Real-Time Estimation and Compensation of the Laser Interferometer in Nano-Scale

Yong-Woo Lee*, Hyun-Seok Choi*, Tong-Jin Park*, Chang-Soo Han*
Tae-hoon Choi**, Nak-kyu Lee**, Hyoung-wook Lee**, and Kyung-hwan Na**
*Department of Mechanical Engineering, Hanyang University, Ansan, Korea
(Tel : +82-31-400-4062; E-mail: cshan@hanyang.ac.kr)
** Korea Institute of Industrial Technology, ChonAn, Korea
(Tel : +82-3-000-0000; E-mail: aarai@japan.ac.jp)

Abstract: In this study, Real-time estimation and compensation procedure are developed for the laser interferometer. This system is designed with homodyne quadrature-phase detection method using the Laser interferometer. The errors in this system are due to noise, disturbance and undefined model dynamics. DSP(Digital Signal Processor) is applied for real time compensation of these errors. This estimator and compensation is verified with measurement test.

Keywords: Homodyne Interferometer, Displacement measurement, real-time estimation, compensator

1. INTRODUCTION

In these days, many researchers have investigated on the high precision displacement measuring system in nanometrology. The laser interferometry is the technique most frequently used for high precision displacement measurement. The laser interferometer provides the resolution much smaller than sub-nanometer, its accuracy is limited by the nonlinearity [1]. The scale division nonlinearity is mainly attributable to the electronic evaluation of two signals for which the assumed 90° phase shift is not maintained [2-4]. Measurement errors in the laser interferometer are due to a polarization mixing, a laser power drift, and a laser beam alignment, an imperfection of the electronic circuit and error motion of the stage. These errors cause the Lissajous trajectory of two phases quadrature signals to be distorted from the ideal circle [2,3].

Heydemann proposed a technique of least-square fitting of experimental data in order to assess and correct for these errors [2]. Frantisek [5] and Chien-ming [6] proposed a technique of elliptical fitting by a least-squares method to correct the nonlinearities resulting from many factors. A similar approach was reported in phase-shifting interferometry to improve the phase measuring accuracy [7,8]. However, these error compensations are very time-consuming procedure, so this method difficult to be used for the real-time measurement [1].

Wansong and Xianyu report the real-time calibration method in phase shift interferometer. It is based on the FFT(Fast Fourier Transform) and adaptive filter [9].

In this paper, a homodyne interferometer with quadrature fringe detection was constructed for displacement measurement. We try to minimize errors of miss-aligned optical parts in this optical system. The kalman filter and the digital signal processing are applied for compensation of phase shift error and real-time measurement in this study. The measurement errors are minimized using the Kalman filter and the real-time signal processing.

2. PRINCIPLE OF DISPLACEMENT MEASUREMENT INTERFEROMETER

A single-frequency polarizing interferometer with phase-quadrature detection method is shown in Fig. 1. Two partially linear polarized beams with the same frequency are split by a beamsplitter(BS) upon which 45° linearly polarized radiation impinges. A retardation plate is placed before the fixed mirror in order to obtain circularly polarized radiation.

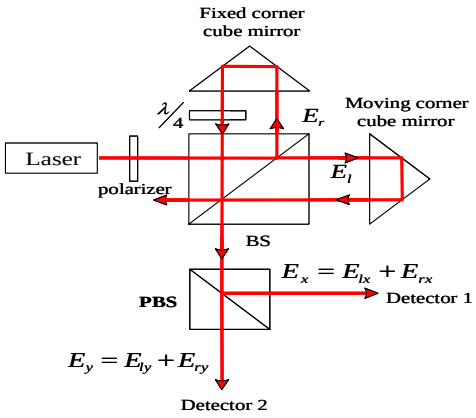

Fig. 1 Configuration of the one-frequency interferometer

The combined radiation, including linear and circular radiations, impinges upon a polarizing beamsplitter(PBS) and then produces two orthogonal and linear radiations with phase difference of $\pi/2$ between them.

