Linearized projector alignment system for structured light illumination systems using the quad target method

Matthew E. L. Jungwirth,* Eric P. Rudd, and Elena Yudovina
CyberOptics Corp, Minneapolis, Minnesota, United States

Abstract. Structured light illumination (SLI) systems can measure a target's two-dimensional height profile. High-quality imaging in SLI systems is typically dependent upon well-focused optical systems that can be difficult to achieve due to the nonlinear nature of defocus. We present the quad target method (QTM), a simple and fast alignment technique that linearizes defocus for SLI projectors. This linear response provides a clear zero crossing along the optical axis to minimize any focus ambiguity and maximize projector focus quality. QTM creates the linearization by sampling the field at two locations along the optical axis using a quad target, which is a diffuse target with two different heights in opposing quadrants. To maximize feedback speed, QTM can estimate the current focus position using only a single image by measuring the contrast of projected sinusoidal fringes on the quad target. QTM was tested using a commercial SLI system and compared with a published phase shifting reconstruction technique. The nonlinear defocus behavior was efficiently linearized as evidenced by focus estimation accuracies over a broad range of QTM variables. [DOI: 10.1117/1.OE.61.11.115101]

Keywords: optical alignment; focus linearization; Fourier filtering; structured light illumination.

Paper 20220845G received Jul. 29, 2022; accepted for publication Oct. 19, 2022; published online Nov. 10, 2022.

1 Introduction

1.1 *Overview*

Many applications find it beneficial to determine the height profile of a sample, producing a two-dimensional map of surface heights. One common method to produce height maps uses structured light illumination (SLI),[1,2] where a pattern of light from a projector illuminates the object and a camera images the scene. The pattern varies spatially and is further distorted by the sample's three-dimensional (3D) shape. Fringe projection in particular has gained prominence due to its noncontact measurement method and high-precision absolute phase maps.[3]

High-quality reconstructions in SLI systems are highly dependent upon well-focused optical systems. As with many optical systems, a prevalent issue in focusing SLI systems is the quadratic behavior of defocus near the focal plane. At this location, the slope of conventional figures of merit used to focus the system (e.g., spot size, wavefront error, and Strehl ratio) goes to zero, which makes precise focus difficult to discern. Linearizing defocus would solve this issue by maintaining the slope through best focus.

The idea of a linearized focus metric is not new. For instance, Volgenau et al.[4] described a technique for focusing an astronomical telescope that involves optimizing a sharpness metric based on the full width half maximum of star images. Because all stars appear at the same focal plane, it is necessary to collect exposures at multiple focus positions. The derivation of the optimal focal point then follows traditional techniques: a model function is fitted to the sharpness metrics through focus, and the peak of the linear model function is calculated. Reynolds et al.[5] described a technique for focusing a photolithography projector. They split the pupil, bring each beam to a focus, and derive a linear focus measure from the lateral disparity between the two images thus formed.

*Address all correspondence to Matthew E. L. Jungwirth, mejungwirth@gmail.com

This paper introduces the quad target method (QTM), an alignment technique for focusing SLI projectors that features a linearized response. With this innovation, the projector can be precisely focused to maximize focus quality. The key to linearizing defocus' nonlinear behavior is to sample the field at two locations along the optical axis achieved here using a quad target. The quad target is a 2×2 array of flat surfaces in pairs of quadrants at two different heights. Then, focus is efficiently linearized by measuring the difference in the chosen figure of merit between the two quad target heights. Linearization of defocus has two main benefits: it creates a clear zero crossing to minimize any ambiguity of the focus location and it provides easier feedback to the user, focusing the projector. Note that adjusting the tip/tilt of the projector's focal plane is outside the scope of this paper.

QTM could employ many different physical aspects for its figure of merit, e.g., the magnitude of the common three-phase reconstruction technique.[6] To maximize feedback speed, our chosen figure of merit is the contrast of projected sinusoidal fringes on the quad target. This simple height reconstruction approach lessens the number of required images from three (with three-phase reconstruction) to a single image, a 3× improvement in speed. The contrast is then quickly calculated with bandpass filtering on the known fringe frequency.

This paper is organized into five sections. The remainder of Sec. 1 gives brief backgrounds on SLI and the published theory we use for experimental comparison. Section 2 covers the QTM, including the layout, the reconstruction technique, the input variables, and the data processing method. Section 3 details the test system, describes the experimental methods and targets, and determines the focus position ground truth. Section 4 experimentally tests QTM over a broad range of input variables and compares the results with published theory.[6] Conclusions are found in Sec. 5.

1.2 *Structured Light Illumination*

One common method to perform 3D profilometry is SLI. A schematic of a general SLI system is shown in Fig. 1.[2] Here, the structured light projector images a particular pattern onto the target. Many patterns are found in the literature, such as binary,[7] gray level,[8] and colored bar patterns.[9] The rectangular target distorts the color bars in Fig. 1 due to its height above the substrate. A camera then images the entire scene. The combination of the camera image and a processing method enables the reconstruction of the target's 3D shape.

1.3 *Three-Phase Reconstruction*

One reconstruction method uses projected patterns with a sinusoidal profile. The fringes are constant along one direction; along the other direction, the gray level value varies sinusoidally. The desired results at each pixel are albedo, modulation, and the phase of the sinusoidal pattern that illuminates the corresponding target location. Because there are three unknowns, fringes in

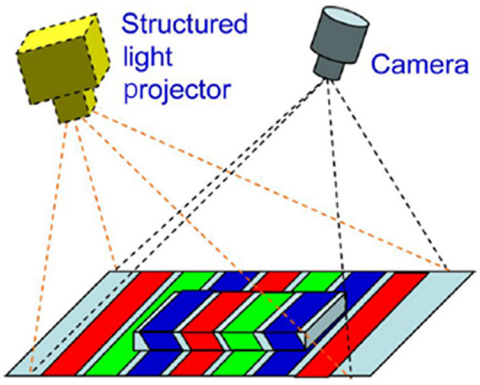

Fig. 1 Schematic of a general SLI system. Notice that the projected color bars are distorted by the target's shape. Reprinted with permission from Ref. 2 © The Optical Society.

three phases are required for unambiguous reconstruction. The three parameters can be computed with simple trigonometry based on the measured values of the images; however, phase shifts of 0 deg, 120 deg, and 240 deg are customary because of their analytical simplicity as well as their optimal numerical condition. Object height corresponds to a difference between the measured phase at a pixel and the phase that would be observed when measuring a flat plane at $Z = 0$.[6] In this paper, we refer to this technique as the "three-phase method."

2 Quad Target Method

2.1 *Layout*

The QTM linearizes defocus to provide a clear zero crossing and therefore minimize any ambiguity in determining the current focus position. The key to QTM is the quad target, a 2×2 array of flat surfaces at two heights, thereby allowing measurement of focus figures of merit at two locations along the optical axis simultaneously. Each of the diagonally opposed pairs of steps of the quad target is at roughly the same height above the substrate, so chosen to evenly sample the projector's field of view (FOV) and minimize the effect of target tilts and latent aberrations within the projector. The quad target is discussed further in Sec. 3.2.

A schematic of QTM is shown in Fig. 2. Two field points (red and blue) are imaged by the projector optics onto the quad target with two of the four steps shown in profile in Fig. 3. Notice that due to the quad target step height h_{step} (greatly exaggerated for ease in viewing), the blue field is in focus on the front step, and the red field defocused on the back step. A camera images the entire scene with the camera's front working distance labeled Z_1. The distance Z_2 is the projector's back working distance that can be physically altered to shift the projector's focus position during alignment.

