RESTORATION OF THE CANTILEVER BOWING DISTORTION IN ATOMIC FORCE MICROSCOPY IMAGES

S. A. Tsaftaris, J. Zujovic, and A. K. Katsaggelos

Department of Electrical Engineering and Computer Science, Northwestern University
2145 Sheridan Rd., Evanston, IL 60208, USA
phone: (+1) 847-491-3039, fax: (+1) 847-491-4455, email: {stsaft, jzu850, aggk}@eecs.northwestern.edu
web: ivpl.ece.northwestern.edu

ABSTRACT

Due to the mechanics of the Atomic Force Microscope (AFM), there is a curvature distortion (bowing effect) present in the acquired images. At present, flattening such images requires human intervention to manually segment object data from the background, which is time consuming and highly inaccurate. In this paper, an automated algorithm to flatten lines from AFM images is presented. The proposed method classifies the data into objects and background, and fits convex lines in an iterative fashion. Results on real images from DNA wrapped carbon nanotubes (DNA-CNTs) and synthetic experiments are presented, demonstrating the effectiveness of the proposed algorithm in increasing the resolution of the surface topography. In addition a link between the flattening problem and MRI inhomogeneity (shading) is given and the proposed method is compared to an entropy based MRI inhomogeniety correction method.

1. INTRODUCTION

AFM functions by bringing a cantilever tip in physical contact with (or close proximity to) the sample, revealing nanometer scale topographical information [1]. Fig. 1 shows a block diagram of an AFM. The repulsive force from the surface applied to the tip bends the cantilever. The amount of bending is measured and fed back to control the vertical movement of the sample in order to keep the contact force constant. The vertical movement follows the surface profile and is recorded as the surface topography. AFM is used to capture images of cells, materials, biomolecules etc. An example of an AFM image of DNA-CNTs is shown in Fig. 2.

Due to the mechanics of the AFM, there is a curvature distortion (bowing effect) present in the acquired images. This can be observed in Fig. 2, where intensities are low in the middle of the image while they are high at the sides. The tip follows arc-like lines in the image, creating a spherical or parabolic shape, depending on the scanner [2]. To compensate for this (known as line flattening or plane fitting), objects in an AFM image are manually labeled to generate an exclusion mask, which usually is rigid (either parallelogram or ellipse) and may not adequately represent the shape of the sought after object. The data of each row in the image outside the mask are fitted by a polynomial, which is subsequently subtracted from all the data values of the line.

With respect to the existing procedures, the processing of AFM images can be substantially improved for the following reasons:

(i) it is labor intensive and its automation is highly desirable;
(ii) the manual labeling of the objects in the image is highly inaccurate, since the objects are not easily distinguishable in the unflattened image and current software only allows for regular object shapes (i.e., circles, squares, etc);
(iii) with the existing line flattening techniques, the fitted polynomials are non-convex, in disagreement with the intrinsic physics, thus reducing the accuracy of the recovered surface topography.

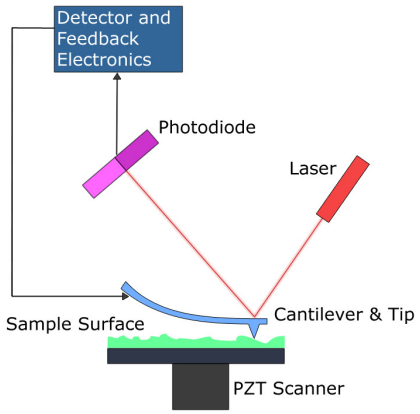

Fig. 1. AFM block diagram (WikiPedia Foundation).

