

HISTORIC MASONRY MONITORING BY MOTION MAGNIFICATION ANALYSIS

VINCENZO FIORITI, IVAN ROSELLI, ANGELO TATÌ & GERARDO DE CANIO

ENEA Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Italy

ABSTRACT

Vibration monitoring of historic structures in urban environment is a relevant issue for health survey and early damaging detection in sustainable and enhanced resilient cities. This study explores the potentialities of vibration monitoring by Motion Magnification Analysis. Motion magnification acts like a microscope for motion in video sequences, but affecting only some groups of pixels. The motion magnification uses the spatial resolution of the video-camera to extract physical properties from images to make inferences about the dynamical behaviour of the object, e.g. to visualize at least the first mode shape, no matter its dimensions, since any point on the surface of the object can be considered a virtual sensor. Recently, a number of experiments conducted on simple geometries like rods and other small objects, as well as on bridges, showed the reliability of this methodology compared to accelerometers and lasers. Researchers have been also very interested in assessing the method's feasibility, since conventional devices are surely more precise, but more expensive and much less practical. In this paper, we give an introduction to MMA and describe its application to the analysis of two full-scale historic masonry walls tested on shaking tables. This is an interesting point, because the size of tested walls is larger than usual small experimental set-ups implemented in MMA testbed until now. Results showed that MMA allowed a visual identification of fractures in advance. Moreover, we performed some conventional calculation for modal analysis of the walls, such as FRF and PSD, on MMA output data. Though the used equipment (camera, tripod and lighting) was of low quality, in order to test the methodology in an unfavourable environment with very high data noise, the estimate of the first modal frequency showed a good agreement with modal analysis by a more conventional optical system used as reference.

Keywords: monuments, monitoring, magnified motion, augmented motion.

1 INTRODUCTION

The historic masonry protection is a primary task in order to preserve monuments and the first step surely is to arrange a suitable monitoring. Besides the usual wear due to age, traffic, weather, we have also to face earthquakes as a major threat. Consequently, researchers are trying continuously to ameliorate the monitoring systems, hopefully in the sense of enlarging and extending their predictive capabilities. To support engineers, a variety of devices has been developed: capacitive sensors, contact accelerometers, optical sensors, lasers, global positioning systems, linear variable differential transformers and video cameras. Costs, dimensions (too small or too large), complexity of the associated electronics, energy requirements, specialized operators, are all disadvantages to be faced during measurement campaigns [1]–[3].

However, the most important problem is the number of devices needed to saturate the surface of the structure to be analysed, at least in some cases. To overcome this issue, optical tools named “motion capture” are available today. Working in the infrared range, these motion capture systems are able to measure displacements of small reflecting markers positioned on the surface of the structure; in this case the motion capture control may run up to hundreds of markers [4], [5]. The 3DVision optical system described in [4], was the first of the kind used in a laboratory to measure displacements of structures stressed by a shaking table, with good results. Clearly, in the outdoor environment things change, since boundary conditions are not manageable anymore, false reflections increase too much, and, of course,



the dimension of the structure simply can be too large to be included in the optical range of the motion capture cameras. Given these circumstances, we propose to resort to a new methodology called the motion magnification analysis (MMA) [6].

The analysis of image sequences in the field of civil engineering is not new. For many years attempts to produce qualitative (visual) and even quantitative analysis using high quality videos of large structures have been conducted, but with poor results. This was because of the resolution in terms of pixels, of the noise, of the camera frame rate, computer time and finally because of the lack of appropriate algorithms able to deal with the extremely small motions related to building displacements.

These and other limitations have restricted the applications of digital vision methodologies to just a few cases. Nevertheless, recently important advances have been obtained by Freeman and collaborators of the Massachusetts Institute of Technology (MIT) [2]. Their algorithms, named motion magnification, seems able to act like a microscope for motion and, more importantly, in a reasonably short computer time. The latter point is crucial, as it is well known that video processing takes a lot of time and resources. Therefore, any viable approach must consider the reduction of the calculation time as an absolute priority.

The basic MMA version looks at intensity variations of each pixel and amplifies them, revealing small motions which are linearly related to intensity changes through a first order Taylor series. To validate this new methodology also in the field of building monitoring, we compare its performance with the 3DVision measurements, obtained during the tests of a stone masonry subjected to a very weak artificial earthquake.

The experiments were performed within the CoBra Project, funded by the Regione Lazio, a program to support important monuments from Lazio. The main objective of the Project CoBra is to develop and disseminate methods, technologies and advanced tools for the conservation of the cultural heritage. The Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) leads the project with its laboratories and several working groups.