Then, the nonlinear behavior of defocus is linearized by tracking the fringe contrast on the front and back steps and applying the normalization[11]

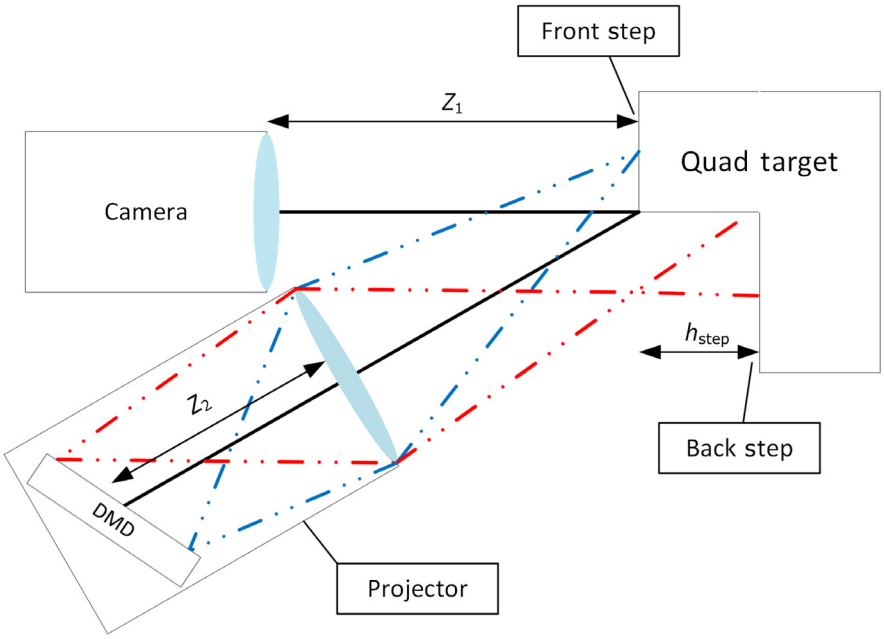

Fig. 2 Cross-section showing a finite conjugate SLI system and two steps of the quad target. The quad target step is exaggerated for ease in viewing. Notice that the blue field is focused on the front step and the red is defocused on the back step. Other aspects: the thick black line is the optical axis, the DMD produces the SLI pattern and is tilted to meet the Scheimpflug condition, distance Z_1 is the camera's target-space working distance, and distance Z_2 is the projector's back working distance.[10]

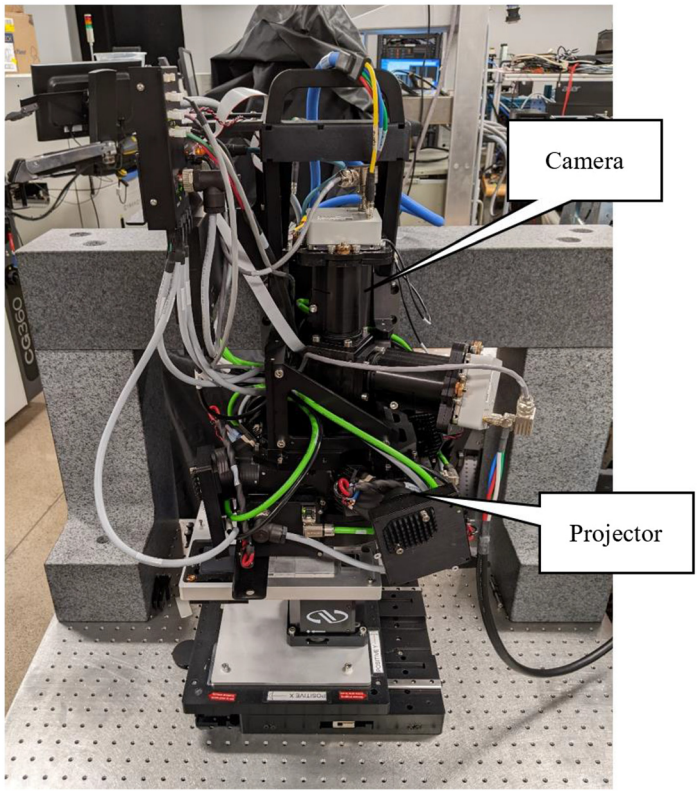

Fig. 3 The SLI system under test. The single projector/camera combination examined is labeled.

$$L = \frac{F - B}{F + B}, \tag{1}$$

where F and B are the figures of merit from the front and back steps, respectively, and L is the QTM linearized value, i.e., a linearized focus metric. As mentioned in Sec. 1.1, the figure of merit could be any physical attribute that can be imaged by the SLI camera, e.g., the phase magnitude from the three-phase method or the contrast from a projected fringe. Choosing the latter maximizes processing speed, as described further in Sec. 2.2.

2.2 *Reconstruction Method*

To determine the values F and B from Eq. (1), we apply a bandpass filter on the known fringe frequency using Gabor wavelets.[12] It is highly desirable to minimize the effects of image and target noise on the contrast estimates $C(R_N)$, where C is the measured fringe contrast and R_n is the region of interest (ROI). Because the spatial frequency of the fringes is approximately known, noise can be greatly reduced by bandpass filtering the image prior to the contrast estimate, which can be conveniently carried out by convolution of the image with a suitable Gabor wavelet. The envelope of the wavelet is smoothly tapered by a function designed to minimize out-of-band response, and the wavelet is also designed to have zero response to constant offsets, so that it responds only to the modulation. The Gabor wavelet has both real and imaginary components, which produce a complex response. The complex phase of the response gives information about the location of the fringes (which is essential for phase profilometry, but which is unnecessary here); however, the absolute value of the response gives us the contrast function $C(R_N)$ that we need for further processing. The techniques used to perform this convolution quickly on a computer are beyond the scope of this paper.

Next, F and B are determined by averaging the values from each pair of steps, specifically,

$$F = \frac{C(R_1) + C(R_3)}{2}, \tag{2}$$

and

$$B = \frac{C(R_2) + C(R_4)}{2}, \tag{3}$$

where R_N and $C(R_N)$ are as previously defined and N = (1, 2, 3, and 4) for the four quadrants of the quad target using the standard Cartesian definition (e.g., northeast is $N = 1$, etc.). The averaging not only reduces noise but also minimizes the effect of field-dependent aberrations. Note that Eqs. (2) and (3) do not assume that the contrasts are constant over R_N. In fact, QTM is robust with small variances within ROIs and can produce accurate focus estimations even for moderate tilts of the projector focal plane, as is shown in Sec. 3.4.

Equations (2) and (3) can be calculated from any image of projected fringes with the resultant values fully satisfying Eq. (1). In other words, only a single image is necessary for QTM to produce an estimate of the projector's focal position. Thus, QTM is 3× faster than three-phase reconstruction because three-phase requires three images, whereas QTM needs only one.

2.3 *Variable Ranges*

The performance of QTM is especially affected by two variables: the step height of the quad target h_{step} and the frequency of the projected sinusoidal fringes ξ_{project}. QTM determines the contrast on front and back steps, so if h_{step} is much greater than the camera's depth of field, the contrast on one of the steps might go to zero and cause poor linearity in Eq. (1). Alternatively, if h_{step} is too much smaller than camera's depth of field, the contrast difference between F and B is small, and the estimate, as provided by Eq. (1), is excessively noisy.

The projected frequencies in our SLI system are limited by the digital mirror device (DMD) that generates the fringes in the system. The maximum projected frequency is limited by the pitch of the DMD pixels to typically about half the DMD's Nyquist frequency. The minimum ξ_{project} is imposed by the decoding requirement that the Gabor wavelet fit completely within a quadrant.

2.4 *Through-Focus Calibration*

To align the projector, L must be calibrated with respect to the desired focus position Z_0. The calibration process shifts the distance Z_1 in Fig. 2 through a range of Z positions to measure the response of Eqs. (2) and (3) through focus. After applying Eq. (1) to each F and B value, L is now known at every point within the Z range. In other words, L is now a function of Z_1 or, proportionally, the projector's back working distance Z_2 because the projector is a finite conjugate system.

2.5 *Data Processing*

After the calibration process, we now have three through-focus curves: $F(Z)$, $B(Z)$, and $L(Z)$. The through-focus curves change shape based on the h_{step} and ξ_{project} combination employed, the residual aberrations of the optical sensor, and the experimental noise due to optomechanical tolerances or environmental effects. To minimize these effects, the raw contrast data are processed with the following method.

