A moving window correlation method to reduce the distortion of scanning probe microscope images

Wei Chu,[1,2] Joseph Fu,[1] Ronald Dixson,[1] George Orji,[1] and Theodore Vorburger[1]
[1]*National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*
[2]*School of Mechanical and Electrical Engineering, Harbin Institute of Technology, Harbin 150001, People's Republic of China*

(Received 27 March 2009; accepted 7 July 2009; published online 31 July 2009)

Many scanning probe microscopes such as the scanning tunneling microscope and atomic force microscope use piezoelectric actuators operating in open loop for generating the scans of the surfaces. However, nonlinearities mainly caused by hysteresis and drift of piezoelectric actuators reduce the positioning accuracy and produce distorted images. A moving window correlation method is proposed in this paper to determine and quantify the hysteresis. This method requires both trace and retrace profiles to be recorded. With a window imposed on each of the profiles, correlations are implemented between the data inside two windows to find corresponding pixel pairs on two different profiles but the same physical positions along the fast scanning axis (x). The x-distances between pixel pairs are calculated and then a simple correction scheme is applied to reduce the distortion. © *2009 American Institute of Physics.* [DOI: 10.1063/1.3189041]

I. INTRODUCTION

Piezoelectric tube (PZT) actuators are widely used in micropositioning applications because of their properties of high stiffness and fast frequency response. Many scanning probe microscopes (SPMs), such as scanning tunneling microscopes and atomic force microscopes (AFMs) use a PZT for generating the scans of the surfaces. However, the piezononlinearities mainly caused by hysteresis and drift of piezoelectric actuators can reduce the positioning accuracy and produce distorted images. Because of the hysteresis, the pixel positions of trace and retrace profiles along the scanning direction will slightly deviate from the real physical positions in opposite directions. Even if the scan size is calibrated in advance for an open-loop SPM, the hysteresis will still distort the image features.

It is important to develop effective methods to eliminate the error that is induced by the nonlinearity of PZTs for the purpose of precision measurement. Different approaches have been explored to attempt to solve this problem. These include research on improvement of piezoelectric material and closed loop control with accurate position sensors.[1–3] Another method applies the electric charge rather than voltage on the PZT actuator.[4] However this method has a cost of reducing the available scan range. Sometimes, a complex nonlinear control signal instead of a typical triangular wave signal is applied to reduce the hysteresis effect. This counteracts the hysteresis and produces an approximate linear stretch or contract on the PZT actuator.[5,6] Unfortunately, it is almost impossible to build an accurate model to cover all variants like scan range, scan size, voltage offset, and so on.

All these techniques require extensive modification of the instruments. Without requiring any modification on hardware, post-image correction methods may be of interest to large number of users of common open-loop SPM instruments that are not applied for metrology purposes. Jørgensen *et al.*[7] presented such a method based on the measurement of a periodic sample. In this paper, a moving window correlation method is proposed to determine and quantify the hysteresis in the fast scanning direction. It does not have limitation for the pattern of the sample. Then similar correlation method is applied to attempt to solve the hysteresis problem in the slow direction.

II. MOVING WINDOW CORRELATION

Theoretically, a pair of trace and retrace profiles would be identical and their x coordinates (along the fast scanning direction) would be equal to the sample's real physical position if the PZT displacement were an exact linear function of its driving voltage. In reality, a separation between trace and retrace always exists because of hysteresis. Figure 1 shows a sketch of a typical hysteresis loop. The ascending branch and the descending branch correspond to trace and retrace scans, respectively. Due to the hysteresis, the same physical location corresponds to slightly different voltages depending on whether the voltage is ascending or descending. For ex-

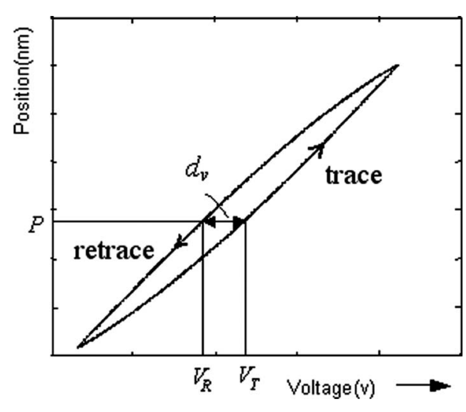

FIG. 1. A sketch of typical hysteresis loop.