The electric fields of radiation both of the moving mirror arm(E_l) and of the fixed arm(E_r) can be expressed in complex notation:

$$E_l = B\exp\left[i\left(\omega t + k\cdot r + \phi_l\right)\right] \quad (1)$$

$$E_r = A\exp\left[i\left(\omega t + k\cdot r + \phi_r\right)\right] \quad (2)$$

Here ϕ_l and ϕ_r are initial constant phases for each electric field. Because vectors k and r are in coincident directions both for the E_l and for the E_r electric field, the magnitudes of the two electric fields can be written as

$$E_l = B\exp\left[i\left(\omega t + 2kL_1 + \phi_l\right)\right] \quad (3)$$

$$E_r = A\exp\left[i\left(\omega t + 2kL_2 + \phi_r\right)\right] \quad (4)$$

After passing through a polarizing beamsplitter, the combined beam, $E_l + E_r$, splits into two components, the x component denoted by E_x and the y component denoted by E_y, with the following relations:

$$E_x = E_{lx} + E_{rx} = B_x\exp\left[i\left(\omega t + 2kL_1 + \phi_l\right)\right] + A_x\exp\left[i\left(\omega t + 2kL_2 + \phi_{rx}\right)\right] \quad (5)$$

$$E_y = E_{ly} + E_{ry} = B_y \exp[i(\omega t + 2kL_1 + \phi_l)] + A_y \exp[i(\omega t + 2kL_2 + \phi_{ry})] \quad (6)$$

Here $\phi_{rx} = \phi_l$ and $\phi_{ry} = \phi_{rx} + \pi/2 = \phi_l + \pi/2$.

The wave intensity being received by the photo-detector is proportional to the square of the electric field, so that the output signal, I_x, of detector 1 is

$$I_x \propto (E_x)(E_x)^* = B_x^2 + A_x^2 + 2A_x B_x \cos[2k(L_1 - L_2)] \quad (7)$$

and the signal of detector 2 is

$$I_y \propto (E_y)(E_y)^* = B_y^2 + A_y^2 + 2A_y B_y \cos\left[2k(L_1 - L_2) - \frac{\pi}{2}\right] \quad (8)$$

By varying the optical path length $L_1 - L_2$ (that is, by moving the measuring arm), both the output of detector 1 and that of detector 2 appear as sinusoidal signals with a phase difference of $\pi/2$ between them. An ellipse of the rotating vector with centre at h of the horizontal axis and k of the vertical axis and radius a and b can be obtained when we feed the phase quadrature signals I_x and I_y.

Here $h = A_x^2 + B_x^2, k = A_y^2 + B_y^2, a = 2A_x B_x$ and $b = 2A_y B_y$.

Assume that $E_x = E_y$ and that the gains of receivers 1 and 2 are equal, then the output figure must be a circle with $h = k$ and $a = b$. Because one revolution of the rotating vector (a phase change of 2π) is equivalent to a change in optical path length of $\lambda/2$, the measurement of optical path length is now performed in terms of measuring the phases of the rotating vector. If one can resolve the phases to within 0.1°, then length differences less than $0.1nm$ can be obtained [1].

In practice, the electric states after the polarizing beamsplitter are subject to cross talk with each other due to imperfections in the polarizing beamsplitter and other reasons. In the case with cross talk

$$E_{lx}^{'} = E_{lx} + pE_{ly} = B_x \exp[i(\omega t + 2kL_1 + \phi_l)] + pB_y \exp[i(\omega t + 2kL_1 + \phi_l)] \quad (9)$$

$$E_{rx}^{'} = E_{rx} + pE_{ry} = A_x \exp[i(\omega t + 2kL_2 + \phi_{rx})] + pA_y \exp[i(\omega t + 2kL_2 + \phi_{ry})] \quad (10)$$

$$E_{ly}^{'} = E_{ly} + pE_{lx} = B_y \exp[i(\omega t + 2kL_1 + \phi_l)] + pB_x \exp[i(\omega t + 2kL_1 + \phi_l)] \quad (11)$$

$$E_{ry}^{'} = E_{ry} + pE_{rx} = A_y \exp[i(\omega t + 2kL_2 + \phi_{ry})] + pA_x \exp[i(\omega t + 2kL_2 + \phi_{rx})] \quad (12)$$

