Motion magnification analysis applied to the dynamic identification of historic constructions

Vincenzo Fioriti, Ivan Roselli, Angelo Tati, Roberto Romano and Gerardo De Canio
ENEA Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Italy
Vincenzo.fioriti@enea.it

Abstract. We applied a new methodology, namely the motion magnification (MM), to the dynamic identification of historic constructions. MM acts like a microscope amplifying small motions in video sequences, therefore tiny motion patterns are made visible with the naked eye. This technique provides advantages of particular interest to the dynamic identification: no wires, no physical contact, simplicity and low costs. In this paper, we investigated the ambient vibration monitoring of historic structures in urban environments, which is a relevant issue for the health surveying and early damage detection, particularly for ancient buildings. We give an introduction to the MM methodology and describe its practical application to the frequency domain through two case-studies: the so-called Temple of Minerva Medica of Rome and the Ponte delle Torri of Spoleto. Since in the outdoor environment the MM is much more prone to noise because of the wind, shadows, atmospheric refraction, light reflection, distance from the object, weak vibration sources, the case study of the Ponte delle Torri represents a hard test-bed for MM. However, the MM estimates of the first modes showed a good agreement with the experimental data of contact velocimeters.

1. Introduction
Vibration monitoring is a key issue for the structural health analysis of the cultural heritage. To this end, today a new image processing tool, the motion magnification (MM) is available. Motion magnification acts like a microscope for micro-motions in videos, but saving the topology of the image sequence and unveiling visual patterns hardly visible with the naked eye. The motion magnification use the spatial resolution of the video-camera to extract physical properties from images to make inferences about the dynamical behavior of the object. Researchers are very interested in assessing the MM analysis (MMA) feasibility, since conventional vibrometers are surely more precise, but expensive and much less practical. Moreover, evaluating the health of large structures, such as historical monuments in a short time span and possibly by contactless devices, may produce a major breakthrough for civil engineering. The usage of video sequences in the field of civil engineering is not new. For many years attempts to produce qualitative and even quantitative analysis by means of videos of large structures have been conducted, but with poor results. This because of the resolution in terms of pixels, of the camera frame rate, computer time and finally because of the lack of proper algorithms, able to deal with the extremely small motions related to building displacements. These and others limitations have restricted in the past the application of digital video methodologies, till the
important advances obtained during the last years at the Massachusetts Institute of Technology (MIT) by Freeman and coworkers [3]. Their algorithm, named motion magnification, works within a reasonably short elaboration time. The point is crucial, as it is well known that video processing takes a lot of time and resources, preventing a viable application to the real world.

Recently, a number of indoor experiments conducted on simple geometries like rods and other small objects at MIT, as well as on masonries tested on the shaking table at the ENEA Casaccia facilities, showed the reliability of this methodology compared to lasers vibrometers or contact sensors. However, in the outdoor environment the analysis is much more difficult.

Moreover, although revealing the smallest movements of the structure is per se an interesting feature, we need a quantitative analysis to assess the health of the monument by the classical frequency domain methods. To this end, we have chosen the environment of the so-called temple of Minerva Medica and the Ponte delle Torri as a demanding outdoor test-bed for MMA, using as a reference the modal analysis provided by [1].

In [1] experimental velocimetry data and the finite element model (FEM) are in strong accordance, representing an appropriate benchmark for MMA. Therefore, in this paper we use extensively the results of [1] to evaluate the performance of the motion magnification in the frequency domain.

2. The Ponte delle Torri of Spoleto
The Ponte delle Torri is a large historical construction that connects Colle Sant’Elia with Mount Monteluco in Spoleto, Italy (Figure 1). It was probably built in the 13th century, possibly on Etruscan or Roman ruins [1]. The bridge superstructure is made up of a pedestrian deck, provided with a water canal on one side, supported by lancet arcades and stone piers known as "towers".

The bridge has a stone structure with shoulders and piers made up of rubble walls in a square matrix assembled by mortar and lime. The structure is only apparently regular. In fact, piers and arches have all different shapes and sizes, as well as the walls towards Colle Sant’Elia and Monteluco are also diverse in the masonry texture, due to their different construction timing and subsequent rebuilding interventions. The bridge has an overall length of about 230 m, while the highest tower is over 70 m. The bridge has a state of damage typical of historic masonry bridges. At the top of the arches, especially those on the slope of Monteluco, there are heavy water infiltrations resulting in losses of mortar binder and wall apparatus skiving. Moreover, a widespread damaged layer, especially in the lower part of the piers, was detected. This has caused the expulsion of some cornerstones of the piers, the disarticulation of the masonry texture and the collapse of portions of the outer layers [2].

