Random laser speckle pattern projection for non-contact vibration measurements using a single high-speed camera

Pablo Etchepareborda, Marie-Hélène Moulet, Manuel Melon

Version of Record: https://www.sciencedirect.com/science/article/pii/S088832702100114X
Manuscript_edabef5741bfddf763f847861504e165

Random laser speckle pattern projection for non-contact vibration measurements using a single high-speed camera

Pablo Etchepareborda[a,*], Marie-Hélène Moulet[b], Manuel Melon[a]

[a]*Laboratoire d'Acoustique de l'Université du Mans, LAUM - UMR 6613 CNRS, Le Mans Université, Avenue Olivier Messiaen, 72085 Le Mans Cedex 9, France*
[b]*CTTM Le Mans, 20 rue Thalès de Milet, 72000, Le Mans, France*

Abstract

3D vision methods are a powerful tool for measuring full-field vibration patterns in the most varied types of surfaces. Digital Image Correlation is typically used to track noticeable features on the surface in order to measure local displacements on the tested system. In this work, a novel randomly structured light system based on laser speckle pattern projection and its calibration procedure are proposed for applying computer vision methods to the measurement of vibration in featureless or reflective objects without modifying them. A simple projector is used based on a laser beam affected by a diffuser element. A stereovision set-up with at least one high-speed camera is used to record video sequences from which the initial shape and evolution of the sample displacement are obtained. The vibration mode behavior of different steel plates are obtained using this full-field vibration measuring technique. The pros and cons of the proposed method are discussed and compared to similar stereovision set-ups.

Keywords: Full-field vibration measurement, Structured light system, Speckle pattern, High-speed cameras, Digital Image Correlation

1. Introduction

Lately, high-speed 3D shape measurement techniques became widely used in fields such as biomechanics, industrial quality control, and human-computer interaction. Stereo-vision systems using two fast cameras are the most popular set-up to measure full-field vibrations of objects by capturing images from different viewpoints [1, 2]. However, as high-speed cameras are expensive, alternative set-ups using mirrors and a single camera have been proposed [3, 4, 5, 6], and used for measuring vibration [7]. Moreover, a simpler set-up with a single high-speed camera and without mirror may be used if only the normal out-of-mean-plane displacement of the sample is required [8].

In a 3D coordinate calculation of displacement based on triangulation, the natural features on the surface of the object or scene are used to find corresponding points between the viewpoints. This process is called stereo-matching and is typically performed by means of Digital Image Correlation (DIC) techniques. Then, tracking the corresponding points along the frames in video files enables the triangulation of the temporal evolution of the observed shape. However, in featureless surfaces, visible artificial patterns need to be added to the sample for the stereo matching procedure, usually by using a customized speckle pattern [9]. Frequently, adding features damages or mistreats the sample, making it a destructive test tool for industry purposes. In addition, adding speckle stickers or painting modifies the vibration behavior of light structures.

Structured light system (SLS) methods use multishot or single-shot approaches for achieving contactless 3D shape reconstruction of featureless objects based on projected light patterns and a camera system. Multishot approaches, such as fringe projection profilometry using phase shifting, are appropriate for full-resolution high accuracy surface

*Corresponding author: manuel.melon@univ-lemans.fr

Abbreviations: DIC, Digital Image Correlation; DOE, diffractive optical element; GGD, ground glass diffuser; LSP, laser speckle projection; POV, point of view; SLS, structured light system; WFC, white flat calibration.

reconstruction of a static object [10]. Nevertheless, single-shot techniques are better suited for vibration measurements due to their higher acquisition rate. A Fourier transform analysis applied to a single image showing a fringe pattern obtains wrapped displacement-related phase maps due to the use of the arctangent function with range $[-\pi, \pi)$. Therefore, a complex phase unwrapping procedure is necessary to eliminate phase ambiguities. In addition, fringe analysis suffers from the spectrum overlapping problem which limits the local maximum slopes or step heights. Some of these limitations can be solved by using convoluted post-processing routines, increased number of patterns, multi-view systems [11], or by optimizing composite embedded patterns without compromising the fringe modulation and the phase measurement accuracy [12]. However, they are still not suitable when simple set-ups are required or when the fringe density is similar to the bandwidth of the involved signal. On the other hand, single-shot approaches for 3D reconstruction of moving objects based on laser speckle projection (LSP) are becoming popular in the consumer electronic market, *i.e.*, Apple iPhone X, Microsoft Kinect, and Intel RealSence series [13]. Laser speckle is being preferred to the conventional projection system using a digital light projector, as it is easy to miniaturize, it provides significantly wider depth of field and brightness control, invisible patterns with infrared laser can be used, and it has configurable spatial coherence which provides adaptability to different sizes of the measured object or camera resolution. Laser projection systems may use diffractive optical elements (DOE) to project known speckle-like patterns [13], but unknown random laser speckles are generated when a laser beam passes through a normal ground glass diffuser (GGD), or through a DOE if the beam diameter is smaller than required by the DOE. The interference of light waves scattered by the DOE (or GGD) generates objective speckles that are projected on the bare surface of the object encoding the necessary features for the correlation procedure. Although a single camera at different positions may be used for measuring static objects with random LSP by establishing dense correspondences between images [14], two simultaneous recordings should be analyzed to locate and track correspondences on the projected speckles in dynamic samples [15]. We note that initial works using LSP used dynamic patterns and temporal correlation schemes [16, 17]. However, they are not well suited for dynamic samples.

SLS may not only be used for patterning in a stereo view system, but also the projector itself may become a different point of view in a single camera system. If the projected pattern is set by an image or video projector, the input image can be used the same way as it was acquired by a camera placed at the same position. Different calibration procedures were proposed for this set-up requiring the projector to use different and dedicated images as input [18, 19, 20]. Differently, laser speckles are random, so the projected image is unknown, and therefore stereo camera systems have been preferably used with LSP despite last efforts to register the 3D intensity of repeatable speckle patterns [21]. In this work, we show how to calibrate an LSP-SLS system using a single viewpoint by obtaining the speckle structure generated by an off-the-shelf DOE illuminated with a not necessarily collimated laser beam which width is smaller than required by the DOE. This article is organized as follows: Section 2 describes the theoretical background and operation of the proposed method, Section 3 reports results from proof-of-concept measurements of three different steel plates and comparisons against other full-field and single-point measurement techniques. Finally, results are discussed and pros and cons of the methods are listed.

2. Methods

2.1. Stereo vision basics

The 3D locations of points in space can be measured when at least two different points of view (POV) provide relevant information for a triangulation calculation. A 3D point in the real world specified in a camera coordinate system $(X, Y, Z)^T$ projects onto a 2D location $(u, v)^T$ over the pixel matrix of a camera according to the pinhole model [22]

$$(uZ, vZ, Z)^T = K(X, Y, Z)^T, \tag{1}$$

with

$$K = \begin{pmatrix} f_u & 0 & p_u \\ 0 & f_v & p_v \\ 0 & 0 & 1 \end{pmatrix}. \tag{2}$$

f_u and f_v are the focal lengths along the $\vec{u}$ and $\vec{v}$ axes, and (p_u, p_v) image plane projection of the optical center of the objective of the camera. Thus, the location $(X, Y, Z)^T$ is necessarily on the line going through its projection $(u, v)^T$

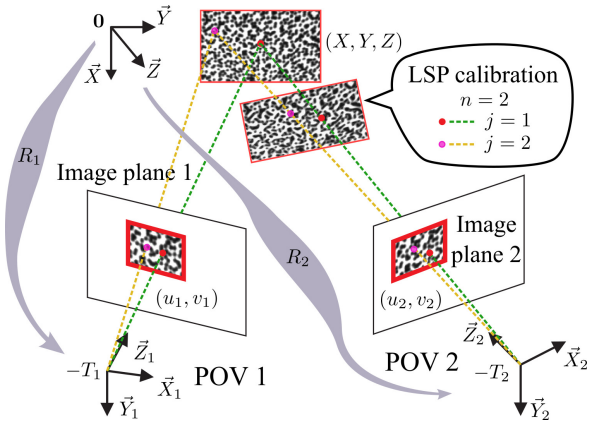


Figure 1: Stereo-vision set-up. Two images of a patterned object obtained by two viewpoints with associated coordinate systems related to world coordinate system by R_i and T_i. Two corresponding pairs are shown. In addition, a second positioning of the target is shown for an LSP-SLS calibration procedure.

and the optical center, as shown by the dotted lines in Fig. 1. These values, coupled with the distortion coefficients, correspond to the intrinsic parameters of a POV or camera.

