Wordlength Estimation of Digital Controller Synthesis for Inkjet Printer Mechanism

Hung-Ming Cheng
Georgia Southern University, Statesboro, GA, USA
Email: hmcheng@georgiasouthern.edu

Georgia T.-C. Chiu
Purdue University, West Lafayette, IN, USA
Email: gchiu@purdue.edu

Abstract— **The effect of finite wordlength on coefficients in implementing discrete-time controllers has been a subject of many recent studies. Especially, this issue is more severe in the low budget consumer electronic products when a faster sample rate is desired. To save the cost of development, most of the manufacturers applied cheap processors on those kinds of products. Therefore, it is desired to develop an estimation method to predict the required wordlength while designing a controller. In this paper, we will focus our study on the controller synthesis for an inkjet printing mechanism and investigate the effect insufficient wordlength. We will also develop an algorithm to estimate the required resolution and apply the method to determine the required bits of wordlength for the controller designed.**

Index Terms—**finite wordlength, digital controller, inkjet printer**

I. INTRODUCTION

As the operating speed of microcontroller becomes faster and faster, more and more industrial applications are switched from analog devices to digital components. With those state-of-the-art digital microprocessors, the sample rate applied is faster than ever. Based on sampling theorem [1], a signal can be completely reconstructed if applied sample rate is fast enough. However, this claim is only valid when the resolution of the applied hardware is sufficient. It is not always true for the synthesis of every digital controllers and filters. Nevertheless, majority of digital controller synthesis techniques in modern control are based on precision computation without the consideration of implementing resolution. Due to the cost and efficiency considerations, the designed controllers are then deployed using microcontrollers with limited resolution. The conversion from a floating point representation to a fixed-point arithmetic is non-trivial and has in many cases become the major impediment for implementing advanced control algorithms. Coefficients, gain constants, and the results of mathematical operations have to be represented in the limited wordlength in decimal fraction and may lose the necessary precision.

According to z-transform, as the sample rate approaches to zero, all stable poles will converge to a single point $(1, j0)$. If the adopted hardware of the controller does not have enough resolution close to this specific point, fast sample rate does not imply a better approximation. Instead, poor rounding scheme due to the limited resolution can distort the controller response and even result in instability. Thus, it is important to develop a technique to estimate the required wordlength of the hardware so that the controller can be kept within the stable region – the unit circle.

In the past two decades, the effect of quantization on a sampled signal has been the topic of many prior researches [2]. Quantization errors can be modeled as a uniformly distributed random measurement noise [3]-[5]. Different controller synthesis techniques have been proposed to minimize the effect of quantization noise. From the results of different research groups, it is well known that without proper consideration for the finite precision implementation of the controller coefficients when selecting a sample rate, a well designed stable closed-loop system may not achieve its desired performance or may even become unstable. Therefore, a common objective in finite precision analysis is to choose a suitable wordlength such that the finite precision implementation is a sufficiently accurate realization of the desired frequency response while minimizing the hardware and software complexity and cost.

Other than the required wordlength, the coupling factor, sample rate, needs to be taken into consideration while designing a controller. Notice that significant numerical issues can arise when the sample rate of a finite precision digital filter is much higher than its cut-off frequency [2]. In this case, higher sample rate no longer implies a better approximation to the desired frequency response. Three earlier papers [6]-[8] have discussed the coupled relationship between the sample rate and the wordlength for second-order systems. Fialho and Georgiou [9] also studied the stability and performance of sampled-data system subjected to wordlength constraint for higher order controllers. Their works concluded that the highest sample rate is restricted by the resolution of the implementing hardware, and the accuracy of a digital controller is also limited by the finite wordlength (FWL) of its implementation.

To illustrate the limitation on sample rate, consider a notch filter design with a notch at 16 rad/sec. For an 8-bit implementation, the maximum sample frequency of this specific filter is around 700 Hz if the controller is implemented using the direct form. In this case, finite precision implementation of the notch filter is realized at sample frequencies of 300 Hz and 800 Hz, see Table I.

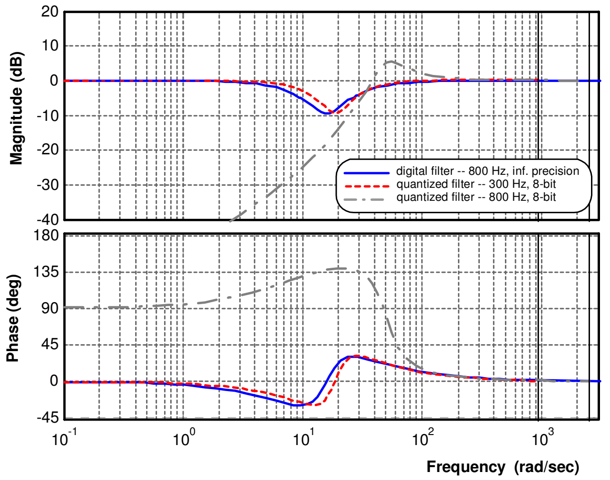

Figure 1. Frequency responses of different realizations of a 2nd order digital filter.

TABLE I.
DIFFERENT REALIZATIONS OF A NOTCH FILTER

	S/R = 800 Hz	S/R = 300 Hz
Floating Point	$\frac{z^2-1.987z+0.9877}{z^2-1.963z+0.9632}$	$\frac{z^2-1.966z+0.9682}{z^2-1.902z+0.9048}$
Quantized (8-bit)	$\frac{z^2-1.988z+0.9883}{z^2-1.961z+0.9648}$	$\frac{z^2-1.965z+0.9688}{z^2-1.902z+0.9063}$

The corresponding frequency responses are shown in Fig. 1. For an 8 bit realization, if the sample rate is higher than the limitation calculated, 800 Hz in this case, the finite precision realization can not properly realize the desired frequency response. Instead, a realization with a slower sample rate, 300 Hz in this case, which is within the upper bound can adequately realize the desired frequency response.

