Measuring Discrete Feature Dimensions in AFM Images with Image SXM

Barry R. Bickmore1, Eric Rufe1, Steve Barrett2, and Michael F. Hochella, Jr.1

1Department of Geological Sciences, 4044 Derring Hall, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 USA
2Surface Science Research Centre, University of Liverpool, Liverpool L69 3BX, United Kingdom
(Received August 2, 1999; Published November 23, 1999)

Abstract

A suite of macros for the freeware image analysis program, Image SXM, are described. These macros are designed to measure the perimeter, horizontal area, and volume of discrete features in AFM images, obtaining accurate and consistent estimates. Directions for using the software and example applications are also given. Such tools allow one to perform tasks which would otherwise be extremely tedious or next to impossible. Examples include calculating reaction rates from time-series images of reacting particles or etch pits with complex shapes, and classifying objects based on their dimensions.

Keywords: Atomic Force Microscopy, image analysis, perimeter, area, volume, Image SXM

Introduction

In the past decade, a significant number of Earth scientists have successfully employed Atomic Force Microscopy (AFM) to characterize the surface morphology and reactivity of minerals and other environmental particles (for reviews, see Nagy and Blum, 1994; Hochella, 1995; Maurice and Lower, 1997). In this technique, the response of a scanned probe is translated into a digital, 3-dimensional map of a surface. Thus, the spatial dimensions of surface features such as etch pits, as well as small particles deposited on a substrate (e.g. clays) can be measured. The ability of the Atomic Force Microscope to operate under fluids, including aqueous solutions, also allows for the real-time characterization of the rates at which such features and particles grow or dissolve. Some features, such as euhedral etch pits, are conveniently characterized using the standard analysis software that comes with AFM systems, but others have complicated shapes, and require more sophisticated analysis.

Using its built-in macro language, we have customized the Image SXM image analysis environment to measure the perimeter, horizontal area, and volume of discrete surface features and particles in AFM images, obtaining accurate and consistent estimates. Image SXM (Barrett, 1997) is a spin-off of the popular public domain NIH Image analysis software (Rasband and Bright, 1995; Barrett et al., 1995; Liner, 1999), designed especially for the analysis of Scanning Probe Microscopy (SPM), Scanning Electron Microscopy (SEM), and Scanning Auger Microscopy (SAM) images. With respect to SPM images, Image SXM is able to read 3-dimensional image files created by most commercial SPM software packages, and subject them

Copyright ©1999 by the Mineralogical Society of America
to various analysis routines. Two aspects of this program are especially convenient for the measurement of irregularly-shaped features in AFM images. First, Image SXM allows the user to utilize rectangular, oval, polygonal, and freehand selection tools to define a region of interest (ROI) of any shape, thus excluding extraneous features. Second, it allows the user to create macros to perform most of the functions of the program automatically, and perform calculations, in a specified sequence.

Users of this software will now be able to quickly obtain reliable measurements of discrete features in AFM images. Such tools allow one to perform tasks which would otherwise be extremely tedious or next to impossible. Examples include calculating reaction rates from time-series images of reacting particles or etch pits with complex shapes (Bosbach et al., 1999; Rufe and Hochella, 1999), and classifying biomolecules and other surface features based on their dimensions (Chen et al., 1996; McMaster et al., 1996).

Algorithms

Perimeter and horizontal area

In order to estimate the perimeter or horizontal area of a discrete feature, the image must be transformed into binary format. That is, a certain height (or pixel intensity) level is chosen as the threshold, and pixels on either side of the threshold are changed either to black or white, depending upon which side of the threshold they fall. The black areas are treated as particles, and the area of each one can be determined by multiplying the number of pixels in the particle by a scaling factor. The perimeter of each particle may be calculated by adding 1 for each edge-touching pixel, and \( \sqrt{2} \) for each corner-touching pixel, and then multiplying by a scaling factor (Russ, 1990, pp. 183-185; Russ, 1995, pp. 520-522).

