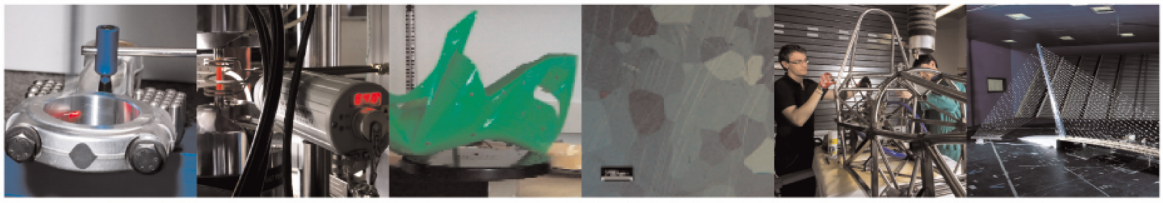
POLITECNICO
MILANO 1863

DIPARTIMENTO DI MECCANICA



Video Motion Magnification to Improve the Accuracy of Vision-Based Vibration Measurements

Perez E.; Zappa E.

This is a post-peer-review, pre-copyedit version of E. Pérez and E. Zappa, "Video Motion Magnification to Improve the Accuracy of Vision-Based Vibration Measurements," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-12, 2022, Art no. 5012712, doi: 10.1109/TIM.2022.3175977

Video motion magnification to improve the accuracy of vision-based vibration measurements

Eduardo Pérez and Emanuele Zappa *Senior Member, IEEE*

***Abstract*—Video motion magnification allows to turn a tiny movement recorded on video into an easily visible displacement. In this work, such techniques are employed to decrease the uncertainty of vision-based motion measurements. Previous works have already combined motion magnification with vision-based measurement techniques, such as digital image correlation. This present study analyzes its applicability on simpler measurement techniques: pattern matching and edge detection, analyzing the accuracy improvements when measuring on a motion-magnified video. In addition, the suitability for these purposes of recent motion magnification algorithms is tested, and the circumstances under which the magnification is beneficial are studied. As a result of the experiments, it is concluded that motion magnification can increase the motion measurement accuracy, especially for small displacements, which could be considered an initial step for future quantifications of its applicability range.**

***Index Terms*—Measuring small movements, motion magnification, optical sensing, vibrometry.**

I. Introduction

In this work, we study the convenience of amplifying the movement of vibrating objects in a video, before taking displacement measurements on that recording. The potential accuracy improvements are analyzed, as well as the conditions under which such improvements occur.

In this first section of the article, we overview the previous works developed by different authors on the field, in order to expose later the goals and aimed contributions of our study. Next, in Section II, it can be found a survey on the theoretical fundamentals of the motion magnification methods that are later employed for our experiments. An explanation on why those methods are used for the experimental analyses can be found at the beginning of Section III, as well as a general explanation of the common methodology for the different experiments. A more detailed explanation of the methodology of each experimental test can be found at the beginning of each subsection of Section III, since we adapt the procedure to the specific approach of each experiment. Finally, the conclusions and thoughts about future lines of work can be found in Section IV.

A. State of the art

Two main topics merge together in this present work: vision-based measurements and video motion magnification.

1) Vision-based measurements: Cameras, containing a cluster of light sensors, can be used to measure the position, size, displacement, etc., of a near element, allowing for a contactless multipoint (or even full-field) analysis with one single device. The multiple techniques for using cameras as a position measuring tool, known as vision-based measurement methods, have been widely used for many different purposes [1], such as structural monitoring [2], [3].

Many algorithms have been developed to perform such techniques. For example, pattern matching has been used for bridge monitoring by Cigada *et al.* [4], [5] and Ye *et al.* [6], [7]. Pattern matching allows to track the displacement of a selected region of interest, known as pattern or template, along the frames of a video. Another very widely used algorithm is Digital Image Correlation (DIC) [8], which is generally more precise than pattern matching, but it commonly requires to set a speckle pattern on the analyzed surface. Finally, another remarkable example to mention is edge detection, used, for instance, by Cigada *et al.* [4] for bridge vibration analysis, and by Son *et al*. [9] to measure the vibration of a pipe.

2) Motion magnification (MM): Video motion magnification, or simply motion magnification, is a group of techniques that allow to amplify the displacements in a video or in a sequence of images. Generally, the motion to amplify is the displacement between each frame and a reference frame, usually the first one. However, some techniques amplify the displacement between consecutive frames [10] or the acceleration [11]. According to the most common definition of MM [12], the displacement gets multiplied by $1+\alpha$, with α being the so-called magnification factor.

The first algorithm developed to magnify motions in a video was published in 2005 by Liu *et al.* [13]. Several advances were made after it, specially by the MIT Computer Science & AI Lab (CSAIL) [14]. A very important step-forward was the method proposed by Wadhwa *et al.* [15] in 2013, called complex steerable pyramid (CSP) method. This technique is sometimes known as the phase-based method for motion magnification, although other phase-based methods have been developed in the following years [16]. These methods are based on the decomposition of the image intensity in wavelets, through the application of convolutional filters. The phase difference between the wavelets in two different frames, an indicator of the displacement between those images, is manipulated to amplify the motion.

3) Vision-based measurements with motion magnification: Since the release of the CSP algorithm in 2013, motion magnification has been used in several works as a preprocessing technique before the vision-based measurement process. All vision-based measurement methods have a certain level of uncertainty. For instance, in pattern matching, the uncertainty is typically around 0.01 pixels [17]. This means a very low reliability when trying to measure tiny displacements. Motion magnification can solve this by amplifying the displacement,

thus decreasing the relative uncertainty. Moreover, band-pass filters can be added to the motion magnification algorithm, allowing to amplify only motions inside a frequency range, which can be useful, for example, when measuring a specific operating shape. An important drawback, however, is that motion magnification typically generates blur and artifacts on the images, which can actually mean an increase of the final measurement error.

In 2017, Molina-Viedma *et al.* [18] used the CSP algorithm in combination with DIC for the characterization of high frequency mode shapes of a cantilever bar. The authors observed that motion magnification allows to register high frequency vibrations without the need to apply high levels of excitation. They also noticed a less noisy estimated displacement profile. However, they observed that the actual level of magnification the algorithm was performing did not match well with the commanded magnification factor. These inaccuracies might lead to low reliability if the goal is to measure the amplitude of the displacement.

Also in 2017, Poozesh *et al.* [19] employed CSP magnification in combination with 3D DIC and 3D point tracking to measure the higher-frequency operating shapes of a cantilever beam and a wind turbine blade. According to the authors, the method allows to extract small vibrational response of structures, by improving the signal-to-noise ratio and obtaining the frequency of low-amplitude vibration modes. Fioriti *et al*. [20] applied motion magnification and image processing techniques to extract the vibration mode frequencies of different ancient constructions of Rome and Istambul.

Some other authors have used modifications of the CSP algorithm to directly estimate the motions in a video [21]–[23]. This allows to by-pass the magnification process that can generate blur and artifacts that later lower the reliability of the final results.

In 2018, Choi *et al.* [24] used CSP motion magnification and Canny edge detection to estimate the mode shape and natural frequencies of damaged and undamaged cantilever beams. They used the results to calculate the location and depth of the cracks of the beam with good accuracy values. However, they did not compare the results to those that would be obtained applying vision-based measurements without MM.

Also in 2018, a learning-based method for motion magnification was released by Oh *et al*. [10]. It employs a convolutional neural network (CNN), which is a deep-learning technique specially used for image processing. This method shows good results in comparison to CSP, according to its authors. Although it is available in the CSAIL website [14], it has not been publicly used as much as CSP for vibration analysis, possibly because of its novelty. Peng *et al.* [25] used learning-based motion magnification to qualitatively analyze the microvibration of high-speed rotating machinery.

Finally, Zhou *et al.* [26] very recently proposed a novel iterative algorithm to measure the vibration signals from a sequence of images.

B. Objectives

The main objective of this present work is to study whether motion magnification can be used to increase the accuracy of simple image processing methods, specifically pattern matching and edge detection. As mentioned before, pattern matching allows to track, along the frames of a video, the displacement of a template or subset. Generally, this template is a region of interest selected from the first frame or from any other image of the video.

Motion magnification has already been used in combination with DIC to obtain the frequencies and deformed shapes of the high-frequency vibration modes of several structures. This present work, on the other hand, tries to assess whether there is some improvement when measuring the whole displacement-time history of a moving object. In addition, speckle patterns, generally used for DIC, might be not optimal for motion magnification, since there is evidence that high spatial intensity frequencies (hard textures) suffer more image distortion during the magnification process than softer textures [27]. Hence, there might be more applicability with methods that do not require the use of any pattern of this kind.

As Poozesh *et al.* [19] mentioned, a drawback of motion magnification is that there is no systematic method to choose an appropriate magnification factor. Theoretically, an increase of the magnification would mean a decrease of the uncertainty. However, motion magnification methods always introduce some level of blur and artifacts in the images, and such image distortion becomes more severe when the magnification factor is increased [10]. A better understanding on the relationship between the magnification factor and the final measurement error is another goal of our work.

