt aggravated by an excessively laconic statement, did unfortunately not allow the participants to achieve as good results as in C1. A solution is nevertheless proposed below to justify the feasibility of the investigation, and to illustrate once again how the rolling shutter effect can be used as an advantage to interpret videos of dynamic phenomena.

As illustrated in Fig 10, a first step to tackle this issue is to study images resulting from the averaging of successive frames. This step helps inferring that the beam is subject to 2 successive (and different) second order bending modes. Then, the post-processing procedures outlined in § 3.2 (gray scale conversion, perspective distortion, and centering step) might be implemented to prepare the solution below, but will not be explicitly detailed in this second case to avoid redundancies.

Let $w[i,t_{ij}]$, be the transverse movement of the beam expressed in pixels, with i the image column index in the rolling shutter scanning direction, and j the frame index. Since the shutter scan is in the direction of the ascending column index (from the right to the left hand side of the picture, this information is obtained from C1), time t_{ij} is defined such as:

$$t_{ij} = (j-1)/f_{fr} + (i-1)/f_{sh}$$

The solution proposed in this paper is divided in three steps detailed hereafter.

1. **Estimate $w[i,t_{ij}]$ the position of the beam.** (the solution proposed here can be improved using classical image processing filters, but is enough to perform a modal analysis of the beam.)
   - Crop selection. Every original frame are reduced to $[L \times C]$ pixels, with L the number of lines, and with C the number of columns. One selection is focused on the direct view of the beam, while an other on its shadow.
   - Convert in grey scale.
   - Select the $y-axis$ index of the min value of each column, since the beam is dark and the background is bright. Subtract the mean index value averaged over all frames for each column, so as to remove DC, this is $w[i,t_{ij}]$.
2. **Unfold the signal frequency.** This part aims at interpolating $w[i,t_{ij}]$ to estimate $w[i,t_{1j}]$: the transverse position that would be observed if all frame columns j were acquired simultaneously, at time t_{1j}. This correction of the rolling shutter effect requires an a-priori knowledge of the instantaneous frequency, whose real value remains

undetermined because of aliasing (wagon wheel effect). The following procedure explains how the rolling shutter effect can help solve this undetermination.
It is well known that a regular sampling of a signal leads to a frequency replication of its spectrum, with a period equal to the sampling frequency. The aliased (apparent) frequency f_{alias} (as well as its symmetric value $-f_{alias}$) is thus infinitely replicated at frequencies $f = S \cdot f_{alias} + N \cdot f_{fr}$, $S \in \{-1,1\}$, and $N \in \mathbb{Z}$. The true frequency is to be found between all these potential combinations of S and N, the span of N being possibly reduced to $N \in \mathbb{N}^*$ if we only look for positive frequencies, and if it is admitted that $N = 0$ (no aliasing) is physically non realistic - which is obvious considering the apparent oddness of the video. The pertinence of a combination $\{S,N\}$ can be assessed by its ability to recompose the ODS of the expected mode: the 2nd order bending mode as observed in Fig 10. More rigorously, it will be verified, for a given $\{S,N\}$, that the oscillation of all points of the reconstructed ODS are well in or out of phase, which should be the case for an ODS of a lightly damped structure around one of its natural frequencies. This analysis is carried out by a first optimisation problem, performed over candidate values of S and N. A second optimization process, based on the same indicators, is eventually proposed to refine the estimation of f_{sh}, which is firstly roughly estimated according to its minimal possible value: $f_{sh} = f_{fr} \cdot N_C = 32400Hz$ (where N_C is the total number of columns), or eventually its value estimated through C1 (see section 3.4): $f_{sh} = f_{fr} \cdot N_C / 0.923 = 35100Hz$. The whole part is detailed in the following:

- Compute the (aliased) analytic signal for each point of the beam $w[i,t_{ij}] = A[i,t_{ij}]exp(\imath\phi[i,t_{ij}])$ using the Hilbert transform, and unwrap the phase. A and ϕ are respectively the instantaneous envelope and unwrapped phase of the aliased beam motion at the column index i.
- Iterate on the candidate $\{S,N\}$ pair (see Fig. 11)
  (a) calculate the true instantaneous phase from the aliased one:

$$\phi_{SN}[i,t_{ij}] = S \cdot \phi[i,t_{ij}] + 2N\pi \cdot j$$

  (b) Interpolate in the time domain $w_{SN}[i,t_{ij}]$ and $A[i,t_{ij}]$ for all pixels ($i \in [1,...C]$) of the beam on the same time sampling grid t_{1j} ($j \in [1,...J]$) so as to obtain synchronous ODS at each frame j :

