Improved accuracy and speed in scanning probe microscopy by image reconstruction from non-gridded position sensor data

To cite this article: Dominik Ziegler et al 2013 Nanotechnology 24 335703

View the article online for updates and enhancements.

Related content

- A monolithic MEMS position sensor for closed-loop high-speed atomic force microscopy
  N Hosseini, A P Nievergelt, J D Adams et al.

- Gwyscan: a library to support non-equidistant scanning probe microscope measurements
  Petr Klapetek, Andrew Yacoot, Petr Grolich et al.

- Advances in the atomic force microscopy for critical dimension metrology
  Danish Hussain, Khurshid Ahmad, Jianmin Song et al.

Recent citations

- Ideal Scan Path for High-Speed Atomic Force Microscopy
  Dominik Ziegler et al

- Rotational scanning atomic force microscopy
  A Ulinas and Š Vaitekonis

- Controlling nanoflake motion using stiffness gradients on hexagonal boron nitride
  Matthew Becton and Xianqiao Wang
Improved accuracy and speed in scanning probe microscopy by image reconstruction from non-gridded position sensor data

Dominik Ziegler\textsuperscript{1}, Travis R Meyer\textsuperscript{2}, Rodrigo Farnham\textsuperscript{3}, Christoph Brune\textsuperscript{4}, Andrea L Bertozzi\textsuperscript{2} and Paul D Ashby\textsuperscript{1}

\textsuperscript{1} Lawrence Berkeley National Laboratory, Molecular Foundry, 1 Cyclotron Road, 94720 Berkeley, CA, USA
\textsuperscript{2} Department of Mathematics, University of California Los Angeles, 520 Portola Plaza, Los Angeles, CA 90095-1555, USA
\textsuperscript{3} Department of Mathematics and Statistics, California State University, Long Beach, 1250 Bellflower Boulevard, Long Beach, CA, 90840-1001, USA
\textsuperscript{4} Department of Mathematics and Computer Science, University of Münster, Einsteinstrasse 62, D-48149 Münster, Germany

E-mail: pdashby@lbl.gov

Received 16 March 2013, in final form 3 July 2013
Published 26 July 2013
Online at stacks.iop.org/Nano/24/335703

Abstract

Scanning probe microscopy (SPM) has facilitated many scientific discoveries utilizing its strengths of spatial resolution, non-destructive characterization and realistic \textit{in situ} environments. However, accurate spatial data are required for quantitative applications but this is challenging for SPM especially when imaging at higher frame rates. We present a new operation mode for scanning probe microscopy that uses advanced image processing techniques to render accurate images based on position sensor data. This technique, which we call sensor inpainting, frees the scanner to no longer be at a specific location at a given time. This drastically reduces the engineering effort of position control and enables the use of scan waveforms that are better suited for the high inertia nanopositioners of SPM. While in raster scanning, typically only trace or retrace images are used for display, in Archimedean spiral scans 100\% of the data can be displayed and at least a two-fold increase in temporal or spatial resolution is achieved. In the new mode, the grid size of the final generated image is an independent variable. Inpainting to a few times more pixels than the samples creates images that more accurately represent the ground truth.

(Some figures may appear in colour only in the online journal)

1. Introduction

The pace of discovery in nanoscience accelerated significantly with the invention of scanning probe microscopy [1, 2]. Scanning probe microscopy has the ability to image material surfaces with exquisite resolution in a large variety of environments from vacuum [3] to high temperature and pressure [4]. Furthermore, careful probe design facilitates nanoscale measurement of specific physical or chemical properties such as electrostatic [5, 6], magnetic [7, 8], or surface energy [9, 10]. While spatial resolution and precision are traditional strengths of scanning probe microscopy, achieving high accuracy is challenging. As scanning probe microscopy matures scientists seek to perform more quantitative measurements and place higher demands on instrument accuracy. For example, measuring the interaction energy landscape for aggregating proteins requires precise intermolecular distance measurements [11].
Unfortunately, the poor accuracy is exacerbated as researchers push instruments to achieve higher temporal resolution. Increasing image accuracy and temporal resolution for the next generation of scientific discovery requires rethinking how we do scanning in scanning probe microscopy.

