Measurement 129 (2018) 375–380

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Motion Magnification Analysis for structural monitoring of ancient constructions

Vincenzo Fioriti*, Ivan Roselli, Angelo Tatì, Roberto Romano, Gerardo De Canio

ENEA, CR Casaccia, via Anguillarese 301, Rome, Italy

ARTICLE INFO

Keywords:
Motion Magnification
Vibration monitoring
Dynamic identification
Historic structures

ABSTRACT

A new methodology for digital image processing, namely the Motion Magnification (MM), allows to magnify small displacements of large structures. MM acts like a microscope for motion in video sequences, but affecting only some groups of pixels. The processed videos unveil motions hardly visible with the naked eye and allow a more effective frequency domain analysis. We applied the MM method to several historic structures, including a 1:10-scale mockup of Hagia Irene in Constantinople tested on shaking table, the so-called Temple of Minerva Medica in Rome and the Ponte delle Torri of Spoleto. MM algorithms parameters were calibrated by comparison with reference consolidated modal identification methods applied to conventional velocimeters data. Encouraging results were obtained in terms of vibration monitoring and modal analysis for dynamic identification of the studied structures, offering a low-cost, viable support to the standard vibration sensing equipment, such as contact velocimeters, laser vibrometers and others.

1. Introduction

Vibration monitoring of historic monuments in the urban environment is a relevant issue for health survey and damaging detection. Today, a new digital image processing method, namely the Motion Magnification Analysis (MMA), allows to magnify small displacements in video motions. Motion magnification acts like a microscope for motion in video sequences, but affecting only some groups of pixels, unveiling motions hardly visible with the naked eye. The motion magnification uses 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 method's feasibility, since conventional devices are surely more precise, but expensive and much less practical. Recently, a number of experiments conducted on simple geometries like rods and other small objects as well as on bridges, have demonstrated the reliability of this methodology compared to contact accelerometers and laser vibrometers [1–3]. In this paper, we extend the MMA to the analysis of a 1:10 scale mockup of the church of Hagia Irene of Constantinople tested on a shaking table, to the so-called Temple of Minerva Medica in Rome and to the Ponte delle Torri of Spoleto. Results show that MMA allows a visual identification of vibration mode shapes and of the most vulnerable elements of the structures. Though our equipment was of low quality in order to test the methodology in an adverse environment, results were very good. Evaluating the health of large structures such as a historical monument in a short time span and possibly by simple devices that do not require expert operators, may be a pivotal issue in civil engineering. Thus, the availability of intuitive methodologies such as those based on a digital acquisition of images may result in a major breakthrough. However, 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 a building displacement. These and others limitations have restricted in the past the applications of digital vision methodologies to just a few cases. Nevertheless, recently important advances have been obtained by Freeman and collaborators at the Massachusetts Institute of Technology [4]. Their algorithm, named motion magnification, seems able to act like a microscope for motion and, more importantly, in a reasonably short elaboration time. The latter point is crucial, as it is well known that image 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, revealing small motions linearly related to intensity changes through a first order Taylor series, for small

* Corresponding author.
E-mail address: vincenzo.fioriti@enea.it (V. Fioriti).

https://doi.org/10.1016/j.measurement.2018.07.055
Received 22 March 2018; Received in revised form 18 July 2018; Accepted 20 July 2018
Available online 21 July 2018
0263-2241/

variations. Since our intention is only to give a general idea of the potentiality of the motion magnification, we will not enter into the formal description of the algorithm. Rather we will propose some practical implementation examples: a laboratory case-study and more importantly, two monuments such as the so-called Temple of Minerva Medica in Rome and the Ponte delle Torri of Spoleto.

The original video files were named:

- "Video Hagia Irene.ppt";
- "MMIII_ORIGINAL_115959";
- "SPOLETO_ORIGINAL_2578_crop".

While the motion magnified video files were named:

- "Video Hagia Irene.ppt";
- "MMIII_magnified_115959";
- "SPOLETO_magnified_2578_crop_alph140".

All above files are downloadable at the following link: https://drive.google.com/drive/folders/0Bz540aXsdKTnbjdsQVl6TzBYbVU?usp=sharing.

2. The algorithm

Here we will describe the Eulerian version of MM [4], although actually we have used the phase based version [5] to process the videos. Videos are made up 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 do-main corresponds to the temporal sequence. We consider a 1-D translating image with displacement $\delta(t)$. $I(x, 0) = f(x)$ at the image-position x and video-time $t = 0$ (for the treatment of the general problem, see [2]). We have:

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

The final expression of its motion magnified by 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 Taylor's first order series around x, at 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 and to δ being non-zero. The intensity change at each pixel can be expressed as:

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

Which, taking into account Eq. (3), becomes:

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

and finally:

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

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

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

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

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

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

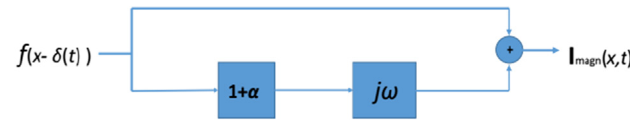

Fig. 1. Temporal filtering applied to each pixel time history. Cut-off frequencies have to be chosen carefully in order to enclose the band of the phenomenon to be analyzed and exclude other frequencies.

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

but Eq. (9) is immediately derived from the first order Taylor's expansion of the magnified motion of Eq. (2).

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 from frame 0 to the current frame at the time t.

If the video is long-lasting, the required computer time may be a major problem. Other physical limitations, such as the ones regarding 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, should also be taken into account in order to achieve good-quality results. In particular, the scene illumination should remain constant, as changing the background light could produce apparent motions. In fact, shadows and the sun light affect severely the MM, since any pixel intensity variation is considered by the algorithm just like a motion variation, but actually the variation is noise, not a real movement.

Finally, we note that the Shannon-Nyquist Theorem has to be respected. In fact, to reproduce correctly a signal it is necessary the condition:

$$f_{sampling} \geqslant 2f_{max} \quad (10)$$

where f_{max} is the maximum frequency of the signal in the temporal domain, $f_{sampling}$ is the sampling frequency. Here (10) becomes:

$$f_{fps} \geqslant 2f_{max} \quad (11)$$

where f_{fps} acts as a sampling frequency. Therefore, using a 28 fps video-camera, the maximum frequency allowed is 14 Hz, frequencies above this threshold will introduce spurious peaks because of the aliasing.

3. Experimental applications

Videos have been recorded by means of low resolution, low frame-rate video-cameras, both in laboratory conditions and outdoor in the urban environment. The laboratory tests have been carried out at the ENEA shaking tables facility of the Casaccia research center, located near Rome, on a mockup of Hagia Irene, an ancient church located in Istanbul [6]. The indoor experiments allowed the assessment of the effects of image noise and, consequently, to point out a strategy for the image noise reduction procedure (e.g. the image skeletonization method, see Fig. 2 and the video Hagia_Irene.ppt). The distance between the camera and the mockup was of 11 m. Also, indoor experimentation helped the comprehension of the effect of filtering and processing parameters to be implemented in the MM algorithm in a controlled environment.