The effect of FWL on the coefficients in implementing digital controllers has been a subject of many studies. A lot of researches have pointed out the apparently high sensitivity between the filter coefficients and the filter characteristics. Good reviews on the effects of FWL in digital filters, such as coefficients and signal quantization, can be found in [10] and [11]. One approach to evaluate the effect of filter quantization is to measure the difference between the realized poles and zeros and the desired poles and zeros of the filters. For a given accuracy requirement in pole/zero location, higher sample rate requires more bits. Rader and Gold [12] showed that for a given filter realization, if the poles (or zeros) are tightly clustered, it is possible that small errors in the denominator (numerator) coefficients can cause large shifts in the poles (or zeros) locations. Mantey [13] has studied the eigenvalue sensitivity of the selected representation of the system model, and he has derived the deviation of eigenvalues due to the perturbation of the state matrix. To reduce the effect due to quantization, the delta operator was introduced by Middleton and Goodwin [14], [15] since 1986. Compared to z-operator, the delta operator has better numerical properties and is more robust to rounding errors. Many research groups [16]-[18] have also developed the techniques of controller synthesis to restrict the response deviation caused by quantization within a bounded norm. The numerical issue due to quantization is more severe in lower-end consumer electronic products. To save the cost of production, a lot of companies utilize low-cost processors on their products. In this paper, we will focus our study on an inkjet printing mechanism.

The expansion of personal computing and the ever-popular multi-medialization of personal computers facilitated the increased demand for affordable printing solution for variety of applications. Among all of the solutions, inkjet printers are the most common type of computer printers for the general consumer due to their high quality of output, capability of printing in vivid color, and ease of use. They offered an attractive feature set that fits well to the consumer's needs at the low end of the market. However, in the earliest phase of controller design, the calculation of controller synthesis methods is based on the assumption that the adopted processors have infinite resolution [19]. The effect of the finite wordlength was never taken into consideration. In this paper, we will investigate the effect of a controller without a proper wordlength and apply the developed estimating method to determine the required resolution of the controller designed.

This paper is organized as following. The second section will discuss the deviation caused by the rounding scheme applied. A method to estimate the required wordlength will be developed in this section. The third section will use an inkjet printing mechanism to illustrate the developed method. Conclusions are given in the last section.

II. ESTIMATION OF REQUIRED WORDLENGTH FOR TRANSFER FUNCTION REPRESENTATION

The z-transform is one of the mathematical means used for the analysis and design of discrete-time control systems. The delay operator z^{-1} and its associated z-transform are synonymous as the method of implementing digital controllers. A discrete-time single input/single output (SISO) linear time invariant controller $C(z)$ can be realized by the following two transfer function representations:

$$C(z)=\frac{N(z)}{D(z)}=\begin{cases}\dfrac{\sum_{k=0}^{M} b_k\cdot z^{-k}}{1-\sum_{k=1}^{N} a_k\cdot z^{-k}} & \text{Direct form}\\[2ex] \dfrac{b_0\cdot\prod_{k=1}^{M}\left(1-z_k\cdot z^{-1}\right)}{\prod_{k=1}^{N}\left(1-p_k\cdot z^{-1}\right)} & \text{Coupled form}\end{cases}, \qquad (1)$$

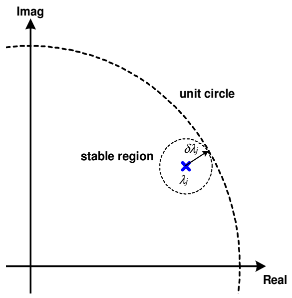

Figure 2. Stable region of the j^{th} quantized pole.

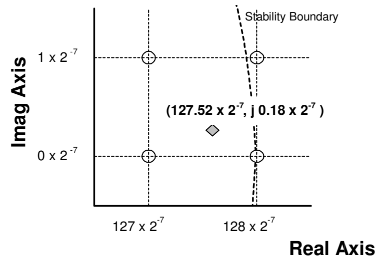

Figure 3. Pole movement due to the rounding of quantization.

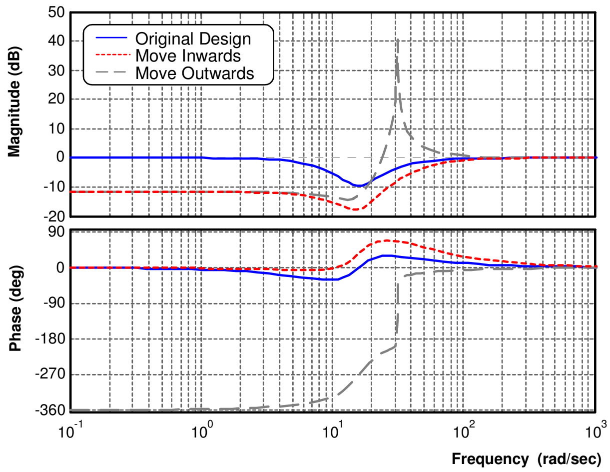

Figure 4. Frequency responses of a notch filter with different rounding methods realization.

where $\{a_k\}$ and $\{b_k\}$ are the coefficients of the denominator and numerator polynomials, $D(z)$ and $N(z)$ respectively. $\{z_k\}$ and $\{p_k\}$ are the zeros and poles of the transfer function, respectively. The first representation, which is the ratio of the numerator and denominator polynomials, is called the *direct form* and the second representation is called the *coupled form.* The sampling period T_s is embedded in the coefficients of the direct form and in the zeros and poles of the coupled form. With most of the controller synthesis techniques, the polynomial coefficients or the poles and zeros are assumed to have infinite precision.

