Available online at www.sciencedirect.com
SCIENCE DIRECT®
Optics Communications 250 (2005) 16–23
OPTICS COMMUNICATIONS
www.elsevier.com/locate/optcom

Optical encoder based on the fractional Talbot effect

Ching-Fen Kao *,[1], Mao-Hong Lu

Department of Photonics and Institute of Electro-Optical Engineering, National Chiao Tung University, Building 16, 321 Section 2, Kuang Fu Road, Hsinchu, Taiwan 30050, ROC

Received 23 September 2004; received in revised form 11 January 2005; accepted 3 February 2005

Abstract

This paper presents a reflective optical encoder based on the Talbot self-image of binary phase grating. The proposed encoder has several advantages in comparison to conventional Moiré effect based encoders, including larger gap between gratings, higher gap tolerance and greater light passage efficiency. This encoder with gap 220 μm can accurately measure the position of linear stages which move at high-speed. The encoder includes a grating scale with period 20 μm and provides a signal period of 20 μm with a frequency response of approximately 200 kHz. This system and the associated electronics provide an interpolation factor of 200 that gives a resolution of 0.1 μm. The repeatability is 0.2 μm for a traveling length of 600 mm.

Keywords: Optical encoder; Talbot effect; Self-image

1. Introduction

Most optical encoders (linear transducer as well as rotational encoder) are based on the Moiré effect. A collimated beam penetrates two gratings, which have equal period. Then, an oscillating signal is generated when a grating moves in the direction which is perpendicular to its ruling. In actual optical encoders, the most commonly used light source is an infrared LED with a typical size of 350 μm. The optics used to collimate the light beam typically comprises a high-numerical-aperture singlet with a focal length of about 10 mm, due to the limitations on the total size of the reading head of the encoder. With this ratio of source size to focal length, the divergent angle of light is around 2°. The contrast decay of the signal can be described

* Corresponding author. Tel.: +886 3 5743887; fax: +886 3 5722383.
E-mail address: alicekao@itri.org.tw (C.-F. Kao).
[1] Also affiliated with Center for Measurement Standards, Hsinchu, Taiwan 300, ROC.

0030-4018/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.optcom.2005.02.006

by $m = (2g/T)\tan\gamma$, where g denotes the gap between two gratings, γ is the divergent angle of light, and T is the grating period [1]. For example, for an encoder mentioned above and with a grating period of 20 μm, the gap between gratings should be roughly 50 μm in order to get a good signal. Such a small gap between gratings could cause a big problem in mechanical tolerance in the manufacture of these encoders. In the Moiré type encoders, a bearing and spring are used to maintain the gap. However, the bearing cannot work for high-speed motion. Meanwhile, in high-speed motion, a reflective type of encoder can be used without bearing. Crespo et al. [2] designed an optical encoder using the Lau effect in a reflective-type encoder to increase the gap. Nevertheless, they used an amplitude-type grating, through which only half of light flux can pass through.

This paper presents an optical encoder using the Talbot effect [3–5]. A phase grating which produces a self-image at a quarter Talbot distance is used in the encoder [6–8], so that all light can penetrate it. This reflective-type encoder with phase grating has the advantages of a large gap and more light efficiency. The remainder of this paper is organized as follows: In Section 2.1, we derive expressions for the self-image of phase grating. In Section 2.2, we discuss a reflective-type optical encoder, using the three-grating system model. In Section 3, we present some measurements made using the reflective-type encoders and evaluate the manufacturing tolerance.

