Error Compensation Using Inverse Actuator Dynamics

Witoon Panusittikorn, Kenneth P. Garrard, Thomas A. Dow
North Carolina State University, Precision Engineering Center
Campus Box 7918, Raleigh, North Carolina USA

A signal processing technique has been developed to dramatically increase both the usable bandwidth and tracking accuracy of an actuator. Traditional feedback controller design makes little or no use of the information content of the driving command signal. Feedforward and look-ahead schemes use a small sliding window of trajectory data to improve tracking performance and to reduce the overshoot of an accelerated axis. If the entire command signal for each axis of a machine is generated off-line, and the dynamic response of each axis is well characterized, then a new command signal can be derived which will drive the axes to follow the correct motion path. This idea is related to reference feedforward control, but by considering the complete command signal *a priori,* performance of the overall machine tool is greatly enhanced.

An inverse dynamics algorithm has been developed that modifies a motion path using an actuator's impulse response to produce a new command signal that counteracts actuator dynamics, yielding a significant reduction in steady-state tracking error. This technique critically depends on knowledge of the actuator dynamics and does not replace the feedback controller. Rather it improves the response of an actuator by compensating for high frequency dynamics that the servo feedback loop cannot correct due to saturation of the drive system.

Background A diamond turning lathe is capable of producing high quality surfaces in a variety of nonferrous materials. However, its use is normally limited to surfaces of revolution that are realized by moving the tool and/or workpiece through one-half of a cross section of the desired shape while the rotation of the workpiece on a spindle causes a symmetric surface to be machined. Advantages of turning are reduced process times as compared to milling due to the continuous high cutting speed achieved via rotation of the workpiece, excellent form fidelity and optical quality surface finish.

It is desirable to produce optical systems with the diamond turning process that have high spatial frequency features or are non-rotationally symmetric (NRS). The addition of a high-bandwidth fast tool servo (FTS) is one technique that has been used to add low amplitude, sub-revolution features to an asphere or sphere. Examples include torics and segmented off-axis conics. Two major complications arise when using a FTS to produce NRS surfaces: static decomposition and dynamic convolution.

Static Decomposition First, the motion commands for each axis of the base machine and the FTS must be generated from a 3D description of the desired part. In some cases the location of the part with respect to the spindle axis and the orientation of the FTS are also considered. The part shape must be decomposed into a rotationally symmetric surface of revolution (i.e., a cross section) for the diamond turning machine (DTM) and a spindle angle dependent motion path for the FTS. The principle difficulty is that while the FTS has a high relative bandwidth it has a limited range of motion. This decomposition process has been solved for a variety of shapes and will not be addressed in this paper.

Dynamic Convolution The second complication is the dynamic behavior of the axes, particularly the FTS. The heterogeneous nature of the axes results in phase errors in the recomposition of the decomposed command signals. The phase errors of the base machine axes can usually be ignored as the axis accelerations are moderate and velocity tracking performance of the control system is quite good. However even a small lag time associated with a FTS will result in poor form fidelity for off-axis surfaces and improper placement of NRS surface features with respect to the rotationally symmetric (RS) base surface and any fiducial. The decomposed FTS motion changes direction at least once on each rotation of the spindle whereas the axis motions are almost always in one direction (the waxicon being a notable exception). A first order correction of the FTS phase error is to apply a constant lead time to its motion commands. The magnitude of this phase advance depends only on actuator bandwidth and spindle velocity. However for motion paths that use a substantial portion of the bandwidth range or contain frequency components close to the bandwidth of the axes this simple technique yields poor results.

