DETECTION OF ELLIPTICAL PARTICLES IN ATOMIC FORCE MICROSCOPY IMAGES

Jiří Sedlář [a,b]   Barbara Zitová [a]   Jaromír Kopeček [c]   Tatiana Todorciuc [c]   Irena Kratochvílová [c]

[a] Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation,
Prague, Czech Republic – sedlar@utia.cz, zitova@utia.cz
[b] Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic
[c] Academy of Sciences of the Czech Republic, Institute of Physics, Prague, Czech Republic –
kopecek@fzu.cz, todorciuct@yahoo.com, krat@fzu.cz

ABSTRACT

In this paper we describe a method for detection and measurement of elliptical particles in atomic force microscopy (AFM) images. AFM imaging is used in physics to scan surfaces; the measured heights are represented by pixel values. Each sample in our project consisted of elliptical particles of principally the same size; the size could be characterized by the average length and width of a number of salient particles. The method we proposed is based on segmentation of undamaged particles and their approximation by ellipses; the major and minor axes provide robust estimates of the lengths and widths of the particles, respectively. The method is robust to distortions typical of AFM images. Its performance was demonstrated on images of pyrroles and compared with manual detection. Results show that the automatic method could be used in place of the time-consuming manual detection.

Index Terms— particles detection, atomic force microscopy (AFM) imaging, watershed segmentation, image moments, approximation by ellipses

1. INTRODUCTION

Image processing is a progressive discipline with applications in many scientific domains, including surface physics. While imaging sensors allow us to scan the surfaces of materials, image processing is utilized for interpretation of the acquired data. Our study focuses on analysis of atomic force microscopy (AFM) images [1].

AFM images are affected by various artifacts. The acquired image is, in fact, a convolution between the cantilever tip and the surface of the sample; only details bigger than the diameter of the tip are observed. The exact shape of the tip is usually unknown; it can even change during the measurement due to breaking or due to adhesion of dust particles or parts of the sample. The scanned image is, furthermore, affected by various noise patterns. The image analysis method should be robust to these distortions.

The goal of our project was to measure the characteristic size of particles in AFM images of phenylpyridyldiketopyrrolopyrrole (PPDP), a new organic semiconducting material for hydrogen sensors [2]. The deposited layers of PPDP consisted of ellipsoidal particles with approximately the same size; the size of particles reflects the quality of the sample.

The existing AFM image analysis methods were not convenient for our purposes. Methods used for automatic detection of particles in AFM images are sensitive to the level of noise, while manual detection is very laborious and time-consuming – the analyzed images contained about one hundred particles each. We took advantage of the elliptical shape of particles to develop an automatic method that is robust to noise; the method approximates the shapes of salient particles in the AFM image by ellipses.

The paper is organized as follows: Section 2 contains definitions of image moments that are used in Section 3, which describes our method for detection and measurement of elliptical particles in AFM images; Section 4 compares results of the proposed automatic method and of a manual detection by an expert; and Section 5 discusses possible future improvements to the proposed method.

2. MATHEMATICAL BACKGROUND

The geometric moment

$$m_{pq} = \iint\limits_{D} f(x,y)x^p y^q dx dy , \qquad p,q \in \mathbb{N}_0$$

is a projection of an image function f with a compact support $D \subseteq \mathbb{R}^2$ onto a polynomial basis, where $(p+q)$ is the order of the moment [3]. Translation of the image by vector $(-x_t, -y_t)$, where $x_t = m_{10}/m_{00}$ and $y_t = m_{01}/m_{00}$ are centroid coordinates, creates a centered image.
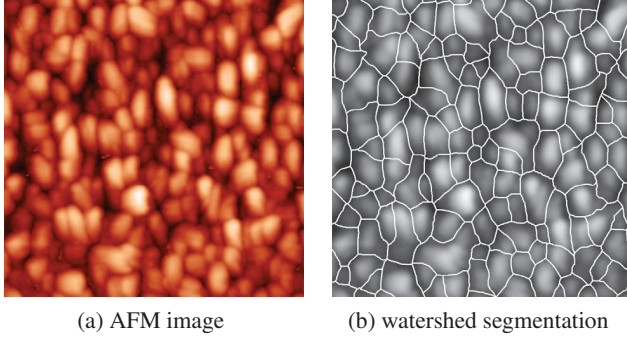
(a) AFM image   (b) watershed segmentation

Fig. 1. *sample A:* (a) Atomic force microscopy (AFM) image of a PPDP sample. (b) Watershed segmentation of the denoised image.

