

Detecting vibration features from remote objects based on high-speed imagery

Gabriela Paunescu,* Peter Lutzmann, Daniel Wegner, and Benjamin Göhler

Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Ettlingen, Germany

Abstract. Remote detection of vibrational features from an object is important for many short-range civil applications, but it is also of interest for long-range applications in the defense and security areas. The well-established laser Doppler vibrometry technique is widely used as a high-sensitivity, noncontact method. The development of camera technology in recent years made image-based methods reliable passive alternatives for vibration and dynamic measurements. We investigate and discuss the potential of high-speed imaging technique for medium- and long-range vibration detection. The sensitivity and the limitations of the method are experimentally investigated in comparison to the well-established Doppler vibrometry technique. As atmospheric turbulence is expected to become a limiting factor for long-range applications, imaging in the short-wave IR (SWIR) to mid-wave IR (MWIR) rather than in the visual range is advantageous due to the longer wavelength. We present experiments on the vibration detection from SWIR and MWIR image sequences, as well as additional experiments on the extraction of vibration signature under strong local turbulence conditions. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: [10.1117/1.OE.60.1.013102](https://doi.org/10.1117/1.OE.60.1.013102)]

Keywords: high-speed imaging technique; laser Doppler vibrometry; vibration information; turbulence; angle-of-arrival fluctuations.

Paper 20201046 received Sep. 4, 2020; accepted for publication Dec. 11, 2020; published online Jan. 6, 2021.

1 Introduction

The vibrational features of an object deliver valuable information for various civil applications, but also for applications in the defense and security areas.¹ Noncontact vibration measurement techniques are highly desirable and have been intensively studied in the last years. At the present, laser Doppler vibrometry is widely used as a high-sensitivity method.^{1,2} With the development of camera technology, the image-based approaches became reliable passive alternatives for vibration and dynamic measurements.^{3–14} The passive recovery of vibration information from high-speed video recordings was reported in applications such as vibrations of buildings⁷ and bridges^{8,12} or voice reconstruction.¹¹ The sensitivity of the method has also been demonstrated in the recovery of audio signals from high-speed video by detecting the small vibrations caused on an object by sound.^{3–5} Moreover, high-speed vision-based sensor systems have been developed that extract vibration signals in real time.¹²

This study mainly focuses on the potential extension of the high-speed imaging technique for vibration detection to long-range applications. Section 3 investigates and discusses the sensitivity and the limitations of the method in comparison to laser Doppler vibrometry. The comparison is based on indoor experiments with a visual camera, using a loudspeaker membrane as a test object. The potential extension of the high-speed imaging method to longer range applications was also experimentally investigated. As atmospheric turbulence is expected to become a limiting factor for long-range applications, imaging in the short-wave IR (SWIR) to mid-wave IR (MWIR) rather than in the visual range is advantageous. Experimental investigations on the

*Address all correspondence to Gabriela Paunescu, gabriela.paunescu@iosb.fraunhofer.de

extent of the high-speed imaging method from visual to SWIR and MWIR spectral range are presented in Sec. 4. The influence of strong local turbulent conditions on the two-dimensional (2D) vibration signature is investigated in Sec. 5.

2 Vibration Detection from High-Speed Image Sequences

The vibration detection from a series of images recorded at a defined frame rate is based on the analysis of the relative displacement of a selected pixel surrounded by a subset as a function of time. One of the images, e.g., the first image of the series, might be used as a reference image. The displacement calculated in pixels can be converted in length units considering the geometry and the parameters of the measurement system. To detect the pixel displacement, a sum of squared difference criteria^{4,6} was implemented. The technique described by Wang et al.⁴ is reviewed below. The image matching process maps the points in one image (reference) to the target image with subpixel accuracy. The first step is to select a reference subset containing sufficient local intensity variations and centered at the interrogated pixel. The image matching technique searches the deformed image for the particular target subset whose intensity pattern is most similar to the reference subset. To detect the subset matching, the process employs a correlation criterion function, which can be written as^{4,6}

$$C = \sum_{i=1}^N [af(x_i, y_i) + b - g(x'_i, y'_i)]^2, \quad (1)$$

where N is the number of pixels in the subset, i is the pixel index, $f(x_i, y_i)$ is the intensity value at the pixel i in the reference subset, $g(x'_i, y'_i)$ is the intensity value at the potential matching pixel in the target subset, a is a scale factor, and b is an offset of intensity. This is called the generalized parametric sum of squared difference⁶ and the subset matching is achieved by minimizing the value of C [Eq. (1)].

A square pattern with $N = (2M + 1) \times (2M + 1)$, centered at the analyzed pixel (x_0, y_0) , will be selected further as a reference subset. The pixels in the corresponding target subset could be expressed by the shape mapping function

$$x'_i = x_i + \xi + \xi_x \Delta x_i + \xi_y \Delta y_i, \quad y'_i = y_i + \eta + \eta_x \Delta x_i + \eta_y \Delta y_i, \quad (2)$$

where (ξ, η) denotes the translation between the centers of the two potential matching subset patterns, $\Delta x_i = x_i - x_0$, $\Delta y_i = y_i - y_0$, and parameters $\xi_x, \xi_y, \eta_x, \eta_y$ represent the displacement gradient.

The C coefficient in Eq. (1) becomes a function of eight unknown parameters, which are included in the vector $\mathbf{p} = \{\xi, \eta, \xi_x, \xi_y, \eta_x, \eta_y, a, b\}^T$. By defining the function

$$\zeta_i(\mathbf{p}) = af(x_i, y_i) + b - g(x'_i, y'_i), \quad (3)$$

Equation (1) can be rewritten as

$$C(\mathbf{p}) = \sum_{i=1}^N [\zeta_i(\mathbf{p})]^2, \quad (4)$$

$C(\mathbf{p})$ may be minimized in various ways. A time effective method is the Gauss–Newton algorithm, which iteratively finds the minimum by applying the governing equation⁴

$$\mathbf{p}_{n+1} = \mathbf{p}_n - \left[\sum_{i=1}^N (\mathbf{J}_i \cdot \mathbf{J}_i^T) \right]^{-1} \cdot \sum_{i=1}^N [\zeta_i(\mathbf{p}) \cdot \mathbf{J}_i], \quad (5)$$

where $n = 0, 1, 2, \dots$, and $\mathbf{J}_i = [\partial \zeta_i(\mathbf{p}) / \partial \mathbf{p}]$ is the gradient of ζ_i . In the case of small displacements, the iteration process may start from the initial parameters $\mathbf{p}_0 = \{0, 0, 0, 0, 0, 0, 1, 0\}^T$. As

Eq. (5) involves the intensities at noninteger pixel positions in the target image, an interpolation algorithm is required such as the bicubic interpolation function.

The resulting values for ξ and η represent the horizontal and vertical displacements at the subset center. To convert these values to length units, the geometry of the measurement arrangement has to be considered. If an object placed at distance D (m) from the camera is imaged onto the sensor with a lens of focal length f (mm), the object movement d_{mm} and pixel displacement d_{pixel} are related by

$$d_{\text{mm}} = d_{\text{pixel}} \times \frac{1}{M} = d_{\text{pixel}} \times \frac{D u_{\text{pixel}}}{f}, \quad (6)$$

where u_{pixel} (μm) is the pixel size and M (pixel/mm) may be termed as optical magnification. In addition, an angle correction should be considered. Defining the incidence angle δ as the angle between the movement direction and the camera line of sight, the displacement d_{mm} should be divided by $\sin(\delta)$.

3 Experiments in Visual Spectral Range: Comparison to Laser Doppler Vibrometry

3.1 Short-Range Experiments

Several short-range experiments were performed to investigate the sensitivity of the described method and the influence of different parameters on the measurement performance. A loudspeaker membrane excited by a sine wave was used as a test object. For comparison, the vibrations of the membrane were simultaneously measured by a laser Doppler vibrometer (LDV).

3.1.1 Experimental setup

The experimental arrangement is shown in Fig. 1. The loudspeaker was placed on a separate optical table in order to avoid direct transfer of vibrations to the camera. As the image matching technique requires varying texture intensities, a pattern was applied on the vibrating membrane. The measuring systems were placed at a distance of about 2.5 m from the loudspeaker and aimed at the object with an incidence angle of about 35 deg. The camera used was a Phantom V711 with a monochrome CMOS sensor acquiring up to 7530 frames per second (fps) at the full resolution of 1280×800 pixels. The camera was equipped with lenses with various focal lengths starting from 50 up to 400 mm. The 632-nm-LDV was a commercially available Polytec model OFV-303 with the modular vibrometer controller OFV-3001.

