7-1-2012

Vision-based tracking of a dynamic target with application to multi-axis position control

Carlos A. Montes
Clemson University

Chan Wong
Clemson University

John C. Ziegert
University of North Carolina

Laine Mears
Clemson University, mears@clemson.edu

CONDITIONS AND LIMITATIONS OF THE LOW-BANDWIDTH VISUAL-SERVO LOOP IN MICRO-POSITIONING APPLICATIONS

Carlos A. Montes[1], John C. Ziegert[2], Chan Wong[1] and Laine Mears[1]
[1]Department of Mechanical Engineering
Clemson University
Clemson, SC, USA
[2]Department of Mechanical Engineering and Engineering Science
University of North Carolina at Charlotte
Charlotte, NC, USA

INTRODUCTION

The present research studies the dynamics of the *single-CAM single-image processing thread* visual servo loop in micro-positioning applications. The study is conducted on a previously introduced control system utilized for the positioning of an XY table. Analysis and experimental results expose fundamental limitations in low-bandwidth vision-based servo-mechanisms when applied to high-resolution position control. Although sensor resolutions on the order of 2.5 μm are reliably achieved through adequate image processing, the overall best achievable resolution for in-plane closed-loop positioning of the stage is found to be on the order of 100's of micrometers. This aspect is attributed to errors in the approximation of the feedback signal; such approximation is required to compensate for intermittencies and delays associated with the slow dynamics of the imaging sensor and the image processing thread. Though this approximation would certainly be considered acceptable for macro-positioning applications, micro-positioning systems might require additional elements to aid in achieving the desired resolution levels. The conditions for achieving smooth output motion in vision-based control systems when integrated in manufacturing equipment are also presented.

EXPERIMENTAL SETUP

Implementation is carried out on an XY-stage as shown in Figure 1. The desired displacement command of the stage in a planar environment is emulated, in one end of the kinematic chain, by a cross-hairs target on a liquid-crystal display (LCD) [2]. On the other end of the kinematic chain a digital camera observes the active target and provides visual feedback information utilized for position control of the tool. This system was first introduced in [1], and further investigated in [2].

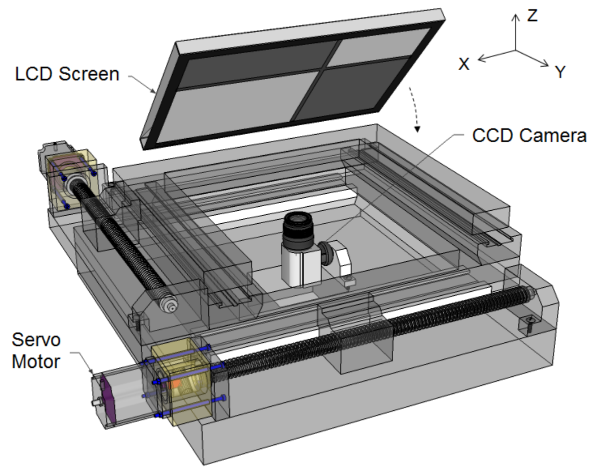

FIGURE 1. 3D CAD model of XY-stage.

The analysis presented next concerns a *single CAM- single image processing thread*-configuration, where a single camera and a single graphic processing unit (GPU) are used to acquire and process the target images.

THE HYBRID COMMAND ISSUING PROTOCOL

In the hybrid command generation protocol, commands equal to a multiple of one LCD pixel are issued by moving the target on the LCD by that specific distance, and setting the reference on the image plane to the origin (i.e. the camera's principal point (PP)). In this case the control system operates as a regular tracking system. For commands smaller than one LCD pixel or not equal to an integer multiple of pixels, the target is displaced by the smallest integer number of pixels that is larger than the desired commanded displacement and the reference signal is set to a value different than PP; the value of the reference is less than 1 LCD pixel size. Figure 2 shows the proposed control structure, where the path plan is generated off-line, block *C* represents the controller, *G* represents the dynamics of a single axis, $\hat{G}$ represents the mathematical model of a single axis, V represents the camera- and image processing thread-dynamics and $\hat{V}$ its corresponding model. Variable x_i^{Target} is the displacement of the target referenced to the LCD frame. The reference signal to the control system is defined as $x^{ref} = x^{Target} - x^{fine}$. The value x^{fine} is the desired displacement command, where $|x^{fine}| < 1$ LCD pixel. Proof of convergence of x (position of the stage) to x^{fine} when setting $x^{ref} = x^{Target} - x^{fine}$, as well as other considerations for implementation of this method, are provided in [3].