Here p stands for the cross talk ratio. Letting $\beta_y = pB_y$, $\alpha_y = pA_y$, $\beta_x = pB_x$ and $\alpha_x = pA_x$ represent the coefficients of the cross talk term, then the two x and y component radiations with polarizing cross talk are as follows:

$$E_x^{'} = E_{lx}^{'} + E_{rx}^{'} = (B_x + \beta_y)\exp[i(\omega t + 2kL_1 + \phi_l)] + A_x \exp[i(\omega t + 2kL_2 + \phi_{rx})] + \alpha_y \exp[i(\omega t + 2kL_2 + \phi_{ry})] \quad (13)$$

$$E_y^{'} = E_{ly}^{'} + E_{ry}^{'} = (B_y + \beta_x)\exp[i(\omega t + 2kL_1 + \phi_l)] + A_y \exp[i(\omega t + 2kL_2 + \phi_{ry})] + \alpha_x \exp[i(\omega t + 2kL_2 + \phi_{rx})] \quad (14)$$

The output signals of detectors 1 and 2 with the cross talk term are denoted by $I_x^{'}$ and $I_y^{'}$:

$$I_x^{'} \propto (E_x^{'})(E_x^{'})^* = B_x^2 + \beta_y^2 + A_x^2 + \alpha_y^2 + 2A_x B_x \cos[2k(L_1 - L_2)] + 2\alpha_y \beta_y \sin[2k(L_1 - L_2)] \quad (15)$$

$$I_y^{'} \propto (E_y^{'})(E_y^{'})^* = B_y^2 + \beta_x^2 + A_y^2 + \alpha_x^2 + 2A_y B_y \sin[2k(L_1 - L_2)] + 2\alpha_x \beta_x \cos[2k(L_1 - L_2)] \quad (16)$$

Letting

$$h^{'} = B_x^2 + \beta_y^2 + A_x^2 + \alpha_y^2 \quad (17)$$

$$k^{'} = B_y^2 + \beta_x^2 + A_y^2 + \alpha_x^2 \quad (18)$$

$$\phi = 2k(L_1 - L_2) \quad (19)$$

and with a and b as defined before, Eq. (15) and Eq. (16) can be re-written in the simpler forms

$$I_x^{'} \propto h^{'} + a\cos\phi + 2\alpha_y \beta_y \sin\phi \quad (20)$$

$$I_y^{'} \propto k^{'} + b\cos\phi + 2\alpha_x \beta_x \cos\phi \quad (21)$$

Eq. (20) and Eq. (21) are similar to Eq. (7) and Eq. (8) but with an extra distortion term in each equation. These equations form another elliptical rotating vector with a new centre at $h^{'}$ and $k^{'}$, the same radius a with an extra modulated term $2\alpha_y \beta_y \sin\phi$ and the same radius b with an extra modulated term $2\alpha_x \beta_x \cos\phi$ [1].

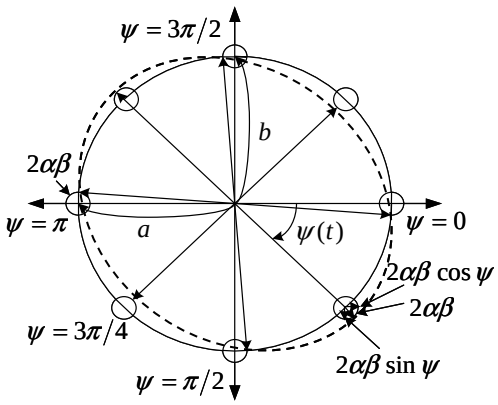

Fig. 2 phase diagram explaining the nonlinear phase error of a one-frequency interferometer

These two ellipses are illustrated in Fig. 2. For simplicity, h', k' and βx are as following; $h^{'} = h$, $k^{'} = k, \alpha_x = \alpha_y = \alpha$ and $\beta_x = \beta_y = \beta$.