0034-6748/2009/80(7)/073709/4/$25.00

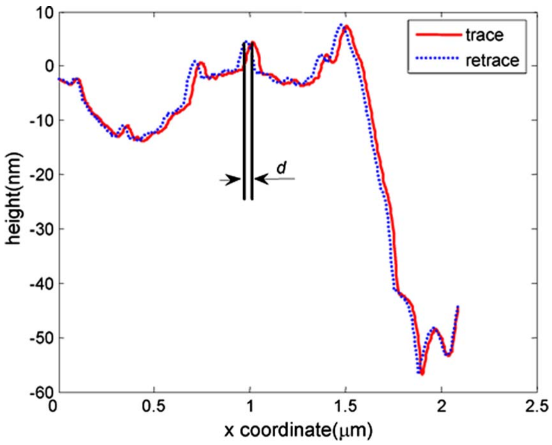

FIG. 2. (Color online) Difference between trace and retrace caused from hysteresis.

ample, for the same physical position P, the driving voltages of the trace and retrace scans are V_T and V_R, respectively. The difference between them is denoted as d_v. Without a displacement sensor or other correction tool, the driving voltages will be proportionally converted to the lateral coordinates of recorded profiles. Therefore, this voltage difference d_v translates into a coordinate difference d between the trace and retrace profiles, as shown in Fig. 2. Plotted in Fig. 2 is a pair of trace and retrace profiles scanned over the surface of a compact disk, with a 2.08 μm scan size, a 0.3 Hz scan rate and a sampling density of 512 points per line. The slow scan (along the y-direction) is disabled to ensure that the scans are performed at the same section of the sample.

As a result, both the data of trace and retrace profiles deviate from each other at the designated point by a distance d and from the true surface position, which is in between them. If the values of d for every pixel position could be evaluated, the hysteresis would probably be corrected. A moving window correlation method is used to determine the values of d for each pixel. This method requires taking both trace and retrace images during the scan. The correlation is performed in line-by-line fashion, using corresponding trace and retrace pairs shown as in Fig. 3.

In Fig. 3, a pair of trace and retrace profiles are plotted

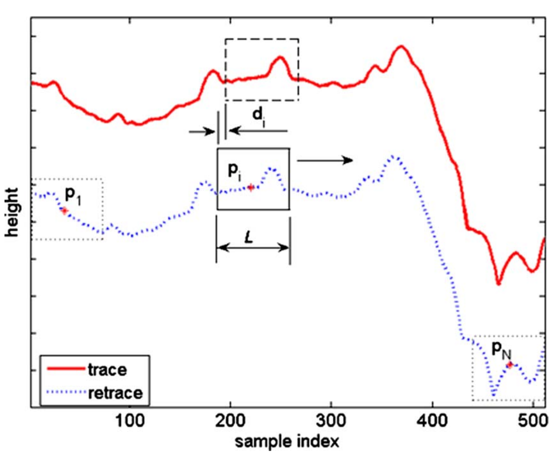

FIG. 3. (Color online) A diagram of the moving window correlation. The profiles are offset vertically for clarity.

as a solid line and a dotted line, respectively, each with the same-sized window superimposed on it. One window is fixed and the other moves around to find a best match position. Then, the fixed window moves to the next pixel position. This process is repeated until the window passes the entire profile from beginning to end, pixel by pixel. For instance, we hold the retrace as a reference, and the window on the retrace profile is fixed and is denoted as R with a width of L. Its center moves from P_1 to P_N (N=512-L+1). Window R is expressed as R_i when its center moves to position P_i ($i=1,2,\ldots,N$) and the data point included in this window is denoted as $R_i(k)$ ($k=1,2,\ldots,L$). For a window R_i, there would be a corresponding same-sized window T moving on the trace profile around pixel position i within a specified range. The data included in window T as $T_j(k)$ ($k=1,2\ldots L$). A cross correlation function will be performed with $R_i(k)$ and $T_j(k)$ according to Eq. (1). Here the subscript j signifies the pixel difference between window R and window T.

