Signal processing method of phase correction for laser heterodyne interferometry

Tao Yang, Liping Yan, Benyong Chen, Yanna Liu, Quihong Tian

Nanometer Measurement Laboratory, Zhejiang Sci-Tech University, Hangzhou 310018, China

ARTICLE INFO

Article history:
Received 20 November 2013
Received in revised form 16 January 2014
Accepted 18 January 2014

Keywords:
Heterodyne interferometry
Phase measurement
Movement direction identification
Phase correction
Displacement measurement

ABSTRACT

A novel signal processing method of movement direction identification and phase correction is presented for laser heterodyne interferometry. Based on the reference signal, four intervals with phase difference of 90° each other are set up. The real-time movement direction identification and the integer fringe counting are realized by detecting the times that the rising-edge of the measurement signal crosses the intervals. The phase correction approach is proposed in detail to solve the phase compensation when the initial phase difference is not equal to the zero phase difference. Three experiments of the stability test, the nanometer and micrometer displacement tests on bi-directional movement were performed to demonstrate the usefulness and feasibility of the presented signal processing method.

1. Introduction

With the rapid development of precision machining and microelectronics technology, urgent requirements are put forward on nanometer displacement measurement [1–6]. Laser heterodyne interferometer with high precision, high speed and long distance has been widely used in displacement, straightness and micro-angle measurements [7–11]. Signal processing for laser heterodyne interferometer is a key technique to realize nanometer measurement accuracy. In the past few decades, many researchers have focused on this study [12–15]. For example, Demarest developed the heterodyne interferometer electronics to provide a position resolution of 0.31 nm at velocities up to 2.1 m s⁻¹ with the designed phase meter consisting of a phase locked loop, a delay line interpolator and a delay-locked loop [8]. Eom et al. described a phase-encoding electronics including the phase demodulating electronics and the nonlinearity compensating electronics [12]. Yim et al. proposed a dual mode phase measurement for heterodyne interferometry whose main idea is combining the high bandwidth frequency counter and the high resolution phase meter with mutually complementary characteristics, and provided a relevant phase correction method of high bandwidth phase counter [14]. Kim et al. presented a phase measurement method that achieves a resolution of 0.15 nm using simple electronics for a target speed of 2.4 m s⁻¹ which adopts a frequency-conversion technique to lower the original beat frequency to 100 kHz by mixing it with a stable reference signal generated from a phase-locked loop [15]. These works have been done to mainly improve phase-measuring electronics to obtain higher resolution or speed. In this paper, a novel movement direction identification and fraction fringe phase correction method is proposed for laser heterodyne interferometry. And its feasibility will be verified by experiments.

2. Phase measurement problem

As it is well known, a laser heterodyne interferometer [14,16] is shown in Fig. 1. An orthogonally linearly polarized beam with different frequencies of \( f_1 \) and \( f_2 \), emitted from a stabilized two-frequency He–Ne laser, is divided into two parts by a beam-splitter (BS). The reflected one passes through a polarizer and projects onto the first detector to generate the reference signal. The transmitted one projects onto a polarizing beam-splitter (PBS) and is split into two beams, i.e. the \( f_1 \) an \( f_2 \) beams. These two beams separately project onto the reference mirror and the moving mirror and return to PBS where they are recombined into one beam, which is reflected by a reflector and passes through another polarizer and projects onto another detector to generate the measurement signal. When the moving mirror is moved by a measured object, the phase meter is used to detect and accumulate the phase difference (\( \Delta \phi \)) between the reference signal and the measurement signal. The measured displacement (\( \Delta L \)) can be
expressed as \( \Delta l = \Delta \phi \lambda / 4\pi \), where \( \lambda \) is the wavelength of light source.