2.1. The OMA identification
In order to identify the dynamic properties of the bridge through experimental data, three seismographs were used to acquire the ambient vibration on site [1].

![Figure 1. The Ponte delle Torri of Spoleto (Northeast view).](image1)

![Figure 2. Positions of velocimeters (in red) in the first experimental setup [1].](image2)
Figure 3. Modal estimation graph by the Frequency Domain Decomposition and the Enhanced Frequency Domain Decomposition methods. Although these two modal calculation methods are slightly different, results are the same.

The data presented here were acquired on 29 May 2017. The used seismographs, SL06 recorders (SARA Instruments) equipped with triaxial electrodynamic velocimeters, were set to 200 Hz sampling frequency. Multiple Test Sets Measurement Procedures (MTSP) were performed with one common measurement point on top of the central pier of the bridge, as reference sensor (Figure 2).

As the reference sensor was staying in the same positions during all setups, it basically measured the mode shapes in this position over and over, while other sensors were moved to different positions on the bridge. Such reference position was determined as the measurement point where the modes of interest were supposed to have the highest response level according to FEM modal shapes. Data were acquired in eight configurations, each representing a test setup in the MTSP. All configurations were acquired for at least 20 minutes.

The Operational Modal Analysis (OMA) techniques implemented in the ARTeMIS Modal Pro software were used to elaborate the experimental data of the velocimeters (Figure 3 and Table 1). Among the several OMA techniques available, the Frequency Domain Decomposition (FDD), the Enhanced Frequency Domain Decomposition (EFDD) and the Stochastic Subspace Identification (SSI) were utilized. EFDD provides also an estimation of modal damping, SSI is a more sophisticated and automatic procedure based on time-domain approach [1, 2].

The results of the identification of the modal frequencies averaged over the OMA techniques are resumed in Table 1. Also a finite element analysis (FEM) has been carried out along with the experimental campaign, furtherly confirming results of Table 1 [1, 2].

| Mode | Averaged modal frequency (Hz) |
|------|------------------------------|
| 1    | 0.632                        |
| 2    | 1.011                        |
| 3    | 1.496                        |
| 4    | 1.975                        |

3. The 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, see [3, 4, 5] for an extensive discussion. Videos are made of a temporal sequence of 2D images, whose pixel intensity is \( I(x, t) \). The
2D array of color 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 [3]). Translating for the quantity \( \delta(t) \), we have:

\[
I(x, t) = f(x - \delta(t))
\]

The final expression of the magnified motion by the constant \( \alpha \) is defined as:

\[
\Delta I = f(x - (1 + \alpha) \delta(t))
\]

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 \( t \):

\[
I(x, t) = f(x) - \delta(t) \left( \frac{\partial f}{\partial x} \right) + \varepsilon
\]

where \( \varepsilon \) is the error due to the Taylor’s approximation. The intensity change at each pixel can be expressed as:

\[
\Delta(x, t) = I(x, t) - I(x, 0)
\]

that taking into account (3), becomes:

\[
\Delta(x, t) = f(x) - \delta(t) \left( \frac{\partial f}{\partial x} \right) + \varepsilon - f(x)
\]

and finally disregarding the error \( \varepsilon \):

\[
\Delta(x, t) \approx - \delta(t) \left( \frac{\partial f}{\partial x} \right)
\]

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

\[
I(x, t) \approx I(x, 0) + \Delta(x, t)
\]

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

\[
I_{magn}(x, t) = I(x, 0) + \alpha \Delta(x, t) \approx f(x) - \delta(t) \left( \frac{\partial f}{\partial x} \right) - \alpha \delta(t) \left( \frac{\partial f}{\partial x} \right) + O(\varepsilon, \delta)
\]

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

\[
I_{magn}(x, t) \approx f(x) - (1 + \alpha) \delta(t) \left( \frac{\partial f}{\partial x} \right)
\]

But (9) is immediately derived from the first order Taylor’s expansion of the exact magnified motion:

\[
\Delta I = f(x - (1 + \alpha) \delta(t))
\]

and therefore:
\[
f(x - (1 + \alpha) \delta(t)) \approx f(x) - (1 + \alpha) \delta(t) \frac{\partial f}{\partial x}
\]  

(11)

It is important to observe that (6) is obtained by a band-pass derivation. 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 \( \alpha \) are too large.

In practice, to remain into the linearity bound, we need slowly changing the images and small amplifications. Moreover, here we do not consider the noise of variance \( \sigma^2 \) 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 \( O(\varepsilon, \alpha, 2\sigma^2 \alpha^2) \) actually.

