Motion estimation of a cantilever beam in the quasi-static case based on the monogenic signal and the Lucas-Kanade optical flow

Jimmy Touzet, Hugo André, Frédéric Bonnardot, Simon Chesné, Olivier Alata

Motion estimation of a cantilever beam in the quasi-static case based on the monogenic signal and the Lucas-Kanade optical flow

J. Touzet [1], **H. André** [1], **F. Bonnardot** [1], **S. Chesné** [2], **O. Alata** [3]
[1] Laboratoire d'Analyse des Signaux et des Processus Industriels, EA3059, Universite Jean Monnet de Saint-Etienne, Universite de Lyon, F-42023 Saint-Etienne, France

[2] Université de Lyon, CNRS INSA-Lyon, LaMCoS UMR5259, F-69621, France

[3] Laboratoire Hubert Curien, UMR CNRS 5516, UJM-Saint-Étienne, IOGS, Univ. de Lyon, 42023, Saint-Étienne, France

July 4, 2022

Abstract

Who doesn't have a smartphone in their pocket? Today, there is a real interest in using video cameras for operational modal analysis or preventive maintenance applications instead of using traditional sensors such as accelerometers or strain gauges. Indeed, the video camera allows a contactless diagnosis, i.e. does not modify the properties of the observed structure and allows synchronous full-field measurement of displacement and strain, all with very few cables and equipment. The potential of video-based vibration monitoring is all the more attractive as it can be achieved through the use of any smartphone. The first aim of this work is to provide a comparison between two different methods of image processing and the results from a classical sensor which is the accelerometer. These methods allow a quick and simple follow-up of the beam's displacement field from the tracking of its contours with a high-speed camera. The other objective is to determine the conditions under which such good results can be obtained with a smartphone camera. To meet these multiple objectives, a comparison study is implemented on a test bench presenting a cantilever beam excited by a piezoelectric actuator and simultaneously observed by a high-speed camera and a smartphone. The processing of the videos is carried out two different methods. The quality of the obtained results is analyzed from a metrological and computational cost point of view and validated by standard accelerometer measurement. The proposed methodology is promising in the case of an embedded beam, and the adaptation of this method to more complex shapes than beams seems feasible as long as the detection of the contours can be done efficiently.

1 Introduction

Today, there is a real interest in using video cameras for operational modal analysis or preventive maintenance applications instead of using traditional sensors such as accelerometers or strain gauges. Indeed, the video camera allows a contact-less diagnosis, i.e. does not modify the properties of the observed structure and allows synchronous full-field measurement of displacement and strain, all with very few cables and equipment. The potential of video-based vibration monitoring is all the more attractive as it can be achieved

through the use of any smartphone. Indeed, the solutions used today are mainly sensors such as stepped piezoelectric beams [1], [2] or strain gauges [3], which are expensive and require a delicate installation, not only because of the difficult access areas to be instrumented but also because of the quantity of cables and data to be processed. Another popular solution which is contact-less, is the laser vibrometer, but its price remains relatively high and does not produce an instantaneous measurement [4], [5]. The advantages of video camera analysis are numerous: for example, the "non-contact" aspect, which does not modify the response of the structure studied; the simplicity of installation with the need for little equipment; the fact that each pixel can be seen as a sensor, which allows a global study in a single acquisition etc.

Early work has proposed methods to reveal the vibrations animating various static objects using a fast camera [6]. This work is based on image amplification techniques that combine the small fluctuations simultaneously observed by several pixels. Other work shows that it is possible to extract mechanical properties (modal analysis) from "static" objects using high-speed cameras [7] and presents video analysis as a potential competitor to traditional measurement techniques.

The aim of this paper is to make a first link between a mechanical approach (operational modal analysis) and a signal and image processing approach by diverting the uses of known methods to use them to perform motion estimation on a cantilever beam in the quasi-static case.

We will start by presenting the two distinct methods we will use, namely the monogenic signal and the optical flow computed with a Lucas-Kanade type method. We will then study the different results obtained for each of the two methods and compare them with the results obtained by a traditional method, i.e. by using an accelerometer. Finally, we will look at small displacements by analysing the results obtained near the embedding.

2 Image processing methods

In this section we present two different methods for obtaining the operational deformation of the beam: a method based on the computation of the monogenic signal [8] [9] and another method based on the optical flow [10], following the Lucas-Kanade approach. Our global framework for motion estimation is given Figure 1.