Image inaccuracy results from raster scanning being the entrenched paradigm for image creation in scanning probe microscopy. The German expressions for AFM and STM, ‘Rasterkraftmikroskopie’ and ‘Rastertunnelmikroskopie’ respectively, show how the raster concept is fundamentally linked to scanning probe techniques. But the idea of raster scanning predates AFM and STM. For applications like analog television, where transmission bandwidth was precious, it was economical that a single data series could create images without using X, Y position data. When AFM and STM were invented in the mid-1980s before digital signal acquisition became commonplace raster scanning facilitated crafting 3D topographs from individual paper scan lines printed by pen plotters [12]. In the digital age, the advantage of raster scanning is that it speeds display and saves memory. By sampling at a constant rate, only a single channel needs to be recorded and each sampling maps directly to a corresponding pixel in the final image. However, achieving non-distorted images requires the tip to be at a specific location at a given time with perfectly linear motion of the scanner. Unfortunately, piezoelectric nanopositioners have notoriously nonlinear displacement response and high inertia with mechanical resonances, which significantly compromises image accuracy. Specifically designed nonlinear output voltages can partially compensate the errors caused by piezo nonlinearities. Open-loop techniques frequently use second order modeling of piezo displacement and a few coupling terms to create a more linear displacement [13] (see figure 1(a)). The results are satisfactory for the fast scan axis but creep is not managed well causing errors in the slow scan axis and poor offset and zoom performance. For recently designed scanning probe microscopes it is more common to operate in a closed-loop configuration where X, Y positions are controlled using feedback [14, 15] (see figure 1(c)). Unfortunately, feedback loops have significantly lower bandwidth than the position sensor signal such that accuracy is maintained only up to scan rates of a few lines per second. Feed-forward, also called adaptive scan, is a mode of operation very similar to open-loop but the piezo model used to transform the scan waveform is developed by measuring the response of the piezo with position sensors [16, 17] in the fast scan direction. As an open-loop technique, feed-forward has high bandwidth performance but creep is not managed well. Combining feedback and feed-forward harnesses the advantages of each correction method but is complicated to implement [18]. The enormous engineering effort to control the piezo position has its roots in the paradigm of raster scanning. In the paradigm, the controller dictates strict position requirements based on the scan parameters. But position inaccuracies of the instrument do not influence how data are received and interpreted. This simplifies image display and the onus is on the instrument to provide accurate positioning even though piezo nanopositioners present formidable physical challenges. Another negative consequence of the raster scan paradigm beyond the unnecessary control of piezo position is that sequential scan lines moving in opposite directions are adjacent to each other. Any delay from either X, Y scanner control or the Z-feedback cause adjacent scan lines to be mismatched. Thus the convention is to discard half the data and only show trace or retrace in one image compromising spatial and temporal resolution.

Figure 1. (a) Open-loop scanning: a raster scan wave is applied in both fast and slow scan directions and they define the pixel positions for image display. (b) Open-loop 15 µm × 15 µm scan of a calibration grating using 512 scan lines (256 trace and 256 retrace). Zoom-ins of the yellow dashed rectangle region display topography and amplitude data. Piezo nonlinearity leads to 1.45 µm mismatch between trace and retrace and creep compresses the feature in the slow scan direction. (c) Closed-loop scanning: a feedback loop is used to control piezo position based on independent position sensor data but pixel positions are still defined by the input scan waveform. (d) Closed-loop 15 µm × 15 µm scan of the same calibration grid as (b). Zoom-ins of the yellow dashed rectangle region display topography and amplitude data. The feedback controller regularizes the scan well but delay in the topography feedback loop as well as the XY position feedback cause 0.23 µm mismatch between trace and retrace.
While it may initially seem trivial, to relax $X, Y$ control and passively measure position sensor data to create images is a much more elegant solution to the problem of poorly behaved piezo nanopositioners. The absence of any feedback in $X, Y$ position results in the high bandwidth of open-loop scanning and greater accuracy than any piezo control system. More importantly, the technique frees us from the raster scan paradigm and enables the use of scan waveforms better matched to the physical limitations of piezoelectric nanopositioners and for which all scan time can be used to create images.

