Received: 8 October 2022 | Accepted: 11 November 2022

DOI: 10.1002/pamm.202200274

Optimization of Mirror Positioning for Single-Camera 3D Displacement Measurements

Max Gille[1,*], **Johannes Maierhofer**[1], and **Daniel J. Rixen**[1]

[1] Chair of Applied Mechanics, Faculty of Mechanical Engineering, Technische Universität München, Boltzmannstr. 15, 85748 Garching

Camera-based 3D displacement measurements are only possible when two different viewing angles of the respective measurement point are available. We earlier introduced a simplified 3D reconstruction algorithm for camera-mirror measurements where the mirror reflects the second view onto the camera sensor to eliminate the need for a second (costly) high speed camera. In this study, we concentrate on formulating guidelines on how to position the mirror with respect to the camera and the measured structure, the goal being to minimize errors arising in the 3D reconstruction procedure. The focus lies on the angle between the optical axes of both camera views. We hereby also take into account limitations in the experimental setup that the special camera-mirror-setup brings with it. This study is conducted in terms of simulations using the animation software BLENDER as well as real experiments on a simple vibrating structure.

© 2023 The Authors. *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH GmbH.

1 Introduction

Measuring displacements with cameras has become a relatively common technique as digital camera sensors improved over the past two decades. While using a camera is a contact-less technique and enables full-field measurements, there are also some drawbacks that need to be taken into account when deciding how to measure vibrations of a given system. One common problem, compared to more traditional measurement techniques such as piezo-electric accelerometers or Laser-Doppler Vibrometry (LDV), is the higher noise level of image sensors. Also, by construction, a usual image sensor is only a two-dimensional measuring device and therefore 3D measurements are not possible right away.

One optionto enable 3D vision is the use of a second camera. While this is a common strategy as described by Baqersad et. al. [1] which is widely available as commercial products combining two-camera setups and the software, it is, however, quite expensive. Especially in use cases where very high frame rates and a low noise level are crucial, suitable cameras are available only in the price range of some tens of thousands of Euros.

So, one might think of a much cheaper solution involving a mirror that reflects a second image of an object onto the image sensor of the available camera. Researchers have worked with similar setups, often consisting of multiple mirrors arranged on an optical table, and proven their functionality, especially for 3D DIC applications, such as Yu and Pan [8] with a four-mirror setup or Genovese et. al. [3] who used a biprism to split the camera view.

Our approach only utilizes one single mirror whose position is not fixed relative to the camera, but can be moved freely. Additionally, we present a simplified 3D reconstruction with some limitations that doesn't need a calibration procedure with several different images of a known pattern, but instead relies on just one image, possibly taken from the measurement recording itself. This simplification comes at the cost of some additional errors as we assume that all light rays reach the camera in parallel, which means that we assume an orthographic camera.This seems to be quite a bolt assumption, but it proved to yield satisfying results in an earlier study in our lab [4].

This work introduces the reconstruction strategy for a freely positioned mirror and experimentally tries to find practical mirror and camera angles for good measurement results.

2 Simplified Reconstruction

2.1 Motivation

Reconstruction of 3D information based on two separate 2D views of one object is typically done by means of triangulation. Software, e.g. in Matlab or Python, is readily available, but involves calibrating the two cameras. This calibration procedure starts with taking several still images of a known reference pattern (usually a checkerboard pattern of given dimensions). Using the gathered information, it is then possible to calculate the cameras' relative position and angle as well as the so called internal parameters like focal length etc.

With our single mirror setup, we are generally very limited in terms of field of view such that the very step of taking several pictures with different configurations of the reference pattern is a first challenge that is not easy

* Corresponding author: e-mail max.gille@tum.de, phone +49 089 289 15212

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

to tackle. Therefore, we propose a simplified approach that assumes an orthogonal projection onto the camera sensor, i.e. the assumption that the light reaching the camera sensor in parallel. This assumption gets better and better as the used focal length – and with it the distance between camera and object – increases.