2 THE EULERIAN MAGNIFIED MOTION ALGORITHM

Since our intention is only to give a general idea of the potentiality of the motion magnification we will not enter into the full formal description of the algorithms.

Videos are made up of a temporal sequence of 2D images, whose pixel intensity is $I(x, t)$. The 2D array of colour intensity is the spatial domain, while the time domain corresponds to the temporal sequence. Here, in order to describe the Eulerian version of the magnification algorithm, we consider only a 1D translating image with displacement $\delta(t)$. At the image-position x and video-time $t = 0$ it is: $I(x, 0) = f(x)$ (for the treatment of the general problem, see [2]). Translating for the quantity $\delta(t)$, we have:

$$I(x, t) = f(x - \delta(t)). \quad (1)$$

The final expression of the magnified motion by the constant α is defined as:

$$\Delta I = f(x - (1 + \alpha) \delta(t)). \quad (2)$$

Now, if the displacement $\delta(t)$ is small enough, it is possible to expand the relation (1) as the Taylor's first order series around x at the time t :

$$I(x, t) = f(x) - \delta(t) (\partial f / \partial x) + \varepsilon, \quad (3)$$

where ε is the error due to the Taylor's approximation. The intensity change at each pixel can be expressed as:

$$\Delta(x, t) = I(x, t) - I(x, 0) \quad (4)$$

that taking into account eqn. (3), becomes:

$$\Delta(x, t) = f(x) - \delta(t) (\partial f / \partial x) + \varepsilon - f(x) \quad (5)$$

and finally disregarding the error ε :

$$\Delta(x, t) \approx -\delta(t) (\partial f / \partial x) \quad (6)$$

meaning that the absolute pixel intensity variation Δ is proportional to the displacement and to the spatial gradient. Therefore, the pixel intensity can be written as follows:

$$I(x, t) \approx I(x, 0) + \Delta(x, t). \quad (7)$$

Magnifying the motion by a given constant α using eqns (3) and (4), simply means that pixel intensity $I(x, t)$ is replaced by the magnified pixel intensity $I_{\text{magn}}(x, t)$ according to the following:

$$I_{\text{magn}}(x, t) \approx I(x, 0) + \alpha \Delta(x, t) \approx f(x) - \delta(t) (\partial f / \partial x) - \alpha \delta(t) (\partial f / \partial x) + O(\varepsilon, \delta), \quad (8)$$

where $O(\varepsilon, \delta)$ is the remainder of the Taylor series. Then, the magnified intensity can be calculated as:

$$I_{\text{magn}}(x, t) \approx f(x) - (1 + \alpha) \delta(t) (\partial f / \partial x) \quad (9)$$

but eqn (9) is immediately derived from the first order Taylor's expansion of the exact magnified motion:

$$\Delta I = f(x) - (1 + \alpha) \delta(t) \quad (10)$$

and therefore:

$$f(x) - (1 + \alpha) \delta(t) \approx f(x) - (1 + \alpha) \delta(t) (\partial f / \partial x). \quad (11)$$

It is important to observe that (6) is obtained by a band-pass derivation, thus the process can be basically summarized as in Fig. 1. Therefore, we can say that to magnify the motion displacement it suffices to add $\alpha \Delta(x, t)$ to $I(x, t)$, as long as the Taylor's expansion (9) is valid, that is until its remainder $O(\varepsilon, \alpha)$ is small.

This limitation depends on the linear approach entailed in the Taylor's expansion, either if the initial expansion (3) or the amplification α are too large. In practice, to remain into the linearity bound, we need slowly changing images and small amplifications. Moreover, here we do not consider the noise of variance σ^2 to be added to the intensity, that is amplified too, resulting in an amplified noise variance $2\sigma^2\alpha^2$, thus the error to be evaluated should be $O(\varepsilon, \alpha, 2\sigma^2\alpha^2)$. Also, it should be noted that the calculation of $\Delta(x, t)$ implies the whole time span needed from frame 0 to the current frame at time t .



Figure 1: Temporal filtering applied to each pixel time history. Cut-off frequencies have to be chosen carefully in order to enclose the band of the phenomenon to be analysed and exclude other frequencies.