$$f_i(j) = \frac{\sum_{k=1}^{L} R_i(k) \times T_j(k)}{\sqrt{\sum_{k=1}^{L} [R_i(k)]^2}\sqrt{\sum_{k=1}^{L} [T_j(k)]^2}}. \quad (1)$$

The best match position for two windows is obtained when the cross correlation function $f_i(j)$ reaches its maximum. In another word, the best matched location for a pair of trace and retrace profiles is indicated by a maximum cross correlation function $f_i(j)$. This can be expressed mathematically as the argument of the maximum D_i.

$$D_i = \arg\max_j \{f_i(j)\}. \quad (2)$$

For a perfect linear system, this best matched location will occur when D_i is equal to zero. A nonzero value of D_i, which represents the coordinate difference between trace and retrace profiles at the ith pixel point, is a result caused by the hysteresis. Noted the range of i and j is less than the number of pixels in each profile, i.e., 512 points; therefore, there are $(L/2)$ pixels with no cross correlation value at the beginning and the ending of a pair of profiles. If we do not want the image size after processing to be reduced, at these regions the hysteresis data should be extrapolated. Theoretically, there is no hysteresis at both ends of a hysteresis loop as shown in Fig. 1.

The width of L significantly affects the calculation; a too small L will cause insufficient information for correlation calculation, and whereas a too large L will lead to a practical average or smooth effect of multiple neighboring points hence reducing the sensitivity. An empirical value of 50 pixels is chosen in this paper.

III. IMAGE RECONSTRUCTION

The moving window correlation method is applied to a pair of trace and retrace profiles and D (set of all calculated D_i) is plotted in Fig. 4(a). In Fig. 1, for a certain physical point, the difference of driving voltage between trace and retrace, d_v, increases with driving voltage monotonically un-

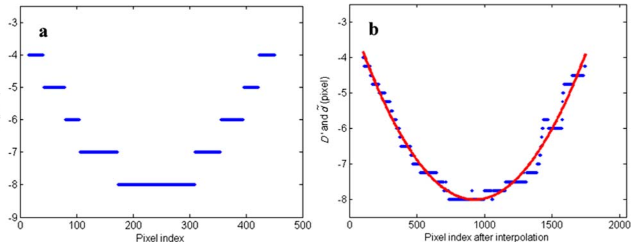

FIG. 4. (Color online) Calculated difference between trace and retrace profiles; (a) directly calculated result D_i using raw data; (b) calculated result D_i' using interpolated data and its fitting curve $\widetilde{d}$ as final hysteresis estimate.

til the midpoint of the driving voltage is reached and then decreases after that. So the x-coordinate difference, d, between the trace and retrace should bear a similar relation between d_v and the driving voltage. The index difference D_i for trace and retrace is plotted in Fig. 4(a), which is consistent with the expected behavior. Although it can be used for the correction directly, it is somewhat coarse and discrete. Therefore, a second order polynomial fit is applied to yield a smooth curve $\widetilde{d}$ as the final hysteresis estimate of d. An interpolation function is also applied to the pair of profiles prior to the correlation to increase the number of data points so as to alleviate the discontinuity. Figure 4(b) shows D' the coordinate difference calculated using same pair of profiles as in Fig. 4(a) but with four times the number of original pixel points, i.e., 4×512, and its second order fitting curve $\widetilde{d}$.

The ultimate goal is to reduce the image distortion caused by the hysteresis, and the previously described moving window correlation method is used to determine and quantify the nonlinearity. Because of the complexity of hysteresis behavior and lacking a mature mathematical model, we assume that the deviation of both trace and retrace from the real physical position is the same and is equal to one half of d which has been determined by the moving window correlation method with an approximate value $\widetilde{d}$.