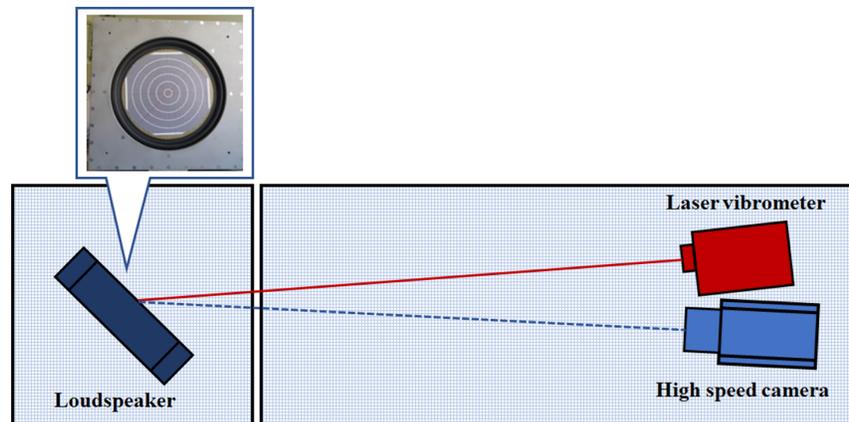


Fig. 1 Experimental setup for optical measurement of loudspeaker vibration in comparison with LDV.

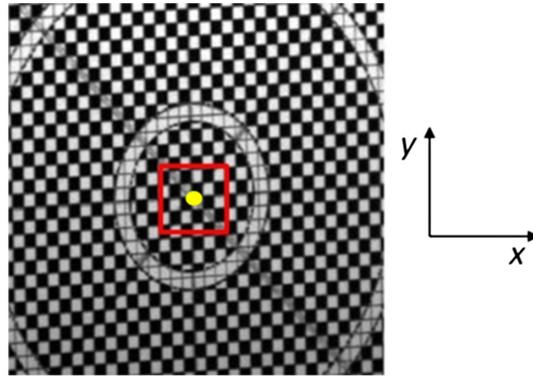


Fig. 2 Image of the loudspeaker membrane recorded with the high-speed camera. The red rectangle marks the subset used for the calculations, the yellow point the investigated pixel.

3.1.2 Reconstruction of the vibration signal

The first test of the optical method was to detect/reconstruct the vibrations of the loudspeaker membrane. The membrane was excited with a sine wave with frequencies between 10 and 700 Hz. The surface was imaged on the high-speed camera using lenses with focal lengths between 50 and 400 mm. For example, Figs. 2 and 3 show the measurement of a signal of 37 Hz. The camera was equipped with a zoom lens set to a focal length of 300 mm. High-speed acquisitions were performed with a resolution of 256×256 pixels at a frame rate of 2000 fps. Figure 2 displays an acquired image with the corresponding pattern. The procedure was applied to a sequence of 1 s (2000 images) for a subset of 41×41 pixels (red square) centered at the middle of the pattern, marked in Fig. 2 by the yellow point. The pixel displacement was evaluated for every image in comparison with a target obtained by averaging all images in the

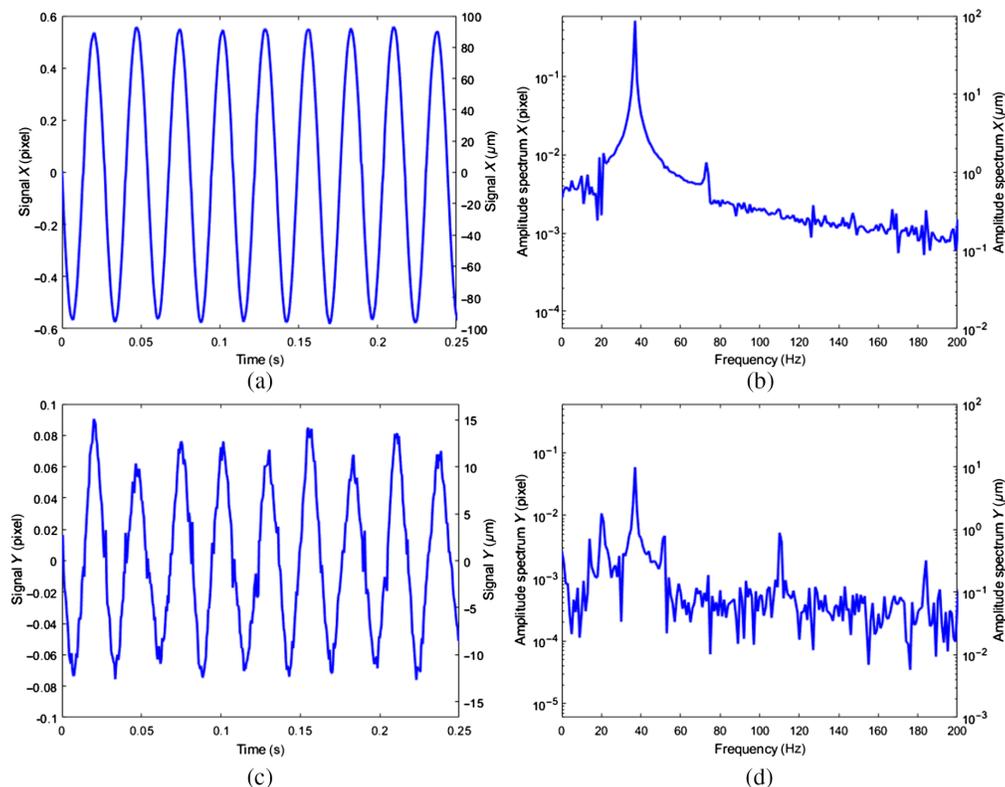


Fig. 3 (a)–(d) Vibration measurement using the high-speed imaging technique. Subpixel displacement over the time along the x and y axes and the corresponding Fourier spectra.

sequence. The convergence tolerance was set to 10^{-5} and the maximum number of iterations to 200. The results are shown in Fig. 3. As the deformations of the membrane were negligible inside the subset, the displacement gradients $\xi_x, \xi_y, \eta_x, \eta_y$ have been set to zero for these calculations. Applying the algorithm for a vector \mathbf{p} with all eight parameters resulted often in a singular matrix $\sum_{i=1}^N (\mathbf{J}_i \cdot \mathbf{J}_i^T)$, that is, one not invertible as required in Eq. (5). Using a different method to minimize the value of C coefficient as a function of eight parameters, such as the “conjugate gradient” method in Mathematica, gave similar results.

Figure 3 shows the calculated pixel displacements along the x and y axes [Fig. 3(a) and (c)] over a time interval of 0.25 s. The method leads to a very accurate reconstruction of the sinusoidal vibration, especially for the x axis, where the incidence angle is about 35 deg. Along the y axis, the camera has a very small incidence angle and the signal is about 10 times lower. The vibration frequency of 37 Hz could be extracted from the both signals, as observed in the Fourier transform spectra for the x and y axes [Fig. 3(b) and (d)].

3.1.3 Sensitivity comparison of the methods

The sensitivity of the high-speed imaging technique was investigated using the experimental arrangement shown in Fig. 1. The loudspeaker membrane was excited by a sinusoidal electrical signal created by a wave generator. The signal amplitude was tuned between 0.1 and 1000 mV. The amplitude of the membrane vibration was measured as a function of the applied voltage, both by laser vibrometry and by high-speed imaging. The vibrometer laser beam was focused on the center of the pattern, which was also considered the center of the investigated subset (Fig. 2). For a comparison between the two methods, the values of the vibration amplitude were estimated in length units. The pixel displacement in micrometers was calculated from Eq. (6). In addition, an angle correction should be applied to both methods. The laser vibrometer exhibits a cosine dependence with respect to the angle of incidence, whereas for high-speed imaging, a sine dependence must be considered. Figure 4 shows the measured vibration amplitude in micrometers as a function of the applied electrical signal before (a) and after (b) the angle correction. Each of the methods shows a linear variation (slope of 1 in the log–log scale) of the measured vibration amplitude with the applied signal, and the values are quite similar after the angle correction [Fig. 4(b)]. Although the LDV could measure the vibration over the full investigated range, the high-speed imaging technique could detect a signal starting from 5 mV ($0.5 \mu\text{m}$), implying a sensitivity at least one order of magnitude lower. The actual difference in the sensitivities of the high-speed imaging technique and the LDV cannot be exactly quantified on the basis of the experiment performed, since the former was not yet at the detection threshold at the smallest amplitude of the electrical excitation signal and could, therefore, detect even smaller amplitudes. The used commercial LDV controller OFV-3001 can theoretically achieve a displacement resolution of 2 nm.