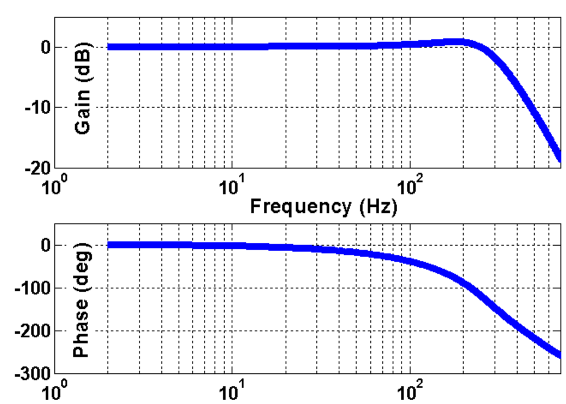
The response of a dynamic system to an applied signal is a function of the frequency of that signal. In general, the response results in an attenuated and delayed motion of the output with respect to the input signal. When the system is a fast tool servo and the dynamic input signal is made up of a variety of frequencies, the result is form error in the machined surface. The actuator mechanism can be considered as a spring-mass-damper driven by a periodic function $x(t)$ with an amplitude A and a frequency of ω radians per second. The steady-state output motion $y(t)$ is

$$\textbf{Input:}\quad x(t)=\sum A_i \sin(\omega_i t)$$
$$\textbf{Output:}\quad y(t)=\sum a_i A_i \sin(\omega_i t+\phi_i)$$
$$\textbf{Modified Input:}\quad x(t)=\sum \frac{A_i}{a_i}\sin(\omega_i t-\phi_i)$$
$$\textbf{Modified Output:}\quad y(t)=\sum a_i \frac{A_i}{a_i}\sin(\omega_i t+\phi_i-\phi_i)$$

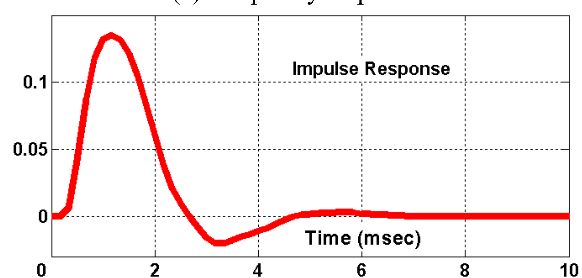
attenuated by a gain factor a and is phase shifted by an angle ϕ.

If the input $x(t)$ contains multiple frequencies, the corresponding output can be determined using *superposition*; that is, the input signal can be decomposed into single-frequency components each of which will be attenuated and delayed. The response of the system $y(t)$ will be the sum of these frequency components. To machine a desired surface profile with high fidelity, the amplitude attenuation and phase related delay must be eliminated. To achieve this, the input signals are modified prior to applying them to the system such that the attenuation is canceled and the phase is compensated.

After the dynamics of the system are identified and the desired motion is decomposed into sinusoids, amplitude (a_i) and phase (ϕ_i) adjustments can be found to construct a modified command signal that moves the actuator in the desired manner. A technique for system identification and an input signal modification algorithm have been developed and applied to the Variform™ FTS to produce the desired output response for high frequency inputs.



(a) Frequency response



(b) Impulse response

Figure 1. Dynamic response of the Variform

Actuator Dynamic Response The dynamic response of the Variform can be presented in either the time domain (impulse response) or the frequency domain (frequency response). They are both shown in Figure 1 and are mathematically equivalent and complete for a linear system. On the left in Figure 1(a), the top graph shows the ratio of the amplitude of the output to the input (in dB units) as a function of frequency. At low frequency, the output is equal to the input and the ratio is unity giving a dB value of zero. As the frequency is increased, the amplitude response peaks at about 200 Hz and then drops rapidly for higher frequencies. The lower graph in Figure 1(a) shows the phase angle between the input and the output. At lower frequencies these signals are nominally in phase, but even a small lag of less than a degree represents a significant length of arc if the actuator is at a large radius from the axis of spindle rotation. As the frequency rises the output increasingly lags the input. At 100 Hz this lag is about 45º and at 340 Hz it is about 180º. The impulse response is shown on in Figure 1(b) and describes the characteristics of the system with respect to time. The Fourier transform of the impulse response is exactly equal to the frequency response and the inverse Fourier transform of the frequency response is the impulse response.