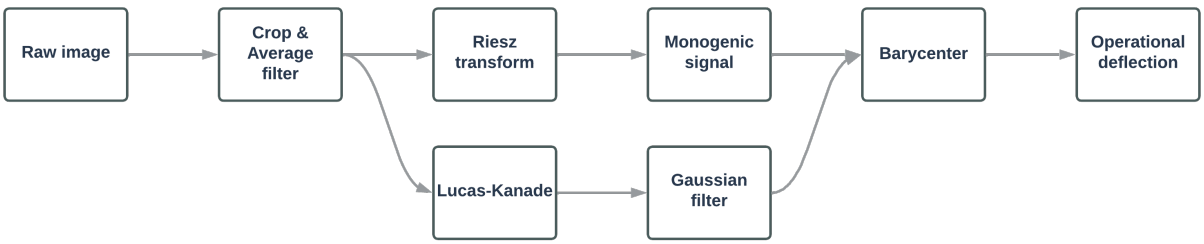

Figure 1: Processing steps

2.1 Position estimation of the cantilever based on monogenic signal

One of the reference methods for surface tracking is the DIC method (Digital Image Correlation), developed by Sutton et al. [11] (1983) and Bruck et al.[12] (1989), which has pros and cons. Indeed, it is an iterative method and therefore computationally intensive. Moreover, it requires the surface studied to be prepared beforehand in order to make a random pattern appear.

In this paper, we are interested in the motion estimation of a cantilever beam by studying the plane normal to the beam which is the one corresponding to the plane of the video. This means that we can see the beam as a line rather than a surface, which can justify not to use the DIC process.

In image processing, it is usual to pre-process images with different 2D filters in order to obtain an image that can be processed more efficiently and quickly later on. There are many "classic" pre-processing filters for contour tracking, such as the Canny filter [13], the Sobel filter [14], the Prewitt filter [15] or the Roberts filter [16]. In this study, we will use a filtering from the monogenic signal of each image, calculated using the Riesz transform.

The Riesz Transform

The Riesz transform, first defined by M. Riesz in 1927 [17], has seen its definition evolve considerably over time [18], [19] and [20] to become what it is today. The Riesz transform is quite similar to the gradient operator and can be seen from a signal processing point of view as a Hilbert transform in 2 dimensions. The Riesz transform $\mathcal{R}_f$ of the 2D signal f can be easily defined from its Fourier transform:

$$\forall \vec{\xi} \in \mathbb{R}^2, \widehat{\mathcal{R}}_f(\vec{\xi}) = -i\frac{\vec{\xi}}{|\vec{\xi}|}\hat{f}(\vec{\xi}) \tag{1}$$

where $\vec{\xi} = (\xi_x, \xi_y)^T$, $\hat{g}$ is the 2D Fourier transform of the 2D signal g and $i^2 = -1$.

This transform produces two 2D signals according to the spatial components x and y. The Riesz transform along the x axis corresponds to the real part and the Riesz transform along the y axis corresponds to the imaginary part of our image. The Riesz transform eliminates the DC component of the image and is used with a band-pass filter. We obtain, with $\vec{q} = (x, y)^T$:

$$\mathcal{R}_f(\vec{q}) = (\mathcal{R}_{fx}(\vec{q}), \mathcal{R}_{fy}(\vec{q}))^T \tag{2}$$

The norm and angle of the resulting signal can then be calculated:

$$\begin{cases} |\mathcal{R}_f(\vec{q})| = \sqrt{\mathcal{R}_{fx}(\vec{q})^2 + \mathcal{R}_{fy}(\vec{q})^2} \\ \theta(\vec{q}) = \arg\{\mathcal{R}_{fx}(\vec{q}) + i\mathcal{R}_{fy}(\vec{q})\} \end{cases} \tag{3}$$

The norm of the Riesz transform can be seen as a contour detector and its phase as a local measure of the contour orientation. An example of the results obtained for this transform is shown in Figure 2.

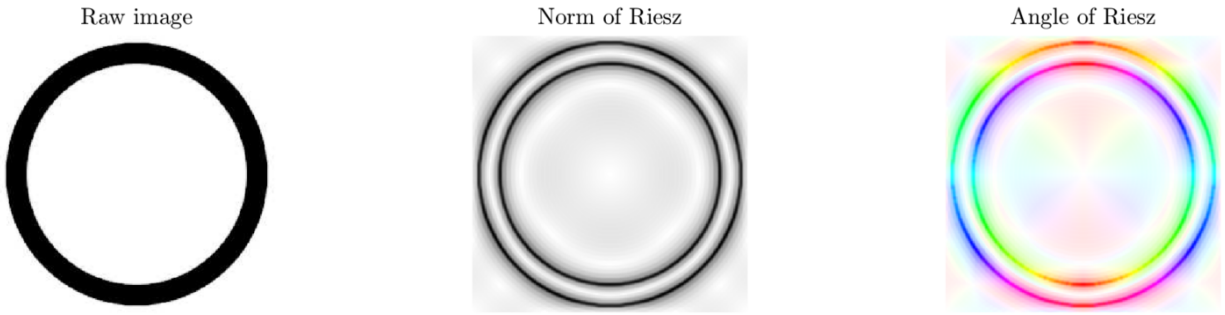

Figure 2: Example of the results obtained from the Riesz transform for a didactic image

The Monogenic Signal

The definition of the monogenic signal was proposed simultaneously by M. Felsberg [8] and O. Larkin [9] in 2001. It is a 2D extension of the analytical signal. The monogenic signal is defined by its amplitude A and its phase φ :

$$\begin{cases} A(\vec{q}) = \sqrt{f(\vec{q})^2 + |\mathcal{R}f(\vec{q})|^2} \\ \varphi(\vec{q}) = \arg\{f(\vec{q}) + i|\mathcal{R}f(\vec{q})|\} \end{cases} \tag{4}$$

The norm of the monogenic signal is an optimal contour detector and its phase is additional contour classification information. It is by definition a phase-invariant contour detector. In this paper, we will only use the norm of the monogenic signal as a contour detector to detect the contour and then determine the neutral line of the beam, and thus the position of each pixel on the beam at each instant. To do this, we perform a barycenter calculation, or statistical mean distribution, on each column of the image in order to obtain a sub-pixel estimate of the position of the neutral line of the studied beam (Section 3.2.3).