To realize displacement measurement with nanometer resolution, the following phase measurement problem should be addressed:

The zero phase difference means that the phase difference between the reference and measurement signals is 0° (position I in Fig. 2). The initial phase difference is the phase difference between the reference and measurement signals at the beginning of measurement (\( \Delta \phi_{\text{0}}, \) position II in Fig. 2). Generally, \( \Delta \phi_{\text{0}} \) is not equal to 0° at the beginning of measurement, and the phase difference will induce a 360° abrupt transition at the zero phase difference during the measurement process. So the measurement error probably occurs if the integer fringe counting adds or subtracts 1 according to crossing the zero phase difference. In fact, the integer fringe counting should be \( \pm 1 \) only when forward or backward change of the measurement signal makes the phase difference exceed the initial phase difference again (position II in Fig. 2). Therefore, when the initial phase difference is not 0°, it is essential to correct the phase difference in order to achieve right measurement result. In addition, the movement direction of the measured object must be also judged in real time.

3. Signal processing method

In order to solve the above mentioned problem, we propose a signal processing method for laser heterodyne interferometry, including the four intervals movement direction judgment approach and the fraction phase correction approach.

3.1. The basic idea of movement direction identification

Four intervals within one period of the reference signal is set up, the rising-edge of the measurement signal is detected in real time. A fringe counter will count once whenever the rising-edge of the measurement signal crosses one interval. While the measurement signal crosses the interval forward, the counter adds 1. Conversely, the counter subtracts 1. When the accumulated number of the counter is greater than 4 or less than \(-4\), that is, when current phase difference exceeds the initial phase difference forward or backward, an integer fringe counting is completed. The movement direction of the measured object can be judged according to the polarity of the integer fringe counting number.

3.1.1. The establishment of the intervals

As shown in Fig. 3, based on the reference signal \( \text{ref}_1 \), three signals \( \text{ref}_2, \text{ref}_3 \) and \( \text{ref}_4 \) are produced by shifting phase of the reference signal with 90°, 180° and 270°, respectively. Then, one period of the reference signal is divided into four intervals \( (S_a, S_b, S_c, S_d) \) by the four rising-edges (A, B, C and D) of the four signals \( \text{ref}_1, \text{ref}_2, \text{ref}_3 \) and \( \text{ref}_4 \). Apart from this, if the duty cycle of 50% is guaranteed, the falling-edges of \( \text{ref}_1 \) and \( \text{ref}_2 \), together with their rising-edges, can be used to form the four intervals. When the rising-edge of the measurement signal arrives, if the levels of the four signals \( \text{ref}_1, \text{ref}_2, \text{ref}_3 \) and \( \text{ref}_4 \) are “1001”, this means that the rising-edge of the measurement signal is in the interval \( S_a \), we define that the measurement signal state is “00”. Similarly, the levels “1100”, “0110” and “0011” of the four signals correspond that the rising-edge of the measurement signal is in the intervals \( S_b, S_c \) and \( S_d \), respectively, and the measurement signal state is “01”, “10” and “11”, respectively. When the measured object moves forward, that is, the phase difference between the reference and measurement signals increases gradually, the measurement signal state will change with the order “00” \( \rightarrow “01” \rightarrow “10” \rightarrow “11” \rightarrow “00” \rightarrow “00” \rightarrow “00” \rightarrow ……” And when the measured object moves backward, that is, the phase difference decreases gradually, the measurement signal state will change with the order “00” \( \rightarrow “11” \rightarrow “10” \rightarrow “01” \rightarrow “00” \rightarrow “00” \rightarrow “00” \rightarrow ……” We then define a 24-bit binary variable \( \text{num} \) that represents the change times of the measurement signal state. When the measurement signal state changes once forward, the variable \( \text{num} \) adds 1, that is, \( \text{num}=\text{num}+1 \). On the contrary, \( \text{num}=\text{num}-1 \). The change sequence of the measurement signal state is shown in Fig. 4. Therefore, during the measurement process, as long as the measurement signal state is detected in real time, the movement direction of the measured object can be judged.

