Total internal reflection for precision small-angle measurement

Aiyu Zhang and Peisen S. Huang

A method for precision small-angle measurement is proposed. This method is based on the total-internal-reflection effect of a light beam at a pair of glass prisms. Angular displacement of the light beam is measured when the intensity change of the reflected beam is detected as a result of the relative phase shift between the *s*- and the *p*-polarized beams. An initial phase shift between the *s*- and the *p*-polarized components is introduced to increase measurement sensitivity. For increased measurement linearity and reduced effect of laser power fluctuation on the output, a differential method is used in which the light beam is split equally into two beams, each reflected at a prism and detected by a photodiode. The output is obtained as the difference of the two detected intensities divided by their sum. A prototype device was built, which demonstrated a nonlinearity error of 1.3% in a measurement range of ±0.6° or 0.4% in ±0.3°. The peak-to-peak noise level was found to be at approximately 0.5 arc sec. This noise level can be reduced further and resolution increased by a reduction of the measurement range.

OCIS codes: 120.0120, 120.3930, 120.3940, 120.5700, 260.3160, 260.5430, 260.6970.

1. Introduction

Optical sensors for precision small-angle measurement have wide applications in many different fields. For example, they can be used for precision alignment of mechanical systems, calibration of machine tools and coordinate measuring machines, etc. Traditional optical instruments for small-angle measurement based on heterodyne interferometry and autocollimation cannot be used as sensors because they are usually bulky and expensive. Huang *et al.* previously have developed several new methods based on the internal-reflection effect, which allowed for much more compact and lower-cost devices to be built.[1–5] These methods have so far been applied successfully to machine tool calibration,[6,7] surface profile measurement of mirrors,[8] calibration of polygon mirrors,[9] and the development of an atomic force microscope[10] and a nanoradian angle sensor.[11]

When this research was performed, the authors were with the Department of Mechanical Engineering, State University of New York at Stony Brook, Stony Brook, New York 11794-2300. A. Zhang is now with the Department of Electrical Engineering, Pennsylvania State University, University Park, Pennsylvania 16802. P. S. Huang's e-mail address is peisen.huang@sunysb.edu.

Received 14 August 2000; revised manuscript received 7 December 2000.
0003-6935/01/101617-06$15.00/0

In this paper we propose another new method based on the total-internal-reflection effect of light for precision small-angle measurement. When a light beam undergoes total internal reflection at an interface of two media, its phase changes with the angle of incidence. The amount of this phase change depends on the polarization state of the light beam. The relative phase shift between the two polarization states is a function of the angle of incidence. Therefore, if this relative phase shift is detected, the angle of incidence and its changes can be measured. Previously, others have used heterodyne interferometry methods to measure this phase shift for small-angle measurement.[12–16] These methods all require a double-frequency laser as the light source, which makes the system expensive and complicated. Also only limited resolution can be achieved because the beating frequency of the interference signal is high and the resolution of the phase meter is limited. In this research we propose a new method that measures the intensity changes caused by this relative phase shift for angle measurement instead of trying to measure the phase shift directly. The result is a much simpler device with higher resolution and better linearity.

In Section 2 we describe the principle of the new method. In Section 3 we discuss the initial phase-shifting method for sensitivity improvement. In Section 4 we present the experimental results of a prototype device. The possible error sources of the

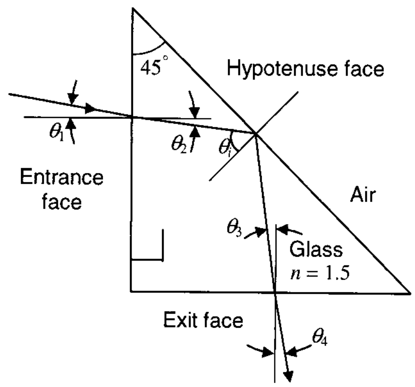

Fig. 1. Light propagation through a right-angle prism.

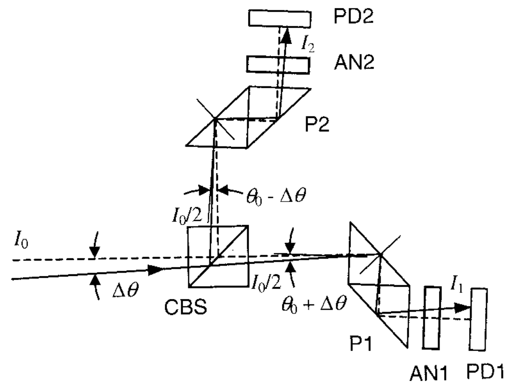

Fig. 3. Principle of the total-internal-reflection method for small-angle measurement.

method are discussed in Section 5, and the conclusions are given in Section 6.