$$\tilde{w}_{SN}[i,t_{1j}] = \tilde{A}[i,t_{1j}]exp(\imath\tilde{\phi}_{SN}[i,t_{1j}])$$

  where the $\tilde{}$ denotes the interpolation.
  (c) Estimate the phase shift $\hat{\alpha}$ magnifying the energy of the real part of the ODS (in a way to standardize the ODS observed in each frame):

$$\hat{\alpha}_j = \underset{\alpha\in[0,\pi]}{\operatorname{argmax}} \left(\sum_{i=1}^{C} \left(\tilde{A}[i,t_{1j}] \cdot \cos(\tilde{\phi}_{SN}[i,t_{1j}] + \alpha) \right)^2 \right)$$

  (d) Evaluate the relevance of the ODS resulting for a given $\{S,N\}$ pair. This can be done two ways:
    – Recognize the physical mode corresponding to the a-priori knowledge that we expect a 2nd bending mode shape of a cantilever beam.
    – Compute the following power ratio:

$$\rho_j = \frac{\sum_{i=1}^{C} \left(\tilde{A}[i,t_{1j}] \cdot \cos(\tilde{\phi}_{SN}[i,t_{1j}] + \hat{\alpha}_j) \right)^2}{\sum_{i=1}^{C} \left(\tilde{A}[i,t_{1j}] \right)^2}$$

    for each frame j. This ratio is the fitness function maximised when most elements (pixels) of the beam are in phase (mod π) between themselves, it should be close to one when the excitation frequency is close to a natural frequency of the structure, when the ODS is essentially resulting from the contribution of the corresponding eigenmode.

  Amongst the five iterations shown in Fig. 11, only the third one ($S = -1, N = 2 \Longrightarrow f = 58.6Hz$) appears to match the classical 2^{nd} bending mode shape. On the other hand, ρ_j is drawn in Fig. 12 (left) for several $\{S,N\}$ pairs, for both the direct view of the beam and its shadow. Clearly, the $S = -1, N = 2$ possibility appears to be the most credible one.

- Once S and N are determined, the ρ indicator may be used to refine the evaluation of the shutter frequency, as shown in Fig 12 (right). To do that, previous steps (a) to (d) must be repeated for different candidate shutter frequencies.

$$\hat{f}_{sh} = \underset{f_{sh} > L \cdot f_{fr}}{\operatorname{argmax}} \left(\sum_{j=1}^{J} \rho_j \right)$$

3. **Identify the natural frequencies**.

   - Localise the frame $\hat{j}$ of maximal vibration, using the space-averaged envelope:

$$\hat{j} = \underset{j}{\operatorname{argmax}} \left(\sum_{i=1}^{C} \tilde{A}_{SN}[i, t_{1\,j}] \right)$$

   The left plot of Fig 13 shows the envelope for both observations of the beam. As expected initially, the direct view shows large oscillations at the beginning of the video (t = 13s) while the shadow view movement is larger in the end ($t = 42s$).

   - Compute the instantaneous frequency using the corrected instantaneous unwrapped phase:

$$f_j = f_{fr} \cdot \left\langle \tilde{\phi}_{SN}[i, t_{1,j}] - \tilde{\phi}_{SN}[i, t_{1,j-1}] \right\rangle_{i_1 \le i \le i_2},$$

   where $\langle x_i \rangle_{i_1 \le i \le i_2}$ stands for the averaging operator of x_i over interval $[i_1, i_2]$. Note that the interval used to average the phase difference may be limited to high amplitudes, the nodal zone has to be excluded because of the poor SNR. The estimated instantaneous frequency is drawn in Fig. 13 (right), together with the true shaker instantaneous frequency, that was "simultaneously" acquired on the smartphone using the the hands-free kit analog input. Note that the view (direct or shadow) used to estimate the frequency is at each frame the one that maximises the averaged envelope.

   - Conclude that the mode observed on the direct view is about $59Hz$ while the mode observed on the shadow view is about $55Hz$.