Fig. 1. A block diagram of laser heterodyne interferometer.

Fig. 2. Schematic of the phase difference change.

Fig. 3. Establishment of four intervals based on the reference signal.
3.1.2. Movement direction identification

Suppose that the change times of the measurement signal state is num_n, that corresponds to the nth measurement, and num_{n+1} to the (n+1)th measurement. The integer part of (num_{n+1} − num_n)/4 is denoted as N, and the fraction part of (num_{n+1} − num_n)/4 is denoted as m. Then m = num_{n+1} − num_n − 4N and −3 ≤ m ≤ 3. For large displacement movement, the movement direction of the measured object is determined according to the polarity of N. For example, if N is positive, this means that the measured object moves forward a certain displacement from the start point. For nanometer displacement movement, the movement direction is determined by the value of m. The specific determination method is as follows:

The waveforms of the reference signal ref and the measurement signal mea are shown in Fig. 5. Δφ0 is the initial phase difference between the reference and measurement signals at the start point. Δφ is the final phase difference at the stop point. If m > 0, this means that the measured object moves forward. And if m < 0, the measured object moves backward. In particular, when m = 0, this means that the measurement signal state is not changed. At this time, the movement direction is determined by the variation of phase difference Δφ − Δφ0. If Δφ − Δφ0 > 0, the measured object moves forward. And if Δφ − Δφ0 < 0, the measured object moves backward.

3.2. Fraction phase correction

As shown in Fig. 5, the phase difference between ref and mea is determined by counting the filled high frequency pulse clk [16]. When the rising-edge of ref arrives, counting the high frequency pulse clk starts until the adjacent rising-edge of mea reaches.

The phase difference is given by

\[ Δφ_0 = \frac{n_0}{N_{\text{period}}} \times 360° \quad \text{and} \quad Δφ = \frac{n}{N_{\text{period}}} \times 360° \]

where \( n_0 \) represents the number of filled clk at the start point, \( n \) represents the number of filled clk at the stop point and \( N_{\text{period}} \) represents the number of filled clk in one period of ref.

According to the movement direction determined by the above mentioned method, when the initial phase difference is not equal to the zero phase difference, the phase correction approach is as follows:

1. when \( num_{n+1} = num_n \)

\( num_{n+1} = num_n \) represents that the rising-edge of the measurement signal does not pass through any interval between two consecutive measurements. At this time, the integer fringe number \( N \) is not changed. Therefore, regardless of the initial phase difference in arbitrary interval, the fraction phase does not need to be compensated. That is: \( Δφ = N \times 360° + Δφ - Δφ_0 \).

2. when \( num_{n+1} > num_n \)

\( num_{n+1} > num_n \) represents that the rising-edge of the measurement signal forward passes through one or several intervals between two consecutive measurements. At this time, for different phase difference relationship of the \( Δφ \) and \( Δφ_0 \), the phase difference compensation is not same. When \( Δφ < Δφ_0 \) and \( Δφ > Δφ_0 \), the phase corrections are discussed below, respectively.

(A) \( Δφ < Δφ_0 \)

There are four cases as follows:

(i) When \( Δφ_0 \) is in the interval \( S_n \), the possible position of \( Δφ \) satisfying the condition of \( 0 < Δφ < Δφ_0 \) is shown in Fig. 6(a). The possible change path of the measurement signal state is plotted with dotted line (herein-after the same). The measurement signal state changes forward, and the rising-edge of the measurement signal passes through four intervals and the zero phase difference. The integer fringe counting \( N \) will add 1. Although the rising-edge of the measurement signal does not pass through a complete cycle actually, it is equivalent to having compensated the fraction phase when calculating the variation of the phase difference adopts this moment value of \( N \). Thus, in this case, the fraction phase is no longer needed to be compensated. That is: \( Δφ = N \times 360° + Δφ - Δφ_0 \).