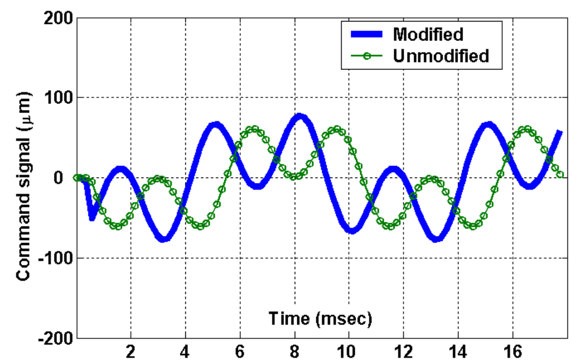
Deconvolution The effect of a dynamic system on an input signal of length N can be determined by applying the convolution operation to the input $x[n]$ and the impulse response $h[n]$ of the system. The convolution operation describes how a linear filter (an actuator) modifies a given input signal (command) to produce an output signal (motion). This process is represented by the * (star) operator and is expressed in Equation (1). Each value in the convolved output signal can be calculated with the summation in Equation (2), where M is the length of the impulse response. The length of $y[n]$ is $N+M-1$. Convolution is mathematically equivalent to polynomial multiplication and can be directly calculated from Equation (2) in the time domain. If the discretely sampled input and impulse signals are transformed to the frequency domain, convolution becomes a straightforward complex multiplication (element by element) of two vectors. The discrete Fourier transform (DFT) is commonly implemented by the fast Fourier transform (FFT) algorithm which exploits the symmetries in Equation (2) to accelerate this calculation. Equation (2) can thus be restated in the frequency domain using an FFT operator as shown in Equation (3). The inverse FFT operation can then be used to find $y[n]$.

$$x[n]*h[n]=y[n] \tag{1}$$

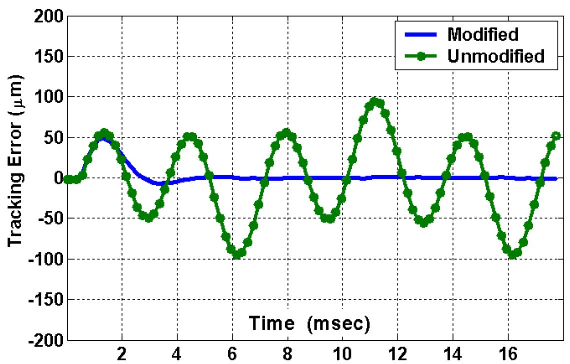
$$y[k]=\sum_{m=0}^{M-1} h[m]\,x[k-m] \tag{2}$$

$$\mathbf{FFT}(x[n])\times\mathbf{FFT}(h[n])=\mathbf{FFT}(y[n]) \tag{3}$$

$$\mathbf{FFT}(x[n])=\frac{\mathbf{FFT}(y_d[n])}{\mathbf{FFT}(h[n])} \tag{4}$$



(a) Command signal



(b) Actuator tracking error

Figure 2. Deconvolution experiment for a high-frequency Variform command signal

Deconvolution is the inverse of convolution and can thus be used to find $x[n]$, given an actuator impulse response $h[n]$ and a desired motion $y_d[n]$. Rearranging Equation (3) gives Equation (4). Application of the inverse FFT operation then yields a signal that should exactly counteract the actuator dynamics so that it follows the desired trajectory, eliminating steady-state following error. This technique critically depends on knowledge of the actuator dynamics as embodied in the impulse response and requires the complete desired motion profile.

Experimental Verification A Variform FTS with 400 µm range and 350 Hz bandwidth has been used to perform experiments to validate the algorithm. The response of this actuator is predominately second order although nonlinearities are present for large amplitude command signals. Figure 2 shows the path error associated with a 160 µm command signal that is the sum of 100 Hz and 300 Hz sine waves. The response of the actuator to the 300 Hz sinusoid is 140 degrees out of phase and attenuated 2.5 dB resulting in a very large (200 µm PV) error. The algorithm uses deconvolution to calculate a new command signal that effectively precompensates for these effects individually at each constituent frequency. The following (or tracking) error is reduced by about 3 orders of magnitude. The response shows virtually no delay from the desired excursion after a startup interval lasting approximately 4 ms. In practice, this startup period will occur before machining begins and thus will not affect the fidelity of a surface.