2.2 The Algorithm

As mentioned in the previous lines, we use an orthogonal projection assumption to circumvent the full camera calibration usually necessary. This leaves only the relative rotations of the cameras (α, β, γ around the x, y and z axes in that order) with respect to some global coordinate system, and a scaling factor κ (in the unit px/mm) to be determined. For this, it is sufficient to have one still frame of some reference pattern along whose edges we define our global coordinate axes. Knowing the representation of just two axes in an image, we can then find the sensor's rotation relative to these axes.

$$ {}_C\boldsymbol{x}_W = \kappa\, {}_C\boldsymbol{A}_W\, {}_W\boldsymbol{x}_W = \kappa\, {}_C\boldsymbol{A}_W \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \kappa\, {}_C\boldsymbol{A}_W[:,1], \tag{1} $$

$$ {}_C\boldsymbol{y}_W = \kappa\, {}_C\boldsymbol{A}_W\, {}_W\boldsymbol{y}_W = \kappa\, {}_C\boldsymbol{A}_W \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \kappa\, {}_C\boldsymbol{A}_W[:,2] \tag{2} $$

using Matlab-like indication of matrix entries in square brackets and where the left index of a vector stands for the coordinate system in which the vector is measured and the rotation matrix ${}_C\boldsymbol{A}_W$ changes the vector's representation from system W to C. So, with a measurement of the world's x and y axes in the sensor coordinate system, the four top left entries of the rotation matrix are known with a common factor of κ. As we can only identify those parts of the reference vectors that are in the $x-y$ plane of the camera sensor, the third line of ${}_C\boldsymbol{A}_W$ will never be measurable. This gives us the following equations for the four unknown parameters α, β, γ, and κ:

$$ \begin{aligned} a_1 &= \kappa(\cos\beta\cos\gamma), & b_1 &= \kappa(\cos\alpha\sin\gamma), \\ a_2 &= \kappa(-\cos\beta\sin\gamma), & b_2 &= \kappa(\cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma) \end{aligned} $$

with a_1, a_2, b_1 and b_2 the measured length of the world's axes along the sensor directions.Finding the parameters requires solving this nonlinear system of equations which is currently done with the function `fsolve` from the optimization toolbox in Scipy [7]. The big problem is, that the result of the used algorithm is highly dependent on the initial guess provided to the algorithm and often becomes some local minimum instead of the actual correct rotation. Therefore, until now, this step requires manual work to check the result for plausibility and if needed start the algorithm again with different initial guesses for the angles. With that, the correct one of two solutions to problem can be chosen and from now on, all parameters of the experimental setup are known. The raw displacement measurements in camera coordinates ${}_C\boldsymbol{\Delta}$ are known in ${}_Cx$ and ${}_Cy$ direction while displacements along ${}_Cz$ are unknown as of now.

$$ \kappa\, {}_W\boldsymbol{\Delta} = {}_W\boldsymbol{A}_{C_1}\, {}_{C_1}\boldsymbol{\Delta}_1 \overset{!}{=} {}_W\boldsymbol{A}_{C_2}\, {}_{C_2}\boldsymbol{\Delta}_2. \tag{3} $$

By solving first the right part of the equation above, the displacements along the ${}_Cz$ axes can be determined and based on that, the true displacements in the world coordinates are easily computed with the left part of the equation. For the first part, solving for the unknown displacements along the camera's z axis, the overdetermined system of equations needs to be solved:

$$ \underbrace{\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}}_{{}_W\boldsymbol{A}_{C_1}} \begin{pmatrix} \Delta x_{C_1} \\ \Delta y_{C_1} \\ \Delta z_{C_1} \end{pmatrix} \overset{!}{=} \underbrace{\begin{pmatrix} j & k & l \\ m & n & o \\ p & q & r \end{pmatrix}}_{{}_W\boldsymbol{A}_{C_2}} \begin{pmatrix} \Delta x_{C_2} \\ \Delta y_{C_2} \\ \Delta z_{C_2} \end{pmatrix} \tag{4} $$