(ii) When \( Δφ_0 \) is in the interval \( S_p \), the possible position of \( Δφ \) satisfying the condition of \( 0 < Δφ < Δφ_0 \) is shown in Fig. 6(b). If \( 90° < Δφ < Δφ_0 \), the measurement signal state changes forward, and the rising-edge of the measurement signal passes through four intervals and the zero phase difference. The integer fringe counting \( N \) will add 1. Similar to the analysis of (i), the fraction phase does not need to be compensated. But if \( 0° < Δφ < 90° \), that is, \( Δφ \) in the interval \( S_n \), the measurement signal state changes forward, and the rising-edge of the measurement signal passes through three intervals and zero phase difference. In this case, because the integer fringe number \( N \) is not changed, the fraction phase should be compensated by \( +360° \). That is: \( Δφ = N \times 360° + Δφ - Δφ_0 + 360° \).

(iii) When \( Δφ_0 \) is in the interval \( S_c \), the possible position of \( Δφ \) satisfying the condition of \( 0 < Δφ < Δφ_0 \) is shown in Fig. 6(c). If \( 180° < Δφ < Δφ_0 \), the rising-edge of the measurement signal forward passes through four intervals and the zero phase difference. The integer fringe counting \( N \) will add 1. Similar to the analysis of (i), the
fraction phase does not need to be compensated. But if \(0^\circ < \Delta \phi < 180^\circ\), that is, \(\Delta \phi \) in the interval \(S_1 \) or \(S_2\), the rising-edge of the measurement signal forward passes through less than four intervals and zero phase difference. In this case, because the integer fringe number \(N\) is not changed, the fraction phase should be compensated by \(+360^\circ\). That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0 + 360^\circ\).

(iv) When \(\Delta \phi_0\) is in the interval \(S_P\), the possible position of \(\Delta \phi\) satisfying the condition of \(0^\circ < \Delta \phi < \Delta \phi_0\) is shown in Fig. 6(d). Similar to the analysis of (iii), if \(270^\circ < \Delta \phi < \Delta \phi_0\), the fraction phase does not need to be compensated. But if \(0^\circ < \Delta \phi < 270^\circ\), that is, \(\Delta \phi\) in the interval \(S_A\) or \(S_C\), the fraction phase should be compensated by \(+360^\circ\). That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0 + 360^\circ\).

(B) \(\Delta \phi > \Delta \phi_0\)

In this case, when \(\Delta \phi_0\) is in arbitrary interval, \(\Delta \phi\) is always \(\Delta \phi_0 < \Delta \phi < 360^\circ\). The rising-edge of the measurement signal does not pass through the zero phase difference. The fraction phase does not need to be compensated. That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0\).

(3) When \(num_{n+1} < num_n\):

num_{n+1} < num_n represents that the measured object moves backward. Using the above similar analysis method, we can get

(A) \(\Delta \phi > \Delta \phi_0\)

There are also four cases as follows:

(i) When \(\Delta \phi_0\) is in the interval \(S_A\), the possible position of \(\Delta \phi\) satisfying the condition of \(\Delta \phi_0 < \Delta \phi < 360^\circ\) is shown in Fig. 7(a). The possible change path of the measurement signal state is plotted with dotted line. If \(\Delta \phi < 90^\circ\), the rising-edge of the measurement signal backward passes through four intervals and the zero phase difference. The integer fringe counting \(N\) will subtract 1. Although the rising-edge of the measurement signal does not pass through a complete cycle actually, it is equivalent to having compensated the fraction phase when calculating the variation of the phase difference adopts this moment value of \(N\). Thus, in this case, the fraction phase is no longer needed to be compensated. That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0\). If \(90^\circ < \Delta \phi < 360^\circ\), the rising-edge of the measurement signal backward passes through less than four intervals and zero phase difference. At this time, the integer fringe number \(N\) is not changed, the fraction phase should be compensated by \(-360^\circ\). That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0 - 360^\circ\).