Finally, as mentioned before, the learning-based algorithm for MM has hardly been used to improve subsequent vision-based measurements. Analyzing the suitability of this method is another objective of this work.

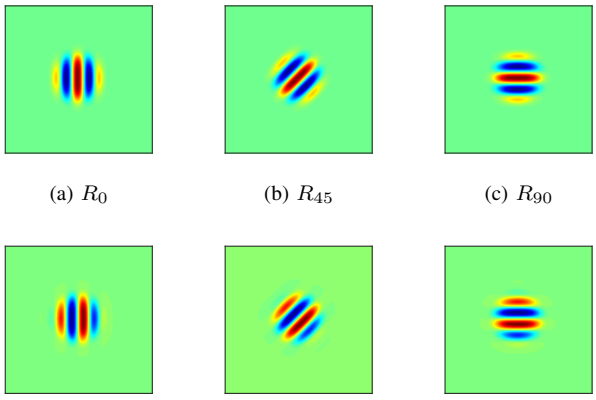
II. Motion magnification methods

A. Lagrangian vs. Eulerian approaches

A first important classification on motion magnification techniques is established between Lagrangian and Eulerian methods. Similarly to the Lagrangian fluid dynamics approach, Lagrangian motion magnification is based on tracking the displacement of image features along time and amplifying them afterwards. This technique entails several problems, like the generation of “holes” in the locations where a displaced feature was located.

On the other hand, Eulerian methods divide the video frames in regions (generally, in a spatially-multiscale way) and analyze the intensity values at each region. By operating with the intensity values, the displacement is amplified without needing to explicitly compute it. Generally, this is achieved by applying several convolutional filters to obtain a displacement-related parameter that is then manipulated to change all the displacements in each frame.

Although there are some previous works on manipulating the motion of specific marker positions, the first method to amplify motions in a video was the Lagrangian technique proposed by Liu *et al.* [13] in 2005. After that, the linear Eulerian method proposed by Wu *et al.* [28] in 2012 showed several advances in computing time and quality of the results.



(a) R_0 (b) R_{45} (c) R_{90}

(d) I_0 (e) I_{45} (f) I_{90}

Fig. 1. Examples of complex Gabor filters in different orientations. R_θ represents the real part and I_θ the imaginary part.

Because of the mentioned advantages of the Eulerian algorithms over the Lagrangian methods, Eulerian techniques have been more used in previous studies, thus developing several algorithms that follow different approaches to perform the magnification. The CSP algorithm is an example of an Eulerian approach that has been widely used in the past, while the learning-based method is a newer technique that has been claimed to generally outperform the CSP method [10]. As further explained later, these two are the techniques applied in this present work.

B. Complex steerable pyramid (CSP) algorithm

The CSP algorithm, proposed by Wadhwa *et al.* [15] in 2013, is an Eulerian approach in which the convolutional filters, applied to the frames of the video, are Gabor (or Gabor-like) filters that perform a spatial decomposition of the image intensity into wavelets. The process is analogous to the Fourier decomposition of a signal, but using self-similar Gabor wavelets (sine waves multiplied by a Gaussian envelope with a window width proportional to the wavelength) as basis functions, instead of sinusoids. Also, since the image intensity is a 2D spatial signal, the decomposition must be repeated for several orientations, as well as several scales, each one corresponding to a specific spatial frequency. Some examples of Gabor filters in different orientations can be observed in Fig. 1.

The result from applying the Gabor filters to each frame at different orientations and scales gives as a result the so-called complex steerable pyramid [29]. Then, every complex value is transformed into amplitude and phase. According to the Fourier shift theorem, the displacement of a wave corresponds to a proportional shift of its phase. This way, for every location (x, y), orientation (θ) and scale (r), the CSP algorithm computes the phase differences between each frame at time t and a reference frame (usually, the first one), and amplifies it:

$$\hat{\phi}(x,y,\theta,r,t) = \phi(x,y,\theta,r,t) + \alpha \cdot [\phi(x,y,\theta,r,t) - \phi(x,y,\theta,r,t_0)] \quad (1)$$

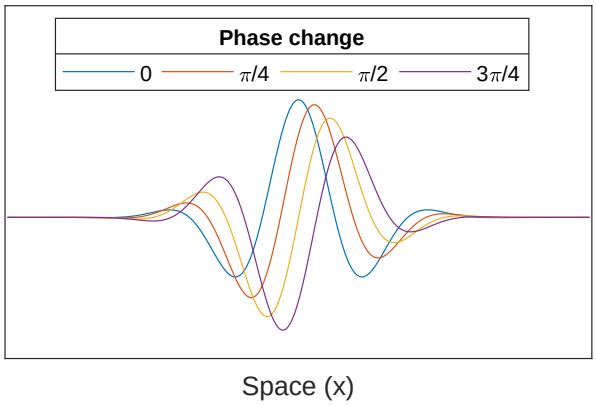


Fig. 2. Approximate translation of a 1D wavelet by means of a phase shift.

Where $\hat{\phi}$ is the corresponding amplified phase that will substitute the original phase value ϕ at time t. Amplitudes are not modified so that illumination changes do not get amplified. In order to obtain one single magnified frame from each pyramid at time t, every pyramid is collapsed by convoluting each one of its levels (each orientation and scale) to the same Gabor filters applied before, and summing together the results. The design conditions for the Gabor filters that allow for such reconstruction are stated in [27].

The reason why the image intensity is decomposed in wavelets instead of performing a simple Fourier decomposition is that sinusoids are not localized. Therefore, amplifying the displacements in a certain area of the image would mean to displace also pixels that are far from there, thus generating artifacts. However, the use of wavelets has an important drawback: when increasing the phase of the wavelet, the sinusoid component gets displaced, but the Gaussian envelope remains in the same location (Fig. 2).

This leads to a deformation of the wavelet, that can be observed as a distortion of the output image, when the displacement is very large in comparison to the width of the window of the Gaussian envelope. Since the frequency of the wavelet is inversely proportional to the width of the window, this distortion is more severe on areas where the image intensity has a high spatial frequency, such as edges and hard textures. In order to avoid this, the size of the Gaussian window is sometimes increased, although it means an increase of the processing time. Moreover, it can also lead to artifacts, since the basis functions are less localized (the decomposition is closer to a Fourier decomposition) if the size of the window is larger.

Another phase-based method for motion magnification is the Riesz pyramid algorithm, proposed by Wadhwa *et al.* [30]. This technique, similar to the CSP method, is much faster, since it does not need to perform the decomposition in multiple orientations.

C. Learning-based motion magnification

The learning-based algorithm for motion magnification was proposed by Oh *et al.* [10] in 2018, with the hypothesis that

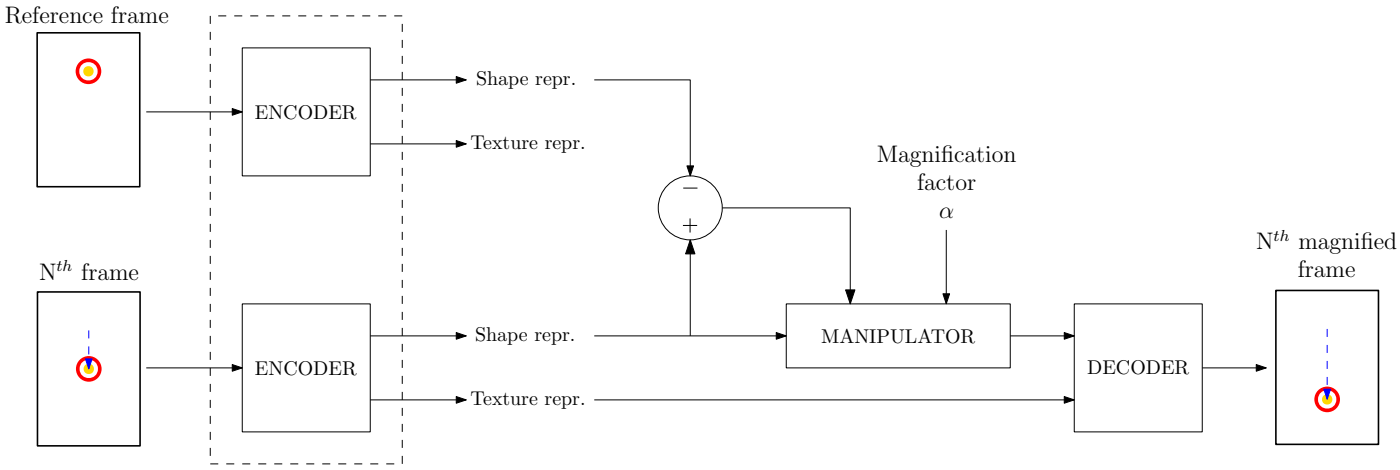


Fig. 3. Simplified overview of the CNN's architecture proposed in [10].

previous methods, based on the use of hand-designed filters, may not be optimal. In this method, the convolutional filters are subsituted by a Convolutional Neural Network (CNN), a deep-learning technique specially used for image processing [31].