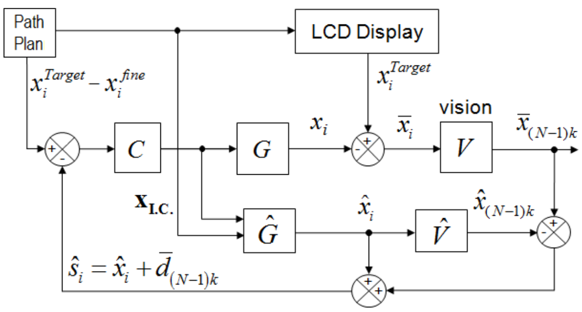

FIGURE 2. Model-aided control system scheme represented in the discrete domain with discrete sampling variable i.

In the proposed hybrid protocol the states of the model $\hat{G}$ are enforced to follow the relative stage position, $\bar{x}_i$, by updating the initial conditions of the model states, $\mathbf{x}_{I.C.}$, every time the position of the target, x_i^{Target}, is displaced on the LCD. This procedure requires the synchronization of the different experimental components, later called processing threads. The conditions for adequate operation of the system in Figure 2 are presented next. Additionally, these conditions are extended to more common visual tracking architectures for position control of NC machines.

CONDITIONS OF OPERATION

Dynamic Target Tracking

The update of the model initial conditions should only occur right after the position of the cross-hairs has changed with respect to the LCD frame. Hence, adequate command issuing can only be achieved if the host thread, in charge of generating and displaying the dynamic target, and the control thread, where the model-based closed-loop system is run, are synchronized (Figure 3).

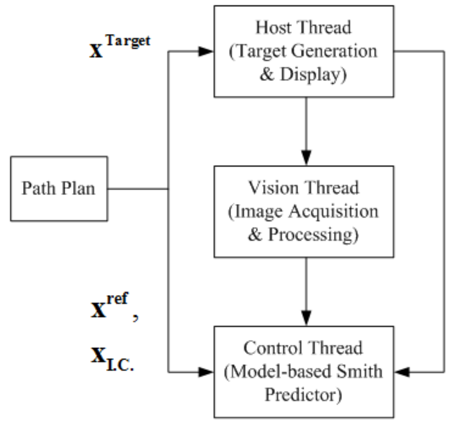

FIGURE 3. Multi-thread communication diagram.

Once a new target imaged has been displayed, the host thread must notify the control thread so that the latter updates both the initial conditions of the model and the reference signal. The time diagram in Figure 4 shows how the three processing threads operate in a synchronized matter. Communication times between threads are considered to be small, with respect to other processing times, and are assumed to be negligible in the analysis. At t_0 the host thread displays a new target image, x^{Target}. Once the image has been updated on the LCD, the vision thread starts the acquisition process at t_0^D, and $\mathbf{x}_{I.C.}$ and x^{ref} are updated at t_0^{update} in the control thread (with $t_0^D = t_0^{update}$). At this point the stage starts moving according to the initial conditions of the model and the control effort tries to bring the model states to the reference position. If the resulting velocity from the control effort is too high the tracking error will be reduced to zero rapidly. In Figure 4 it takes the control system Δt_0^{fine} seconds to bring the tracking error to zero, according to the model feedback signal. The stage then waits motionless for Δt_0^{wait} seconds until a new target image is displayed at t_1. This wait time is actually a drawback as it introduces discontinuities in the axis output velocity.

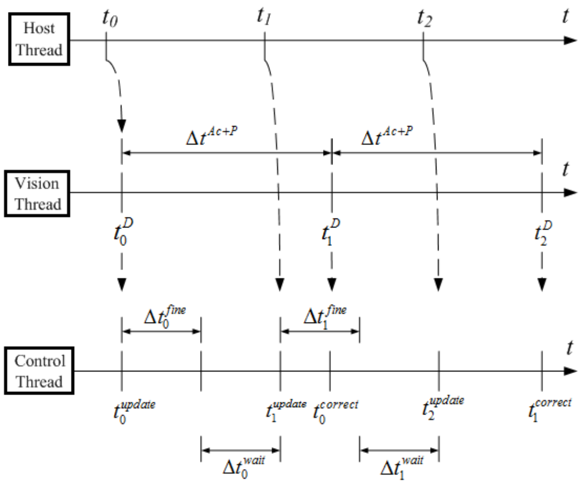

FIGURE 4. Multi-thread time diagram.