In the Fig. 2, The large full ellipse with radius a and b represents, in the absence of cross talk, the ideal measurement signal which rotates about the origin with $\phi(t)$ as the difference in optical path changes.

The nonlinearity of a one-frequency interferometer can be obtained by comparing Eq. (7) and Eq. (8) to Eq. (20) and Eq. (21). Fig. 2 provides a clear basis for determining the nonlinearity. The large solid ellipse with radius a and b describes, in the absence of cross talk, the ideal measurement signal which rotates about the origin with $\phi(t)$ as the difference in optical path changes; the small full circles of radius $2\alpha\beta$ describe the distortion which is caused by polarization cross talk; and the broken ellipse describes the real measurement signal in the presence of cross talk. To account for the center shift effect caused by $h^{'}$ and $k^{'}$, Fig.

2 tells us that the nonlinearity is an asymmetrical, second-order phase error of two cycles per fringe. If the center shift by h' and k' is omitted; that is, we let $h' = h$ and $k' = k$. As shown in Fig. 2, the maximum phase leadings are at $\phi = 0$ and at $\phi = \pi$, with the same magnitude as following;

$$\Delta\phi = \tan^{-1}\left(\frac{\alpha\beta}{A_x B_x}\right) \tag{22}$$

The maximum phase laggings occur at $\phi = \pi/2$ and $\phi = 3\pi/2$ with the same magnitude

$$\Delta\phi = \tan^{-1}\left(\frac{\alpha\beta}{A_y B_y}\right) \tag{23}$$

At $\phi = \pi/4, 3\pi/4, 5\pi/4$ and $7\pi/4$, there are no phase errors.
Then,

$$\Delta L = \frac{\lambda}{4\pi}\cdot\Delta\phi \tag{24}$$

Where $\Delta L = L_1 - L_2$, $\Delta\phi$ is phase difference with fixed arm and moving arm.

3. DESIGN OF THE ESTIMATOR AND COMPENSATOR IN REAL-TIME

3.1 Estimator : Model-based Kalman filter

Consider a nonlinear system:

$$\begin{aligned} \dot{x} &= f(x,t) + w, \quad w \sim N(0,Q), \\ z &= h(x,t) + u, \quad u \sim N(0,R). \end{aligned} \tag{25}$$

f is the dynamics of the state vector x(t). Since f is used to estimate the state of the system, it is referred to as the predictor. h is the model of measurements z. It is used to correct the prediction and referred to as the corrector. w and u are the zero-mean Gaussian white system and measurement noise. The Extended Kalman filter calculates the best estimation of the state $\hat{x}$ and its covariance P:

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}) + PH^T R^{-1}(z - h(\hat{x})), \\ \dot{P} &= FP + PF^T + Q - PH^T R^{-1} HP, \end{aligned} \tag{26}$$

with the local linearizations $F = \left.\frac{\partial f(x)}{\partial x}\right|_{x=\hat{x},t}$ and $H = \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x},t}$.

The application is to integrate the dynamic model into f. This implementation is shown in Fig. 3. If the model depends on additional control parameters, these parameters are also entered in the predictor.

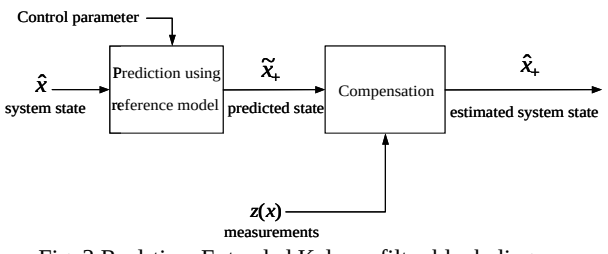

Fig. 3 Real-time Extended Kalman filter block diagram

Dynamic model is used to calculate the prediction of the next state $\tilde{x}_+$. The measurements z(x) are used to correct this prediction.