2. Theory

2.1. Fractional Talbot effects of gratings

Self-imaging properties were first observed by Talbot in 1836. During the 1960s, Winthrop and Worthington investigated many theoretical and practical aspects, especially for fractional Talbot planes. In this study, we applied the theory of Winthrop and Worthington in a one-dimensional case. The transmittance of a one-dimensional periodic structure shown in Fig. 1 with unit cell transmittance $t_p(x)$ can be described by

$$t(x) = t_p(x) \otimes \mathrm{comb}\left(\frac{x}{p}\right), \tag{1}$$

where p denotes the spatial period in the x-direction and $\otimes$ represents convolution. From Fresnel diffraction [9], the optical field in an arbitrary plane at a distance z_F behind the periodic structure is given by

$$U(x, z_F) = C(z_F) \times t_p(x) \otimes \left\{ \sum_{m=-\infty}^{\infty} \exp\left[-i\pi\lambda z_F \left(\frac{m}{p}\right)^2\right] \exp\left[i2\pi\left(\frac{m}{p}\right)x\right] \right\}, \tag{2}$$

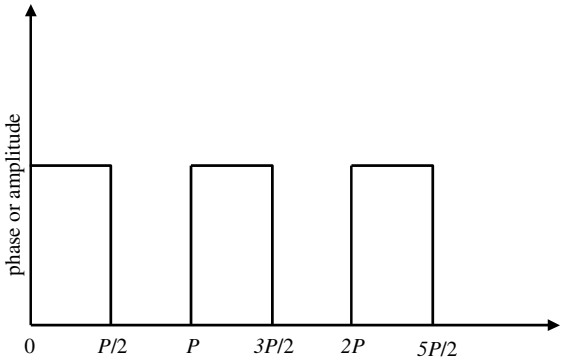

Fig. 1. One-dimension binary grating with period P.

where $C(z_F) = (\exp(ikz_F)/i\lambda z_F) \cdot \sqrt{2z_F/k}\exp(i\pi/4)$ and λ is the light wavelength. In this study, the proposed encoder is based on the fractional Talbot effect and only the diffraction amplitude on the plane at $z = z_T/4$ is considered, where z_T denotes the Talbot distance and equals $2p^2\lambda$. Accordingly, Eq. (2) can be rewritten as

$$U\left(x,\frac{z_T}{4}\right) = A \times \left\{ t_p(x) \otimes \sum_{m=-\infty}^{\infty} \exp\left(-i\pi\frac{m^2}{2}\right)\exp\left[i2\pi\left(\frac{m}{p}\right)x\right]\right\}, \tag{3}$$

where $A = \sqrt{2}/(ip\sqrt{\pi})\exp[i\pi(p/\lambda)^2 + i\pi/4]$. By setting $m = 2n_0 + n_1$, the summation term on the right-hand side of Eq. (3) can be rearranged as

$$\sum_{m=-\infty}^{\infty} \exp\left(-i\pi\frac{m^2}{2}\right)\exp\left[i2\pi\left(\frac{m}{p}\right)x\right] = \sum_{n_0=-\infty}^{\infty}\sum_{n_1=0,1} \exp(-i2\pi n_0^2)\exp(-i2\pi n_0 n_1) \times \exp\left[-i\pi\left(\frac{n_1^2}{2}\right)\right]\exp\left[i4\pi\left(\frac{n_0}{p}\right)x\right]\exp\left[i2\pi\left(\frac{n_1}{p}\right)x\right]. \tag{4}$$

The terms $\exp(-i2\pi n_0^2)$ and $\exp(-i2\pi n_0 n_1)$ in Eq. (4) equals one because n_0 and n_1 are integers. Thus, Eq. (4) can be further simplified as

$$\sum_{m=-\infty}^{\infty} \exp\left(-i\frac{\pi}{2}m^2\right)\exp\left[i2\pi\left(\frac{m}{p}\right)x\right] = \left\{1 + \exp\left[i\frac{2\pi}{p}\left(x - \frac{p}{4}\right)\right]\right\} \times \sum_{n_0=-\infty}^{\infty}\delta\left(x - n_0\frac{p}{2}\right) = \sum_{n_0=-\infty}^{\infty}\left\{1 + \exp\left[i\pi\frac{(2n_0-1)}{2}\right]\right\}\delta\left(x - n_0\frac{p}{2}\right), \tag{5}$$

where δ represents the Dirac delta function. Substituting Eq. (5) into Eq. (3), the amplitude distribution $U(x, z_T/4)$ becomes

$$U\left(x,\frac{z_T}{4}\right) = A\sum_{n_0=-\infty}^{\infty}\left\{\left[1 + \exp\left(i\pi\frac{(2n_0-1)}{2}\right)\right] \times t_p\left(x - \frac{n_0 p}{2}\right)\right\}. \tag{6}$$