The difficulty in perimeter and horizontal area calculations is deciding at which height level to set the threshold, because, especially in the case of irregularly shaped features, the effect of altering the threshold even slightly can be drastic. In order to estimate consistent values from feature to feature and from image to image, one must decide upon some optimum threshold level at which to make the measurements. Therefore, our Image SXM macros threshold the image at 254 of the 256 possible gray levels (excluding white and black), and calculate the perimeter and area of the particles in the selected area at each setting. The perimeter vs. threshold height curve is then subjected to a 3, 5, 7, or 9 point (user-defined) smoothing routine, and the derivative of the perimeter vs. threshold height curve is calculated at each height level from the smoothed curve. The resulting derivative curve is essentially a map of image complexity vs. threshold setting. That is, perimeter values where the feature boundaries are the least complex will change the least from threshold level to threshold level, and so will produce values near zero in the derivative curve (Russ, 1990, pp. 108-115). Optimum threshold levels are selected where the absolute values of a string of 5 or more consecutive derivative points fall below a user-defined tolerance level.
Volume

The volume of a feature can easily be calculated by first defining a baseline height, and then multiplying the area of the selected ROI by the average pixel height (relative to the baseline height) within the ROI (Russ, 1995, p. 541). This method of volume calculation implicitly assumes that any noise in the image is distributed symmetrically about the true height values. It also ignores effect of AFM tip geometry and other artifacts on the perceived volume, although a user could modify the macro code to account for such things for specific applications.

The volume calculation macros define the baseline height by first creating a histogram of the number of pixels at each height level within the ROI. The histogram is then subjected to a 5-point smoothing routine, after which the height level of the baseline is determined by identifying the maximum of the first large peak from the bottom (in the case of particles) or top (in the case of pits). If there is a significant deviation from the overall baseline in an area within the ROI, a smaller peak in the histogram might interfere with the baseline determination. Therefore, the user is asked to define a minimum (particles) or maximum (pits) height level for the baseline to exclude such anomalies.

Directions for use

Image SXM macros are stored in text files, and are loaded into the program by selecting the “Load Macros” command under the “Special” menu, and selecting the desired file. The macros described here are in a text file named “pavmacro.txt”. The following macros are included in the file: 1) “Prepare Temp Window”, 2) “Particle Perim-Area”, 3) “Pit Perim-Area”, 4) “Plot Particle Perim”, 5) “Plot Particle Area”, 6) “Plot Particle P-Deriv”, 7) “Plot Pit Perim”, 8) “Plot Pit Area”, 9) “Plot Pit P-Deriv”, 10) “Particle Volume”, 11) “Pit Volume”, and 12) “Plot Smoothed Histogram”. This section describes procedures which can be used to utilize these macros to measure discrete features in AFM images.

Image preparation

In preparation for measurement, an image must be subjected to three processes, all of which can be performed using Image SXM or any standard AFM software. First, images must be subjected to a flattening routine, a least-squares polynomial fit to remove unwanted features from the scan lines. This removes variations in the baseline height from scan to scan in an image. Second, if the image is tilted with respect to the x-y plane, a tilt correction routine must be applied. Third, a median filter, which assigns each pixel the median height value of its immediate neighborhood (9 or 25 pixels) must be applied to reduce random noise. While operations such as lowpass filters may flatten out the edges of a feature and thus distort its shape, median filters do not, although some blurring does occur (Russ, 1990, pp. 46-48). Within Image SXM, flattening and tilt correction routines may be found under the “Compensation” category in the “SPM” menu. [Note: It is best to hold the “Shift” key down before using the mouse to pull down the “SPM” menu and select the tilt correction routine. This causes the program to readjust the z-scaling so that objects in the image are not cut off by being moved above or below the original z-scale.] A 3x3 pixel median filter routine can be found under the “Rank Filters” category in the “Process” menu.
Measuring perimeter and area

In order to run macros in Image SXM, one must first choose the “Load Macros” function under the “Special” menu and then select the desired macro text file. The names of the macros in the file then appear in the “Special” menu, and will run when selected.

To measure perimeter/area, one opens the desired image file and then creates a blank temporary image window with dimensions identical to those of the image, by running the “Prepare Temp Window” macro. The image window is automatically brought to the front. One of the selection tools is then used to define a ROI around the feature of interest. The user then runs either the “Particle Perim-Area” or “Pit Perim-Area” macro. The program then measures the perimeter and area of the feature at every height level.

After the perimeter/area calculations have been performed, the user is asked to provide values for the smoothing degree; i.e. the number of points averaged to smooth the perimeter vs. height curve before calculating the derivative. Smoothing degrees of 3, 5, 7, or 9 may be used. The user is also prompted for a perimeter picking tolerance. For instance, a tolerance of 0.005 (the default) means that only perimeter derivative values of 0±0.005 will be considered when the computer picks suitable height levels to measure perimeters and the corresponding areas. (The default values for smoothing degree and tolerance of 7 and 0.005, respectively, usually work well for reasonably simple features within images up to a few µm on a side. A lower smoothing degree might be required if, for instance, the user would like the program to pick a larger number of suitable height levels. A higher tolerance might be desirable if the horizontal area of the image is very large, and hence the derivative reflecting even relatively small changes in the calculated perimeter curve would also be correspondingly larger.) When the tolerance is entered, the program picks any suitable height levels at which to accurately estimate perimeter values, and then prints the corresponding heights and perimeters in the “Info” window. The user is asked whether the results should be saved, and if so the perimeter picks, as well as the area, perimeter, and perimeter derivative values at each height level are saved in a text file in a form easily accessible to a spreadsheet program for further analysis. Whether the file is saved or not, the user will be asked whether the computer should recalculate the results, and if so, the user is prompted for a new smoothing degree and perimeter picking tolerance, after which the derivative curve is recalculated and new perimeter picks are made. The user then has the option of saving the new results file.