It reduces the average least-square error between the real state x and the estimated state $\hat{x}$. If the dynamic model differs from the real dynamics, this is no longer true. If f' represents the deviation of the model from the real dynamics eq. (26) now reads as follows:

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}) + P'H^T R^{-1}(z - h(\hat{x})) + f'(\hat{x}) \\ \dot{P} &= FP' + P'F^T + Q - P'H^T R^{-1} HP' + F'P' + P'F'^T \end{aligned} \tag{27}$$

$\hat{x}'$ and P in eq. (27) includes additional terms which have been underlined. These terms are perturbations of the real filter dynamics. Obviously, there is no compensation mechanism which is able to reduce this perturbation.

3.2 Compensator in real-time

In order to stabilize the Kalman filter, the perturbations in equation (27) have to be compensated. This is done by introducing another Kalman filter which estimates the errors of the state filter. The complete design is shown in Fig. 4. . The state filter is boxed. The measurements z are split between these two filters into z_1 and z_2 to avoid statistically dependent measurements. z_2 is used to measure the difference of the estimate $\hat{x}'$ and the real state.

This difference corrects the prediction of the error δx. The whole process can be described by:

$$\begin{aligned} \dot{\delta x} &= F\delta x + w_\delta, \quad w_\delta \sim N(0,Q_\delta), \\ z_\delta &= h(\delta x) + u_\delta, \quad u_\delta \sim N(0,R_\delta). \end{aligned} \tag{28}$$

The filter equations for this second filter are:

$$\begin{aligned} \dot{\delta\hat{x}} &= F\delta\hat{x} + P_\delta H_\delta^T R_\delta^{-1}(z_2 - \hat{x} - H_\delta\delta\hat{x})), \\ \dot{P}_\delta &= FP_\delta + P_\delta F^T + Q_\delta - P_\delta H_\delta^T R_\delta^{-1} H_\delta P_\delta \end{aligned} \tag{29}$$

This filter is able to correct the effect of the perturbations in eq. (27) since measurements of δx include these effects. In case of a mismodeled system, the dynamics of the mismodeled error estimation reads as follows:

$$\begin{aligned} \dot{\delta\hat{x}}' &= F\delta\hat{x}' + P_\delta' H_\delta^T R_\delta^{-1}(z_2 - \hat{x}' - H_\delta\delta\hat{x}')) + F'\delta\hat{x}' \\ \dot{P}_\delta' &= FP_\delta' + P_\delta' F^T + Q_\delta - P_\delta' H_\delta^T R_\delta^{-1} H_\delta P_\delta' \\ &\quad + F'P_\delta' + P_\delta' F'^T \end{aligned} \tag{30}$$

Eq. (30) also contains perturbations but, in this case, they can be easily handled since Q_δ is a measure of the quality of the error model. By increasing the weight of the measurements, one can reduce the influence of the perturbation.

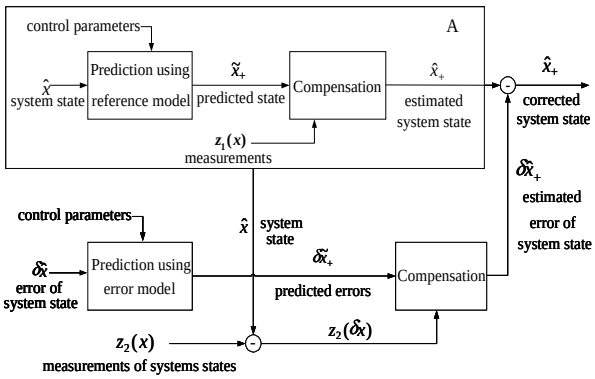

Fig. 4 Stabilized model-based Kalman filter

In the Fig. 4, This filter consists of two Kalman filters: The

boxed filter(A) estimates the state. It corresponds to the standard application of an Extended Kalman filter (Fig. 3) The second filter estimates the error of the boxed filter(A). The available measurements are split between these two filters to guarantee statistically independent measurements.