The parity of n_0 determines the value of the phase term in Eq. (6). If n_0 is odd, then $\{1 + \exp[i\pi(2n_0 - 1)/2]\} = (1 + i)$ and $n_0 = 2n + 1$. On the contrary, if n_0 is even, then $\{1 + \exp[i\pi(2n_0 - 1)/2]\} = (1 - i)$ and $n_0 = 2n$. These results lead Eq. (6) to

$$U\left(x,\frac{z_T}{4}\right) = A\sum_{n=-\infty}^{\infty}\left[(1-i)t_p(x - np) + (1+i)t_p\left(x - np - \frac{p}{2}\right)\right] = \sqrt{2}A\exp\left(-i\frac{\pi}{4}\right)\left[t(x) + \exp\left(i\frac{\pi}{2}\right)t\left(x - \frac{p}{2}\right)\right]. \tag{7}$$

Eq. (7) indicates that the field distribution in the quarter Talbot plane [7–9] consists of two parts. The both parts are the same as the transmittance of the periodic object. But there exist a displacement of a half period and a phase shift of $\pi/2$ between the two parts.

In this study, the proposed encoder contains two different periodic structures: binary-phase grating (BPG) and binary-amplitude grating (BAG). Supposing that these gratings have the same periodicity, the phase step of the BPG is ϕ, and the open ratio of the BAG is 0.5. The transmittances of the BPG and BAG can be expressed as follows:

$$t_{BPG}(x) = \left\{\exp(i\phi)\mathrm{rect}\left(\frac{x - \frac{p}{4}}{\frac{p}{2}}\right) + \mathrm{rect}\left(\frac{x - \frac{3}{4}p}{\frac{p}{2}}\right)\right\} \otimes \mathrm{comb}\left(\frac{x}{p}\right), \tag{8}$$

$$t_{\mathrm{BAG}}(x) = \mathrm{rect}\left(\frac{x-\frac{p}{4}}{\frac{p}{2}}\right) \otimes \mathrm{comb}\left(\frac{x}{p}\right). \tag{9}$$

If the input object is a binary-phase grating, i.e., $t(x) = t_{\mathrm{BPG}}(x)$ and the phase step $\phi = \pi/2$, then the field distribution on the $z_{\mathrm{T}}/4$ plane, obtained by substituting Eq. (8) into Eq. (7), is

$$U_{\mathrm{BPG}}\left(x, \frac{z_{\mathrm{T}}}{4}\right) = 2\sqrt{2}A \exp\left(\mathrm{i}\frac{\pi}{4}\right) \mathrm{rect}\left(\frac{x-\frac{p}{4}}{\frac{p}{2}}\right) \otimes \mathrm{comb}\left(\frac{x}{p}\right). \tag{10}$$

According to Eq. (9), the above expression shows a binary-amplitude grating located at a quarter Talbot distance. This lens-like imaging effect can be reversed if the input object is a binary-amplitude grating, i.e., $t(x) = t_{\mathrm{BAG}}(x)$. By substituting Eq. (9) into Eq. (7), the field distribution on the quarter Talbot plane is

$$U_{\mathrm{BAG}}\left(x, \frac{z_{\mathrm{T}}}{4}\right) = \sqrt{2}A \exp\left(-\mathrm{i}\frac{\pi}{4}\right) \left[\mathrm{rect}\left(\frac{x_0-\frac{p}{4}}{\frac{p}{2}}\right) + \exp\left(-\mathrm{i}\frac{\pi}{2}\right) \mathrm{rect}\left(\frac{x_0-\frac{3p}{4}}{\frac{p}{2}}\right)\right] \otimes \mathrm{comb}\left(\frac{x}{p}\right). \tag{11}$$

2.2. Three-grating model [10]

Fig. 2 depicts the structure of the optical linear encoder presented in this study. The binary-phase and binary-amplitude gratings are the index and main scales, respectively. Both gratings have the same period and their separation is a quarter of the Talbot distance. The collimated light passes through the index scale and produces a BAG image on the main scale, as described in Section 2.1. The light then is reflected by the main scale and goes through the index scale again. Finally, a photo-detector collects the reflected light. The output current of the detector changes with the displacement between the index grating and the main scale.