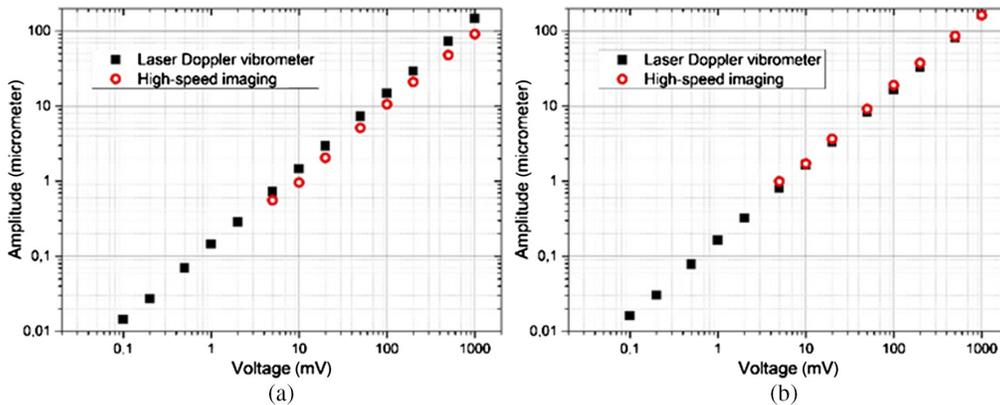


Fig. 4 Vibration amplitude of the loudspeaker membrane as a function of applied signal: (a) before and (b) after the angle correction. The red points show the measurements with the high-speed camera, the black points those with the LDV.

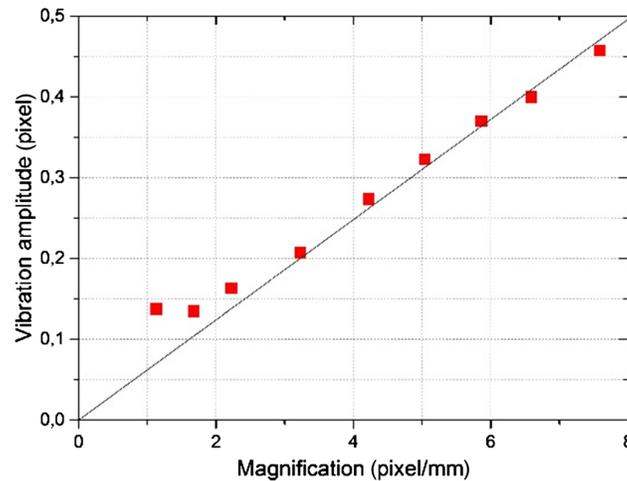


Fig. 5 Variation of the calculated vibration amplitude in pixels as a function of the magnification.

3.1.4 Influence of optical magnification on vibration analysis

For the same experimental geometry (Fig. 1), the variation of the focal length induces a directly proportional variation of the magnification. Consequently, for constant vibration amplitude, the value of the pixel displacement should be directly proportional to focal length or magnification, respectively [Eq. (6)]. For improving the experimental setup and verifying the reliability of the calculations, the influence of the imaging optics was experimentally investigated. Measurements were performed using lenses with various focal lengths in the range between 50 and 400 mm. The results are displayed in Fig. 5. The calculated vibration amplitude values in pixels fit closely to a line intersecting the origin. A deviation from the line is observed for the focal length of 50 mm only. This may be because the pattern is not well resolved at the corresponding magnification.

3.1.5 Contrast influence on vibration analysis

The applied procedure requires intensity variations over the investigated subset, so the illumination of the object and the contrast in the acquired images are very important for measurement accuracy. The influence of illumination conditions was experimentally investigated. The loud-speaker membrane vibrating at 16 Hz was illuminated at different light levels and the vibration amplitude was estimated as a function of the image contrast (Fig. 6). The modulation contrast was calculated according to $(I_{\max} - I_{\min}) / (I_{\max} + I_{\min})$, where I_{\max} and I_{\min} are the maximum

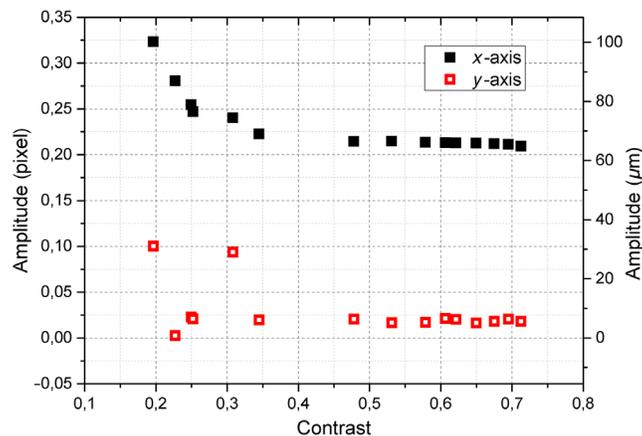


Fig. 6 Measured vibration amplitude as a function of modulation contrast for the investigated subset.

and minimum intensities in the investigated subset of 61×61 pixels. The calculated vibration amplitude along the x axis is about 10 times higher than the vibration amplitude along the y axis due to the different incidence angles. Above a certain threshold of the contrast, the measured vibration amplitude remains constant for both axes, while in the lower range, the values show strong variations. The measurements under low-contrast conditions allow the detection of real vibration frequency, but the obtained vibration amplitude is not reliable. For accurate amplitude measurements, a certain contrast (with this experiment a modulation value of about 0.4) is required.

3.2 Medium-Range Measurements

The potential of the high-speed imaging method for medium-range applications was experimentally investigated. We used a visible high-speed camera for our first experiments, but the development of camera technology in recent years also allows transferring this technique to other wavelength ranges (i.e., SWIR or MWIR), particularly as the requirements for frame rate will be less stringent. For many applications, a frame rate between 100 and 200 fps will be sufficient.

The test target—a small truck (Unimog)—stood outside at a distance of 67 m away from the position of the sensors, which were operated from the third floor of one of our laboratories (Fig. 7). The vibrational features were estimated by the same method used for short-range experiments. The images were taken with full sensor size of 1280×800 pixels with a frame rate of 2000 fps. The camera was equipped with a zoom lens with a range of focal length between 50 and 500 mm. A focal length of about 400 mm was chosen in order to adapt the field of view to the entire target.

Figure 8 shows an example of measured power spectra with the engine idling (about 1980 rpm) and turned off. The red square in the visible image [Fig. 7 (b)] marks the investigated subset of 33×33 pixels. The main frequency component at 33 Hz and additionally the first harmonic at the vertical y axis can be clearly observed when the engine is turned on. The images were acquired with a focal length of ~ 400 mm corresponding to a magnification of 0.3 pixels/mm according to Eq. (6) ($20\text{-}\mu\text{m}$ pixel size of the high-speed camera). The region of interest (ROI) used as example for the one-dimensional (1D) vibration signature is a region of the lateral window of the truck. The window surface was not modified for the measurements. As the window has a high transmissivity in the visible range, the contrast is simply given by the objects in the cabin. The signal detected for this region indicates vibrating objects inside the cabin.

Apart from 1D signatures from a target, the high-speed recordings offer the possibility to generate spatially resolved two-dimensional (2D) vibration signatures. An example is shown in Fig. 9 for the characteristic main frequency at 33 Hz. Vibration amplitudes were extracted from 39×24 subsets, with one single subset consisting of 33×33 pixels. The 2D maps of the

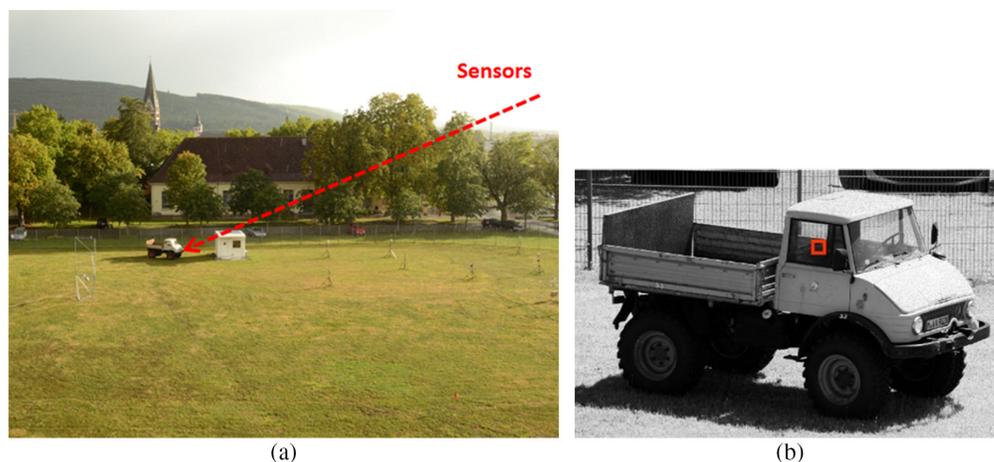


Fig. 7 (a) View from the sensors to the target and (b) visible image recorded with the high-speed camera from the sensors' location.