Step 1. The contrast values are measured for the four quadrants through the range of target plane positions (i.e., QTM calibration in Sec. 2.4).
Step 2. The raw $F(Z)$ and $B(Z)$ contrast curves are determined with Eqs. (2) and (3), respectively.
Step 3. Fifth-order polynomials are fitted to the $F(Z)$ and $B(Z)$ curves.
Step 4. The curves are normalized by the peak-fitted contrast.
Step 5. L is determined via Eq. (1) using the fitted $F(Z)$ and $B(Z)$ polynomials.
Step 6. A fifth-order polynomial is fitted to $L(Z)$.
Step 7. Nonlinear (higher-order) terms of $L(Z)$ are set to zero.
Step 8. Coefficients are saved for run-time correction.

The data processing above assumes that $Z_0 = 0$. When $Z_0 \neq 0$, Steps 6 and 7 must be modified to account for the displacement, thereby ensuring that the tangent line is computed at Z_0.

2.6 *Alignment Process*

After the calibration and raw data processing, we can estimate the current focus position using the linear fit of $L(Z)$

$$L \approx m_L Z_{\text{focus}} + b, \tag{4}$$

where m_L and b are the coefficients for linear slope and bias, respectively, garnered from the polynomial fit. By definition of QTM, the current focus position is where $L = 0$, so the current focus position Z_{focus} is estimated as

$$Z_{\text{focus}} = -\frac{b}{m_L}. \tag{5}$$

When altering Z_2 to align the projector, Z_{focus} has now shifted to a new position. Real-time estimates of Z_{focus} are produced with values determined in Sec. 2.5: the raw F and B contrasts are measured, the raw values are normalized by the peak fitted contrast in step 4, and the L value is determined with Eq. (1). Next, Z_{focus} is determined by solving for Z_{focus} in Eq. (4) or

$$Z_{\text{focus}} = \frac{L - b}{m_L}. \tag{6}$$

Finally, the user continues to adjust Z_2 to drive Z_{focus} to Z_o, thereby aligning the projector. Note that Z_o must be within the calibrated Z range from Sec. 2.4 for QTM to function.

3 Experimental Methods

3.1 *Layout*

A commercially available SLI system was used to experimentally test QTM. The commercial SLI sensor[13] shown in Fig. 3 has two cameras and four projectors, meaning that there are eight total SLI channels. Throughout this paper, however, only a single projector/camera combination was examined for brevity. A listing of the optical attributes of the sensor is found in Table 1. Note that the SLI system design is outside the scope of this paper.

In Table 1, two FOVs are listed—3D and raw. The 3D FOV is the lateral region, where the sensor can measure and generate 3D height maps, whereas the raw FOV is the full FOV of the camera. The raw FOV is slightly larger than the 3D FOV to ensure good alignment across the 3D FOV.

3.2 *Quad Target*

The essential experimental piece for QTM is the quad target itself. A manufactured quad target with $h_{\text{step}} = 500\ \mu$m is shown in Fig. 4, where the front surfaces are the first and third quadrants and the back surfaces are the second and fourth quadrants (following the standard Cartesian definition). The target was milled from a 50 mm × 50 mm aluminum block and was sand blasted to create a smooth diffuse surface.

The coordinate plane of QTM is shown in Fig. 5. The alignment ROIs (green boxes) used by Eqs. (2) and (3) are slightly larger than the 3D FOV (red box) to ensure good alignment across the 3D FOV. Also, the four alignment ROIs sample 51.6% of the raw FOV to minimize the effect of field-dependent aberrations on the focus estimation. Figure 5 also shows an example of a raw

Table 1 List of optical attributes for the custom SLI sensor.

Attribute	Value
3D FOV	25.0 mm × 25.0 mm
Raw FOV	26.3 mm × 26.3 mm
Target space camera numerical aperture	0.050
Target space projector numerical aperture	0.040
Tested wavelengths	440 to 460 nm
Camera depth of field	±200 μm
Maximum projected fringe frequency	13.7 cy/mm
Lateral pixel resolution	5.0 μm
Optical axis resolution	0.1 μm
Number of cameras	2
Number of projectors	4
Triangulation angle	30 deg

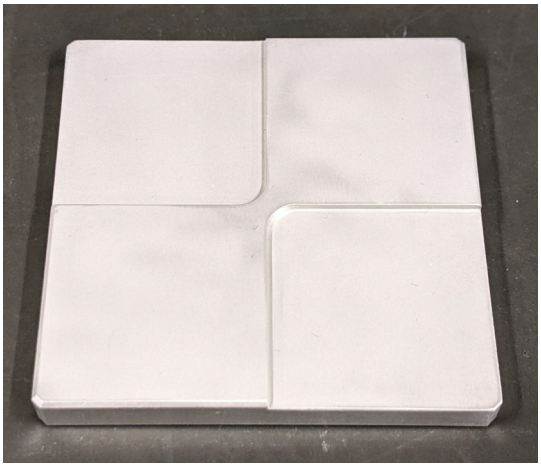

Fig. 4 Picture of quad target with $h_{\text{step}} = 500$ μm. Notice that the surface is diffuse.

image of the quad target illuminated with 3.4 cy/mm fringes. In this case, the front surfaces are in focus as evidenced by the specular glints from the granular diffuse surface of the quad target.

3.3 *Experimental Process*

To test QTM, experimental measurements of Z_{focus} were taken using the test system in Sec. 3.1. As shown in Table 2, nine different combinations of step heights and spatial frequencies were examined to sample the two QTM variables across wide ranges. Specifically, h_{step} covers 1×, 3×, and 5× of the camera's depth of field and ξ_{project} covers 0.25×, 0.50×, and 1× of the projector's maximum spatial frequency. For ease, the remainder of this paper will use qualitative descriptions to reference any of the nine combinations, e.g., $h_{\text{step}} = 200$ μm and $\xi_{\text{project}} = 13.7$ cy/mm in the upper left-hand corner of Table 2 will be referred to as {h-low, ξ-high}.

Figure 6 shows two of the three tested quad targets resting below the test sensor. The targets are displaced in Z to change the distance Z_1 in Fig. 2. For all experimental results in Sec. 5,

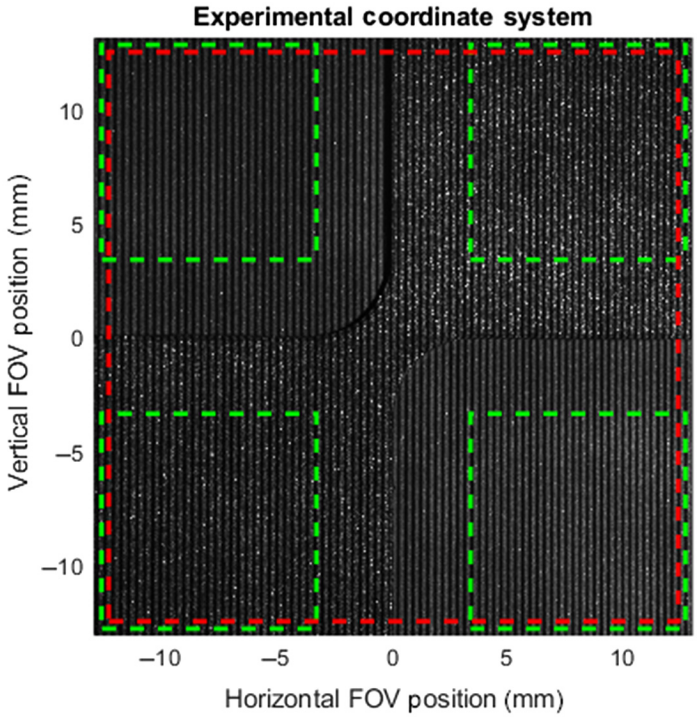

Fig. 5 Coordinate plane for experimental testing of QTM with 3.4 cy/mm fringes in focus on the front surface. The four green boxes are the four alignment ROIs, and the red box is the 3D FOV.