In the encoder, the light received by the detector passes through the index grating twice and is reflected by the main grating once. Fig. 3 illustrates that the optical configuration is equivalent to a three-grating system, where G_1 and G_3 are the index grating and G_2 is the main grating. A monochromatic plane wave is diffracted by G_1 and propagates to G_2. Because the distance between G_1 and G_2 is $z_{\mathrm{T}}/4$, the diffraction filed in front of G_2 is given by Eq. (10),

$$U_{2\mathrm{f}}(x) = 2\sqrt{2}A \exp\left(-\mathrm{i}\frac{\pi}{4}\right) \mathrm{rect}\left(\frac{x-\frac{p}{4}}{\frac{p}{2}}\right) \otimes \mathrm{comb}\left(\frac{x}{p}\right). \tag{12}$$

The above equation shows that the field diffracted by G_1 is the same as the transmittance of grating G_2. The self-image of G_1 overlaps on grating G_2, therefore, the Moiré fringe is generated when one grating moves in

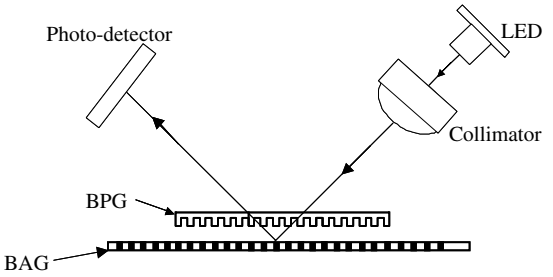

Fig. 2. Schematic diagram for the reflective optical encoder.

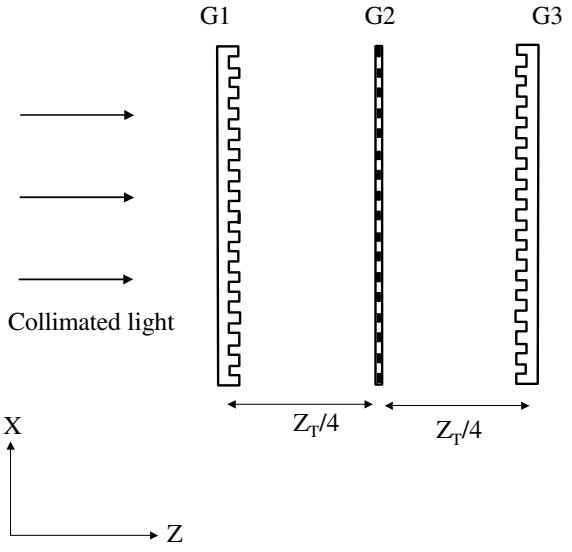

Fig. 3. The three-grating system.

the x-direction, as shown in Fig. 3. If a displacement Δx occurs between G_1 and G_2, the diffraction field just behind G_2 is

$$U_{2b}(x) = U_{2f}(x) \times t_{BAG}(x) = \text{rect}\left(\frac{x - \frac{p}{4} - \Delta x}{\frac{p}{2} - \Delta x}\right) \otimes \text{comb}\left(\frac{x}{p}\right). \tag{13}$$

The Moiré fringe given by Eq. (13) seems like the field that a plane wave just passes through a binary-amplitude grating laterally moved by Δx. Additionally, the open ratio of the grating is changed to $(0.5 - \Delta x/p)$. The wave $U_{2b}(x)$ keeps propagating and arrives at G_3. The description of the wave $U_{3f}(x)$ before G_3 is similar to Eq. (11) because the travel distance equals $z_T/4$. Hence, $U_{3f}(x)$ is given by

$$U_{3f}(x) = A_2 \exp\left(-i\frac{\pi}{4}\right)\left[\text{rect}\left(\frac{x - \frac{p}{4} - \frac{\Delta x}{2}}{\frac{p}{2} - \Delta x}\right)\exp\left(-i\frac{\pi}{2}\right) + \text{rect}\left(\frac{x - \frac{3p}{4} - \frac{\Delta x}{2}}{\frac{p}{2} - \Delta x}\right)\right] \otimes \text{comb}\left(\frac{x}{p}\right). \tag{14}$$