In section 2 we discuss the difficulties of raster scanning to display trace and retrace in a single image in greater detail. Section 3 introduces our new sensor inpainting technique to reconstruct images from sensor data. Section 4 highlights the results for a constant velocity Archimedean spiral, and section 5 presents conclusions.

2. Raster scan pattern

2.1. Open-loop scanning

Figure 1 illustrates the difficulties of raster scanning to generate accurate images from trace and retrace scan lines in a single image. The performance of conventional open and closed-loop configurations are compared. The schematic of an open-loop scan mode, the most basic positioning technique for scanning probe microscopy, is shown in figure 1(a). The scan parameters (image size, resolution, and speed) define scan waveforms that drive the piezo actuator and delineate the pixel positions in the image. Figure 1(b) shows topography data, where 256 trace and 256 retrace lines are displayed in the same image. All data presented in this paper were collected on a MFP-3D (Asylum Research, Santa Barbara) using amplitude modulation AFM in air with a free amplitude of 30 nm and an amplitude set-point of 24 nm. The cantilever had a nominal resonance frequency and stiffness of 70 kHz and 30 nm and an amplitude set-point of 24 nm. The cantilever had a nominal resonance frequency and stiffness of 70 kHz and 30 nm and an amplitude set-point of 24 nm.

The scan pattern was a triangular raster scan without using model based correction nor using overscan. The $X, Y$ positions are the applied piezo voltage scaled by the first order coefficient of piezo sensitivity. The total acquisition time was 205 s, with 512 scan lines and 15 $\mu$m scan size, resulting in an average tip velocity of 37.5 $\mu$m s$^{-1}$. The sample features are isolated 6 $\mu$m wide squares with a spacing of 3 $\mu$m and height of 100 nm (calibration grating by Bruker Nano). The edges of the calibration steps in figure 1(b) clearly show that trace and retrace scan lines do not overlay. The multi-domain structure of high sensitivity piezoelectric ceramics causes sensitivity to increase as field increases and hysteresis when field reverses such that the same applied voltage does not result in the same position. Thermally activated alignment of domains causes additional displacement or creep along the slow scan axis, such that a larger scan is compressed into the image. The zoom-ins of the yellow dashed rectangle region in figure 1 display topography and amplitude data and focus on a particle defect. This same area will be used throughout the paper for comparing all the methods discussed. Using open-loop scanning the mismatch between trace and retrace is up to 1.45 $\mu$m for a 15 $\mu$m scan or 10%. The amplitude image shows the alternating dark and light features typical for descending and climbing the step on the calibration grating. For trace and retrace they clearly do not occur at the same location. This large mismatch is mainly due to hysteresis.

2.2. Closed-loop scanning

Another common mode of operation is closed-loop scanning, where feedback loops control piezo position based on independent position sensor data. The pixel positions are still defined by the input scan waveform (figure 1(c)). Closed-loop scanning not only significantly improves image accuracy by compensating hysteresis of the piezo material (figure 1(d)) but also corrects for creep enabling excellent reproducibility for zooming and large offsets. Furthermore, active monitoring of the sensor allows the instrument to respond to unique mechanical characteristics of each scanner and measured drift and slip. While the large scale 15 $\mu$m × 15 $\mu$m images appear to be correct, the zoom-ins reveal a remaining discrepancy of 0.23 $\mu$m or 1.5% in the closed-loop image. Any delay from either $X, Y$ scanner control or the Z-feedback still causes this mismatch. The result clearly demonstrates that using raster scan lines that move in opposite directions necessitates throwing away half the data for image creation, even when closed-loop operation is used.