The CNN is divided into an encoder, a manipulator and a decoder. Two images are always the input to the system. One image is the frame to magnify, and the other one is the reference with respect to which the displacement is amplified. That reference frame can be the first one in the video (static mode) or the previous frame (dynamic mode), as set by the user.

First, the encoder extracts from the images a position-related component called shape representation. On the other hand, the non-position-related counterpart is called texture representation. The shape representation is analogous to the phase in the CSP algorithm, and the texture representation is analogous to the amplitude. Then, the encoder modifies the shape representation to magnify the motion in the corresponding image, according to the commanded magnification factor and the difference between the shape representations of the image to magnify and the reference frame. Finally, the decoder reconstructs the image with the modified shape representation and the original texture representation (Fig. 3). The process is repeated for every frame of the video.

The CNN was trained by the authors with a large dataset of artificial videos, with their corresponding true-magnified versions as expected output. During the training, the CNN is adjusted to minimize a loss function that considers, among other factors, the difference between each frame magnified by the CNN and the corresponding ground-truth magnified frame.

A drawback of this technique is the opacity of its network, since it is not based on hand-designed filters that follow a known analytical model. For a better understanding of the inner working of the network, Oh *et al.* [10] analyzed the response of the encoder when passing various test images. The authors concluded that the extraction of the shape representation can be approximated as the application of Gabor-like and Laplacian-like filters, suggesting that the working of that section of the network is similar to the CSP algorithm.

III. Experiments and Results

As stated before, the main purpose of this work is to study whether motion magnification can be used to improve subsequent vision-based measurements taken on the processed video. This topic has been analyzed in the past, using the CSP algorithm for MM in combination with digital image correlation [18], [19], although focusing mainly on measuring the vibration frequency. Choi *et al.* [24] used CSP with edge detection for structural health analysis, but they did not quantify the benefit of applying MM. The learning-based algorithm, proposed by Oh *et al.* [10], generally outperforms the CSP algorithm in terms of image distortion, according to its own authors. The only exceptions, related to the algorithm not magnifying some displacements, were observed when a band-pass filter was applied.

For this reason, the main MM algorithm used in the following experiments is the learning-based method without band-pass filter, using also the CSP algorithm developed by Wadhwa *et al.* [15] in some of the tests to compare the results. The algorithm for learning-based MM developed by Oh *et al.* can be found online [14], including the already-trained CNN, employed for the experiments of this article. The full algorithm for CSP by Wadhwa *et al.* is also available online on the same site [14], including the Gabor filters, for which the size of the Gaussian window with respect to the period of the wavelet (full octave, half octave or quarter octave) can be chosen. We decide to use the half-octave bandwidth filters as the authors of the algorithm stated that it generally gives better results [15].

The Riesz pyramid algorithm can be discarded since its main advantage over the CSP algorithm is a lower processing time, which is not a priority for our purposes. Furthermore, although the Riesz pyramid method has been proved to provide similar image quality comparing to the CSP method with an octave bandwidth pyramid [30], the CSP algorithm generally gives better results with a half-octave bandwidth pyramid, as stated before.

For the measurement stage, we apply pattern matching and edge detection, as explained in the Objectives chapter of

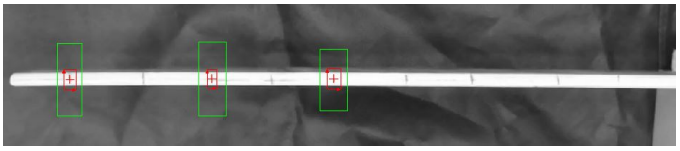
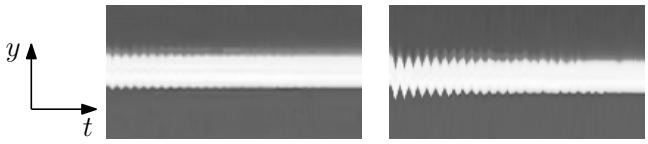


Fig. 4. Templates (red) and search areas (green) for pattern matching on the wooden beam.



(a) Original (b) Magnified 5×

Fig. 5. Motion of the tip of the beam.

this article (Section I-B). For that purpose, we use National Instruments built-in algorithms for all the experiments. Further explanation on the input parameters and the functioning of the employed algorithms can be found in the following subsections and in [32].

Four experiments are carried out, among which the first three are developed to obtain a first validation of the suitability of MM, while the last one is aimed to be a more complete and deep analysis on the topic.

A. Motion magnification to compensate for low reliability conditions for pattern matching

In this experiment, it is studied whether motion magnification can be used to improve the accuracy of pattern matching and to compensate for two low reliability conditions: selection of pattern matching algorithm parameters that prioritize computing speed over accuracy, and low image resolution.

To perform both tests, the transient vibration of a wooden cantilever beam is recorded. The beam has several vertical marks to ease the pattern matching process (Fig. 4). Pattern matching is applied on 3 of those marks.

To record the vibrating beam, a Huawei p30 lite smartphone camera (1280×720) [33] is used. It is important to remark that the goal of the experiment is not to obtain a very low measurement error, but to study whether, in the conditions of this experiment, the error decreases after applying MM. Hence, the characteristics of the camera and the quality of the acquired images do not affect the generality of the results.

1) Test 1: changing the pattern matching parameters: The software used along this project to perform pattern matching is Vision Assistant by National Instruments. The employed built-in algorithm is based on the computation of the cross correlation matrices between the template and each region of a search area defined by the user. An explanation on the theoretical basis of pattern matching (named "2-D prototype matching") and implementation issues can be found [34]. Also, technical documentation on the specific functioning of pattern matching in Vision Assistant can be found in [32].

Vision Assistant allows to use 4 different preset parameters with different levels of accuracy, named according to the priority: accurate, fast, very fast and default. These parameters are related to the subpixel accuracy, the number of iterations and the matching score computation, among other factors. A description of some of these factors, and their relationship with respect to the final accuracy, are provided in [32], [35].

Assuming that the "accurate" parameters lead to much more reliable results, it can be considered as a "ground truth" in comparison to the "very fast" pattern matching. Therefore, the pattern matching process is performed first in the "very fast" mode for both the magnified and the original video. Then, the "accurate" pattern matching is applied only on the original video. If the results from the "very fast" mode on the magnified video match better the "accurate" results than the original, it can be concluded that, at least when using the "very fast" parameters, magnifying the video can lead to more accurate results.

After recording the video, its motion is amplified by 5 (Fig. 5), by means of the learning-based algorithm. This magnification factor was chosen after several trials to find a value that produces a significant motion magnification while avoiding a too high image distortion, this is, assessing qualitatively that the features of the search area are well preserved and that the pattern matching algorithm is capable of finding the template along all the video, despite the artifacts introduced after applying MM. No band-pass filter is applied.

After applying pattern matching on the three areas for both videos, the displacements with respect to the first frame are computed and the results coming from the magnified video are re-scaled to obtain real displacement. The results can be observed in Fig. 6. It can be observed that the measurements coming from the magnified video match better the "accurate" pattern matching than the results from the original video with "very fast" pattern matching.

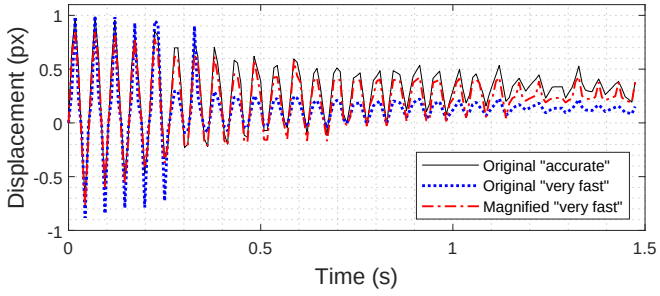
In order to quantify the resulting differences, the root-mean-square deviation (RMSD) for both results from "very fast" pattern matching are computed with respect to the "accurate" pattern matching. The results can be seen in Table I. The normalized RMSD (dividing the RMSD by the mean absolute displacement according to the "accurate" pattern matching) is shown in brackets.

It can be observed that, for the three considered templates, the magnified video gives results that are closer to the "accurate" pattern matching. The RMSD reduction increases going from left (free end) to right (constrained), indicating that magnifying the motion might be specially useful when the displacements are smaller. It can be concluded that applying motion magnification seems to partially compensate for the selection of pattern matching algorithm parameters that prioritize computing speed over accuracy. In other words, at least when using the "very fast" pattern matching, magnifying the motion may produce more accurate results.