2.2 Optical flow motion estimation

Here we will use a method that is classically used when it comes to the study of optical flow: Lucas-Kanade. It is a differential method developed by Bruce D. Lucas and Takeo Kanade [21] which assumes that the displacement of a point from one image to another is small and approximately constant in a neighbourhood of that point. The local velocity vector (V_x, V_y) must then, for each pixel, satisfy the optical flow equation:

$$I_x V_x + I_y V_y = -I_t \tag{5}$$

With I_x, I_y and I_t the partial derivatives of the image I according to the variables of space x, y and time t. The Lucas-Kanade method proposes to solve the optical flow equation by the method of least squares considering the velocity of objects homogeneous by zone, which makes it possible to obtain the displacement field of all the points from one image to another by solving the equation for each pixel q_i around the central pixel of the studied zone:

$$V = \begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(q_i)^2 & \sum_i I_x(q_i) I_y(q_i) \\ \sum_i I_x(q_i) I_y(q_i) & \sum_i I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_x(q_i) I_t(q_i) \\ -\sum_i I_y(q_i) I_t(q_i) \end{bmatrix} \tag{6}$$

In order not to give the same importance to all the pixels q_i of the window and to favour the pixels closest to the pixel under study, we weight the least squares equation with weights w_i, following a spatial Gaussian window centered on the studied pixel. We must then solve the equation :

$$V = \begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i w_i I_x(q_i)^2 & \sum_i w_i I_x(q_i) I_y(q_i) \\ \sum_i w_i I_x(q_i) I_y(q_i) & \sum_i w_i I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i w_i I_x(q_i) I_t(q_i) \\ -\sum_i w_i I_y(q_i) I_t(q_i) \end{bmatrix} \tag{7}$$

We will study the norm of this displacement field:

$$|V| = \sqrt{V_x^2 + V_y^2} \tag{8}$$

In concrete terms, this matrix gives us information about each pixel and whether it has moved a lot or a little from one image to the next. We then make the assumption that the set of points that have moved the most correspond to the points of the beam, the rest being considered as noise beyond a pre-established threshold. Once the thresholded matrix is obtained, we apply a horizontal Gaussian filter to smooth and assume continuity in the direction of the beam. Then, like the monogenic signal, we will proceed to a barycenter calculation on each column of the image and thus obtain what is assumed to be the centre of the beam (Section 3.2.3).

3 Experimental results

3.1 The cantilever beam

In this study, we are interested in a cantilever beam (Figure 3) on which an accelerometer and two piezoelectric patches are glued back to back near the clamping: one "transmitter" patch whose role is to excite the beam and one "sensor" patch whose role is to measure this deformation. The beam is excited at a frequency

$f_{excit} = 6.38Hz$, which corresponds to its first vibration mode previously determined. The video acquisition is carried out using a "fast" camera with a frequency of 2000 frames per second.

Figure 3: Studied cantilever beam

In order to obtain a satisfactory contrast, the beam was painted white and a black background was used. A high brightness is necessary to obtain a good contrast on each image because of the speed of acquisition of the camera.

3.2 Pre-processing

First, we do a grey scale crop acquisition, which allows us to divide by more than 3 the amount of data received and we can directly use the pre-processing methods studied previously. A filtering of the images using an average filter of kernel

$$h = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{9}$$

is then applied. The purpose of this filter is to smooth the acquired image and avoid problems that may arise due to the Bayer matrix of the camera's CMOS sensor.

3.2.1 Monogenic Signal

The acquired are then pre-processed, using the methods studied previously and after a second more precise crop, we obtain the norm of the Riesz transform as well as the norm of the monogenic signal, Figure 4. The band-pass filter of the monogenic signal must be carefully defined in order to see, depending on the size of the filter, the beam as a line or as a contour of the beam. In the case of a line, the monogenic signal gives us a single contour, which corresponds to the center of the beam, whereas in the case of a contour, we will have the upper and lower contours. The norm of the monogenic signal obtained is then, by definition, well centred on the contours. We choose in this study to study the case of the two contours.

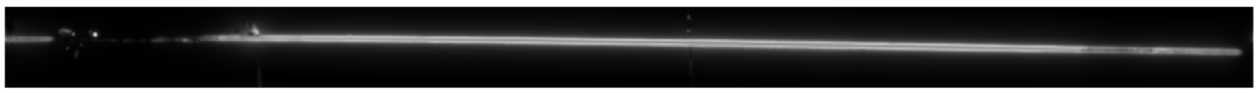

Figure 4: Norm of monogenic signal

3.2.2 Lucas-Kanade optical flow

The study of the optical flow is done directly on the filtered image rather than the Riesz transform norm or the monogenic signal norm because the results obtained were no significantly better. We then obtain, after horizontal Gaussian filtering, the norm of the velocity matrix $|V|$, plotted on Figure 5.