which leads to the following equation:

$$ \begin{pmatrix} c & -l \\ f & -o \\ i & -r \end{pmatrix} \begin{pmatrix} \Delta z_{C_1} \\ \Delta z_{C_2} \end{pmatrix} = \begin{pmatrix} j & k \\ m & n \\ p & q \end{pmatrix} \begin{pmatrix} \Delta x_{C_2} \\ \Delta y_{C_2} \end{pmatrix} - \begin{pmatrix} a & b \\ d & e \\ g & h \end{pmatrix} \begin{pmatrix} \Delta x_{C_1} \\ \Delta y_{C_1} \end{pmatrix} \tag{5} $$

The left matrix is obviously built up from the third columns of the two rotation matrices side-by-side, which represent the respective camera's z axis. So, the Moore-Penrose pseudoinverse used to solve for $\Delta z_{C_{1,2}}$ can be

© 2023 The Authors. *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH GmbH.

written as

$$\left(\begin{bmatrix} {}_W\boldsymbol{z}_{C_1}^T \\ -{}_W\boldsymbol{z}_{C_2}^T \end{bmatrix}\begin{bmatrix} {}_W\boldsymbol{z}_{C_1} & -{}_W\boldsymbol{z}_{C_2} \end{bmatrix}\right)^{-1} = \begin{pmatrix} 1 & -\cos\psi \\ -\cos\psi & 1 \end{pmatrix}^{-1} \tag{6}$$

with the angle between the both cameras' z axes ψ, i.e. the the angle between the optical axes. Obviously, the matrix that needs to be inverted gets ill-conditioned as the cameras are more and more aligned, so it seems desirable to have the two sub-images as different from each other as possible. For example, this condition number drops from ≈ 130 to ≈ 32 as ψ is increased from 10° to 20°. Moreover, for collinear cameras and in the absence of noise the right hand side of eq. (5) reaches exactly zero (both cameras deliver the same image), so there should be more useful signal when the two views differ significantly.

3 Simulation and Experiment

In the last section, we saw that from a numerical point of view, a larger relative viewing angle between the two views is preferable. This observation shall now be tested, both in simulated environment free of disturbances with the rendering software BLENDER [2] and via actual experiments with a real structure and a camera.

3.1 Experimental Setup

The experimental setup comprises a brass structure that consists of three solid parts soldered together. The main movements of the structure, however, only come from motions of the plate-like top which has a relatively soft connection to the rest of the structure. This brass structure is firmly screwed to a table and excited with a force hammer commonly used for experimental modal analysis. We equipped the plate with two accelerometers at the bottom of the top plate to have reference measurements while the plane's top surface was prepared with a checkerboard pattern to enable optical measurements, see fig. 1 .

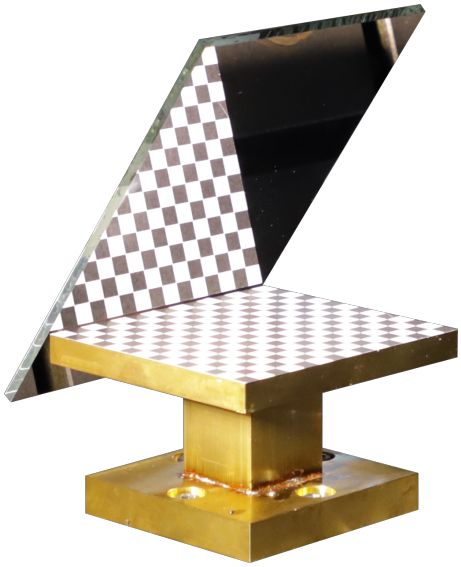

Fig. 1 Example structure with checkerboard pattern on top. The mirror angle φ is roughly 50°. Two accelerometers are fixed to the bottom of the plate, one in the middle of far side close to the mirror and one in the far left corner. The exitation is achieved with a single hit with an impulse hammer which triggers the acceleration measurement.