A. Rounding of Coefficients

The realizable grids on the complex plane are determined by the applied wordlength. The quantized grids on the complex plane are different from each other if different realization scheme is used. Once the resolution of the adopted hardware is selected, the total amount of the implementable controllers is also determined. Therefore, it is important to estimate the required wordlength in order to estimate the acceptable deviation on the complex plane to prevent the poles from being perturbed into the unstable region or to keep the poles within an acceptable neighborhood of the desired location. As shown in Fig. 2, if a pole location is close to the border of the unit circle, improper quantization may perturb a stable pole away from the unit circle and result in an undesired controller.

For example, if the notch filter mentioned in the previous section is implemented using the coupled form and the sample rate used is 4 kHz, the resulting transfer function is

$$C_z(z)=\frac{z-0.9987+j0.0038}{z-0.9963+j0.0014}\cdot\frac{z-0.9987-j0.0038}{z-0.9963-j0.0014}$$

and the poles before quantization are

$$p_{1,2}=127.52\times2^{-7}\pm j0.1775\times2^{-7},$$

which are very close to the stability boundary. As shown in Fig. 3, there are four possible rounding values to implement the complex poles. The quantized filter will have different responses if different rounding methods are used. It will be preferable to choose a rounding scheme that shifts the resulting poles inwards to $(127\times2^{-7}, j1\times2^{-7})$ rather to the nearest location, which is the (1, j0) point. The corresponding frequency responses are shown in Fig. 4.

However, it is not trivial to estimate the response change caused by rounding methods. Different realizations of the same transfer function have different mappings on to the complex plane, and the corresponding rounding schemes will change the responses differently.

For instance, From Eq. (1), the characteristic polynomial of a 2nd order system can be implemented in the direct and the coupled forms. They are

$$D(z)=1+a_1z^{-1}+a_2z^{-2},$$

and

$$D(z)=\left(1+p_1z^{-1}\right)\cdot\left(1+p_2z^{-1}\right)$$

respectively. For the direct form, when the coefficients of the polynomial are quantized to $\hat{a}_1$ and $\hat{a}_2$, we have

$$a_k-2^{-B}\le\hat{a}_k\le a_k+2^{-B},$$

where a_k is the original coefficient and $\hat{a}_k$ is the quantized value. The corresponding quantized poles $\hat{p}_{1,2}$ can be computed by

$$\hat{p}_{1,2}=\frac{-\hat{a}_1}{2}\pm\sqrt{\frac{\hat{a}_1^{\,2}}{4}-\hat{a}_2}.$$

If the transfer function is in the *coupled form*, then the poles can be quantized directly, i.e.

$$\hat{p}=\mathrm{Im}\left(\hat{p}_k\right)+\mathrm{Re}\left(\hat{p}_k\right),\begin{cases}\mathrm{Re}\left(p_k\right)-2^{-B}\le\mathrm{Re}\left(\hat{p}_k\right)\le\mathrm{Re}\left(p_k\right)+2^{-B}\\ \mathrm{Im}\left(p_k\right)-2^{-B}\le\mathrm{Im}\left(\hat{p}_k\right)\le\mathrm{Im}\left(p_k\right)+2^{-B}\end{cases}$$

After the quantized poles in the z-plane are known, the achieved damping ratio ξ and the natural frequency ω_n can be found by

$$\omega_n = \frac{1}{T_s}\sqrt{\ln(\hat{p}_1)\cdot\ln(\hat{p}_2)}, \text{ and } \xi = \frac{\ln(\hat{p}_1)+\ln(\hat{p}_2)}{\sqrt{\ln(\hat{p}_1)\cdot\ln(\hat{p}_2)}}.$$

Thus, the deviation of natural frequency and damping ratio caused by quantization in different sample rates with the same resolution can be estimated. Notice that this estimation is only valid for 2nd order systems. For higher order systems, a different mapping technique is required.

B. Deviation of Poles due to Quantization – the Direct Form

The deviation of a pole is defined as the expected change in the location of the pole of the characteristic equation from changes in the characteristic equation coefficients. Such coefficient variations are the results of round-off when implementing the controller/filter in a fixed precision processor with finite wordlength. In this section, we will study the deviation of the poles of a higher order filter using sensitivity analysis.

If a controller is represented in the *direct form*, the corresponding characteristic equation can be written as

$$P(z) = 1+\alpha_1 z^{-1}+\alpha_2 z^{-2}+\cdots+\alpha_n z^{-n} = 0 \tag{2}$$

where $\{\alpha_k\}$, $k = 1, 2, \ldots, n$, are the coefficients of the polynomial. This equation has roots at $\lambda_1, \lambda_2, \ldots, \lambda_n$. Assume that one of the coefficients α_k is subject to a perturbation $\delta\alpha_k$. If α_k is changed to $\hat{\alpha}_k+\delta\alpha_k$, where $\hat{\alpha}_k$ is the quantized value of α_k, then the j^{th} pole λ_j will also be perturbed to a "quantized" root $\hat{\lambda}_j+\delta\lambda_j$, where $\hat{\lambda}_j$ is the j^{th} quantized pole and $\delta\lambda_j$ is the perturbation of the pole due to the changes in the characteristic coefficient. The quantized polynomial $\hat{P}(z)$ can be written as

$$\hat{P}(z) = 1+\hat{\alpha}_1 z^{-1}+\hat{\alpha}_2 z^{-2}+\cdots+\hat{\alpha}_n z^{-n} \tag{3}$$