Therefore, if the video is a long-lasting one, the computational time may be a major problem. On the other hand, the Nyquist sampling theorem requires the frame rate of the video f_{fps} to be at least double than highest frequency of interest (f_{max}). In fact, to reproduce correctly a signal it is necessary that:

$$f_{sampling} \geq 2f_{max} \quad (12)$$

where f_{max} is the maximum frequency of the signal in the temporal domain, $f_{sampling}$ is the sampling frequency. Then (11) becomes:

$$f_{fps} \geq 2f_{max} \quad (13)$$

f_{fps} acting as a sampling frequency. Using a 28 fps video camera, the maximum frequency allowed is 14Hz: above this threshold it will be introduced spurious frequencies because of the aliasing. Furthermore, since in our elaborations the frequency resolution is 0.1Hz, the video duration must be at least 10 seconds.

Other physical limitations, such as illumination, shadows, camera unwanted vibrations, poor pixel resolution, low frame rate, presence of large motion, distance from the object, decrease severely the quality of the motion magnification and should be taken into account in order to achieve high-quality results.

3 THE GENERAL METHODOLOGY

The first step is to record a video of the structure we are interested in, taking good care to avoid large motions such as people passing by in front of the camera, swinging cables or non-fixed mechanical parts. The presence of large motions is the most significant source of noise for the MM, requiring the experimenter to isolate part of the image, although usually this is not a feasible option. The issue is still open.

In our case, we recorded by means of a commercial tablet, pixel resolution 720x1280, frame rate 28 fps, and then videos were processed by the MM algorithm [1]. The basic methodology is to take advantage of the large number of pixels contained in an image. Theoretically, we could have 921600 “virtual sensors”, meaning that each pixel has a time history of intensity variation (colour or grey scale), from the first frame to the last one. These time series contain the information about the displacements of the physical points related to the pixels (but they are not real displacements). Of course, it would be too cumbersome to analyse all the virtual sensors and moreover not all the surface of the structure generates useful information, therefore we identify a small area of the surface with high signal-to-noise ratio (SNR). On the left of Fig. 2 is showed the masonry, enclosed by safety beams to prevent risks for the operators, tested on the shaking table. We select a high SNR part of the image

(see Fig. 3) and follow its pixel variation along the video. Typically, an area of 19×78 pixel is selected as in Fig. 3, providing 1482 virtual sensors. Note on the right the small 19×79 pixels area (enlarged) and three markers of the 3DVision motion capture system. The system has available ten infrared cameras positioned around the shaking table to receive light reflection from markers arranged on the structure. Two cameras are able to triangulate and track the position of a marker with a precision of 0.01mm, therefore velocity and acceleration can be derived by numerical methods. These signals from the markers provide the reference for the comparison with the MMA.

It is also to be noted that usually the presence of edges or texture is helpful for the MMA, but unfortunately, the area in the red boxes of Fig. 3 is rather homogeneous. These circumstances produce a certain amount of noise, to be added to other disturbances. Signals from the magnified motion technique does not provide directly the displacements, although they could be recovered by means of the constant contours method [2]. On the other hand, they may be used to calculate the power spectral density (PSD) or the FFT, allowing the modal analysis and the calculation of the frequency response function (FRF). Hence, the 3D Vision system is a convenient reference to evaluate the MMA performance in the frequency domain.

4 MASONRY WALLS SHAKING TABLE TEST RESULTS

The optical movement detection 3DVision system tracks the displacements of selected points on the masonry in the dynamic laboratory tests at the shaking table earthquake test facilities of the ENEA Casaccia Research Center, located near Rome, Italy (Fig. 4).



Figure 2: The tested walls masonry, typical of central Italy, a region prone to earthquakes [8], [9]. On the left: lateral view with the safety framework, right: frontal view. The size of the two panels are 3.45m height and 3.31m width.