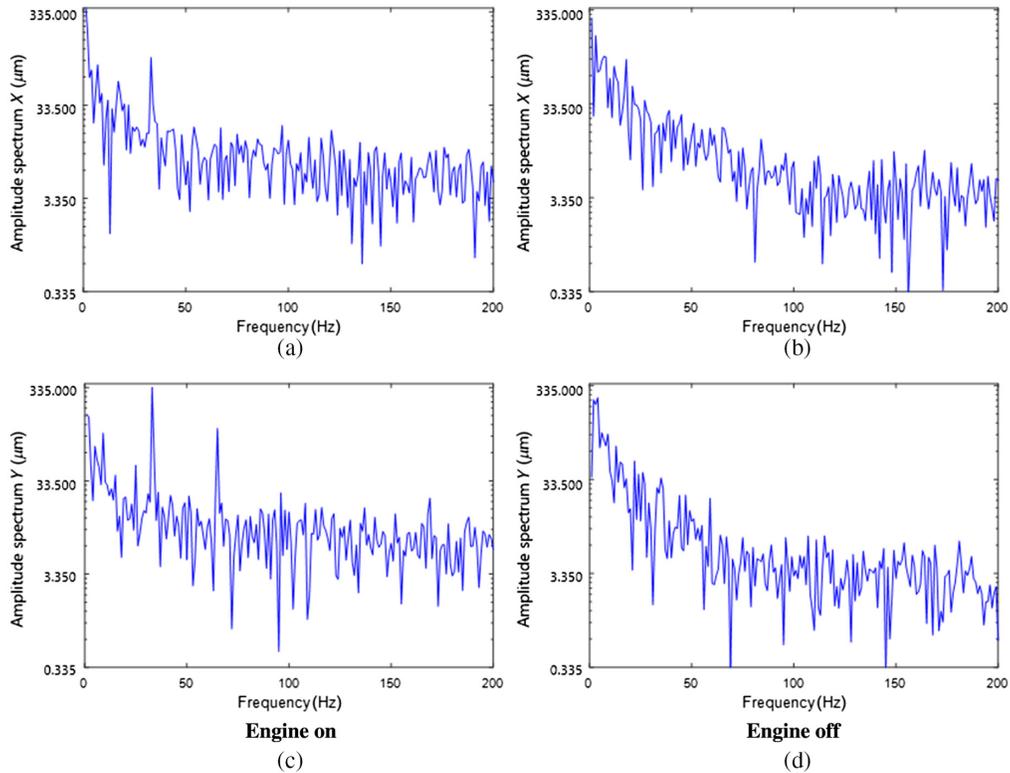


Fig. 8 High-speed camera: vibrational spectra of the small truck acquired from the position marked by a red square at Fig. 7(b) [(a), (c) engine on and (b), (d) engine off].

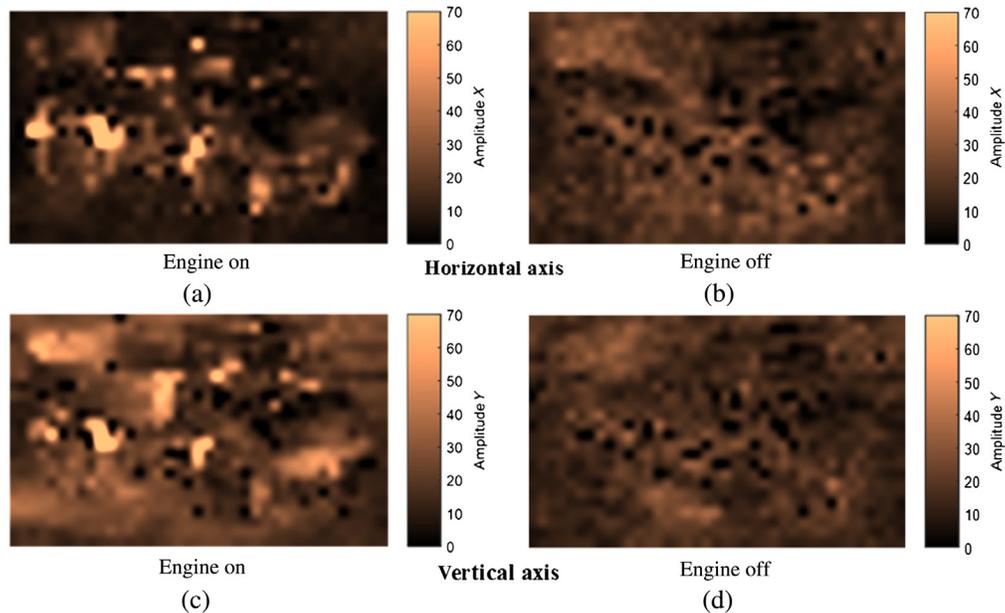


Fig. 9 2D-maps of vibration amplitude (39×24 pixels) along the horizontal axis and vertical axes recorded with the high-speed camera for (a), (c) the engine idling at the fundamental frequency of 33 Hz and (b), (d) with the engine turned off.

vibration amplitude along the x and y axes, respectively, are displayed. Note that a separate chart for the horizontal/vertical displacement direction will make different characteristics of the target more prominent. Mapping the vibrations along the horizontal x axis resulted in an easily noticeable contour of the wheels and of the back part of the truck [Fig. 9(a) and (c)]. In the vertical

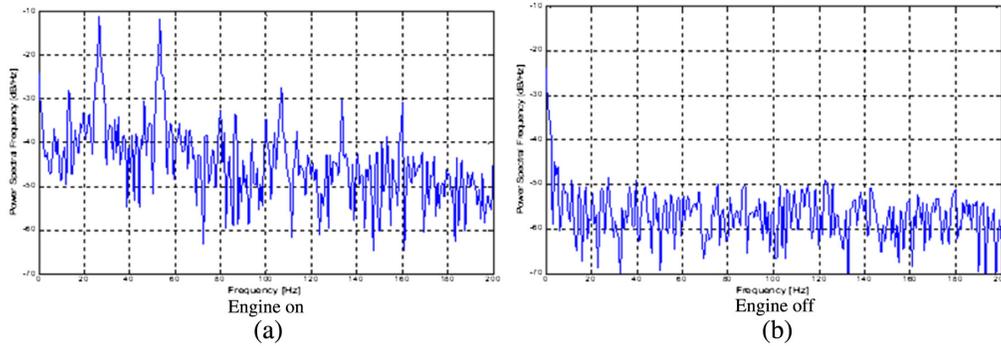


Fig. 10 LDV: vibrational spectra of a small truck [with (a) engine on and (b) engine off] acquired from an arbitrary position marked by a white circle at Fig. 11.

y axis vibration map, other features could be recognized, including the truck shadow. For comparison, the same measurements were performed with the engine turned off [Fig. 9(b) and (d)].

The comparison with the second investigated technique—laser Doppler vibrometry—was performed with an in-house-developed system.² The sensor is all-fiber-based and operates at a wavelength of 1.5 μm . With an output power of several hundreds of mW, a spectral line width smaller than 1 kHz and a field of view of 75 μrad , this system is optimized for vibration measurements at distances of up to one kilometer. The vibrometer system was mounted on a computer-controlled pan/tilt-head, and the laser beam was scanned step by step across the target to build a vibration image with 39×27 pixels.

Figure 10 shows the vibrational spectra of the small truck from an arbitrary single pixel of the target with idling engine and engine off. The fundamental frequency component and its first harmonic are easily recognizable. The idling frequency agreed reasonably well with the camera measurement (27 Hz instead of 33 Hz).

The spatially resolved vibration signature at 27 Hz is shown in Fig. 11 for both cases, engine on and off. Note the very clear segmentation of the vibrating targets from the background.