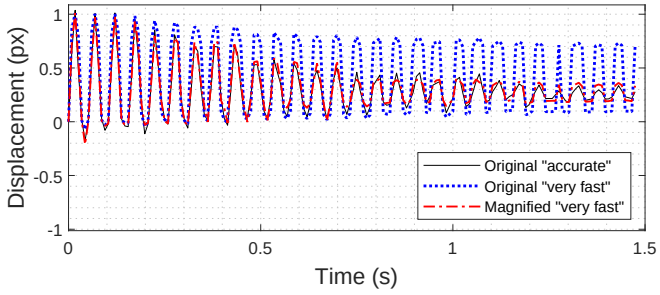
Note that applying MM and then the "very fast" pattern matching is more time-consuming than just using the "accurate" pattern matching. In this sense, it is important to remark that the aim of this first experiment is not to find a practical use case for MM, but to obtain a first validation of MM to enhance pattern matching measurements. For the following experiments, the accuracy parameters are always set

TABLE I
RMSD OF THE MEASUREMENTS WITH "VERY FAST" PATTERN MATCHING WITH RESPECT TO "ACCURATE" PATTERN MATCHING.

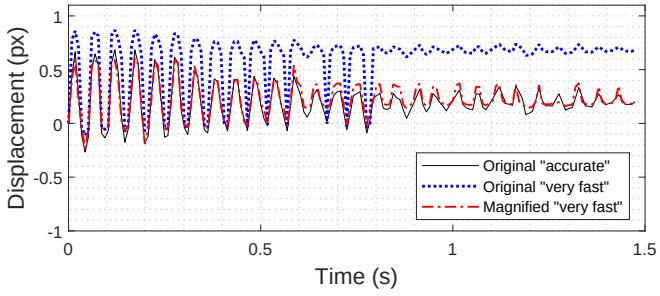
Video	Left	Middle	Right
Original	$212{\cdot}10^{-3}$ px (60.5%)	$260{\cdot}10^{-3}$ px (88.3%)	$425{\cdot}10^{-3}$ px (191%)
Magnified 5×	$80.7{\cdot}10^{-3}$ px (23.0%)	$42.1{\cdot}10^{-3}$ px (14.3%)	$52.3{\cdot}10^{-3}$ px (23.4%)



(a) Left template.

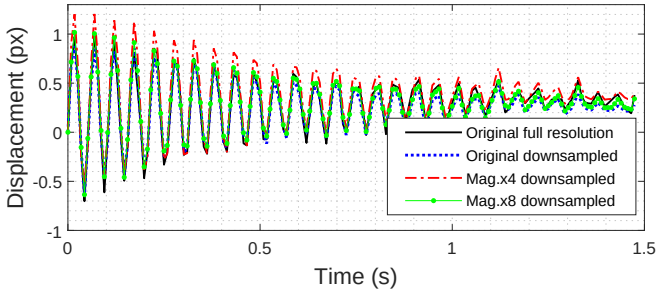


(b) Middle template.

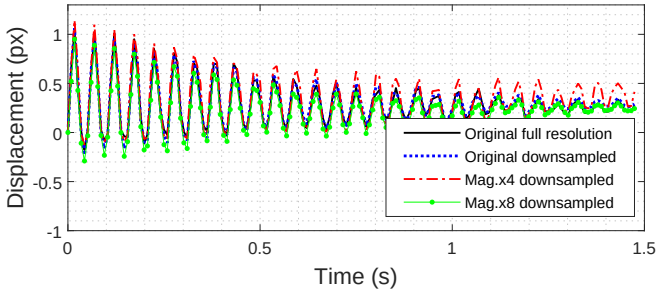


(c) Right template.

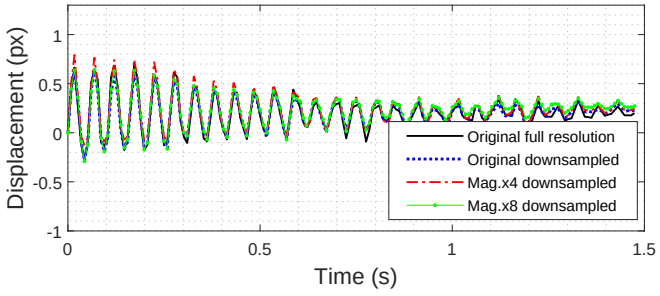
Fig. 6. Measured displacement of the three templates on the three studied conditions.



(a) Left template.



(b) Middle template.



(c) Right template.

Fig. 7. Measured displacement of the three templates on the four studied conditions. All measurements have been re-scaled to compensate for the resolution change and the magnification.

to "accurate."

2) Test 2: changing the image resolution: A lower image resolution decreases the reliability of pattern matching measurements [17]. For this test, all the recorded frames are downsampled by 4 in both height and width. Then, two magnified videos are generated, multiplying the displacements by 4 and 8. No band-pass filter is applied.

Pattern matching with parameters set to "accurate" is used on the original and the two magnified videos, on the same three points from the previous test. The displacements hereby obtained are re-scaled to compensate for the resolution change and the motion magnification, in order to compare the results to the measurements obtained from the original video with no resolution change. This way, if the displacements coming from the magnified videos match better the measurements coming from the original video with full resolution, it could be concluded that, at least with low resolution recordings, magnifying the videos can increase the accuracy.

The measured displacements can be visualized in Fig. 7. It can be observed that, in this case, the results from the non-magnified low-resolution video match better the reference (original, full resolution) than when changing the pattern matching parameters. In order to compare the results, the RMSDs of the resulting displacements in comparison to the reference were computed, as it can be observed in Table II. It

TABLE II
RMSD OF THE MEASUREMENTS FROM THE DOWNSAMPLED VIDEOS WITH RESPECT TO THE ORIGINAL FULL-RESOLUTION VIDEO.

Low-resolution video	Left	Middle	Right
No MM	$66.7{\cdot}10^{-3}$ px (19.0%)	$40.2{\cdot}10^{-3}$ px (11.5%)	$41.4{\cdot}10^{-3}$ px (11.8%)
Magnified 4×	$110{\cdot}10^{-3}$ px (31.4%)	$98.4{\cdot}10^{-3}$ px (28.1%)	$51.8{\cdot}10^{-3}$ px (14.8%)
Magnified 8×	$54.2{\cdot}10^{-3}$ px (15.4%)	$78.8{\cdot}10^{-3}$ px (26.7%)	$53.4{\cdot}10^{-3}$ px (23.9%)

is noticeable that, in this case, it cannot be concluded that magnifying the motion in low resolution videos generally increases the measurement accuracy. A possible explanation is that a lower resolution affects also negatively the motion magnification process.

It is also worth to mention that Popek and Iskander [36] stated that motion magnification could lead to incorrect frequency response when dealing with multiple-frequency vibration, which is the case of the two tests exposed within this first experiment (transient vibration). However, they only mention CSP and other phase-based techniques, not the learning-based method here employed. For all of the following experiments, the measured motion is a single-frequency vibration.

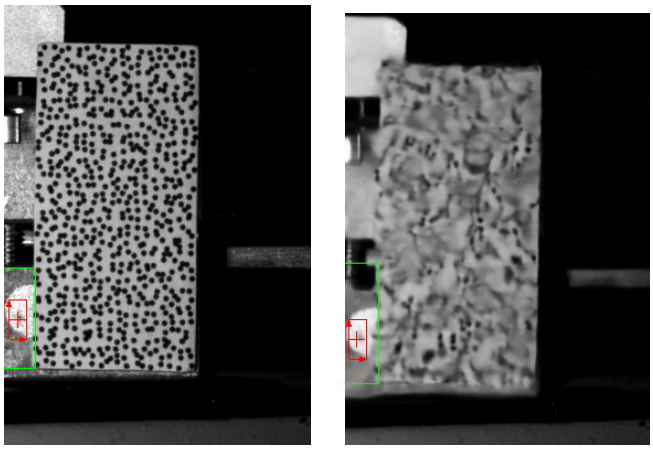
B. Analysis of pattern matching accuracy improvements with various MM methods

In this experiment, it is studied whether motion magnification can increase pattern matching accuracy under already good reliability conditions. Moreover, three methods are used for the magnification, in order to compare their results: CSP algorithm, and learning-based algorithm with static mode and with dynamic mode.

The video for this analysis is a laboratory recording that shows a shaker making a mass vibrate with constant frequency (28Hz) and amplitude. The video had been previously recorded for a different project where it was used to apply digital image correlation (DIC), for which there is a speckle pattern located on the mass. The camera used to record the video is a MIKROTRON EoSens mini2. The lens is a Zeiss ZE/ZF.2 Makro-Planar T* 50mm. The image resolution is 1696×500 pixels and the frame rate is 300 fps. The employed section of the video is composed of 490 frames (1.63s). An accelerometer (1.8 kHz sampling rate) is mounted on the mass. Its results are used as reference for the analysis after being integrated twice, dividing by the angular frequency squared.