3. Sensor inpainting

The enormous engineering effort to control piezo position has its roots in the paradigm of raster scanning. In the paradigm, based on the scan parameters, the controller dictates strict position requirements. Sensor inpainting relaxes this control and uses advanced image processing techniques to create gridded images from non-gridded sensor data (figure 2(a)). Inpainting is a class of digital image processing methods used to solve missing data problems [19]. Traditionally it has been used for such problems as digital restoration of films, artwork restoration such as old frescos [20], and removal of occlusions such as text from photographs. Special effects in the movie industry can also make use of inpainting algorithms, e.g. for removing objects/people from movies, while reasonably filling in the background [21]. Recently inpainting has also been used in 3D fluorescence microscopy or tomography to address low z-axis resolution and gaps between slices [22]. Many inpainting algorithms are based on partial differential equations [21, 23–25] or variational minimization approaches [26]. One of the most basic inpainting methods is heat equation inpainting (also called harmonic inpainting). It has the same functional form as diffusion problems in physics and when applied to image processing it linearly diffuses the known data to unknown regions. More advanced methods better maintain edge sharpness by using total variation (TV) priors [26–28] representing nonlinear diffusion, or use similar regions (patch comparisons) elsewhere in the image to inform the regions of interest (Non-Local Means, NLM) [29–32]. Those nonlocal...
and nonlinear inpainting approaches are often based on nonlocal derivatives or dictionary learning techniques.

In the scanning probe microscopy application, the missing data are the values of the pixels in a gridded image. The collection of these unknown, not-measured pixels is called the inpainting domain. Figures 2(b)–(e) present the steps for image generation from non-gridded data using heat equation inpainting. Figure 2(b) shows the measured $X, Y$ positions of non-gridded sensor data. The topography data recorded at each point are represented by the color of each square. To redistribute the non-gridded data back to the grid of the desired image we use linear binomial interpolation. The height information of each data point is spread to the four nearest points on the grid (figure 2(c)). Furthermore, we attribute to each point a weighting factor, which describes the confidence of the data, and is given by the distance from the data point to the grid. When only one data point contributes to the pixel the height value is simply copied and the weighting saved for use in the inpainting algorithm. When more than one data point contributes to the same pixel, the weights are used to linearly interpolate height information from the contributing data points to determine the value (figure 2(c)) and a composite weighting value is saved for the inpainting algorithm. Hence, for large data sets and coarse grids this first step might be sufficient to attribute a value to each pixel and thereby generate a full image. But pixels might remain empty when sparse data sets are projected on a fine grid. In this case, heat equation inpainting (figure 2(d)) diffuses the existing weighted data out to the entire grid filling empty data points while denoising. (e) Final rendered image.

(f) Inpainted result from the open-loop data in figure 1(b). Despite hysteresis and creep, a correct and non-distorted image is generated (g) Zoom-in of dashed area in (f) shows good overlap of forward and backward scan lines without any control of $X, Y$ piezo position. Mismatch is only due to $Z$-feedback delay. (h) A delay correction can be used to improve the mismatch but subtle inaccuracies remain from raster scan lines moving in opposite directions.

$$\min_{u} E(u) = \int_{\Omega} (\nabla u)^2 + \int_{D} \lambda(u - f)^2$$

is minimized to compute the inpainted result, $u$, from the collected data, $f$, which was distributed on the grid (figure 2(e)). The domains $D$ and $\Omega$ represent the collected data area and area to be inpainted respectively. While $\lambda$ is a scalar based on the weightings used to create the gridded data set, $f$. The equation includes a gradient term to produce a smooth result and difference terms for fidelity to the original measured data. Since the functional is minimized over the whole image, the relative contribution of the gradient term determines the amount of smoothing of the data during inpainting. Sensor inpainting of open-loop data from figure 1(b) produces an accurate result in figure 2(f). The square shape of the features and the fact that the final resulting image is elongated in the $Y$-axis are evidence that hysteresis and creep are accommodated properly. A full image can be restored using all the data but the zoom-ins still reveal a mismatch between trace and retrace scan lines. In the closed-loop configuration (figure 1(d)) the $0.23 \mu m$ discrepancy was partially due to $X, Y$ control delay. Sensor inpainting removes all $X, Y$ delay however an offset of $0.17 \mu m$ remains from $Z$-feedback delay (figure 2(g)). Identifying the $Z$-feedback loop as a persistent source of delay between topography values and their position enables compensation of the delay by offsetting the data before generating the image using inpainting. Figure 2(h) shows the result with a 5 ms offset, which corrects for line mismatch. However, subtle differences between trace and retrace due to hysteresis of the $Z$ piezo as well as effects from the $z$ feedback loop overshooting remain. Even while using sensor inpainting, these unavoidable artifacts result from persisting
with the raster scan paradigm. Fortunately, sensor inpainting enables use of scan waveforms better matched to the physical limitations of piezoelectric nanopositioners.