In some cases the program will not pick any height levels at which to measure the perimeter, or will not pick any in the specific height region of interest to the user. In that case, the perimeter, area, and perimeter derivative data can be plotted vs. height level by running the “Plot Particle Perim”, “Plot Particle Area”, “Plot Particle P-Derv”, “Plot Pit Perim”, “Plot Pit Area”, and “Plot Pit P-Derv” macros. These take the data obtained by running the “Particle Perim-Area” or “Pit Perim-Area” macro and create an x-y plot in a separate window. One may obtain the exact coordinates of any point on a plot by running the mouse cursor across the plot window. The x value of the cursor position, and the y value of the plot corresponding to that x value, are displayed in the “Info” window. The user can employ these plots to find the most suitable heights at which to measure perimeter and area by eye.
Often it is useful to obtain the perimeter and area of a feature in a time-series. For instance, the authors have used this software to analyze lateral etch pit growth and clay mineral dissolution in time series of AFM images. Growth/dissolution rates can be measured and related to specific components of the surface area (e.g. the area of step edges). Since the assigned height of the baseline level in an image can vary somewhat due to noise, image drift, etc., it is not useful to pick a height at which to measure the perimeter/area and then measure those values at the same height level throughout the series of images. Rather, it is better to pick a suitable height level to make the measurements in the first image, note the specific feature of the perimeter vs. height plot corresponding to that height level, and then find the same feature in the perimeter vs. height plots in subsequent images to determine where to make the measurements.

**Measuring volume**

Volume measurements are made by selecting a ROI around the feature of interest, and running either the “Particle Volume” or “Pit Volume” macro. The user is then prompted to define a minimum (particles) or maximum (pits) height level for the baseline. After this is done, the baseline level chosen by the computer and the calculated volume are displayed in the “Info” window. It is useful to then run the “Plot Smoothed Histogram” macro, which creates an x-y plot of the smoothed number of pixels vs. height histogram. Again, by running the mouse cursor across the plot, the user can find the plotted value for the x value of the cursor position, which is displayed in the “Info” window. Using this plot, the user may determine whether the program estimated a reasonable value for the baseline height. If not, the volume macro may be run again, and a minimum or maximum baseline height may be entered to exclude a false result.

**Examples**

In this section, a few examples will be cited to illustrate both the usefulness of the software described here, and the nature of the results that can be obtained therefrom.

**Square pits**

Figure 1a shows our first example, which is a synthetic AFM image, created with the drawing and calibration tools included with Image SXM. The base level (i.e. the level above the pits) in this image is at 81 nm. In the center there are two concentric, square pits, the larger one at 48 nm height and the smaller at 11 nm. The larger pit is 2.00 µm on a side, while the smaller is 1.00 µm on a side. After drawing the pit pattern, we also used a macro to add random noise in a normal distribution about the original pixel values. In this case, the standard deviation of the noise distribution about the original pixel values is 5% of the total z-range (black to white, 100 nm). The dashed line around the pits marks the selected ROI for our calculations. Figure 1b shows a histogram of the number of pixels in the ROI corresponding to each height level.

Since this is a synthetic image, the values one should expect to measure with the macros are already known. When the “Pit Perim-Area” macro was run (smoothing degree = 7, tolerance = 0.005), the program picked two height levels at which to measure perimeters and areas: 35.3 nm and 68.2 nm. Figure 2 shows plots of the perimeter, perimeter derivative, and area vs. threshold height obtained by running the plotting macros described above. The dashed lines in the figure
denote the height levels the program picked. These plots show that the program picked height levels for measurement near the upper edges of the pits, but where the perimeter (and area) values change very little. Figure 3 illustrates this principle even more clearly. Here a cross section of the pits is shown, and again the dashed lines indicate the height levels picked. It is readily seen that heights were chosen near the upper edges of the pits, but just below where the noise in the image would complicate the measurements.

The program should be expected to measure perimeters and areas for the smaller pit in the neighborhood of 4 µm and 1 µm², and indeed, the measured values were 3.97 µm and 1.00 µm². The measured values for the larger pit, 8.02 µm and 4.00 µm², are similarly in agreement with the expected values, 8 µm and 4 µm².

