5-2009

Method to Measure Planar Displacement Using Centroid Calculation

Carlos A. Montes
Clemson University

Laine Mears
Clemson University, mears@clemson.edu

John C. Ziegert
Clemson University

METHOD TO MEASURE PLANAR DISPLACEMENT USING CENTROID CALCULATION

Carlos A. Montes
International Center for Automotive Research, Clemson University
Clemson, South Carolina

Laine Mears, Ph.D., P.E.
Assistant Professor,
Mechanical Engineering
International Center for Automotive Research, Clemson University
Clemson, South Carolina

John C. Ziegert, Ph.D., P.E.
Professor & Timken Chair in Design
International Center for Automotive Research, Clemson University
Clemson, South Carolina

KEYWORDS

CNC, Positioning, Vision, Digital, Curve Fitting

ABSTRACT

This paper presents a new sub-pixel resolution approach to measure displacements of a planar motion stage using a vision acquisition sensor which images an actively controlled target generated on a planar display such as an LCD screen. In this paper we examine the resolution with which we can detect a target image defined by the intersection of two curves. The points that generate each curve are identified by using an intensity weighting of the pixels that define the curve. Once the data points are collected, two curves are fit to the data using a least-square approximation. Finally the intersection of the two curves provides a robust and reliable point location that can be tracked for position feedback in manufacturing equipment. Experimental results show that although the display which generates the target image has pixel size greater than 200µm, this procedure can reliably detect stage motions as small as 5µm.

INTRODUCTION

Computer Numerical Control (CNC) equipment is widely used in mass production to enable high-volume manufacturing with high accuracy. One of the determining factors for the resolution and accuracy of such equipment is the feedback sensing system used for axis positioning, e.g. typically linear or rotary encoders. However, in multi-axis motion systems, the feedback devices do not directly sense the position of the control point. Instead, the spatial position of the control point is estimated using the outputs of the position feedback sensors and a kinematic model of the machine. Inevitably, the kinematic model does not exactly describe the actual machine due to imperfect straightness of the axis guideways, non-squareness of their motion directions, and thermal variations with time, resulting in positioning errors. These errors have traditionally been compensated using inverted error mapping applied to axis commands. This is a complex and expensive approach, and suffers from the fact that the error map is static and cannot compensate time-dependent effects. Alternative approaches suggest the use of other sensors, such as the system introduced in [Fan, Wang et al.], which uses a 3D laser ball bar to measure the error of multi-axis machines.

In [Wong, Montes et al.], a new approach to multi-axis position feedback is presented,

whereby a vision feedback system is implemented to directly sense the tool position rather than through the use of the traditional kinematic model. This system utilizes a ground-based camera trained on an active pixel display fixed to the stage being controlled. Target images on the active display are generated and then acquired by the camera, with the difference between the desired and actual target position on the camera image plane used to generate the error vector for drive commands (see Figure 1).

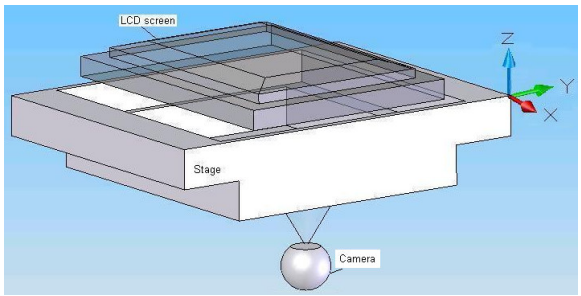

FIGURE 1. 2-AXIS STAGE CONTROL THROUGH FIXED CAMERA ACQUIRING PIXEL ARRAY IMAGE.

In this work, the achievable resolution of the described system is investigated when using a target image consisting of two intersecting curves displayed in pure black/white form, i.e. no grayscale or color modulation of the target image is considered. The camera images this target, and analytic functions of the proper form are best-fit to the pixel data using a least-squares technique. The intersection of these functions is defined as the target location.

BACKGROUND

Manufacturing equipment must perform very small displacement measurements for feedback purposes in position control, typically in the micron and sometimes sub-micron range. In order to perform these measurements using vision acquisition devices imaging a target displayed on a flat panel display, sub-pixel resolution algorithms must be developed, since the typical pixel size of a camera or liquid crystal display screen is one or two orders of magnitude higher than the desired resolution. A brief review of vision-based motion control research is presented.