After passing G_3, the field distribution becomes

$$U_{3b}(x) = U_{3f}(x) \times t_{BPG}(x) = A_2 \exp\left(-i\frac{\pi}{4}\right)\left[\text{rect}\left(\frac{x - \frac{p}{4} - \frac{\Delta x}{2}}{\frac{p}{2} - \Delta x}\right) + \text{rect}\left(\frac{x - \frac{3p}{4} - \frac{\Delta x}{2}}{\frac{p}{2} - \Delta x}\right)\right] \otimes \text{comb}\left(\frac{x}{p}\right). \tag{15}$$

Finally, the light wave reaches the detector. Since the size of the detector is much larger than the grating period, the detector output can be expressed by

$$S(\Delta x) \propto \int_0^p |U_{3b}(x)|^2 \, dx \propto \Lambda\left(\frac{\Delta x}{p}\right), \tag{16}$$

where Λ(argument) is the triangle function. Eq. (16) shows that the detector output is a triangle function of Δx, which is the lateral displacement between G_1 and G_2. However, in practice the shape of the output signal is similar to a sinusoidal function owing to the non-perfect collimated light and finite grating structure.

3. Experimental setup

Fig. 2 schematically depicts a setup for the reflective-optical encoder. In this experiment, an infrared light emitted diode (LED) is used as light source. Though laser diode (LD) has better high properties than LED, we consider the compactness of the head and the LED meets all the requirements for signal to noise ratio (SNR) and contrast in this case.

The LED light source with a wavelength of 0.88 μm is collimated by a collimating lens with focal length of 9 mm. The collimated beam passes through the phase grating and is reflected by grating G_2. The phase-grating G_1 and reflective-grating G_2 have the same period 20 μm. The phase grating is divided into four segments. The aperture of segment is 1 mm; consequently, the pupil size covers the period number of 50. The number of periods is sufficient to product the self-imaged effect. The spacing between adjacent segments is $(20N + 5)$ μm, where N is the integer, as shown in Fig. 4. The four segments in the phase grating G_1 correspond to four sinusoidal outputs with phase shifts by 90° from four photo-detectors. The phase grating G_1 is illuminated with the collimated light and self-imaged on grating G_2. A quarter of Talbot distance is 0.22 mm. The luminous dimension of LED is about 160 μm. With this ratio of luminous dimension to focal length, the divergent angle of light source is about 1°. Fig. 5 shows the contrast variation of the Moiré signal due to the gap change. The normalized contrast m can be defined by

$$m = \frac{\mathrm{AC}(g)}{\mathrm{AC}(0.22)}, \tag{17}$$

where $\mathrm{AC}(g)$ denotes the amplitude of the encoder signal when the gap between two gratings is g and $\mathrm{AC}(0.22)$ represents the amplitude of the encoder signal when the gap is 0.22 mm. For getting a good signal (the minimum contrast must be greater than 0.6), the tolerance of gap is approximately ±70 μm.

From Eq. (7), the binary phase grating is converted into a binary-amplitude grating image at the quarter of the Talbot distance when the phase step is ϕ. The intensity distribution of the image is [11]

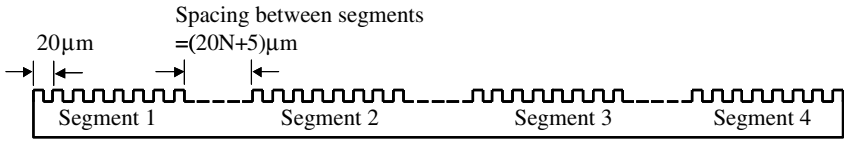

Fig. 4. Scheme of four-segmented phase gratings. The zone spacing between adjacent segments is $(20N + 5)$ μm, where N is integer. There is a 90° phase shift between adjacent segments.