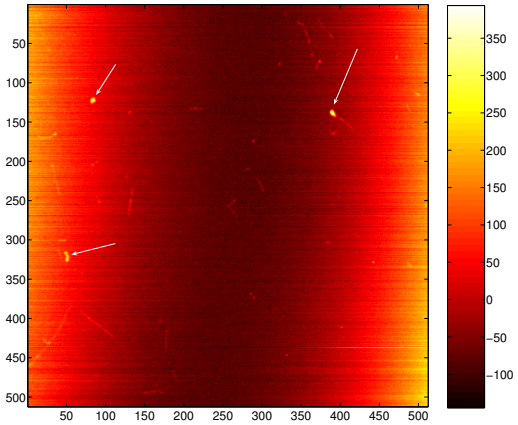

Fig. 2. AFM image of DNA-CNTs (indicated with 3 arrows).

In this paper, a method to automatically detect and exclude object points in a line and fit convex polynomials is presented. It is superior in comparison to the techniques used so far because it speeds up the process, uses less human resources, and produces more accurate results. In addition, a connection is made between the AFM problem and the shading (also referred to as MRI inhomogeneity, bias) correction encountered in magnetic resonance imaging [11].

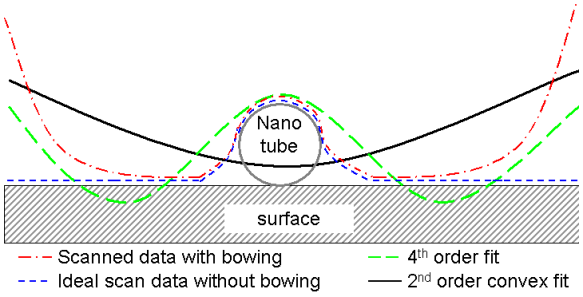

Fig. 3. Bowing effect illustration.

This paper is organized as follows: Section 2 provides the problem formulation and discusses possible solutions; Section 3 discusses the proposed algorithm in detail. Section 4 offers the results and discussion while Section 5 concludes with summary comments and possible future extensions.

2. BACKGROUND AND PROBLEM FORMULATION

To demonstrate the significance of convexity, an illustrative synthetic example of a nanotube imaged by an AFM is shown in Fig. 3. The dash-dot line represents the actual data recorded by the AFM, containing the bowing artifact, and the short-dashed line represents the recording if there were no curvature distortion. Due to tip convolution [1], there is blurring at the object boundaries. In an ideal scenario, the recorded height of the tube should be equal to the recorded width but due to the tip convolution problem the width is highly inaccurate. Therefore, the goal is to recover the height of the object while removing the bowing effect. The dashed line represents a non-convex polynomial least squares (LSQ) fitted to the recorded curve. Subtracting this polynomial from the recorded line would completely exclude the object in the flattened image, and the information of its height would be lost. Enforcing convexity will ensure a better approximation (solid line) of the curvature distortion, and also will avoid possible exclusion of the object in the flattened image. An example of a line from the image in Fig. 2 is shown in Fig. 4. Notice the high curvature issue in this image.

To the best of the authors' knowledge there is no direct literature on the topic of automatic object detection in AFM images; however, principles from other application areas could be applied. For example, if object points are treated as outliers, then model-fitting methods with outlier detection could be utilized. RANSAC (RANdom SAmples Consensus) [3] has the capability of dealing with a small number of outliers, but could easily fail when there are more object points (outliers) than background along the line. At the same time the method in [4], requires *a priori* knowledge of the expected maximum number of outliers, which implies knowledge of an object's shape, information not always available.

An apparent global 2D solution would be to fit parabolic surfaces on the whole image [5]. However, some artifacts in AFM images are intrinsic to the particular moment of operation (i.e. the scanning of each individual line), like vertical scanner drift, its internal non-linearities etc. Due to the presence of differing distortions and noise properties on each individual scan line, the bowing effect cannot be accurately modeled with a single 2D parabolic surface. In [6] a Gaussian Mixture Model is employed to model a smooth, and distinguishable from the signal, background. In AFM, the background is neither smooth (the necessity of processing each line is stressed previously), nor can it be assumed that the background and foreground pixels are distinguishable only by their values (as can be seen in the original AFM image in Fig. 2).