(a) normalized brightness   (b) core parts of particles

Fig. 2. *sample A:* (a) Normalization of brightness within watershed regions. (b) Core parts of particles after thresholding.

The central geometric moment is defined as

$$\mu_{pq} = \iint\limits_{D} f(x,y)(x - x_t)^p (y - y_t)^q \, dx \, dy$$

Rotation around zero by angle

$$\alpha = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right)$$

(provided that $\mu_{11} \neq 0$ or $\mu_{20} - \mu_{02} \neq 0$) transforms the centered image to a normalized position. The values of second order moments in the normalized position satisfy conditions $\mu'_{11} = 0$ and $\mu'_{20} \geq \mu'_{02}$; their values can be expressed as

$$\mu'_{20} = \frac{1}{2}\left(\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\right)$$

$$\mu'_{02} = \frac{1}{2}\left(\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\right)$$

3. METHOD

The aim of our method is to detect salient particles in an AFM image (see Fig. 1a) and measure their size. The method works with the assumption that particles in the image are approximately elliptical and of principally the same size; the shape of an elliptical particle can be described by its length and width. The size of particles in the image can thus be characterized by the average length and width of a statistically significant number of particles. The method is based on detection of separate particles and approximation of their shapes by ellipses; the major and minor axes provide robust estimates of the lengths and widths of the particles, respectively. Particles that are partly overlapped or significantly damaged by noise are not included in the measurements so as not to skew the results.

The method consists of the following steps: first, the AFM image is preprocessed to obtain a grayscale image with suppressed noise; the image is then segmented by watershed, and regions with overlapped or damaged particles are removed;
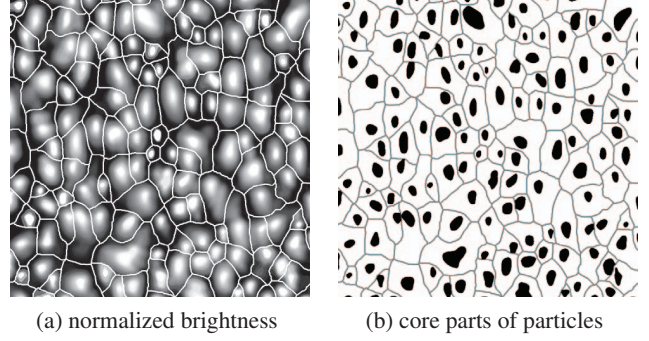
in the remaining regions, the shape of the particle is approximated by an ellipse, and the error of this approximation is estimated; finally, the lengths and widths of particles with low approximation errors are measured.

3.1. Preprocessing

AFM images contain information about the measured height of the scanned surface. Some imaging software applications, however, yield artificially colored images where the correspondence between the height and the color is coded by a color map. In such a case, the method transforms the colors in the image to brightness according to their position in the color map; the brightness values in the resulting grayscale image represent the measured heights of the surface.

For the purposes of segmentation, the level of noise has to be suppressed. The most critical disruption is high-frequency noise, which can be removed by a low-pass filter, e.g. convolution with a Gaussian mask. Although a part of high-frequency data information is discarded, the proposed method is robust to this loss.

3.2. Segmentation

The denoised image is segmented into regions corresponding to separate particles (see Fig. 1b). The segmentation is performed by a watershed algorithm [4]. High frequencies in the noise would cause over-segmentation, but the smoothening with a low-pass filter significantly reduces the number of falsely detected regions. Particles that overlap one another are, however, merged together.

The next step of the method consists of detecting regions that should be excluded from measurements. Brightness in the denoised image is normalized separately in each region (see Fig. 2a), and core parts of particles are computed by thresholding with a threshold value common to all regions (see Fig. 2b). Very small cores correspond to noise artifacts or particles deep in the scanned layer; cores joined with watershed or image borders belong to partly overlapped particles.
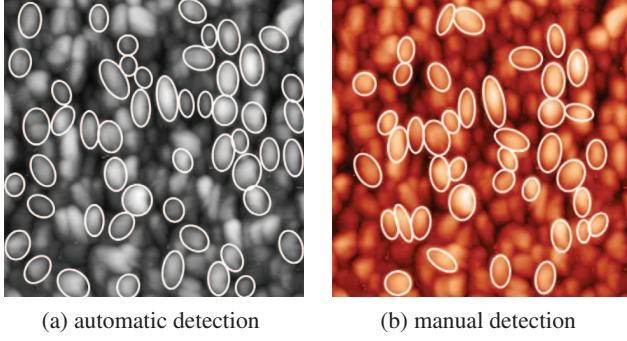
| (a) automatic detection | (b) manual detection |