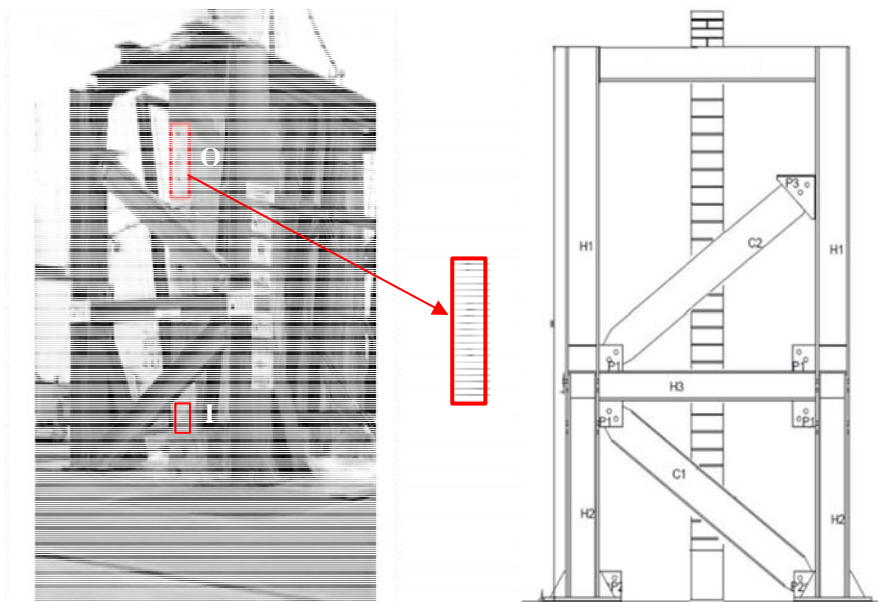


Figure 3: A frame from the magnified motion video (left): note how the image is blurred because of the noise. O: output markers, I: input markers of the FRF function. Right: the surface of the virtual sensors with some markers of the 3DVision system (red boxes) and the side view of the stone masonry.



Figure 4: The ENEA Casaccia seismic test facility: the two shaking tables.

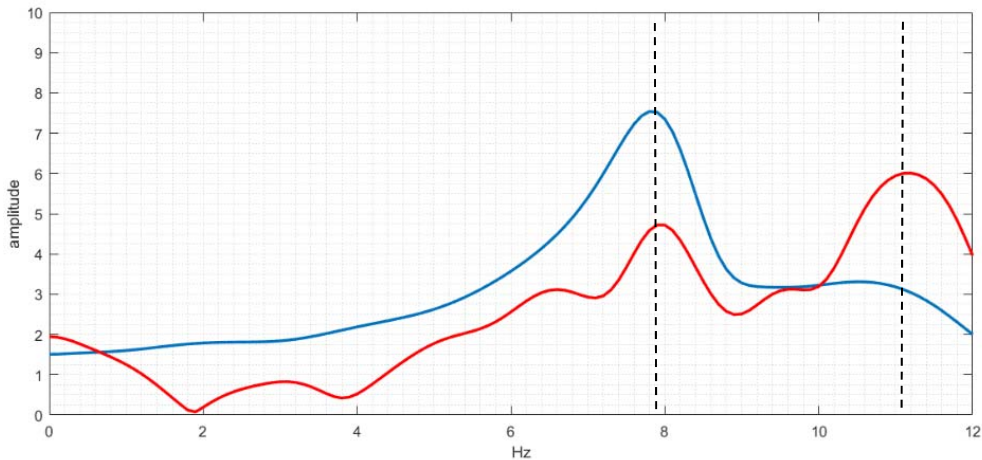


Figure 5: In blue the FRF from the 3D Vision, red the analogous from the magnified motion technique. Signals are averaged to produce just one signal, resulting in a low pass filtering action and successively normalized. The first peak is recovered within an absolute error of 0.1Hz and the second within 0.4Hz.

The object tracking and the 3D optical movement detection have been experimented to track the motion during the low frequency (up to 30Hz) vibrations typical of the shaking table tests. Ten high resolution (up to 4MPixel) infrared cameras are able to measure accurate 3D positions of hundreds of markers placed on the structure during the seismic tests [4]. This monitoring technique allows measuring 3 axial absolute displacements with easy and fast test set-up, high accuracy and the possibility to link the 3D-motion time histories of the tracked markers with CAD drawings of the structure and validate the finite element model (FEM) in real time experimental data assimilation [5].

The intensity variation of the 19x79-pixel area from the motion magnification (MM) and those of the 3D Vision are averaged to produce just one signal, in order to reduce noise, since the average is somewhat like a low pass filtering procedure.

At this point, we calculate the FRF for both the 3D Vision and for the MM signals. The frequency response function is a mathematical representation of the relation between two points on a structure (the input and the output) in the frequency domain. The 3D Vision output signal is indicated with O and the input signal of markers positioned at the basis of the wall is indicated with I in Fig. 3. Measuring the input and the output accelerations we describe a functional relation between these points. To calculate the FRF the experimental EMA SI-MO methodology was used:

$$H_I(\omega) = S_{io}(\omega) / S_{ii}(\omega), \quad (14)$$

where S_{io} is the FFT spectrum of the input-output cross-correlation function, S_{ii} is the FFT spectrum of the auto-correlation function. Modal parameters allow to evaluate the seismic vulnerability of a structure since a significant variation before and after a certain period of time may indicate a damage.