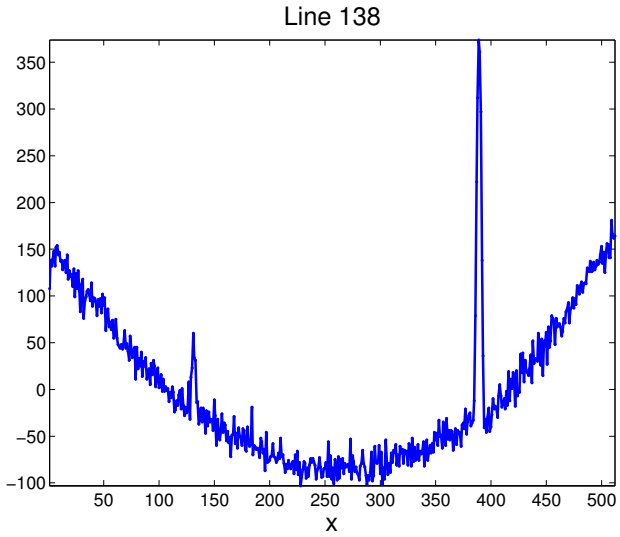

Fig. 4. Line 138 from the image in Fig. 2.

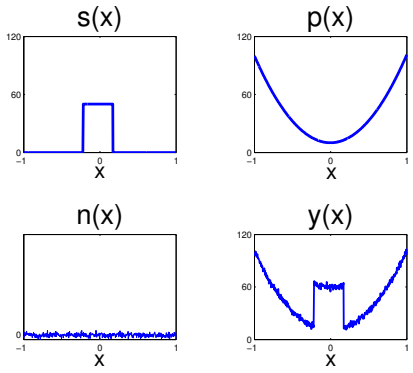

Fig. 5. Signal s(x), polynomial p(x), noise n(x) and y(x).

There is also an obvious relation to the MRI shading removal problem also known as MRI inhomogeneity [10],[11]. However, in [7], [8] either the noise is ignored, or a Gaussian signal distribution is assumed, and hence these methods are not applicable here. The algorithm presented in [9] can be applied to the 1D case (to accommodate the AFM problem), but histogram-based entropy estimation is not robust in the case of large object coverage, and the optimization routine does not enforce any convexity constraints. Nevertheless, in this paper the proposed method is compared with the method in [9] that has been modified to accommodate the AFM 1D case.

Each observed scan line can be modeled as:

$$y_i(x) = s_i(x) + p_i(x) + n_i(x),\ i=1,\ldots,N, \qquad (1)$$

where $y_i(x)$, $s_i(x)$, $p_i(x)$, $n_i(x)$ represent the raw data, signal, (convex) polynomial, and noise, respectively; x is the horizontal coordinate, i corresponds to the line number in the image, and N is the total number of lines. An example is shown in Fig. 4.

The overall objective is to

$$\text{estimate } s_i(x) \text{ given } y_i(x). \qquad (2)$$

To find signal points in the line one could resort to image based object detection or segmentation. However, the presence of noise and the limited *a priori* knowledge of the object's shape and size render both methods not applicable as general solutions.

Since $s_i(x)$ cannot be explicitly estimated, the problem can be best formulated as

$$\text{estimating } p_i(x) \text{ given } y_i(x), \qquad (3)$$

where $p_i(x)$ is a convex polynomial.