In a stereo set-up, the global 3D coordinate system is related to the corresponding coordinate system of each of two POV by a rotation matrix R_i and a translation vector T_i with $(i = 1, 2)$, called the extrinsic parameters of each POV. The location $(X, Y, Z)^T$ in the ith POV coordinate system and image plane is found by doing

$$(k_i u_i, k_i v_i, k_i)^T = K_i(X_i, Y_i, Z_i)^T = K_i R_i[(X, Y, Z)^T + T_i], \tag{3}$$

where k_i is a parameter that determines the location of the point $(X, Y, Z)^T$ within the line going through the optical centre and the location (u_i, v_i) in the image plane. In a stereo POV system producing two images with two corresponding 2D locations (matching points), two lines in the 3D space are defined, allowing the triangulation of $(X, Y, Z)^T$ as the crossing point of these lines (see Fig. 1). This process starts by reversing Eq. (3) to obtain an expression of the first line parametrized by k_1. The projection of this line into the second image plane gives the *epipolar* line, and finding the nearest point of this line to the corresponding point in the second image allows defining the 3D location of the analyzed point.

The intrinsic and extrinsic parameters are obtained by a calibration process. Typically, several images of a flat calibration target are used as input to well known numerical calibration routines in the OpenCV Library [23] for getting the intrinsic parameters of each camera, and simultaneous pictures of the pattern taken by two cameras are used for getting the extrinsic parameters of a stereo-camera set-up. Targetless calibration was made possible as well by using an optimization solver for a geometrical model of corresponding points in an LSP based stereo-vision scheme similar to the one employed in this work [24].

Once the intrinsic and extrinsic parameters are known, the measurement of the shape of a sample is obtained by selecting a particular set of points with noticeable features in one of the images and finding their corresponding pairs by maximizing cross-correlation with the second image. Triangulation gives the 3D location of the set of points. When the images belong to sequences obtained by high-speed cameras, the tracking of these points by means of DIC, allows finding the evolution of the shape of the object.

2.2. *Structured light systems*

In a stereo-vision measurement of shape or vibration amplitude of a sample without natural features or texture, some noticeable features need to be added artificially. The most straightforward method is to add a pattern to the surface by drawing or sticking a printed pattern so that the cameras may see how this pattern is deformed by the surface of the object. However, it is usually preferable to project a visible pattern to avoid modifying, corrupting, or destroying the sample. The projected pattern is used for finding corresponding pairs of points between the images from different POV, and a point cloud representation of the shape of the sample is obtained by triangulation. A time evolution of this

shape can be measured with SLS if simultaneous video sequences are analyzed. However, this measurement is not a 3D local displacement assessment as obtained when the tracking is performed using measurement points fixed to the sample. When using SLS, the measurement points are always moving over the surface of the sample as the projected pattern varies due to the surface movement, despite of using a static projector. The obtained point clouds may be resampled using a transformed coordinate system that has its $\vec{Z}_\perp$ axis aligned with the normal direction to the mean plane of the sample, defined as the best fitting plane to the initial shape; and fixing the sample positions in the $(X_\perp, Y_\perp)$ plane to a rectangular grid. The analysis of the evolution of the measured shape is simplified as only the $Z_\perp$ values have to be considered, and normal displacement fields may be defined as the difference between deformed and initial states. In what follows, the normal displacement fields obtained by using SLS will be simply called as displacement measurements.

2.2.1. Single camera and projector-camera set-up

The structured light technique can be further exploited by replacing one of the cameras of the stereo-vision system with a projector and analyzing how the object geometry distorts the shined structured pattern using the images acquired from another perspective. In such a system, the features in the projected pattern need to be calibrated in order to perform the correspondence procedure the same way as done in a stereo-vision system.

The calibration of the system consists in accurately determining the relationship between a location in world coordinates and its projection onto the camera and projected image planes. As the projector cannot capture images like a camera, the process of calibration is more complex than the usual stereo-vision system calibration. The most established method is similar to the camera calibration procedure and requires the camera to capture images *for* the projector and then transforming the images onto the projector image plane, so that the camera pixels can be corresponded to projector pixels [18]. Different phase shifting, coding procedures, and multiple image projection over dedicated calibration targets were proposed for improving or simplifying the calibration of the projector [20].

In all of these methods, the projected image is an input to the system so a known pattern with known features is projected in order to find the correspondence between a projector and a camera pixel. There are commercial products using random patterns designed to encode locally unique features so that any given point on the camera image can be uniquely matched to a point in the pattern [2]. Thus, stereo matching is straightforward, and miniaturized implementations are simpler as only one single image is needed. However, spatial resolution might get limited by the low-cost projector device. Also, out-of-focus projection is usually a problem when using a video projector, and a well designed and more complex calibration procedure should be implemented, especially when the object is much larger and is more distant than the calibration target [25].

2.2.2. System accuracy

The uncertainty of the cited methods depends on factors as the size of the sample, the angles of the point of views with respect to the sample, the correctness of the depth of field, the resolution of the cameras or projector, the sensitivity of the camera (especially in high frame rate and high depth of field situations), the shutter speed to signal bandwidth ratio, the level of displacements in the sample, the spatial correlation width or feature uniqueness of the pattern, the size of the analysis window of the DIC procedure, the settings of the applied sub-pixel correlation method, and the accuracy of the system calibration. See Refs. [1, 7], and references therein, for detailed and systematic uncertainty assessments considering these commonly shared factors, and approaches to improve the measurement accuracy.

As mentioned in Section 1, the aspects related to the statistical properties of the pattern, the light intensity and the resolution of the projected pattern are addressed by using LSP-SLS. A known issue in LSP is the subjective speckle that is also formed and captured by the camera when imaging the optical coherent waves scattered by the rough surface of the test object. The images acquired from different perspectives will be affected by very different subjective speckle distributions, leading to a reduction in the surface reconstruction quality unless a wide aperture of the imaging lens is used [26] or a color illumination and acquisition system is introduced [27] for filtering out subjective speckle. Due to the high frame rate involved in vibration measurements, a wide aperture is preferred in order to maximize the use of available illuminance in low exposure times. In addition, another way of coping with subjective speckles is by transversely shifting cameras [28] in order to preserve the advantages of LSP such as access to optimized measurement points selection and matching algorithms [29, 30].

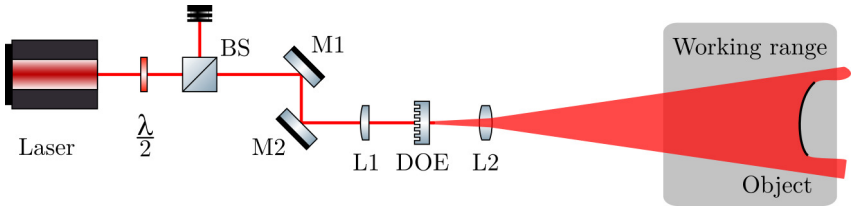


Figure 2: LSP optical set-up. The intensity of the laser beam is controlled by means of a half-wave plate and a polarazing beam splitter. Two mirrors are used to orientate the beam towards the object. Lens L1 is used to control the beam width at the diffractive optical element, and lens L2 allows expanding the beam wider if needed.