the above equation has roots at $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n$. By taking the first order approximation of the original polynomial at the j^{th} quantized pole $\hat{\lambda}_j$, the quantized polynomial can be written as

$$\hat{P}(\hat{\lambda}_j) = P(\lambda_j)+\left.\frac{\partial P}{\partial z}\right|_{z=\lambda_j}\delta\lambda_j+\sum_{k=1}^{n}\frac{\partial P}{\partial\alpha_k}\delta\alpha_k+\Delta P_{hot}(\hat{\lambda}_j) = 0 \tag{4}$$

where $\Delta P_{hot}(\hat{\lambda}_j)$ are the higher order terms. Assuming the higher order terms, ΔP_{hot}, is negligible and noting the fact that $P(\lambda_j) = 0$, then the deviation of quantized pole, $\delta\lambda_j$, of the polynomial can be approximated by

$$\delta\lambda_j = -\left.\frac{\sum_{k=1}^{n}\frac{\partial P}{\partial\alpha_k}\delta\alpha_k}{\frac{\partial P}{\partial z}}\right|_{z=\lambda_j} = -\frac{\sum_{k=1}^{n}\lambda_j^{n-k}\cdot\delta\alpha_k}{\prod_{k=1,k\neq j}^{n}(\lambda_k-\lambda_j)} \tag{5}$$

C. Deviation of Poles due to Quantization – the Coupled Form

If the controller is implemented in the *coupled form*, the characteristic equation is represented as

$$P(z) = (z-\lambda_1)\cdot(z-\lambda_2)\cdots(z-\lambda_n) = 0 \tag{6}$$

If the characteristic equation is quantized at $z = \lambda_j$, the quantized polynomial is written as

$$\hat{P}(z) = (z-\hat{\lambda}_1)\cdot(z-\hat{\lambda}_2)\cdots(z-\hat{\lambda}_n) = 0, \boldsymbol{\lambda}_q\in\mathbb{C}^n \tag{7}$$

In this case, the change of the j^{th} pole λ_j due to quantization is simply given by

$$\delta\lambda_j = \lambda_j - \hat{\lambda}_j \tag{8}$$

D. Estimation of Required Wordlength

If the controller is implemented in the direct form, Eq. (5) shows the deviation of the j^{th} pole due to quantization. If the required wordlength of the j^{th} pole of the controller is N_j bits, the upper bound of quantization error for the coefficients is $|\delta\alpha_k| < 2^{-N_j}$. By taking the magnitude of the complex numbers of Eq. (5), we have

$$|\delta\lambda_j| < \left|\frac{\sum_{k=1}^{n}\lambda_j^{n-k}}{\prod_{k=1,k\neq j}^{n}(\lambda_k-\lambda_j)}\right|\cdot 2^{-N_j} \tag{9}$$

If the acceptable deviation of each pole is specified by Δ_j, from the above inequality, the required wordlength of the j^{th} pole is

$$2^{N_j} < \frac{\left|\sum_{k=1}^{n}\lambda_j^{n-k}\right|}{\left|\prod_{k=1,k\neq j}^{n}(\lambda_k-\lambda_j)\right|}\cdot\frac{1}{|\Delta_j|}. \tag{10}$$

By taking the base 2 logarithm of the above inequality, the required wordlength can be derived as

$$N_j = \left\lfloor \log_2\left(\frac{\left|\sum_{k=1}^{n}\lambda_j^{n-k}\right|}{\left|\prod_{k=1,k\neq j}^{n}(\lambda_k-\lambda_j)\right|}\cdot\frac{1}{|\Delta_j|}\right)\right\rfloor+1 \tag{11}$$

where the operator $\lfloor\cdot\rfloor$ denotes taking the greatest integer within the bracket.

Similarly, if the controller is implemented in the coupled form, Eq. (8) shows the deviation of the pole due to quantization. The quantization errors $\delta\lambda_j$ of the j^{th} pole should be smaller than the quantization error, and the bound on the deviation of the j^{th} pole is

$$|\mathrm{Re}(\delta\lambda_j)| < 2^{-N_j} \text{ and } |\mathrm{Im}(\delta\lambda_j)| < 2^{-N_j} \Rightarrow |\Delta_j| < \sqrt{2}\cdot 2^{-N_j} \tag{12}$$

By specifying the acceptable deviation of the j^{th} pole to be Δ_j, the required wordlength of the j^{th} pole can be calculated by

$$N_j = \left\lfloor \frac{1}{2}+\log_2\left(\frac{1}{|\Delta_j|}\right)\right\rfloor+1 \tag{13}$$

The required wordlength of the characteristic equation for both forms can be derived by taking the maximum integer of $\{N_j\}$, i.e.

$$N_r = \max_j\{N_j\} \tag{14}$$

Eq.(14) provides a conservative estimation of the required wordlength. Regardless of the rounding scheme applied to the polynomial, the required wordlength calculated guarantees that quantized poles are located within a circle with radius $|\Delta_j|$ centered on the desired pole location.

III. Experimental Inkjet Printer System

The experimental platform is a paper advance system used in inkjet printer as shown in Fig. 5. It consists of a DC motor mounted with two helical pinions as the actuator. The DC motor is driven by a current mode power amplifier, and the two helical gears are used to engage with the pinions to transmit motion. A 31.75 mm diameter polymer drive roller is used to advance the paper. The traction force needed to advance the paper is provided by two pivot bars with rollers on one end that press against the drive roller. The pivot bars are designed by using a compliant mechanism approach that provides 9.807 N of normal force. As the paper exits the print zone, a second roller supplies a slight tension on the paper to maintain a flat paper profile in the print zone. Paper position is derived from a rotary encoder disk mounted at the drive roller shaft. It should be noted that the paper position is calculated based on a perfect round roller with an ideal radius. In an actual printer, the eccentricity of the encoder, the gear train, and the roller as well as the tolerance on the roller diameter and themanufacturing tolerance on the encoder disk will affect the final media positioning accuracy.