3. PROPOSED APPROACH

Each line in an AFM image is flattened iteratively in a two ste process: (i) K-means based classification into object and back ground points, followed by (ii) convex polynomial fitting on th background points. In more detail the steps are:

1. For x in *Background*, fit a convex polynomial $p_i^{(k)}(x)$ (is the iteration number) and subtract it from the data
$$z_i^{(k)}(x) = y_i^{(k)}(x) - p_i^{(k)}(x), \quad (4$$
to get the vector $z_i^{(k)}(x)$. Initially $y_i^{(1)}(x) \equiv y_i(x)$, an all points are considered as background.
2. Assuming a known noise variance σ_i^2, decide if the lin has objects based on:
$$\mathrm{var}(z_i^{(k)}(x)) > \alpha \cdot \sigma_i^2, \quad (5)$$
where α is a user given parameter and $\mathrm{var}(z_i^{(k)}(x))$ is the variance of $z_i^{(k)}(x)$ from step 1. If Eq. 5 is satisfied continue, otherwise go to step 6.
3. For all x, using the K-means algorithm, cluster the signal
$$c_i^{(k)}(x) = y_i(x) - p_i^{(k)}(x), \quad (6)$$
into *Object* and *Background* classes.
4. Increase the iteration number k, and for x in *Background*
$$y_i^{(k)}(x) = y_i(x). \quad (7)$$
5. Repeat steps 1 to 4.
6. No object data further detected, output the polynomial from the last iteration
$$p_i^*(x) = p_i^{(k)}(x), \quad (8)$$
and the x belonging to the final *Object* class.

For $i+1$ we could use $p_i^*(x)$ as the initial polynomial thus skipping the fitting part of step 1, however we have witnessed cases where the polynomial coefficients differ significantly from line to line, thus this initialization could be rather destructive. This also strengthens our argument as to the need for 1D processing.

In the following steps 1 and 3 are detailed, and noise variance estimation methods for step 2 are discussed.

3.1. Fitting Convex Polynomials

For fitting convex polynomials the following methods were tested: (a) *Constrained optimization* with Sequential Quadratic Programming (SQP) [12], or sum-of-square (SoS) polynomials [13]; however, SQP might not converge if the initial guess is not close to the solution, and SoS is computationally very intensive. (b) *Fitting ellipses* [14] which are easy and fast to implement; however, they are restricted to only second degree curves and hence may not adequately fit the boundary points. (c) *Direct least square fitting of convex polynomials* as in [15]; this method, however, did not yield accurate results. The SQP method was finally chosen for its accuracy and computational efficiency.

3.2. K-means clustering

The variance test in step 2 is chosen since it can detect even small objects. The larger the object, the worse the first fit, the greater the variance, thus detection confidence grows with the object size, which is expected. The flattened line in Eq. 6 is clustered in *Background* and *Object* classes. This is a 1D rather than a 2D clustering procedure, given that the points are clustered only by their $c_i^{(k)}(x)$ value, not their x position. The background points should have lower values than the object points, and therefore the cluster with the smallest centroid is marked as "background", while the one with the largest is labeled as "objects". Since K-means clustering is sensitive to initialization, the min and max of $c_i^{(k)}(x)$ are chosen to be the initial centroids. The points marked as objects are excluded from the signal and a new polynomial is fitted.

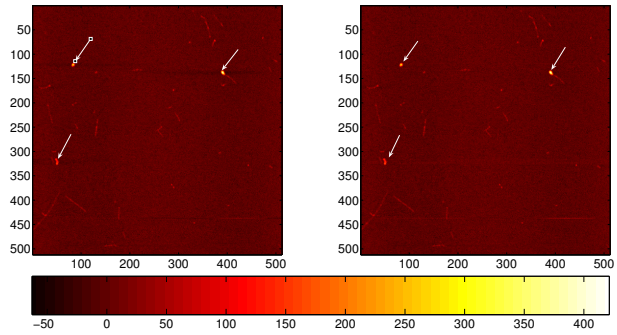

Fig. 6. Image in Fig. 2 (*left*) flattened with 3rd degree polynomials without any object detection; (*right*) flattened with 3rd degree polynomials using the proposed algorithm.