Control of positioning through camera input has been accomplished in the past to a certain degree of accuracy. In Figure 2, the image capture and thresholding processing techniques capture a hand image, determine its outline, and quantify that information as an input to path planning for a motion system [See A., 2005]. However, only lower accuracy positioning (in the order of 50µm) has been achieved with independent vision servoing.

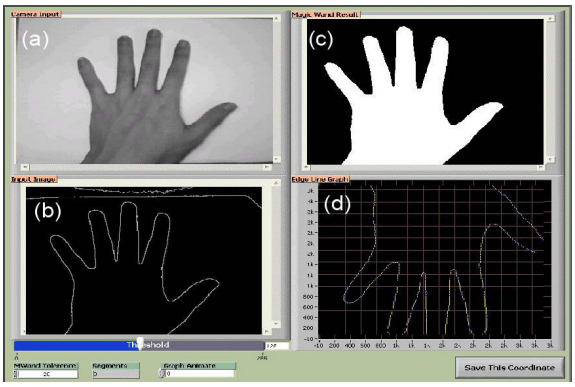

FIGURE 2. VISION-BASED CONTROL. (A)IMAGE ACQUISITION, (B) EDGE DETECTION, (C) PROCESSING, (D) TRAJECTORY GENERATION. [SEE A., 2005]

Vision input has been used as an aid to machine control. Xie proposes a modified Smith Predictor and new control architecture to reduce common time delays in visual-aided control systems (see Figure 3) [Xie, Sun et al.].

The time delay problem is addressed by taking the visual delay out of the control loop. Hence the closed loop system does not suffer from the image acquisition and image processing delays, yet the visual information is still used when available.

Although the control loop operates independently from the rates of the visual component, in most of the operation cycles the feedback signal is generated by the prediction algorithm. This of course possesses an error that is only corrected after the vision process is accomplished. The production costs for this implementation are also affected by the use of the encoders.

These compensation and vision augmentation systems have the potential to be applied to position control for machine tools or other manufacturing equipment. However, vision systems that require only static passive images

(uncontrolled targets) require additional processing time to identify object locations and edges. Vision inputs are characterized as slow and inaccurate, and have heretofore been used only to augment conventional feedback control for precision systems. In this approach, we want to avoid the processing overhead of edge detection and thresholding by leveraging the *a priori* knowledge of the image geometry. Such an approach has the potential to improve processing speed and accuracy to free machine control from complex error compensation schemes associated with kinematic model assumptions. In order to do this, we must gain an understanding of digital image processing for identifying displacement.

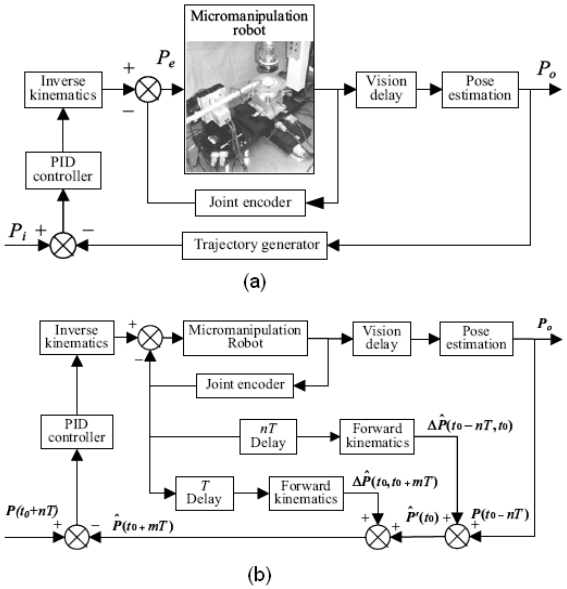

FIGURE 3. SMITH PREDICTOR ARCHITECTURE. (A) SINGLE-FEEDBACK SERVO, (B) PREDICTOR ARCHITECTURE. [XIE, 2005]

Digital Image Correlation

Digital Image Correlation (DIC) uses the measure of image correlation to optimize the parameters of transformation matrices between an acquired image and the known image geometry.

Sutton initially described this approach in his establishment of a method to obtain full-field in-plane deformations of an object [Sutton, Mingqi et al.]. In his description, Sutton highlights the importance of pixel grayscale intensity in image digitization, and a transformation to a continuous representation form.