2. Principle

Figure 1 shows how light propagates through a right-angle prism with the angle of incidence at the hypotenuse face greater than the critical angle. In this case, light is totally and internally reflected. According to Fresnel's law of reflection and refraction, the reflectance of light is unchanged but the phase changes with the angle of incidence under the condition of total internal reflection. This phase shift depends on the angle of incidence and the polarization state of the incident light. The relative phase shift ϕ between the p- and the s-polarized components can be written as

$$\phi = 2 \tan^{-1}\left[\frac{(\sin^2 \theta_i - 1/n^2)^{1/2}}{\tan \theta_i \sin \theta_i}\right], \tag{1}$$

where θ_i is the angle of incidence and n is the index of refraction. Figure 2 is a plot of ϕ as a function of θ_i when $n = 1.5$. It can be seen that, in the vicinity of the critical angle (41.8°), ϕ is highly sensitive to changes in θ_i. If the refraction at the entrance face of the prism is also considered, the following equation can be derived based on Snell's law:

$$\theta_i = \frac{\pi}{4} + \sin^{-1}\left(\frac{\sin \theta_1}{n}\right), \tag{2}$$

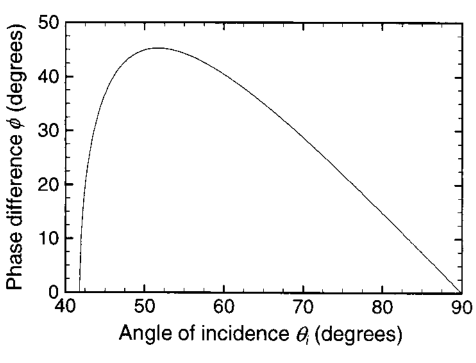

Fig. 2. Relative phase shift between the s- and the p-polarized components.

where θ_1 is the angle of incidence at the entrance face. When θ_i is in the vicinity of the critical angle, θ_1 is small and hence the transmittance is virtually unchanged.

The relationship between the relative phase shift ϕ and the angle of incidence θ_i can be used for precision angle measurement. Previously, several methods have been proposed that measure ϕ as a function of θ_i by use of heterodyne interferometry. By measuring the phase, one has the advantage of being able to avoid the effects of ambient light and the instability of the laser output power. However, it also has the disadvantage of making the instrument more complicated, bulky, and expensive because a double-frequency stable laser source has to be used. The size of a double-frequency laser prohibits the phase measuring technique from being used to build sensors for small-angle measurement. In this research we propose a new method that measures the intensity change of the light beam caused by this relative phase shift. With this method, a simple laser diode can be used as the light source. Analysis showed that this new method could provide much better linearity and measurement sensitivity. The ambient light and laser instability problems can be solved easily when the sensor is mounted in an enclosed housing and a differential method is used.

Figure 3 shows the principle of the newly proposed angle measurement sensor. A linearly polarized laser beam from a laser diode is used as the light source. The direction of polarization is set at 45° from the horizontal plane (the plane of the paper). The laser beam can be considered as the combination of s- and p-polarized beams with equal intensity and phase. The beam is split equally into two beams at a nonpolarizing cube beam splitter (CBS). Each beam is then totally and internally reflected twice at a prism (P1, P2) and passes through an analyzer (AN1, AN2) with its axis located at 45° from the axes of both the s- and the p-polarization directions. Each beam is finally detected by a photodiode (PD1, PD2) for its intensity. The prisms are each made of two right-angle prisms. The function of the analyzer is to bring the two polarized components to interfere with each other. As a result of this interference, the beam intensity changes with the relative phase shift ϕ and therefore the angle of incidence θ_i. Assume that the intensity of the incident beam is I_0. The intensities detected by photodiodes 1 and 2 are of the forms