The dynamics of the standard Kalman filter is a rotated ellipse with an amplitude of about 1.5. The filter performance is less accurate than the unfiltered sensor information, which clusters around the green ellipse. The reason for this misbehavior is the fact that the corrector weights the predicted state against the measured state, but both lie on different ellipses. Since the period of both signals also differs, the result is not necessarily a ellipse between these two ellipses. The filtered values still cluster around the real ellipse, there is no obvious deviation.

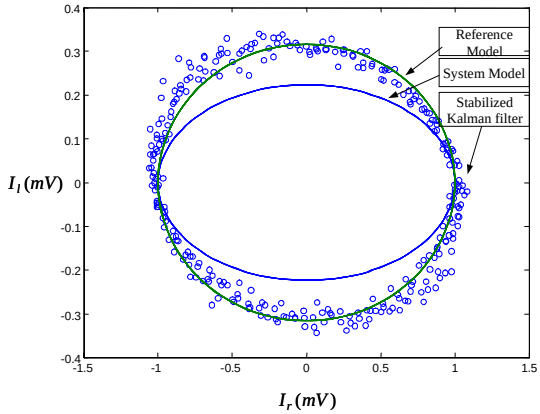

Fig. 5 Performance of the compensator

4. EXPERIMENT CONFIGURATION AND MEASUREMENT METHOD

This chapter described the configuration and measurement method of ultra-high precision measurement system.

4.1 Experiment configuration

The experimental set-up is shown in Fig.6-7. As a beam source, a vertical polarized He–Ne laser at the wavelength 0.6329μm is used. In the path of the laser beam, a half-wave plate first rotates the polarization from vertical to about 45°, so that the polarizing beam splitter that follows can split the beam into two nearly equal components, one directed to the reference mirror and the other to the object mirror, which is fixed on a flexure micro-stage being actuated by a PZT. Both beams from two mirrors travel back and are interfered at the polarizing beam-splitter. The non-polarizing beam splitter divides the beam into two nearly equal components. One of these components is sent directly to a polarizing beam splitter, which produces a pair of interference patterns in phase opposition. The other component is made to pass through a quarter-wave plate at 45°, so that an extra $\lambda/4$ delay is added to the optical path difference. The final polarizing beam splitter then produces a pair of patterns in phase opposition, but with $\pi/2$ phase lag with respect to the previous ones. Overall, four quadrature signals are available at the photo-detectors. The interference signals detected by four photo-detectors have the phases of 0°, 90°, 180° and 270°, respectively.

Two signals $I_x + I_y$ with a phase difference of 90° are obtained by subtracting the signal with phase of 180° from the signal of 0° and the signal with phase 270° from the signal of 90°.

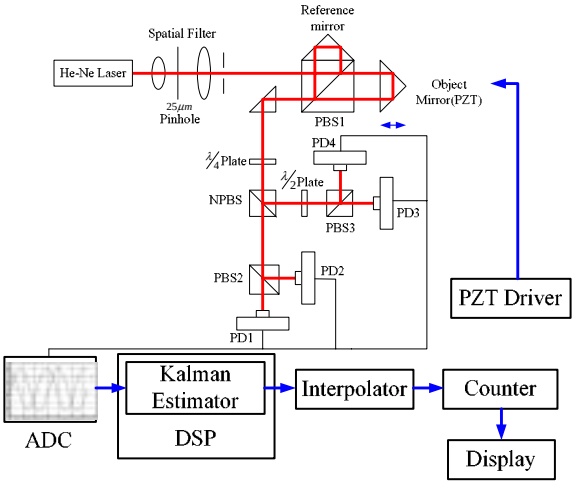

Fig. 6 Schematic diagram of experiment

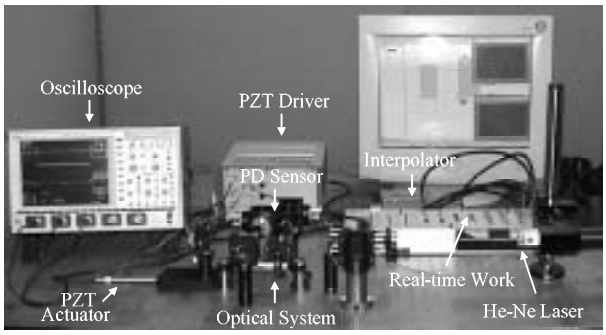

Fig. 7 Picture of optical system

Each detector consists of a PIN-photodiode with integrated amplifier. These detectors are connected to the ADC which allows these signals to be transferred to a DSP for further signal processing. A set of elliptical parameters used in calculating the displacement.