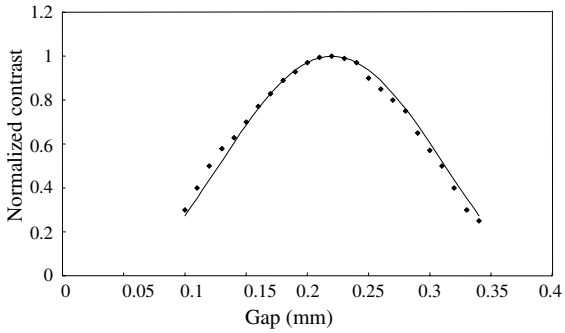

Fig. 5. The contrast variation of signal due to the gap change.

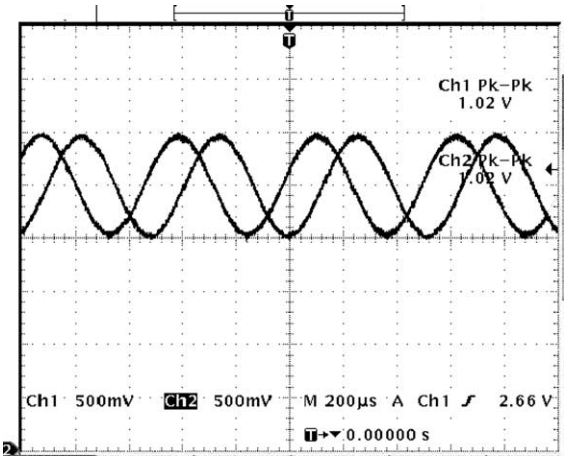

Fig. 6. Two orthogonal signals (sine and cosine) from an oscilloscope.

$$I = \left|U\left(x, \frac{z_T}{4}\right)\right|^2 = \left|\sqrt{2}A \exp\left(-i\frac{\pi}{4}\right)\left[t\left(x - \frac{p}{2}\right) \exp\left(i\frac{\pi}{2}\right) + t(x)\right]\right|^2$$
$$\propto \text{rect}\left(\frac{x - \frac{p}{4}}{\frac{p}{2}}\right) \cdot (1 + \sin\phi) + \text{rect}\left(\frac{x - \frac{3p}{4}}{\frac{p}{2}}\right) \cdot (1 - \sin\phi). \quad (18)$$

The step height of phase grating is $h = \lambda\phi/2\pi(n-1)$. If manufacturing error Δh occurs for the step height, the phase shifts with $\Delta\phi$. The phase shift can be written as

$$\Delta\phi = \frac{2\pi(n-1)\Delta h}{\lambda}. \quad (19)$$

The visibility of the grating pattern can be calculated by

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \sin(\phi + \Delta\phi) \quad (20)$$

If the phase step is $\phi = \pi/2$, the step height of phase grating is $h = \lambda\phi/2\pi(n-1) = 430$ nm, and the visibility can be expressed by $\cos(\Delta\phi)$. The step depth of phase grating measured by the atomic force microscopy approximately equals $h = 480$ nm and the error of manufacture is 50 nm, which corresponds to a phase shift of 0.0575π. According to Eq. (20), the visibility of self-image, V, equals 98%. The contrast of the Moiré signal between phase grating G_1 and reflective grating G_2 decreases from 100% to 98%. For a good visibility (>95%), the step depth tolerance of the phase grating is ±87 nm.

A linear positioning system with a scanning range of 60 cm was used to measure the accuracy of the optical encoder. The linear positioning system is a linear motor with position feedback control by laser interferometer. In this study, a laser interferometer is used as a standard for measuring the encoder error. Fig. 6 shows two orthogonal signals (sine and cosine) with a signal period of 20 μm. The resolution could be improved to 0.1 μm using interpolation circuit with a factor of 200. Fig. 7 shows the three-path (back and forth) repeatability measurement. The vertical axis represents the measuring difference between the optical encoder and the laser interferometer. Meanwhile, the horizontal axis is the travel distance of the linear stage. In this work, a repeatability error of 0.5 μm is obtained that includes the backlash of the linear stage. We can divide the total repeatability into two parts – forward and backward. Both of them have repeatability error of less than 0.2 μm.