DIC was successfully applied to large-scale planar deformation for measuring linear displacement using an efficient computation method [Sutton, Mingqi et al.]. However, in direct strain measurement using DIC, the method is found to have large variability in the computational result.

Bruck enhanced the method using partial differential correction via a Newton-Raphson numerical approximation, and improved the ability to accurately measure strain [Bruck, McNeill et al.].

Intensity-Weighting Algorithms

In order to improve the system accuracy beyond the size of a discrete unit (pixel), forms of intensity-weighted distributions and bilinear interpolation have been implemented to discrete systems in order to represent the data in a continuous form.

Sutton performed fundamental work with digital image correlation to represent image data in a continuous form which accounts for individual pixel intensity weighting [Sutton, Wolters et al.]. This intensity-based representation allows for applications of mapping functions to transform the data for optimal correlation with the undeformed data set.

Chen estimates the center of a discrete data set using an adaptive estimation algorithm, which improves the result of previous approaches to identification of the center of intensity of a noisy signal [Chen and Lin]. Such results are directly applicable to the centroid calculation on the image of imperfect screen pixel geometry.

Finally, Fosu et al. determines the location of star images acquired by a discrete Charge-Coupled Device (CCD) microcircuit [Fosu, Hein et al.] using a point spread function and centroid calculation. Such results gained in the field of astronomy for accurately locating the positions of stars are readily applicable to determining the accurate location of manufacturing equipment position.

METHODOLOGY

Given some *a priori* information, the position of a known target is determined and tracked by the motion system over time. The target is represented by the intersection of two curves.

The individual locations of the points, to which the curve functions are fit, are obtained by calculating the intensity weighted center of mass on each row of the digital image. The effects of noise are then minimized by using a least-squares technique to find the best two functions that fit the acquired data. Finally the intersection of the functions is computed and used to calculate the error vector of the motion system.

The procedure for locating the target is summarized in three steps:

Step 1. Coarse point location: the sets S_1 and S_2, containing the points later used to estimate the functions, are identified.

Step 2. Fine location: a more accurate location of the points in S_1 and S_2 is performed, according to grayscale intensity of the pixels. Two new sets are defined: S_{C1} and S_{C2}.

Step 3. Best fit line: using the points in S_{C1} and S_{C2}, the two best fit curves, C_1 and C_2, are determined using a least-squares approximation. The intersection of the curves is calculated analytically using the best fit approximations.

Step 1: Identification of the Sets

The initial step of the target detection process is the identification of the individual points on the image that are later used to calculate the equations of the curves (Step 1). Furthermore, these points have to be classified in two sets associated with the two curves whose intersection is to be determined:

$$S_1 = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$$
$$S_2 = \{(\hat{x}_1, \hat{y}_1), (\hat{x}_2, \hat{y}_2), \ldots, (\hat{x}_n, \hat{y}_n)\} \quad (1)$$

where $m,n \in \mathbb{R}$. Figure 4 (A) shows a typical 8-bit target image, where pixels with intensities closer to 255 (white) are considered to contain relevant information, and low pixel intensities (black) are related to background. Hence, a coarse point location based on intensity peaks can be determined.

The digital image is processed row-wise from left to right, and top to bottom. Two constraints on the target image are established:

1. If two points $p_1 \in S_1$ and $p_2 \in S_2$ are present in the first significant row (first row from top to bottom, containing relevant information) of the image, these points are separated by at least 50 horizontal pixels. The same constraint applies for the second significant row.
2. The intersection of the sets occurs at least 50 pixels below (Y-direction on the image plane) the first significant row, and there is only one intersection point in the image.

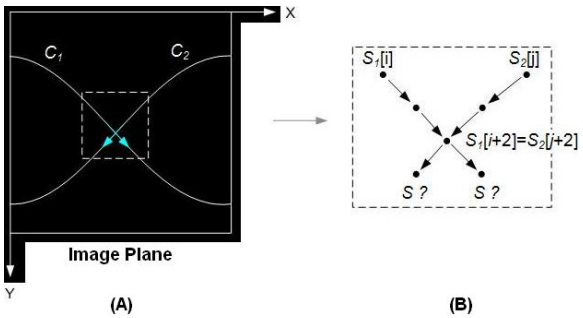

FIGURE 4. (A) INTERSECTION OF SINUSOIDAL CURVES ON IMAGE PLANE, (B) AREA OF INTERSECTION, USING INVERTED COLORS FOR ANALYSIS PURPOSES.