When the “Pit Volume” macro was run, it correctly picked the baseline level at 81 nm (see Figure 4), and calculated a pit volume of 0.168 µm³. The expected value was 0.169 µm³. Thus, all of the measured values for perimeter, area, and volume are within 1% of the expected values. Certainly the more complicated shapes and imaging artifacts found in real AFM images would degrade the accuracy of these measurements, but this example serves to show that their most limiting factor is likely the quality of the data itself. That is, scanner drift, tip-sample convolution, and poor resolution of small features in an image would likely contribute more to the measurement error than anything associated with the computerized image analysis routines. For a discussion of various perimeter, area, and volume calculation routines, see Russ (1990; pp. 99-125, 181-191), and Russ (1995; pp. 507-522, 541-545).

Irregularly-shaped clay particle

Figure 5a is an AFM image of a montmorillonite clay particle fixed to a polyethyleneimine-coated mica substrate, taken under deionized water (see Bickmore et al., 1999). The image was subjected to flattening and median filter routines to remove variations in scan line height and random noise, respectively. The baseline height of the image around the particle is ~2 nm, and it exhibits two distinct terraces at ~8 nm and ~14 nm.

When the “Particle Perim-Area” macro was run, the program picked two heights at which to make measurements: 6.7 and 12.9 nm. Figures 5b and 5c show the binarized image, thresholded at these height levels. Careful comparison of these with Figure 5a reveals that if one were to trace the outline of the particle for measurement by eye, the true perimeter and area of the particle would likely be significantly overestimated. Figure 6 shows a cross-section of the particle, with the height levels picked by the computer algorithm marked by dashed lines. This figure illustrates the fact that the measurement routine picks height levels near the top of the terraces, to minimize the error associated with edge-broadening due to the pyramidal shape of the AFM probe tip, but just below where the rounded end of the probe tip, random noise, etc., begin to complicate the image.

The experimental artifacts just mentioned are made apparent in Figure 6, and hence it also serves to illustrate potential sources of error in the volume calculations. Although we cannot be certain of the “true” value of the particle volume, an “expected” volume was approximated using the measured heights and areas of the terraces, and the “Particle Volume” macro was run.
Whereas a volume of 1.21 x 10^{-3} \mu m^3 was expected, the routine calculated a volume of 1.35 x 10^{-3} \mu m^3, representing an error of ~12%. (The volume of the two small, white bumps on top of the particle in Figure 5a amounted to ~1 x 10^{-5} \mu m^3, and was taken into account in the calculation of the “expected” volume.) If errors of this magnitude are unacceptable, users may wish to customize this measurement routine (or create new routines) to account for artifacts specific to their images. For instance, the following references describe methods to remove tip-sample interaction artifacts, which may add significantly to the measured volume (Keller, 1991; Keller and Franke, 1993; Bonnet et al., 1994; Markiewicz and Goh, 1994, 1995; Wilson et al., 1995; Villarrubia, 1997). A simple routine of this type has been incorporated into Image SXM (menu SPM/Images/Tip Locus Effect), and is described by Barrett et al. (1999). In this routine the effect of a paraboloid tip with a user-defined radius of curvature is deconvoluted from the captured image. The image of the clay particle in Figure 5a was deconvoluted using this routine, assuming a radius of curvature of 30 nm for the tip. The calculated volume of the particle in the deconvoluted image was 1.26 x 10^{-3} \mu m, representing an error of only ~4%.

**Dissolving mica etch pit – time series**

Figure 7 is an animation of a time series of AFM images of some etch pits on a phlogopite mica surface, that are dissolving in pH 2 HCl (see Rufe and Hochella, 1999). A ROI was selected around the largest pit in each of these images, and the “Pit Perim-Area” macro run. Figure 8 consists of perimeter vs. threshold height plots for these images, in sequence from top to bottom. The dashed lines mark the threshold heights the computer program picked to measure perimeter and area values. This figure clearly shows that the baseline height of the image can change significantly in such a sequence, but the feature corresponding to the ideal measurement height in each perimeter vs. threshold height plot is easily identified.

Figure 9 is a plot of the pit volume (calculated by multiplying the measured pit area by the step height of 1.0 nm) vs. time. The lack of scatter in these data illustrates plainly the utility of having measurement tools to estimate accurate and consistent values for discrete feature dimensions in AFM images. In addition, phlogopite mica is a 2:1 phyllosilicate, which is known to dissolve inward from edge surfaces, rather than etching of the basal surfaces. The information gained from subjecting images such as these to the analysis routines described here can be used to obtain dissolution rates for these minerals, normalized to the reactive (edge) surface area. In fact, such information, obtained in a much more tedious fashion, has already been used in this way (Rufe and Hochella, 1999; Bosbach et al., 1999).