Saturation In general, the dynamic response of a linear actuator is a function of the frequencies of the applied signals. Though the dynamic response results in an attenuated and delayed motion, the output of a linear system contains only the frequencies of the applied signals. Since the Variform FTS contains hardware limitations in which the velocity is constrained, the output profile is reshaped. The effects of this distortion on the operating range of the Variform as well as the measured impulse response employed in the deconvolution technique can be thoroughly investigated using frequency spectrum analysis. At frequencies near actuator bandwidth even a saturated input signal will not produce a full magnitude excursion. This situation is entirely predictable for a given command signal and is easily incorporated into the algorithm.

In Figure 3, the deformed shape of the FTS response to a high frequency input is illustrated. The distorted response comprises not only the frequency components of the desired trajectory (200 Hz), but also the 3rd and the 5th harmonics, 600 and 1000 Hz, respectively. The constituent sinusoid associated with the frequency of the desired path is in-phase and the 3rd harmonic is out-of-phase, transforming the output to a triangle-like profile. Note that the amplitude of this constituent sinusoid is smaller than the amplitude of the distorted output excursion.

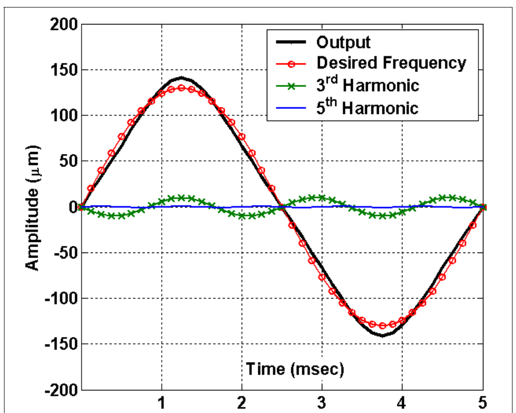


Figure 3. Actuator saturation

The amplitude difference has a direct effect on the accuracy of the impulse response measurement using a spectrum analyzer. The analyzer generates a varying frequency input (i.e., swept sine wave) with fixed amplitude to the FTS, and analyzes the attenuation and phase of the actuator motion at each input frequency as measured by an LVDT. If the output signal is velocity saturated, its frequency component associated with the input frequency has smaller amplitude than the actual output signal. As a result, the frequency response of the Variform shows severe attenuation. Figure 4 depicts the attenuation of the Variform as a function of frequency and amplitude. The dynamic response shows significant attenuation where the input signal contains large amplitudes with high frequencies. This is because the maximum velocity of this FTS is limited to 140 mm/sec.

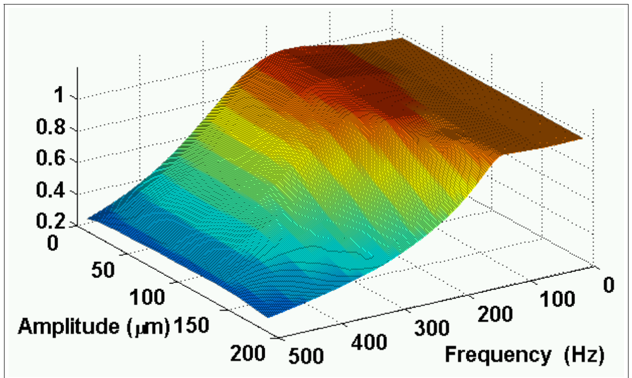


Figure 4. Expanded frequency response

Real Time Deconvolution The desired path of the FTS can be defined as a function of the spindle speed and the axis cross-feed rate of the DTM. As either velocity changes, the time period of the desired path is also altered. Since deconvolution over the entire command signal is only valid for immutable machining parameters, a more robust approach would be to perform the deconvolution in *real time* over a moving window into the desired trajectory. As shown in Figure 1b, the impulse response of the FTS is essentially zero after 10 ms. The modified actuator command signal at any given instant depends only on a window of the command signal that extends forwards and backwards in time over a 20 ms interval. A more advanced deconvolution algorithm has been developed to perform the deconvolution in *real time*; albeit at a sub-servo rate.