If these constraints are met, the first two elements of each set can be identified by selecting the two highest intensity peaks, on the first and second significant rows, that are separated by at least 50 horizontal pixels. The next points of each set are distinguished according to their intensity, their proximity to the other points of the same set and the direction vector $\overrightarrow{P_{i-1}P_i}$ of the previous 2 points, P_{i-1} and P_i, of the set. The direction vector is calculated for every two consecutive points in a set; then, if a new point is being analyzed to determine if it belongs to the set, the direction vector between itself and the nearest point on the set is calculated and compared with the previous direction vector of the last 2 elements of the same set. If the angle between the two vectors is sufficiently small, these vectors can be thought of as approximations to the tangent vectors at nearby points on a continuous differentiable line, hence the new point is included as part of the set. On the other hand, if

the angle is large, this suggests that the new point probably belongs to a different set and is therefore discarded. The angle between direction vectors is calculated using the inner product of the vectors.

Figure 4 (B) shows an example of the point selection principle. It is clear that at the intersection point both sets share the same point ($S_1[i+2]=S_2[j+2]$). Afterwards, the direction vector between adjacent points is used to determine which point belongs to which set.

Step 2: Fine location

A more accurate measurement of the points' coordinates on the image plane is achieved by calculating the horizontal intensity weighted center of mass, or centroid.

For a continuous 2-dimensional real function, $f(x,y)$, with $x \in [a,b]$, $a,b \in \mathbb{R}$, and y fixed, i.e. $y=j$, $j \in \mathbb{R}$, the horizontal centroid (along the X-axis), denoted by $X_{C,j}$, can be obtained from

$$X_{C,j} = \frac{\int_a^b xf(x,j)\,dx}{\int_a^b f(x,j)\,dx} \tag{2}$$

For a discrete function, $I(x,y)$, where x varies in a discrete manner over a horizontal array of pixels from 0 to m, and y is fixed to the current row under analysis, (2) is rewritten as

$$X_{C,j} = \frac{\sum_{i=0}^{m} x_i I(x_i,j)}{\sum_{i=0}^{m} I(x_i,j)} \tag{3}$$

By calculating the horizontal centroid for each one of the points in S_1 and S_2 using (3), a more accurate location of these points is determined. This procedure is performed by selecting a defined window of pixels that enclose the point of maximum intensity along the X-axis of the image plane, as shown in Figure 5.

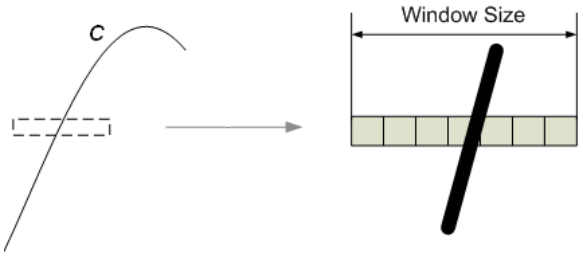

FIGURE 5. PIXEL WINDOW FOR HORIZONTAL CENTROID CALCULATION.

The outcome of the fine location procedure is the redefinition of the sets S_1 and S_2, as S_{C1} and S_{C2}. Using a matrix notation, S_{C1} and S_{C2} are expressed as

$$\begin{aligned} S_{C1} &= \begin{bmatrix} \boldsymbol{X_C} & \boldsymbol{Y} \end{bmatrix} \\ S_{C2} &= \begin{bmatrix} \hat{\boldsymbol{X}}_{\boldsymbol{C}} & \hat{\boldsymbol{Y}} \end{bmatrix} \end{aligned} \tag{4}$$

where

$$\boldsymbol{X_C} = \begin{bmatrix} x_{C1} \\ x_{C2} \\ \vdots \\ x_{Cm} \end{bmatrix} \quad \hat{\boldsymbol{X}}_{\boldsymbol{C}} = \begin{bmatrix} \hat{x}_{C1} \\ \hat{x}_{C2} \\ \vdots \\ \hat{x}_{Cn} \end{bmatrix} \tag{5}$$

The vectors in (5) contain the X-axis centroid data of each one of the points in S_1 and S_2. The column vectors $\boldsymbol{Y}$ and $\hat{\boldsymbol{Y}}$, containing the Y-axis coordinates of the points in S_1 and S_2, are kept the same for S_{C1} and S_{C2}, respectively.