5. EXPERIMENT RESULT AND SIMULATION

An ideal Lissajous pattern is produced by two identical interferometric signals(Fig. 8) that are in perfect quadrature and that have zero dc offsets.

If these signals are used to produce a Lissajous patten, a rotating vector is obtained which describes a circle of radius R: $R = I_r^2 + I_l^2$

This is illustrated in Fig. 9, where we can see that one revolution of the vector is equivalent to an optical phase change of 2π. The instantaneous phase θ is usually obtained by the arctangent of I_l / I_r. The fringe fraction is given by θ/2π:

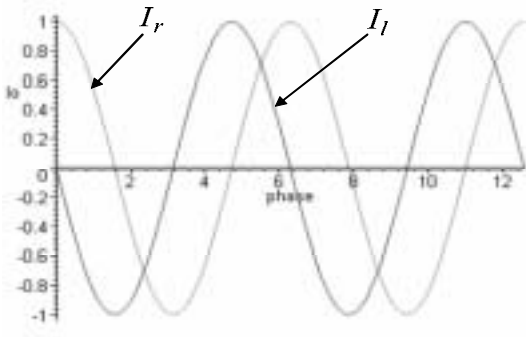

Fig. 8 Ideal phase quadrature interferometer signals

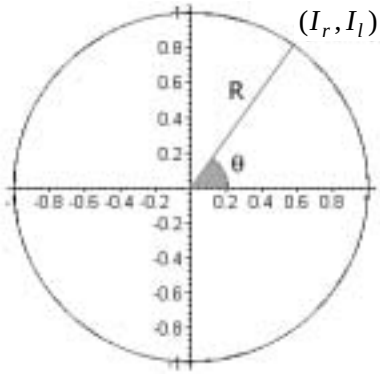

Fig. 9 Ideal Lissajous pattern

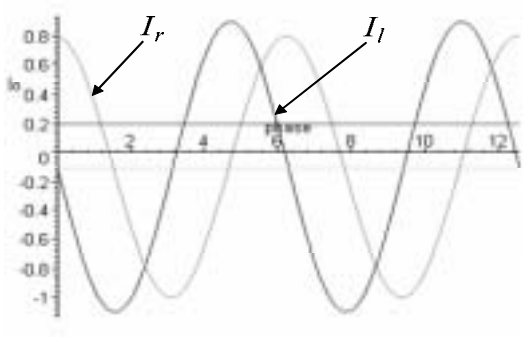

Fig. 10 Simulated realistic signals

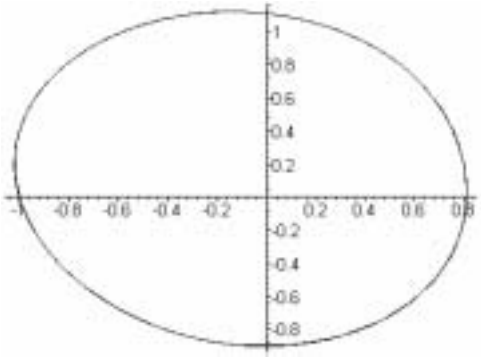

Fig. 11 Realistic Lissajous pattern

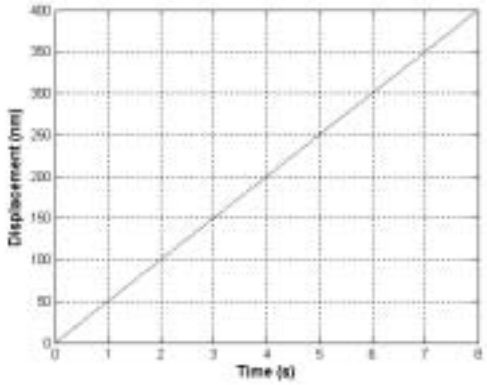

Fig. 12 Linear response to piezo device

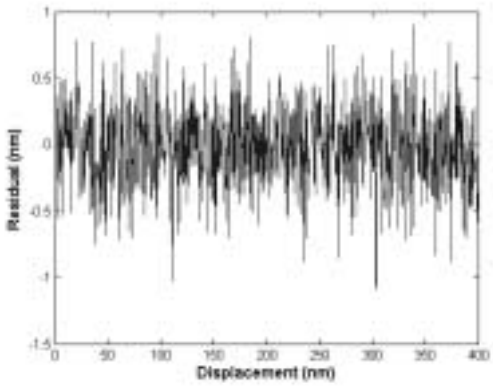

Fig. 13 residual deviation for the interferometer signals

We also checked the cyclic error by analyzing the response of the interferometers in the single-axis moving. The displacement of the mirror is plotted as a function of time, as shown in Fig. 12. Almost linear response of the applied voltage to the piezo device with a constant increase was obtained. The slight curve actually exhibits nonlinear motion of the piezo device. Fig. 13 shows a plot of the residual deviation. The residual deviation is quite small with an amplitude of less than 50.1 nm. Its period of sinusoidal waveform of about 40nm indicates that the deviation is caused by the interferometer rather than the motion error of the mirror.

6. CONCLUSIONS

The high precision displacement measurement system using the real-time estimation and compensation system was researched in this paper. The conclusions of this paper are as follows:

(1) The high precision displacement measurement system is designed and constructed in order to reduce outer disturbances.