(ii) When \(\Delta \phi_0\) is in the interval \(S_B\), the possible position of \(\Delta \phi\) satisfying the condition of \(\Delta \phi_0 < \Delta \phi < 360^\circ\) is shown in Fig. 7(b). Similar to the analysis of (i), if \(\Delta \phi < 180^\circ\), the fraction phase does not need to be compensated. If \(180^\circ < \Delta \phi < 360^\circ\), the fraction phase should be compensated by \(-360^\circ\). That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0 - 360^\circ\).

(iii) When \(\Delta \phi_0\) is in the interval \(S_C\), the possible position of \(\Delta \phi\) satisfying the condition of \(\Delta \phi_0 < \Delta \phi < 360^\circ\) is shown in Fig. 7(c). Similar to the analysis of (i), if \(\Delta \phi < 270^\circ\), the fraction phase does not need to be compensated. If \(270^\circ < \Delta \phi < 360^\circ\), the fraction phase
should be compensated by \(-360^\circ\). That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0 - 360^\circ\).

(iv) When \(\Delta \phi_0\) is in the interval \(S_{D_0}\), the possible position of \(\Delta \phi\) satisfying the condition of \(\Delta \phi_0 < \Delta \phi < 360^\circ\) is shown in Fig. 7(d). Similar to the analysis of (i), the fraction phase does not need to be compensated. That is: \(\Delta \phi = N \times 360^\circ + \Delta \phi - \Delta \phi_0\).

(B) \(\Delta \phi < \Delta \phi_0\)

In this case, when \(\Delta \phi_0\) is in arbitrary interval, \(\Delta \phi\) is always \(\Delta \phi < 360^\circ\). The rising-edge of the measurement signal...
backward passes through four intervals and the zero phase difference. The fraction phase does not need to be compensated. That is:

\[ \Delta \phi = \frac{N}{C_2} \times 360^\circ + \Delta \phi_0. \]

Integrating all the above analyses, the flow chart to correct the fraction phase is shown in Fig. 8.

### 4. The realization of signal processing

The proposed signal processing method is implemented by a field programmable gate array (FPGA) module and a computer as shown in Fig. 9. The FPGA (EP2C20Q240C8, Altera Corporation) was used to realize the integer and fraction fringe counting. When the synchronous control signal from the computer arrives, using delay method shifts the phase of the reference signal \( \text{ref} \). Four same frequency signals with phase difference of 90° each other are then generated. Using the rising-edges of the four signals as boundaries, the rising-edge of the measurement signal crossing these boundaries is detected in real time. Then, a 24-bit boundary detection data is obtained. At the same time, the phase difference between the reference signal \( \text{ref} \) and the measurement signal \( \text{mea} \) is determined by counting the filled high frequency pulse generated from the phase-locked loop. The filling pulse counting value is a 16-bit data. After transmitting the output latch, these data are sent to the computer to process. The computer is responsible for the overall control of the measurement system, the movement direction identification of the measured object and the fraction phase correction, and then calculates the measured displacement.

### 5. Experiments and results

To verify the feasibility of the proposed signal processing method, an experimental setup of laser heterodyne interferometer was constructed as shown in Fig. 10. The laser source was a dual-frequency He–Ne laser (5517C, Agilent Corporation) which emits a pair of beams with the frequency difference of 2.7 MHz and a wavelength of \( \lambda = 632.991354 \) nm. The nano-positioning stage (P-752.1CD, Physik Instrumente Corporation) whose movement range is 15 \( \mu \)m with 0.1 nm resolution was used to provide nanometer displacement to be measured. And the linear movement stage (M-521.DD, Physik Instrumente Corporation) whose movement range is 200 mm with 0.1 \( \mu \)m resolution was used to provide large displacement to be measured. The developed signal processing board based on the proposed signal processing method was used to handle the phase difference between the reference and measurement signals detected by two photodetectors (PD1 and PD2).