3.3. Noise estimation

Estimation of the noise variance is of critical importance. In the current incarnation of the algorithm, the noise variance is estimated on a per line basis. After the first fit, the polynomial is subtracted from the original data, and local variances are found using a sliding window. The location of the peak in the local variances histogram is used as the estimated noise variance. This method may not perform well if there is large object coverage in a line. In the opposite case a global approach is utilized, where each line in an image is first fitted with a polynomial, then the variance of each line is measured, and the line with the minimum variance is used with the above procedure to find an estimate of the noise variance. Alternatively, an approach similar to [16] can be used; however in this case difference operators are used to identify edges in images which are very susceptible to the noise in AFM images. As a final option, the user can supply a variance estimate.

4. RESULTS AND DISCUSSION

To obtain the results the algorithm of section 3 was implemented in MATLAB (The Mathworks, Inc). The optimization toolbox and the function *fmincon* (which utilizes an SQP solver) were used to fit convex polynomials of degree two to five. Initially a polynomial was LSQ fitted to the data. If this polynomial was not convex the *fmincon* routine was used to find a convex polynomial with initial starting coefficients of the LSQ fitted polynomial. The latter is used to ensure a good starting point, and improve the convergence of the SQP algorithm. Parameter α was chosen to be 2 (see section 4.2). In the subsequent sections, results on flattening Fig. 2 are first presented, followed by tests on synthetic raw line data to highlight the robustness of the algorithm. The proposed method is also compared with the mutual information method of [9] that has been modified. The modifications are presented in section 4.2.

4.1. Flattening of AFM images of DNA-CNTs

The image of Fig. 2 was used as a test image. 3rd degree polynomials were fitted to each line without any object detection (automatic or manual) and subtracted from each line, as shown in Fig. 6 (left). The height of the object at pixel location (139,390) (rightmost arrow) was 398 in the raw image while it was 413.8 in the flattened image. (Please notice in Figs. 2 and 6 that the minimum values are negative.) In Fig. 6 (right) the result of applying the proposed algorithm, with 3rd degree polynomials is shown.

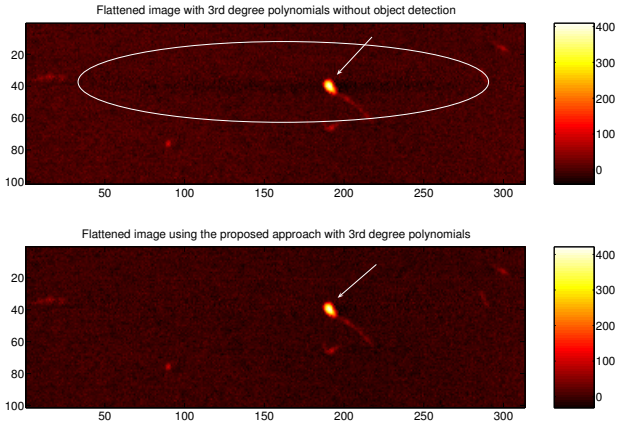

Fig. 7. (*top*), (*bottom*) detail of left and right image of Fig. 5 respectively, to highlight the ability of the proposed algorithm to segment the data and eliminate negative contrast regions (ellipse).

The height of the object at (139,390) is 425. The dark patches around the object in the top image, Fig. 6, are an artifact of bad fitting and introduce a non-existent dent-like deviation of the background. On the bottom image, Fig. 7, it is shown that these dark patches have disappeared, due to the automated object detection, thus better preserving the (relative) object height.

4.2. Flattening of synthetic AFM line data

The entropy minimization in MRI imaging is usually performed using Powell's multidimensional directional set method and Brent's one-dimensional optimization algorithm [9]. Since it is important to process the images line-by-line, as stressed previously, this algorithm was adapted to fit the 1-D case.

The objective function for minimization is the entropy of the line that resulted from subtracting the estimated polynomial from the original line. The optimization parameters are the coefficients of the polynomial. However, this method can determine the desired parameters up to the additive constant, since entropy takes into account only the distribution of values, and not the values themselves. To overcome this, another constraint had to be imposed. After finding the optimal parameters, the estimated polynomial is shifted so that the mean of the resulting line is positive, and the mean of absolute values of the resulting line is minimal. This ensures that the shifting constant is correct, even in the cases of large signal coverage.