Table 2 Table of h_{step} and $\xi_{project}$ combinations tested both experimentally and in simulation.

		Experimental variable combinations		
Fringe frequency $\xi_{project}$	High	{200 μm, 13.7 cy/mm}	{500 μm, 13.7 cy/mm}	{1000 μm, 13.7 cy/mm}
	Mid	{200 μm, 6.9 cy/mm}	{500 μm, 6.9 cy/mm}	{1000 μm, 6.9 cy/mm}
	Low	{200 μm, 3.4 cy/mm}	{500 μm, 3.4 cy/mm}	{1000 μm, 3.4 cy/mm}
		Low	Mid	High
			Step height h_{step}	

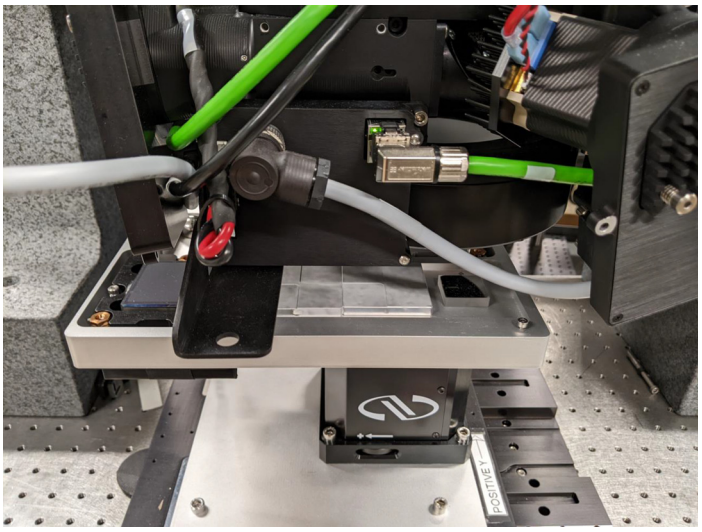

Fig. 6 Picture of quad target mounted on Z stage.

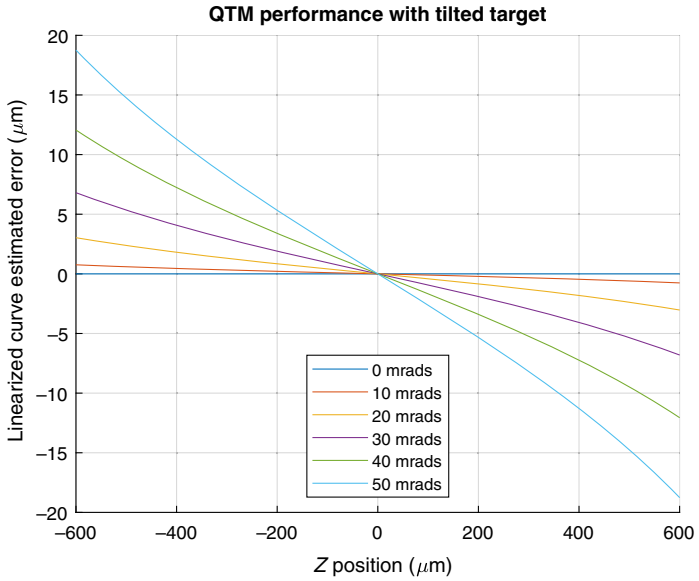

Fig. 7 Height errors for $L(Z)$ with varying target tilts across a ± 600 μm Z range. Notice that, due to symmetry, all height errors go to zero at $Z = 0$.

the range of Z values was ± 600 μm with a 50 μm step size. The stage is a Newport GTS30V and is controlled via a Matlab script.[14] Finally, the test sensor is rigidly mounted to a granite frame to reduce vibrations, as shown in Fig. 3.

3.4 *Target Tilt Sensitivity*

Because the tilt of a physical target will never be exactly zero, we conducted a theoretical study to see how QTM was affected by quad target tilt. Assuming an unaberrated lens system, but a realistic $C(Z)$ model based on the diffraction MTF of defocused sinusoidal fringes,[15] we computed the output of QTM for various target tilts. Figure 7 shows the estimated height difference (i.e., height error) of $L(Z)$ curves for a tilted quad target with respect to an untilted target over the Z range from Sec. 3.3. As can be seen from Fig. 7, tilting the target has minimal effects for tilts well beyond those likely to be encountered in a carefully constructed calibration fixture. In fact, a simple symmetry argument suffices to show that the curves all go through the origin, which is the most important factor in QTM's ability to accurately indicate proper focus. Thus, QTM is insensitive to small tilts of the target relative to the projector's focal plane.

3.5 *Quad Target Flatness*

Height images generated by the commercial test system of all three quad targets are shown in Fig. 8 in the same orientation as Fig. 4. The nominal step heights have been removed to enhance the visibility of small nonidealities. Notice that the quad targets are neither optically flat nor equally nonflat (Table 3), and that tooling marks are visible throughout the target surfaces and within each ROI. However, as Sec. 4 shows, QTM is relatively robust with respect to these minor manufacturing errors for several reasons: the effect of tooling marks is small because linearized function $L(Z)$ is computed from averages over large alignment ROIs, the nonflatness is small compared with the camera's depth of field, and the calibration process in Sec. 2.4 removes any constant bias induced by the tilts.

3.6 *Three-Phase Focus Estimation*

For the experimental results, the ground truth of the focus location was determined by three-phase reconstruction,[6] specifically using $\{h\text{-mid}, \xi\text{-high}\}$. The modulation from the three-phase