To setup the experimental platform, a Compaq Pentium II 400 MHz PC with 128 MB memory was used for data acquisition and implementation of the control algorithm. A ComputerBoards DAS-1602-16 I/O board with a 12-bit resolution D/A unit was used to send out the command signals to the motor driver. The I/O board was also used for acquiring the measurement of the angular velocity of the roller with a 16-bit resolution of A/D. Paper position can be derived from a rotary encoder disk mounted at the drive roller shaft. The resolution of the encoder is 2048 pulse/rev. The encoder is connected to a ComputerBoards PCI-QUAD04 encoder interface board.

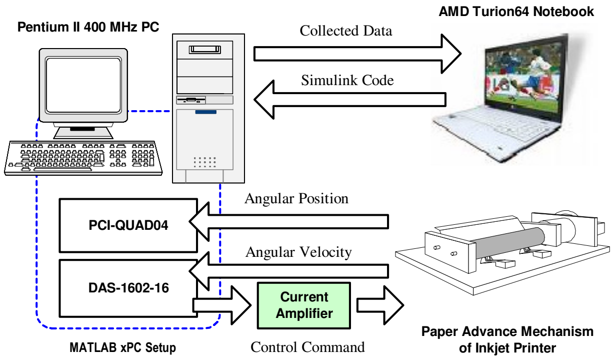

Figure 5. Schematic diagram of the control system.

The data acquisition and control algorithm was developed using MATLAB Simulink. The Simulink model was compiled into executable binary code using a Notebook computer with a 1.6 GHz AMD Turion64 microprocessor and 1 GB memory. The real-time kernel supplied with the xPC Toolbox that comes with MATLAB 2006a is used.

A. System Dynamic and Modeling of the Mechanism

To design the controller for the experimental system, an accurate plant model is necessary. The nominal plant model can be obtained using well established system identification techniques, e.g. frequency response, step response and direct calculation. In this section, the experimental system was modeled by the block diagram shown in Fig. 6. Table II summarizes the parameters of the plant.

For simplicity, it was assumed that the analog current controller is well designed and that the current loop is fast enough so that the dynamics of the current loop can be

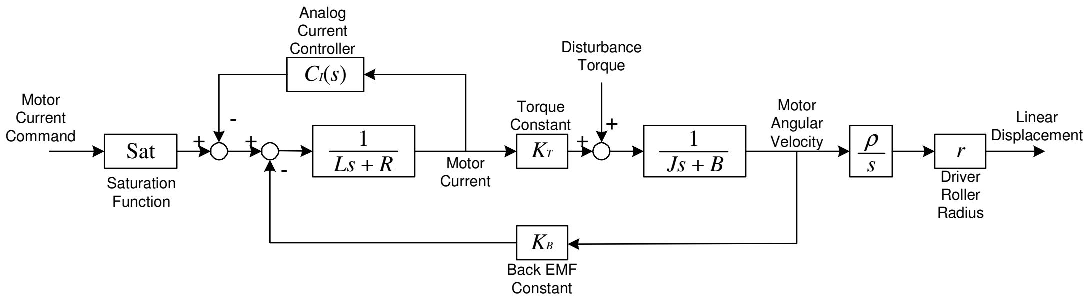

Figure 6. Block diagram of the experimental media advance system.

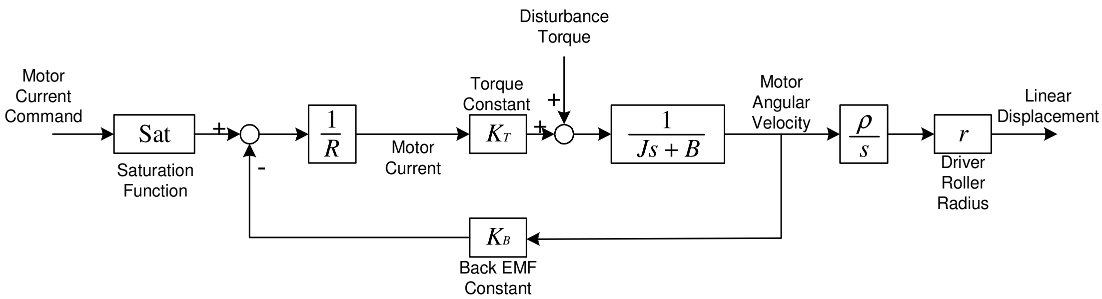

Figure 7. Simplified plant model.

TABLE II.
CHARACTERISTICS OF THE EXPERIMENTAL SYSTEM

Description	Symbol
Effective rotational inertia	J
Effective viscous friction coefficient	B
Motor armature inductance	L
Motor armature resistance	R
Motor torque constant	K_T
Motor back-EMF constant	K_B
Gear ratio	ρ
Drive roller radius	R

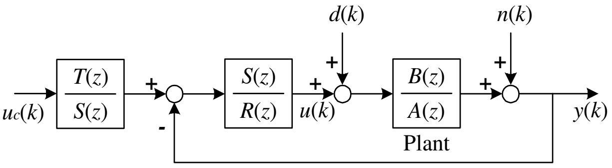

Figure 8. Schematic diagram of the control system.