Comparing the results of the two sensors—high-speed camera and LDV—reveals some essential differences. The high-speed camera technique is sensitive to any contrast changes, and therefore, other features could be incorrectly detected as vibrational features of the target, e.g., the shadow of the truck [Fig. 9(c)]. Every movement of the image pixels—regardless of the cause—may lead to a vibrational image. These pixel shifts could be caused by turbulence effects, vibrations introduced to the camera cooling system, or building vibrations, depending on where the camera is placed. However, the images show distinctive characteristics at the typical real vibrational features (some specific sharp frequency components). Note that the vibration images, and therefore, their interpretation are differently for the two techniques. With high-speed imaging, small movements along the line of sight cannot be detected, whereas with laser

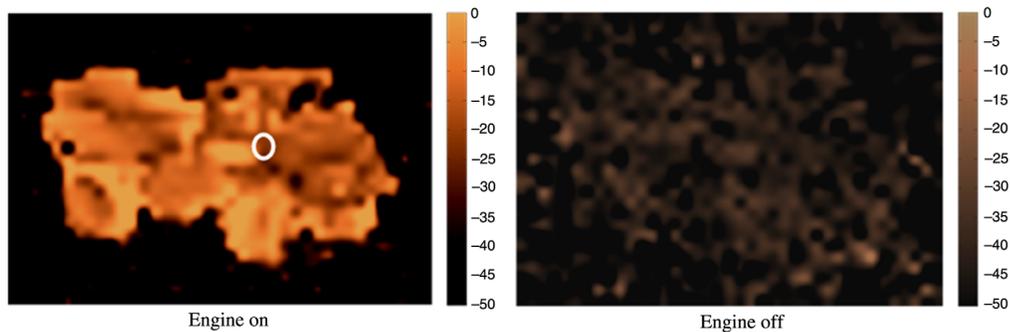


Fig. 11 2D-vibration image (spatially resolved vibration signature) for the small truck at the fundamental frequency component at 27 Hz. Data were recorded using the 1.5- μm LDV with an image of 39×27 pixels at a range of 67 m.

Table 1 Technical data of the sensor systems used for infrared image acquisitions.

	MWIR camera	SWIR camera
Detector	HgCdTe-FPA	InGaAs-FPA
Spectral range (μm)	1.5 to 5.3	0.4 to 1.7
Cooling	Linear Stirling cooler	None
Elements	256 \times 256	640 \times 512 (256 \times 256 used)
Pitch (μm)	40	15
Data (bit)	14	14
Frame rate	884 Hz	400 Hz with full frame 700 Hz with 256 \times 256 subframe
Integration time (ms)	1	1
Optics	100 mm F/2	35 mm F/3.5
Spatial resolution (mrad)	0.4	0.428
Filters spectral range (μm)	1.5 to 1.8 2.0 to 2.5 3.4 to 5.3 4.8 to 5.3	1.4 to 1.7 0.9 to 1.7

vibrometry they are identified due to the Doppler shift along the laser beam (equivalent to the line of sight).

4 Measurements in IR Spectral Range

4.1 Sensors

The recovery of vibration information from high-speed images in the SWIR and MWIR spectral range was investigated using an experimental arrangement similar to that described in Sec. 3.2. The target—a small truck (Unimog), as shown in Fig. 7—was placed outdoors at a distance of ~ 65 m from the imaging systems located at the third floor of the institute building. Infrared images were acquired using two cameras, one in the MWIR and one in the VIS/SWIR spectral range. The main specifications of the systems are listed in Table 1. The MWIR camera is based on a cooled 256 \times 256 HgCdTe-FPA for the 1.5- to 5.3- μm spectral range. It enables 14-bit digital image acquisition at a full frame rate of 884 Hz. Using different spectral filters inserted between lens and detector array, the effects of emitting and reflecting contributions to the target signatures can be investigated with respect to their significance for the vibration detection of the target. The second camera is based on an InGaAs-FPA detector sensitive in the VIS/SWIR spectral range from 0.4 to 1.7 μm . A 400-Hz frame rate at the full image size of 640 \times 512 can be captured with 14 bit depth. Selecting a subframe of 256 \times 256 pixels enables a frame rate of 700 Hz to be used during the measurements. Additionally, the spectral band has been limited within the 0.9- to 1.7- μm range using different bandpass or high-pass filters. An important characteristic of this system is that the detector is uncooled. Thus, no vibrations from the system disturb the measurements as it was found for the cooled system. In order to assure approximately the same spatial resolution, the HgCdTe camera was equipped with 100-mm focal length optics, resulting in an instantaneous field of view (IFOV) of 0.4 mrad and the InGaAs camera with 35 mm focal length optics giving a 0.428 mrad IFOV. The selected optics and camera parameters enabled a comparable image acquisition for both systems.

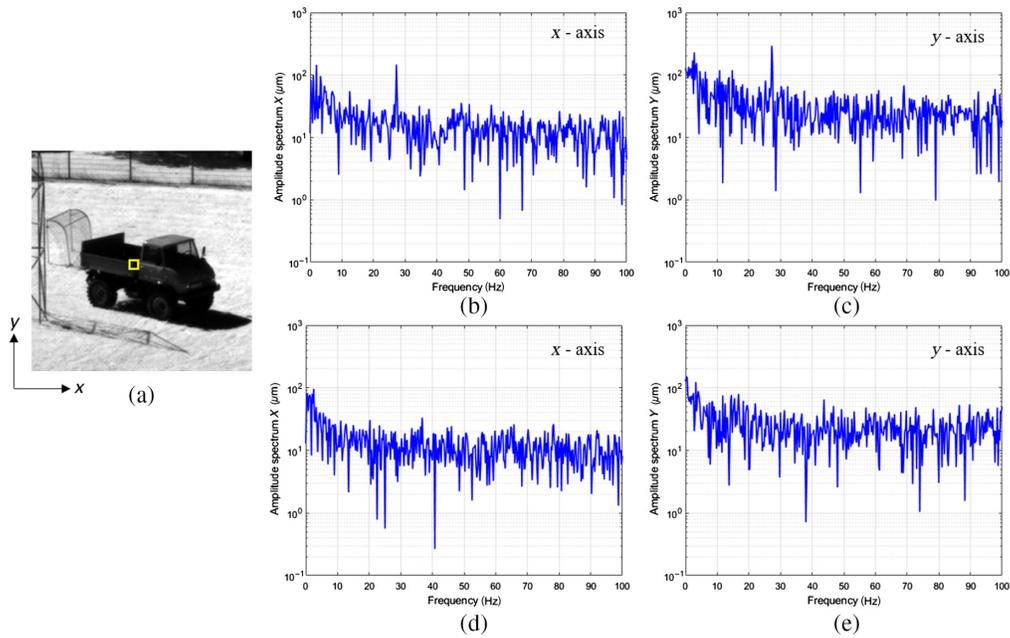


Fig. 12 Infrared image of the small truck recorded with the InGaAs camera in the spectral range 0.9 to $1.7 \mu\text{m}$ (a) and the vibrational spectra acquired from the position marked by the yellow square with (b), (c) engine on and (d), (e) engine off.

4.2 1D Vibration Signature

The method described in Sec. 2 may be used to determine the vibration spectrum from a selected region in the image frames—so called one-dimensional (1D) vibration signature—but also to generate 2D vibration signatures at specified frequencies.³ The technique was applied to various image sequences recorded with the two cameras presented in Table 1. The vibration of the small truck with engine idling could be detected for all investigated spectral ranges. Using the HgCdTe camera, a 1-s recording was sufficient to obtain a good signal-to-noise-ratio of the vibration line, whereas for the InGaAs camera, a 4-s sequence was necessary in order to achieve an appropriate signal. An example for the 1D vibration signature is shown in Fig. 12. The images have been recorded with the InGaAs camera in the spectral range 0.9 to $1.7 \mu\text{m}$ at a frame rate of 700 Hz with a resolution of 256×256 pixels. The pixel displacements along the x and y axes were calculated for every image in a series of 4 s (2800 images) for a subset of 7×7 pixels, marked by the yellow square in Fig. 12. The vibration frequency of 26.25 Hz could be detected for both axes [Fig. 12(b) and (c)] as is obvious in comparison to the same measurements with the engine off [Fig. 12(d) and (e)].