Raw AFM lines were synthesized following a similar construct as in Fig. 5. The initial polynomials are of second degree, in accordance to the AFM physics [2]. Results obtained by the proposed algorithm on a synthetic $y(x)$ are shown in Fig. 8. In this figure it can be seen that: (i) $p^*(x)$ of 5th degree, calculated using only entropy minimization method (red solid curve) does not provide a very accurate estimate of the original $p(x)$ (green solid); (ii) the convexity constraint is necessary, since the 5th degree unconstrained curve (dotted line) completely obfuscates the object; (iii) the polynomial $p^*(x)$ fitted by the proposed algorithm of 5th degree with imposing convexity (dashed line) is relatively close to the original polynomial, and performing entropy minimization on that estimated polynomial gives a slightly better result (dash-dot line) and (iv) performing K-means classification after subtracting the estimated and adjusted polynomial (dash-dot) from the original data (solid black) accurately detects the object ('+' markers).

For more exhaustive experiments, various lines were generated with $p(x)$ a binomial $p(x)=ax^2+bx+c$, with (a, b, c) taking values from the set {(140.5, -0.5, -100), (60.5, -0.5, -50), (80.5, -70.5, -50)}, and a pulse $s(x)$ of height A {20, 75, 200, 400}, width

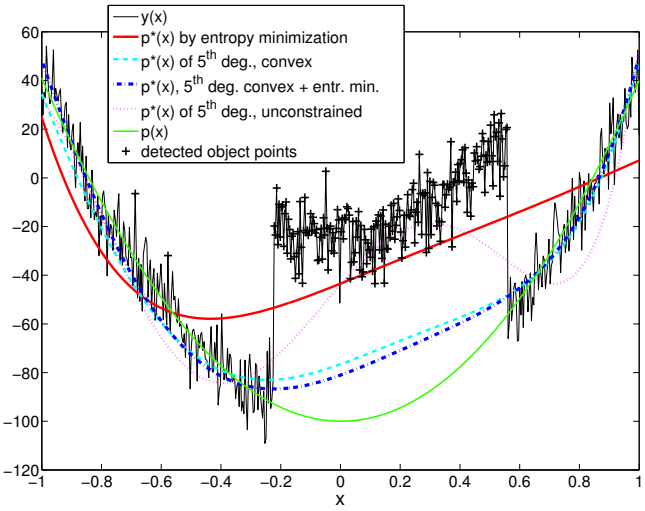

Fig. 8. An example line fitted with various polynomials.

w {10, 20, 100, 300} and delay L {50, 100, 200}. Gaussian random noise of zero mean and variance σ^2 {4, 16, 100, 400} was also added. For a given noise variance, while the other parameters remained the same, the algorithm was run 20 times by fitting 2nd and 5th degree convex and 5th degree unconstrained polynomials, as well as polynomials obtained with entropy minimization algorithm, initialized with LSQ fitted polynomial of 5th degree, and the 2nd and 5th degree convex polynomials that resulted from the proposed method. At each trial the height of the recovered signal defined as $s^*(x)= y(x)-p^*(x)$ for x in *Object*, was estimated by its average $A^*=E\{s^*(x)\}$. The final estimate A^* was the average of all 20 trials $A^*=E\{A^*\}$. Although the results are not shown the value of α=2 minimizes the ratio error and this value was used for all experiments.

In Fig. 9 the ratio of estimated height and pulse height (A^*/A) is shown as function of the ratio of pulse height and noise standard deviation (A/σ) fitted with the five polynomial types. It is evident that enforcing convexity is important since the unconstrained polynomials fail to recover the signal and hence register large errors. Also, it should be noted that the polynomials obtained by the entropy minimization method, when initialized with the resulting polynomial from our proposed method, perform better fitting in comparison to the proposed approach alone. However, this is not the case if the entropy minimization method is initialized by a less optimal solution.