ignored. Based on that assumption, the block diagram can be simplified as shown in Fig. 7. If the saturation function is ignored, the transfer function from the motor current command to the linear displacement of the advanced paper can be represented by the following transfer function:

$$G_P(s)=\frac{K_0}{s\cdot(1+\tau\cdot s)} \tag{15}$$

where

$$K_0=\frac{K_T\rho r}{BR+K_TK_B}. \tag{16}$$

$$\tau=\frac{JR}{BR+K_TK_B}$$

In this case, the only parameters that need to be experimentally identified are K_0 and τ. Due to the existence of nonlinear frictions, the parameters of the system identification of the experimental results change in a range of values for each of the parameters. The mean values of K_0 and τ are used for the nominal plant model, which can be written as

$$G_P(s)=\frac{57.25}{s\cdot(1+0.08s)} \tag{17}$$

Using zero-order-hold transformation, the pulse transfer function of the nominal plant can be represented as

$$G(z)=\frac{B(z)}{A(z)}=\frac{K(z-b)}{(z-1)(z-a)} \tag{18}$$

where

$$K=K_0(\tau\cdot e^{-T_s/\tau}+T_s-\tau),$$

$$a=e^{-T_s/\tau},\ b=\frac{(\tau\cdot e^{-T_s/\tau}+T\cdot e^{-T_s/\tau}-\tau)}{(\tau\cdot e^{-T_s/\tau}+T-\tau)},$$

and T_s is the sampling period. In this case, the sample rate picked was set at 1 kHz.

B. Controller Synthesis

Fig. 8 shows the block diagram of a general two-degree of freedom (TDOF) control system, where $B(z)/A(z)$ is the pulse transfer function of the plant, $u_c(k)$ is the reference input, $y(k)$ is the measurement output, $u(k)$ is the control signal, and $d(k)$ and $n(k)$ are disturbance and measurement noise, respectively. The control law can be written as

$$u(k)=\frac{T(z)}{R(z)}u_C(k)-\frac{S(z)}{R(z)}y(k) \tag{19}$$

and the output of the closed-loop system is

$$\begin{aligned} y(z)&=G_{u_cy}(z)u_C(z)+G_{dy}(z)d(z)+G_{ny}(z)n(z)\\ &=\frac{BT}{AR+BS}u_C(z)+\frac{BR}{AR+BS}d(z)+\frac{AR}{AR+BS}n(z)\end{aligned} \tag{20}$$

The design problem is to find the polynomials $R(z)$, $S(z)$, and $T(z)$ so that the input/output behavior between the command input u_c and the plant output y satisfies the desired specifications and the closed-loop system is robust with respect to external disturbances and measurement noise. To incorporate an integrator in the controller to compensate for constant disturbances such as Coulomb friction, an internal model of the form $(z - 1)$ is incorporated into $R(z)$ as a factor, i.e.

$$R(z)=(z-1)\cdot(z+r_1) \tag{21}$$

Since the polynomial $A(z)$ is second order, we can conclude that to be able to arbitrarily assign the closed-loop characteristic polynomial $A_{CL}(z)$, we need to have $\deg[A_{CL}]\geq 4$. Let

$$A_{CL}(z)=z^4+a_{C1}\cdot z^3+a_{C2}\cdot z^2+a_{C3}\cdot z+a_{C4} \tag{22}$$

and define

$$S(z)=s_0\cdot z^2+s_1\cdot z+s_2.$$

Then a standard model matching design is employed to find the polynomials $R(z)$, $S(z)$, and $T(z)$:

1) Factor polynomial $B(z)$ into cancelable and un-cancelable parts $B(z) = B^+(z)\cdot B^-(z)$, where $B^+(z)$ is monic and contains all the cancelable zeros of $B(z)$ while $B^-(z)$ contains all the un-cancelable zeros of $B(z)$. In this case, $B^+(z) = 1$ and $B^-(z) = K(z - b)$.
2) Determine the desired I/O model transfer function $G_m(z)$, i.e.

$$G_m(z)=G_{u_cy}(z)=\frac{B(z)\cdot T(z)}{A(z)\cdot R(z)+B(z)\cdot S(z)}=\frac{B_m(z)}{A_m(z)} \tag{23}$$

where $A_m(z) = A_{CL}(z)$ is given by Eq. (22), and $B_m(z)$ is chosen as

$$B_m(z)=\frac{(z-b)\cdot(1+a_{C1}+a_{C2}+a_{C3}+a_{C4})}{1-b} \tag{24}$$

such that the gain of the closed-loop system is 1.
3) Solve the Diophantine equation

$$\begin{aligned}&A(z)(z-1)(z+r_1)+K(z-b)(s_0z^2+s_1z+s_2)\\&=z^4+a_{C1}\cdot z^3+a_{C2}\cdot z^2+a_{C3}\cdot z+a_{C4}\end{aligned} \tag{25}$$

4) Compute $T(z)$, using the following equation

$$T(z)=B_m(z)=\frac{(z-b)\cdot(1+a_{C1}+a_{C2}+a_{C3}+a_{C4})}{1-b} \tag{26}$$

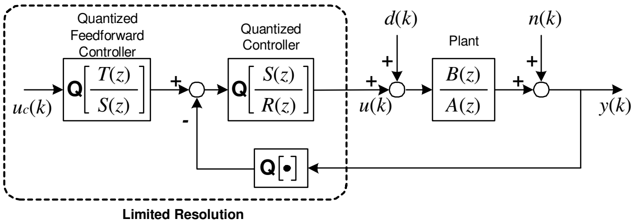

Figure 9. Block diagram of a quantized control system.