The first experiment was stability test. The nano-positioning stage and the linear movement stage were stationary. At this time, the changes of the reference and measurement signals near the initial phase difference would reflect the fluctuation of environmental parameters (such as the refractive index of air, vibration). In this case, the initial phase difference, the real-time phase difference with a time interval of 5 s, the integer and fraction fringes were recorded. According to the fraction phase correction method described in Section 3, the experimental result with the initial phase difference of 106.21° in the interval \( SB \) is shown in Fig. 11. It indicates that the average error is 4.12 nm, the standard deviation is 6.84 nm, the maximum error is 18.25 nm and the minimum error is –12.93 nm.

The second experiment was nanometer displacement test. The P-752.1CD nano-positioning stage performed positive or reverse direction stepper motion with nanometer displacements. The rising-edge of the measurement signal would probably pass through four intervals and the zero phase difference. In this case, for different initial phase difference, the fraction phase should be

![Fig. 9. The block diagram of the signal processing.](image)

![Fig. 10. The photograph of the experimental setup.](image)

![Fig. 11. The stability experimental result.](image)
correspondingly compensated according to the fraction phase correction method described in Section 3. The experimental results are shown in Fig. 12. Fig. 12(a) is the experimental result of stepping 50 nm with the initial phase difference of 162.25° in the interval $S_A$. The average error is $-5.59$ nm and the standard deviation is $5.14$ nm. Fig. 12(b) is the experimental result of stepping 100 nm with the initial phase difference of 42.84° in the interval $S_B$. The average error is $0.37$ nm and the standard deviation is $0.5620$ nm. The average error is $0.3376$ μm and the standard deviation is $0.3867$ μm. And Fig. 12(d) is the experimental result of stepping 500 nm with the initial phase difference in $S_C$. The average error is $-5.62$ nm and the standard deviation is $7.00$ nm.

The third experiment was micrometer displacement test. The M-521-DD linear movement stage performed positive or reverse direction stepper motion with micrometer displacement. For different initial phase difference, the fraction phase should be correspondingly compensated according to the fraction phase correction method described in Section 3. The experimental results are shown in Fig. 13. Fig. 13(a) is the experimental result of stepping $10$ μm with the initial phase difference of $44.81°$ in the interval $S_A$. The average error is $0.2338$ μm and the standard deviation is $0.5649$ μm. Fig. 13(b) is the experimental result of stepping $10$ μm with the initial phase difference of $218.29°$ in the interval $S_C$. The average error is $0.3771$ μm and the standard deviation is $0.5620$ μm. Fig. 13(c) is the experimental result of stepping $100$ μm with the initial phase difference of $292.72°$ in the interval $S_B$. The average error is $0.3376$ μm and the standard deviation is $0.3867$ μm. And Fig. 13(d) is the experimental result of stepping $1000$ μm with the initial phase difference of $122.29°$ in the interval $S_C$. The average error is $0.6684$ μm and the standard deviation is $1.0961$ μm.

The above experimental results are basically consistent with the technical parameters of the nano-positioning stage and the linear movement stage. These experiments show that the proposed signal processing method can correctly identify movement direction of the measured object and implement precision displacement measurement.

6. Conclusion

In this study we analyzed the characteristics of measuring the phase difference in laser heterodyne interferometry and presented a signal processing method of movement direction identification and phase correction to achieve precision displacement measurement with nanometer resolution. The realization approaches of the movement direction identification and the phase correction were described in detail. And the implementation of the signal processing method was designed by a FPGA device and a computer. The bi-directional movement experiments verify the feasibility of the presented signal processing method for laser heterodyne interferometer. The stability
test and nanometer displacement test experiments show that the proposed system can reach a measurement accuracy of 6.84 nm.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) under grant numbers 51375461, 51205365 and 90923026 and the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT).

References