Vibration frequency measurement using a local multithreshold technique

Belen Ferrer,[1] Julian Espinosa,[2] Ana B. Roig,[3] J. Perez,[2] and D. Mas[2,*]

[1]*Department of Civil Engineering, Univ. Alicante P.O. Box, 99, 03080 Alicante Spain*
[2]*Inst. of Physics Applied to the Sciences and Technologies, Univ. Alicante P.O. Box, 99, 03080 Alicante Spain*
[3]*Department of Optics, Univ. Alicante P.O. Box, 99, 03080 Alicante Spain*
[*]*david.mas@ua.es*

Abstract: In this paper, we demonstrate the use of a video camera for measuring the frequency of small-amplitude vibration movements. The method is based on image acquisition and multilevel thresholding and it only requires a video camera with high enough acquisition rate, not being necessary the use of targets or auxiliary laser beams. Our proposal is accurate and robust. We demonstrate the technique with a pocket camera recording low-resolution videos with AVI-JPEG compression and measuring different objects that vibrate in parallel or perpendicular direction to the optical sensor. Despite the low resolution and the noise, we are able to measure the main vibration modes of a tuning fork, a loudspeaker and a bridge. Results are successfully compared with design parameters and measurements with alternative devices.

OCIS codes: (100.2000) Digital image processing; (120.7280) Vibration analysis; (040.7290) Video.

References and links

1. J. J. Lee, Y. Fukuda, M. Shinozuka, S. Cho, and C. Yun, "Development and application of a vision-based displacement measurement system for structural health monitoring of civil structures," Smart Struct. Syst. **3**(3), 373–384 (2007).
2. H. N. Nassif, M. Gindy, and J. Davis, "Comparison of laser Doppler vibrometer with contact sensors for monitoring bridge deflection and vibration," NDT Int. **38**(3), 213–218 (2005).
3. E. Valero, V. Micó, Z. Zalevsky, and J. García, "Depth sensing using coherence mapping," Opt. Commun. **283**(16), 3122–3128 (2010).
4. F. Hild and S. Roux, "Digital Image Correlation: from displacement Measurement to identification of elastic properties – a review," Strain **42**(2), 69–80 (2006).
5. J. C. Trinder, J. Jansa, and Y. Huang, "An assessment of the precision and accuracy of methods of digital target localization," ISPRS J. Photogramm. **50**(2), 12–20 (1995).
6. D. Mas, J. Espinosa, A. B. Roig, B. Ferrer, J. Perez, and C. Illueca, "Measurement of wide frequency range structural microvibrations with a pocket digital camera and sub-pixel techniques," Appl. Opt. **51**(14), 2664–2671 (2012).
7. D. Teyssieux, S. Euphrasie, and B. Cretin, "MEMS in-plane motion/vibration measurement system based CCD camera," Measurement **44**(10), 2205–2216 (2011).
8. D. Mas, B. Ferrer, J. T. Sheridan, and J. Espinosa, "Resolution limits to object tracking with subpixel accuracy," Opt. Lett. **37**(23), 4877–4879 (2012).
9. J. L. Barron, D. J. Fleet, and S. Beauchemin, "Performance of optical flow techniques," Int. J. Comput. Vis. **12**(1), 43–77 (1994).
10. D. Mas, B. Domenech, J. Espinosa, J. Pérez, C. Hernández, and C. Illueca, "Noninvasive measurement of eye retraction during blinking," Opt. Lett. **35**(11), 1884–1886 (2010).
11. B. Ferrer, J. Espinosa, J. Perez, S. Ivorra, and D. Mas, "Optical scanning for structural vibration measurement," Res. Nondestruct. Eval. **22**(2), 61–75 (2011).
12. J. Espinosa, B. Ferrer, D. Mas, J. Perez and A. B. Roig, "Método y sistema para medir vibraciones," Patent pending nº P201300498 (05–23–2013).
13. Casio Europe at http://www.casio-europe.com/euro/exilim/exilimzrserie/exzr1000/ (visited on 09/03/2013).
14. AOS Technologies at, *http://www.aostechnologies.com/high-speed-imaging/products-high-speed/* (visited on 10/09/2013).

1. Introduction

Vibration measurement and analysis is an important topic in many scientific fields, like structural engineering, acoustics, biotechnology, entertainment devices, security and surveillance. Although there exist different methods for measuring vibrations, accelerometers are usually preferred. These devices register the local acceleration of the specific point of the specimen where they are attached. Since they are contact devices, they are difficult to use when the specimen is not accessible due to the object itself or the surrounding conditions (inaccessibility, damage risk, etc.) [1].

