

# The Subpixel Resolution of Optical-Flow-Based Modal Analysis

Jaka Javh<sup>a</sup>, Janko Slavič<sup>a,\*</sup>, Miha Boltežar<sup>a</sup>

<sup>a</sup>*University of Ljubljana, Faculty of Mechanical Engineering, Aškerčeva 6, 1000 Ljubljana, Slovenia*

---

Cite as:

Jaka Javh, Janko Slavič and Miha Boltežar  
The Subpixel Resolution of Optical-Flow-Based Modal Analysis  
Mechanical Systems and Signal Processing, Vol. 88, p. 89–99, 2017  
DOI: 10.1016/j.ymsp.2016.11.009

## Abstract

This research looks at the possibilities for full-field, non-contact, displacement measurements based on high-speed video analyses. A simplified gradient-based optical flow method, optimised for subpixel harmonic displacements, is used to predict the resolution potential. The simplification assumes a constant image-gradient, producing a linear relation between the light intensity and the displacement in the direction of the intensity gradient. The simplicity of the method enables each pixel or small subset to be viewed as a sensor. The resolution potential and the effect of noise are explored theoretically and tested in a synthetic experiment, which is followed by a real experiment. The identified displacement can be smaller than a thousandth of a pixel and subpixel displacements are recognisable, even with a high image noise. The resolution and the signal-to-noise ratio are influenced by the dynamic range of the camera, the subset size and the sampling length. Real-world experiments were performed to validate and demonstrate the method using a monochrome high-speed camera. One-dimensional mode shapes of a steel beam are recognisable even at the maximum displacement amplitude

---

\*Corresponding author

*Email address:* `janko.slavic@fs.uni-lj.si` (Janko Slavič)

of 0.0008 pixel (equal to  $0.2 \mu\text{m}$ ) and multiple out-of-plane mode shapes are recognisable from the high-speed video of a vibrating cymbal.

*Keywords:* gradient-based optical flow, modal analysis, full-field non-contact displacement measurement, subpixel resolution, operational displacement shapes, photogrammetry

---

## 1. Introduction

The vibrations associated with structures are commonly measured with piezoelectric accelerometers. These accelerometers are attached to the structure and measure the acceleration from the pressure that a seismic mass generates on the piezoelectric. Accelerometers are precise, but their added mass can change the structure's dynamics. However, non-contact measurement methods are often preferred, as they do not influence the structure's dynamics. Most non-contact measurement systems use monochromatic light and the Doppler effect or interferometry. Laser Doppler vibrometry uses the Doppler effect to measure the speed of the object's vibrating surface. A laser vibrometer has good resolution, but measures only one point at a time and only the speed in the direction of the laser beam. To measure a full-field response the measured point has to be moved and successive measurements performed, much as with accelerometers. Scanning systems are used to scan the whole surface and measure the full field [1]. Interferometric systems such as electronic speckle pattern interferometry (ESPI) [2] and holographic interferometry [3] have already been used to measure vibrations. These methods determine the out-of-plane displacements based on the interference fringes of a reference light and the light reflected from the surface. Since the whole surface is illuminated, a full-field, out-of-plane response is measured. Also, stress mode shapes have previously been successfully measured via thermoelasticity and filming with a high-speed thermal camera [4]

Full-field optical techniques using high-speed cameras are increasingly being used in structural engineering as a means of measuring displacements and have recently made their way into modal analysis [5][6]. One study even showed that sound can be reconstructed from the structural vibrations of an object filmed using a high-speed camera [7].

The most common image-processing technique used to determine full-field displacements in structural engineering is called digital image correlation (DIC) [8]. DIC is based on calculating the correlation between subsets

in sequential images (also known as template matching). It determines the displacements in an image by calculating the locations with the highest correlation [8].

Subpixel registration can be achieved by curve-fitting the integer pixel correlation coefficients and using the curve-fitted peak position as the subpixel displacement. Such a subpixel registration combined with an integer pixel correlation calculation in the Fourier domain [9] is relatively fast, but it does not offer the best accuracy [10]. The subpixel displacements can alternatively be determined iteratively by using a root-finding method (e.g., Newton-Raphson), which works by interpolating the subset and calculating the correlation until the desired convergence is achieved. The Newton-Raphson produces an increased accuracy at the expense of the processing time [10]. An inverse composition Gauss-Newton algorithm (introduced by [11]) produces a matching accuracy to the Newton-Raphson at a higher computational efficiency [12]. The DIC technique is well established in engineering, because it provides good results and a high (subpixel) displacement resolution, typically quoted at a hundredth of a pixel [13].

A phase-based optical flow method [14] has also been successfully used in full-field modal analysis [15]. Phase-based optical flow determines the displacements by tracking the constant phase contours of the images. Local phase and amplitude information are determined by filtering an image with different filters.

Gradient-based (also known as differential) optical flow approaches [14] such as the Lucas-Kanade method [16] and the Horn-Schunck method [17] are fast and often used in real-time operations [18]. Gradient-based optical flow methods are based on the optical flow equation [17], and hence on the intensity gradient determining the velocity from the difference in the pixel intensity between sequential images [19]. The optical flow equation is an undetermined system and can only provide information about the velocity in the direction of the intensity gradient (the aperture problem [20]). This is usually solved with information from the surrounding pixels using smoothness constraints [17], by imposing a least squares condition on a subset [21], or with multiple constraints such as RGB colours [22].

Many studies have previously analysed various errors of optical flow methods (e.g., bias and effect of noise [10, 12, 21, 23]). Since the first implementations of optical flow [16] it is known that high contrasts are required for more accurate results and various research was done to determine the adequacy of the patterns painted on the observed surface [24, 25]. Despite

the extensive research an overview of the theoretical resolution potential is lacking. This study analyses the displacement resolution potential on a simplified gradient-based optical flow method for small harmonic displacements expected in the modal analysis. Individual pixels or small subsets can act as individual one-degree-of-motion displacement sensors, providing a dense full-field.

The text is organized as follows: Section 2 covers the basics of the gradient-based optical flow approach. Section 3 discusses the maximum resolution potential. Section 4 uses the displacement amplitudes to produce the operational shapes. Section 5 tests the method in a synthetic experiment at the resolution limit and with high noise. A validation with a piezoelectric accelerometer is presented in Section 6, followed by full-field experiments in Section 7.