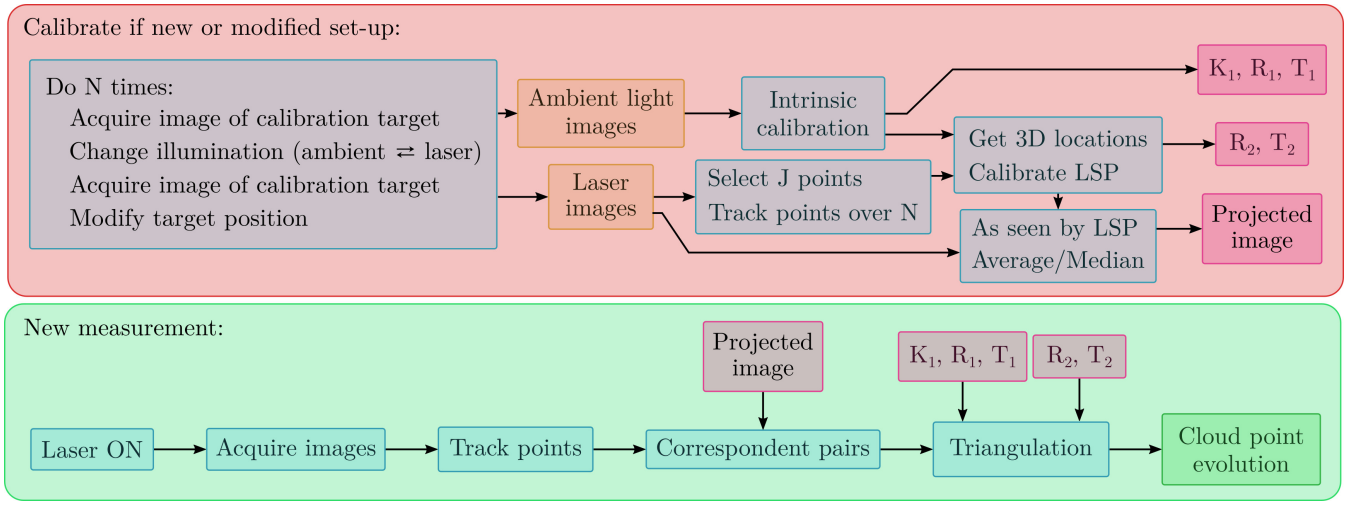


Figure 3: Flowchart of the complete vibration measurement procedure including the calibration of the proposed system.

2.3. *Random projected image calibration*

In this work, we propose using an LSP system as shown in Fig. 2. A beam is directed towards a DOE with controlled width and intensity. Note that the DOE is simply used a holographic diffuser that produces a divergent beam with a random objective speckle pattern. The position of the lens L1 with respect to the DOE plane and its focal length determines the beam width at the DOE. The smaller the focus spot on the diffuser the larger is the speckle size and vice versa, see Ref. [28] for more details in relevant statistical properties of laser speckle. Lens L2 optionally allows wider expansion of the pattern to completely illuminate the object. Our model assumes that there is a working range of the beam where the divergent pattern projected by the DOE has an almost constant spatial distribution along the propagation axis away from the DOE. A pinhole model describes the 3D distribution of intensity. The image plane can be defined at an arbitrary distance $Z_p = 1$ in the projector coordinate system $(X_p, Y_p, Z_p)^T$, which relates to real world points according to Eq. (3) and $(X_p, Y_p, Z_p)^T = (k_p u_p, k_p v_p, k_p)^T$. Here, (u_p, v_p) is the location of points in the unknown random projected image. The intrinsic matrix is assumed to be the identity matrix $K_p = \mathbb{I}_3$ without loss of generality. R_p and T_p are the extrinsic parameters describing the orientation of the coordinate system and the position of the projector focus. Note that the calibration of the system implies estimating the extrinsic parameters, and also the projected image at a user defined resolution. The flowchart in Fig. 3 shows the complete procedure of calibration and vibration measurement as proposed in this work. The following brief analysis introduces the calibration procedure based on simple geometrical constraints given by the proposed model.

A particular point in the projected image plane $(u_p, v_p)_j$ propagates along a line coming from the projector optical center. Suppose this point is surrounded by a particular feature kernel at the projector image plane, a scaled version of this kernel Γ_j would be repeated almost the same into an all-white flat calibration (WFC) target if this target was presented in front of the beam at N slightly different tilting angles and distances from the optical center. Let us use a calibrated stereo-vision system to capture images of the $N \geq 2$ projected patterns over the WFC target, and arbitrarily

select J points over the first image ($n = 1$) of a particular POV. The surrounding pixels of these points define the kernels Γ_j, with $j = 1, \ldots, J$. As example, three images from both cameras are shown in Fig. 4(a) when using a patterned target for calibrating the stereo set-up, and in Fig. 4(b), when using different placements of the WFC target illuminated by the LSP for calibrating the SLS. The cross-correlation of the kernels (extracted from $n = 1$) with the other images ($n = 2, \ldots, N$) allows finding the matching points between images of different positions of the target, as shown in the zoomed areas in Fig. 4(b) for a particular j. Therefore, the sets of points $\mathcal{X}_j = \{\mathbf{x}_{j,n} \in \mathbb{R}^3, n = 1, \ldots, N\}$ are defined, one for each Γ_j. Note that the 3D location of the points is found by triangulation using their matching pair in the image from the second camera. A fitting line $\mathbf{l}_j(\alpha) = \alpha\mathbf{v}_j + \mathbf{c}_j$ is found for each set by using the singular value decomposition of the matrix $A_j = U_j\Sigma_j V_j^T \in \mathbb{R}^{N\times 3}$, where $\mathbf{l}_j, \mathbf{v}_j, \mathbf{c}_j \in \mathbb{R}^3$, $\mathbf{c}_j$ is the mean point of the set, $U_j \in \mathbb{R}^{N\times 3}, \Sigma_j \in \mathbb{R}^{3\times 3}$ is diagonal, $V_j \in \mathbb{R}^{3\times 3}$, and the nth row of A_j is $\mathbf{x}_{j,n} - \mathbf{c}_j$. $\mathbf{v}_j$ is the column of V_j corresponding to the index of the maximum singular value in Σ_j [31]. Note that in the case of $N = 2$, $\mathbf{l}_j$ correspond to the lines intersecting the pair of points in $\mathcal{X}_j$. All of the J fitting lines should pass through each corresponding $(u_p, v_p)_j$ in the image plane and should finally converge to the projector optical center $-T_p$, which is estimated as the point that minimizes the overall distance to the lines $D = \sum_j \|\alpha_j\mathbf{v}_j + \mathbf{c}_j + T_p\|^2$, by solving

$$\begin{bmatrix} N\mathbb{I}_3 & V \\ V^T & \mathbb{I}_N \end{bmatrix} \begin{bmatrix} T_p \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} -\sum_j \mathbf{c}_j \\ -\mathbf{w} \end{bmatrix},$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_J)^T$, and $\mathbf{w} = (\mathbf{v}_1^T\mathbf{c}_1, \ldots, \mathbf{v}_J^T\mathbf{c}_J)$. Fig. 1 schematizes this model by showing a situation with $N = 2$ and $J = 2$, where the intersection of the dotted lines defines the origin of POV2 coordinate system. Then, the rotating matrix R_p is set with a yaw and pitch angles (no roll) to orientate the $\vec{Z}_p$ axis of the projector coordinate system (centered in -T_p) towards the mean point of the employed sets of points $\mathbf{c}$. The orientation vector $\mathbf{o} = (o_x, o_y, o_z)^T = \mathbf{c} + T_p$ defines the yaw (θ_x) and pitch (θ_y) angles as $\theta_x = \arctan(o_y/o_z)$, and $\theta_y = -\arctan\left(o_x/\sqrt{o_y^2 + o_z^2}\right)$, so that $R_p = R_yR_x$, where R_x, and R_y are formed by using the Rodrigues equation using θ_x and θ_y and their respective rotation axis [23]. Thus, the points $(u_p, v_p)_j$ on the projector image should be regularly distributed over the $Z_p = 1$ plane and around $(0, 0)$.

Once the extrinsic parameters are found, it only remains to interpolate the projected image at user defined resolution by using the intensity values registered at different locations in space over the same WFC target and taken to their corresponding locations in the projector image plane (u_p, v_p), by means of Eq. (3). This can be performed by using a single image where the entire pattern was projected over the WFC target, or applying this same transformation to get all N images as if they were *seen by the projector* and averaging them to obtain the final values of the projected image, as shown in Figs. 4(c,f). It is worth noticing that this *projected image* can be used as the image, or high-frame rate video, acquired from a second camera that does not present subjective speckle effects due to the averaging operator.