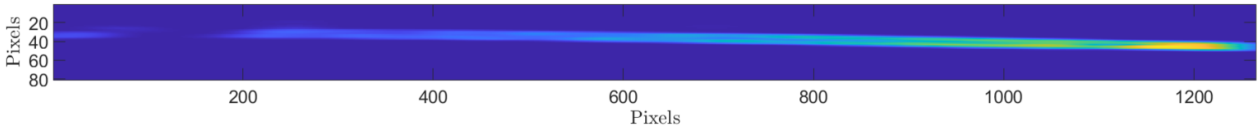

Figure 5: Norm of filtered speed matrix

In our case, the motion is almost completely vertical, so we get the velocity field matrix:

$$|V| = \sqrt{V_x^2 + V_y^2} \approx |V_y| \tag{10}$$

3.2.3 Barycenter calculations

After having obtained the norm of the monogenic signal and the norm of the velocity matrix from the optical flow method, we can estimate the position of the beam at each instant using the barycenter. We obtain the barycenter for a vector $x = (x_1, ..., x_n)$ by associating it with the corresponding sample vector: $y = (1, ..., n)$. The barycenter G is then :

$$G = \frac{\sum_i x_i y_i}{\sum_i x_i} \tag{11}$$

We obtain, for one image, the positions given Figure 6. There is an estimation error at the beginning of the beam as the area is less illuminated and this is where the piezoelectric patches are glued. The accelerometer is located around the 250 pixels.

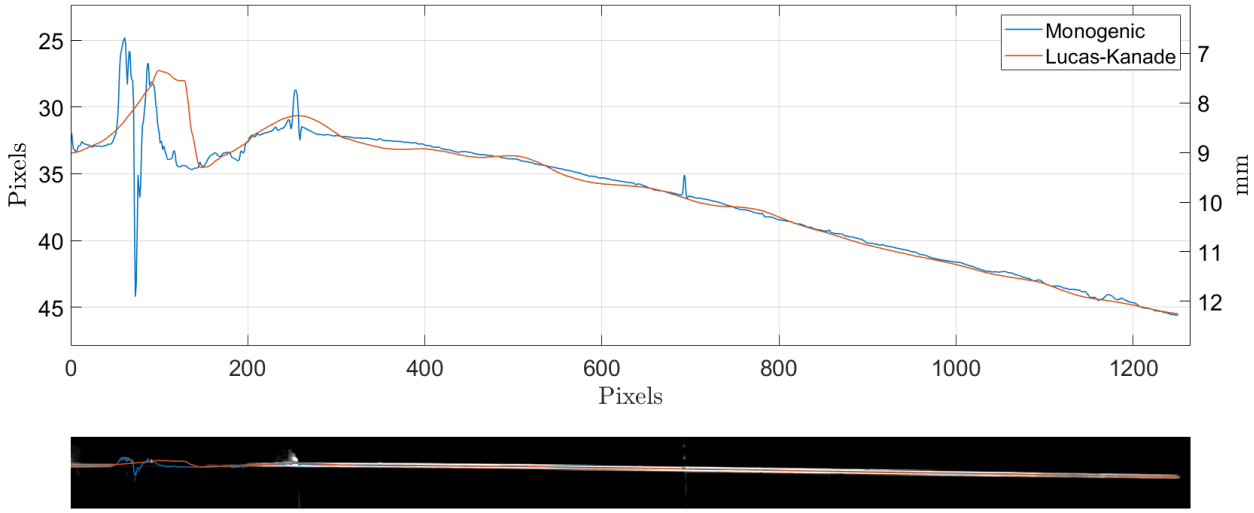

Figure 6: Results for one image

3.3 Comparison of the results obtained with the different methods

Operating Deflection Shape

We obtain the position of the neutral fibre of the beam by following the two methods presented in the section Optical flow motion estimation. We then have the position of each point of the beam at each frame. In the following, the results obtained with the video methods with those obtained with the accelerometer are compared. We consider here that the results provided by the accelerometer give us an accurate ground truth. We will study the displacement at the accelerometer, which corresponds to a defined area of 11 pixels in the video. These 11 pixels are then averaged to obtain an image of the displacement at the accelerometer level as seen by the camera. The results obtained are shown in Figure 7 and the corresponding spectrums are shown in Figure 8. The spectrum is plotted in the frequency range [0-30] Hz in order to focus only on what is happening around the excited eigenmode and its potential harmonics. This spectrum is obtained by performing a direct Fourier transform without any windowing. All the signals do not have the same acquisition time (5 seconds for the video methods and 60 seconds for the accelerometer), we divide each FFT by the number of points of each signal.