Hence, there exists a need to either reduce the output velocity, which would in turn increase Δt_0^{fine} and reduce Δt_0^{wait}, or increase the rate at which the target image is displaced on the LCD, or both. If, for a given desired trajectory, the frequency at which the target image is displaced on the LCD is too high, the control system would operate based mainly on the synthetic feedback signal over the entire path, not taking full advantage of the direct sensing capabilities of the vision sensor. On the other hand, low target update frequencies would cause long wait times making the output motion intermittent and yielding discontinuities in the velocity curve. Ideally, the position of the target on the LCD should be updated at a similar rate as the vision block update frequency. In Figure 4 the time required by the vision block to acquire and process a target image is assumed to be constant and has a value of Δt^{Ac+P} seconds; hence, the vision block update frequency is approximately $1/\Delta t^{Ac+P}$ Hz. In such case, when tracking the target over a predefined path plan it is desired to enforce

$$\frac{f_{T.P.}}{f_{V.B.}} \approx 1 \quad (1)$$

where $f_{T.P.}$ is the frequency at which the target position is updated on the LCD, and $f_{V.B.} = 1/\Delta t^{Ac+P}$. In practice, $f_{V.B.}$ is not constant and an average can be calculated from experimental data and used for this purpose. In addition to condition (1), the velocity of the output motion must also be regulated to increase the value of Δt_0^{fine}, in Figure 4, and maintain continuous motion, i.e. avoid discontinuities in the output velocity. If these two conditions are enforced, the output motion and output velocity are expected to be smooth curves over a predefined path plan.

Extension of Conditions to Regular Vision-based Control in NC Machines

Conditions for adequate operation of the proposed visual-servo loop can easily be extended to regular visual servo control. Figure 5 shows a typical visual servo loop utilized for position control of an NC machine. The target being imaged is no longer dynamic (cross-hairs target on LCD), but is a real world object; i.e. the machine tool or the robot's end effector. If the frequency of the required motion commands signal, $f_{R.C.}$, is considered before interpolation the conditions for adequate operation can be redefined as

1. Frequency ratio: $f_{R.C.}/f_{V.B.} \approx 1$
2. Velocity of the output motion must be regulated to avoid intermittent output motion. This can be achieved through an increase in the proportional gain of the PID controller or through feedforward velocity control.

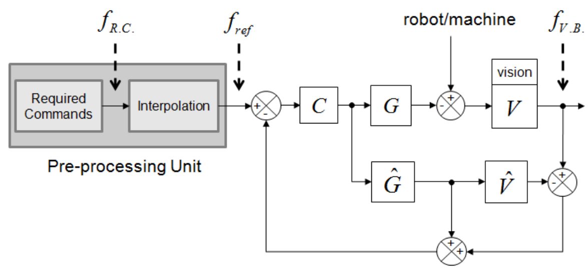

FIGURE 5. Typical visual servo configuration used in robot manipulators an applied here to position control of an NC machine.

The diagram in Figure 5 could also be thought of as a coarse/fine command issuing scheme. Condition 1 guarantees that the update frequency of the coarse command signal , $f_{R.C.}$, is close to the vision block update frequency, while the interpolated points, updated at a slightly higher rate, f_{ref}, are processed by the control loop according to the synthetic feedback signal yielding continuous motion in between camera updates.