Fig. 3. *sample A:* (a) Automatic detection and approximation of particles by ellipses. (b) Manual detection of particles by a physicist.
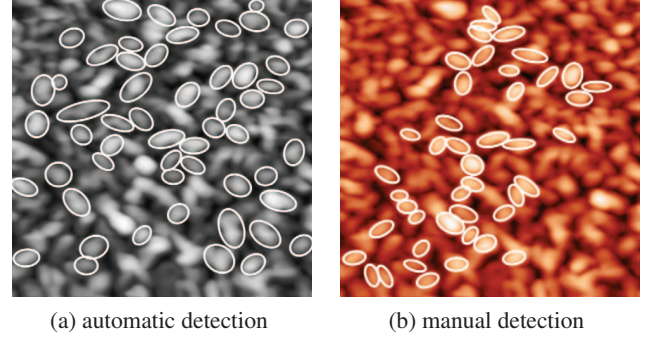

| (a) automatic detection | (b) manual detection |

Fig. 4. *sample B:* (a) Automatic detection and approximation of particles by ellipses. (b) Manual detection of particles by a physicist.

Regions with such cores are therefore eliminated from further processing.

3.3. Approximation by ellipses

The particle in each of the remaining regions is approximated in the coordinate system of the normalized position of the region by an ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \, , \qquad a \geq b > 0$$

– we take advantage of the elliptical shapes of particles in the analyzed images. The semiaxes are computed as

$$a = 2\sqrt{\mu'_{20}} \, , \qquad b = 2\sqrt{\mu'_{02}}$$

with the brightness of the denoised image as the image function f, and the region as its support D (see Section 2); this approximation is similar to the reference ellipse in image moments theory [3]. The major axis $2a$ and minor axis $2b$ provide robust estimates of the length and width of the particle, respectively. The parts of the region outside the ellipse could skew the measurements because they usually do not belong to the particle. In order to refine the shape and position of the ellipse, its parameters are computed again, this time over the intersection of the region and the original ellipse (see Figs. 3a and 4a).

3.4. Approximation by ellipsoids

The degree to which a particle is damaged can be described by its difference from an ellipsoid; we take advantage of the ellipsoidal shapes of particles in the scanned samples. The surface of the segmented particle is therefore approximated by the upper half of an ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \, , \qquad a, b, c > 0 \,\,\&\,\, z \geq 0$$

(see Fig. 5) where the semiaxes a and b lie in the image plane and equal the major and minor semiaxes, respectively, of the

approximating ellipse. The position $[x_t, y_t]$ and orientation α of the ellipsoid in the image plane are also determined by the approximating ellipse. The length of the semiaxis perpendicular to the image plane is estimated as

$$c = f(x_t, y_t)$$

The height of the approximating ellipsoid can be expressed as

$$z(x, y) = \begin{cases} c\left(1 - \left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)\right)^{\frac{1}{2}} & \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \\ 0 & \frac{x^2}{a^2} + \frac{y^2}{b^2} > 1 \end{cases}$$

The approximation error is computed as

$$\epsilon = \frac{1}{\sum\limits_{[i,j] \in R} z_{ij}} \sum_{[i,j] \in R} z_{ij}(f_{ij} - z_{ij})^2 \qquad (3.1)$$

where f is the brightness of the denoised image within the processed region R. If the error is large, the ellipsoid is significantly different from the measured surface, usually because of overlaps or noise artifacts. Only particles with low approximation errors are included in measurements.

3.5. Measurements

The length and width of a particle are estimated by the length of the major axis $2a$ and the length of the minor axis $2b$ of the approximating ellipse, respectively. The size of particles in the AFM image can be robustly represented by the average length and width of particles with approximation errors ϵ (see (3.1)) lower than a threshold T_ϵ. The value of the error threshold T_ϵ can be computed adaptively so that a sufficient number of particles is used in the measurements.

4. RESULTS

We tested the proposed method on AFM images of PPDP (see Section 1). To avoid damage of the soft organic material, the semicontact (tapping) mode was selected; the measurements were performed on the Ntegra Prima apparatus (NT-MDT, Russian Federation). The images were stored in a 24bit
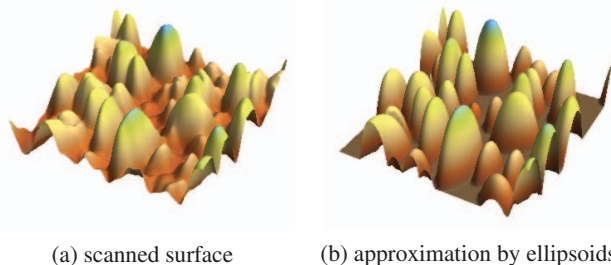
(a) scanned surface   (b) approximation by ellipsoids

Fig. 5. (a) The surface of a blurred AFM image. (b) Approximation of salient particles in the image by ellipsoids.