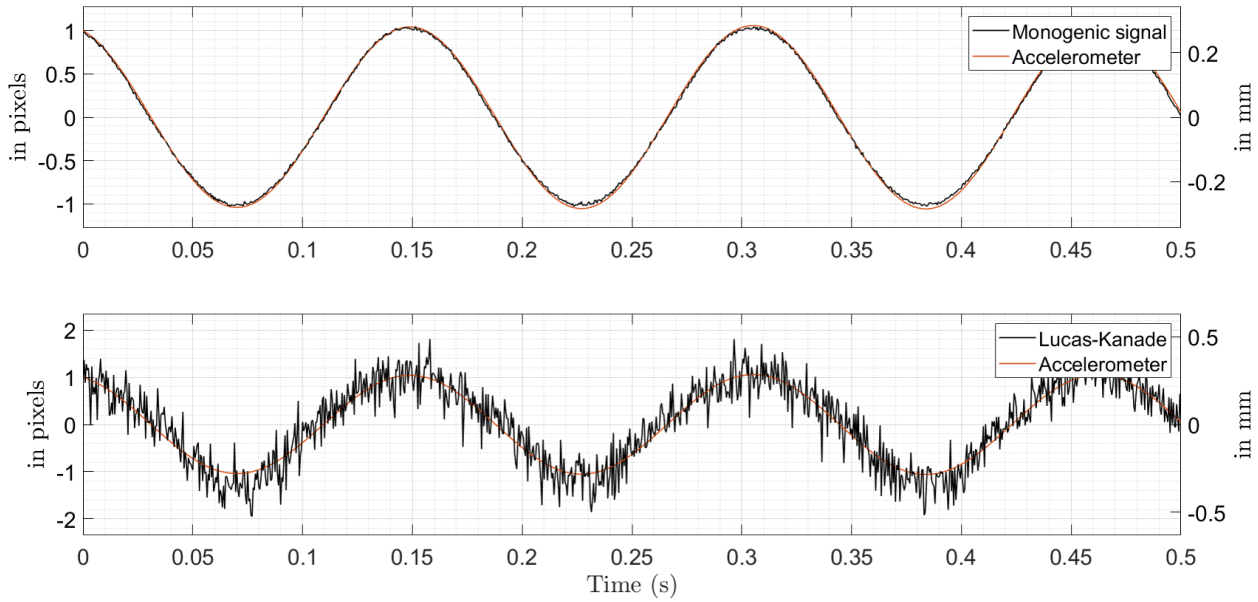

Figure 7: Motion estimation at the accelerometer

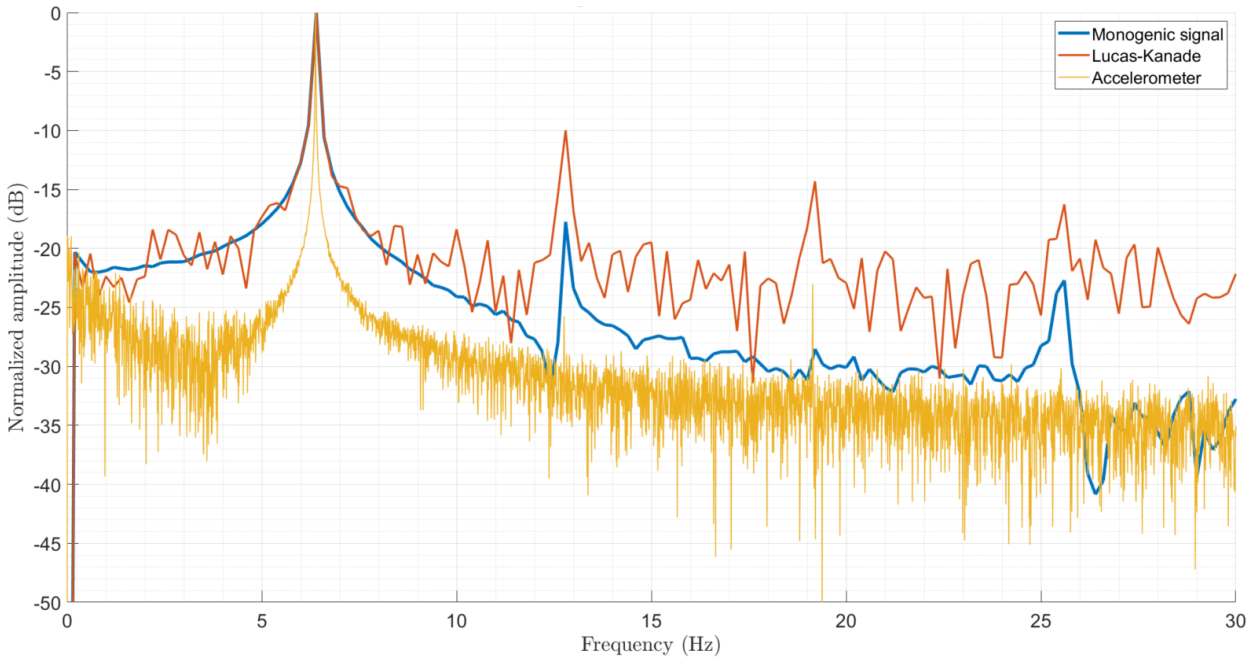

Figure 8: Normalized spectrum at the accelerometer

The sinusoidal trend is clearly observed for both methods around the excitation frequency of the first mode at 6.38Hz. The first harmonics of this first mode are also observed for both methods. Wide-band noise is observed on each signal and the signal-to-noise ratio appears to be much better in the case of the monogenic signal compared to the Lucas-Kanade case. To see the influence of the video acquisition frequency on the results, the SNR between the accelerometer and the two video methods is calculated using the following calculation:

$$SNR = \frac{A^2}{\frac{1}{N-1} \cdot \sum (x_{acc} - x_{vid})^2} \tag{12}$$

with :