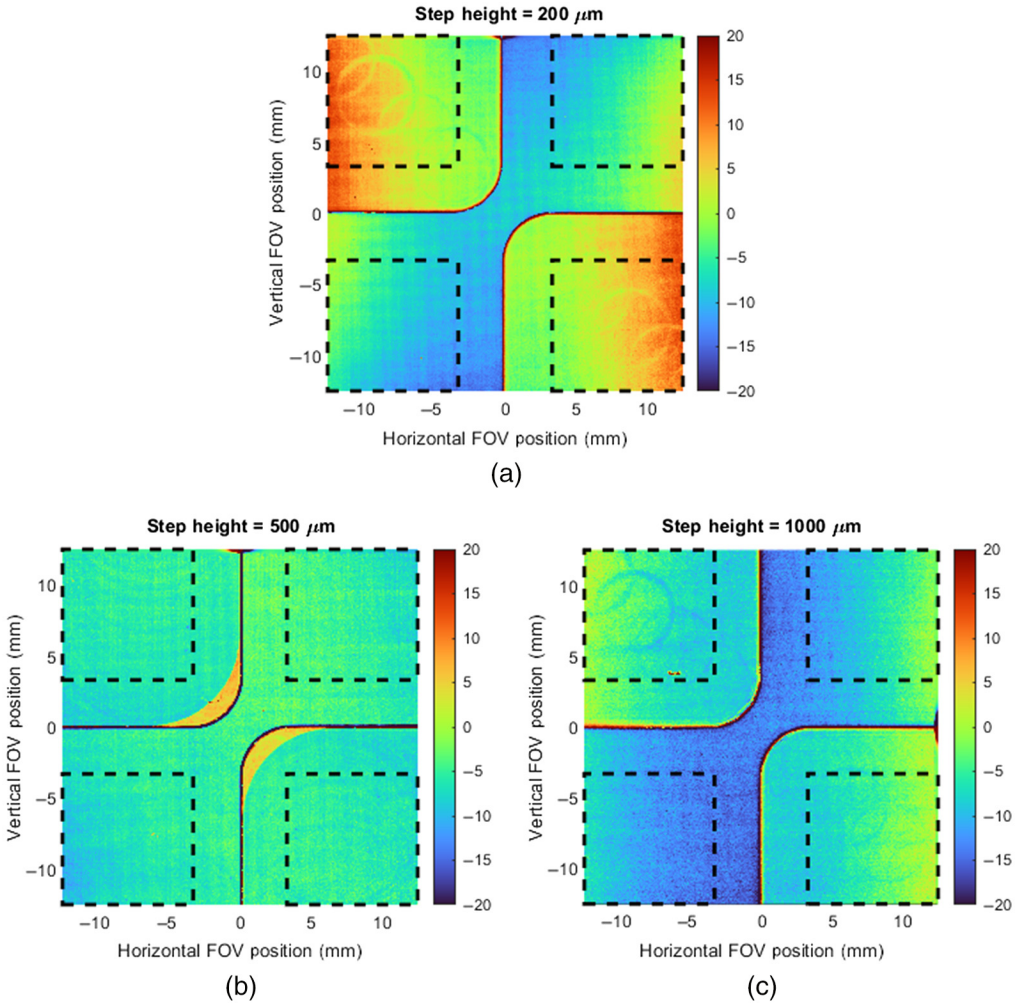

Fig. 8 Height maps of the three quad targets shown as deviation from the nominal step height: (a) $h_{\text{step}} = 200$ μm, (b) 500 μm, and (c) 1000 μm.

Table 3 Heights and surface flatnesses of test quad targets. The flatnesses are only calculated for the four ROIs.

Step height (μm)	Root mean square flatness in four ROIs (μm)
200	6.7
500	1.5
1000	4.0

method was measured on each of the four quadrants through focus and normalized by albedo. An example of the curves from one of the four trials is shown in Fig. 9(a). The peak modulation (i.e., the step height) was calculated by fitting these curves to a polynomial with the results in the four steps heights shown in Fig. 9(b). The ground-truth focus position is the mean of the four trials, determined to be -44.2 μm, the uncertainty is taken to be the standard deviation of the four trials, determined to be 1.5 μm.

Figure 9 also shows that the quad target was titled by 0.44 deg from the second to fourth quadrants as evidenced by the 138.3 μm height difference that is not observed in Fig. 8(b).

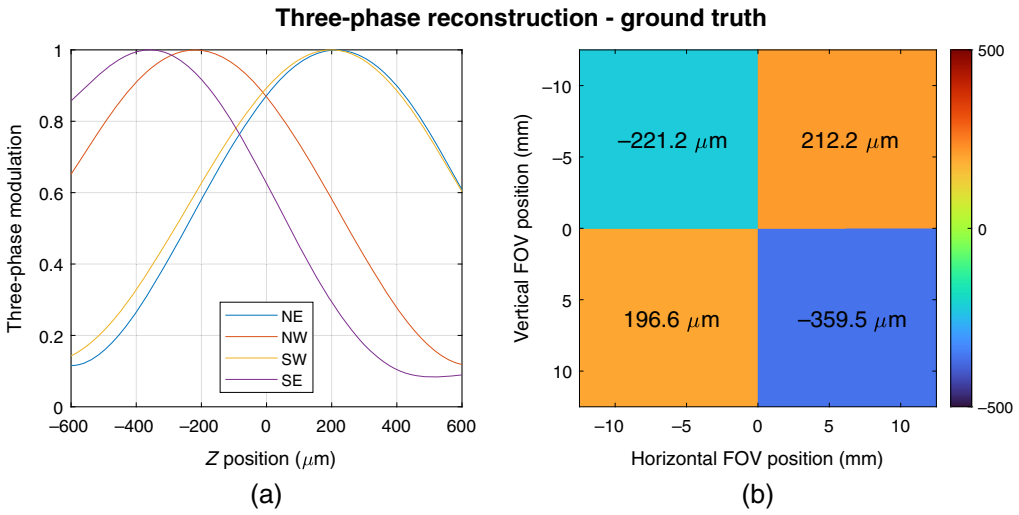

Fig. 9 Curves and focus positions from three-phase method using the {h-mid, ξ-high} combination: (a) through-focus curves and (b) quad target heights.

Again, as shown in Sec. 3.4, QTM is insensitive to small tilts so the quad target tilt errors are ignored.

4 Experimental Results

4.1 *Fitting Raw Values*

After the procedure in Sec. 3.3, Fig. 10 shows the raw and fitted data from {h-mid, ξ-mid} using QTM. The error bars are $\pm 2\sigma$ of the normalized contrast values over four experimental trials. The fitted lines track the raw data well as the fits are mainly inside the error bars except near the peak contrasts. Here, small contrast differences are difficult for Eqs. (1)–(3) to discern likely due to the specular glints visible in Fig. 5. This highlights the need for the polynomial fitting in Sec. 2.5 as it reduces the noise from these glints and other sources.

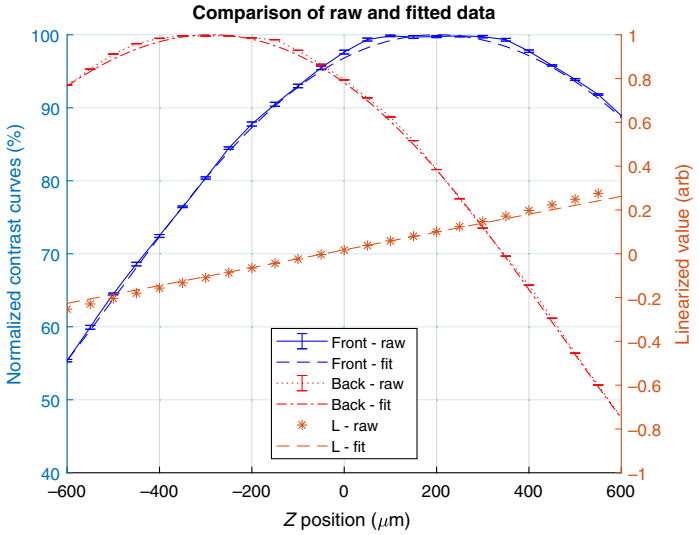

Fig. 10 Experimental data for {h-mid, ξ-mid} with raw data and fits following the data processing method in Sec. 2.5. The error bars are $\pm 2\sigma$.