The vision set-up was further simplified in this work by removing one of the cameras in the system. The intrinsic parameters of the remaining camera were estimated using a calibration target containing a grid of Gaussian points with known locations within the target, i.e. only the left images in Fig. 4(a) are used. Gaussian points were recommended in Ref. [7] for improved estimation of the location of pattern features and more accurate intrinsic calibration. As usual, room light was used for the illumination of each of these scenes. For each shot, before moving the target, a new picture was taken with the room lights off and the LSP illumination on. Fig. 4(d) shows the LSP illumination images corresponding to the images in Fig. 4(a). Note that the shined zone of the printed pattern changes when the target is moved. The usual calibration functions in OpenCV using left images in Fig. 4(a) as input have the location and orientation of the calibration target as output, represented by the rotation matrix R_t and translation vector T_t. Therefore, the 2D locations over the plane of the target $(x_t, y_t)^T$ are also defined by 3D positions in the camera coordinates as $\mathbf{x}_c = R_t(x_t, y_t, 0)^T + T_t$. Thus, the target plane is set in terms of a normal vector $\mathbf{n}$ so that $\mathbf{n}^T\mathbf{x}_c = 1$, and which is obtained by solving $\mathbf{n}^T(T_t, \mathbf{r}_{t1}, \mathbf{r}_{t2}) = (1, 0, 0)$, being $\mathbf{r}_{t1}$ and $\mathbf{r}_{t2}$ the first two columns of R_t. So, using this condition in Eq. (1) results in the following transformation of each pixel (u, v) corresponding to a point in the target to a 3D location

$$\mathbf{x_c} = \frac{K^{-1}(u, v, 1)^T}{\mathbf{n}^T K^{-1}(u, v, 1)^T}.$$

As the target was kept immobile for both illumination schemes, only the matching procedure between images of different positions of the target has to be performed when working with images in Fig. 4(d).

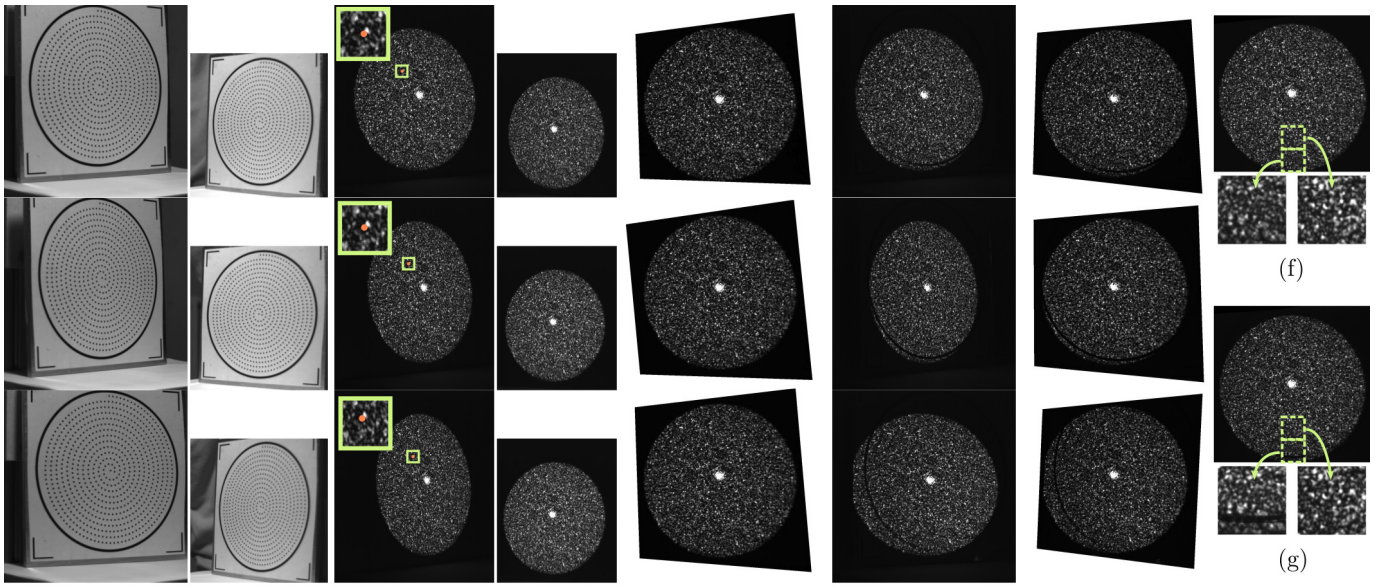


Figure 4: Three images from left and right cameras (L and R, R reduced) (a) changing the position of a target with Gaussian points, and (b) changing the position of a WFC target being shined by the LSP, with magnified zones of matching points. (c) transformed L camera images in (b) as were seen by the projector, (d) LSP illumination of the L images in (a), (e) transformed L camera images in (d) as were seen by the projector. And calibrated LSP images (with magnified boxes) using: (f) stereo system and WFC target, (g) single camera and LSP over patterned target.

It is worth mentioning that the reduced reflectivity inside the Gaussian points of the target was not an obstacle for performing the matching procedure of kernels Γ_j between the N images (usually more than 3). Other calibration patterns such as the two-color chess board in Ref. [18] could be used to reduce the effects of the underlying pattern in the laser speckle calibration images. However, an initial guess of T_p and R_p was performed by using only two positions of the target, being those the most parallel oriented pair but still being at different planes. The reduced deformation of the kernel in this pair of images eases the matching procedure. The initial calibration allows transforming the images as if they were *seen by the projector* so that the matching procedure for all images is significantly simplified. Once the extrinsic calibration is performed, the transformation of the images is corrected for interpolating the projected image. Note in Fig. 4(e) that the transformed images present features corresponding to the printed pattern on the target. The median values from the images are used instead of the average to minimize the effects of these outlying low intensity pixels, the result is shown by the magnified zones in Figs. 4(f,g) comparing the calibrated LSP images from both procedures.

3. Results

The proposed method for using a random LSP system as a viewpoint in an SLS stereo vision set-up was applied in the measurement of normal displacements of a steel flat plate without surface features and excited by a shaker (10 N max. output force at 10 V input). A conventional stereo vision system was simultaneously applied for comparing the results and assessing the quality of the proposed calibration procedure. The plate needed to be covered with a blank sticker to make its surface diffusive and be able to use the SLS 2-camera set-up. The uncovered plate was also measured to prove the convenience of using the proposed camera-projector system in high reflective surface cases, where the set-up of two cameras at separated locations is impractical. Finally, a curved diffusive plate with features over the surface was analyzed by using the proposed approach, the 2-camera set-up with SLS, and the more conventional 2-camera set-up with white light illumination. In all cases, a Polytec CLV-2534 laser vibrometer was used as reference in a one-point measurement. A discussion of results is provided to show the limitations of the proposed simplified set-up requiring a single camera.

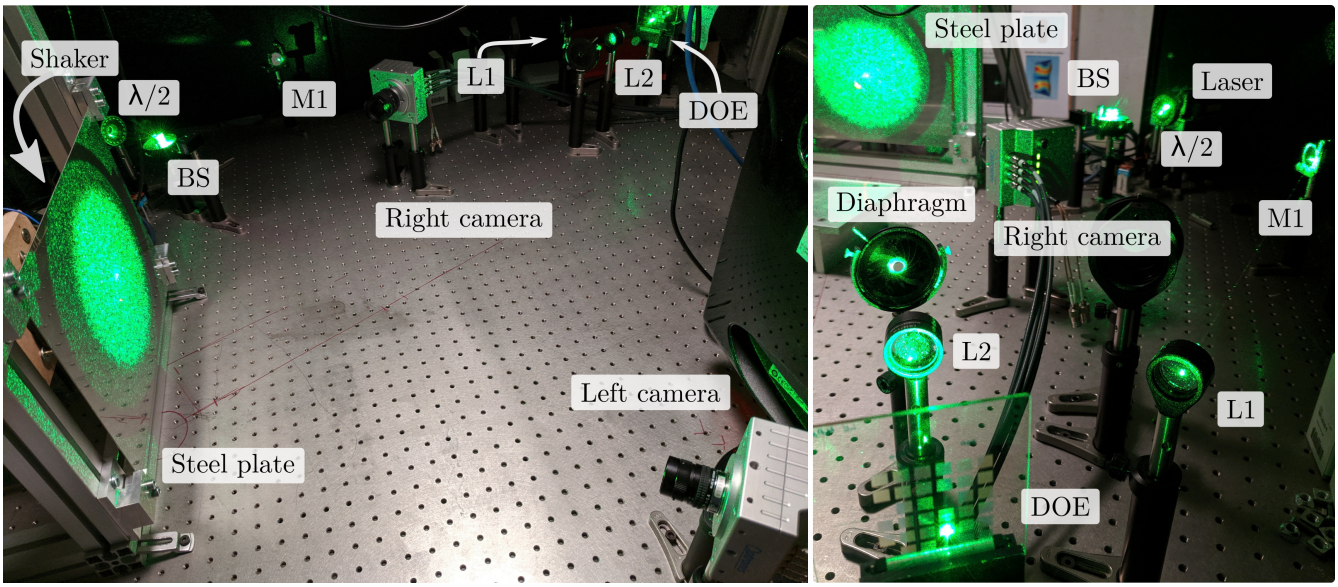


Figure 5: Views of the set-up for measuring a steel plate comprising the LSP and the stereo-vision arrangement with two cameras.