- A: the amplitude of the fundamental of the accelerometric signal,
- N: the number of samples in x_{acc} and x_{vid}.

The results are shown in Figure 9 in log-log scale. Two different trends can be observed for the two curves: In the case of the monogenic signal, the SNR will increase as the video sampling rate is increased. In the Lucas-Kanade case, a local maximum is observed around 100 Hz. Before this value, the Lucas-Kanade SNR is better than that of the monogenic signal. Beyond this value, we observe a slight decrease to an SNR around 21dB for higher sampling frequencies. These results are still being analysed but we assume that this is because Lucas-Kanade assumes a movement that is neither too small nor too large between two successive images. It can be seen that the method based on a Lucas-Kanade type approach performs better than the monogenic signal when the acquisition frequency is low (below 140 Hz) but the relative difference between the 2 methods remains quite small.

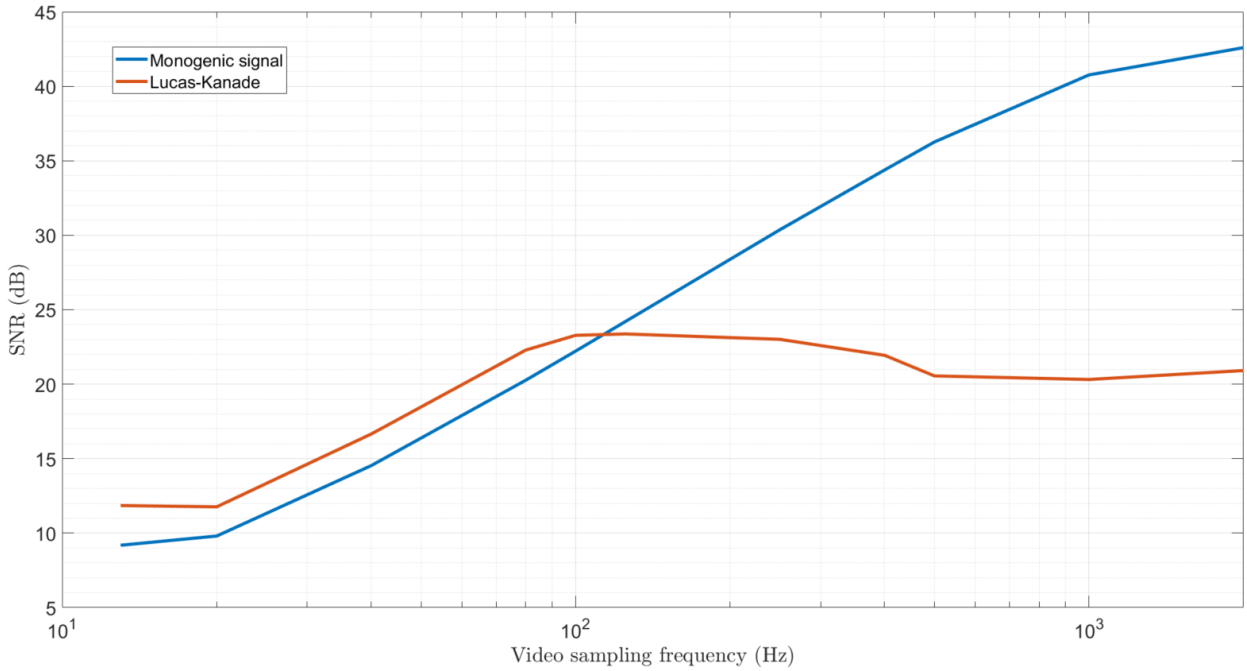

Figure 9: Signal to Noise Ratio

Displacement close to clamping

To check if a method allows us to monitor small amplitudes without being too sensitive to noise, we are interested in the results obtained when we track the pixels that are close to the embedding and therefore have the smallest displacement amplitude (no displacement in theory with a perfect installation). The results obtained are presented in Figure 10 and Figure 11. The monogenic signal appears to provide a pseudo-sine

signal, whereas Lucas-Kanade does not appear to provide a usable result in the time domain. There is also an amplitude ratio of 2 between the two curves in the time domain and a high level of noise is also observed for Lucas-Kanade in the frequency domain. The peak corresponding to the excitation frequency of the first mode at $f_{excit} = 6.38Hz$ is observed for both methods. It seems that the use of the monogenic signal is more robust to small displacements than Lucas-Kanade. However, the use of Lucas-Kanade could be useful in the case where the background of the video is not so homogeneous because the background will be considered as noise during processing.