The motion of the video is amplified by 2.5. In order to calibrate the images, a grid of blobs with 20mm spacing had been placed in the scene. Using Vision Assistant, this grid is used to estimate the scale to be 0.1869 mm/pixel. After that, pattern matching is applied on both the original and the magnified videos. Generally, it is better to use hard textures, such as the speckle pattern, as template for pattern matching [37]. However, in this case, it is observed that, in the original video, the results match better the reference[1] when using the template shown in Fig. 8. Moreover, in the magnified videos,

[1]Specifically, the measurement error is 17.4% higher when using the speckle pattern as template.



(a) First frame. (b) Last frame.

Fig. 8. Pattern matching on the vibrating mass. Video magnified with the learning-based algorithm, static mode.

the distortion is specially severe on the speckle pattern (Fig. 8). The reason is probably the high spatial frequency of the image intensity in that region of the image, which means a higher distortion for the same displacement, as explained in the previous section II-B.

The estimated displacements can be visualized in Fig. 9. The motion amplitude measured by the accelerometer is 1.127 mm. The estimated motion amplitudes[2] from the videos, as well as the errors with respect to the accelerometer data, can be observed in Table III (no rotation). It can be noticed that, with any of the three MM methods, the results are more accurate than before the magnification. The best results are obtained with the learning-based algorithm with static mode. Applying this method reduces the error from 172 μm to 30.1 μm, a decrease of about 80%.

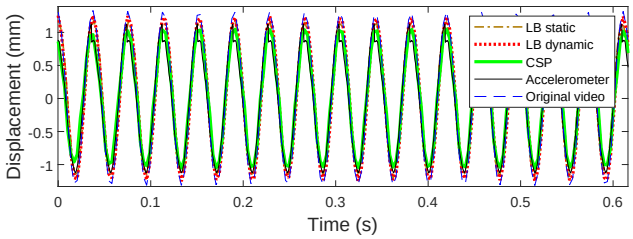
Then, the same video and the accelerometer estimations are used for a brief analysis on whether the performance of MM depends on the orientation of the displacement. To accomplish that, the video is rotated 60 degrees counterclockwise and then magnified employing the same algorithms and parameters from the previous test. Pattern matching is applied on the rotated magnified and non-magnified videos, using the same template.

The results can be observed in Table III (60° rotation). Although this just attempts to be a preliminary analysis, the

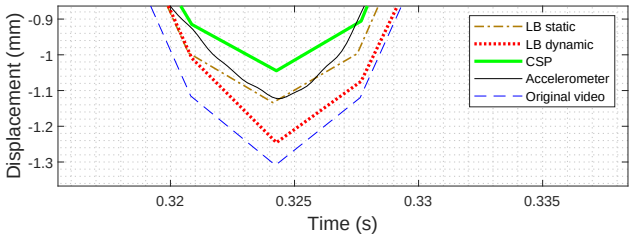
[2]The amplitude of the accelerometer displacement is directly computed as half the mean peak-to-peak value, while, for the estimations on the videos, the single-frequency signal is reconstructed through least-squares fitting to avoid inaccuracies due to the lower sampling rate.

TABLE III
COMPARISON IN TERMS OF ACCURACY OF THE ESTIMATED AMPLITUDES IN THE ORIGINAL AND MAGNIFIED VIDEOS.

	No rotation			60° rotation		
	Estimated amplitude (mm)	Error (mm)	Error change with respect to "no MM"	Estimated amplitude (mm)	Error (mm)	Error change with respect to "no MM"
No MM	1.299	0.172	.....	1.240	0.113	.....
LB static	1.157	0.030	-82.56%	1.110	-0.017	-85.16%
LB dynamic	1.243	0.116	-32.56%	1.172	0.045	-60.73%
CSP	1.037	-0.090	-47.67%	0.983	-0.144	+27.05%

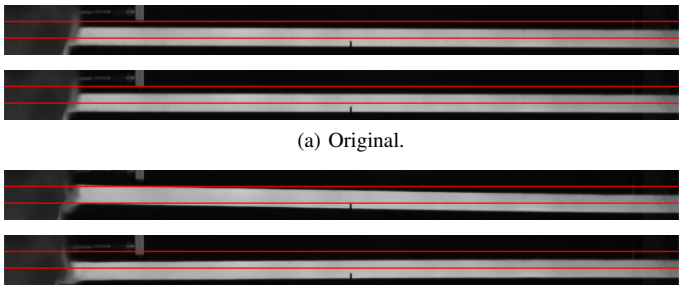


(a) Time history.



(b) Detailed view.

Fig. 9. Estimated displacement of the mass.



(a) Original.

(b) Magnified (10×).

Fig. 10. Frames at highest and lowest position of the beam.

effect of the direction does not seem to be very relevant for the performance of the learning-based algorithm with static mode, used in all the experiments of our work. In all the other tests of the manuscript we analyze the vertical motion of the targets, although some minor horizontal displacement can be present because of the imperfect alignment of the camera and due to the flexibility of the recorded objects.

C. A first analysis of MM to improve edge detection displacement estimations

Edge detection is a vision-based technique that allows to locate the edges of an object. With two video frames, the displacement profile along the edge between both instants can be computed. The purpose of this experiment is to study whether MM can improve the reliability of displacement measurements obtained through edge detection.

TABLE IV
AVERAGE FITTING RMSE FOR THE UPPER EDGE DISPLACEMENT.

Polynomial fitting	RMSE (px) before MM	RMSE (px) after MM	Difference
2nd degree	0.0349	0.0197	-43.6%
3rd degree	0.0336	0.0177	-65.2%
4th degree	0.0331	0.0163	-50.8%

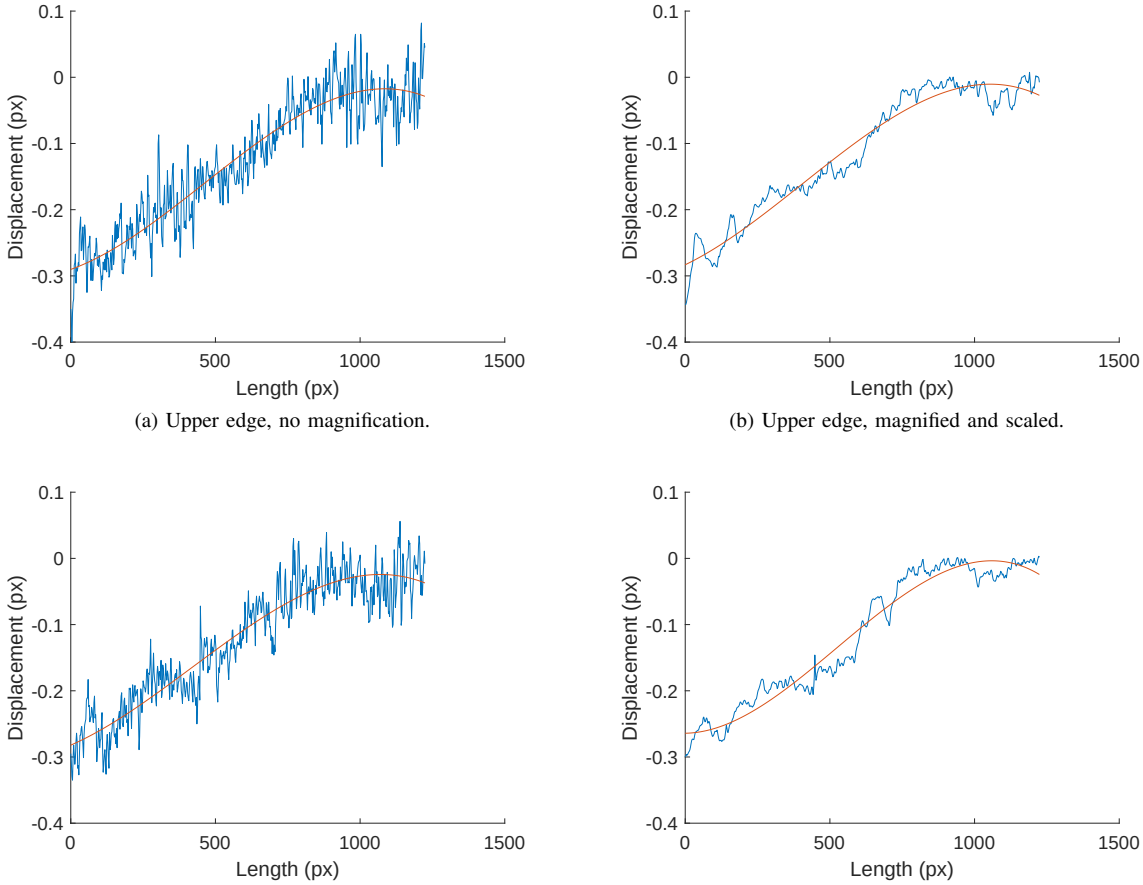
TABLE V
AVERAGE FITTING RMSE FOR THE LOWER EDGE DISPLACEMENT.