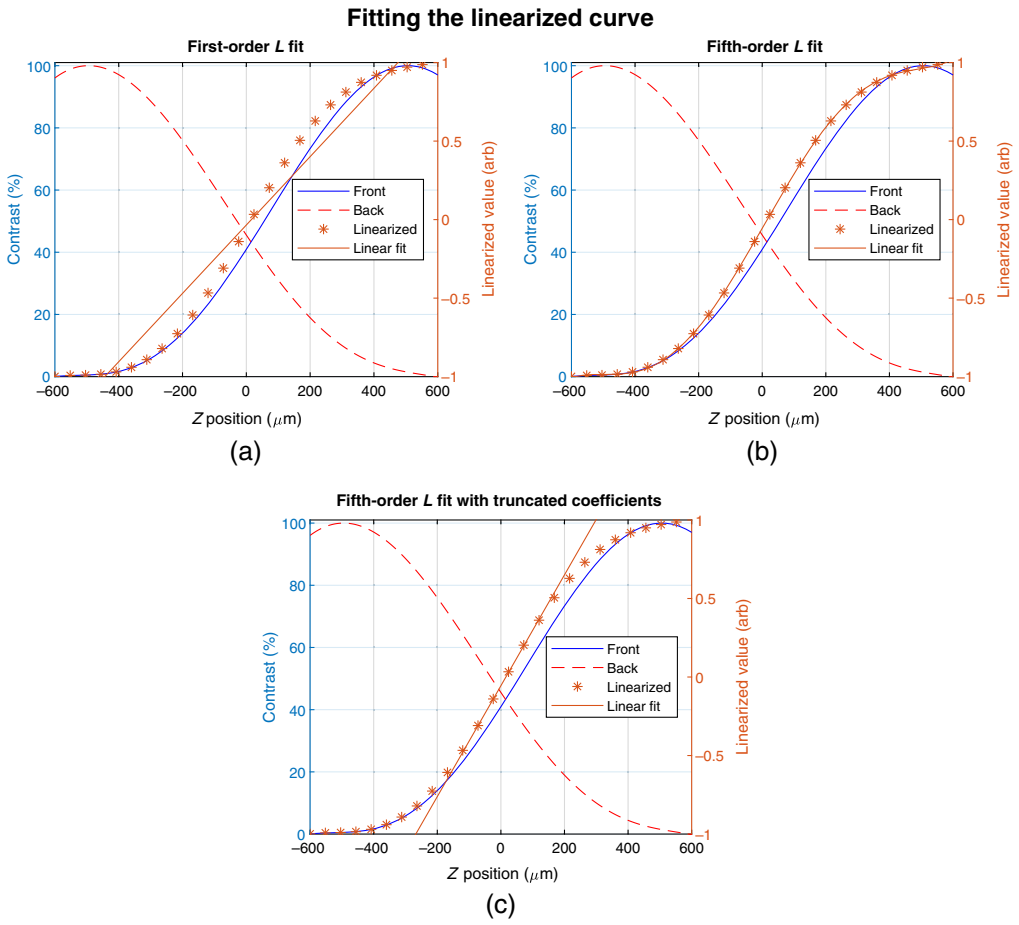

Fig. 11 Fitted data for {h-high, ξ-high}: (a) using only a linear regression for $L(Z)$, using a (b) fifth-order fit, and (c) set nonlinear coefficients of fifth-order fit to zero.

4.2 *Linearized Curve Fit*

The quality of the $L(Z)$ curve fit is critical to the accuracy of the focus estimation. Specifically, step 6 in Sec. 2.5 ensures that the regression of $L(Z)$ is accurate. Figure 10 visually shows that $L(Z)$ is well approximated by a linear fit for {h-mid, ξ-mid}. However, changing to {h-high, ξ-high}, as shown in Fig. 11(a), we see that $L(Z)$ (orange asterisk) is not linear over the tested Z range in all cases. Thus, if the regression is purely linear (orange line), the shape of $L(Z)$ is not well captured and could lead to focus estimation errors. When a fifth-order polynomial is applied, as shown in Fig. 11(b), the shape of $L(Z)$ is preserved throughout the Z range.

Then, step 7 in Sec. 2.5 ensures that $L(Z)$ gives a linear response near the desired focus at Z_0. The fifth-order fit in Fig. 11(b) does an excellent job of tracking the shape of $L(Z)$. However, in terms of alignment, this means that a user will observe nonlinear adjustment feedback for $|Z| > \sim 200$ μm. By setting the second- to fifth-order coefficients to zero, as shown in Fig. 11(c), linear feedback is maintained for the usable range of alignment, i.e., where the F and B values are greater than zero, as discussed in Sec. 2.3.

4.3 *QTM Variable Extremes*

The variable ranges of h_{step} and ξ_{project} are further probed in Fig. 12, where Figs. 12(a)–12(d) visually display the extremes of the QTM variables. For higher ξ_{project}, the center of the Z range lies, where the $F(Z)$ and $B(Z)$ contrast curves are not only less suitably sloped but small. In fact, the contrast goes to zero in Fig. 12(b) at the edges of the tested Z range. These problems cause the $L(Z)$ curve to be not only nonlinear but also with a slope that is almost zero beyond a certain

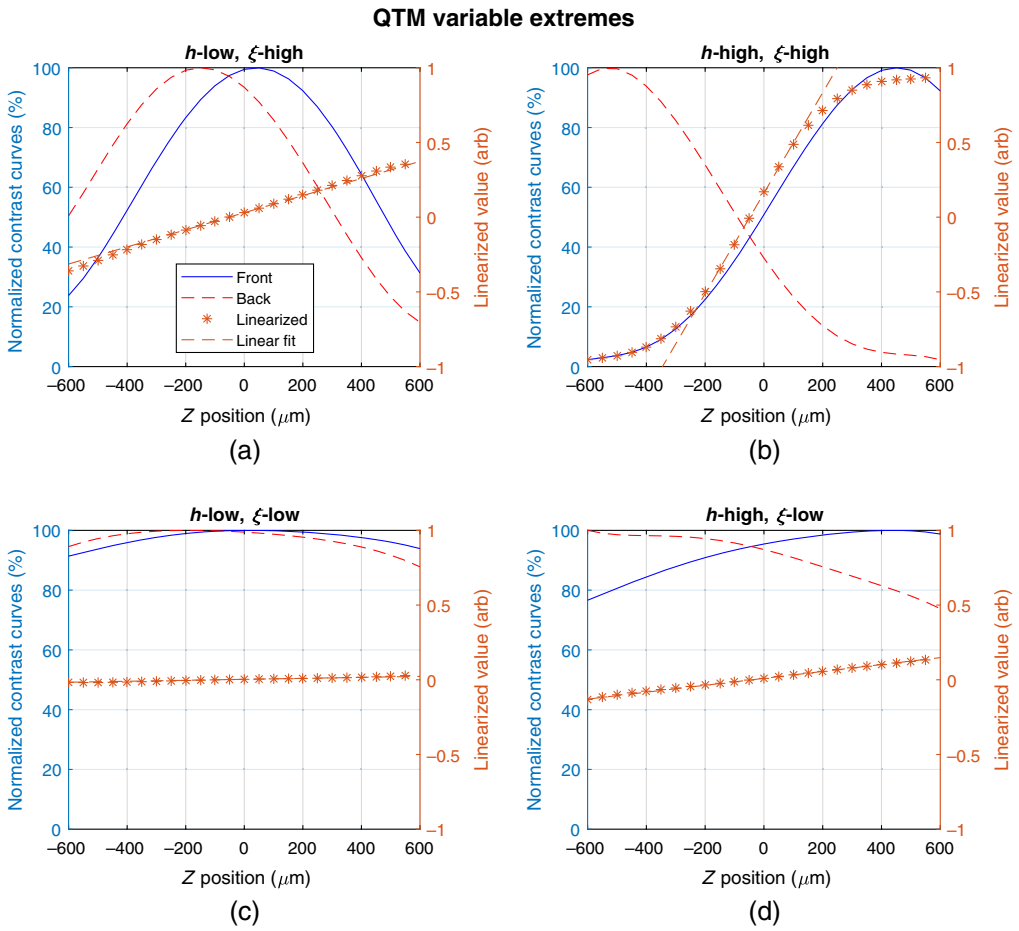

Fig. 12 Comparison of experimental results for variable extremes. Panels (a)–(d) are labeled according to Table 2 and follow the legend in panel (a).

point. This point marks the edge of the usable Z range for the $\{h$-high, ξ-high$\}$ combination in Fig. 12(b) to $|Z| \sim 300$ μm.

Figure 12(c) shows another combination, where the slope of $L(Z)$ is low. Here, due to the small dynamic range of $F(Z)$ and $B(Z)$ contrasts, the slope is nearly zero and also noisy even near $Z = 0$. The result is a $L(Z)$ function that is unsuitable for determining focus position.

Overall, despite the data processing method in Sec. 2.5, Fig. 12 shows that $L(Z)$ is only truly linear for specific variable combinations, such as the $\{h$-low, ξ-low$\}$, in Fig. 12(c). For the other three parts of Fig. 12, the linearized values (orange asterisk) diverge from the fit line (orange dotted line) as $|Z| \to 600$ μm. However, QTM does not require that $L(Z)$ be exactly linear, but only monotonic, which reduces the determination of desired focus to a zero-finding problem rather than the more difficult sharpness optimization problem. Additional advantages accrue if $L(Z)$ is close to linear because correcting the estimated Z error will rapidly converge to desired focus.