LSP was implemented using a 300 mW 532 nm CW laser in a set-up as presented in Fig. 2. In this work, we used a DOE producing a wide beam with rectangular or circular shape. DOEs are designed to be illuminated by a collimated beam to produce the desired shape at a particular distance. However, our set-up illuminates the DOE with a convergent spot obtaining a stable speckle pattern with higher intensity inside the shape produced by the DOE. Fig. 5 shows the implemented set-up. The vibrometer is not shown in the pictures as it is located further away and pointing perpendicular to the plate.

3.1. Number of calibration images and resolution assessment

A patterned calibration target was placed in 8 different positions and with different illumination conditions as described in Section 2.3. The stereo-camera system was calibrated with the white light images, and the single camera LSP-SLS was then calibrated using only the left camera images of the set. The latter calibration process was repeated 7 times while using sets of different amounts of images $(2, 3, \ldots, 8)$, obtaining 7 different sets of calibration parameters and interpolated projected images. Afterwards, a WFC target was presented in 7 different positions to the calibrated stereo setup, and images were acquired by both cameras with the laser speckle pattern projected over the target. The shape of the flat target was obtained by using 150 measurement points in average for each calibrated system and each position of the target using the stereo-camera system (1 calibration x 7 WFC target positions) and the single-camera LSP-SLS (7 x 7). A simple cross correlation procedure was used to find corresponding points between the viewpoints, with very low user intervention. The measured shapes were compared to a flat plane fitted to the shape obtained by the conventional stereo-camera system. The root mean square error and the median error were used as statistical quality indices, being the median error less sensitive to outlier measurement points that could be avoided by changing to a more reliable correlation technique, as used in Ref. [5]. Fig. 6 shows the error indices, averaged over the 7 different positions of the target, for each calibration of the proposed system. The mean fitting error values from the stereo camera system measurements (calibrated using the 8 pairs of images) are given as reference with horizontal dashed lines. The obtained resolution of the proposed technique was 40% higher than the resolution of the more expensive stereo-camera system, and had similar magnitude to other systems in the cited references using similar approaches to assess resolution [12, 16, 17, 28]. It is worth noticing that due to space constraints, the stereo system had a wider angle between viewpoints than the single-camera LSP-SLS, producing better accuracy as ground truth in this validation assessment. Both indices show the convenience of using 5 images for this procedure of calibration of the single-camera system. The decreased performance at higher number of images is explained by a blurring effect

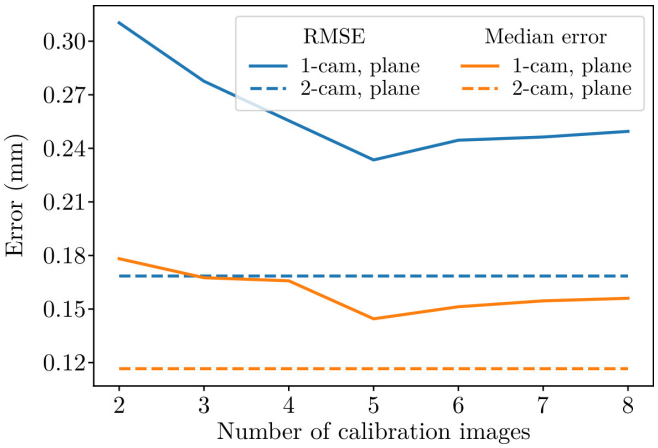


Figure 6: Root mean square error and median error as resolution quality indices of the single camera LSP-SLS technique and their dependence with the amount of images used for the calibration (flat target at different positions). The errors are measured against a plane fitting to the measurement points obtained by a conventional stereo system. The fitting errors are given as reference using dashed lines.

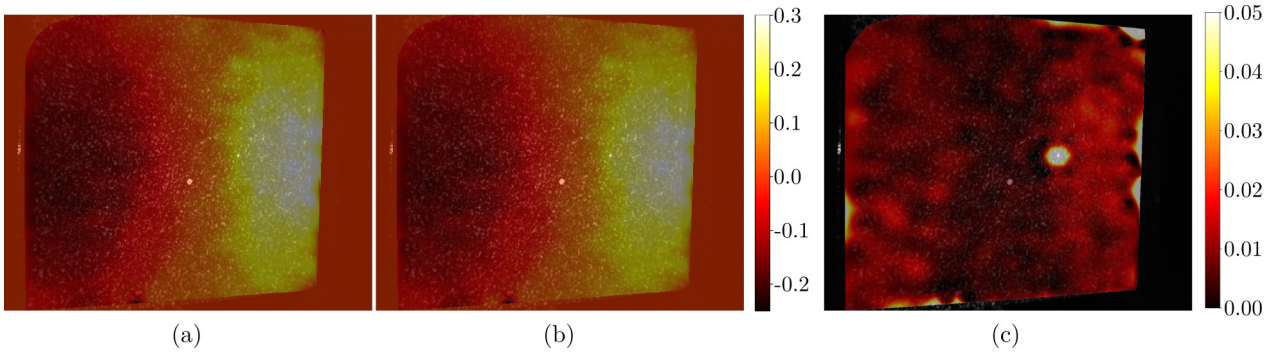


Figure 7: Amplitude measured in mm of sinusoidal normal displacement of flat plate with shaker excitation at 63 Hz using set-up with (a) 1 camera, and (b) 2 cameras. The standard deviation error is shown in (c).

that was observed on the obtained projected image when adding more images to the median based fusion operation explained in Section 2.3.

3.2. Comparison between stereo and camera-projector SLS systems

A flat steel plate held by the corners and with an A4 sized blank sticker paper introducing a rough surface was inspected by using the described LSP system and two cameras placed so that their viewpoints are symmetrically opposed with respect to the normal direction of the plate. The direction of propagation of the LSP system was located in between the direction normal to the plate and the camera with viewpoint closer to the right side of the plate (right camera). The left camera was used as the single camera in the case of using the proposed method, and both cameras were used in a conventional SLS set-up for comparison purposes. Clearly, a conventional stereo-vision measurement is not suitable due to the lack of visible features on the sample.

First, an AC sweep was used to drive the shaker to find particular modal frequencies at which the behavior of the vibrating plate would give interesting and comprehensible visual comparisons. Besides, different torques were used for tightening the holders at the corners of the plate to intentionally produce asymmetries in the vibration modes that could only be assessed by full field measurements. Then, bursts of sinusoidal signals were used as excitation to the shaker at 8 V amplitude. The recording of pictures begun 1.5 s after the first periods of excitation to avoid measuring transient effects. Fig. 7(a,b) shows the amplitude of the Fourier component of the displacement signal at the frequency of excitation $f_e = 63$ Hz, which was measured at a frame rate of 450 fps using the proposed single camera method and the 2-cameras method with SLS setup, respectively. The results are plotted over the initial image of the sequence,