Alternatively to the accelerometers Doppler, vibrometers are used as non-contact devices. A test laser beam is sent against a moving surface. The device collects the light scattered from it and, then, it is interfered with a reference beam. Although they provide very accurate result, their measuring distance may remain short due to attenuation. Additionally, they result really expensive and no-cost effective for many applications [2].

Recently, Valero et al. presented a method for motion detection with special application in sound registration [3]. The method is based on the detection and processing of a defocused coherent speckle pattern. Results are accurate but, under our point of view and despite the usefulness and applicability of the method, it presents several drawbacks: reflectivity of the scattering surface has to fulfill some requirements of roughness and there is a need of a laser probe that will attenuate its intensity with the distance.

In the last years, image based methods are becoming a reliable alternative to non-contact measurement of movement and vibrations. Most common methods are based in object recognition and tracking through digital image correlation [4]. Although they are easy to implement, they require of high-resolution cameras, which may not be always available. Other methods allow measurements with an improvement of the physical resolution of the acquiring device and are known as sub-pixel method [5]. Sub-pixel techniques have been shown to produce accurate results even with low cost devices, being a real alternative to traditional methods, since they may increase the theoretical resolution in more than 50 times [6].

One of the simplest approaches to the sub-pixel techniques consists of analyzing the response of a single detector (pixel) when detecting an edge. Light diffused by an object surface usually presents smooth transitions, but proper illumination conditions can induce abrupt brightness changes between different areas defined by a different slope, texture, color, height or any other property. Discrete size of the sensors in the CCD or CMOS arrays makes that, eventually, light coming from both parts of an object border (in and out) falls onto the same sensor thus giving a response that depends on the value of each region and the area covered by them. As we will show below, these pixels are very sensitive to changes in the scene and proper analysis of this "border region" may provide useful information about object movements. Based on this principle, the authors in [7] propose to use the energy gradient produced by optical diffraction in external border regions to accurate measure the in-plane vibrations of micro devices. Achieved resolutions are of the order of 0.001 px, although the method seems to be limited to micro-objects.

In [8], the authors present a simple that only considers binary levels and counting dots. Their method was based on the probability of detecting a single pixel change in the whole scene and show that the motion detection accuracy could be increased in several orders of magnitude, up to 10^{-6} px. The proposal is proven through a theoretical experiment by using an object consisting on a sparse random cloud of dots. The problem is that those analyzed objects are difficult to implement, so optimum resolutions are hard to achieve in practical applications. Despite this improvement, this method still needs a target to be attached to the specimen under analysis, so it is not a general valid solution for inaccessible objects.

In this paper, we propose a sub-pixel technique for measuring frequency vibrations of standard objects without using external elements like targets, physical probes or optical beams. The method is based on searching differences between successive frames and quantifying them. Although the philosophy is well known, convenient arrangement of the

information provides an efficient solution for frequency measurement. Moreover, it is implemented with a simple setup composed by a pocket camera, a tripod and a standard computer running with Matlab.

From a captured frame, our method consists of taking a small region of interest where light intensity variation due to object vibration is expected to happen. Instead of considering small variations of intensity, the analysis is done at different thresholded levels so only binary information is considered. Variation in the number of white pixels inside the region is tracked in order to obtain the frequency of the vibration movement.

In some sense, the method is related to determination of optical flow like in [9], where the time-evolution of the pixels in a video sequence is analyzed. Nevertheless, our approach is simpler and does not consider voxels but image blocks. Additionally, no estimation of velocity or movement direction is done thus resulting in a faster and more efficient algorithm. Although this may seem a small advantage, the reader should notice that high-speed sequences of relatively long processes (several seconds) produce a huge amount of images that have to be stored and analyzed. With our proposal, analysis of here presented sequences (~1000 frames) is done in less than one minute, which is very convenient for many applications.

The manuscript is developed as follows: in the next section, we describe the main methods, materials and experiments that have been used. In Section 3, we present the results of our experiments and compare them with design parameters and data measured with standard instruments. Discussion about results is also included in that section. Finally, we summarize the main conclusions in the last section.

2. Material and methods

2.1 General principles of sub-pixel tracking

Objects in nature present a wide variety of illumination levels. Binary thresholding on a digital image provides the number of pixels whose luminance value is equal or lower than a selected value. Result of this operation is a binary image with white blobs on a black background. The appearance of this image is directly connected with the object geometry and the illumination structure but, in the discretization and binarization processes, continuous shadows and profiles are broken and obtained borders show a pseudo-arbitrary profile.