Polynomial fitting	RMSE (px) before MM	RMSE (px) after MM	Difference
2nd degree	0.0305	0.0190	-37.7%
3rd degree	0.0298	0.0172	-42.3%
4th degree	0.0294	0.0162	-44.9%

The video for this test shows a notched cantilever beam being subjected to a vibration generated by a hammer bumping on a point close to the free extreme with constant frequency. The learning-based algorithm is used to magnify the motion of the 100 video frames, amplifying the displacements by 10 (Fig. 10). No band-pass filter is applied.

The software used to apply edge detection and to track the edges of the beam is LabVIEW, specifically the block "IMAQ Rake 3", which includes a built-in function for locating the edges through parallel rake search. The algorithm finds edge points, based on contrast values along a set of parallel straight lines defined inside a rectangular region of interest. An edge point is assumed to be located where the contrast value along the search line is higher than a threshold defined by the user, considering also an internal hysteresis value [38]. A more detailed explanation on the algorithm and the contrast calculation is available in [32]. Theoretical insight on edge detection, including this approach and more complex methods, can be found in [34].

The algorithm parameters are set to extract one point for each edge every two pixels in the longitudinal direction. The region of interest is adjusted to include both the upper edge and the lower edge, avoiding the hammer. The other algorithm parameters are adjusted through iteration and qualitative analysis to ensure that continuous edge profiles are found for the original non-magnified video. Then, the same parameter values are used when finding the edges in the magnified video. After estimating the edge points, the results are smoothed by means



(a) Upper edge, no magnification.

(b) Upper edge, magnified and scaled.

(c) Lower edge, no magnification.

(d) Lower edge, magnified and scaled.

Fig. 11. Third-degree polynomial fitting of the edge displacements at the frame number 60.

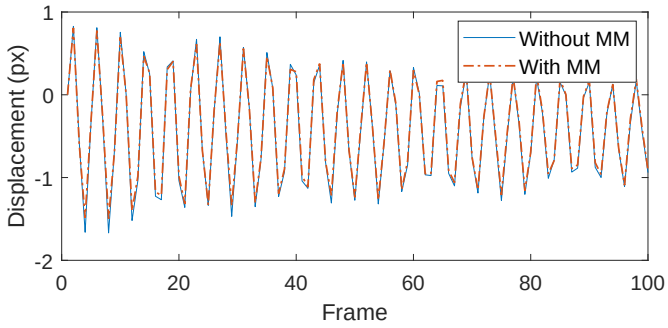


Fig. 12. Vertical displacement of the upper edge's 4^{th} pixel (counted from the leftmost extreme).

of a 5-point running average. 614 points were estimated for each edge, from which 4 points were discarded later to remove the notch. The results coming form the magnified video are re-scaled dividing the displacement by the magnification.

Complex expressions have been proposed to model the deflected shape of a vibrating cantilever beam, such as the Euler-Bernoulli formula [39]. However, due to limitations on the knowledge of the characteristics of the beam, and for the sake of simplicity, we consider the displacement profile of the notched beam to be a polynomial function, in order to assess the reliability of the measurements. A similar assumption was also made by Molina-Viedma *et al.* [18] to analyze the reliability of MM-enhanced DIC measurements on a cantilever beam. A discussion on the accuracy of employing a 4th degree polynomial function as the estimated displacement profile of a cantilever beam can be found in [40].

This way, the polynomial fitting error can be used to estimate the reliability of the results. In Fig. 11, an example of 3rd-degree polynomial fitting of the estimated displacements is shown. It can be observed that the displacement profile after applying MM is much less noisy.

For every frame of both videos, the polynomial fitting error (RMSE) is computed. After that, the average error is obtained for each video. The resulting average fitting errors, for several polynomial degrees, are shown in Tables IV and V. It can be noticed that, regardless the employed polynomial, the fitting error is lower after applying MM, with error reductions of 65.2% and 42.3% for 3rd-degree polynomial fitting. .

It could be discussed whether the displacement profiles after MM might be less noisy, but also less accurate. For example, a discrepancy between the commanded magnification factor and the actual magnification would produce inaccurate results despite the less noisy profile. To analyze if this is true, the displacements from both videos are compared in Fig.

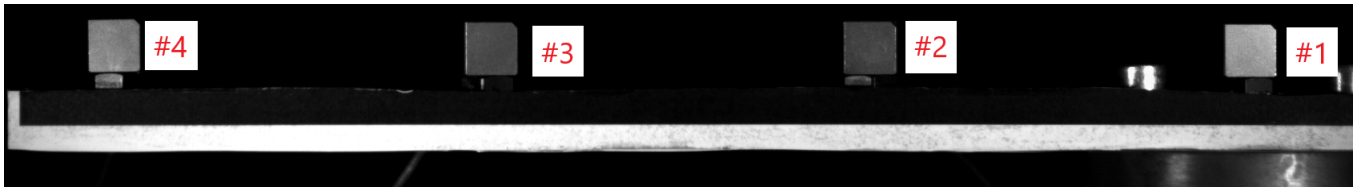


Fig. 13. Aluminum beam and accelerometers used for the experiments.

Fig. 14. Setup for the aluminum beam experiments.

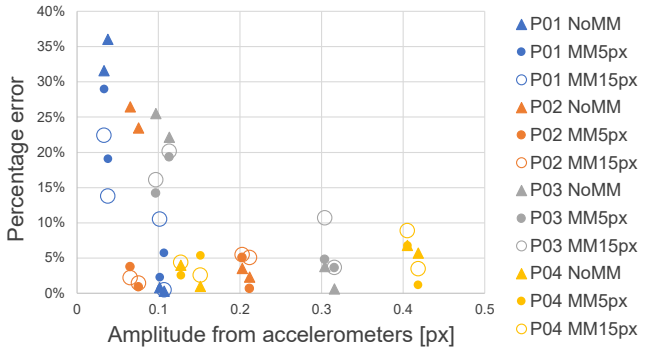


Fig. 15. Relative error as a function of the amplitude in px before magnification.

12, which shows the motion estimated at a randomly chosen location (upper edge, 4^{th} pixel from left to right). Regardless which estimation is more accurate, it can be observed that the differences are not very significant, compared to the different noise levels.

D. Analyses on the conditions for MM applicability

1) Study of the relationship between displacement amplitude and MM suitability with edge detection: After observing in the previous analysis that MM may improve edge detection displacement measurements, the goal of this present experiment is twofold: analyzing further the suitability of MM to improve edge detection estimations, and studying whether this improvement is greater when the original displacement is smaller, as observed in Section III-C for pattern matching.

To accomplish this, a 200×30×10 mm vibrating aluminum beam is recorded with the same model of MIKROTRON highspeed camera used for the experiment explained in Section III-C and the same lens with aperture set to f/2. The recording rate is 600 fps. The beam is supported on a shaker on the right hand side of the images, and it is free to vibrate at the other end. The shaker vibrates with constant frequency and amplitude. Four accelerometers are mounted on the beam, numbered from #1 (rightmost, above the shaker), to #4 (leftmost, free end of the beam). They are used as ground truth data for the experiment. Fig. 13 shows the aluminum beam and the 4 accelerometers, and Fig. 14 includes a picture of the setup. A grid with known distance between circle holes is used to estimate the scaling factor as 0.1470 mm/px. A black and white piece of paper is attached to the beam to later apply edge detection on the border line (Fig. 13).

Four tests, named P01, P02, P03 and P04, are performed making the beam vibrate at increasing amplitudes. The vibration frequency along all the tests is 80Hz. For each of the 4 corresponding recordings, 2 magnified videos are generated using different magnification factors. According to [10], for a given video, the image distortion generated by MM depends on the magnified displacement amplitude. Therefore, in order to generate similar quality videos regardless the original amplitude of the vibration, the applied magnification factors are chosen to generate a certain amplified amplitude. Specifically, each of the 4 recordings is amplified to the point at which the vibration amplitude of the 4th accelerometer (free end of the beam) is equal to 5 px (first magnified video) and 15 px (second magnified video).

Edge detection is used to estimate the vibration amplitude at 4 points corresponding to the horizontal positions of the accelerometers on the beam, along each one of the 4 original recordings and their magnified versions. Then, the measurement errors are computed using the accelerometer data as reference. It is important to consider that amplitudes at positions #2 and #3 are significantly smaller than #1 and #4, allowing for a more complete analysis on the relationship between displacement amplitude and MM suitability.

Fig. 15 represents the results from this analysis, showing the relative measurement errors as a function of the vibration amplitude before applying MM. The amplitude is expressed

TABLE VI
VIBRATION AMPLITUDES AS MEASURED ON THE ORIGINAL IMAGES (NO MM) AND BY THE ACCELEROMETERS, AS WELL AS PERCENTAGE ERRORS BEFORE AND AFTER MM, AND OPTIMAL MAGNIFICATION FACTOR (α OPT) FOR MINIMUM ERROR. ORDERED FROM SMALLEST TO LARGEST AMPLITUDE.