TABLE III.
CHARACTERISTICS OF THE EXPERIMENTAL SYSTEM

	$S(z)$	$R(z)$	$T(z)$
Original roots	0.9631±0.0225j	1, 0.63923	0.7
Quantized roots (16-bit)	0.9631±0.0230j	1, 0.6392	0.7
Quantized roots (9-bit)	0.9629±0.024j	1, 0.6387	0.6992
Quantized roots (8-bit)	1, 0.9258	1, 0.6367	0.6992

By picking a desirable closed-loop transfer function $G_m(z)$, a set of controllers can be synthesized:

$$\frac{T(z)}{S(z)} = 0.00622502 \cdot \frac{z-0.7}{z^2-1.9262z+0.9281} \quad (27)$$

$$\frac{S(z)}{R(z)} = 75.6211355 \cdot \frac{z^2-1.9262z+0.9281}{z^2-1.63923z+0.63923}$$

C. Estimation of Required Wordlength

The controllers in the previous section were designed without the consideration of the effect of quantization. Fig. 9 shows the modified block diagram of the control system, where **Q**[·] denote the quantization and round-off operation. To ensure a stable and minimal phase controller as desired, all of the poles and the zeros of the two controllers need to be kept within the unit circle. From Eq. (12), the required bits for $R(z)$, $S(z)$ and $T(z)$ are 2, 9, and 1, respectively. By taking the maximum value of the individual required bits, the required wordlength B_r is

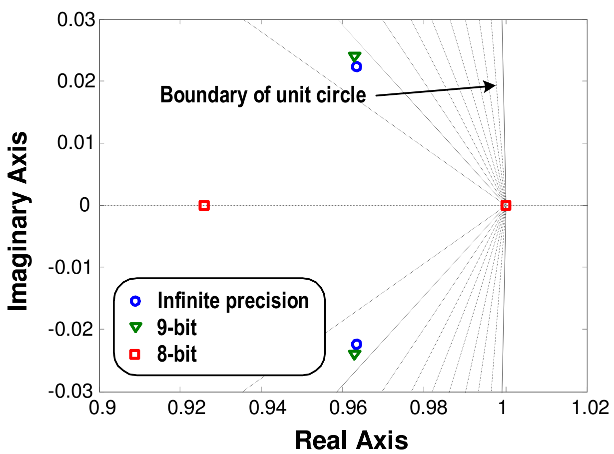

Figure 10. Pole movement of $S(z)$ due to quantization.

9 bits. If the resolution of the processor is less than 9 bits, then the quantized controllers might be unstable. Table III shows the deviation of the roots of the polynomials due to the quantization process, and Fig. 10 demonstrates the pole movements of $S(z)$ due to quantization. It is clear that the 8 bit implementation will have very different characteristics as the original design. After quantization, the poles of $S(z)$ are shifted onto the real axis, and one of them acts as an integrator.

D. Experimental Results

Two experiment scenarios, sufficient resolution and insufficient resolution, were investigated to verify the effect of the coefficient quantization. The controllers were first implemented with a resolution higher than the required wordlength. In this case, the controllers were implemented in 9 bits and 16 bits, respectively. Fig. 11 shows the experimental step response with sufficient resolution, where the reference position changed from 0 to 20.32 mm. The control commands of the simulation and experimental results are shown in Fig. 12. If the

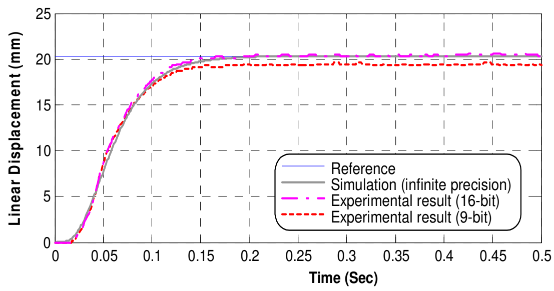

Figure 11. Experimental results for a single paper advance with different resolutions.

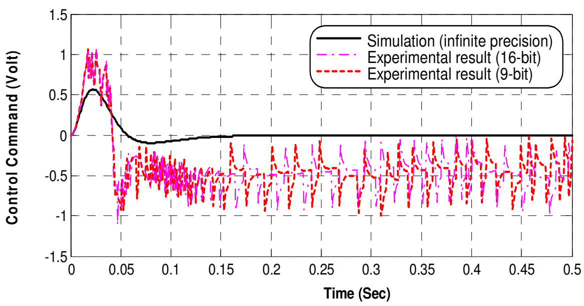

Figure 12. Control commands for a single paper advance with different resolutions.

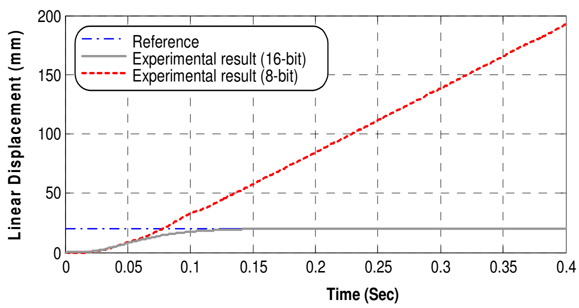

Figure 13. Linear displacement for a single paper advance with insufficient resolution

controller is implemented with a resolution higher than the required wordlength, the outputs closely match that of the infinite precision output, where is approximated better by the high resolution (16 bits).

To demonstrate the impact of insufficient wordlength, the controller is implemented with a resolution of 8-bit, which is one bit less than estimated. Instead of placing the pole in the designed location, the quantization process moved one of the complex conjugate poles to (1, *j*0), which is equivalent of adding an integrator in the controller, and the closed loop system becomes unstable with two integrators, see Fig. 13. The quantized controller can no longer achieve the desired response.