Any movement of the object, even the smallest one, will change the scene brightness distribution and, consequently, the borders in the binarized digital image provided that the resolution of the pixel array is high enough. Detection of the changed pixels provides information of movement at a sub-pixel scale [8].

Let us consider a binary object like the one in Fig. 1. When the object is captured by a discrete camera sensor array, empty sensors are not giving response (pixel value = 0) while object areas that fall completely inside a pixel region give full response (pixel value = 1) and are detected. Notice that object borders represented by gray areas in Fig. 1(a) are not grid-shaped and only occupy a part of the sensor area. The response of these sensors is set to 0 or 1 depending whether the majority of the area is empty or occupied.

In Fig. 1(b), we depict the situation after discretization. Object profile has been clearly degraded and, although the contour resembles to the original object, the exact form is lost. When the object slightly moves, the occupied area ratio in some border pixels may change and so will do the contour. In Fig. 1(c), we depict the situation of the object in Fig. 1(a) after a lateral shift of 0.25 pixels. Notice that even with such small displacement, changes in the object shape between Figs. 1(b) and 1(c) are evident.
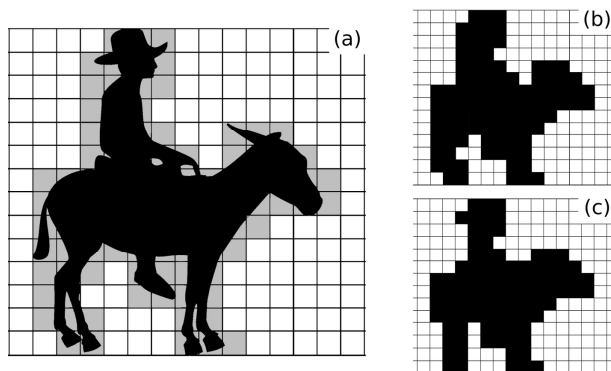
Fig. 1. (a) Binary object scene. Border pixels are marked in gray; (b) scene in (a) after spatial quantization; (c) scene in (a) after a horizontal shift of 0.25 px.

Detection of these changes may be useful to detect movements in the object. Unfortunately, real objects are not as clean as the one here depicted. Image binarization may produce a lot of noise because of sparkles, surface texture or non-uniform background. Moreover, changes in the binary images can be due to many factors, like shadows or illumination changes, and not only to object movements.

However, if the object is not randomly moving but vibrating, it exists a periodic pattern that can be detected and extracted from the noise and thus, the frequency of the movement can be obtained.

2.2 Multilevel thresholding

Up to this moment, we have only referred to binarized objects, while scenes are often captured in gray-scale or color. The simplest solution for this problem is using binary scenes. This can be accomplished by attaching a binary target to the object or using saturated objects in the scene as targets. These techniques were successfully used in [10,11]. In those cases, the image borders were clearly located so information about the object trajectory was also obtained.

In general, these solutions are not always possible. Objects are usually more complex and present a relatively wide dynamic range. Reduction of dynamic range in images is always problematic since usually appear false contours than can distort the original target shape and mask the movement. Thus, some criteria must be followed in order to obtain a suitable binary image. By simplicity, we will consider gray scale images, although the concepts can be easily extended to color scenes.

The first approach to gray scale images consist of reducing them to binary images by selecting a proper thresholding level according to the histogram, statistical values (maximum, minimum, median…) or any other criterion. Nevertheless, we would like to underline that we are looking for sub-pixel movements, so changes may be only perceptible in small bright sparkles, middle grays or dark areas. Therefore, it is difficult to guess which gray levels will be affected.

Instead of making models and predictions about illumination changes due to movement, we propose here to explore pixel changes at different levels simultaneously. Therefore, a multilevel threshold would be a convenient approach.

If the object vibrates, it is very likely that several brightness levels are affected by this vibration and thus, redundant information can be obtained from them. The remaining levels will carry useless information so that the final effect is a noise background that degrades the power spectrum. An adequate composition of the power spectra from all levels may enhance the frequency peaks corresponding to the vibration and clearly detect the vibration frequency.

Therefore, our proposal is to threshold each frame in different levels. The variation in the number of active pixels with respect to the first frame is analyzed for all the levels and the

obtained information is then analyzed and composed in order to determine the main vibration modes and cancel out the noise [12].