## 2. Optical flow in the direction of the intensity field gradient

A gray image is a 2D matrix of  $m \times n$  pixels containing light-intensity values. The discrete 2D matrix can be interpreted as a 2D function of the intensity values over the camera’s projection coordinates  $x$  and  $y$ , and is denoted as  $I(x, y)$ . The intensity field is a feature of the object’s surface. When the object moves, the intensity field in the camera’s projection  $x, y$  shifts, thereby changing the pixel intensities (Figure 1). The object moves with time; therefore, the intensity values are time  $t$  dependent  $I(x, y, t)$ .

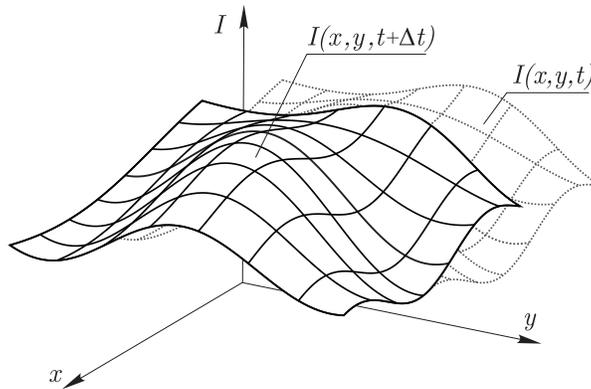


Figure 1: An example of a light-intensity field as it shifts

Optical flow assumes that there is brightness conservation [19]. Furthermore, the light-intensity of a point  $(x_j, y_k)$  on the object that moved to

$(x_j + \Delta x, y_k + \Delta y)$  is assumed to be constant:

$$I(x_j, y_k, t) = I(x_j + \Delta x, y_k + \Delta y, t + \Delta t) \quad (1)$$

Most gradient-based optical flow methods approximate the image-intensity function with a Taylor series as:

$$\begin{aligned} I(x_j + \Delta x, y_k + \Delta y, t + \Delta t) &= I(x_j, y_k, t) + \frac{\partial I}{\partial x} \Delta x \\ &+ \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t + \dots, \end{aligned} \quad (2)$$

where  $\partial I/\partial x$ ,  $\partial I/\partial y$  are the derivatives in their individual directions  $x$ ,  $y$ , and  $\partial I/\partial t$  is the derivative of time. For a small deformation and short time intervals the higher-order terms in Equation (2) can be neglected. Taking into account the optical flow condition (1), Equation (2) takes a new form as the optical flow equation [16][17]:

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (3)$$

For a discrete problem  $\partial I/\partial t \Delta t$  is interpreted as the change in the intensity at a pixel between sequential pictures:

$$\frac{\partial I}{\partial t} \Delta t = I(x_j, y_k, t + \Delta t) - I(x_j, y_k, t) \quad (4)$$

Equation (3) can now be written as:

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y = I(x_j, y_k, t) - I(x_j, y_k, t + \Delta t) \quad (5)$$

Since the object's intensity field  $I$  is known, the gradients  $\partial I/\partial x$  and  $\partial I/\partial y$  can be determined. The resulting Equation (5) has two unknowns,  $\Delta x$ ,  $\Delta y$ , and therefore cannot be solved (because of the aperture problem). The aperture problem is usually solved using smoothness constraints [16][17] that employ information from the other pixels, some of which are expected to have a different gradient direction, thereby providing displacement information in the  $x$  and  $y$  directions. The proposed method limits the observation to displacements in the direction of the intensity gradient, only. Consequently, Equation (5) is rewritten here in the direction of the gradient [16]:

$$|\nabla I| \Delta s = I(x_j, y_k, t) - I(x_j, y_k, t + \Delta t) \quad (6)$$

$|\nabla I|$  is the known scalar value of the intensity gradient and is calculated from the partial derivatives of the first image:

$$|\nabla I| = \sqrt{\frac{\partial I^2}{\partial x} + \frac{\partial I^2}{\partial y}} \quad (7)$$

$\Delta s$  is the unknown scalar value of the displacement in the direction of the intensity gradient. The partial derivatives  $\partial I/\partial x$ ,  $\partial I/\partial y$ , are calculated numerically for every pixel based on the surrounding pixels using a gradient filter (Sobel operator [26], Prewit operator [27] or any other normalized gradient filter).

Figure 2 can be used to help visualise the relation written in Equation (6). The figure shows a small displacement of the object producing a intensity-field shift. A point on the object carrying the intensity field moves from point 1 to point 1' and the intensity intensity in the camera pixel  $(x_j, y_k)$  changes from point 1 to point 2.

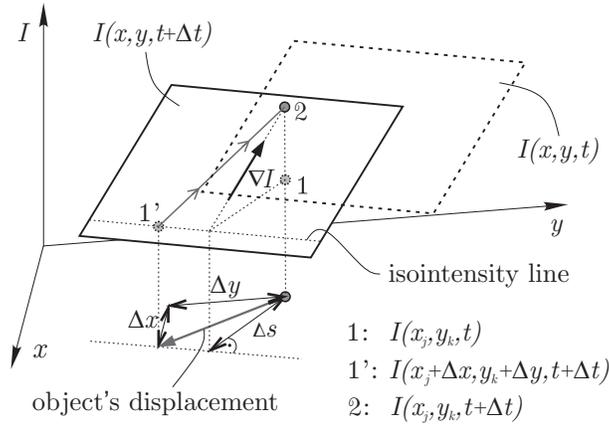


Figure 2: Displacement of the intensity field and the intensity change at the selected pixel  $(x_j, y_k)$

Normally, the video would be analysed frame by frame, determining the gradient  $|\nabla I|$  at each step and deducing the displacements between the first image  $I(x, y, t)$  and the image that follows  $I(x, y, t + \Delta t)$ . This step would then be repeated for every subsequent pair of images. In contrast, the proposed method assumes small (subpixel) displacements that stay within the area of an approximately constant intensity gradient. The displacement  $s$  can be observed in terms of the initial (reference) image, by monitoring the

pixel intensity between the reference image  $I_0(x, y)$  and subsequent sequential images  $I(x, y, t)$ :

$$s(x_j, y_k, t) = \frac{I_0(x_j, y_k) - I(x_j, y_k, t)}{|\nabla I_0|}, \quad (8)$$

where  $\nabla I_0$  is the intensity gradient of the reference image and  $s(x_j, y_k, t)$  is the displacement between the reference and the current image.