The camera used for this study is the Photron Fastcam Nova S6 with 8 GB of memory and 1024 × 1024 pixel resolution. It has a monochrome 12 bit sensor. The light source is a 360 W LED that provides 36 000 lm. A first-surface mirror is used to provide the second view.

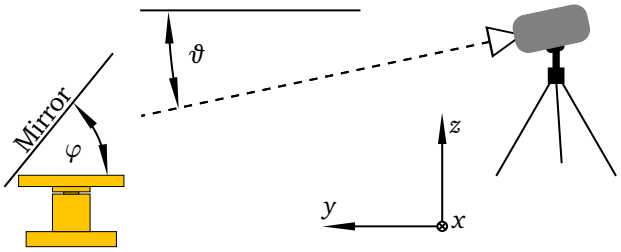

Fig. 2: Schematic side-view of the experimental setup with mirror and camera angles. The height of the camera was manually chosen such that the structure and its mirrored image were roughly in the middle of the image.

In this study, six different camera/mirror configurations as depicted in fig. 2 were investigated with mirror angles from 15° to 45° and camera angles between 0° to 30° according to table 1. For both simulation and

© 2023 The Authors. *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH GmbH.

		φ		
		15°	30°	45°
ϑ	0°	1	2	4
	10°		3	5
	30°			6

Table 1 Configurations tested in this study and their respective numbers used for referencing a configuration.

experiment, the plate's apparent 2D displacements in the sensor plane are found by applying the Lucas-Kanade method [5] as implemented in the Python package pyIDI [6] to the recordings at the observed points.

3.2 Simulation

At first, the example structure that would be later used in the lab, was modeled in BLENDER. There, a rectangular plane with a checkered surface was placed in front of a camera with the same optical parameters as the real one in terms of sensor size, pixel count and focal length. The plane is mirrored by a second plane representing the mirror and the angle between the measurement surface and the mirror plane is varied as well as the inclination and position of the camera. The plane was then moved harmonically with known amplitude and frequency along all three coordinate axes such that a perfect reference was available with the following trajectories:

$$\Delta\{x,y,z\}(i_{\text{frame}}) = 0.05\,\text{mm} \cdot \sin(f_i \cdot i_{\text{frame}}) \quad \text{with } f_i = \{0.5, 0.8, 1.2\} \text{ for } \{x,y,z\}$$

where i_{frame} is the current frame number, starting at $i_{\text{frame}} = 0$. In BLENDER, everything is animated with frames as the unit of time and they are only translated into a real time as soon as an animation is rendered and exported at a specific frame rate. For this application, the actual playback speed doesn't play a role as we only compare the simulated optical measurements with the actual displacement frame by frame, no matter the real time. The results for such a reconstruction, compared to the reference signal is shown in fig. 3 as an example for the configuration $\varphi = 30°$, $\vartheta = 10°$. Obviously, the reconstruction is not perfect, however the overall kinematics