Obviously, the maximum information is obtained when all levels in the image are used. Nevertheless, this leads to inefficient algorithms both in time and memory resources. Proper selection of the levels to be analyzed will accelerate the algorithms and also increase the signal to noise ratio. The easiest option is to take a reduced number of levels between the minimum and the maximum. In all cases here analyzed, we selected 8 levels equally spaced. Some trials were made with more levels, but no relevant information was added. We also tried with 4 levels, but in some experiments the vibration was lost, so finally 8 levels were selected.

Selection of the number of levels is arbitrary and depends very much on the scene illumination and object reflectivity. One can take advantage from a special illumination or from some details that are more likely to change with the movement and, therefore, make a specific multilevel threshold for that scene: in controlled environments one can always make some areas brighter or force some sparkling points on the moving surface in order to facilitate the analysis. Unfortunately, this is not always possible and thus the multilevel thresholding has to be adapted to the particular experimental setup.

Once the thresholding is accomplished, the process consists of counting the white pixels in each binarized level and compare this number with the corresponding one of the same level in the first frame. This way we obtain the pixel variation through the sequence. At the end of the process, the different signals obtained for each level are composed in order to obtain a single result from the measurement. This composition makes the technique more robust since enhances very much the common peaks while cancels out random noise. Therefore, the method provides accurate results even in presence of strong noise.

2.3 Selection of a region of interest (ROI)

The proposed method only pretend to detect the vibration frequency, but not tracking the object. Since it is based on the summation of all the pixel values inside a binarized area specific location of each pixel position os lost. Additionally, only few pixel changes are due to the movement and noise or undesired changes in the scene may mask the vibration signal.

For this reason, we recommend taking a ROI as small as possible around a point where the movement has more probabilities to be detected which implies making some general hypothesis about the object dynamics. This way, we exclude all parts in the scene that do not add relevant information and the ratio between pixel changes due to movement and background noise is also increased. Notice that in the experiments shown below, all ROIs are smaller than 40x40 pixels.

Since a multilevel threshold will be performed on a relatively small area, one must be sure that in the area selected the histogram is wide enough to guarantee the presence of several gray scale levels carrying different information can be analyzed. As it happened above, proper illumination will contribute very much to the success in the measurement.

2.4 Materials and experimental procedure

The method just proposed has been proven by two lab experiments consisting of the detection of vibrations parallel and perpendicular to the image plane, respectively. The first case was accomplished with a tuning fork and, for the second, we used a loudspeaker membrane. In Figs. 2(a) and 2(b), we show picture of the specimens measured. Images were taken in color but only the green channel was considered. As we explained before, only 8 binarized levels are considered here.

In order to show the performance of the method in a more general case, we took the setup outside the lab and measured the main vibration mode of a small structure consisting of a 6.6 m long bridge passage connecting two parts of a building, as can be seen from Fig. 2(c).

In all cases, we used a CASIO Exilim EX-ZR1000 pocket camera [13], which was able to record video at different high-speed rates from 120 to 1000 frames per second (fps), with varying resolutions from 640×480 for the lowest temporal rate to $224 \times 64px$ for the highest. This camera stores the video sequence in JPEG-AVI format, so it introduces some noise in the

sequence that will appears as random dots after binarization and thus will not affect to the vibration measurement. In order to avoid vibrations coming from pressing and releasing the camera shutter, the first three seconds of the sequence have been always rejected.

The frequency response of the camera was checked with a stroboscopic light. We observed that there are some inaccuracies in the acquisition speed, so the peaks present a wide basis. Anyway, the error in the peak location was below 1% so we considered this camera accurate enough for our proposal.

For the lab experiments, the camera was set 75 cm far from the objects to be measured. Illumination in both cases was accomplished with a 50 W LV halogen lamp with dichroic reflector. The lamp was connected to a stabilized DC power supply in order to avoid the detection of the AC cycle (100 Hz). The illumination was oblique from an angle of 30-45 degrees with respect to the object-camera axis. The camera was set to 1000 fps for both experiments.
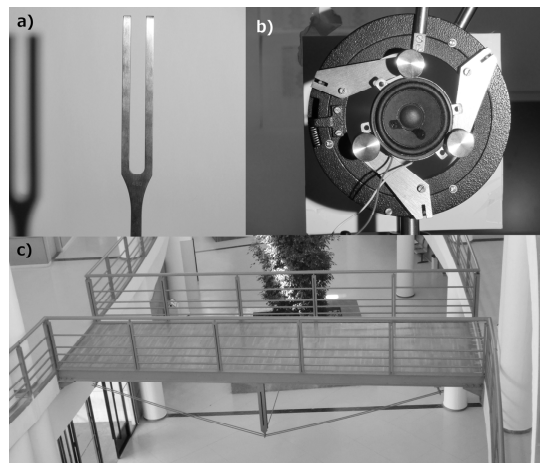
Fig. 2. Pictures of the specimens under analysis: (a) tuning fork; (b) loudspeaker membrane; (c) bridge passage.