Equation (8) provides a linear relation between the displacement and the intensity level. The change in the intensity of a selected pixel is proportional to the displacement  $s$  and the intensity gradient  $|\nabla I_0|$  in that pixel.

The gradient-based optical flow is sensitive to noise. This is mainly due to the gradient being calculated on noisy images. The proposed method uses the gradient of a reference image. This reference image  $I_0$  is best obtained by averaging the pixel intensities over time, thereby generating an average-value image that is less affected by the noise. Such averaging is possible, as small harmonic displacements do not change the average intensity level of a pixel.

Assuming small displacements, the displacements are not expected to exceed the area of the constant gradient; however, in the case of larger displacements (in the range of a few pixels), pixel shifting can be used to stay in the range of the constant gradient:

$$s(x_j, y_k, t) = \frac{I_0(x_j, y_k) - I(x_j + \Delta x_L, y_k + \Delta y_L, t)}{|\nabla I_0|} \pm \sqrt{\Delta x_L^2 + \Delta y_L^2}, \quad (9)$$

where the sign in front of the root is determined by the gradient direction and  $\Delta x_L$   $\Delta y_L$  are integer displacements:

$$\Delta x_L = \text{round}\left(\frac{\frac{\partial I}{\partial x}}{|\nabla I_0|} s(x_j, y_k, t - \Delta t)\right) \quad (10)$$

$$\Delta y_L = \text{round}\left(\frac{\frac{\partial I}{\partial y}}{|\nabla I_0|} s(x_j, y_k, t - \Delta t)\right) \quad (11)$$

### 3. Theoretical displacement-resolution potential

Structural vibrations usually exhibit small displacements that are barely visible, if at all. This is especially true for higher frequency vibrations. A good displacement resolution is therefore required. This section takes a closer look at the achievable pixel-displacement resolution of the suggested method.

### 3.1. Single-pixel resolution potential

When filming with an 8-bit intensity depth camera, a line with the values 0 (gray-scale black) on the left-hand side and 255 (gray-scale white) on the right-hand side, passing over a single pixel, will produce a steady increase in the intensity values in this pixel from 0 to 255 over all 255 intensities, giving a  $1/255$  or  $\approx 0.004$  pixel-displacement resolution. This concept is shown in Figure 3. The resolution potential when observing one pixel is therefore a function of the intensity depth in bits  $n$ :

$$\Delta x_{\min} = \frac{1}{(2^n - 1)} \quad (12)$$

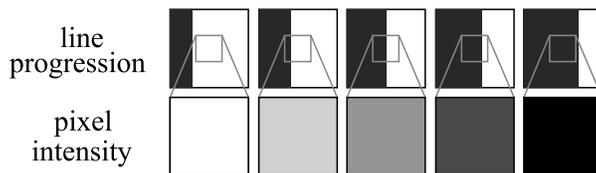


Figure 3: The upper fields show the progression of the colour line over the pixel; the lower fields show the registered intensity in the pixel

### 3.2. Subset-resolution potential

However, the potential is even greater when observing a larger subset with a size of  $d \times d$  pixels due to different pixels changing their values at different displacements. This can be demonstrated when the line moves over the observed subset and the individual pixels change their values independently. Since the line is moving from left to right, the values in the pixels can only increase. From left to right the line crosses  $d$  pixels. During the crossing  $(2^n - 1)d^2$  states are theoretically observed. For a  $n = 2$  bit depth and a subset size of  $d = 3$  this produces 9 observable states for a one-pixel move, see Figure 4. The resolution potential when observing a subset is therefore:

$$\Delta x_{\min} = \frac{1}{(2^n - 1)d} \quad (13)$$

The developed method calculates the sum of the subset, thereby accounting for all  $(2^n - 1)d$  states. The subset sum is calculated for every pixel,

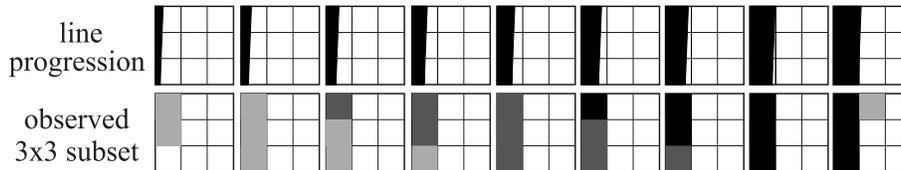


Figure 4: All 9 states observed in a  $3 \times 3$  subset during the one-pixel move of a line at an angle for a 2-bit gray-scale image

producing a filtered image  $I_f$ :

$$I_f(x_j, y_k, t) = \sum_{p=-q}^q \sum_{r=-q}^q I(x_{j+p}, y_{k+r}, t), \quad (14)$$

where  $q = (d - 1)/2$  and the subset size  $d$  is an odd number. Such filtering smooths the image and thereby increases the area of constant gradient. The filtered image  $I_f$  replaces  $I$  in Equation (8) to produce displacements. The reference image is now  $I_{f0}$  instead of  $I_0$  and is still calculated as the average of the intensities over time, only this time the filtered image sequence  $I_f$  is used in the averaging over time.

### 3.3. Effect of noise on the resolution

Since the light intensity linearly relates to the displacement, the camera noise maps to the displacement noise. The resolution potentials from 3.1 and 3.2 are only possible when no noise is present; however, cameras usually produce very noisy measurements, therefore the frame-to-frame displacement resolution for optical flow measurements is typically in the range of a hundredth of a pixel [13].

Better resolution can be achieved when using a larger subset, as the subset displacement is determined based on the change of intensities of all the pixels in the subset. The best results are expected when all the subset pixels have a large intensity gradient, in fact Pan, *et al* [24] propose a parameter to quantify the optimal subset size based on the Sum of Square of Subset Intensity Gradients (SSSIG). A larger subset displacement is analogous to averaging the neighbouring displacement data and therefore reducing the spatial resolution of the measurement.