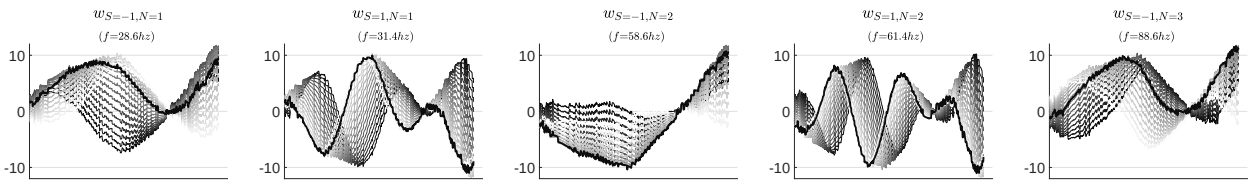

Figure 11: ODS reconstruction of 1 frame assuming different values for S and N. Each plot superposes ODS obtained for several $\alpha \in [0, \pi]$, though the curve corresponding to the optimized phase $\hat{\alpha}$ is plot in bold. $f = 58.6Hz$ presents the expected ODS of a second order bending mode.

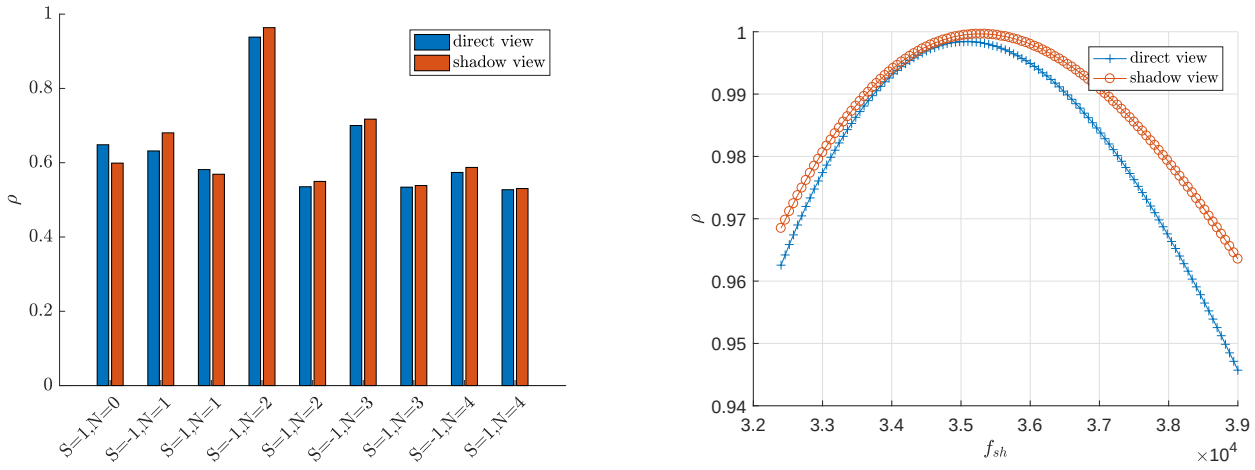

Figure 12: ρ function is maximised to optimize $[S,N]$ on left plot, and f_{sh} on right plot. Direct view is focused on the frames from the beginning period while the shadow view is focused on frames from the ending period, as detailed in Fig 10.

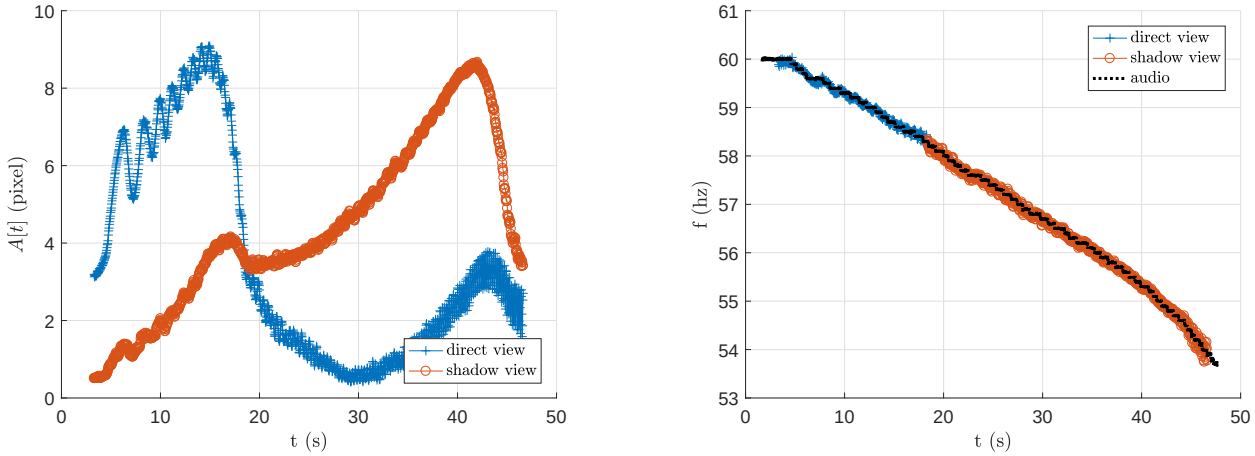

Figure 13: Left: Envelope averaged on the whole beam as a function of the time. Right: Estimation of the instantaneous frequency from both view points (direct and shadow view) and from the shaker signal (audio).