(a) lengths   (b) widths

Fig. 6. *sample A:* Distributions of (a) lengths and (b) widths of automatically (solid) and manually (dashed) detected particles.

RGB color format of 721×721 pixels; the corresponding 24bit RGB color map consisted of 256 levels. We applied the method described in Section 3 to the images. For the purposes of comparison, the same AFM images were analyzed by a physicist who manually approximated salient particles by ellipses (see Figs. 3b and 4b).

For each AFM image, we compared the lengths and widths of the automatically detected ellipses with the lengths and widths of the ellipses manually detected by the expert (see Fig. 6). Table 1 illustrates the similarities in the results of these two methods on two examples. The lengths and widths are characterized by their mean values $\overline{2a}$, $\overline{2b}$, respectively, with standard deviations $\text{std}(2a)$, $\text{std}(2b)$, respectively; n denotes the number of measured particles, i.e., in case of the automatic method, particles with $\epsilon < T_\epsilon$ (see Section 3.5). The ellipses detected by the automatic method were very similar to the manually detected ones – for each image, the differences between the mean values from the automatic and the manual method were within the corresponding standard deviations. The results indicate that the automatic method could be used in place of the manual measurements.

sample, method	n	$\overline{2a} \pm \text{std}(2a)$	$\overline{2b} \pm \text{std}(2b)$
A – automatic	52	74 ± 14	50 ± 8
A – manual	41	79 ± 14	51 ± 8
B – automatic	54	66 ± 15	42 ± 7
B – manual	43	65 ± 14	36 ± 6

Table 1. Comparison of the lengths $2a$ and widths $2b$ (in pixels) of particles detected by the proposed algorithm (automatic) and by a physicist (manual) in sample A and in sample B.

5. CONCLUSIONS

We proposed a method for detection and measurement of elliptical particles in atomic force microscopy (AFM) images. The detection is based on watershed segmentation and approximation of particles by ellipses; the length and width of a particle are estimated by the major and minor axes of the ellipse, respectively. Before the segmentation, high-frequency noise in the AFM image was suppressed by a low-pass filter.
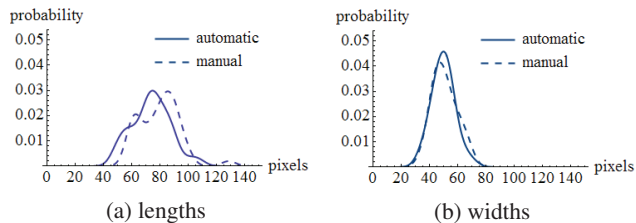
As the size of particles was principally the same within each sample we analyzed, the average length and width could be computed just from salient particles; regions with overlapped particles or particles deep in the scanned layer were therefore removed. The remaining particles were approximated by ellipses, whose parameters were estimated by means of image moments. Particles with non-ellipsoidal shapes were excluded from the measurements.

Performance of the method was demonstrated on AFM images of phenylpyridyldiketopyrrolopyrrole (PPDP), a semi-conducting material. The sizes of automatically detected particles were compared with the sizes of particles detected manually by a physicist; the differences between the mean values were within the corresponding standard deviations. The results showed that the proposed automatic method could replace the laborious, time-consuming manual measurements. The main advantage of the proposed method is its robustness to high-frequency noise. Possible future improvements include further research on the approximation error and on the shape of the approximating ellipsoid.

6. REFERENCES

[1] V.L. Mironov, *Fundamentals of SPM*, Russian Academy of Sciences, Institute for Physics of Microstructures, Nizhniy Novgorod, 2004.

[2] O. Salyk, J. Vyňuchal, I. Kratochvílova, T. Todorciuc, J. Pavluch, and P. Toman, "Study of phenylpyridyldike-topyrrolopyrrole interaction with hydrogen in gas and in acids," *Dyes and Pigments*, vol. 207, no. 10, pp. 2327–2333, 2010.

[3] J. Flusser, T. Suk, and B. Zitová, *Moments and Moment Invariants in Pattern Recognition*, John Wiley & Sons Ltd, Chichester, 2009.

[4] V. Osma-Ruiz, J.I. Godino-Llorente, N. Sáenz-Lechón, and P. Gómez-Vilda, "An improved watershed algorithm based on efficient computation of shortest paths," *Pattern Recogn.*, vol. 40, pp. 1078–1090, 2007.