## 4. Operational displacement shapes

A modal analysis is usually performed in the frequency domain. A discrete Fourier transform of the pixel-displacement time series (9) produces complex displacement amplitudes:

$$S(x_j, y_k, f) = FFT(s(x_j, y_k, t)) \quad (15)$$

The displacement amplitudes for a given frequency produce an operational deflection shape. An operational shape at the resonant frequency is a close match to the mode shape of the structure [28].

### 4.1. Displacement amplitude resolution

The discrete Fourier transform enables the detection of signals even below the discrete theoretical displacement-resolution potential from Section 3. A signal below the discrete resolution can be detected if it is superimposed on a carrier, where a carrier can be another signal or even noise. The carrier signal helps the weak signal to cross the quantization level. If the weak signal is coherent and the noise is not, the amplitude noise level will decrease with the square root of the sample size. The described effect is known as noise dithering [29].

In the given case of this study the weak signal is a displacement smaller than the resolution limit from Equation (13) and the sample size is the length of the displacement time series, which is equal to the number of captured images.

Provided that the intensity gradient is optimal, the resolution is therefore mainly limited by the noise level. To reduce the noise level, a large number of captured images is preferred.

## 5. Synthetic experiments

The following are synthetic experiments that demonstrate the high-resolution potential and noise robustness. Three simulations were performed: (a) a simple reconstruction, (b) a high-noise test and (c) a high-resolution test.

### 5.1. An optimal pattern

The method works with a random texture, but gives better results with patterns of contrasting intensities. A pattern of black and white lines was used (Figure 5), because such a pattern has the maximum contrast on the

edges of the lines and is easily reproduced in the real world. The limitation of such a pattern is that it produces visible displacements on the line edges, only.

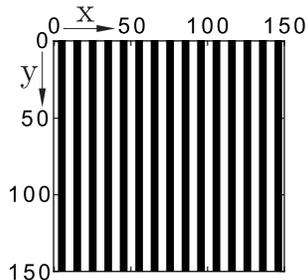


Figure 5: A pattern of black and white lines

The simple reconstruction (a) and the high-resolution test (c) were performed on a pattern that had a dynamic range of 250 intensity values and a size of  $150 \times 150$  pixels. The high-noise test (b) was performed on a pattern with a dynamic range of only 125 intensity values to accommodate the higher noise.

### 5.2. Simulating displacements

Every pixel was assigned a harmonic in-plane displacement  $s_{sim}$  in the  $x$  direction with an angular frequency  $\omega$  normalized to 1 Hz:

$$s_{sim}(x, y, t) = D(x, y) \sin(\omega t - 0.61), \quad (16)$$

where  $D(x, y)$  is a simulated mode shape and was calculated as:

$$D(x, y) = a \left( \left( \frac{x}{m} - 0.75 \right)^2 x - \left( \frac{y}{n} - 0.75 \right)^2 y \right), \quad (17)$$

where  $m \times n$  is the size of the image in pixels and  $a$  is a factor used to scale the amplitudes to a desired value. The choice of the function (17) and the phase shift in (16) is irrelevant. The simulated mode shape is seen in Figure 6.

The displacements (17) were normalized to a maximum amplitude of 0.04 pixel for the simple reconstruction (a) and the high-noise test (b), and 0.0004 pixel for the high-resolution test (c).

The pattern was distorted according to the displacements (16) using a bilinear interpolation [30] for every time instance.

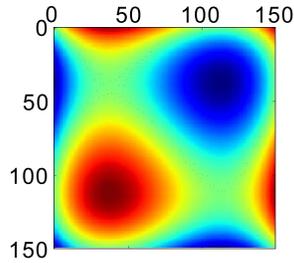


Figure 6: The simulated mode shape

Gaussian noise was added to the intensity values after the bilinear interpolation, when the value is still decimated and not yet rounded to an integer value.

The standard deviation of the Gaussian noise is 0.2 intensity values for the simple reconstruction (a) and the high-resolution test (c), and 25 intensity values for the high-noise test (b). Figure 7 shows an image from the high-noise sequence (b).

A sequence of 900 images was generated for each test, covering 30 periods of the harmonic.

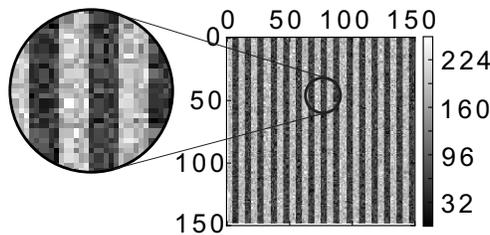


Figure 7: An image from the high-noise sequence

### 5.3. The reconstruction

The displacement amplitudes were reconstructed with a program that calculated the displacements from the simulated sequence using Equation (15).

The size of the subsets used was the width of a single line, which is  $d = 5$  pixels. A larger subset would produce a gradient with a smaller range of linearity, also a larger subset size blurs the neighbouring displacements more. The subset  $d = 5$  gives a resolution limit of 0.0008 pixel for the pattern with a dynamic range of 250 values.

The gradient direction of the selected pattern is in the  $x$  direction and is interchanging between the edges. The reconstructed displacement amplitudes were corrected for the gradient interchanging; therefore, they indicate amplitudes in the  $x$  direction, see Figure 8.



Figure 8: The direction of the gradient for the simulated pattern and the correction

The reconstructed mode shapes are compared to the true simulated values and can be seen in Figure 9. The white areas are where the values were cropped because the insufficient image gradient does not produce a valid reconstruction. The reconstruction for the simple reconstruction test (a) is good, because the displacements are above the resolution limit. The simulated mode shape is still evident at the resolution limit (the high-resolution test (c)), thanks in part to the noise dithering. The displacement shape is also recognisable in the high-noise test (b).