5. Discussion on the acquisition frequency

The works presented in this paper opens up the possibility to do dynamic analysis from videos at frequencies which are much higher than the classical Shannon's limit (Nyquist frequency, equal to half the video frame rate). In C1, even considering only a specific radius, the angular position of the fan is estimated at $B = 10$ distinct instants during the time laps between two frames. In C2, the eigenfrequencies of the beam are identified even though they are about 4 times higher than the Nyquist frequency. This possibility is induced by the rolling shutter effect which allows, under certain conditions, to observe an object at multiple distinct instants during the analysis of a single frame. It seems relevant to conclude this work by detailing the conditions favouring this effect.

The first condition to be met is that the object must be visible on several columns of the same frame (columns are considered as along the image axis perpendicular to the rolling shutter). The number of columns on which the object is observed corresponds to the number of instants when the object is observed. In the extreme case of a column wise object, only one position is determined during the frame acquisition, the frequency analysis is subject to classical aliasing issues. Fig 7 shows 10 distinct samples whose acquisition time are included in $\pm 0.01s$, while the frame sampling period is $1/f_{fr} > 0.03s$. Considering several radii in C1 would allow to involve more columns in the estimation of the angular position, and would directly increase the temporal resolution of the velocity signal. Note that this operation would necessitate to determine precisely the geometry of the blade edges, whose angular position may slightly change with the radius. This can be done from static pictures of the fan at the beginning of the video. In C2, the orientation of the beam regarding the frame column direction has a direct impact on the number of column involved. If the phone had been tilted by 90°, the shape of the mode would not have been distorted by the rolling shutter, and the vibration frequency would have been much more difficult to unfold. In order to illustrate these consideration, an additional video is provided together with the two contest videos. This video is capturing the same scene as the C2 video, a cantilever beam on a shaker. The difference is that the smartphone was in an horizontal position, as illustrated in Fig. 14. In the video taken horizontally, the beam axis is perpendicular to the rolling shutter direction, the rolling shutter effect is not present and the shape of the beam on the picture corresponds to the true mode shape. However, as it can be seen on the video, the video frame rate does not allow to catch the real frequency and a classic stroboscopic effect is observed: the true frequency is not identifiable. In the video taken vertically (the one provided for C2), the beam is parallel to the rolling shutter direction and a strong rolling shutter effect is present, that allows finding the true frequency following considerations developed in section 4.

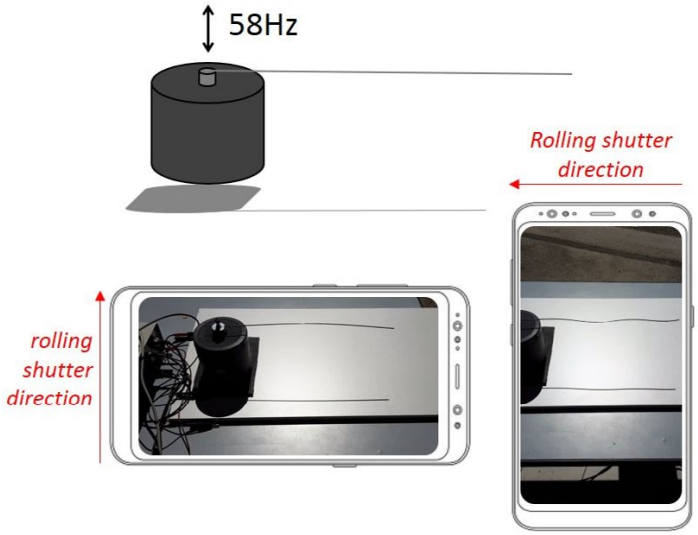

Figure 14: C2 video acquisition setup with the smartphone in horizontal of vertical position. Displayed frames are taken from the C2 video (horizontal position) and from another video taken with the phone oriented horizontally.