(2) The output signal parameter is estimated using the Kalman filter in real-time process.

(3) The reference model is set to the perfect circular trajectory in Lissajous graph for designing the compensator.

(4) It is verified that the proposed error compensator is effective method for decrease error in high precision displacement measurement system.

ACKNOWLEDGMENTS

This work has been sponsored by MOCIE(Ministry of Commerce, Industry and Energy) of Korea as a part of the project of "Development of Micro optical and thermofluidic devices with high functionality".

REFERENCES

[1] T. B. Eom and J. Y. Kim, “The dynamic compensation of nonlinearity in a homodyne laser interferometer”, KSPE, Vol. 18, No. 9, 2001.

[2] Heydemann LM., “Determination and correction of quadrature fringe measurement errors in interferometers”, Applied Opt, Vol. 20, No. 19, pp. 3382–4, 1981.

[3] Norman Bobroff, “Recent advances in displacement measuring interferometry”, Meas. Sci. Technol, Vol. 4, pp. 907-26, 1993.

[4] Chien-ming Wu and Ching-shen Su, “Nonlinearity in measurements of length by optical interferometry”, Meas. Sci. Technol., Vol. 7, pp. 62– 6, 1996.

[5] Frantisek Petru and Ondrej Cip, “Problems regarding linearity of data of a laser interferometer with a single-frequency laser”, Elsevier Sci. Inc., Vol. 23, pp. 39–50, 1999.

[6] Chien-ming Wu, Ching-Shen Su and Gwo-Sheng Peng, “Correction of nonlinearity in one-frequency optical interferometry”, Meas. Sci. Technol, Vol. 7, pp. 520–4, 1996.

[7] C. T. Farrell and M. A. Player, “Phase step measurement and variable step algorithms in phase-shifting interferometry”, Meas. Sci. Technol., Vol. 3, pp. 953-8, 1992.

[8] C. T. Farrell and M. A. Player, “Phase-step insensitive algorithms for phase-shifting interferometry”, Meas. Sci. Technol., Vol. 6, pp. 648-52, 1994.

[9] Wansong Li and Xianyu Su, “Real-time calibration algorithm for phase shifting in phase-measuring profilometry”, SIPE, Vol. 40, pp. 761-6, 2001.

[10] Karl Johan Astrom and Bjom Wittenmark, “Adaptive control (second edition)”, chapter, Addison-Wesley Publishing Company, 1995.