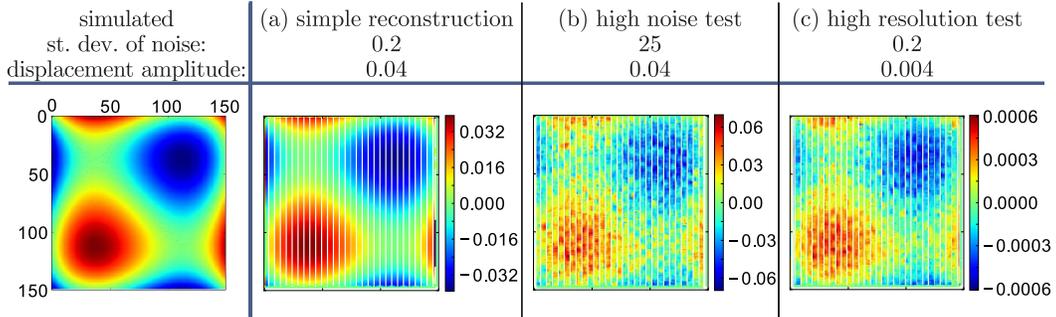


Figure 9: Simulation amplitude results for a simple reconstruction (a), a high-noise test (b) and for the high-resolution test (c) compared to the simulated displacements

## 6. Validation test

A real-world experiment was performed to validate the camera measurement system.

A piezoelectric accelerometer was attached to a cymbal's edge. The cymbal used is a Sabian 14"/35 cm B8 Thin Crash cymbal. A sticker with a black and a white area was placed on the accelerometer and filmed with a Fastec HiSpec 4 monochrome 8-bit high-speed camera. The cymbal was excited with the strike of a modal hammer; this triggered the acquisition of the acceleration data from the accelerometer and the filming of the camera. The cymbal set-up and a frame from the camera are shown in Figure 10. The frame size is  $304 \times 260$  pixels, the dynamic range is approximately 165 intensity values, the frame rate is 10,850 frames per second and a total of 27,096 frames in the bitmap (bmp) file format were captured. The acceleration from the accelerometer was sampled at a frequency of 50,000 Hz and 100,000 samples were recorded.

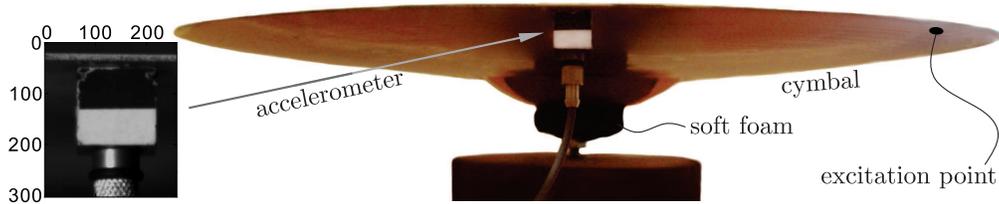


Figure 10: The validation experiment set-up and a frame from the high-speed camera

The image sequence was analysed using a subset size of  $3 \times 3$  pixels and the displacement of the centre point at the accelerometer was compared to the data from the accelerometer (Figure 11). The accelerometer data was integrated in the frequency domain to displacements and converted to the displacement in pixels using a conversion of 15.1 pixel per millimetre. The conversion was calculated from the known dimensions of the accelerometer and the number of pixels it covered. From Figure 11 it is clear that the peaks match well at lower frequencies where the displacements are larger than the noise floor, except at 100 Hz, where the flickering of the light produces an additional peak. The flickering can be filtered out.

## 7. Full-field modal test

Full-field, high-speed measurements of a steel beam and the cymbal from the validation test were conducted to measure the operational deflection shapes.

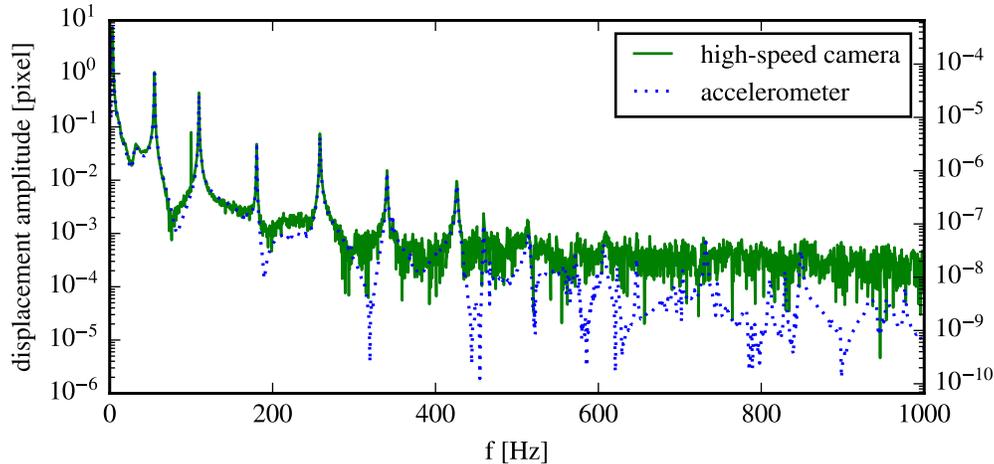


Figure 11: A comparison of the displacement amplitudes from the camera and the accelerometer

### 7.1. Free-free beam

A solid  $15 \times 30 \times 500$  mm steel beam was suspended on long ropes and hit with the modal hammer. A sticker with black and white lines was stuck to the thinner edge of the beam and filmed with the high-speed camera, enabling the camera to observe the displacements of the beam in the vertical direction. The experimental set-up is shown in Figure 12 and a frame from the camera is seen in Figure 13. The frame size is  $1710 \times 128$  pixels, the dynamic range is approximately 170 intensity values, the frame-rate is 3770 frames per second and 9784 frames were captured. The hammer was struck at the centre of the beam; the black tip of the modal hammer striking the beam is seen on the frame (Figure 13). It should be noted that a better approximation to the free-free boundary condition would be achieved if the excitation was in the direction perpendicular to the ropes and not in line with them. Such a set-up was used, because it produces smaller displacements, which do not exceed the area of constant gradient in the image.

The analysis was performed using a subset size of  $5 \times 5$  pixels, giving a theoretical displacement resolution limit of 0.0013 pixel.