Step 3: Line fitting

At this point, the information collected in the preceding steps is utilized to completely define the shape of the curves C_1 and C_2, through a least-squares approximation. This final step aims to minimize the effects of noise, associated with imperfections in the measurement process, refreshment rates of the LCD screen and external light sources. In general, the form of the first curve is defined as

$$C_1 = K_0 + K_1 x + K_2 x^2 + \ldots + K_p x^p \tag{6}$$

where $K_0, K_1, \ldots, K_p, p \in \mathbb{R}$ and x is the X-coordinate of one of the points in S_{C1}. If the X-

coordinates of all the m points in S_{C1} are considered, an m x (p+1) Vandermonde matrix can be constructed and (6) takes the form

$$\boldsymbol{C_1} = \begin{bmatrix} 1 & x_{C1} & x_{C1}^2 & \cdots & x_{C1}^p \\ 1 & x_{C2} & x_{C2}^2 & \cdots & x_{C2}^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{Cm} & x_{Cm}^2 & \cdots & x_{Cm}^p \end{bmatrix} \begin{bmatrix} K_0 \\ K_1 \\ \vdots \\ K_p \end{bmatrix} \quad (7)$$

Finally, the coefficients $K_0,K_1,\ldots,K_p$ are obtained by finding E, such that

$$E = \min_{K_i} \left| \boldsymbol{Y}\text{-}\boldsymbol{C_1} \right|^2, \; i \in [0, p] \quad (8)$$

C_2 is calculated in the same way as C_1, but using the data in S_{C2}. Once the equations of the curves are known, the intersection of the functions is easily computed analytically.

EXPERIMENTAL VALIDATION

Experiments are conducted using BMP images, in order to minimize information losses due to image compression techniques. Furthermore, the curves used for testing are based on the most basic shape for a 2-dimensional function: a straight line. Hence, the target is represented by the intersection of two diagonal lines.

Fixed-Target

The first task is the identification of the optimal combination of pixel intensities for the generation of the target. On an 8-bit display, pixel intensities vary from 0 (black) to 255 (white). Based on this range of intensities, and considering that a strong distinction between the points on the diagonals and the background is desired, two types of formats are tested.

Figure 6 shows two types of targets: black target over white background (A), and white target over black background (B). In order for the centroid algorithm to provide accurate data, high contrast between the pixels of interest and the background is required. In general, the contrast that a given imaging system is able detect can be described on a percentage basis, and is calculated from (see [Edmund Optics])

$$Contrast = \frac{(I_{\max} - I_{\min})}{(I_{\max} + I_{\min})} \cdot 100\% \quad (9)$$

The contrast percentages obtained for our test bench, corresponding to the two target formats, are summarized in Table 1. These results suggest that the optimal configuration is achieved by displaying a white target over black background. Hence, this format is used for the remaining experiments.

TABLE 1. CONTRAST PERCENTAGES FOR TWO DIFFERENT TARGET FORMATS.

	Black Target	White Target
Contrast	65.36%	84.67%

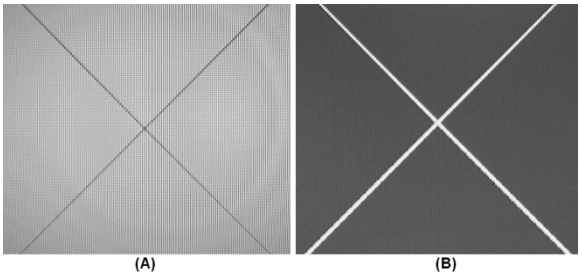

FIGURE 6. COMPARISON OF TARGET FORMATS, WITH LCD-CAMERA DISTANCE d=36.02mm, (A) BLACK TARGET OVER WHITE BACKGROUND, (B) WHITE TARGET OVER BLACK BACKGROUND.

Time-Varying Target

The ability of the system to detect displacements along the X-axis of the stage-image system is tested for 4 different camera-LCD distances. Changes in the distance, d, between the camera and the LCD change the effective magnification of the imaging system. When the camera is close to the LCD, the field of view is small, but each LCD pixel covers many pixels in the camera image plane, suggesting high resolution. However, this higher resolution in location of individual pixels is counteracted by the fact of the camera image containing a smaller portion of the target image. In contrast, when the camera is further from the LCD, the resolution with which the location of each pixel can be located is reduced, but a larger portion of the target image is included in the curve fitting routine.