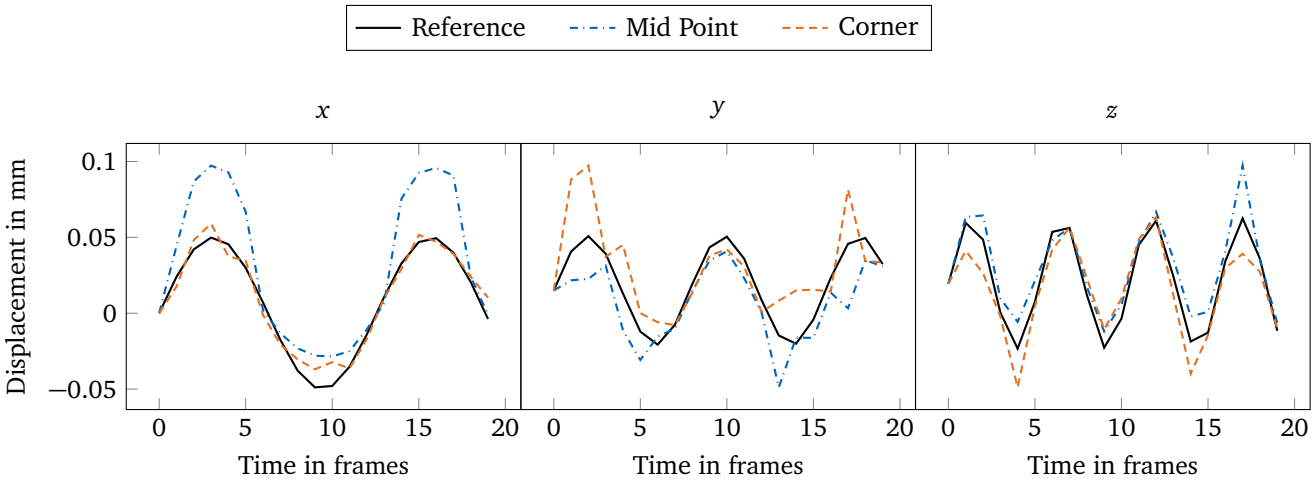

Fig. 3: Comparison of displacement measurements in all three directions for the rendered video. The reference is the known truth that was prescribed in BLENDER for the animation. To cover the influence of the measurement's position, both the plate's mid point as well as a corner were used to optically measure the displacements. The mirror/camera configuration was $\varphi = 30°$, $\vartheta = 10°$.

are preserved and the movement of the structure is captured. Despite apparent problems in the amplitude of the measurement, it is obvious that the reconstruction technique is able to separate the three directions of the plate's movement when looking at the frequencies of the measured signal. Instead of plotting the time series for all six configurations, we now compare a global measure for the error of the respective measurements compared to the reference. For this, we choose the relative root mean square error e:

$$e = \sqrt{\frac{\sum(\Delta_{\text{meas}} - \Delta_{\text{ref}})^2}{\sum \Delta_{\text{meas}}^2}} \tag{7}$$

where each summation symbol sums over all frames of the recording. We then get an error for each configuration and each measurement position (middle/corner).

fig. 4 shows the average of all relative RMS errors corresponding to each configuration. We can see two things: firstly, the values are all in the same range such that from this comparison it is hard to give meaningful

© 2023 The Authors. *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH GmbH.

advice concerning a preferable configuration. Secondly, the values seem pretty high. However, as even the very good behavior of the orange line in the left-most subplot of fig. 3 has a relative RMS error of over 0.2, we are not overly concerned by the seemingly high errors.

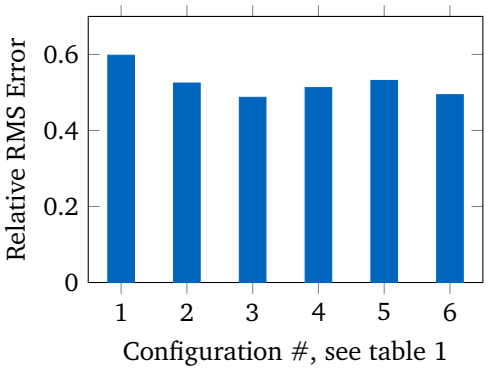

Fig. 4 Relative root mean square errors for different configurations averaged over all directions and both measurement points for the simulated measurement.

3.3 Experiment

Additionally to the simulated experiments, real measurements of the structure mimicked by the blender model, were conducted.

The brass structure mentioned before was impacted in a front corner with a force hammer. Its measured output was only used for triggering the actual acceleration measurement recorded with two piezo-electric transducers. These accelerometers were mounted on the bottom side of the top plate on the far edge – one in the middle, one in the left corner.