The real parts of the displacement amplitudes for two rigid-body and three structural natural frequencies are plotted with respect to the beam length in Figure 14. The plots consist of the measured displacement amplitudes for the upper and lower lines on the beam. The mode shape is still

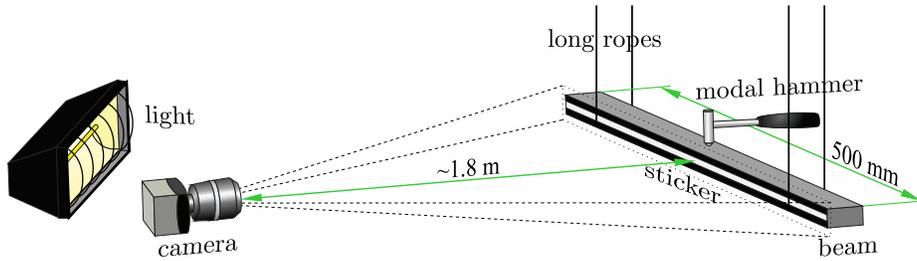


Figure 12: The set-up for the full-field beam experiment

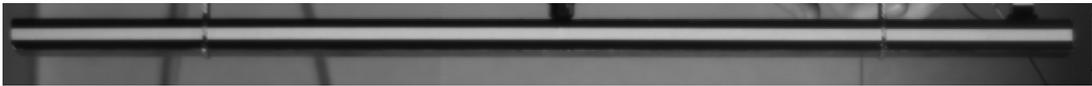


Figure 13: A frame from the full-field beam experiment

recognisable at 1671.15 Hz with a maximum amplitude of only 0.0008 pixel, or  $0.2 \mu\text{m}$  using a 3.328 pixel/mm conversion. The amplitude resolution is greater than the maximum amplitude of 0.0008 and greater than the 0.0013 pixel displacement limit as well, thanks to the noise dithering (Section 4.1).

## 7.2. Cymbal

A second measurement was performed, where the full-field cymbal's out-of-plane displacements were measured. The bottom side of the cymbal was painted with a pattern of black and white lines. The painted cymbal was filmed at an angle, as seen in Figure 15. Filming at such an angle enables the camera to see the cymbal's out-of-plane vibrations. The vibrations were induced with a shaker using a pseudo-random signal with the spectral power concentrated at higher frequencies. The cymbal was filmed with a Fastec SA-Z camera, using an 8-bit acquisition setting. A frame from the camera is shown in Figure 16. The frame size is  $1024 \times 592$  pixels, the dynamic range is approximately 200 intensity values, the frame-rate is 28000 frames per second and 112000 frames were captured, resulting in 64 GB of data.

The analysis was performed using a subset size of  $3 \times 3$  pixels on 90112 points. The processing was done on a laptop with an Intel i7-4510U processor with 8 GB of RAM and a solid-state drive. The calculation of the displacement time series for all the points took 9.5 minutes, while a further 15 minutes were needed for the Fast Fourier Transform of all the time series. The com-

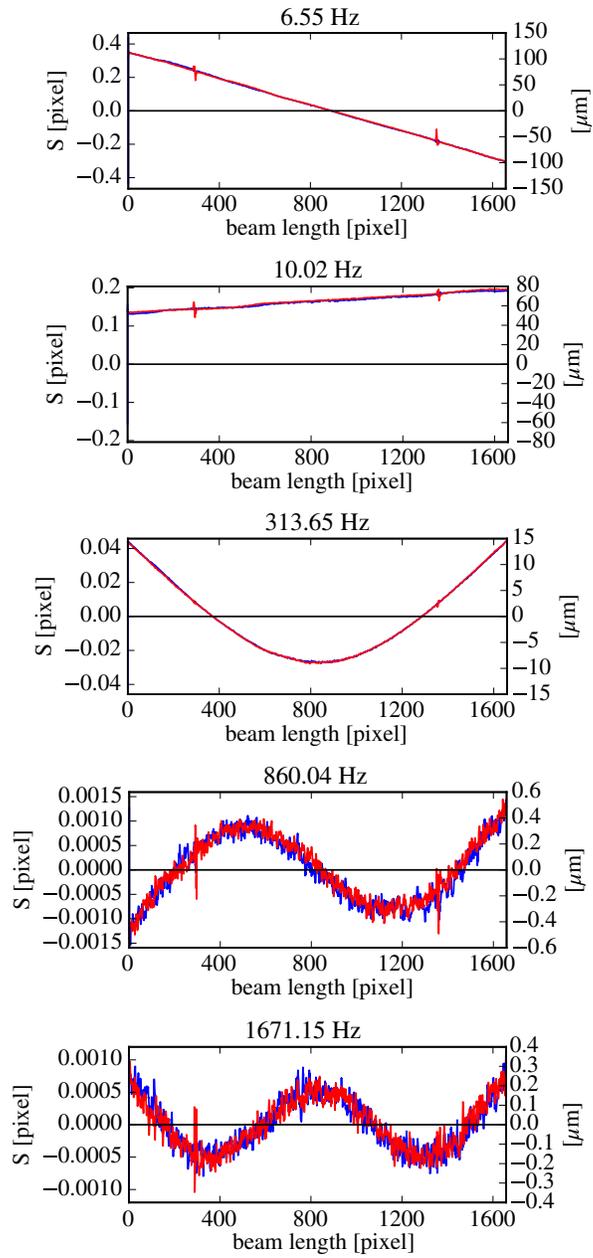


Figure 14: Displacement amplitudes of multiple lines from the analysis of the beam (blue - upper, red - lower line on the beam)