TABLE 2. STANDARD DEVIATION OF THE IMAGE SAMPLES, FOR 4 DIFFERENT LCD-CAMERA DISTANCES

d [mm]	Standard deviation in X [pxls]									Average Std [pxls]
	0um	5um	10um	15um	20um	25um	30um	35um	40um	
16.02	0.0318	0.0342	0.0299	0.0301	0.0313	0.0318	0.0293	0.0282	0.0301	0.0308
26.02	0.0123	0.0142	0.0148	0.0135	0.0151	0.0130	0.0161	0.0187	0.0159	0.0148
36.02	0.0102	0.0123	0.0144	0.0123	0.0174	0.0132	0.0135	0.0148	0.0119	0.0133
46.02	0.0133	0.0135	0.0114	0.0130	0.0140	0.0133	0.0121	0.0114	0.0131	0.0128

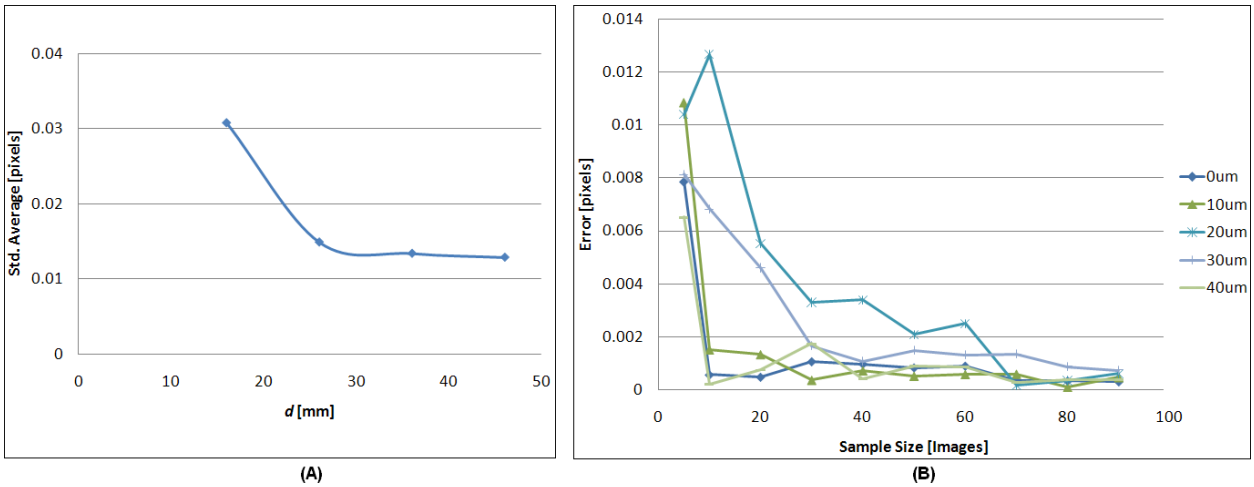

FIGURE 7. (A) STANDARD DEVIATION AS A FUNCTION OF LCD-CAMERA DISTANCE, (B) ABSOLUTE ERROR AS A FUNCTION OF SAMPLE SIZE, FOR FIXED STEP DISPLACEMENTS OF 5µm AND *d*=36.02mm.

We tested this effect by moving the stage in 5µm steps over a total distance of 40µm. At each step, 100 image samples are collected, and the location of the target point is computed. The mean of these points is considered as the "best" estimate of the measured target motion. The standard deviation of target position estimates was also computed. This procedure was repeated for 4 different camera-LCD distance, *d*. The results for standard deviations are tabulated in Table 2.

As the LCD-camera distance grows to a feasible limit, the magnification between LCD pixels and camera pixels approaches 1. From Table 2, this seems to provide a much sharper mapping; that is, as *d* grows the standard deviation of the target position decreases. This means that a *thin* diagonal line on the display appears also as a *thin* line on the image plane. Another representation of the data in Table 2 is presented in Figure 7 (A), where the variation of the standard deviation is plotted as a function of *d*. As *d* grows, the standard deviation of the samples decreases and the system is less affected by noise. Figure 7 (B) shows the effect of increasing the sample size on the variation in the estimated position of the intersection point. This chart shows the absolute difference between displacements, calculated for the entire sample of 100 images, minus the average obtained using a smaller number of images. Although only shown for *d*=36.02mm, the absolute error for all tested values followed a similar pattern as the one displayed in Figure 7 (B). It can be seen that using more than about 40 or 50 images does not significantly decrease the variation in the position estimate.