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 \( f_{max} \). Thus, to reproduce correctly a signal it is necessary that:

\[
f_{sampling} \geq 2f_{max}
\]  

(12)

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

\[
f_{fps} \geq 2f_{max}
\]  

(13)

with \( f_{fps} \) acting as a sampling frequency. Using a 50 fps video camera, the maximum frequency allowed is 25 Hz: above this threshold spurious frequencies are introduced because of the aliasing. Furthermore, since in our elaborations the video duration is about 10 seconds, the theoretical frequency resolution is equal to 0.1 Hz. It must also be said that the short time-span of the video of course has a negative impact on the PSD calculations.

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. All these difficulties are present in the case-study of the Ponte delle Torri.

4. MMA modal analysis

Firstly, we recorded a video of the monument during the campaign of the 29 May 2017, taking care to avoid large motions such as people passing by in front of the camera and swinging objects. The presence of large motions is one of the most important sources of noise for the MM, requiring the experimenter to isolate part of the image, although usually this is not a feasible option. We recorded by means of a commercial camera, (pixel resolution 360 x 445, frame rate 50 fps), and then videos were processed by the MM algorithm [3]. The basic methodology is to take advantage of the large number of pixels contained in an image. Theoretically, we could have 160200 “virtual sensors”, meaning that each pixel provides 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 (although they are not real displacements). Of course, not all the surface of the structure generates useful information, therefore we identify a small area possibly with high signal-to-noise ratio (SNR). The identification of this region of interest (ROI) is a crucial point of the procedure. The angle of the camera is much less than 90° (Figure 4). This circumstance is unfavorable, since the magnification will be affected by a distortion. Also the air refraction and the wind provide a substantial source of noise, affecting the algorithm. No attempt to compensate for these distortions has been made.
The main source of vibration is the wind, although a road flanks the monuments on the right side. In any case, man-made vibrations may be considered of low intensity. The camera is positioned on the road near the monument as in Figure 5, hence wind and traffic affect its stability too.

To run the MIT algorithm, we need to set some parameters. In particular, the magnification band and the amplification factor are the more important. We chose a band of 0.5 – 2.5 Hz and an amplification factor of 140. Similar values would be also acceptable, nevertheless, since the MM acts as a pass-band filter, choosing a band close to the actual modes gives better results. In our case we have been supported by the previous analysis [1], but generally one must rely on experience to decide a suitable band.

After the MIT phase based magnification step [3, 4, 5], we have to select the ROI of the virtual sensors, meaning to pick some pixels whose time history of color variation will be translated into the frequency domain. We choose the area as in the red box of Figure 5, because it is morphologically different from the wall and provides some edges to the algorithm. The presence of marked edges or texture would be very helpful for the MMA, but unfortunately the surface of the monument is rather homogeneous. These circumstances produce a large amount of noise, to be added to other disturbances.

The ROI is part of a lamppost, in-built to the wall about 40 m from the camera; however only the red box pixels are able to supply low-noise information, because are close to the structure. In fact, the upper part of the lamppost spreads spurious frequencies, due to its own resonances. Other choices of the ROI have not improved the measurements. Signals provided by MM contain displacement information, but they cannot be used immediately as real displacements. Anyway, acceleration may be calculated just like it can be done with the velocimeter displacement measurements, and the PSD obtained as well. Every pixel provides a signal, therefore an average over all the PSDs is calculated.

4.1. Results
The ROI area of Figure 4 is identified qualitatively by trial and error considering the number of peaks in the PSD and their inter-distance; a dense presence of peaks suggests a high level of noise and of course is misleading, therefore only PSD spectra with identifiable frequencies are acceptable. In practice, to avoid a biased choice, one selects manually the area several times and repeats the calculation procedure. In our case, we find that ten iterations of the MM-PSD procedure is a good trade-off between speed and precision (intended as closeness to the mean value).

Table 2 contains the MM modal frequencies calculated over ten runs, compared to the experimental results by OMA. Table 2 shows a good agreement with the experimental measurements obtained both from the FEM method and data collect by standard velocimeters [1].

In Figure 5 a typical MM-PSD is shown with passband filter in the range 0.5 – 2.5 Hz. The smooth appearance of the MM-PSD depends on the short time span of the video that decreases the PSD frequency resolution.
Table 2. Numerical values of first 4 modal frequencies by OMA and by magnified motion analysis. The last column is the percentage error.

| Mode | Averaged OMA frequency (Hz) | Averaged MM frequency (Hz) | error % |
|------|-----------------------------|-----------------------------|---------|
| 1    | 0.632                       | 0.685                       | -8.39 % |
| 2    | 1.011                       | 1.014                       | -0.30 % |
| 3    | 1.496                       | 1.311                       | 12.37 % |
| 4    | 1.975                       | 2.068                       | -4.71 % |

Figure 6 illustrates all the values obtained over the ten runs of the MM-PSD procedure, along with their averages and the OMA values. Note that some outliers are evidently present, therefore removing them by statistical methods would improve these results.