DYNAMICS OF THE VISUAL SERVO LOOP

In the absence of delays and in an ideal single rate system the feedback signal, S, would be $s_i = \hat{x}_i + (\bar{x}_i - \hat{x}_i) = \bar{x}_i$. However, the actual position of the stage obtained from the camera is available only every *k* samples (where *k* is a constant), and provides information about the plant *k* samples late. Then for instance, at *i* = *k* the estimated feedback signal, $\hat{S}$ in Figure 2, is $\hat{s}_k = \hat{x}_k + \left(\bar{x}_0 - \hat{x}_0\right)$. The approximate correcting factor $\left(\bar{x}_0 - \hat{x}_0\right)$ can be applied during the entire interval $i \in \left[k, 2k\right)$ until a new feedback point is generated by the vision block at $i = 2k$. Let the correcting factor be defined in the discrete domain as $\bar{d}_i = \bar{x}_i - \hat{x}_i$. Then in general $\hat{s}_i = \hat{x}_i + \bar{d}_{(N-1)k}$, where $i \in \left[Nk, (N+1)k\right)$ and $N \in {}_0$. The error between the ideal and actual feedback signals is defined as

$$\begin{aligned} &\bar{e}_i = s_i - \hat{s}_i \\ &\Rightarrow \bar{e}_i = \bar{x}_i - \left(\hat{x}_i + \bar{d}_{(N-1)k}\right) \\ &\Rightarrow \bar{e}_i = \bar{d}_i - \bar{d}_{(N-1)k} \end{aligned} \quad (2)$$

Discrepancies between the ideal and actual feedback signals in model-aided vision-based control, given by (4), are responsible for inaccuracies in the output displacement of the tool. It is important to stress that even if the vision block was capable of operating at the controller's frequency ($k = 1$), i.e. without intermittencies, the error between the ideal and actual feedback signals would still be different from zero due to the processing delay. In such case, $\bar{e}_i = \bar{d}_i - \bar{d}_{(N-1)}$ for $i \in \left[N, N+1\right)$ and $N \in {}_0$. Error $\bar{e}_i$ would only be zero if no intermittencies and no delays were present in the vision block feedback signal; in such case, however, no modeling would be required. This is a fundamental limitation associated with delays and intermittencies in *single CAM- single processing thread*-visual servoing, when utilized in high-resolution positioning applications.

EXPERIMENTAL RESULTS

A basic hysteresis-path plan for fine in-plane displacement is generated (Figure 6, top). Figure 6, top, shows the actual stage output motion plotted on top of the desired in-plane path. Figure 6, bottom, shows the desired constant velocity pattern for *X*-axis motion when enforcing the frequency and velocity conditions.

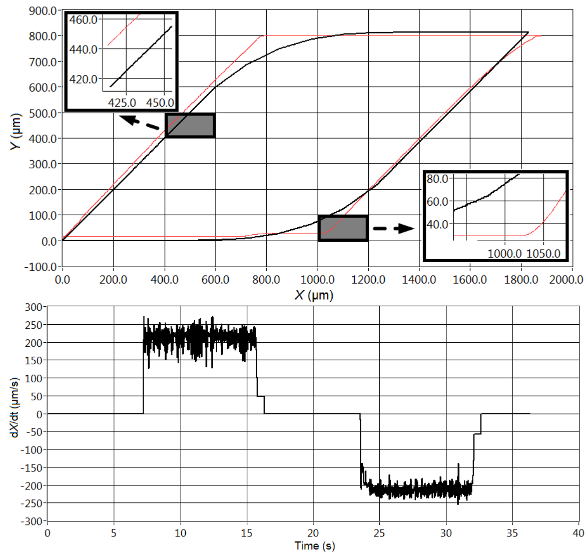

FIGURE 6. Experimental results: (top) Desired and experimental in-plane motion plotted on same graph. (bottom) Output velocity, X-axis.

CONCLUSION

Overall best achievable resolution for in-plane positioning is on the order of 100 μm. For the *single CAM- single image processing thread*-configuration, accuracy errors are associated with the feedback signal, which is approximated through modeling and old sensed data.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 0800507. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Wong, C., Montes, C.A., Mears, L. and Ziegert, J.C. A New Position Feedback Method for Manufacturing Equipment. ASME, Evanston, IL, USA, 2008.

[2] Montes, C.A., Ziegert, J.C., Wong, C. and Mears, L. 2-D Absolute Positioning System for Real Time Control Applications. 25th ASPE Annual Meeting. Atlanta, GA November 5, 2010.

[3] Montes, C.A., Wong, C., Ziegert, J. and Mears, L. Hybrid Command Issuing in a 2-DOF Servomechanism Operated Under Vision-Based Feedback Control. Proceedings of the International Conference on Sustainable Automotive Technologies. Greenville, SC, USA, 2011.