For the bridge measurement, the camera was located at the same level than the bridge, at the right lateral corridor around 15 m from its center, outside the photo limits in Fig. 2(c). The oscillation was recorded by using the 120 fps mode [13]. In this case, the ambient illumination was enough for the experiment so no additional light sources were used.

3. Results and discussion

3.1 Measurement of vibrations in a controlled environment

As we pointed before, for the first experiment, the vibration of a tuning fork depicted in Fig. 2(a) was registered at 1000 fps. We hit the fork with a rubber hammer and then captured the sequence. The first 3 seconds of the video were ignored to avoid interferences from the vibrations coming from the camera operation, and the following 1000 frames were processed. The captured video is shown in Media 1. In order to reduce the noise and accelerate the calculation, the analysis was performed on a small region of interest (ROI), as can be appreciated in Fig. 3.
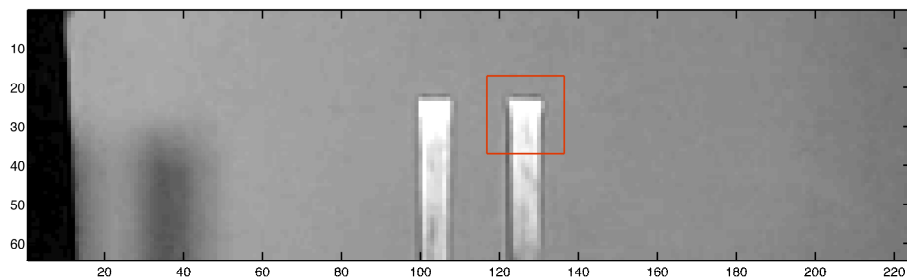
Fig. 3. Arms of a vibrating tuning fork and ROI analyzed. Notice that ROI size is 18 × 17 px. (Media 1)

The image inside the ROI is thresholded by eight different equidistant levels from the minimum to the maximum brightness levels and so obtained binary sequences are independently analyzed. In Fig. 4, we show the 8 binary levels for the first frame of the analyzed sequence.
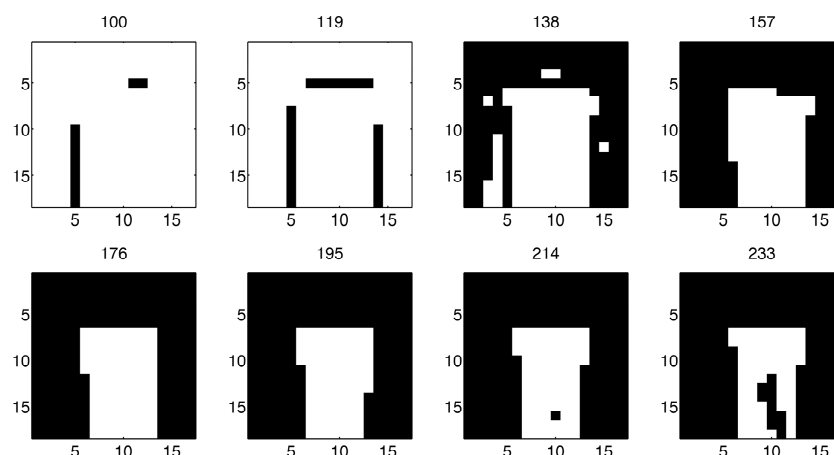

Fig. 4. Binary levels thresholded for the ROI in Fig. 2. Above each image, we write the gray threshold level.

The analysis consists of taking each of the binary sequences, counting the variation of active of pixels with respect to the first frame and performing a Fourier transform. In Fig. 5(a), we show the percentage of pixel variations relative to the ROI size (18x18 px) for each thresholded level. Notice that in all levels except level 2, the variations are less than 5% i.e. 16 px. In Fig. 5(b), we show the Fourier transform of each of the eight signals. In majority of levels, a frequency of 483 Hz is clearly detected. The tuning fork has a design frequency of 480 Hz, but measurements with a microphone showed that its real frequency peak is 481,9. Considering that the accuracy of our measurement is of ± 1 Hz and inaccuracies in the acquisition rate from the low-end camera we can conclude that our measurement provides a correct result, with are error below 0.5%.