$$I_1 = (I_0/2)(1 + \cos \phi_1), \tag{3}$$

$$I_2 = (I_0/2)(1 + \cos \phi_2), \tag{4}$$

where ϕ_1 and ϕ_2 are the relative phase shifts for the light beams reflected at prisms 1 and 2, respectively. According to Eq. (1) and considering the fact that each light beam is reflected twice at its corresponding prism, we can obtain ϕ_1 and ϕ_2 as

$$\phi_1 = 4 \tan^{-1}\left[\frac{(\sin^2 \theta_{i1} - 1/n^2)^{1/2}}{\tan \theta_{i1} \sin \theta_{i1}}\right], \tag{5}$$

$$\phi_2 = 4 \tan^{-1}\left[\frac{(\sin^2 \theta_{i2} - 1/n^2)^{1/2}}{\tan \theta_{i2} \sin \theta_{i2}}\right], \tag{6}$$

where θ_{i1} and θ_{i2} are the angles of incidence on the hypotenuse faces of prisms 1 and 2, respectively. Assume that the two beams have the same initial angle of incidence θ_0. From the optical layout of the system, we can see that if one light beam undergoes an angular displacement of $\Delta\theta$, the other will undergo an angular displacement of $-\Delta\theta$. Therefore we have

$$\theta_{11} = \theta_0 + \Delta\theta, \tag{7}$$

$$\theta_{12} = \theta_0 - \Delta\theta. \tag{8}$$

Here θ_{11} and θ_{12}, the angles of incidence at the entrance faces of the prisms, can also be related to θ_{i1} and θ_{i2} according to Eq. (2). The output signal S of the sensor is defined as

$$S = \frac{I_1 - I_2}{I_1 + I_2} = \frac{\cos \phi_1 - \cos \phi_2}{2 + \cos \phi_1 + \cos \phi_2}. \tag{9}$$

This method of using $S = S(\theta_0, \Delta\theta)$ as the output instead of the beam intensity itself is called the differential method, so called because the two intensities always change in opposite directions. The purpose of our using this differential method is twofold: (1) to make the output insensitive to the fluctuations of the laser output power and (2) to improve output linearity. Computer simulations showed that, if the initial angles of the prisms were properly set, the output S could be highly linear with respect to the angle of incidence in a range of several degrees. In Fig. 4 we show some simulation results with the initial angle of incidence θ_0 changing from $-4.5°$ to $0°$. Here negative θ_0 means that the incident light is below the surface normal of the entrance face of the prism in Fig. 1. The measurement range depends on the initial angle of incidence; the closer its corresponding angle on hypotenuse is to the critical angle, the smaller the measurement range will be. High linearity in the central regions of all the curves is evident. The calculated nonlinearity errors of the

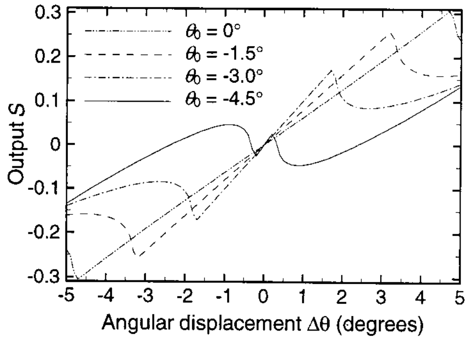

Fig. 4. Output versus angular displacement at differential initial angles of incidence.

curves are 0.006%, 0.052%, 0.299%, and 1.368% for the initial angles of incidence of $-4.5°$, $-3.0°$, $-1.5°$, and $0°$, which are 42°, 43°, 44°, and 45° on the hypotenuse face, respectively. The linearity is improved as the initial angle of incidence moves toward the critical angle.

3. Sensitivity Improvement

As can be seen from Fig. 2, the relative phase shift ϕ changes rapidly with θ_i in the vicinity of the critical angle, which means that measurement sensitivity is high when θ_i is close to the critical angle. Meanwhile, Eq. (9) shows that the output is a function of $\cos \phi_1$ and $\cos \phi_2$, which intrinsically have the highest rate of changes when ϕ_1 and ϕ_2 are at the proximity of 90° or 270° rather than 0° or 180°. Therefore, to improve the measurement sensitivity, we propose to introduce an initial fixed phase shift between the s- and the p-polarization components. When combined with the phase shift ϕ caused by total internal reflection, this initial phase shift will result in a total phase shift close to 90° or 270°.