Let m be a point position on the retrace, then with the hysteresis estimate $\widetilde{d}$, there would be a corresponding point positioned at $(m+\widetilde{d})$ on the trace. Therefore, the corrected pixel position for trace and retrace will be

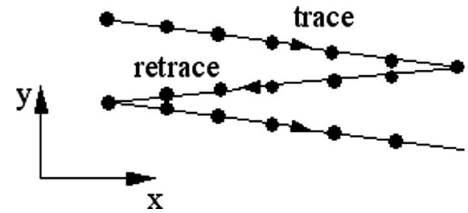

FIG. 6. Schematic description of the tip movement during an AFM scanning.

$$m_c = m + \widetilde{d}(m)/2, \tag{3}$$

where m is the original pixel position $(m=1,2,\ldots,512)$; $\widetilde{d}(m)$ is the estimated hysteresis value, sampled at mth position of $\widetilde{d}$; m_c is the modified pixel position after hysteresis correction. The value of m is an integer, but apparently the result of m_c could be a noninteger value because $\widetilde{d}(m)$ is sampled from a continual curve. Therefore an interpolation process is required to find the height value at the integer pixel location for maintaining the integrity of the data structure.

AFM images consist of sequential scan lines; the image reconstruction process can be implemented through line-by-line correction of hysteresis. Figure 5 shows the reconstructed result of the profile pair in Fig. 3; they match well.

The measurement shown in Fig. 2 was obtained with the slow scan being turned off. In reality, when the slow scan is on, the data acquired from trace and retrace scans are not at exactly the same horizontal section of the sample. Figure 6 describes the movement of our PZT scanner during an AFM scan. Generally the y-spacing between scans is small relative to feature sizes on the sample so that the topography information between corresponding trace and retrace profile is correlated.

Figure 7 shows one of profiles with the slow scan enabled. After image reconstruction [Fig. 7(b)], the x coordinates for the same lateral position point in trace and retrace are identical. However there still is a height difference between them because of the random error. In measurements,

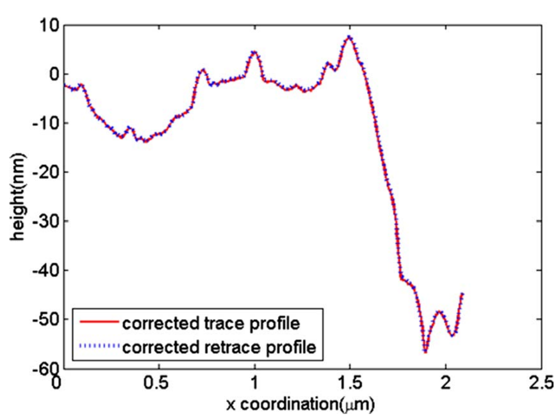

FIG. 5. (Color online) Reconstructed trace and retrace profiles.

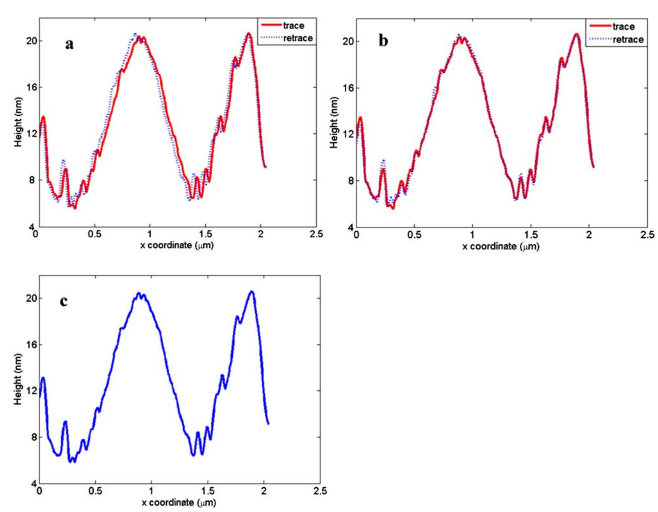

FIG. 7. (Color online) (a) Raw data of a pair of trace and retrace profiles after slow scan enabled; (b) reconstructed profiles; (c) averaged profile.