Results obtained from the Fourier transforms can be combined in order to enhance the peaks and reduce or even cancel the noise. In what follows, we will just consider the sum of the peaks, although more sophisticated combinations can be proposed and developed.

In the second experiment, the vibration of a loudspeaker membrane is determined. The loudspeaker in Fig. 6 is connected to a computer and a two-frequency sound (317 and 412 Hz) is generated. The vibration of the loudspeaker membrane is captured with a microphone and the camera (see Media 2). Each result from the microphone and the camera represented in Fig. 7 have been normalized to its respective peak value at 317 Hz to allow better comparison between power spectra.
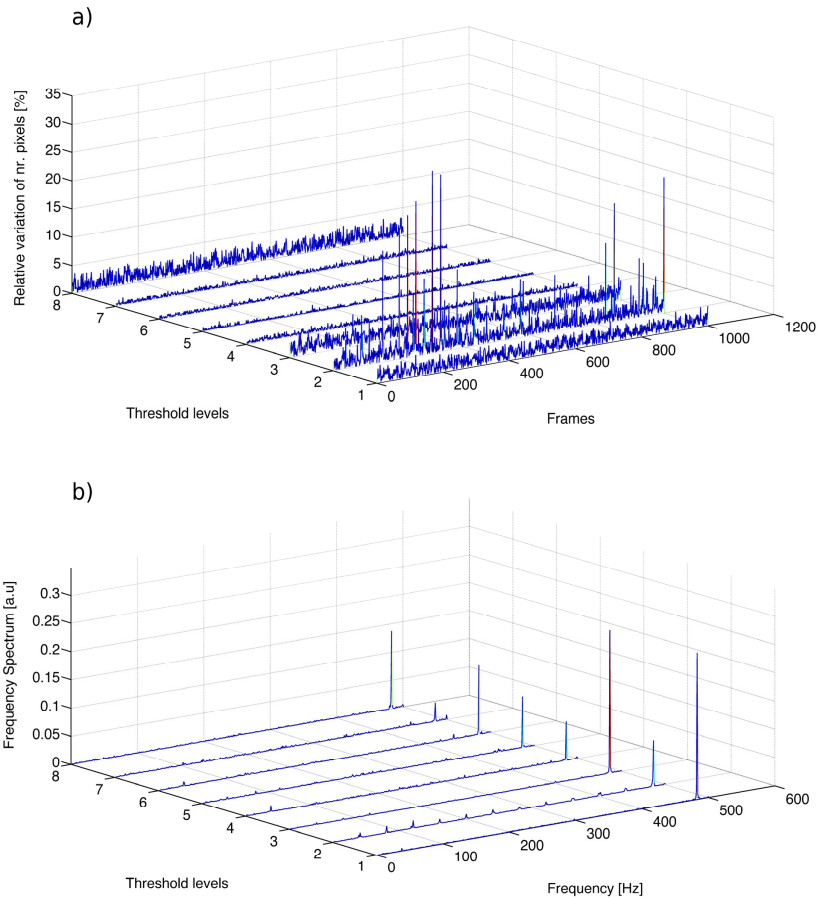
a)

b)

Fig. 5. a) Relative variation with respect to the first frame (absolute value) of the pixel number. b) Fourier transform of the signals in a).
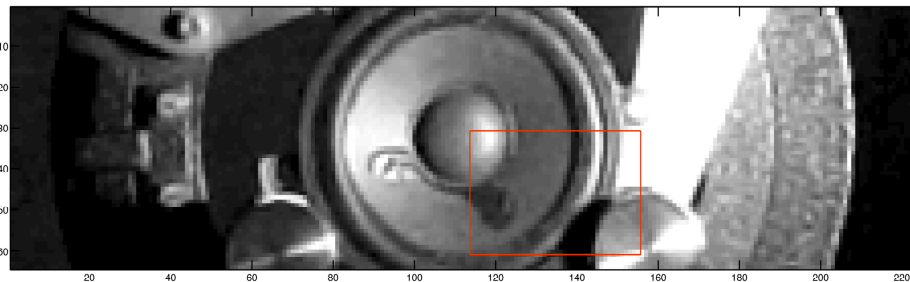

Fig. 6. Membrane of a vibrating loudspeaker and ROI analyzed. Notice that ROI size is 17 × 36 px (Media 2)