4.3 2D Vibration Signature

Apart from the vibrational spectrum of a target, i.e., the 1D vibration signature, the high-speed recordings offer the possibility to generate spatially resolved 2D vibration signatures. The IR cameras described in Table 1 have been equipped with various filters transmissive in spectral bands between 0.9 and $5.3 \mu\text{m}$, covering both reflective and emissive dominated regimes. For each investigated spectral range, several image sequences of the target have been recorded at different times after the engine was turned on. The 2D vibration signature was extracted using 1-s recordings with the MWIR camera working at 884 Hz and full resolution of 256×256 pixels and 4-s recordings from the SWIR camera working at 700 Hz with a 256×256 pixels resolution.

Figures 13 and 14 display the recorded intensity images (a) and the corresponding extracted 2D vibration signatures (b) for the investigated spectral ranges. All 2D vibration signatures have a resolution of 63×63 pixels. A subset of 7×7 pixels was considered for each investigated pixel. The displayed 2D vibration signatures present the intensities of the engine main frequency as a combination of vibration amplitudes along the horizontal and vertical image axes. The

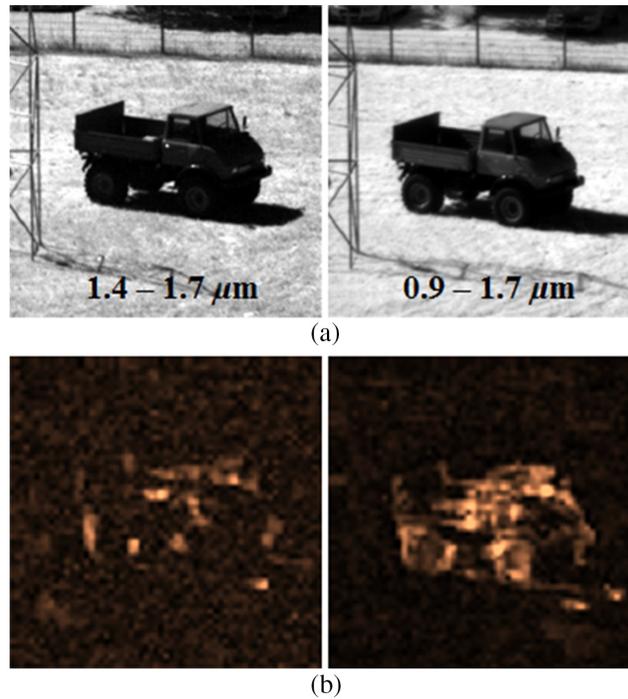


Fig. 13 (a) Sample images and (b) corresponding 2D vibration signatures of the idling engine small truck recorded with the InGaAs camera in two spectral ranges at the engine main frequencies of 27.25 and 26.25 Hz, respectively.

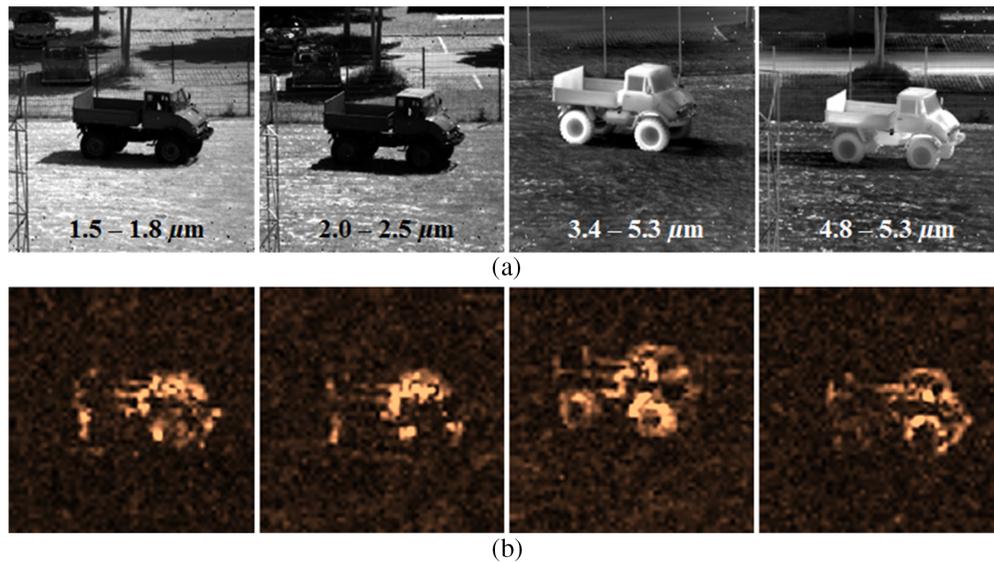


Fig. 14 (a) Sample images and (b) corresponding 2D vibration signatures of the idling engine small truck recorded with the HgCdTe camera in various spectral ranges at the engine main frequencies between 27 and 29 Hz.

engine main frequency differs slightly for different measurements because the revolution per minute (rpm) of the engine was not absolutely reproducible.

The images show that the 2D vibration signature reproduces the target features in all investigated spectral ranges. However, with a narrow spectral band ($\sim 0.3 \mu\text{m}$) in the reflective dominated regime, a very weak signal was obtained, as displayed in Fig. 13(a). A clearly higher signal was measured for recordings with a broader spectrum of 0.8 and 0.9 μm , in both reflective dominated (0.9 to 1.7 μm , Fig. 13) and emissive dominated (3.4 to 5.3 μm , Fig. 14) regimes, respectively.

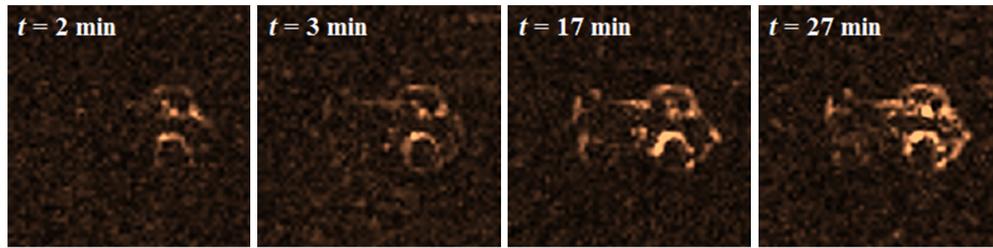


Fig. 15 2D vibration signature of the idling engine small truck recorded with the HgCdTe camera in the spectral range 4.8 to 5.3 μm for different engine run time durations.

The measurements in the spectral range 0.9 to 1.7 μm reveal an interesting feature such as the two points on the ground caused by the mirrors' shadows [Fig. 13(b)]. The high-speed camera technique is sensitive to any contrast changes. Therefore, other features could be incorrectly detected as vibrational features belonging apparently to the target, e.g., the shadow of the vibrating mirrors on the ground. Every defined movement of image pixels—regardless of the cause—may lead to a contribution to the vibrational image.

4.4 Robustness of the Vibration Signature

Recovery of vibration information from high-speed video recordings depends on the image contrast and how detailed the image is (Sec. 3.1.5). Reflective as well as emissive target signatures could change by different factors. The emissive signatures change by the various environmental influences during the day (i.e., air temperature, moisture, wind speed, and heat loading by the Sun), but also by the operating condition of the target. Figure 15 shows the change of 2D vibration signatures of the small truck with engine idling with increasing time after engine start. The HgCdTe camera was equipped by a filter open for the spectral range from 4.8 to 5.3 μm in the emissive dominated spectral region. The 2D vibration signature is improved with engine operating time caused by the increasing image contrast.

5 Influence of Atmospheric Turbulence

The results in Secs. 3.2, 4.3, and 4.4 demonstrate a very good extraction of the 2D vibration signature of an object placed outdoors at distances of 65 to 67 m from the sensors. In the corresponding experimental configuration, the line of sight goes from 15 m altitude (sensors position) to ground level (object). However, it may be expected that for longer distances, the atmospheric turbulence might deteriorate or even destroy the recovery of the vibration signal. An experiment has been performed where maintaining the same distance sensor object, the turbulence strength was increased by the use of a heating plate.

5.1 Experimental Arrangement

The experimental arrangement is displayed in Fig. 16. The InGaAs camera equipped with the 0.9- μm high-pass filter was used to record images of the Unimog target located 65 m from the camera. The heating plate was placed in front of the imaging optic. Various turbulence strength conditions have been created by three measurement scenarios: (1) natural atmospheric turbulence with the heating plate turned off, (2) low extra turbulence at <1 min after the heating plate was turned on, and (3) high additional turbulence after about 8 min of heating with the plate. A series of images of the vibrating object recorded under these conditions has been used to extract the 2D vibration signature, but also to estimate the refractive index structure parameter C_n^2 during the measurements.