The entropy minimization method depends greatly on the initialization parameters. Powell's search method and Brent's algorithm are not converging to the global minimum, since the search for optimal parameters is conducted only within the vicinity of the initial parameters. It is very likely that this optimization algorithm will converge to one of the local minima, and this is illustrated in Fig. 10. The mean square error between the original polynomial and estimated one was used as a measure for comparison, as a function of the ratio of pulse height and noise standard deviation (A/σ). The solid line represents the ratio of MSE for the polynomials found with entropy minimization, initialized with the resulting polynomial from our proposed method, and the MSE for polynomial estimated by the proposed method (denoted as p_k* on the graph). The dashed line represents the ratio of MSE for the polynomial estimated by entropy minimization method, but initialized with polynomial coefficients that are obtained by an LSQ polynomial fit, and MSE for p_k*. It is apparent that the MSE

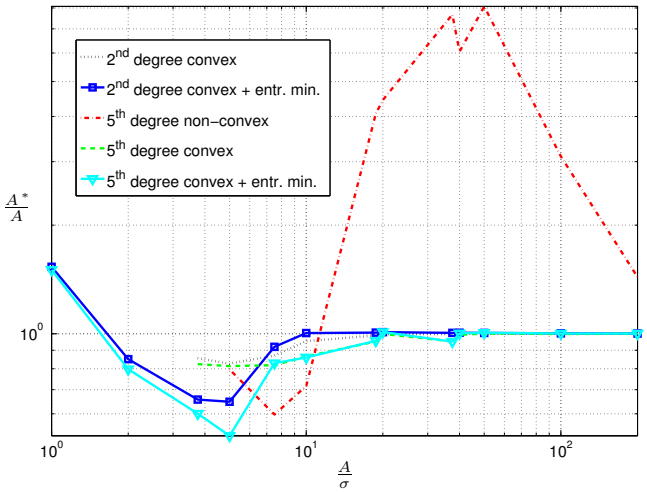

Fig. 9. The ratio A^*/A, for various polynomials, as a function of A/σ, fitted with 2nd and 5th degree convex, 5th degree unconstrained polynomials, and entropy minimization method initialized with 2nd and 5th degree convex polynomials

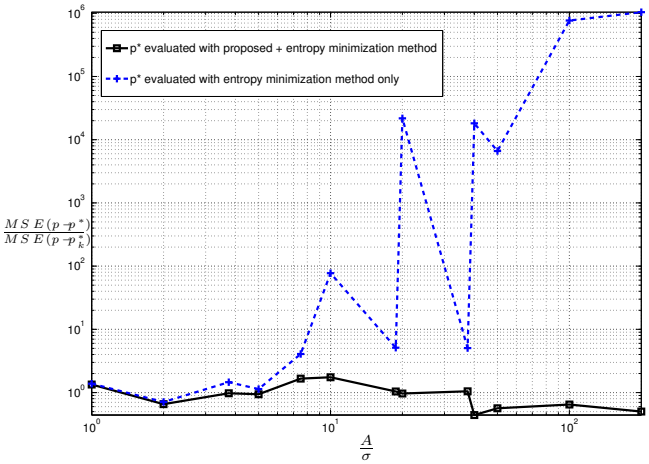

Fig. 10. The ratio A^*/A, for various polynomials, as a function of A/σ, fitted with 2nd and 5th degree convex, and 5th degree unconstrained polynomials.

depends on the initialization parameters and that in the case of bad initial guess, the procedure fails to yield satisfactory results.

Another disadvantage of the entropy minimization method is that it does not enforce convexity of the fitted polynomials, and thus does not follow the intrinsic characteristic of AFM imaging.