**Obtaining the software**

Links for downloading the latest version of Image SXM and the macros described in this paper from the World-Wide Web can be found at the following address:

[http://www.geocities.com/Athens/Parthenon/2671/Macros.html](http://www.geocities.com/Athens/Parthenon/2671/Macros.html)
Acknowledgements

This work was funded by grants from the National Science Foundation (EAR-9527092, EAR-9628023) and the Petroleum Research Fund of the American Chemical Society (PRF 31598-AC2, 34326-AC2). Additional funding was supplied by the National Science Foundation Graduate Fellowship Program and the American Federation of Mineralogical Societies Scholarship Foundation. Christopher Daniel and Frank Spear provided helpful reviews of the manuscript. E.R. would also like to thank K. Rosso for programming instruction during the early stages of this project.
References

Figure 1. a) Synthetic 256 x 256 pixel AFM image of two concentric square pits, 1 µm and 2 µm on a side. The z-range of the image (white to black) is 100 nm, and the three terraces are located at 11, 48, and 81 nm. Random noise has been added to the image in a normal distribution about the original pixel values. The dashed line indicates the ROI upon which the macro calculations were performed. b) Histogram of the number of pixels within the ROI corresponding to each grey level. Both a) and b) are digital captures of windows in the Image SXM program, taken during the measurement procedure.
Figure 2. a) Plot of perimeter vs. threshold height for the calculations performed on the ROI in Figure 1a. Dashed lines represent the threshold heights the program picked as ideal levels to measure perimeter and area. b) Plot of perimeter derivative vs. threshold height. c) Plot of area vs. threshold height. a), b), and c) are digital captures of windows in the Image SXM program, taken during the measurement procedure, with the dashed lines being added afterward.
Figure 3. The dashed line in a) denotes the area from which the data for the cross section of the square pits plotted in b) is taken. Both a) and b) are digital captures of windows in the Image SXM program, taken during the measurement procedure, with the dashed lines in b) being added afterward.
Figure 4. Plot of the same histogram as in Figure 1b, after being subjected to a five point smoothing routine. The “Pit Volume” macro generated the smoothed histogram, and used it to calculate the “baseline” height (dashed line) from which to calculate the volume of the pits. This figure is a digital captures of the data plot window in the Image SXM program, taken during the measurement procedure, with the dashed line being added afterward.
Figure 5. a) 1.7 x 1.7 μm (256 x 256 pixel) AFM image of a montmorillonite clay particle under deionized water. The particle has well-defined terraces at ~8 and ~14 nm height, with the baseline at ~2 nm. The dashed line represents the ROI used for the measurement routines. b) Binarized version of a), thresholded at 6.7 nm, the first level picked by the perimeter/area measurement routine. c) Thresholded at 12.9 nm, the second level picked. a), b), and c) are digital captures of image windows in the Image SXM program, taken during the measurement procedure.
Figure 6. The dashed line in a) denotes the area from which the data for the cross section of the montmorillonite clay particle plotted in b) is taken. The dashed lines in b) indicate the threshold levels at which the perimeter and area were measured. Both a) and b) are digital captures of windows in the Image SXM program, taken during the measurement procedure, with the dashed lines in b) being added afterward.
Figure 7. Animation of a sequence of four 880 nm x 880 nm (512 x 512 pixel) AFM images of the surface of a phlogopite mica crystal, taken under pH 2 HCl. The surface was pre-etched with HF, and the images were taken at 0 hrs., 14 hrs., 39 hrs., and 63 hrs. The perimeter and area of the large, 1 nm deep pit at the top of the image were measured through the sequence to determine the volume change of the pit over time. Click on the frame to play the animation.
Figure 8. Perimeter vs. threshold height plots for the large etch pit near the top of each of the frames in Figure 7. Dashed lines indicate the threshold height at which perimeter and area were measured. Clearly the baseline height of the images, and hence the “absolute” height of the ideal measurement threshold vary from frame to frame, due to changing imaging conditions, image drift, etc. However, the flat terrace in the perimeter vs. threshold height plots, corresponding to the edges of the pit, is clearly recognizable in each. a), b), c), and d) are digital captures of data plot windows in the Image SXM program, taken during the measurement procedure, with the dashed lines being added afterward.
Figure 9. Volume vs. time plot for the large pit near the top of the frames in Figure 7. The volume was calculated by multiplying the area of the pit by the 1.0 nm step height.