5.2 Estimation of Turbulence Strength

The turbulence strength has been calculated using the recorded images. The angle-of-arrival fluctuations variance σ_a^2 is linearly dependent on the refractive index structure parameter

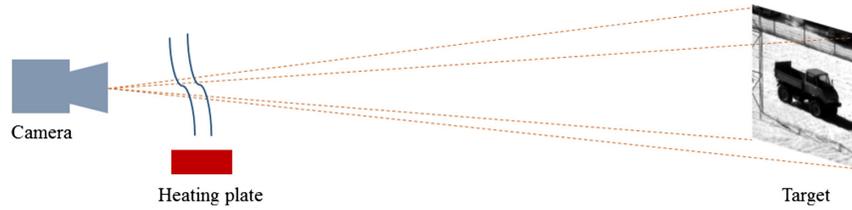


Fig. 16 Experimental setup used to investigate the influence of strong local turbulence on the 2D vibration signature.

C_n^2 .^{15,16} The explicit dependency is given, for example, by Eq. (83) in Chapter 6 of Andrews and Phillips¹⁵ or Eq. (4) of Zamek and Yitzhaky¹⁶

$$\sigma_\alpha^2 = 2.914 \times D^{-1/3} \begin{cases} \int_0^L C_n^2(z) dz & l_0 \ll D \ll \sqrt{\lambda L} \text{ plane wave} \\ \int_0^L C_n^2(z) (\frac{z}{L})^{5/3} dz & \sqrt{\lambda L} \ll D \ll L_0 \text{ spherical wave} \end{cases} \quad (7)$$

Here, D is the optic aperture, L is the optical path, l_0 and L_0 are the turbulence inner- and outer-scale, and λ is the wavelength of the electromagnetic radiation. The condition $l_0 \ll D \ll \sqrt{\lambda L}$ corresponds to the plane wave approximation, whereas the case $\sqrt{\lambda L} \ll D \ll L_0$ complies with the spherical wave approximation.

To determine the angle-of-arrival fluctuations, a region of interest outside the object was selected, containing sufficient intensity variations. The procedure described in Sec. 2 was applied to detect the pixel displacement in comparison with the first image and the variance was calculated. As the InGaAs camera is not cooled, the angle-of-arrival fluctuations are mainly caused by the air turbulence, but not disturbed by camera vibrations. The angle-of-arrival variance σ_α^2 is obtained by multiplying the image displacement single-axis variance σ_{img}^2 with the squared IFOV

$$\sigma_\alpha^2 = \sigma_{img}^2 \times \text{IFOV}^2. \quad (8)$$

To prove the reliability of the C_n^2 calculations, another experiment has been performed in which the turbulence strength was measured simultaneously by a commercially available scintillometer Scintec BLS 900 and by the angle-of-arrival fluctuations method. The scintillometer recorded the C_n^2 values for every minute, whereas the InGaAs camera recorded a 4-s image sequence every 10 min. Figure 17 shows the calculated values of the C_n^2 in the plane wave approximation (blue line with square symbols) and spherical wave approximation (red line with circle symbols) in comparison with the scintillometer measurements (black line). The C_n^2 values

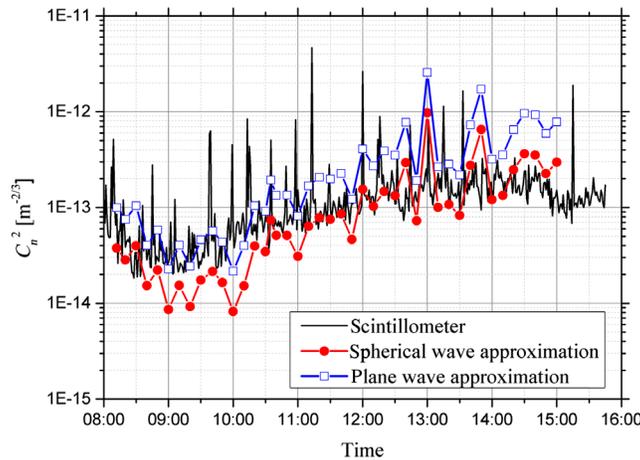


Fig. 17 The structure parameter C_n^2 of the refractive index measured by the Scintec BLS 900 scintillometer and by the angle-of-arrival-fluctuations method.

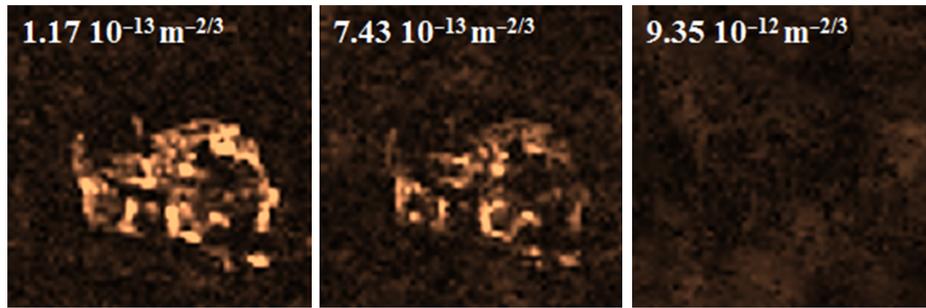


Fig. 18 2D vibration signature of the idling engine small truck recorded with the InGaAs camera in the spectral range 0.9 to 1.7 μm under different turbulent conditions. The inserted values give the determined C_n^2 .

have been calculated from Eq. (7) with the assumption of a constant C_n^2 along the optical path. In spite of several drawbacks of the method, e.g., the sensitivity to the camera movements, the results are in very good agreement with the values delivered by the scintillometer.

5.3 Measurements

The 2D vibration signatures under different turbulent conditions are displayed in Fig. 18. The increase of the path-averaged C_n^2 by a factor of about seven somehow disturbs the extracted signal, but the target features are still clearly reproduced. Under very strong turbulence conditions, with the calculated path-averaged C_n^2 around $10^{-11} \text{ m}^{-2/3}$, no vibration signal could be detected.

6 Discussions

The high-speed imaging technique applied in this study yielded reliable measurements of small-amplitude vibrations. Compared with the laser Doppler vibrometry, the technique reveals both advantages and drawbacks, which will be discussed further. An overview of the main features of the investigated methods is given in Table 2. One benefit of the high-speed imaging method is that it is a passive technique that does not require an additional laser illumination of the object. This may be an advantage, especially for medium and long-range applications, where high-laser

Table 2 High-speed camera technique compared with laser vibrometry (scanning).

Technique	Laser vibrometry	High-speed imaging
System operation	Active	Passive
Operation time	Day/night	Day + night (MWIR)
Acquisition time (same pixel number)	$\sim 1 \text{ h}^{\text{a}}$	$< 1 \text{ s}$
Processing time	Below 1 s	Tens of minutes or hours
Vibration component	Along the line of sight	Perpendicular to the line of sight
Object structure	No contrast required	High contrast required
Range	$> 1 \text{ km}^{\text{b}}$	Several 100 m
Sensitivity	Several nm	Several μm
Simultaneously measured points	No	Yes

^aDepends on the number of measured points, typical value 1 s/point.

^bIn principle suitable for long ranges, but ensuring laser safety becomes a serious issue.

output power may be needed for Doppler vibrometry. To comply with laser safety regulations, the required laser power can be a critical issue.

On the one hand, a high-speed camera requires very good illumination of the scene in the visible spectral range. On the other hand, the image adjustment algorithm requires high-contrast structures on the object surface, which are associated with local intensity variations across the target. Because of this, measurements at night as well as in poor light conditions are not really feasible with a visual camera. This drawback may be overcome using cameras in SWIR and especially MWIR spectral ranges (Sec. 4). The experiments demonstrated that the vibration signature could be extracted from image sequences recorded in both reflective and emissive dominated spectral ranges. It should be noted that the emissive structure of an object can change over the course of the day or depends on the operating condition and thus influences the quality of the vibration feature extraction.

While the laser vibrometer detects the radial component of the object motion in the line-of-sight direction, the high-speed imaging method detects the tangential component. This leads to a sinusoidal dependence on the angle of incidence between the line-of-sight (laser beam direction) and the object surface in the case of laser vibrometry and a cosine dependence on the angle between the line-of-sight and the object surface in the other case.

LDVs have demonstrated a high sensitivity for a very broad frequency range. The high-speed imaging approach used in this study showed a sensitivity at least one order of magnitude lower than the laser vibrometer's. For diffuse reflecting objects, however, the minimum detectable vibration amplitude can be reduced due to the additional noise contribution caused by speckle. The accessible frequency range is limited by the camera frame rate or by the large rate of data acquisition needed for a high frame rate.