The second condition to be met is the knowledge of a-priori information on the shape of the observed object. Within one frame, the areas of the object that are observed from one column to another are not necessarily the same. In C1 and C2, the analysed objects are successively observed through different points, and a-priori information is necessary to infer the position of the entire objects from the observation of each point. In C1, the B estimates of the angular position of the rotor are based on the observation of the B blades. Assuming a-priori that the blades are regularly spaced and do not vibrate, each blade can be used to deduce the angular position of the rotor. In C2, the a-priori hypothesis that the beam deforms according to a classical bending mode is necessary to correctly unfold the modal frequency.

Last but not least, the knowledge of the rolling shutter frequency is necessary to precisely estimate the object's instants of observation. Concerning the smartphone used during this challenge, this frequency appeared significantly greater than its rule of thumb estimation ($f_{sh} \approx 1.08 \cdot f_{fr} \cdot C$). In C1, the rolling shutter frequency was determined by comparing the single frame and the frame to frame approaches. In C2, this difficulty was lifted by the assumption made a-priori on the modal behaviour of the beam (i.e. every beam elements are in phase mod π).

6. Conclusion

Survishno is a joint organization of the conferences Surveillance, VISHNO (Vibration Shocks and Noise) and EVA (Experimental Vibration Analysis), which aims at gathering researchers, scientists and industry professionals, in the fields of vibrations, health monitoring and structural dynamics. Each edition of these conferences was the subject of a technical challenge open to all the attendees, and this paper presents the contest organized during SURVISHNO 2019, in Lyon.

The contest was made of two challenges based on signals acquired using the same acquisition system: the integrated camera of a regular smartphone. The first challenge focused on determining the Instantaneous Angular Speed of a fan approximately installed on a hand operated drill, while the second challenge aimed at performing the first steps of the modal analysis of a cantilever beam. Technical limitations inherent to such universal and low-cost sensors have been first presented: image distortions, wagon-wheel effect and rolling shutter effect. Then, the successful approaches proposed by the candidates regarding the speed reconstruction problem have been synthesized. This problem required a deeper understanding of the influence of the rolling shutter on the video, and offered the opportunity to transform this disturbance into an advantage, somehow increasing the acquisition frequency by a factor of about one thousand. Finally, the contest organizers proposed a formalized method to perform the modal analysis asked in the second challenge, which appeared more difficult to deal with than the first one. Once again, the rolling shutter effect could be used advantageously to precisely ascertain the natural frequencies lying beyond the video frame rate.

Just like in the medical world, where smartphones already seem to be regularly diverted from their original use to assist doctors in diagnosis, this paper highlights the possibility of using smartphone cameras to estimate fundamental characteristics of mechanical structures. The techniques allowing this diversion seem very valuable in the field of structure monitoring, which is dear to the readers of MSSP. In particular, the rolling shutter effect that shifts the acquisition of each pixel in a more or less controlled manner, hitherto considered as a nuisance inseparable from cheap image sensors, offers unexpected opportunities that are beyond the reach of more advanced and expensive systems.

[1] Joonchul Shin, Soocheol Kim, Taehee Yoon, Chulmin Joo, and Hyo-Il Jung. Smart Fatigue Phone: Real-time estimation of driver fatigue using smartphone-based cortisol detection. *Biosensors and Bioelectronics*, 136:106–111, 7 2019.

[2] Yekta Said Can, Bert Arnrich, and Cem Ersoy. Stress detection in daily life scenarios using smart phones and wearable sensors: A survey. *Journal of Biomedical Informatics*, 92:103139, 4 2019.

[3] John E. Madias. A proposal for monitoring patients with heart failure via "smart phone technology"-based electrocardiograms. *Journal of Electrocardiology*, 49(5):699–706, 9 2016.

[4] Helena A. Watson, Rachel M. Tribe, and Andrew H. Shennan. The role of medical smartphone apps in clinical decision-support: A literature review. *Artificial Intelligence in Medicine*, 100:101707, 9 2019.

[5] Sebastián Castellanos-Toro, Mario Marmolejo, Johannio Marulanda, Alejandro Cruz, and Peter Thomson. Frequencies and damping ratios of bridges through Operational Modal Analysis using smartphones. *Construction and Building Materials*, 188:490–504, 11 2018.

[6] Amir H. Alavi and William G. Buttlar. An overview of smartphone technology for citizen-centered, real-time and scalable civil infrastructure monitoring. *Future Generation Computer Systems*, 93:651–672, 4 2019.

[7] Marco Grossi. A sensor-centric survey on the development of smartphone measurement and sensing systems. *Measurement*, 135:572–592, 3 2019.

[8] Didier Rémond, Jérôme Antoni, and Robert B. Randall. Editorial for the special issue on Instantaneous Angular Speed (IAS) processing and angular applications, 2 2014.