Position and test	Amp before MM [px]	Amp [mm]	%err before MM	α opt (i.e. min %err)	%err with MM (α opt)
#3.P05	0.00158	0.00023	801.3%	10	2.9%
#3.P06	0.01160	0.00171	82.9%	10	33.3%
#3.P07	0.01905	0.00280	46.1%	5	1.3%
#4.P05	0.06577	0.00967	2.7%	30	1.1%
#2.P05	0.06911	0.01016	7.1%	30	1.5%
#1.P05	0.13299	0.01956	5.0%	15	0.4%

Fig. 16. Aluminum beam with circle targets and pattern matching marks.

in px for a higher generalizability of the results. It can be observed that, for amplitudes smaller that 0.07 px, MM strongly reduces the measurement error. From 0.1 px to 0.3 px, mixed results are observed, while from 0.3 px on, MM seems to actually increase the error. Therefore, the evidence suggest that it is more convenient to apply MM on small displacements.

2) Study of the relationship between accuracy gain and magnification factor with pattern matching: As stated in [19], a drawback of MM is that there is no well-established method to decide what magnification factor to apply. An increase of the displacement to measure would theoretically decrease the uncertainty, but MM can also generate some blur and artifacts, which may cause measurement errors. According to Oh *et al.* [10], a factor that increases the distortion is the size of the magnified displacement. The complex relationship between MM and image distortion, and the effect of this distortion on the final measurement error, make it difficult to predict whether or not applying a specific magnification factor would be beneficial.

The goal of this final experiment is to have a first idea of what amplification factor to apply, depending on the original true motion amplitude. For this purpose, the same experimental set used for the previous experiment is employed, making the aluminum beam vibrate with 100Hz frequency while recorded in a first test named P05. With this frequency, the vibration amplitude at point #3 is much smaller than the other three (see Table VI) because of the excited mode shape. Hence, two more tests are carried out to cover the data gap, making the beam vibrate with larger amplitudes and the same 100Hz frequency. In these tests, named P06 and P07, only point #3 is analyzed.

In order to apply pattern matching, a paper with circular targets is placed along the beam, as shown in Fig. 16. The motion at each accelerometer position is estimated through a linear interpolation of the two closest targets. Each recording is magnified using several magnification factors α ranging from 0 (no expected amplification) to 50.

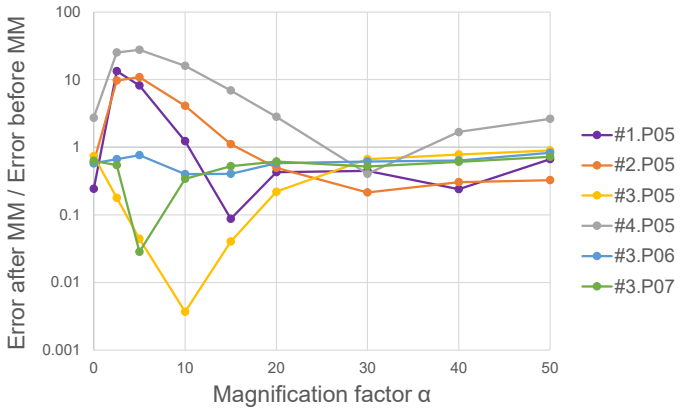


Fig. 17. Measurement error change as a function of the applied magnification factor.

Table VI shows the amplitudes measured on the original images, the accelerometer measurements (reference), and the percentage error of the vision-based amplitude measurements, both in case of the original movie (i.e. no MM) and with MM with optimal α value (minimum error). Fig. 17 shows the ratio of the error after applying MM over the error before applying it. It can be observed that applying MM with $\alpha = 0$ produces different results than without MM. This may suggest that the tasks performed by the encoder and the decoder of the convolutional neural network (decomposition and re-composition of the images) are not perfectly symmetrical.

For any of the analyzed points, there is at least one magnification factor that decreases the measurement error. However, the trends shown by the error as a function of the magnification factor are very different depending on the original displacement size. For amplitudes larger than 0.06 px before MM (#1.P05, #2.P05, #4.P05), applying a low magnification factor seems to generally increase the error. The error seems to decrease when increasing further the magnification, finding, for the three studied cases, a minimum value when the amplified displacement is around 2 px. For instance, the relative error at #1.P05 goes from 5% (no MM) to 0.4% ($\alpha = 15$), as shown in Table VI. Nevertheless, since a randomly chosen magnification factor may increase the error in these cases, applying MM would not be recommended under

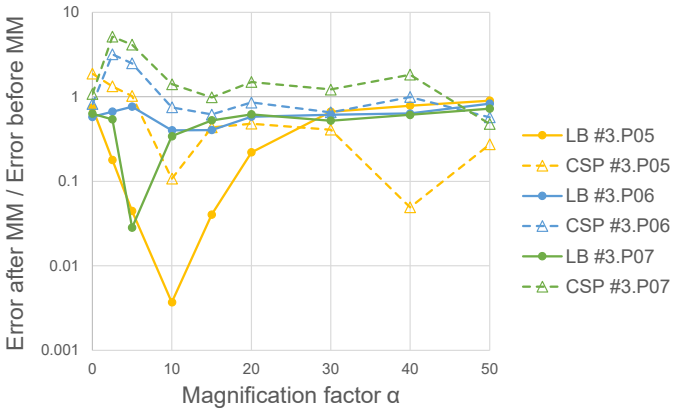


Fig. 18. Comparison between learning-based (LB) and complex steerable pyramid (CSP) algorithms in terms of accuracy gain with pattern matching.

these circumstances.

On the other hand, when the original amplitude is less than 0.02 px (i.e. #3.P05, #3.P06, and #3.P07), applying MM decreases the error with any of the considered magnification factors. In these cases, the uncertainty decreases until a point at which, probably due to image distortion, the error starts to increase. Under these circumstances, the best trade-off is magnification ranging from 5 to 10. The improvement is very significant with point #3.P05, at which the extremely small original displacement causes a relative error of 801% before applying MM. With $\alpha = 10$ magnification, the error decreases to 2.9%.

These results reinforce the findings from the previous experiments: MM is much more effective to reduce the uncertainty if the amplitude of the original vibration is small. In our experiments, the maximum amplitude for MM suitability with pattern matching is found to be between 0.02 px and 0.06 px, while with edge detection it is around 0.1 px to 0.12 px. A possible explanation to this difference is the different sensitivity of the measuring techniques to the artifacts and blur generated by MM. It must be considered, however, that these values may vary depending on the video and the region of interest to measure. As previously observed with the speckle pattern in Fig. 8 and the theory explained in [15], hard textures seem to be more prone to suffer from image distortion when applying MM, thus limiting its applicability.

Finally, to assess again whether the employed learning-based algorithm is more effective to reduce the measurement uncertainty than the CSP algorithm, the whole previously described process is repeated using the CSP magnification, but analyzing just point #3, which is the most suitable for MM, as previously observed. The results, together with the previous outcomes, can be observed in Fig. 18. All the minimum errors are lower when applying the learning-based algorithm instead of the CSP method. Moreover, for almost all the analyzed cases, results of LB are better than CSP ones. Therefore, the learning-based algorithm seems to be more suitable for our purposes than CSP.

IV. Conclusions and Future Lines of Work

In this study, we analyze the potential accuracy improvements when applying motion magnification before a vision-based measurement process. This way, the main contribution of this work is the evidence of measurement error reductions when applying MM before pattern matching and edge detection in some scenarios that are later analyzed. It has been firstly observed that magnifying the motion in a video can produce up to a 80% error reduction when measuring with pattern matching. When applying it before edge detection, the obtained displacement profile is much less noisy.

After this evidence of accuracy improvements, we study the conditions under which MM seems beneficial. We observe that, both when using pattern matching or edge detection, MM suitability increases with small displacements. Bellow a certain vibration amplitude, MM increases the accuracy with any amplification factor ranging from 0 to 50, while mixed results are observed with larger displacements, with some error reductions with specific magnification factors and error increments with other amplification values. It is important to mention that the learning-based algorithm by Oh *et al.* [10] seems to be the best method for these purposes, outperforming the CSP algorithm by Wadhwa *et al.* [15] in all the developed comparisons.

The maximum amplitude values found for MM suitability with the learning-based algorithm (0.02-0.06 px for pattern matching and 0.1-0.12 px for edge detection) may vary depending on some input image variables, such as the contrast and image noise. In fact, we found that hard textures seem to be more prone to suffer from blur and artifacts after applying MM. Therefore, a very interesting line of work would be to study further how these image variables affect the output image distortion, analyzing then the effect of this distortion on the final measurement error. With this information, it should be possible to predict when it is beneficial to apply MM before pattern matching, edge detection, or other vision-based measurements under any conditions.

References

[1] S. Shirmohammadi and A. Ferrero. “Camera as the instrument: The rising trend of vision based measurement,” *IEEE Instrumentation & Measurement Magazine*, vol. 17, no. 3, pp. 41-47, Jun. 2014.

[2] D. Feng and M. Q. Feng, “Experimental validation of cost-effective vision-based structural health monitoring,” *Mechanical Systems and Signal Processing*, vol. 88, pp. 199-211, May 2017.