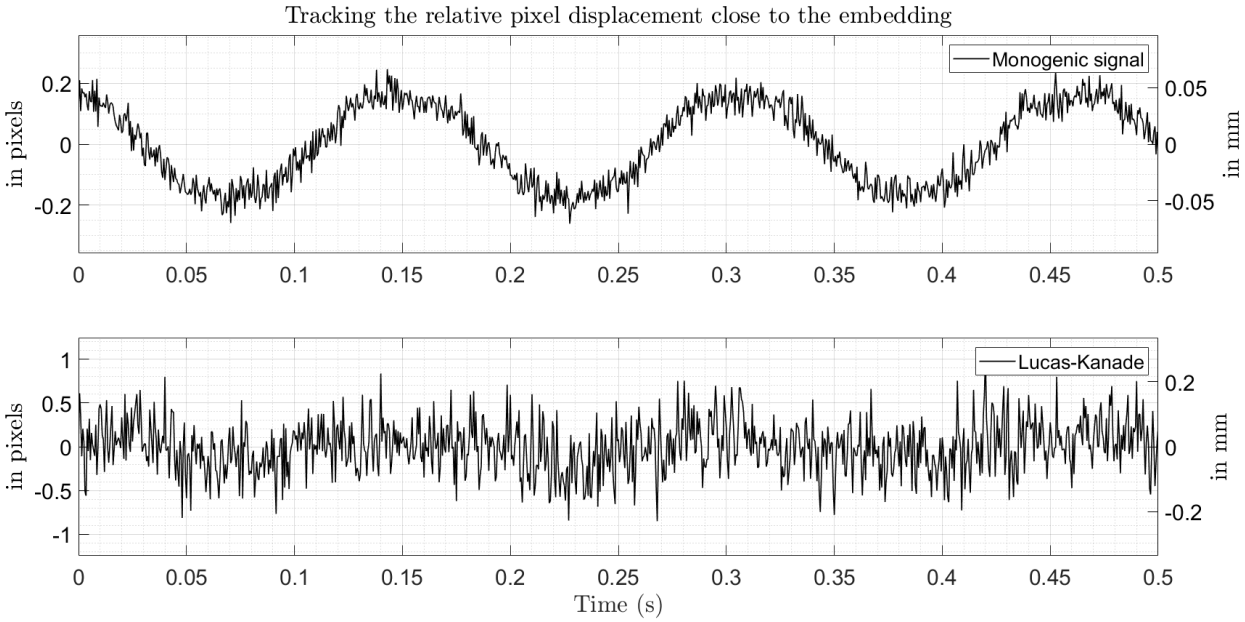

Figure 10: Motion estimation near the embedding

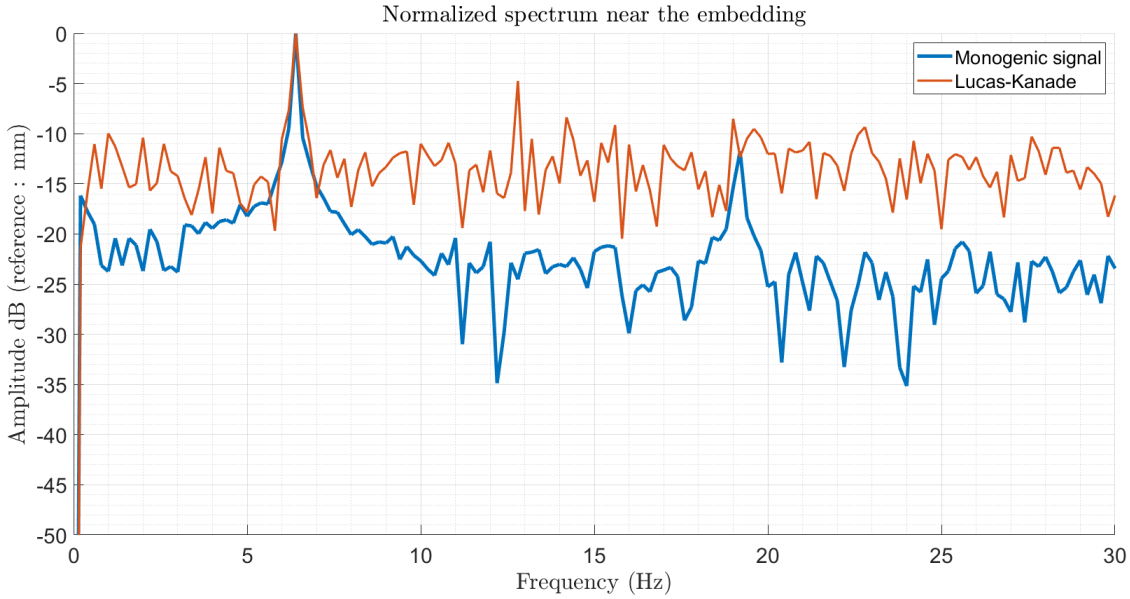

Figure 11: Spectrum near the embedding

4 Conclusion

In this paper we have seen how to perform motion estimation using two completely different approaches. We have compared these two methods with each other and with a proven reference: the accelerometer. We made a temporal and spatial analysis of the results obtained from these two methods. The conclusion is that the monogenic signal seems preferable in the case of a free embedded beam in the quasi-static case with a homogeneous background. This may no longer be the case if only one of these parameters changes. The upcoming objectives of this work are to identify the conditions that favour the use of one method over the other, such as the straightness of the contour, the homogeneity of the background, the displacement in a direction that is not only vertical.