In addition to measurement sensitivity, nonlinearity error is also an important performance measure. High sensitivity is useless if nonlinearity error is large to the degree that it is not acceptable. To find an optimal initial phase shift, which would result in high sensitivity and low nonlinearity error, we change the initial phase shift from 0° to 130° and calculate the sensitivity and nonlinearity error at each angle. Figure 5 shows the results when the incident angle is $\theta_0 = -4.2°$ at the entrance face and $\theta_i = 42.2°$ at the hypotenuse face. We can see from Fig. 5 that when the initial phase shift is 111°, the nonlinearity error is the regional minimum. Therefore we regard 111° as the optimal initial phase shift. The measurement sensitivity is increased to 1.187 (1/deg) from 0.119 (1/deg) before the introduction of the initial phase shift, which is an improvement of ten times. However, the nonlinearity error is increased to 1.6% from less than 0.1%. Although the initial phase shift enlarges the maximum nonlinearity error compared with that without an initial phase shift, fortunately the nonlinearity error can be reduced greatly if the measurement range is narrowed properly. This is because the maximum nonlinear-

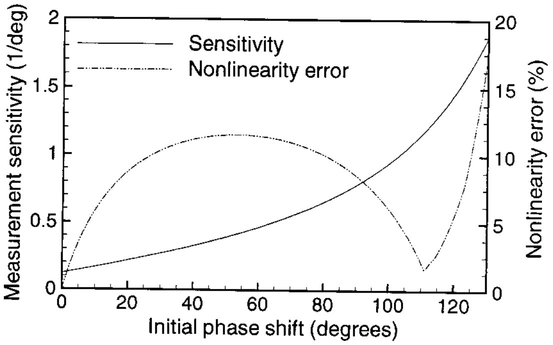

Fig. 5. Measurement sensitivity and nonlinearity error versus initial phase shift when $\theta_0 = -4.2°$.

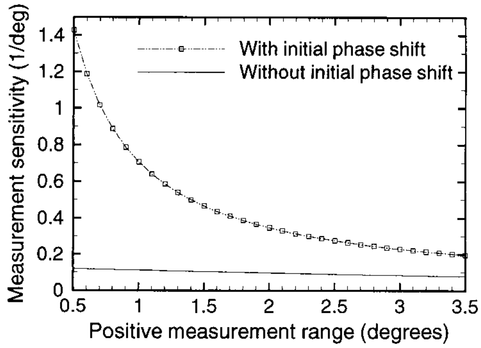

Fig. 7. Comparison of measurement sensitivity with and without the initial phase shift.

ity error always occurs at both ends of the measurement range.

Figure 6 shows the calculated results of the optimized initial phase shift and the associated maximum nonlinearity error as a function of the full measurement range. The increment of the initial phase shift during calculation is 0.1°. The nonlinearity error, which maintains at a level between 1.55 and 1.65%, does not change much with the measurement range. Figure 7 compares the measurement sensitivity before and after the introduction of the optimal initial phase shift. It is clear that the measurement sensitivity increases exponentially with the decrease of the measurement range. Or in other words, this method of initial phase shift for sensitivity improvement is much more effective when the measurement range is small.

Table 1 shows the measurement sensitivity and nonlinearity error for reduced measurement ranges along with the initial angle, full measurement range, and optimized initial phase shift. It can be seen that when the measurement range is reduced to approximately three-quarters of the full measurement range, the measurement sensitivity is more or less the same, but the nonlinearity is reduced dramatically from approximately 1.6% to less than 0.1%.

In summary, by introducing an optimal initial phase shift between the s- and the p-polarized components of the laser beam and appropriately reducing the measurement range, we can significantly improve measurement sensitivity while keeping the nonlinearity error at a low level.

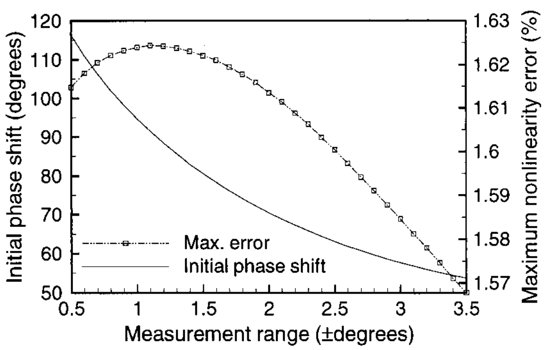

Fig. 6. Optimized initial phase shift and maximum nonlinearity error versus measurement range.