[9] Jérôme Antoni, Julien Griffaton, Hugo André, Luis David Avendaño-Valencia, Frédéric Bonnardot, Oscar Cardona-Morales, German Castellanos-Dominguez, Alessandro Paolo Daga, Quentin Leclère, Cristián Molina Vicuña, David Quezada Acuña, Agusmian Partogi Ompusunggu, and Edgar F. Sierra-Alonso. Feedback on the Surveillance 8 challenge: Vibration-based diagnosis of a Safran aircraft engine. *Mechanical Systems and Signal Processing*, 97:112–144, 12 2017.

[10] Quentin Leclère, Hugo André, and Jérôme Antoni. A multi-order probabilistic approach for Instantaneous Angular Speed tracking debriefing of the CMMNO'14 diagnosis contest. *Mechanical Systems and Signal Processing*, 81:375–386, 12 2016.

[11] Frédéric Bonnardot. Influence of speed fluctuation on cepstrum. *Mechanical Systems and Signal Processing*, 119:81–99, 3 2019.

[12] Cédric Peeters, Quentin Leclère, Jérôme Antoni, Peter Lindahl, John Donnal, Steven Leeb, and Jan Helsen. Review and comparison of tacholess instantaneous speed estimation methods on experimental vibration data. *Mechanical Systems and Signal Processing*, 129:407–436, 8 2019.

[13] Charles r. Farrar, Timothy w. Darling, Albert Migliori, and William e. Baker. Microwave interferometers for non-contact vibration measurements on large structures. *Mechanical Systems and Signal Processing*, 13(2):241–253, 3 1999.

[14] Anthony B. Stanbridge, David J. Ewins, and A. Z. Khan. Modal Testing Using Impact Excitation and a Scanning LDV. *Shock and Vibration*, 7:527389, 2000.

[15] Justin G. Chen, Neal Wadhwa, Young-Jin Cha, Frédo Durand, William T. Freeman, and Oral Buyukozturk. Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration*, 345:58–71, 6 2015.

[16] Davide Scaramuzza, Agostino Martinelli, and Roland Siegwart. A toolbox for easily calibrating omnidirectional cameras. *IEEE International Conference on Intelligent Robots and Systems*, 10 2006.

[17] Zhongwei Tang, Rafael Gioi, Pascal Monasse, and Jean-Michel Morel. High-precision camera distortion measurements with a "calibration harp". *Journal of the Optical Society of America. A, Optics, image science, and vision*, 29:2134–43, 10 2012.

[18] Toby Terpstra, Seth Miller, and Alireza Hashemian. An evaluation of two methodologies for lens distortion removal when exif data is unavailable. *SAE Technical Paper*, (2017-01-1422), 03 2017.

[19] Denis Zorin and Alan H. Barr. Correction of geometric perceptual distortions in pictures. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH 95, pages 257–264, New York, NY, USA, 1995. Association for Computing Machinery.

[20] M. A. Tehrani, A. Majumder, and M. Gopi. Correcting perceived perspective distortions using object specific planar transformations. In *2016 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10, 2016.

[21] Y. Takezawa, M. Hasegawa, and S. Tabbone. Camera-captured document image perspective distortion correction using vanishing point detection based on radon transform. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 3968–3974, 2016.

[22] The MathWorks, Inc., 1 Apple Hill Drive Natick, MA 01760-2098. *Matlab's Computer Vision Toolbox User's Guide*, 2020.

[23] Shudong D. Yu and Xuan Zhang. A data processing method for determining instantaneous angular speed and acceleration of crankshaft in an aircraft engine–propeller system using a magnetic encoder. *Mechanical Systems and Signal Processing*, 24(4):1032–1048, 2010.

[24] Derek Bradley and Gerhard Roth. Adaptive Thresholding using the Integral Image. *Journal of Graphics Tools*, 12(2):13–21, 1 2007.

[25] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[26] Alessandro P. Daga and Luigi Garibaldi. GA-Adaptive Template Matching for Offline Shape Motion Tracking Based on Edge Detection: IAS Estimation from the SURVISHNO 2019 Challenge Video for Machine Diagnostics Purposes. *Algorithms*, 13(2):33, 1 2020.

[27] Alessandro Rivola and Marco Troncossi. Zebra tape identification for the instantaneous angular speed computation and angular resampling of motorbike valve train measurements. *Mechanical Systems and Signal Processing*, 44(1-2):5–13, 2014.