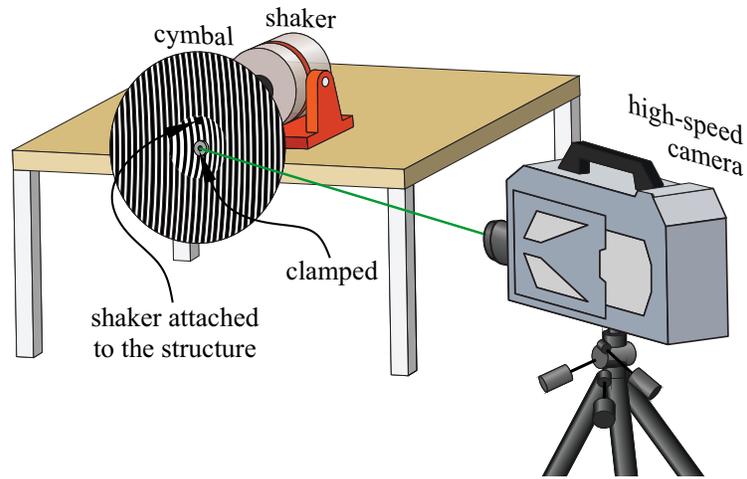


Figure 15: The cymbal full-field experiment set-up

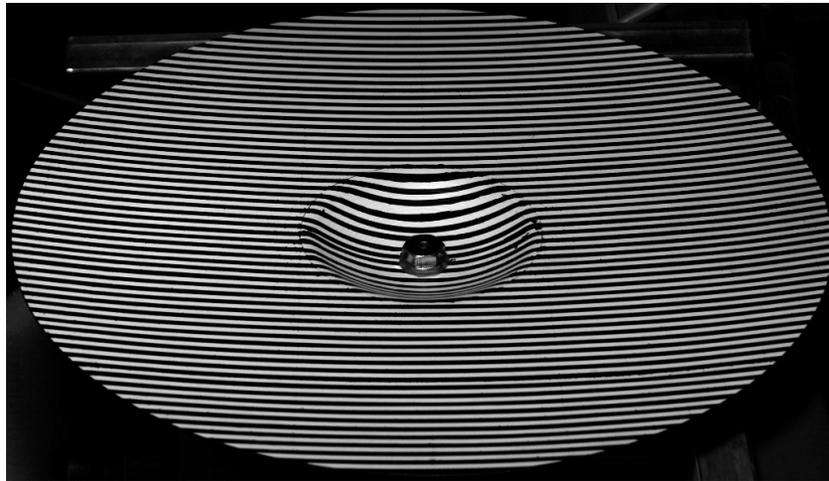


Figure 16: A frame from the full-field cymbal experiment

putational speed was limited by the solid-state drive.

The image gradient runs perpendicular to the lines and since out-of-plane motion is prevailing the displacements measured are the cymbal's out-of-plane displacements. Some recognisable mode shapes from the analysis are shown in Figure 17, where absolute displacement amplitudes in the vertical direction are plotted. The results of the analysis contain approximately 50 mode shapes ranging up to 2370 Hz. The noise floor is approximately

0.0002 pixel and a shape with a maximum absolute amplitude of 0.00036 pixel is still recognisable at 1851.75 Hz.

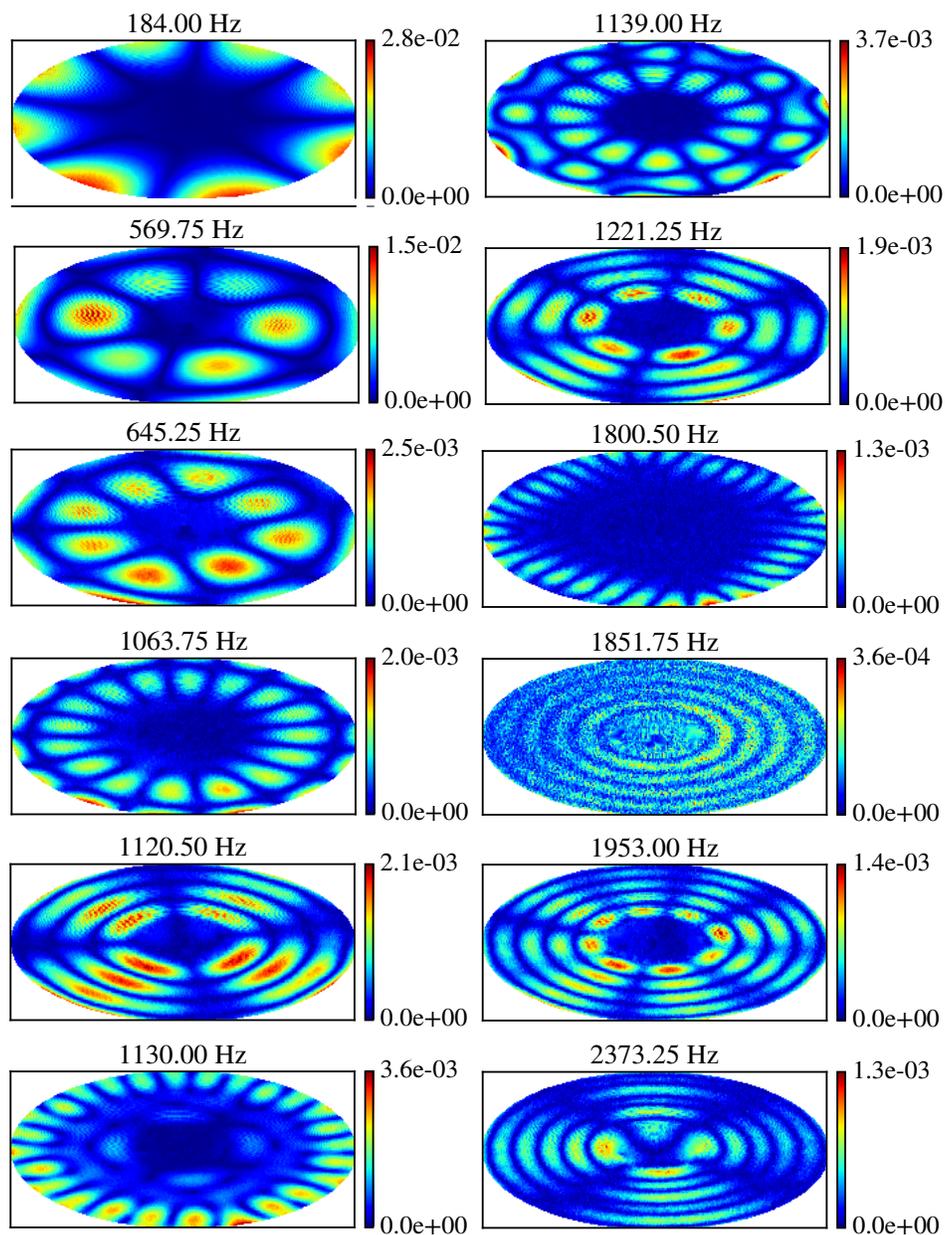


Figure 17: Evident mode shapes from the cymbal full-field experiment in pixels

The conversion factor from the displacements in pixels to the cymbal's out-of-plane displacements in millimetres is approximately 0.3 mm/pixel, depending on the depth from the camera and the angle of the surface.