4. Experiments

Figure 8 shows the experimental setup. A fiber-coupled laser diode is used as the light source. The wavelength is 670 nm, and the output power is approximately 1 mW. The fiber is a single-mode polarization-maintaining fiber. The output beam is collimated with a diameter of 2 mm and is linearly polarized with an extinction ratio of approximately 10:1. To increase the extinction ratio, a polarized beam splitter (PBS) is used with its axis oriented at 45° from the X axis. To introduce an initial phase shift for the improvement of measurement sensitivity, a polarization phase shifter consisting of a half-wave ($\lambda/2$) plate and a quarter-wave ($\lambda/4$) plate is placed after the PBS.[17] The azimuth of the $\lambda/4$ plate is 45° with respect to the X axis. It should be noted here that the p-polarization direction is aligned with the X axis. Rotating the $\lambda/2$ plate and making its azimuth an angle of φ with respect to the X direction, we can introduce a phase shift of $\pi - 4\varphi$ between the s- and p-polarized components. The angle sensing optics is placed on top of a rotary stage. A nonpolarizing dielectric CBS is placed after the adjustable phase shifter. Equal intensities and equal phase shifts introduced by the beam splitter are expected in the two beams. The photodiodes are connected to a signal-processing circuit for amplification of the signals. The signals are then sent to a personal computer for calculation and display of sensor output signal S. As the calibration reference, a laser interferometer (HP5529A) is used that measures the rotation angle of the rotary stage at a resolution of 0.1 arc sec and an accuracy of ±0.2%. The initial angle of incidence θ_0 was set at approximately −4.2°. Its corresponding θ_i is 0.4° over the critical angle. This resulted in a measurement range of ±0.6°. For higher sensitivity, an initial phase shift of approximately 111° between the s- and p-polarized components was introduced. Figure 9 shows the calibration result in the full range of ±0.6°. The dots represent the experimental data and the solid line is the linear fit of the dots. The measurement sensitivity and nonlinearity error are calculated to be

Table 1. Optimized Initial Phase Shift, Measurement Range, Sensitivity, and Nonlinearity Error for Different Values of the Initial Angle

Initial Angle θ_0 (deg)	Full Measurement Range (deg)	Optimized Initial Phase Shift (deg)	Reduced Measurement Range (deg)	Measurement Sensitivity (1/deg)	Nonlinearity Error (%)
−4.3	±0.5	104.3	±0.375	1.085	0.093
−3.8	±1.0	79.0	±0.750	0.525	0.096
−3.3	±1.5	62.7	±1.125	0.340	0.095
−2.8	±2.0	51.4	±1.520	0.250	0.099
−2.3	±2.5	43.3	±1.900	0.198	0.094
−1.8	±3.0	38.2	±2.310	0.166	0.094
−1.3	±3.5	35.4	±2.731	0.145	0.094

1.108 (1/deg) and 1.3%, respectively. These data agree well with those predicted by calculation. If we reduce the measurement range in half to ±0.3°, the nonlinearity error would become 0.4%, still much larger than the predicted value. The increased nonlinearity error could be due to the error of the reference laser interferometer (±0.2%) or the error of adjustment which is discussed further in Section 5. The noise level of the device output was found to be at approximately 0.5 arc sec (peak to peak).

5. Possible Error Sources

A. Initial Angle Alignment

The calculation results are based on the assumption that the initial angles at the two prisms are equal. If the two initial angles are not equal because of alignment errors, both sensitivity and nonlinearity error will change. Figure 10 shows the output curve in the full measurement range when an initial phase shift of 111° is applied. We assumed the initial angle for prism 1 to be $\theta_0 = -4.2°$ and that for prism 2 to be $\theta_0 + \Delta\theta_0$ with $\Delta\theta_0$ being 0°, 0.05°, and 0.1°, respectively. The three curves in Fig. 10 show that, with the two initial angles being unequal, the nonlinearity error is increased, even though the measurement sensitivity does not suffer much. Therefore, when building angle sensors based on this method, careful adjustment of the prisms is necessary to make sure that the two initial angles are close and the nonlinearity error is minimized.