An analogous experiment was carried out on 19 April 2017 at the temple of Minerva Medica in Rome [6]. In Figure 7 we can see the south side of the temple, a velocimeter in the blue circle, and the ROI area of the virtual sensors in the red box. A tramway flanks the monument at only 2 m, producing very strong vibrations [7]. In this preliminary analysis only the first mode was analyzed. The estimated value of the modal frequency by MMA was 1.950 Hz, which is roughly close to the value obtained through OMA of velocimeters data (2.057 Hz) that leads to an error of 5.2 %.

To explain why the MMA appears less precise for the Ponte delle Torri, it should be taken into account the adverse acquisition conditions of the Ponte delle Torri, that prevented a correct recording angle for the camera, the distance from the target, the absence of suitable edges and of course the weak vibration sources. This last point seems the decisive factor. The MM is able to amplify sufficiently even the micro-displacements caused by the wind, but the signal-to-noise ratio (SNR) remains low if the vibration sources are weak, since further increasing the amplification factor would produce artifacts in the magnified video. In other words, the displacements can be recovered, but amid a relatively high noise level, that affects the MM-PSD elaborations.

In the case of the temple of Minerva Medica case-study, the displacements were two orders of magnitude larger, producing a much more favourable SNR for the ROI. Moreover, the camera angle was about 90° and the distance reduced to 8 m.

Yet, having tested the outdoor MM over four modes with an error less than 12.5% in the worst case, results can be regarded as a very promising step toward an outdoor digital video modal analysis.
5. Conclusions
We have investigated the application of the magnified motion, a novel video processing technique, to the Ponte delle Torri of Spoleto and to the temple of Minerva Medica, in order to develop a video driven modal analysis. We have discussed the role of noise and of the virtual sensors, that require careful boundary conditions to the video recording. Given the adverse scenario produced by a low SNR, the estimates of the first four modes for the Ponte delle Torri are more than satisfying, while for the temple of Minerva Medica case a better SNR allowed an almost exact identification. When a dense sensor monitoring is needed or it is impossible to deploy laser vibrometers and velocimeters, the magnified motion technique offers a viable, contactless, low cost solution. Many issues still deserve to be deepened: noise reduction, operative guidelines for outdoor analysis and finally the design of a hardware customization for real-time applications. However, it is probable that all these problems will be fixed over the next few years.

Acknowledgments
The authors wish to thank dr. Valerio Cirulli, for useful discussions and support, and the CO.B.RA. project funded by Lazio Region, within which this work was financed.

References
[1] Roselli I, Malena M, Mongelli M, Cavallagli N, Gioffrè M, De Canio G and De Felice G 2018 Health assessment and ambient vibration testing of the "Ponte delle Torri" of Spoleto during the 2016-17 Central Italy seismic sequence Int. J. Civil Structural Health Monitoring 8(2)
[2] De Canio G, Mongelli M, Roselli I, Tati A, Addessi D, Nocera M and Liberatore D 2016 Numerical and operational modal analyses of the “Ponte delle Torri”, Spoleto, Italy Proc. Int. Conf. in Structural Analysis of Historical Constructions SAHC16 (Leuven) pp 752–7
[3] Wadhwa N, Wu H, Davis A, Rubinstein, M, Shih E, Mysore G, Chen J, Buyukozturk O, Guttag J, Freeman W and, Durand F 2017 Eulerian Video Magnification and Analysis Communications of the ACM 60 pp 87-95
[4] Yu-Wu H et al. 2017 Eulerian Video Magnification for Revealing Subtle Changes in the World, https://people.csail.mit.edu/mrub/papers/vidmag.pdf. Accessed on 2 March 2017
[5] Wadhwa N et al. 2017 Motion microscopy for visualizing and quantifying small motions PNAS 114 44 pp 11639–44
[6] Barbera M, Magnani Cianetti M and Barrano S 2014 Da Massenzio a Costantino: le indagini in corso nel c. d. tempio di Minerva Medica Proc. Int. Conf. of CISEM - La villa restaurata e i nuovi studi sull’edilizia residenziale tardoantica (in Italian)
[7] Roselli I, Fioriti V, Bellagamba I, Mongelli M, Tati A, Barbera M, Magnani Cianetti M and De Canio G 2017 Urban transport vibrations and cultural heritage sites in Rome: the cases of the temple of Minerva Medica and of the Catacomb of Priscilla WIT Transactions on Ecology and the Environment 223 pp 335–43