It is worth comparing this to the behavior of the three-phase method, which aims to identify the location of the peak in the contrast curves. Figures 13(b) and 13(d) show that, at high projected frequencies, the contrast curves have clean and symmetric peaks, and the contrast curves measured with QTM (Gabor wavelets) and the three-phase method are nearly identical. (The two measurements start diverging at low contrast values, but that is not important to the estimation.) Because the peaks are symmetric, the midpoint between the peak locations is at the same Z value as the intersection of the contrast curves.

In contrast, at low projected frequencies [Figs. 13(c) and 13(d)], the peaks no longer look symmetric, and small differences in the two methods of estimating contrast become significant. Estimating the position of the peak reliably becomes difficult; for example, in Fig. 13(d), the

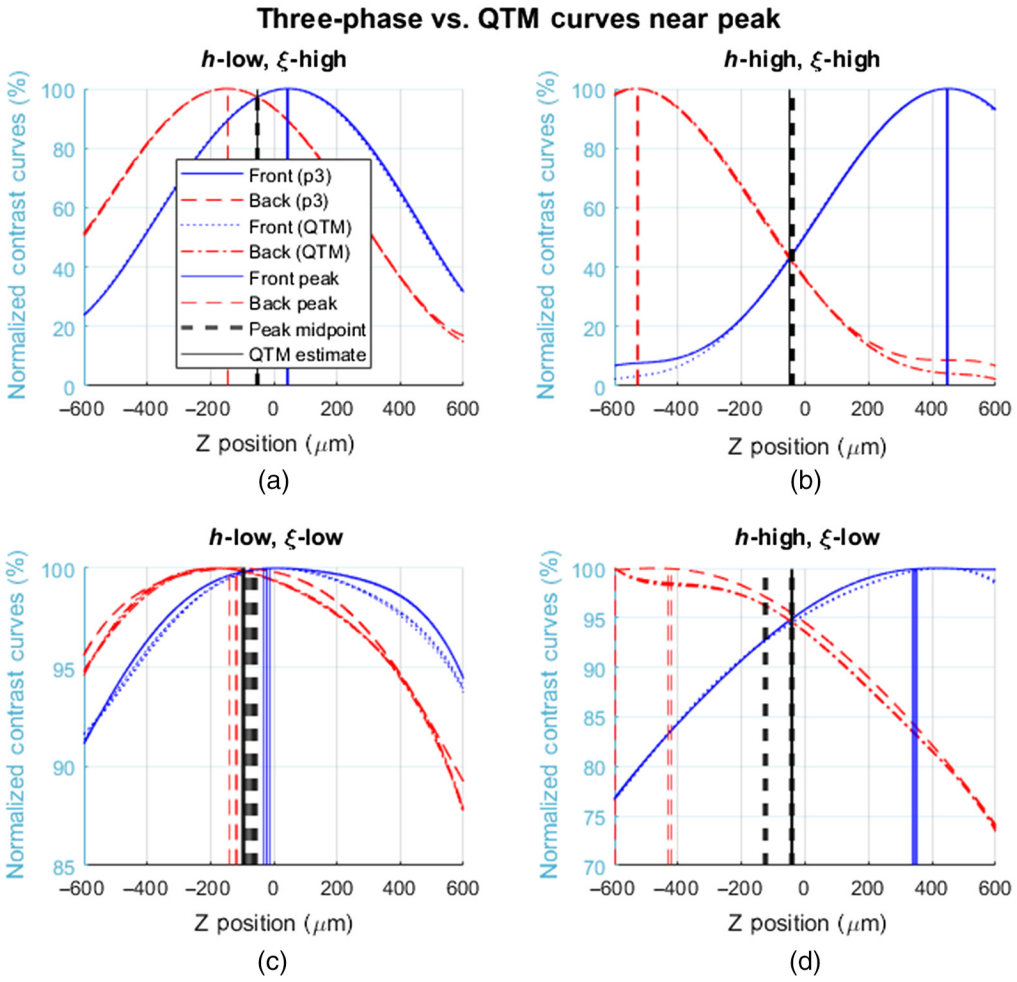

Fig. 13 Comparison of contrast curve shapes and peak estimate positions. Panels (a)–(d) are labeled according to Table 2 and follow the legend in panel (a). Note that the *y*-axes in the four plots are unequal. Note that in panel (c), one of the estimated peaks for the three-phase method is ending up at the bottom of the depth range (−600 μm).

back peak position is identified as −600 μm on one of the four trials of the data. When the separation between the true peaks is large enough, as shown in Fig. 13(c), the QTM is still able to deliver a reasonable estimate of the current focus position, even though the three-phase method would fail in this regime. However, if both the projected frequency ξ_{project} and the separation between the two peaks h_{step} are small, as shown in Fig. 13(d), QTM will also fail.

Note that there are two major differences between the three-phase method and QTM. First, a single image is used in QTM to obtain a contrast measurement at each height rather than three images for the three-phase method. Second, QTM looks for the zero crossing of a (nearly) linear function, whereas the three-phase method we have described looks for the location of the maximum of a nonlinear function. It is certainly possible to use linearization on the contrast measurements obtained with the three-phase method. This would have the robustness and convergence benefits of QTM but would miss out on the speed improvements from using one image per Z position.

Because the usual three-phase reconstruction yields a phase that varies approximately linearly with Z, one might wonder whether such a phase estimate might be a better measure of defocus than QTM. The phase from a three-phase reconstruction does indeed produce excellent measures of Z differences, but this potential scheme is intrinsically unworkable for establishing best focus because the zero point of the phase is determined by the lateral registration of the fringes, which is arbitrary; thus, there is nothing to tie a phase reference to the Z location of best focus.

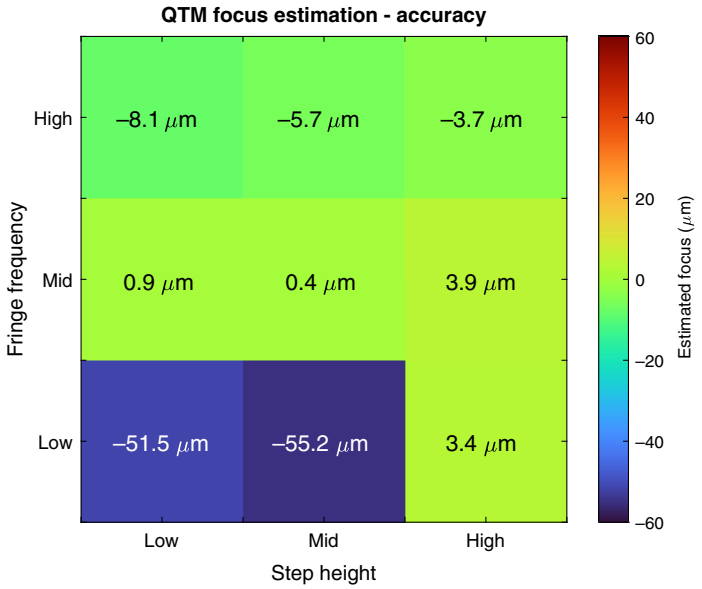

Fig. 14 Comparison of the QTM estimated current focus locations for the nine runs in Table 2. The three-phase estimated focus for {*h*-mid, ξ-low} from Sec. 3.6 was subtracted from all QTM values.

4.4 *Comparison to Three-Phase Reconstruction*

To test QTM's accuracy across variable ranges, the current focus position Z_{focus} was estimated using Eq. (5) for all nine runs in Table 2 and compared with the three-phase value from Sec. 3.6. Figure 14 shows the difference between the QTM and three-phase Z_{focus} values, and Table 4 lists the differences. For all nine runs, the mean difference is -12.9 μm, although this value is dominated by the {*h*-low, ξ-low} and {*h*-mid, ξ-low} combinations. For the other seven runs, the agreement with published theory is very good with -1.3 μm mean difference. Thus, Fig. 14 suggests that the {*h*-low, ξ-low} and {*h*-mid, ξ-low} combinations are outside the reasonable operating range of QTM.