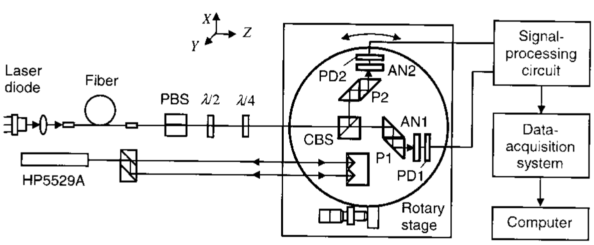

Fig. 8. Experimental setup for the calibration of the prototype device.

B. Transmittance and Reflectance of the Beam Splitter

The CBS plays an important role in the measurement. To minimize the effects of polarization, a laser-line nonpolarizing dielectric CBS designed for the wavelength of 670 nm was used in the experimental setup. This beam splitter has a specified transmittance of 50 ± 5% for any polarization with the *s* and *p* components matched to within 3%. The absorption is less than 0.3%. The small intensity difference between the reflected and the transmitted beams that is due to the difference in the reflectance and transmittance of the beam splitter needs to be compensated for, otherwise an erroneous output signal and therefore measurement error would result. In this research, a proportional compensation method is used in data processing. Assume that the initial angles are identical for the two prisms. When the angular displacement $\Delta\theta = 0$, the phase shifts introduced by total internal reflection at the prisms are identical. Theoretically the intensities $I_1(\Delta\theta = 0)$ and $I_2(\Delta\theta = 0)$ should be equal. But in reality, they

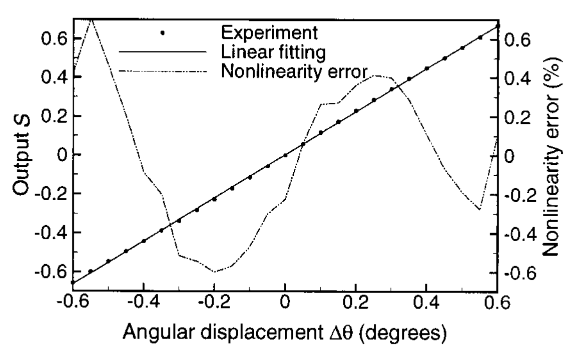

Fig. 9. Experimental result of calibration.

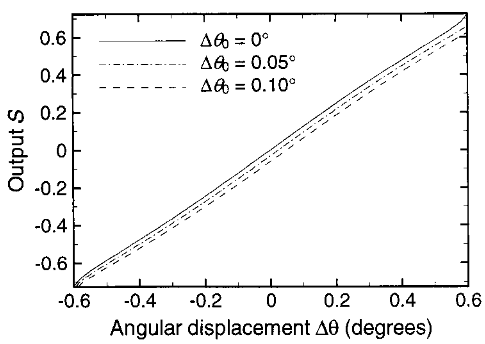

Fig. 10. Change of the output curves that is due to the difference in the initial angles.

are not because of the difference in the reflectance and transmittance of the beam splitter. To compensate for this difference, we simply multiply the measured intensity of the reflected beam by the ratio $I_1(\Delta\theta = 0)/I_2(\Delta\theta = 0)$.

C. Phase Shift by the Beam Splitter

Different phase shifts in the reflected and transmitted beams caused by the beam splitter may affect the measurement result significantly. Fortunately, as long as the phase shift is fixed and identical for both the reflected and the transmitted beams, the phase-shift error can be compensated for. This is because an initial phase shift is introduced by a pair of $\lambda/2$ and $\lambda/4$ plates for improved sensitivity in the experiment. Adding more or less initial phase shift, depending on whether the phase shift caused by the beam splitter is positive or negative, can compensate for the phase shift caused by the beam splitter. If the phase shift is fixed but different in the reflected and the transmitted beams, the experimental setup can be changed to have two phase shifters, one placed in the transmitted beam path and the other in the reflected beam path after the reflection prisms, instead of having just one placed before the beam splitter. Each phase shifter can be a $\lambda/4$ plate plus an analyzer, which already exists in the current setup. The $\lambda/4$ plate is again set at an azimuth of 45° with respect to the X axis. Phase shift between the s- and the p-polarized components is obtained by rotation of the analyzer. If the azimuth of the analyzer is φ with respect to the X axis, the introduced phase shift will be $\pi/2 - 2\varphi$.[17] This phase shift is half that introduced by the phase shifter consisting of a pair of $\lambda/2$ and $\lambda/4$ plates. The detected intensity in this case is $(I_0/2)[1 + \cos(\phi + \pi/2 - 2\varphi)]$, which differs from that for the current setup only in the introduced phase shift. Therefore the proposed method still works if only the azimuth of the analyzer φ is properly set to provide the optimized initial phase shift.