We assume that the movements on both sides of the top plate are more or less identical such that the acceleration signal can be used as a reference for the displacements on the top side with negligible error. Therefore, the acceleration time signal is integrated twice using the trapezoidal rule which yields a displacement reference against which the optical measurements can be tested. Both signals, from the camera and the accelerometers are band-pass filtered before the comparison to remove a drift coming from the time integration as well as to clean the signal from any high-frequency components beyond the recordings' bandwidth. Also, as the camera and our LMS measurement system don't offer the same sampling frequency in the range of interest, the acceleration-based displacements are resampled such that both signals have a sampling frequency $f_S = 2000\,\mathrm{Hz}$ and thus can be compared in terms of the mentioned RMS error. However, the respective beginning of a measurement had to be moved by hand as LMS and the camera are not triggered simultaneously. As the signal mainly contains vibrations at around 400 Hz (all three observable resonance frequencies are located in the between 375 Hz and 440 Hz), the sampling rate is high enough to enable a satisfyingly accurate adjustment of the common start of the time axes.

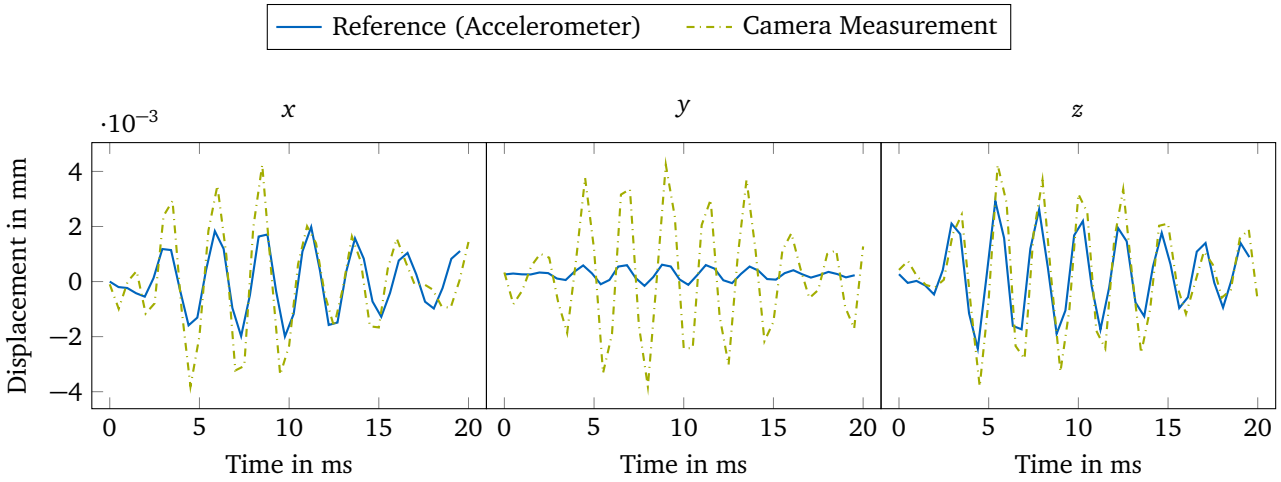

Fig. 5: Comparison of displacement measurements for the plate's mid point in all three directions for the video recorded in the lab. The reference is an acceleration measurement at the plate's bottom which was integrated twice. The mirror/camera configuration was $\varphi = 30°$, $\vartheta = 10°$.

For the experiment, the deviations between measurement and (measured) reference are obviously larger than in the simulated case. Still, the reconstruction apparently can separate the directions, especially for the x and

© 2023 The Authors. *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH GmbH.

z components, the strategy works well in the depicted case. Again, to get an overview over all configurations, we analyzed the relative RMS error, whose averages for each configuration are shown in fig. 6. This time, the

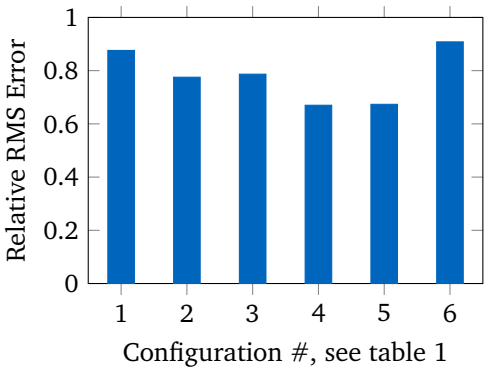

Fig. 6 Relative root mean square errors for different configurations averaged over all directions and both measurement points for the lab measurement.