In Fig. 5 the first resonance peak at 7.9Hz corresponds to the 3DVision at 7.8Hz (up to the frequency resolution 0.1Hz), while the second peak at 11.1Hz is not aligned to the 3DVision at 10.7Hz.

However, it should be considered that the 3DVision has a sampling rate of 200Hz and the video camera has a frame rate of 28Hz. Moreover, the two signals are not well synchronized in time, meaning the start recording instants do not coincide exactly. In particular, the different sampling rate has introduced in the MM FRF spurious peaks below 8Hz and above 12Hz. To overcome this problem, it suffices a better recording device.

To confirm further the 7.9Hz (3DVision) and the 7.8Hz (MM) results, we compare them with an independent estimation of the FRF carried out in [8], by means of a standard accelerometer equipment, that is 7.89Hz for the first FRF peak. Considering reliable the accelerometer value, the 3DVision system yields an error of -0.127% and the MM an error of -1.141%.

Another interesting feature of the MMA is the visual detection of vulnerable points of the structure. In Fig. 2 the dotted red line in the frontal view indicates a crack produced by a strong simulated earthquake that occurred exactly where the MM video showed the largest displacements during a low-intensity identification test on the shaking table. In this sense, MMA provides a predictive capability.

5 CONCLUSIONS

The continuous developments in digital vision technologies are very promising. Even if till a few years ago such kind of methodologies were not viable, recent advances in the digital image and video processing have opened the door to applications to the analysis of vibrations, in particular by motion magnification strategies.

Actually, during our laboratory tests, MMA was effective in amplifying subtle motions in videos, making distinguishable the tiny displacements of structures exposed to mechanical disturbances and facilitating the evaluation of the building stability within a reasonable time. Advantages are many: a large number of “virtual sensors” available, no wires, reduced amount of data storage, no physical contact, simplicity, low costs and predictive capabilities. Unfortunately, for a number of reasons, noise is a pervasive obstacle to MMA, especially when the recording device is a low-quality video camera and indoor laboratory parameters are not optimal. Nevertheless, we showed that the frequency domain results obtained previously for small mechanical structures are extensible to masonry structures for seismic applications.

ACKNOWLEDGEMENTS

Authors would like to acknowledge the CoBra Project funded by Regione Lazio and the Soprintendenza Archeologica di Roma.

REFERENCES

- [1] Hwang, J., Yun, H., Park, S.K., Lee, D. & Hong, S., Optimal methods of RTK-GPS/accelerometer integration to monitor the displacement of structures. *Sensors*, **12**, pp. 1014–1034, 2012.
- [2] Wadhwa, N. et al., Eulerian Video Magnification and Analysis. *Communications of the ACM*, **60**(1), pp. 87–95, 2017.
- [3] Yu-Wu, H. et al., Eulerian Video Magnification for Revealing Subtle Changes in the World, <https://people.csail.mit.edu/mrub/papers/vidmag.pdf>. Accessed on: 2 Mar. 2017.



- [4] De Canio, G., Mongelli, M. & Roselli, I., 3D Motion Capture Application to Seismic Tests at ENEA Casaccia Research Center: 3D Vision System and DySCo Virtual Lab. *WIT Transactions on the Built Environment*, **134**, pp. 803–814, 2013.
- [5] Mongelli, M. et al., Experimental tests of reinforced concrete buildings and ENEA DySCo Virtual Laboratory. *Proceedings of 5th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII-5)*, 2011, Cancùn, Mexico.
- [6] Fioriti, V., Roselli, I., Tati, A. & De Canio, G., Motion magnification for urban buildings. *Proceedings of 12th International Conference on Critical Information Infrastructures Security*, 2017, Lucca, Italy.
- [7] Sinou, J., *A review of damage detection and health monitoring of mechanical systems from changes in the measurement of linear and non-linear vibrations. Mechanical Vibrations: Measurement, Effects and Control*, Nova Science Publishers, Inc. **702**, pp. 643–645, 2009.
- [8] Focaccetti E., *Prove su tavola vibrante di pareti in muratura rinforzate con materiali compositi a matrice inorganica comprendenti tessuti in acciaio e reti in fibra di basalto applicate con malte a base di calce idraulica naturale*. Thesis, supervisor prof. G. de Felice, Roma Tre University, 2017.
- [9] De Canio, G. et al., Passive 3D motion optical data in shaking table tests of a SRG reinforced masonry wall. *Earthquakes and Structures*, **10**(1), pp. 53–71, 2016.