A long tool path must be divided into short pieces that are individually modified using the deconvolution technique. While the Variform follows the current input command, the next piece (or window) of the desired path is modified using the most recent parameters. This *real time* deconvolution can be achieved by applying a window function[1] to a long signal. In general, the window is short in length for sensitivity to high frequency signals.

Real time deconvolution transforms these short windowed pieces to the frequency domain where the operations of inverse dynamics occur. The transform using window functions is called the short-time Fourier transform (STFT) [1]. The reverse process, an overlap-add method is used to reconnect these short pieces after deconvolution. Reconstruction of short overlapped adjacent windows is illustrated in Figure 5. Note that the reconstructed signal is distorted at both ends as a result of the window shape and the overlap-add method, but is an exact copy of the unwindowed signal where the overlap is complete (i.e., between 8 and 32 time units).

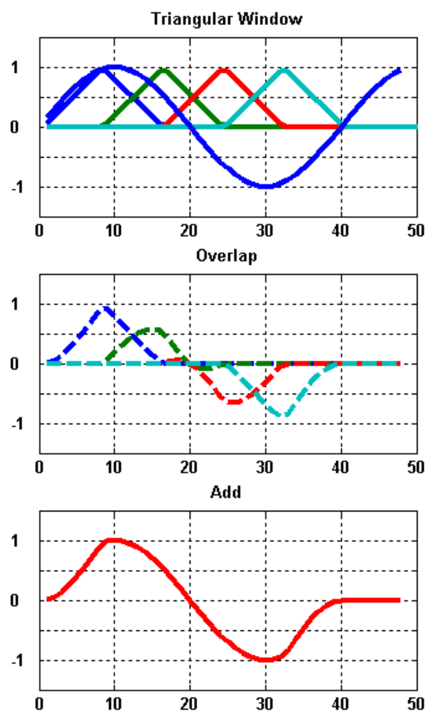


Figure 5. Overlap-add reconstruction

Conclusions Figure 6 shows the result of a *real time* overlap-add deconvolution experiment with the Variform. After a startup interval that is related to the window size and amount of window overlap, the tracking error is reduced to nearly zero. Machining experiments will begin in the near future to obtain practical validation of both deconvolution techniques. A small, concave, off-axis sphere machined into an otherwise flat surface has been selected for these tests. This surface was chosen because the command trajectory is aperiodic and its frequency content can be changed by modifying the off-axis distance and/or spindle velocity. Most important though is that the *f-number* of the spherical surface can be selected so that the results are easily characterized with an interferometer.

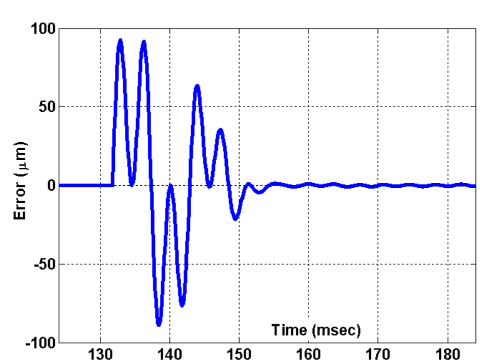


Figure 6. Real time deconvolution tracking error for a 90 μm amplitude command with 90 Hz and 270 Hz sinusoidal components

Reference
[1] Quatieri, T.F. *Discrete-Time Speech Signal Processing*. Prentice-Hall (2001).

[1] A window function is a vector of weighting coefficients which are zero everywhere outside of a certain interval.