4. Spiral scan pattern

The image artifacts associated with raster scan lines moving in opposite directions necessitates throwing away half the data for image creation. Scan waves that direct the scanner to move in the same direction for adjacent scan segments enable the display of 100% of the scan data without artifacts. Scanning the perimeter of consecutively smaller concentric squares would satisfy this condition. However, stopping and starting the massive scanner is challenging, as the present need for overscan of triangular raster scan waveforms attests. As sensor inpainting conveniently renders non-gridded data, following a grid is not important when creating scan waveforms. Scanning a smooth spiral allows adjacent scan segments to move in the same direction and does not have sharp turns with high acceleration making it preferable for high inertia scanners. Spiral techniques are common to many data storage techniques on spinning mediums (vinyl records, hard drives, compact disks, or DVDs). But the spiral scan concept only recently found appearance in scanning profilometers [33], nanoscale data storage [34] or scanning probe microscopy, where spiral [35–37], cycloid [38], Lissajous [39], and various other non-raster scan patterns [40] have been demonstrated. Most of these non-raster scan attempts use sensors to steer the probe over the sample in closed-loop. Spiral scanning has been shown to be useful for fast scanning [35–37]. The narrow frequency spectrum of sinusoidal scan trajectories has been shown to require less bandwidth of the feedback loop [39]. As sensor inpainting uses no feedback at all, its bandwidth is simply given by the performance of the position sensor itself. In figure 3 we show the results for a constant velocity Archimedean spiral which has the simple waveform as a function of time, $t$, $X = \alpha \sqrt{t} \sin(\beta \sqrt{t})$ and $Y = \alpha \sqrt{t} \cos(\beta \sqrt{t})$, where the frequency decreases as $1/\sqrt{t}$ while the amplitude increases as $\sqrt{t}$. $\alpha$ and $\beta$ are coefficients derived from the scan parameters: number of loops, scan size, and scan speed. Topography and amplitude data for the sensor inpainting result of a 256 loop spiral are shown in figures 3(a) and (b) respectively. For illustrating the scan direction a spiral scan pattern with few loops is overlaid onto the amplitude image. The scanned area is approximately equal to the previous $15 \mu m \times 15 \mu m$ raster scans so 256 loops result in a similar tip velocity ($37.5 \mu m s^{-1}$) and spacing between adjacent scan lines as used in figures 1(b) and (f). Sensor inpainting diffuses data to the edges of the square grid circumscribing the collected data and pixels outside the scan region do not accurately depict sample properties. Figure 3(c) contains zoom-ins of the yellow dashed rectangle. Since all adjacent points are scanned in the same direction the quality for the reconstructed step on the feature is outstanding. The amplitude data further confirm the fidelity of the image as each scan segment has similar climbing or descending characteristics without offsets. No artifacts from the physical limitations of the $X, Y$ scanner are evident and 100% of the scan data are displayed. Using sensor data for sensor inpainting guarantees accurate image generation but it is important to mention that it is equally possible to perform inpainting algorithms on model based position data instead of measured position data from the sensors. This is analogous to open-loop scanning and applicable to any waveforms. Figure 3(d) shows the inpainted result when the piezo output voltage scaled by the first order piezo sensitivity is used as the position information to create the image. Using an Archimedean spiral, piezo hysteresis results in slight dilation of the center of the image and rotation along the scan direction. But still 100% of the scan time is used to create the image increasing temporal resolution by over a factor of two compared to raster scanning.