differences are little bit more apparent than for the simulation and the plot suggest either configuration 4 or 5 as the ones with the best overall performance. The errors are quite a bit higher which is to be expected given (a) the very low level of displacement amplitude and (b) the occurence of sensor noise and other adverce influences in the real experiments.

Lastly, the experiments showed that not only perform configuration 4 and 5 slighlty bettter in capturing the motion, but they are also two of only three configutations for which our mirror was large enough to reflect the complete measurement surface onto the image sensor. For configurations in which the camera's optical axis hits the mirror plane under a rather small angle, in this case configurations 1, 3 and 6, one needs much mirror surface for a given size of target surface.

4 Conclusions

The presented method to reconstruct 3D motion from video from two different camera angles proved to work in principle, however, the experimental results were not yet satisfying. There are several reasons why the experimental results fall off by quite some margin compared to the simulations in BLENDER:

- lighting conditions may not be optimal,
- sensor noise is present in the real measurements,
- displacements are much smaller (factor 10) compared to the simulation,
- reference is also a measurement with its own flaws and needed to be integrated twice for comparison.

Nevertheless, the main objective of this study was to investigate the effect and feasibility of altering the mirror and camera angles in such a setup. For the numerical analysis, the differences between the setups are not significant. This might not be such a big surprise: while it was shown in section 2.2 that the condition of the matrix to be inverted gets worse as both views are more and more aligned, the condition numbers resulting from the investigated setups in this study are far from beeing problematically high. The second reasoning for a large angle between the views was that the sensor noise plays a smaller role as the difference between the images is larger. This argument doesn't play any role for the simulation as sensor noise was not modeled in BLENDER.

On the other hand, the experiments speak very clearly: only with a few combinations φ/ϑ (configurations 2, 4 and 5) a good measurement of the whole surface was practically possible.

To conclude, it should be said that the experimental results did not give as clear insights as we hoped for, but still the simple experience of trying out different configurations showed that the described setup is quite limiting by construction and one should primarily take care of good lighting conditions of the test objects under consideration. For the presented case, we recommend roughly 45° as a mirror angle and a camera pointing slightly downwards. Also, that's not far from what was initially done in the earlier study mentioned before in the text [4] where good results were gained.

Future work should compare classic 3D reconstruction with the method presented here to give insights in whether the simpler calibration process justifies the inherent additional measurement errors.

Acknowledgements Open access funding enabled and organized by Projekt DEAL.

© 2023 The Authors. *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH GmbH.

References

[1] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, Mech. Syst. Signal Process. **86(B)**, 17–34 (2017)
[2] Blender Foundation, Blender - a 3D modelling and rendering package, http://www.blender.org (2022)
[3] K. Genovese, L. Casaletto, J. A. Rayas, V. Flores, and A. Martinez, Opt. Lasers Eng. **51**, 278–285 (2013)
[4] M. Gille, M. R. W. Judd, D. J. Rixen, Rotating Machinery, Optical Methods & Scanning LDV Methods, Volume 6 (Springer, Cham, 2022), pp. 97-–104
[5] B. D. Lucas and T. Kanade, Proceedings of the International Joint Conference on Artificial Intelligence **81**, 674–679 (1981)
[6] K. Zaletelj, D. Gorjup, and J. Slavič, ladisk/pyidi: Release of the version v0.23 (2020)
[7] P. Virtanen et. al., Nat.Methods, **17(3)**, 261-272 (2020)
[8] L. Yu and B. Pan, Opt. Lasers Eng. **87**, 120–128 (2016)

 © 2023 The Authors. *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH GmbH.