[3] X. W. Ye, C. Z. Dong, and T. Liu, “A review of machine vision-based structural health monitoring: methodologies and applications,” *Journal of Sensor*s, 2016, Art. no. 7103039.

[4] A. Cigada, P. Mazzoleni, and E. Zappa, “Vibration monitoring of multiple bridge points by means of a unique vision-based measuring system,” *Experimental Mechanics*, vol. 54, no. 2, pp. 255-271, Feb. 2014.

[5] A. Cigada, P. Mazzoleni, E. Zappa, and M. Franzi, “Cameras as displacement sensors to get the dynamic motion of a bridge: Performance evaluation against traditional approaches,” in *6th International Conference on Bridge Maintenance, Safety and Management (IABMAS 2012)*, pp. 2835-2841, 2012.

[6] X. W. Ye, Y. Q. Ni, T. T. Wai, K. Y. Wong, X. M. Zhang, and F. Xu, “A vision-based system for dynamic displacement measurement of long-span bridges: algorithm and verification,” *Smart Structures and Systems*, no. 12, vol. 3, pp. 363-379, Sep. 2013.

[7] X. W.Ye, T. H. Yi, C. Z Dong, T. Liu, and H. Bai, “Multi-point displacement monitoring of bridges using a vision-based approach,” *Wind & structures*, vol. 20, no. 2, pp. 315-326, Feb. 2015.

[8] H. W. Schreier and M. A. Sutton, “Systematic errors in digital image correlation due to undermatched subset shape functions,” *Experimental Mechanics*, vol. 42, no. 3, 303-310, Sep. 2002.

[9] K. S. Son, H. S. Jeon, G. S. Chae, J. S Park, and S. O Kim, “A fast high-resolution vibration measurement method based on vision technology for structures,” *Nuclear Engineering and Technology*, vol. 53, no. 1, pp. 294-303, Jan. 2021.

[10] T. H Oh, R. Jaroensri, C. Kim, M. Elgharib, F. E. Durand, W. T. Freeman, and W. Matusik, “Learning-based video motion magnification,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 633-648, 2018.

[11] Y. Zhang, S. L. Pintea, and J. C. Van Gemert, “Video acceleration magnification,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp. 529-537, 2017.

[12] A. Cat Le Ngo and R. C.-W. Phan, “Seeing the Invisible: Survey of Video Motion Magnification and Small Motion Analysis,” *ACM Comput. Surv.*, vol. 52, no. 6, Nov. 2020, Art. no. 114.

[13] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson, “Motion magnification,” *ACM Trans. Graph.*, vol. 24, no. 3, pp. 519–526, Jul. 2005.

[14] *CSAIL (MIT), Video Magnification*. Accessed: Dec. 30, 2021. [Online]. Available: http://people.csail.mit.edu/mrub/vidmag/

[15] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, “Phase-based video motion processing,” *ACM Trans. Graph.*, vol. 32, no. 4, Jul. 2013, Art. no. 80.

[16] M. P. Popek, M. E. Danielewska, and D. R. Iskander, “Assessing the feasibility of the use of video motion magnification for measuring microdisplacements,” *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2329–2336, Sep. 2017.

[17] Y. Q. Wang, M. A. Sutton, H. A. Bruck, and H. W. Schreier, “Quantitative error assessment in pattern matching: effects of intensity pattern noise, interpolation, strain and image contrast on motion measurements,” *Strain*, vol. 45, no. 2, pp. 160-178, 2009.

[18] A. J. Molina-Viedma, L. Felipe-Sesé, E. López-Alba, and F. Díaz, “High frequency mode shapes characterisation using Digital Image Correlation and phase-based motion magnification,” *Mechanical Systems and Signal Processing*, vol. 102, pp. 245-261, Mar. 2018.

[19] P. Poozesh, A. Sarrafi, Z. Mao P. Avitabile, and C. Niezrecki, “Feasibility of extracting operating shapes using phase-based motion magnification technique and stereo-photogrammetry”, *Journal of Sound and Vibration*, vol. 407, pp. 350-366, Oct. 2017.

[20] V. Fioriti, I. Roselli, A Tatì, R. Romano, and G. De Canio, “Motion Magnification Analysis for structural monitoring of ancient constructions,” *Measurement*, vol. 129, pp. 375-380, Dec. 2018.

[21] J. G. Chen, N. Wadhwa, Y.-J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk, “Modal identification of simple structures with highspeed video using motion magnification,” *J. Sound Vib.*, vol. 345, pp. 58–71, Jun. 2015.

[22] A. Sarrafi, Z. Mao, C. Niezrecki, and P. Poozesh, “Vibration-based damage detection in wind turbine blades using phase-based motion estimation and motion magnification,” *J. Sound Vib.*, vol. 421, pp. 300–318, May 2018.

[23] Z. Shang and Z. Shen, “Multi-point Vibration Measurement for Mode Identification of Bridge Structures using Video-based Motion Magnification,” *arXiv preprint arXiv:1712.06566*, 2017.

[24] A.J. Choi, J.-H. Han, “Frequency-based damage detection in cantilever beam using vision-based monitoring system with motion magnification technique,” *Journal of Intelligent Material Systems and Structures*, vol. 29, no. 20, pp. 3923–3936, Sep. 2018.

[25] C. Peng, C. Zeng, and Y. Wang, “Phase-Based Noncontact Vibration Measurement of High-Speed Magnetically Suspended Rotor,” *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4807–4817, Jul. 2020.

[26] J. Zhou, H. Li, L. Zhang, X. Wang, and Y. Li, “Vibration Measurement With Video Processing Based on Alternating Optimization of Frequency and Phase Shifts”. *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1-13, 2021.

[27] N. Wadhwa, “Revealing and Analyzing Imperceptible Deviations in Images and Videos,” Ph.D. thesis, Department of Mathematics, Massachusetts Institute of Technology, USA, 2016.

[28] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, “Eulerian video magnification for revealing subtle changes in the world,” *ACM Trans. Graph.*, vol. 31, no. 4, pp. 1–8, Jul. 2012.

[29] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, “Shiftable multiscale transforms,” *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 587–607, Mar. 1992.

[30] N. Wadhwa, M. Rubinstein, F. Durand, and W. Freeman, “Riesz pyramids for fast phase-based video magnification,” in *Proc. IEEE Int. Conf. Comput. Photogr.*, vol. 2014, pp. 1–10.

[31] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1-6.

[32] C. G. Relf, “Image acquisition and processing with LabVIEW,” CRC press, 2003.

[33] *Huawei p30 lite specifications.* Accessed: Dec. 30, 2021. [Online]. Available: https://consumer.huawei.com/levant/phones/p30-lite/specs/

[34] R. C. Gonzalez, and R. E. Woods, “Digital Image Processing,” forth ed., Prentice Hall, Englewood, Cliffs, NJ, 2018.

[35] *NI Vision 2019 for LabVIEW Help, Pattern Matching Techniques*. Accessed: Dec. 30, 2021. [Online]. Available: https://zone.ni.com/reference/en-XX/help/370281AG-01/nivisionconcepts/pattern_matching_techniques/

[36] M. P. Popek and D. R. Iskander, “A New Approach to the Phase-Based Video Motion Magnification for Measuring Microdisplacements,” *IEEE Trans. Instrum. Meas.*, vol. 69, no. 2, pp. 354–361, Feb. 2020.

[37] Y. Q. Wang, M. A. Sutton, H. A. Bruck, and H. W Schreier, “Quantitative error assessment in pattern matching: effects of intensity pattern noise, interpolation, strain and image contrast on motion measurements,” *Strain*, vol. 45, no. 2, pp. 160-178, Mar. 2009.

[38] H.G. Breunig, A. Uchugonova, A. Batista, and K. Konig, “Software-aided automatic laser optoporation and transfection of cells,” *Scientific Reports*, vol. 5, pp. 1–11, Jun. 2015.

[39] T. G. Chondros, and A. D. Dimarogonas, “Vibration of a cracked cantilever beam,” *Transactions of the ASME*, vol. 120, pp. 742-746, Jul. 1998.

[40] S. Megahed, and A. Nada, “Dynamic modeling of a flexible cantilever beam: an experimental technique,” *Proceedings of the ASME Design Engineering Technical Conference*, pp.449–456, Sep. 2003.

Eduardo Pérez received in 2021 both the M.Sc. degree in mechanical engineering from Politecnico di Milano, (Milan, Italy) and the M.Sc. degree in industrial engineering at Politécnica de Madrid (Madrid, Spain) through a double degree program. His current research interests include vision-based measurements, image processing, and computer vision.

Emanuele Zappa (M’10, SM’21) is full professor in Mechanical and Thermal Measurements at Politecnico di Milano, Italy. He is Associate Editor in Chief of the IEEE Transaction on Instrumentation and Measurement. The scientific activity is mainly in the field of mechanical and image-based measurements.