Notice that both the camera and the microphone detect the two main frequencies in the loudspeaker and even a third one that may be due to membrane distortions. Nevertheless, the relative height between peaks registered from the microphone is different than that from the camera. The video technique only counts pixels so peak height is not directly connected to any physical magnitude. Therefore, sensitivity to frequencies may be different to those obtained by other methods.
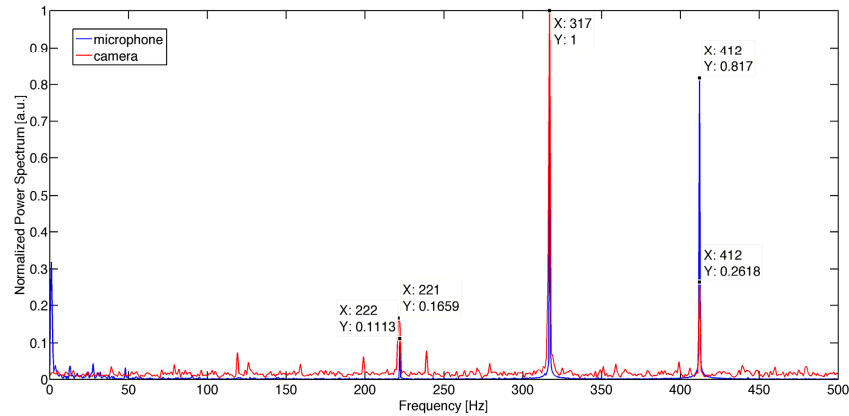
Fig. 7. Normalized frequency spectra obtained by analyzing the number of pixels variation (camera) in red and the sound intensity (microphone) in blue measured from the loudspeaker membrane.

As we said in the previous section, we must remind here that successful results depend very much of the ROI selection. In the case of the tuning fork, the selection is obvious since it is expected that at the fork end the movement amplitude to be maximum. Notice also that the selected ROI also offers high contrast and the eight levels carry different information, thus increasing the possibilities of a correct vibration measurement. In the case of the loudspeaker, the contrast through the diaphragm is relatively low. Only in the lower part, where we can compare the membrane movement with the static holder, which appears brighter, the correct frequencies were found.

Images with low contrast usually present few differences between levels, and thus, the redundancy is lost. In those cases, the presence of the image noise will probably mask the vibration frequency. In the second experiment, this happens for the highest frequency that is lost except for the ROIs selected in the half-lower part of the scene.

3.1 Measurement of vibrations in a real structure

After these lab experiments, the method was also tested on a passage bridge inside one building of the University of Alicante, as can be seen in Figs. 2(c) and 8(a). The bridge was excited by a person jump in the center and its main vibration mode was measured with the camera (see Media 3) and a monoaxial accelerometer located at the center of the bridge. We must point that the signal from the accelerometer was approximately selected two and a half seconds after the impact in order to avoid transient effects and to clearly distinguish the main frequency.

A small ROI in the center of the bridge is selected as can be seen in Fig. 8(a). Since the number of pixels that may change their status is very low, a very small window of 13x15 pixels must be selected in order to improve the signal-to-noise ratio. In Fig. 8(b), we also show the obtained accelerometer signal, together with the temporal window analyzed.

In Fig. 9, we present the frequency measured with the accelerometer and the camera. Results show that frequency registered by both devices is nearly the same. There are some discrepancies regarding the peak basis, where the accelerometer detects some secondary peaks that are lost in the video analysis. However, the main mode is clearly detected.
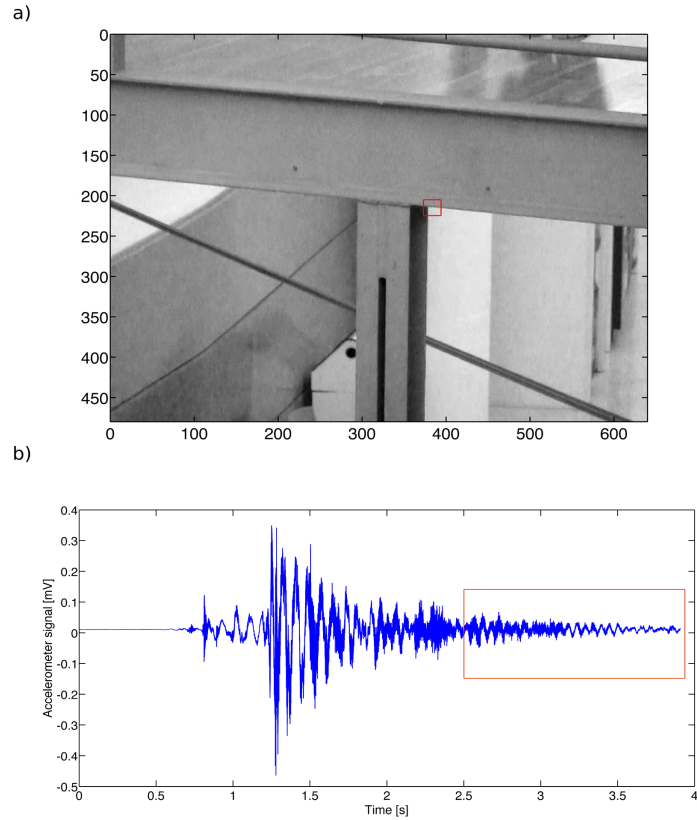
a)