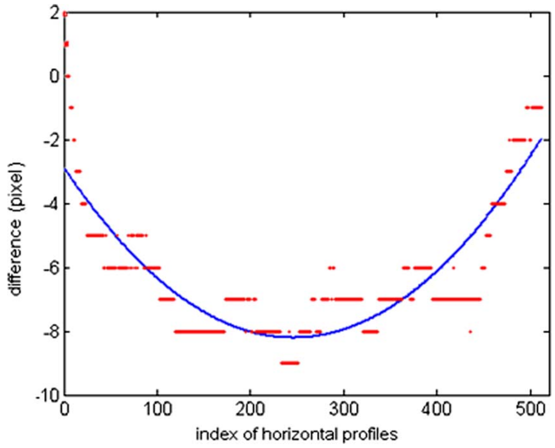

FIG. 8. (Color online) Discrete and fitted pixel differences for each horizontal profile.

an average processing is a common and meaningful approach to eliminate the random error when the position is aligned. So the averaged profile of the corrected trace and retrace profile pair is plotted in Fig. 7(c).

IV. CORRECTION OF *Y*-DIRECTION HYSTERESIS

Analogously, the hysteresis in the slow scanning direction can also be determined by correlating two trace (or two retrace) images with opposite stepping in the *y*-direction, upward and downward. This is more complicated because the drift of PZT scanner and other unknown factors could become mixed up with hysteresis during the longer scan time. Matching up each fast scan line in two corresponding images via a similar cross correlation calculation (no window superimposed on), the index difference between two matched lines, which is the effect of hysteresis and drift, can be obtained. Let image A and B be scanned upward and downward, respectively, and select image A as a reference. Each horizontal profile a_i $(i=1,2,\ldots,512)$ will be taken to calculate the cross correlation function with the horizontal profiles in image B to find out a best matched profile b_k. Here i and k are profile indexes in the upward and downward images, respectively. Figure 8 is a plot of the differences between i and k along with a second order polynomial fit. The rest of the reconstruction process is also similar to the correction scheme for the fast scan in *x*-direction. The difference is that, only one hysteresis loop exists in the slow scan direction (y) whereas 512 hysteresis loops exist in the fast scan direction (x). Notice the parabolic line is the least-squares fit for all data points obtained as the index differences for every profile. Mostly, the difference is the contribution of hysteresis, and the difference of the first data point (at 2 pixels) and the last point (at −1 pixel) is due to the drift, approximately 3 pixels for a pair of scanning downward and upward images.

V. CONCLUSIONS

For an open-loop SPM, even those images with the scan size calibrated may have distorted features within them due to the hysteresis. We find the hysteresis between trace and retrace as large as 8 pixels (32 nm) for an image of about 2.1 μm scan size with 0.3 Hz scan rate. The moving window cross correlation can be implemented to quantify the hysteresis and therefore a post-image correction on the acquired image can be performed. For SPM devices that have hysteresis correction functions implemented by driving the piezoelectric actuator with a nonlinear voltage, due to the complexity of the hysteresis problem, the nonlinear driving voltage may not be sufficiently accurate. The method described here can examine the efficacy of nonlinear correction and may provide an input for setting the model parameters.

Conceivably, initial scanned profiles of trace and retrace can be analyzed to determine the hysteresis function and this quantity can be taken into account for subsequent scanning line as a semi-real time correction.

[1] E. Kouno and P. A. McKeown, CIRP Ann. **33**, 369 (1984).
[2] Y. Okazaki, Precis. Eng. **12**, 151 (1990).
[3] R. C. Barrett and C. F. Quate, Rev. Sci. Instrum. **62**, 1393 (1991).
[4] C. Newcomb and I. Flinn, Electron. Lett. **18**, 442 (1982).
[5] K. Dirscherl, J. Garnæs, L. Nielsen, J. Friis Jørgensen, and M. P. Sørensen, J. Vac. Sci. Technol. B **18**, 621 (2000).
[6] Digital instruments Nanoscope II instruction manual. Version 5. Digital Instruments Inc.
[7] J. F. Jørgensen, K. Carneiro, L. L. Madsen, and K. J. Conradsen, J. Vac. Sci. Technol. B **12**, 1702 (1994).