5. CONCLUSION

In this paper a novel algorithm for the automatic line flattening of AFM images was presented. The algorithm offers distinct advantages when compared to previous solutions, since it accommodates convex polynomials thus increasing the accuracy of the fit, and can automatically detect the presence of objects and segment them using a K-means algorithm. Results on real as well as synthetic data were presented that showed the robustness of the proposed method and its ability to resolve the relevant features of the objects. Line flattening is an integral part of AFM image analysis and precedes any other operation such as tip deconvolution, feature extraction, etc. Tests with more images will be performed in the future. In addition a new approach utilizing Bayesian estimation theory is under investigation [17].

5. ACKNOWLEDGMENTS

The authors would like to thank Prof. Hersam and Dr. Ostojic from the Dept. of Materials Science at Northwestern for providing the images used in this paper.

6. REFERENCES

[1] D. Sarid, *Scanning Force Microscopy*, New York: Oxford University Press, 1991.

[2] D. Ricci, P.C. Braga, "Recognizing and avoiding artifacts in AFM imaging," *Methods in Molecular Biology,* vol. 242, pp. 25-37, 2004.

[3] M. A. Fischler, R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp 381-395, 1981.

[4] S. Arora, S. Khot, "Fitting algebraic curves to noisy data," *J. Comput. Syst. Sci.,* vol. 67, pp. 325-340, 2003.

[5] M. Dai, T. S. Newman, C. Cao, "Least-squares-based fitting of paraboloids," *Patt. Rec.,* vol. 40, no. 2, pp. 504-515, Feb 2007.

[6] A. Saranli, "A Gaussian-mixture based approach to spatial image background modeling and compensation," in *Proc. of EUSIPCO,* 2007, pp. 1457-1461.

[7] J. Sled, A. Zijdenbos, A. Evans, "A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data," *IEEE Trans. on Medical Imaging*, vol. 17, no. 1, pp. 87-97, Feb 1998.

[8] A. Noe, S. Kovacic, J. Gee, "Segmentation of Cerebral MRI Scans Using a Partial Volume Model, Shading Correction and an Anatomical Prior," in *Proc. of SPIE*, 2001, pp. 1466-1477

[9] Z. F. Knops, J. B. A. Maintz, M. A. Viergever, J. P. W. Pluim, "Normalized Mutual Information Based Registration Using K-means Clustering and Shading Correction," *Medical Image Analysis*, vol. 10, issue 3, pp. 432-439, June 2006.

[10] M. Garcia-Sebastian, E. Fernandez, M. Grana, and F. J. Torrealdea, "A parametric gradient descent mri intensity inhomogeneity correction algorithm," *Pattern Recognition Letters*, vol. 28, no. 13, pp. 1657-1666, October 2007.

[11] Z. Hou, "A Review on MR Image Intensity Inhomogeneity Correction," *International Journal of Biomedical Imaging*, vol. 2006, Article ID 49515, 11 pages, 2006.

[12] R. Fletcher, *Practical Methods of Optimization*, Chichester: John Wiley and Sons, 1987.

[13] A. Magnani, S. Lall, S. Boyd, "Tractable fitting with convex polynomials via sum-of-squares", in *Proc. of the 44th IEEE CDC-ECC*, 2005, pp. 1672-1677.

[14] A. Fitzgibbon, M. Pilu, R. M. Fisher, "Direct least square fitting of ellipses," *IEEE Tran. on Pattern Analysis and Machine Intelligence,* vol. 21, no. 5, pp. 476-480, May 1999.

[15] D. J. Hudson, "Least-squares fitting of a polynomial constrained to be either non-negative, non-decreasing, or convex," *J Roy. Stat. Soc. B*, vol. 31, no. 1, pp. 113-118, 1969.

[16] K. Rank, M. Lendl, R. Unbehauen, "Estimation of image noise variance," in *IEEE Proceedings – Vision, Image, and Signal Processing,* vol. 146, no. 2, April 1999.

[17] W. M.Wells III,W. E. L. Grimson, R. Kikinis , F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429–442, 1996.