6. Conclusions

In this paper a new method for precision small-angle measurement is proposed based on the total-internal-reflection effect of a laser beam at a glass prism. It uses a simple laser diode as the light source instead of more complicated, bulky, and expensive double-frequency laser sources used by others in previous research. With the introduction of an optimized initial phase shift by a pair of half- and quarter-wave plates, it has been shown that the measurement sensitivity could be increased dramatically without a significant increase of nonlinearity error. A prototype device has been built, and experiments have taken place. Calibration of the device showed a nonlinearity error of 1.3% and a peak-to-peak noise level of approximately 0.5 arc sec in a measurement range of $\pm 0.6°$. The nonlinearity error for the central $\pm 0.3°$ of the measurement range was smaller at 0.4%. The noise level of the device can be reduced further and resolution increased when the device is designed to have a smaller measurement range. Compared with the previous methods based on heterodyne interferometry, this new method has distinctive advantages in terms of compact device size, high resolution, and high linearity.

References

1. P. S. Huang, S. Kiyono, and O. Kamada, "Angle measurement based on the internal reflection effect: a new method," Appl. Opt. **31,** 6047–6055 (1992).
2. P. S. Huang and J. Ni, "Angle measurement based on the internal-reflection effect and the use of right-angle prisms," Appl. Opt. **34,** 4976–4981 (1995).
3. P. S. Huang and J. Ni, "Angle measurement based on the internal-reflection effect using elongated critical-angle prisms," Appl. Opt. **35,** 2239–2241 (1996).
4. P. S. Huang and Y. Li, "Small-angle measurement by use of a single prism," Appl. Opt. **37,** 6636–6642 (1998).
5. P. S. Huang, "Use of thin films for high-sensitivity angle measurement," Appl. Opt. **38,** 4831–4836 (1999).
6. P. S. Huang and J. Ni, "Angle measurement based on the internal-reflection effect and its application in measurement of geometric errors of machine tools," in *Proceedings of the American Society for Precision Engineering Annual Meeting* (American Society for Precision Engineering, Raleigh, N.C., 1993), pp. 350–353.
7. P. S. Huang and Y. Li, "Laser measurement instrument for fast calibration of machine tools," in *Proceedings: American Society for Precision Engineering Annual Meeting* (American Society for Precision Engineering, Raleigh, N.C., 1996), pp. 112–115.
8. P. S. Huang and X. Xu, "Design of an optical probe for surface profile measurement," Opt. Eng. **38,** 1223–1228 (1999).
9. S. Zhang, S. Kiyono, Y. Uda, and M. Mito, "Development of a measurement system of the angular profile of the polygon mirror surface," Int. J. Jpn. Soc. Precis. Eng. **30,** 349–350 (1996).
10. S. Kiyono, X. Shan, and H. Sato, "Development of an AFM using a critical angular sensor," Int. J. Jpn. Soc. Precis. Eng. **27,** 373–378 (1993).
11. S. Zhang, S. Kiyono, and Y. Uda, "Nanoradian angle sensor and *in situ* self-calibration," Appl. Opt. **37,** 4154–4159 (1998).
12. M.-H. Chiu and D.-C. Su, "Angle measurement using total-internal-reflection heterodyne interferometry," Opt. Eng. **36,** 1750–1753 (1997).
13. M.-H. Chiu and D.-C. Su, "Improved technique for measuring small angles," Appl. Opt. **36,** 7104–7106 (1997).
14. W. Zhou and L. Cai, "Interferometer for small-angle measurement based on total internal reflection," Appl. Opt. **37,** 5957–5963 (1998).
15. W. Zhou and L. Cai, "An angular displacement interferometer based on total internal reflection," Meas. Sci. Technol. **9,** 1647–1652 (1998).
16. W. Zhou and L. Cai, "Improved angle interferometer based on total internal reflection," Appl. Opt. **38,** 1179–1185 (1999).
17. M. P. Kothiyal and C. Delisle, "Shearing interferometer for phase shifting interferometry with polarization phase shifter," Appl. Opt. **24,** 4439–4442 (1985).