For the {*h*-low, ξ-low} and {*h*-mid, ξ-low} combinations, QTM provides a much less accurate estimate of focus position, as predicted by Sec. 4.3. In these runs, QTM is suffering from a lack of contrast dynamic range and specular glints from the quad target. Figure 13(c) shows the raw unfitted curves for {*h*-low, ξ-low} and shows the deterioration when the parameters are unfavorable. There is a peak in the curves but the slope of the curves is so low that the peak is difficult to distinguish. Even though the acquisition and data reduction are, in some sense, still "correct," in this instance, the low spatial frequency of the pattern causes so much uncertainty that the sign of the step in the quad target is not even certain.

The results in Fig. 14 highlight the inherent trade-off in QTM: high resolution, low Z range or low resolution, and broad Z range. Certain SLI systems that desire a broad Z range and can handle looser focus position tolerances would find these two runs useful. However, most

Table 4 Listing of the differences from of the estimated current focus locations from Fig. 14.

		Focus estimation differences (μm)		
Fringe frequency ξ_{project}	High	–8.1	–5.7	–3.7
	Mid	0.9	0.4	3.9
	Low	–51.5	–55.2	3.4
		Low	Mid	High
			Step height h_{step}	

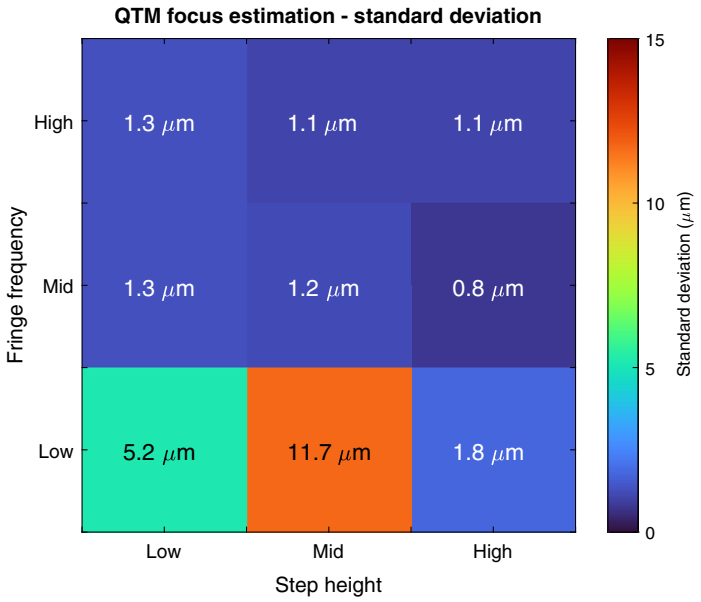

Fig. 15 Comparison of the QTM estimated current focus locations standard deviations for the nine runs in Table 2.

SLI systems require high precision and thus should avoid the {h-low, ξ-low} and {h-mid, ξ-low} variable combinations.

For the other seven variable combinations, the small mean difference in focus estimation is further proof of QTM's robustness. Despite the quad targets being nonflat (Table 3) and tilted (Fig. 9), QTM can match published theory to only 0.6% of the camera's depth of field. Therefore, the results suggest that the quad target can be inexpensively fabricated and also does not require high mounting tolerances while still providing accurate focus position estimations.

One method to detect a low accuracy variable combination is by looking at the standard deviation of several experimental trials, as shown in Fig. 15. Here, the {h-low, ξ-low} and {h-mid, ξ-low} combinations have $\sigma > 5.2$ μm for the four trials. For the other seven runs, the next lowest standard deviation is 1.8 μm. Thus, a focus estimation standard deviation >2 μm seems to indicate a variable combination outside the reasonable QTM operating regime.

5 Conclusions

Focus linearization for SLI systems has been realized using QTM. QTM samples the field at two points along the optical axis using the quad target, an array of flat surfaces at two different heights, to estimate the focal position using only a single image. Thus, QTM is a fast, accurate, and linear focus technique to maximize the projector focus quality. The algorithm tracks the projected fringe contrast to determine focal position and utilizes polynomial fits to minimize the effects of noise and experimental nonlinearities. Using a commercial SLI system, QTM efficiently linearized defocus across a range of quad target step heights and projected fringe frequencies. QTM agreed well with a published phase shifted height reconstruction method with a mean positional accuracy of -1.3 μm over the usable range of QTM. Finally, the QTM is broadly applicable because the quad target itself can be inexpensively fabricated.

Acknowledgments

The authors would like thank J. Konicek, B. Christensen, S. Schmidt, and J. Sapp for fabricating the quad targets. The authors would like to acknowledge E. Jungwirth, T. Skunes, and M. Schoeneck for helpful comments on this manuscript. The authors have no conflicts of interest to report. This paper is an expanded version of Ref. 10.

References

1. J. Wang and Y. Liang, "Generation and detection of structured light: a review," *Front. Phys.* **9**, 688284 (2021).
2. J. Geng, "Structured-light 3D surface imaging: a tutorial," *Adv. Opt. Photonics* **3**, 128–160 (2011).
3. W. Li et al., "Absolute phase retrieval for a single-shot fringe projection profilometry based on deep learning," *Opt. Eng.* **60**(6), 064104 (2021).
4. N. Volgenau et al., "Improving the focus of Las Cumbres Observatory's 1-meter telescopes," *Proc. SPIE* **11449**, 1144929 (2020).
5. G. Reynolds, M. Hammond, and L. A. Binns, "Camera correlation focus: an image-based focusing technique," *Proc. SPIE* **5752**, 1033–1041 (2005).
6. J. Salvi, S. Fernandez, and T. Pribanic, "A state of the art in structured light patterns for surface profilometry," *Pattern Recognit.* **43**(8), 2666–2680 (2010).
7. I. Ishii et al., "High-speed 3D image acquisition using coded structured light projection," in *IEEE/RSJ Int. Conf. Intell. Rob. and Syst.*, pp. 925–930 (2007).
8. S. Inokuchi, K. Sato, and F. Matsuda, "Range-imaging for 3-D object recognition," in *Int. Conf. Pattern Recognit.*, pp. 806–808 (1984).
9. K. L. Boyer and A. C. Kak, "Color-encoded structured light for rapid active ranging," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**(1), 14–28 (1987).
10. M. E. L. Jungwirth and E. Rudd, "Linearization of defocus for projector optical alignment in structured light illumination systems," *Proc. SPIE* **11488**, 1148806 (2020).
11. M. Watanabe, S. K. Nayar, and M. Noguchi, "Real-time computation of depth from defocus," *Proc. SPIE* **2599** (*1996*).
12. D. Gabor, "Theory of communication. Part 1: the analysis of information," *J. Inst. Electr. Eng.-Part III: Radio Commun. Eng.* **93**(26), 429–441 (1946).
13. *Multi-Reflection Suppression™ (MRSTM) Sensor Technology*, CyberOptics Corp, Minneapolis, Minnesota. www.cyberoptics.com.
14. *Matlab, Version 9.12.0.1927505 (R2022a) Update 1*, Mathworks, Natick, Massachusetts (2022).
15. H. C. Lee, "Review of image-blur models in a photographic system using the principles of optics," *Opt. Eng.* **29**(5), 405–421 (1990).

Matthew E. L. Jungwirth received his PhD from the University of Arizona in 2012, studying adaptive optical zoom using carbon fiber mirrors. Currently, he is working as a senior optical scientist at CyberOptics, Minneapolis, Minnesota, United States. His patents and publications cover a diverse array of topics, such as barcode sensing, quantum information processing, and star tracking technology. He is an SPIE Senior Member and Associate Editor of both *Optical Engineering* and the *Spotlight* book series.

Eric P. Rudd received his PhD in electrical engineering from the University of Minnesota in 1986 and currently holds 27 patents in the fields of optics, algorithms, and electronics. Recently, he is retired from CyberOptics, where he worked in the fields of electrical engineering, algorithm development, and optics.

Elena Yudovina received her PhD in mathematics from the University of Cambridge in 2012. Currently, she is working as an algorithm scientist at CyberOptics.