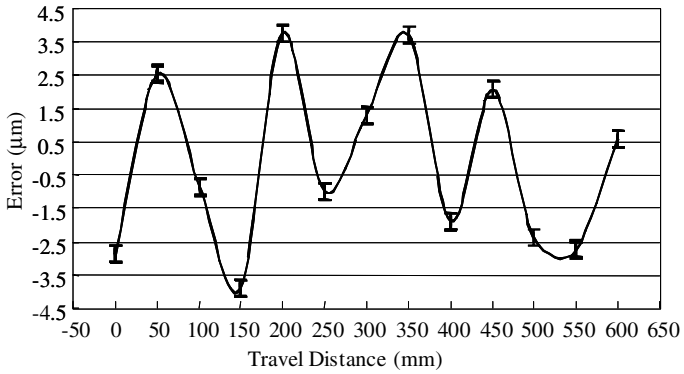

Fig. 7. Three-path (back and forth) repeatability measurement. The vertical axis was the measuring difference between optical encoder and laser interferometer. The horizontal axis was the travel distance of linear stage. We have obtained a repeatability of 0.5 μm.

4. Conclusion

In this work, we develop a reflective optical encoder using Talbot image of the binary phase grating. In comparison with conventional Moiré encoder, the Talbot effect provides a larger gap between two gratings. The phase grating has more light to be passed and gives higher signal intensity. This investigation provides the theoretical expressions for the gap tolerance and manufacture tolerance of phase grating. Experimental results show a good agreement with theoretical results. For high moving speed, the optical encoders have sinusoidal output signals particularly with levels of 1 V at a −3 dB cutoff frequency of approximately 200 kHz. For sub-micro resolution measuring steps, an electronic interpolation with a factor of 200 leads to a measuring resolution of 0.1 μm. This study obtains excellent repeatability of 0.5 μm and large gap of 0.22 mm between two gratings with a pitch of 20 μm for 600 mm traveling range.

Further efforts will focus on improving the resolution and accuracy by using a fine grating period, such as 10 μm. The gap between two gratings with a period of 10 μm is about 60 μm. In order to have a large gap, the system should work in a higher-order Talbot zone. The laser diode source will be used for higher brightness, better wave-front quality and narrower spectral bandwidth.

Acknowledgment

The authors thank the Ministry of Economic Affairs of the Republic of China, Taiwan, for financially supporting this research.

References

[1] Howard C. Epstein, Optical and mechanical design trade-offs in incremental encoders, Hewlett–Packard Report, 1998.
[2] Daniel Crespo, Toma's Morlanes, Eusebio Bernabeu, Opt. Eng. 39 (2000) 817.
[3] J.T. Winthrop, C.R. Worthington, J. Opt. Soc. Am. 55 (1965) 373.
[4] Victor Arrizon, Juan G. Ibarra, Adolf W. Lohmann, Opt. Commun. 124 (1996) 229.
[5] Krzyaztof Patorski, Prog. Opt. XXVII (1989) I3.
[6] Yih-Shyang Cheng, Ray-Chung Chang, Appl. Opt. 33 (10) (1994) 1863.
[7] Piotr Szwaykowski, Victor Arrizon, Appl. Opt. 32 (1993) 1109.
[8] Victor Arrizon, E. Tepichin, M. Ortiz-Gutierrez, Adolf W. Lohmann, Opt. Commun. 127 (1996) 171.
[9] H. Hamam, J.L. de Bougrenet de la Tocnaye, J. Opt. Soc. Am. A 12 (9) (1995) 1920.
[10] D. Crespo, J. Alonso, E. Bernabeu, Appl. Opt. 39 (2000) 3805.
[11] Zbigniew Jaroszewicz, Andrzej Kolodziejczyk, Optik 111 (2000) 207.