References

[1] S. Yang and Y. Lee, "Modal analysis of stepped beams with piezoelectric materials," *Journal of Sound and Vibration*, vol. 176, no. 3, pp. 289–300, 1994. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022460X84713770

[2] C. Maurini, M. Porfiri, and J. Pouget, "Numerical methods for modal analysis of stepped piezoelectric beams," *Journal of Sound and Vibration*, vol. 298, no. 4, pp. 918–933, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022460X06004986

[3] L. Yam, T. Leung, D. Li, and K. Xue, "Theoretical and experimental study of modal strain analysis," *Journal of Sound and Vibration*, vol. 191, no. 2, pp. 251–260, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022460X96901194

[4] A. STANBRIDGE and D. EWINS, "Modal testing using a scanning laser doppler vibrometer," *Mechanical Systems and Signal Processing*, vol. 13, no. 2, pp. 255–270, 1999. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327098912092

[5] D. A. Ehrhardt, M. S. Allen, S. Yang, and T. J. Beberniss, "Full-field linear and nonlinear measurements using continuous-scan laser doppler vibrometry and high speed three-dimensional digital image correlation," *Mechanical Systems and Signal Processing*, vol. 86, pp. 82–97, 2017, full-field, non-contact vibration measurement methods: comparisons and applications. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S088832701500552X

[6] H. André, Q. Leclère, D. Anastasio, Y. Benaïcha, K. Billon, M. Birem, F. Bonnardot, Z. Chin, F. Combet, P. Daems, A. Daga, R. De Geest, B. Elyousfi, J. Griffaton, K. Gryllias, Y. Hawwari, J. Helsen, F. Lacaze, L. Laroche, X. Li, C. Liu, A. Mauricio, A. Melot, A. Ompusunggu, G. Paillot, S. Passos, C. Peeters, M. Perez, J. Qi, E. Sierra-Alonso, W. Smith, and X. Thomas, "Using a smartphone camera to analyse rotating and vibrating systems: Feedback on the survishno 2019 contest," *Mechanical Systems and Signal Processing*, vol. 154, p. 107553, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327020309390

[7] J. Javh, J. Slavič, and M. Boltežar, "High frequency modal identification on noisy high-speed camera data," *Mechanical Systems and Signal Processing*, vol. 98, pp. 344–351, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327017302637

[8] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, 2001.

[9] K. Larkin, D. Bone, and M. Oldfield, "Natural demodulation of two-dimensional fringe patterns. i. general background of the spiral phase quadrature transform," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 18, pp. 1862–70, 09 2001.

[10] J. J. Gibson, *The Perception Of The Visual World*. Boston: Houghton Mifflin, 1950.

[11] M. Sutton, W. Wolters, W. Peters, W. Ranson, and S. McNeill, "Determination of displacements using an improved digital correlation method," *Image and Vision Computing*, vol. 1, no. 3, pp. 133–139, 1983. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0262885683900641

[12] H. A. Bruck, S. R. McNeill, M. A. Sutton, and W. H. Peters, "Digital image correlation using newton-raphson method of partial differential correction," *Experimental Mechanics*, vol. 29, no. 3, pp. 261–267, Sep 1989. [Online]. Available: https://doi.org/10.1007/BF02321405

[13] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

[14] I. Sobel, "An isotropic 3x3 image gradient operator," *Presentation at Stanford A.I. Project 1968*, 02 2014.

[15] P. J. M. S., "Object enhancement and extraction," *Picture Processing and. Psychopictorics*, 1970. [Online]. Available: https://cir.nii.ac.jp/crid/1572824500045263872

[16] L. G. Roberts, "Machine perception of three-dimensional solids," Ph.D. dissertation, Massachusetts Institute of Technology, 1963.

[17] M. Riesz, "Sur les fonctions conjuguées," *Mathematische Zeitschrift*, vol. 27, no. 1, pp. 218–244, Dec 1928. [Online]. Available: https://doi.org/10.1007/BF01171098

[18] A. P. Calderon and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88, no. none, pp. 85 – 139, 1952. [Online]. Available: https://doi.org/10.1007/BF02392130

[19] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*. Princeton, N.J. : Princeton University Press, 1971.

[20] M. N. Nabighian, "Toward a three-dimensional automatic interpretation of potential field data via generalized Hilbert transforms; fundamental relations," *Geophysics*, vol. 49, no. 6, pp. 780–786, 06 1984. [Online]. Available: https://doi.org/10.1190/1.1441706

[21] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision (ijcai)," vol. 81, 04 1981.