## 8. Conclusion

This research analyses the potentials of optical measurements with high-speed cameras in the field of modal analysis. The conclusions are derived on a simplified gradient-based optical flow allowing for every pixel to act as an individual sensor, thereby providing the minimal blurring of spatial information compared to using larger subset sizes in template-matching approaches (DIC), where subsets are commonly in the range 11-71 pixels [24]. By assuming small displacements a linear relation between the light intensity and the displacement in the direction of the intensity gradient is obtained, simplifying the displacement identification and shortening the processing time. The research on the achievable resolution demonstrates a high displacement resolution. This displacement resolution is dependent on the intensity depth, the size of the observed area (subset) and the image pattern. The findings suggest subpixel displacement resolutions of less than thousandths of a pixel, given the appropriate conditions. Furthermore, even smaller displacement amplitudes are resolvable due to noise dithering. In reality the noise limits the resolution to the range of thousandths or ten-thousandths of a pixel. Local averaging, a good reference image and a long image sequence, provide good robustness to noise. The simulations and real-world experiments confirm these findings.

The conclusions on the achievable resolution should in general apply to all optical flow methods (DIC, gradient-based, phase-based) as they all use the same data.

- [1] A.B. Stanbridge, D.J. Ewins, Modal testing using a scanning laser Doppler vibrometer, *Mechanical Systems and Signal Processing* 13(2) (1999) 255-270.
- [2] O. Lekberg, Electronic speckle pattern interferometry, *Physics in technology* 11(1) (1980).
- [3] C.M. Vest, *Holographic interferometry*, John Wiley & Sons, New York, 1979.
- [4] R. Marsili, M. Moretti, G.L. Rossi, Thermoelastic Modal Stress Analysis, In: *Proc. of 26th International Modal Analysis Conference (IMAC XXVI)* (2008).
- [5] M.N. Helfrick, C. Niezrecki, P. Avitabile, T. Schmidt, 3D digital image correlation methods for full-field vibration measurement, *Mechanical Systems and Signal Processing* 25(3) (2011) 917-927.
- [6] W. Wang, J.E. Mottershead, T. Siebert, A. Pipino, Frequency response functions of shape features from full-field vibration measurements using digital image correlation, *Mechanical systems and signal processing* 28 (2012) 333-347.
- [7] A. Davis, M. Rubinstein, N. Wadhwa, G.J. Mysore, F. Durand, W.T. Freeman, The visual microphone: passive recovery of sound from video, *ACM Transactions on Graphics(TOG)* 33(4) (2014).
- [8] W.H. Peters, W.F. Ranson, Digital imaging techniques in experimental stress analysis, *Optical Engineering* 21(3) (1982).
- [9] S.K. Mitra, J.F. Kaiser, *Handbook for Digital Signal Processing*, New York: Wiley, 1993.
- [10] B. Pan, H. Xie, B. Xu, F. Dai, Performance of sub-pixel registration algorithms in digital image correlation, *Meas. Sci. Technol.* 17 (2006) 1615-1621.
- [11] S. Baker, I. Matthews, Equivalence and efficiency of image alignment algorithms, *Proc. IEEE Conf. Comp. Vis. Pattern Recognition* 56 (2001) 1090-1097

- [12] B. Pan, K. Li, W. Tong, Fast, Robust and Accurate Digital Image Correlation Calculation Without Redundant Computations, *Experimental Mechanics* 53(7) (2013) 1277-1289
- [13] E. Zappa, P. Mazzoleni, A. Matinmanesh, Uncertainty assessment of digital image correlation method in dynamic applications, *Optics and Lasers in Engineering* 56 (2014) 140-151.
- [14] J.L. Barron, D.J. Fleet, S.S. Beauchemin, Performance of optical flow techniques, *International journal of computer vision* 12(1) (1994) 43-77.
- [15] J.G. Chen, N. Wadhwa, Y.J. Cha, F. Durand, W.T. Freeman, O. Buyukozturk, Modal identification of simple structures with high-speed video using motion magnification, *Journal of Sound and Vibration* 345 (2015) 58-71.
- [16] B.D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, *International Joint Conference on Artificial Intelligence* (1981) 674-679.
- [17] B.K. Horn, B.G. Schunck, Determining optical flow. In: 1981 Technical Symposium East. International Society for Optics and Photonics (1981) 319-331.
- [18] H. Liu, T.H. Hong, M. Herman, T. Camus, R. Chellappa, Accuracy vs efficiency trade-offs in optical flow algorithms, *Computer vision and image understanding* 72(3) (1998) 271-286.
- [19] C.L. Fennema, W.B. Thompson, Velocity determination in scenes containing several moving objects, *Computer graphics and image processing* 9(4) (1979) 301-315.
- [20] E. Hildreth, The measurement of visual motion. PhD thesis, MIT (1983).
- [21] C.Q. Davis, D.M. Freeman, Statistics of subpixel registration algorithms based on spatiotemporal gradients or block matching, *Opt. Eng* 37(4) (1998) 1290-1298.
- [22] N. Ohta, Optical flow detection by color images. In: *Proc. Tenth International Conference on Pattern Recognition* (1989) 801-805.

- [23] Y.Q. Wang, M.A. Sutton, H.A. Bruck, H.W. Sreier, Quantitative Error Assessment in Pattern Matching: Effects of Intensity Pattern Noise Interpolation, Strain and Image Contrast on Motion Measurements, *Strain* 45 (2009) 160-178.
- [24] B. Pan, H. Xie, Z. Wang, K. Qian, Z. Wang, Study on subset size selection in digital image correlation for speckle patterns, *Optics Express* 16(10) (2008)
- [25] W. Zhiyong, L. Li'an, W. Shibin, Optimization of Speckle Size in Digital Image Correlation Method.
- [26] I. Sobel, G. Feldman, A 3x3 isotropic gradient operator for image processing, a talk at the Stanford Artificial Project (1968) 271-272.
- [27] J.M.S. Prewitt, Object enhancement and extraction, *Picture processing and Psychopictorics* 10(1) (1970) 15-19.
- [28] N.M.M. Maia, J.M.M. Silva, Theoretical and experimental modal analysis. Taunton: Research Studies Press, 1997.
- [29] J. Tsui, Digital techniques for wideband receivers, SciTech Publishing, Raleigh, 2004.
- [30] R.C. Gonzales, R.E. Woods, Digital Image Processing, 2-nd Edition. Upper Saddle River, NJ, Prentice Hall, 2002.