Another advantage of sensor inpainting over raster scanning is that sparse data can be more accurately depicted using inpainting. The number of loops, scan speed, and scan size determine the pitch of data spacing and the time required to collect each spiral scan image. Increasing the sampling rate increases the data density along the scan path but does not provide information between loops. In contrast with raster scan, the pixel size of the final rendered image is not determined by the data collection so it becomes important to determine the amount of inpainting required to most...
Figure 4. Comparison of ‘ground truth’ high resolution image (a) with images created by inpainting sparse data (c)–(e). The vertical columns are for three different sampling rates (I, II, III) leading to different numbers of data points for the images. (b) is the raw data with each point represented with a rectangle. (c), (d), and (e) are inpainted to $17 \times 17$, $34 \times 34$, and $68 \times 68$ pixels respectively. The image that most closely represents the ground truth is (d)II.

accurately render the data for sparse datasets. Figure 4(a) shows a high resolution spiral scan image with dense enough loops and samples to make it evident that the image is tip convolution limited. For this study, this image is ‘ground truth’. Subsequent images are from a spiral scan using eight times fewer loops over the same region and contain the features within the white square of the ground truth. Figure 4(b) shows the raw data at three different sampling rates. The number of data points used to create the images are 77, 207, and 831 in I, II, and III respectively. These images are not inpainted and instead a rectangle as wide as the spacing between samples and as tall as the pitch of the spiral scan represents each data point. This rendering is most analogous to the pixels in a raster scan image. The scan path from upper left to lower right is evident from the tilt of the rectangles. The three data sets with different sampling rates are inpainted to three different resolutions (figures 4(c)–(e)) in the following nine images for comparison to ground truth. Figure 4(c) uses $17 \times 17$ resolution and the pixels are too large leading to rough edges that do not capture the curvature of the features. Figures 4(d)I and (e)I have few data points and the positions of individual data points are evident in the image with blur in between. This is an artifact of the heat equation inpainting algorithm pinning the result where the data is good (fidelity term) anddiffusing information from the surrounding lower valued samples in between high valued samples causing a ‘tent pole’ effect. Tent poling is also evident in the high resolution image of moderately sampled data (figure 4(e)II). Figures 4(d)II and (e)II have dense sampling along the scan path so the tent poling is effectively a line and the surface looks like a sheet draped over two clothes lines. In all of these images the sagging between the data points does not capture the smooth curvature of the ground truth image. The middle image (figure 4(d)II) best represents the features shapes by capturing the smooth curvature and not misleading the eye with tent poling. Interestingly, the inpainted result which most closely resembles the ground truth has many more pixels than samples in the original dataset.

5. Conclusions

The raster scan paradigm severely limits scanning probe microscopy by dictating scan patterns and operation that is not well suited for piezoelectric nanopositioners. The results are significant expenditure of engineering effort and still a loss of at least half of the data when making images. Sensor inpainting breaks the raster scan paradigm by rendering accurate images from position sensor using missing data image processing algorithms and provides a software solution to a challenging hardware problem. Since most instruments of recent design have high-speed position sensors built into the scanner, implementation of sensor inpainting is simple. It enables the display of 100% of the scan data and alternate scan waveforms, like Archimedean spirals, that are best suited for the physical characteristics of the scanner. Sensor inpainting allows choosing the amount of pixels in the generated final image. Sampling data a factor of two higher in the fast scan direction and displaying on a grid with around twice as many pixels as samples produces the best representations of the data. Together these innovations will facilitate quantitative scientific investigation and discovery.

Acknowledgments

We gratefully acknowledge helpful discussions with Yifei Lou, Nen Huynh, Alex Chen, and Jen-Mei Chang. This work was supported by the National Science Foundation Cyber Enabled Discovery and Innovation under Contract No. 940417. AB is also supported by the W M Keck
Foundation and CB acknowledges support by the German Research Foundation DFG through project BU 2327/6-1. Data collection and instrumentation support funded by Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

[27] Zhang X and Chan T F 2010 Inverse Problems Imaging 4 191
[34] Kotsopoulos A G and Antonakopoulos T A 2010 Mechatronics 20 273
[38] Yong Y K, Moheimani S O R and Petersen I R 2010 Nanotechnology 21 365503