IV. CONCLUSION

In this paper, we presented a framework to determine the necessary resolution for a discrete-time controller with the sample rate specified. Thus, with this method, one can determine the operating speed and the required resolution simultaneously before actually implement the controller physically. A suitable processor can then be selected accordingly to leverage the cost and the applicable hardware. However, this framework can only passively predict the required property to find suitable processors. In many cases, the engineers would need to design their controllers based on the resource in hand. Hence, a more general framework to design a controller based on the given wordlength will be investigated in the future.

REFERENCES

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Tech. J.*, vol. 27, pp. 379-423 and 623-656, July and October, 1948.

[2] B. Liu, "Effect on Finite Word Length of Digital Filters - a Review," *IEEE Trans. on Circuit Theory*, Vol. CT-18, no. 6, pp. 670-677, November 1971.

[3] C. L. Philips and H. T. Nagle, *Digital Control System Analysis and Design*. 3rd ed., Englewood Cliffs, NJ: Prentice-Hall, 1994.

[4] D. Williamson, "Finite Word Length Design of Digital Kalman Filters for State Estimation," *IEEE Trans. on Automatic Control*, vol. 30, pp. 30-39, September 1985.

[5] P. Moroney, A. S. Willsky, and P. K. Houpt, "The Digital Implementation of Control Controllers: The Coefficient Wordlength Issue," *IEEE Trans. on Automatic Control*, vol. 25, pp. 621-630, August 1980.

[6] H.-M. Cheng and G. T.-C. Chiu, "Finite Precision Controller Implementation - Limitation on Sample Rate," in *Proc. of 2003 IEEE/ASME International Conf. on Advanced Intelligent Mechatronics*, Kobe, Japan, July 2003, pp. 634-639.

[7] H.-M. Cheng and G. T.-C. Chiu, "Finite Precision Controller Implementation - Explore the Coupling between Sample Rate and Wordlength," in *Proc. of 3rd IFAC Symposium on Mechatronics*, Sydney, Australia, September 2004.

[8] R. C. Agarwal, and C. S. Burrus, "New Recursive Digital Filter Structures Having Very Low Sensitivity and Roundoff Noise," *IEEE Trans. on Circuits and Systems*, vol. CAS-22, no. 12, pp. 921-927, December 1975.

[9] J. Fialho and T. T. Georgiou, "On Stability and Performance of Sampled-Data System Subject to Wordlength Constraint," *IEEE Trans. of Automatic Control*, vol. 39, pp. 2476-2481, December 1994.

[10] G. F. Franklin, J. D. Powell, M. L. Workman, and D. Powell, *Digital Control of Dynamic Systems*. Prentice-Hall. Englewood Cliffs, NJ. 1997.

[11] V. Oppenheim and R. W. Shafer, *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ. 1989.

[12] M. Rader and B. Gold, "Digital Filter Design Technique in the Frequency Domain," *Proc. of IEEE*, vol. 55, pp. 149-171, February 1967.

[13] P. E. Mantey, "Eigenvalue Sensitivity and State Variable Selection," *IEEE Trans. on Automatic Control*, vol. 13, no. 3, pp. 263-269, December 1968.

[14] R. H. Midleton and G. C. Goodwin, "Improved Finite Word Length Characteristics in Digital Control Using Delta Operators," *IEEE Trans. on Automatic Control*, vol. 31, no. 11, pp. 1015-1021, November 1986.

[15] R. H. Midleton and G. C. Goodwin, *Digital Estimation and Control - A Unified Approach*. Prentice-Hall, Englewood Cliffs, NJ. 1990.

[16] D. Williamson and K. Kadiman, "Optimal Finite Wordlength Linear Quadratic Regulation," *IEEE Trans. on Automatic Control*, vol. 34, no. 12, pp. 1218-1228, December 1989.

[17] K. Liu, R. E. Skelton, and K. Grigoriadis, "Optimal Controllers for Finite Wordlength Implementation," *IEEE Trans. on Automatic Control*, vol. 37, pp. 1294-1304, September 1992.

[18] T. Song, E. G. Collins, Jr., and R. H. Istepanian, "Improved Closed-Loop Stability for Fixed-Point Controller Realizations Using the Delta Operator," *Intl. J. of Robust and Nonlinear Control*, vol. 11, no. 1, pp. 41-57, January 2001.

[19] C.Y. Chen, C.C. Cheng, and G. T. -C. Chiu, "Adaptive Robust Control of Media Advance Systems for Thermal InkJet Printers," *Intl. J. of Mechatronics*, vol. 10, iss. 1-2, pp. 111 – 126, 2000.

Hung-Ming Cheng received his Ph.D. in Mechanical Engineering from Purdue University, USA in December, 2005. He got his B.S. and M.S. degree in Mechanical Engineering from National Sun Yat-sen University, Taiwan in 1994 and 1996. Currently, he is an Assistant Professor of Mechanical Engineering Technology at Georgia Southern University. His research and teaching interests include: controller synthesis with the consideration of finite wordlength, precision motion control, fast imaging of atomic force microscope, microprocessors and embedded controllers, data acquisition units and PLC hardware. He is a member of the IEEE and a member of the ASME.

George T.-C. Chiu received the B.S.M.E. degree from National Taiwan University, Taipei, Taiwan, R.O.C., and the M.S. and Ph.D. degrees from the University of California, Berkeley, in 1985, 1990, and 1994, respectively. From 1994 to 1996, he was an R&D Engineer with Hewlett-Packard Company, developing high-performance color inkjet printers and multifunction machines. Since 1996, he has been an Assistant Professor, an Associate Professor and a Feddersen Fellow in the School of Mechanical Engineering, Purdue University, West Lafayette, IN. His current research interests are design and control of imaging and printing systems, signal and image processing, adaptive and optimal control, integrated design and control of electromechanical systems, and mechatronics.