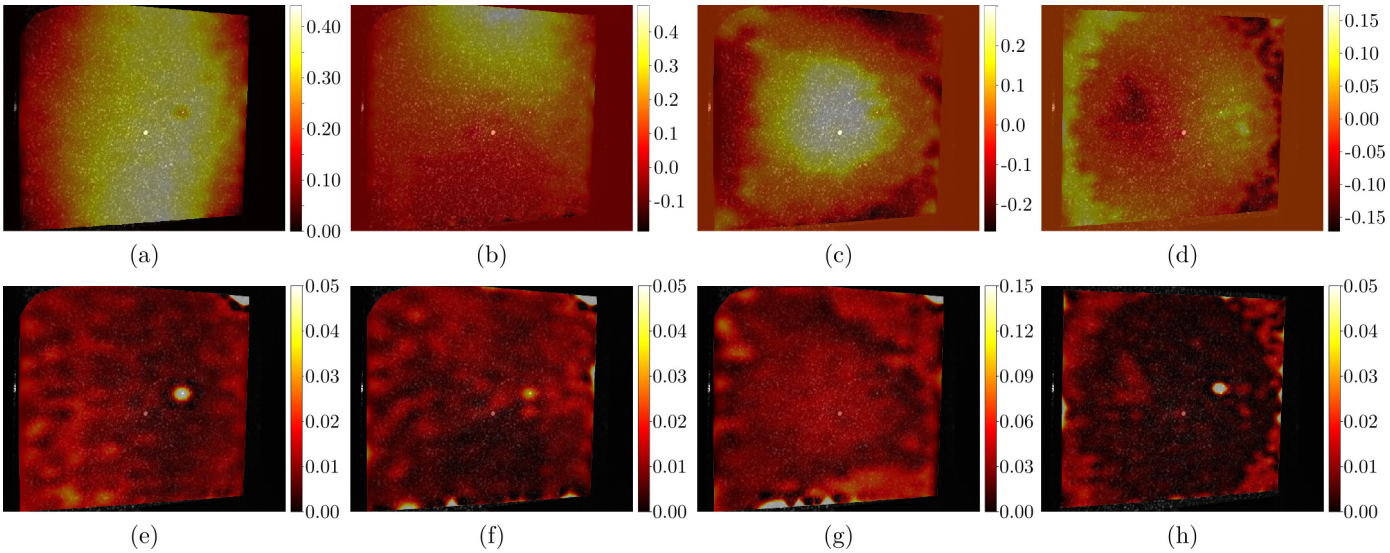


Figure 8: (a-d) Amplitude measured in mm of sinusoidal normal displacement of flat plate with shaker excitation at 38 Hz, 49.5 Hz, 142 Hz, and 231 Hz using the proposed single camera setup. The respective standard deviation error maps against results from the 2-camera set-up with SLS are shown in (e-h).

with a corrected perspective. The error map between the measured displacement signals at each measurement point is presented in Fig. 7(c). As the measurement points at the border of the projected pattern and points around the laser vibrometer spot may have significant error values, the median value of error $e^{\text{med}} = 8.9$ μm is more reliable than the mean value $\overline{e} = 13.4$ μm. The displacement measured by the proposed method and error respect to the SLS 2-camera method are shown in Fig. 8 for the excitation frequencies of 38 Hz, 49.5 Hz, 142 Hz, and 231 Hz, with frame rates 250 fps, 350 fps, 1000 fps, and 1000 fps, and obtained median errors of 10 μm, 9.3 μm, 36 μm, and 3.7 μm, respectively. In these pictures, different mode shapes and asymmetries can be observed plainly, even in the case where the maximum displacement is less than 150 μm.

In Fig. 9, the measured signals at a measurement point close to the laser vibrometer spot, corresponding to the sinusoidal burst excitation at 38 Hz (a,b) and a step response (c,d) are shown. A portion of the displacement signal is shown in Fig. 9(a), and the absolute values of the Fourier transform of the signals are shown in Fig. 9(b,d). A frame rate of 250 fps was used in the sinusoidal case and 1000 fps in the step response, but both cases used 0.8 ms exposure time. The f-number of the cameras was kept low (close to 4) to improve the intensity levels at this low exposure time, so the mean size of the subjective speckle was always smaller than the pixel and therefore, its effect was filtered out. A good correspondence is found between the vision measurement techniques and the single-point laser vibrometer result, being the signals from the vision methods very alike. As expected, the noise level observed in vision methods spectra is higher than the noise level of the laser vibrometry technique, being the standard deviation of noise in the stereo-camera and the proposed systems of 10 µm and 11 µm, respectively. Low frequency noise components observed with vision methods were probably due to the fact that the excitation element was attached to the same support as the vision system, producing undesired vibrations. This should not happen in real-life measurements where the sample can be mechanically isolated from the projector and camera.

3.3. Measurements on a reflective plate

A second set of measurements were performed over the same plate but without the sticker, so that a reflective surface has to be measured. In this situation, the intensity reaching the right camera is much smaller than the intensity reflecting towards the left camera. So, it is impossible to perform the correlation procedures at high frame rates due to the low exposure time involved. Therefore, only a single camera technique is suitable to measure dynamic or transient phenomena on this kind of samples without affecting their surface by any type of mechanical patterning. It is worth noticing that many products in industry possess smooth, and consequently, reflective surfaces as happens with the measured plate.

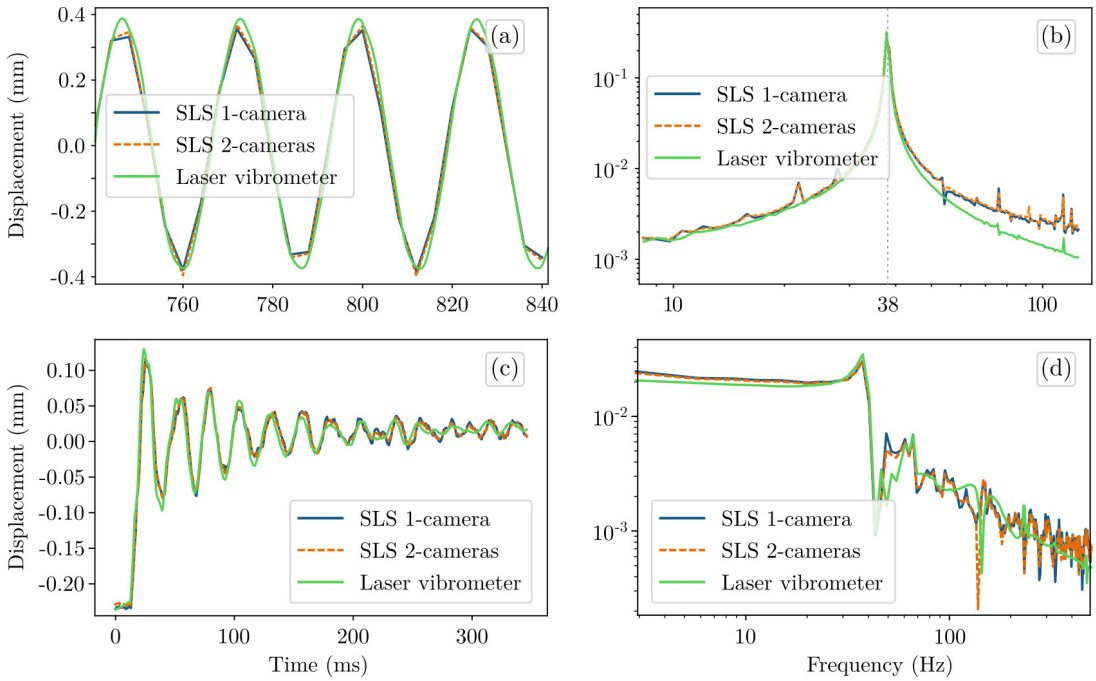


Figure 9: Normal displacement over a measurement point close to the laser vibrometer spot on a flat plate using the proposed single camera set-up and the 2-camera set-up with SLS at: (a, b) frame rate 450 fps and with sinusoidal shaker excitation at 63 Hz, and (c, d) frame rate 1000 fps and 8 V step excitation. Temporal signals are shown in (a, c) and absolute value of the Fourier transform coefficients are shown in (b, d).

Fig. 10 shows the correspondence between the temporal evolution of the step response at a measurement point close to the laser vibrometer spot and the measurement from the laser vibrometer. A measurement made two days later corresponding to the same calibrated setup, excitation, and plate is also shown in Fig. 10. An offset is introduced for visualization purposes (illustrated in the bottom left corner of the plot). The resemblance between the obtained signals shows the stability of the system, which does not require to repeat the calibration procedure unless any position of elements is changed or any configuration of camera lenses is adjusted. Instants at 0 ms, 29 ms, 94 ms, and 335 ms are marked corresponding to the pictures in Fig. 11 that show the initial shape of the plate and the following displacement maps measured with respect to the initial position and using the same color scale. The instant at 29 ms corresponds to the maximum displacement state, whereas at 94 ms some asymmetries can be observed, and the instant at 335 ms corresponds to a final state situation with respect to the initial shape.

3.4. LSP-SLS over curved and patterned plate

A different curved plate was used as sample in a vibration measurement. This plate was covered by an A4 sized sticker with a printed pattern of spots (see Fig. 12(b)) to study the degree of demeaned performance when visible features are present in the surface of the sample. Real-world samples might have slightly textured zones that may produce reflections detected by the camera that are more related to features on the surface rather than the projected pattern. These reflections cannot always be avoided by changing the angle of the surface with respect to the illumination. So, it is important that the measurement can still be made under this noise conditions.