Alternatively, and as expected from the pinhole camera model [Wong, Montes et al.], it is harder for the sensor to detect small displacements as the pixel magnification decreases (i.e., as $d \to \infty$). This phenomenon is shown in Figure 8, where the real displacement, measured in µm, is compared with the sensed displacement also given in µm.

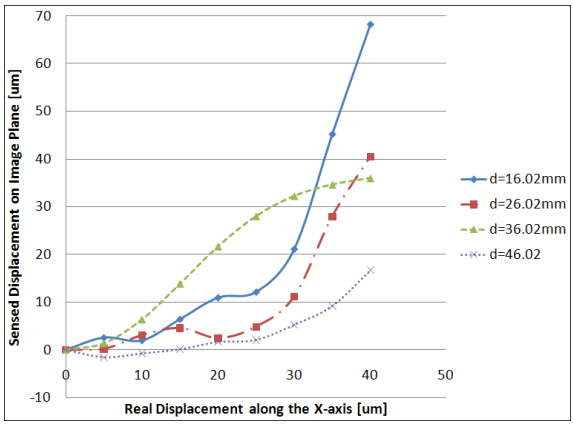

FIGURE 8. SENSOR SENSITIVITY: SENSED DISPLACEMENT VS. REAL DISPLACMENT, FOR 4 DIFFERENT LCD-CAMERA DISTANCES AND A SAMPLE SIZE OF 50 IMAGES.

From Figure 8 it appears that the optimal working distance for this particular camera, is *d*=36.02mm. Ideally we would want the displacement estimated by the sensor to be a purely linear function of the known displacement of the stage. When a linear fit to the triangular marked line representing the results with 36.02mm working distance is performed, the linear regression coefficient, r^2, is 0.9664 indicating a reasonably good linear relationship. This suggests that displacements in the order of 5µm can be detected easier for *d*=36.02mm, using the proposed implementation.

CONCLUSION

A method to measure planar displacements of a target on an LCD display was presented. Experimental results on a vision-controlled testbed revealed a tradeoff between noise rejection and the ability of the system to detect displacements parallel to the image plane. Test data also demonstrated that displacements as small as 5µm can be detected by collecting and processing samples of at least 40 images. The best target configuration was found to be a white target over black background, at an LCD-sensor working distance of 36.02mm.

ACKNOWLEDGEMENT

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

Bruck, H. A., S. R. McNeill, et al. (1989). "Digital Image Correlation using Newton-Raphson Method of Partial Differential Correction." *Experimental Mechanics* Vol. 29(3), pp. 261-267.

Chen, C. and M. Lin (2006). "An improved adaptive centroid estimation algorithm". *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, Hong Kong, China.

Edmund Optics. (2008). "Electronic Imaging Resource Guide." Retrieved 13Oct2008, from http://www.edmundoptics.com/TechSupport/DisplayArticle.cfm?articleid=286.

Fan, KC, H. Wang, et al. (2003). "Development of a 3D Laser Ball Bar for the Volumetric Error Measurement of Multi-axis Machines." *Transactions of the North American Manufacturing Research Institution of the Society of Manufacturing Engineers (NAMRI/SME)* Vol. 31, pp. 249-256.

Fosu, C., G. W. Hein, et al. (2004). Determination of Centroid of CCD Star Images. XXth ISPRS Congress. Istanbul.

See, A. (2005). "Rapid prototyping design and implementation of a motion control integrated with an inexpensive machine vision system". *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, Ottawa, ON, Canada.

Sutton, M. A., C. Mingqi, et al. (1986). "Application of an optimized digital correlation method to planar deformation analysis." *Image and Vision Computing* Vol. 4(3), pp. 143-50.

Wong, C., C. Montes, et al. (2008). "A New Position Feedback Method for Manufacturing Equipment." *Proceedings of the ASME International Manufacturing Science and Engineering Conference 2008* Vol. MSEC2008, pp. 111-120.

Xie, H., L. Sun, et al. (2005). "Visual servoing with modified Smith predictor for micromanipulation tasks". *2005 IEEE International Conference on Mechatronics and Automation (IEEE Cat. No. 05EX1044)*, Niagara Falls, Ont., Canada.