With appropriate laser power, a vibrometer could be used for short, medium, and long-range applications.^{1,2} The experiments presented in Secs. 3 and 4 show the image-based method yields consistent measurements at around a 70-m range. By an appropriate selection of the imaging optics, longer range measurements should be possible. However, the experiments described in Sec. 5 show that high strength of air turbulence over a long distance will become a limiting factor.

One major advantage of the high-speed imaging method concerns spatially resolved vibration measurements. For laser vibrometry, spatial resolution may be achieved using a scan device or a multielement receiver. However, most common laser vibrometers scan an object point by point, and therefore, require a long acquisition time. In contrast, a single high-speed camera recording under 1 s may be used to extract the vibration amplitude at different positions. Moreover, the simultaneous measurement of all points allows the reconstruction of vibration modes.³ A laser vibrometer requires an additional reference measurement to synchronize the measured points. In recent years, several types of multipoint or full-field LDVs have been reported but only for short ranges.^{17,18} By measuring vibrations at different object points simultaneously, such systems considerably lower the acquisition time and still can measure transient events. This is in contrast with "stop-and-stare" systems or continuously scanning systems. On the other side, a drawback for the high-speed imaging method is the high processing time required by the image matching technique.

7 Conclusions

The high-speed imaging technique for vibration and displacement measurement was experimentally investigated and its potential for long-range applications evaluated. Short-range measurements in the VIS spectral range have been performed in comparison with laser Doppler vibrometry. The basis for the methods comparison was the investigation of a vibrating loudspeaker membrane. The experiments showed that the measured vibration frequency and amplitude were consistent with the values obtained from the laser vibrometer. However, the detection threshold was over one order of magnitude higher than for the laser vibrometer.

The applicability of the method to medium ranges up to 100 m was confirmed by measuring the vibrations of a small truck at idle. The medium-range measurements using the high-speed imaging technique provided reliable results for the characteristic frequencies, as well as for the

spatially resolved vibration signature. In spite of lower sensitivity, the method is advantageous because of synchronous recording of different points and thus a significantly lower acquisition time.

The high-speed imaging technique for vibration measurement was also experimentally investigated in the SWIR and MWIR spectral range. Measurements at a range up to 65 m have been performed using the small truck as a test object. The vibrations of the idling engine have been detected from infrared images recorded in different spectral bands between 0.9 and 5.3 μm . The spatially resolved vibration signature could be extracted from image sequences recorded in both reflective and emissive dominated spectral ranges.

The influence of strong local air turbulences on the vibration signature was studied experimentally for the spectral range 0.9 to 1.7 μm . In the test setup, it was found that only turbulences with uncommonly high values of the refractive index structure parameter completely destroy the vibration information. Such conditions may be found when doing measurements in the vicinity of jet engines, etc. This result indicates that under natural turbulence conditions, the method may be extended to much longer ranges than the ones used here. Future work concerns measurements comparing the influence of atmospheric turbulence at different optical wavelengths, as well as experiments on the impact of atmospheric turbulence on the LDV.

The procedure implemented for the vibration recovery was also used to extract the angle-of-arrival fluctuations from the recorded image sequences and to calculate the refractive index structure parameter. Simultaneous turbulence measurements with a scintillometer showed very good agreement with the determined values.

Acknowledgments

The authors would like to acknowledge the financial support by the “Bundeswehr Technical Center for Weapons and Ammunition” (WTD 91) in Meppen. The authors express their thanks to Frank Willutzki, Norbert Hannas, and Simon Brunner for their tireless assistance and support in carrying out the numerous experiments cited in this report, to Chris Hill (Malvern Lidar Consultants) and Uwe Adomeit for helpful discussions, and to Eric Sucher for support in carrying out the turbulence measurements. The authors declare no conflicts of interest regarding the publication of this article.

References

1. P. Lutzmann et al., “Laser vibration sensing at Fraunhofer IOSB: review and applications,” *Opt. Eng.* **56**(3), 031215 (2016).
2. “Non-contact vibration measurement,” <https://www.polytec.com/uk/vibrometry/>.
3. A. Davis et al., “The visual microphone: passive recovery of sound from video,” *ACM Trans. Graphics (Proc. SIGGRAPH)* **33**(4), 1–10 (2014).
4. Z. Wang, H. Nguyen, and J. Quisberth, “Audio extraction from silent high-speed video using an optical technique,” *Opt. Eng.* **53**(11), 110502 (2014).
5. D. Zhang et al., “Efficient subtle motion detection from high-speed video for sound recovery and vibration analysis using singular value decomposition-based approach,” *Opt. Eng.* **56**(9), 094105 (2017).
6. B. Pan, H. Xie, and Z. Wang, “Equivalence of digital image correlation criteria for pattern matching,” *Appl. Opt.* **49**(28), 5501–5509 (2010).
7. D. Mas et al., “Measurement of wide frequency range structural microvibrations with a pocket digital camera and sub-pixel technologies,” *Appl. Opt.* **51**(14), 2664–2671 (2012).
8. D. Mas et al., “Image based subpixel techniques for movement and vibration tracking,” in *11th Eur. Conf. Non-Destruct. Testing (ECNDT 2014)*, Prague (2014).
9. Y.-N. Jeng and C.-H. Wu, “Frequency identification of vibration signals using video camera image data,” *Sensors* **12**, 13871–13898 (2012).
10. B. Ferrer et al., “Vibration frequency measurement using a local multithreshold technique,” *Opt. Exp.* **21**(22), 26198–26208 (2013).

11. M. Akutsu, Y. Oikawa, and Y. Yamasaki, "Extract voice information using high-speed camera," *Proc. Meetings Acoust.* **19**, 055019 (2013).
12. G. Zhang et al., "A high-speed vision-based sensor for dynamic vibration analysis using fast motion extraction algorithms," *Sensors* **16**(4), 572 (2016).
13. G. Paunescu et al., "Comparison of high speed imaging technique to laser vibrometry for detection of vibration information from objects," *Proc. SPIE* **9649**, 96490D (2015).
14. G. Paunescu and P. Lutzmann, "Detection of object vibrations from high speed infrared images," *Proc. SPIE* **9988**, 998809 (2016).
15. L. C. Andrews and R. L. Phillips, *Laser Beam Propagation through Random Media*, SPIE Press, Bellingham, Washington (2005).
16. S. Zamek and Y. Yitzhaky, "Turbulence strength estimation from an arbitrary set of atmospherically degraded images," *J. Opt. Soc. Am. A* **23**(12), 3106–3113 (2006).
17. J. Kilpatrick and V. Markov, "Multi-pixel (matrix) laser vibrometer," in *OSA Meeting, Appl. of Lasers for Sens. and Free Space Commun. (LS&C)* (2010).
18. "MPV-800 multipoint vibrometer," https://www.polytec.com/fileadmin/d/Vibrometrie/OM_DS_MPV-800_E_42389.pdf (2020).

Gabriela Paunescu studied physics at the University of Bucharest and received her PhD in physics from Friedrich Schiller University of Jena in 2006. After PhD studies, she worked on femtosecond laser atmospheric propagation at German Aerospace Center (DLR), Stuttgart, Germany. She joined the Optronics Department of Fraunhofer IOSB in 2012. Her current activity focuses on active sensors technologies and applications.

Peter Lutzmann received his diploma degree in physics from the University of Ulm in 1985. In 1985, he joined the Optronics Department, Research Institute for Optics (FFO), Tübingen, Germany. FFO became FOM in 2000 and Fraunhofer IOSB in 2010. Since 2010, he has been the leader of the Laser Sensors Group. His research area comprises coherent laser radar and laser imaging techniques. He is a member of several bilateral and NATO working groups.

Daniel Wegner received his MS degree in physics from Karlsruhe Institute of Technology, Germany, in 2012. He has been a research assistant at Fraunhofer-Institute IOSB since 2014. His current research interests are imaging simulation of degradational effects occurring in optronic systems and, in particular, imaging techniques for atmospheric turbulence effects. The latter can have significant impact on image quality for remote sensing applications. Another field of interest is image-based imager assessment by machine learning techniques.

Benjamin Göhler received his diploma degree in mathematics from the Technical University of Karlsruhe in 2007. In 2007, he joined the Optronics Department at the Research Institute for Optronics and Pattern Recognition (FOM), Ettlingen, Germany. In 2010, FOM became Fraunhofer IOSB, where he is working in the Laser Sensors Group. His research area covers modeling of laser systems and processing of field trial data. He is a member of several bilateral and NATO working groups.