Here, a visual comparison of results is performed between a conventional stereo 3D vision measurement using a white LED spot for illumination, and another using a single camera and LSP over the already existing printed pattern. The same shaker from previous sections was used to excite the plate with a burst input of sinusoidal 75 Hz and 8 V amplitude. Fig. 12 shows the measured initial shapes and amplitude of displacement using both methods. As expected, results from the stereo-vision system are smooth and accurate. The results from the proposed method are noisy in comparison to the stereo-vision system but it is worth noticing that the initial shape and displacement amplitude are similar to the expected results. Thus, the proposed method is suitable for measuring shape and displacement in complex samples having visible features, smooth diffusive surfaces and reflective zones in part of their surfaces.

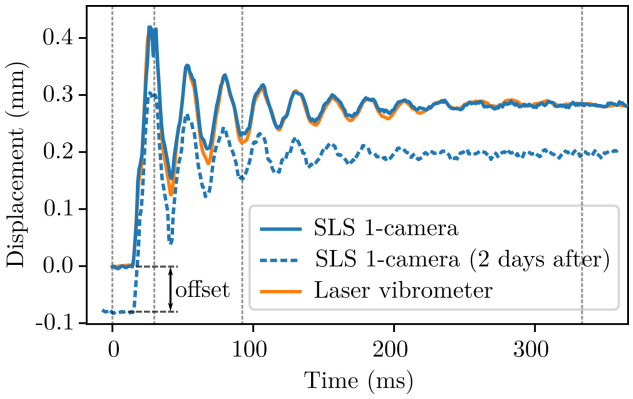


Figure 10: Normal displacement of a measurement point close to the laser vibrometer spot on a reflective flat plate using the proposed single camera set-up at frame rate 1000 fps and 8 V step excitation on the shaker. The absolute value of the Fourier transform coefficients are also shown.

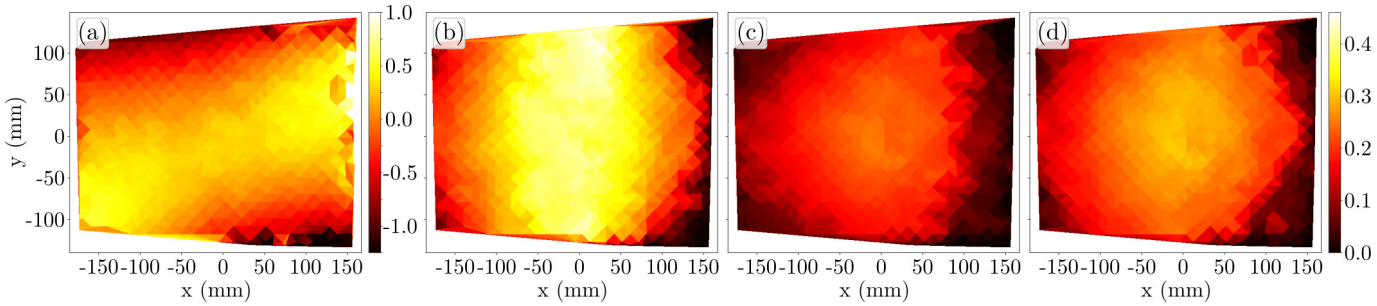


Figure 11: (a) Initial shape (mm) for a measurement on a reflective flat plate using the proposed single camera set-up at frame rate 1000 fps and 8 V step excitation on the shaker. Subsequent normal displacements in mm with respect to the initial position at times (b) 29 ms, (c) 94 ms, and (d) 335 ms. (b-d) pictures share the same color scale shown for (d).

4. Conclusion

A 3D vision method for measuring fast full-field vibration patterns using a single high-speed camera was proposed based on a laser speckle structured light system. The measurement of featureless, or even reflective, objects is made possible by the proposed technique without the need of sticking or forcing features on the surface of the sample. The proposed LSP-SLS technique uses a low-cost set-up as it requires a single high-speed camera and a projector mainly composed by a DOE acting as a holographic diffuser and a laser source, and the alignment is trivial. Additional lenses may be added to control the beam width and size of speckles for adjusting illuminance on the sample and improving correlation procedures, respectively. The proposed calibration procedure is fully explained, and it allows finding the image that would obtain the same speckle pattern if a video projector with known input image was used instead. A stereo-matching algorithm between an image from the camera and the calibrated projected image allows getting the initial shape of an object, and Digital Image Correlation is used to track the speckle pattern over the surface of the moving object. The performance of the proposed system and its calibration was validated and compared against conventional full-field vision techniques measuring a flat target. Vibration modes of a flat and a curved steel plates were obtained using this full-field vibration measuring technique when excited by a shaker. The results agreed with reliable laser vibrometer measurements. The robustness of the proposed method when features are already present in parts of the sample was analyzed by comparing results with a stereo-vision set-up using a LED spotlight instead of laser projection. The obtained results are promising regarding measurement of complex surfaces and future work towards assessing the performance of the proposed method in more complex and heterogeneous textured objects is encouraged.

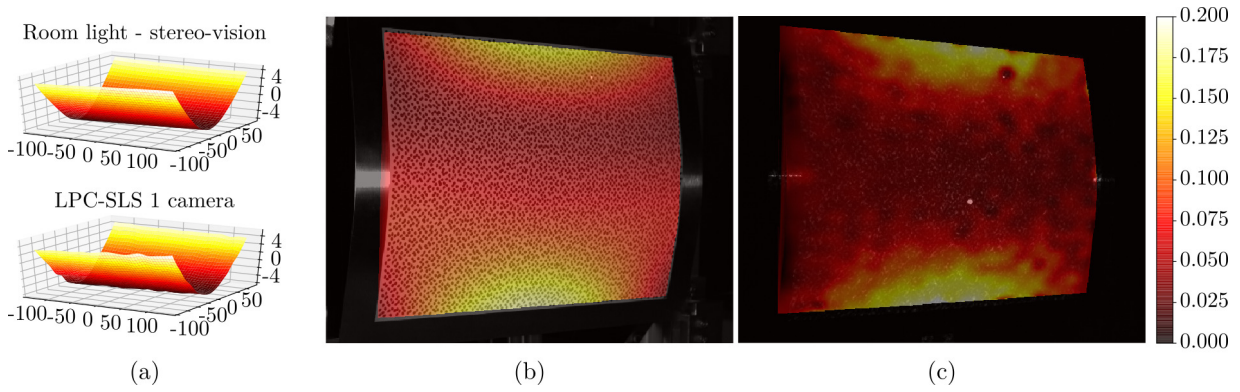


Figure 12: (a) Initial shapes (mm) and (b-c) normal displacement amplitude measured on a curved plate with visible features using conventional stereo-vision system and the proposed LCP-SLS single camera setup. A sinusoidal excitation of 75 Hz was applied to the shaker with amplitude 8 V and studied at 1000 fps frame rate.

References

[1] J. Baqersad, P. Poozesh, C. Niezrecki, P. Avitabile, Photogrammetry and optical methods in structural dynamics–a review, Mech. Syst. Signal Pr. 86 (2017) 17–34.

[2] S. Zhang, High-speed 3d shape measurement with structured light methods: A review, Opt. Laser Eng. 106 (2018) 119–131.

[3] M. Inaba, T. Hara, H. Inoue, A stereo viewer based on a single camera with view-control mechanisms, in: Proc. Int'l Conf. Robots and Systems, volume 3, 1993, pp. 1857–1865 vol.3.

[4] R. Wang, X. Li, Y. Zhang, Analysis and optimization of the stereo-system with a four-mirror adapter, J. Eur. Opt. Soc-Rapid 3 (2008).

[5] T. Durand-Texte, M. Melon, E. Simonetto, S. Durand, M.-H. Moulet, 3d vision method applied to measure the vibrations of non-flat items with a two-mirror adapter, in: J. Phys. Conf. Ser., volume 1149, IOP Publishing, 2018, p. 012008.

[6] J. Li, X. Dan, W. Xu, Y. Wang, G. Yang, L. Yang, 3d digital image correlation using single color camera pseudo-stereo system, Opt. Laser Technol. 95 (2017) 1–7.

[7] T. Durand-Texte, E. Simonetto, S. Durand, M. Melon, M.-H. Moulet, Vibration measurement using a pseudo-stereo system, target tracking and vision methods, Mech. Syst. Signal Pr. 118 (2019) 30–40.