b)

Fig. 8. a) Frame view of the bridge and ROI selected 13 × 15 px. (Media 3); b) Accelerometer signal and time window analyzed.

We would like to emphasize that the video sequence was taken just using the camera and the tripod, with no wires or additional lamps. On the contrary, the setup for the accelerometers required connection to a power supply, an acquisition module connected to a laptop computer and a wired accelerometer. Although the setup for this experiment is not very complex, it results, indeed, much more ostentatious than our proposed setup.
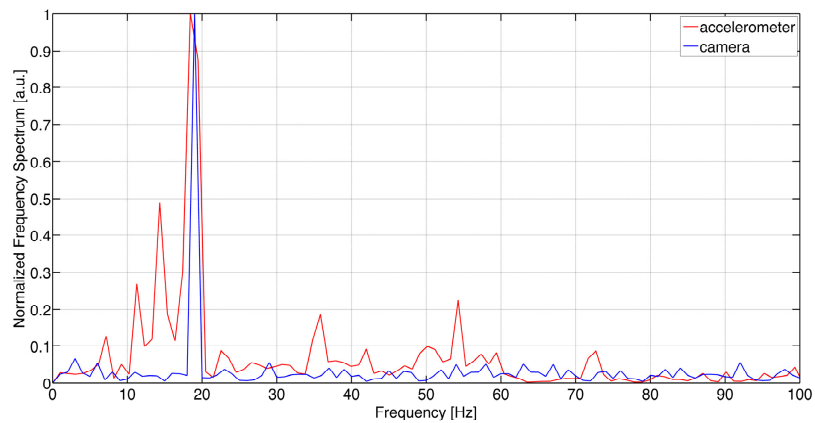

Fig. 9. Normalized frequency spectrum obtained from the camera and the accelerometer.

4. Conclusions

In all exposed experiments, we show that it is possible to measure the vibrations of an object provided that the temporal video resolution is high enough. Notice that the spatial resolution does not play an important role, since the managed ROIs are really small. Thus, the method can be easily implemented with a low cost camera, as we have also demonstrated here.

The method provides good results despite being limited to only measure the main frequency peaks. Although other devices can provide more complete results, it is also true that the setup and calculation complexity makes our method really cost-effective even for preliminary frequency tests.

Our proposal can be very much improved by using a better camera. The one we used here has limited capabilities since it is low-end device. Majority of laboratory cameras record video sequences without compression, and thus, images are less noisy. Additionally, they allow increasing the acquisition speed at the cost of reducing the resolution in real-time [14]. This way, one can adapt the Nyquist limit of the camera to the maximum expected frequency without changing the algorithm and avoid aliasing effects. In the camera here used both options (compression algorithms and acquisition speed) were limited by the hardware. In any case, results were satisfactory in all experiments.

It is also possible further explore and optimize the multi-level threshold and the combination between the different levels. An algorithm for automatically determining the size and location of the region of interest will also improve the method very much. Additionally, future developments could include statistical analysis of the image in order to obtain information about displacement of the pixels in the ROI, and, from that, estimate the vibration amplitude of the specimen under analysis.

Regarding the applications, we have shown in the lab the two most typical cases of vibrations (bar and membrane vibration). We have also shown that the method is easily scalable and can be used to measure the main vibration mode of a real structure without using any contact device and from several meters of distance.

In the presented case, we used a pocket camera but professional cameras with large zoom lenses can extend the range of application up to several hundreds of meters. Our proposal does not require of wires, laser probes or object attachments in the specimen to be measured. It also does not require special illumination conditions, provided that image contrast is high enough. Because of this and according to our results, we believe that the method is versatile and can be a reliable alternative to traditional vibrometry methods.