[8] T. Durand-Texte, M. Melon, E. Simonetto, S. Durand, M.-H. Moulet, Single-camera single-axis vision method applied to measure vibrations, J. Sound Vib. 465 (2020) 115012.

[9] Z. Chen, X. Shao, X. Xu, X. He, Optimized digital speckle patterns for digital image correlation by consideration of both accuracy and efficiency, Appl. Optics 57 (2018) 884–893.

[10] S. Feng, L. Zhang, C. Zuo, T. Tao, Q. Chen, G. Gu, High dynamic range 3d measurements with fringe projection profilometry: a review, Meas. Sci. Technol. 29 (2018) 122001.

[11] T. Tao, Q. Chen, S. Feng, Y. Hu, M. Zhang, C. Zuo, High-precision real-time 3d shape measurement based on a quad-camera system, J. Opt-UK 20 (2017) 014009.

[12] W. Yin, S. Feng, T. Tao, L. Huang, M. Trusiak, Q. Chen, C. Zuo, High-speed 3d shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system, Opt. Express 27 (2019) 2411–2431.

[13] Z. Song, S. Tang, F. Gu, C. Shi, J. Feng, Doe-based structured-light method for accurate 3d sensing, Opt. Laser Eng. 120 (2019) 21–30.

[14] D. Sims-Waterhouse, S. Piano, R. Leach, Verification of micro-scale photogrammetry for smooth three-dimensional object measurement, Meas. Sci. Technol. 28 (2017) 055010.

[15] M. Dekiff, P. Berssenbrügge, B. Kemper, C. Denz, D. Dirksen, Three-dimensional data acquisition by digital correlation of projected speckle patterns, Appl. Phys. B 99 (2010) 449–456.

[16] M. Schaffer, M. Grosse, R. Kowarschik, High-speed pattern projection for three-dimensional shape measurement using laser speckles, Appl. Optics 49 (2010) 3622–3629.

[17] M. Schaffer, M. Grosse, B. Harendt, R. Kowarschik, High-speed three-dimensional shape measurements of objects with laser speckles and acousto-optical deflection, Opt. Lett. 36 (2011) 3097–3099.

[18] S. Zhang, P. S. Huang, Novel method for structured light system calibration, Opt. Eng. 45 (2006) 083601.

[19] D. Moreno, G. Taubin, Simple, accurate, and robust projector-camera calibration, in: 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, IEEE, 2012, pp. 464–471.

[20] L. Nie, Y. Ye, Z. Song, Method for calibration accuracy improvement of projector-camera-based structured light system, Opt. Eng. 56 (2017) 074101.

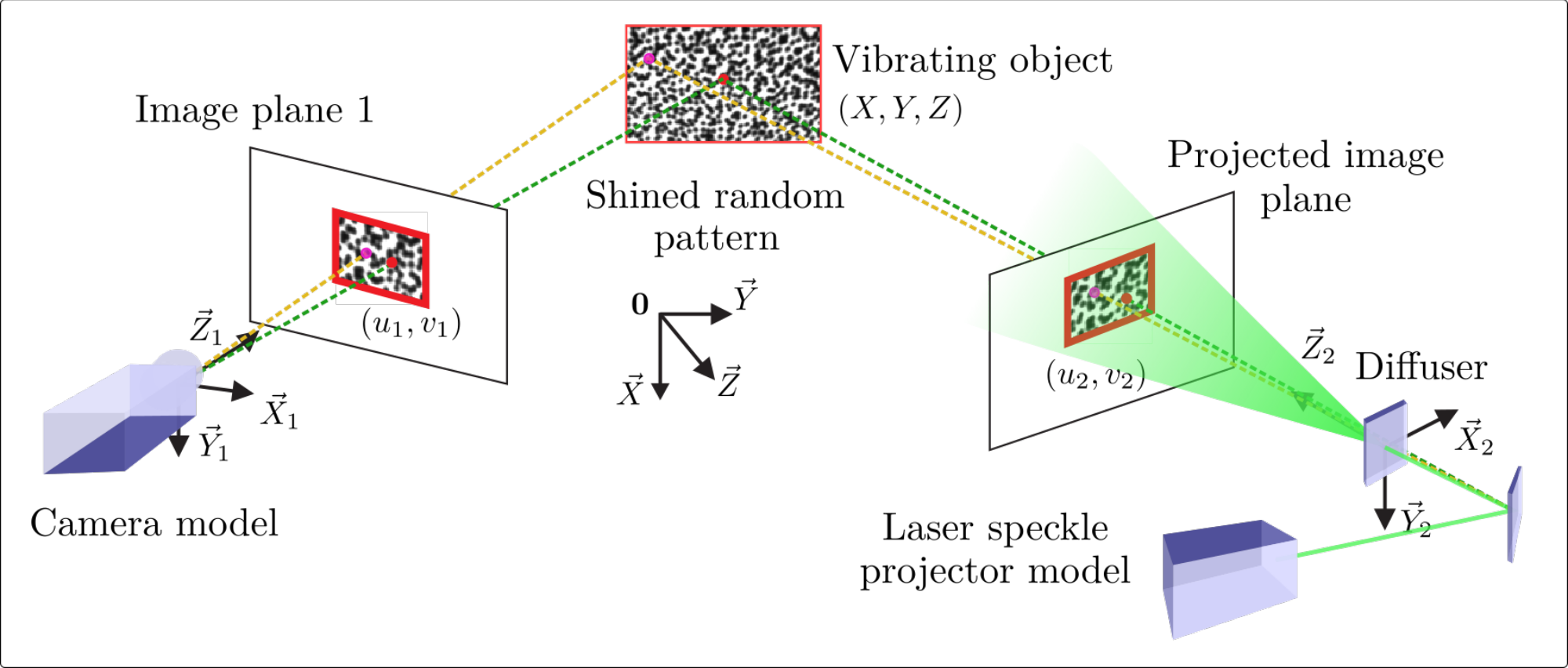
[21] A. Stark, E. Wong, D. Weigel, H. Babovsky, R. Kowarschik, Repeatable speckle projector for single-camera three-dimensional measurement, Opt. Eng. 57 (2018) 120501.

[22] Z. Zhang, A flexible new technique for camera calibration, IEEE T. Pattern Anal. 22 (2000) 1330–1334.

[23] G. Bradski, A. Kaehler, Learning OpenCV: Computer vision with the OpenCV library, " O'Reilly Media, Inc.", 2008.

[24] X. Shao, X. Dai, Z. Chen, Y. Dai, S. Dong, X. He, Calibration of stereo-digital image correlation for deformation measurement of large engineering components, Meas. Sci. Technol. 27 (2016) 125010.

[25] Y. An, T. Bell, B. Li, J. Xu, S. Zhang, Method for large-range structured light system calibration, Appl. Optics 55 (2016) 9563–9572.
[26] D. Sims-Waterhouse, P. Bointon, S. Piano, R. Leach, Experimental comparison of photogrammetry for additive manufactured parts with and without laser speckle projection, in: Proc. SPIE, volume 10329, Soc. Photo-Opt. Ins., 2017, p. 103290W.
[27] F. Zhong, R. Kumar, C. Quan, Rgb laser speckles based 3d profilometry, Appl. Phys. Lett. 114 (2019) 201104.
[28] A. W. Stark, E. Wong, D. Weigel, H. Babovsky, T. Schott, R. Kowarschik, Subjective speckle suppression in laser-based stereo photogrammetry, Opt. Eng. 55 (2016) 121713.
[29] D. Khan, M. A. Shirazi, M. Y. Kim, Single shot laser speckle based 3d acquisition system for medical applications, Opt. Laser Eng. 105 (2018) 43–53.
[30] K. Fu, Y. Xie, H. Jing, J. Zhu, Fast spatial–temporal stereo matching for 3d face reconstruction under speckle pattern projection, Image Vision Comput. 85 (2019) 36–45.
[31] W. Gander, J. Hrebicek, Solving problems in scientific computing using Maple and Matlab®, Springer Science & Business Media, 2011.


Vibrating object
(X, Y, Z)
Image plane 1
Shined random
pattern
Projected image
plane
$\vec{Z}_1$
(u_1, v_1)
$\vec{X}_1$
$\vec{Y}_1$
0
$\vec{Y}$
$\vec{X}$
$\vec{Z}$
$\vec{Z}_2$
Diffuser
(u_2, v_2)
$\vec{X}_2$
$\vec{Y}_2$
Camera model
Laser speckle
projector model