Chapter 7
Data Representation and Image Processing

Scanning probe microscopy data usually have the form of a matrix where the topography (height) or some other signal such as the tunneling current, or dI/dV is measured as a function of the lateral xy-position on the surface. Data representation is the task to map the measured heights (DAC values) to gray levels in an image in an optimal way. Image processing is used in order to enhance the image representation further, i.e. by removing image artifacts such as high-frequency noise, noise pixels or noise lines.

7.1 Data Representation

A data representation using 8-bit or 256 gray levels (ranging from 0 (black) to 255 (white)) is more than sufficient, since the human eye can distinguish only less than one hundred gray levels. These data are displayed as an image of typically 512 × 512 pixels.

The original data on the height of the tip (z-output signal of the digital feedback loop) are usually acquired by digital-to-analog converters (DAC) with a certain resolution. In the following, we consider 16-bit converters as an example (≈65,000 levels), while nowadays 24-bit DACs are available. The task for data representation is now to efficiently map the data, which cover a certain range of the 65,000 levels (DAC units), to the 265 gray levels. This task is also called background subtraction. As an example, we will discuss this first for one scan line. However, the same strategies apply for a whole image. As a convention for the gray levels black is assigned to the lowest height and white to the highest. If one were to map the 16-bit data range linearly from the lowest level to the highest level to the 8-bit gray scale from black to white not much of the surface structure would be visible. One scan line usually covers only a small range of the 65,000 levels. As an example, the scan line shown in Fig. 7.1 contains a range of about 800 height levels (DAC units). If the 256 gray levels were mapped to the complete range of 65,000 digital-to-analog converter (DAC) levels, (level 0 is black and level 65,000 is white) a range of 65,000/256 = 256

B. Voigtländer, *Scanning Probe Microscopy*,
NanoScience and Technology, DOI 10.1007/978-3-662-45240-0_7

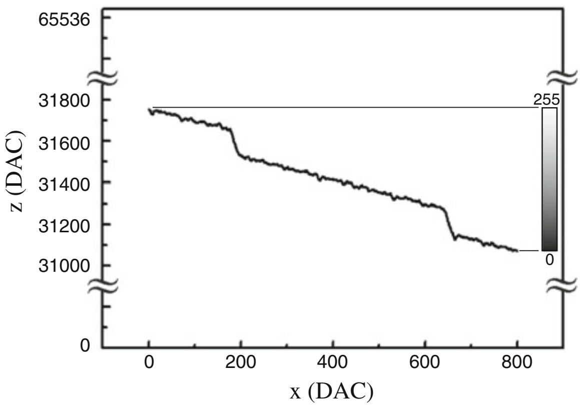

Fig. 7.1 For a good data representation the 256 *gray levels* have to be mapped to the 65,000 DAC levels in a proper way

height levels would be mapped to one gray level. It is clear that most of the information contained in the original data is lost by this poor mapping. For our scan line in Fig. 7.1, the 800 height levels in which the image information is contained would be mapped to only 3 gray levels ($800/256 \approx 3$). Therefore, the gray scale should be mapped to a smaller range of the 65,000 digital-to-analog converter (DAC) levels which contain the (height) data of the scan line, as shown in Fig. 7.1.

Another effect is that the actual topographic data are often hidden due to the quite large slope of a scan line. This slope arises because the scanning plane is usually tilted slightly with respect to the sample. This tilt occurs due to an imperfect alignment of the sample relative to the coordinate system of the scanning piezo element. In the following, we term this the scanning slope, which can be as large as several degrees. This scanning slope shows up as a tilted base line in the data as shown in Fig. 7.1. Usually, and specifically in atomically resolved images, the measured height range is very small (only a few Å), and the range of the measured height data is dominated by the scanning slope. Here we give two quantitative examples in which we consider a relatively large tilt angle between surface and scanner of 3°. If we consider an image of the size of 1 μm the height difference induced by this slope across the image is $\Delta h = \Delta x \tan \alpha \approx 500$ Å. This 500 Å on an image size of 1 μm corresponds to a scanning slope which will be present in all images. If we consider, on the other hand, that we have as the image signal, for instance, 5 atomic steps, each of 3 Å height, the image signal we want to measure (15 Å) resides on a scanning slope of 500 Å. This means that the background height change due to the slope is 30 times larger than the image signal (the steps). In a second example, we take an atomically resolved image of a size of 500 Å, corresponding to a height difference due to the background slope of 26 Å. If the atomic corrugation on a single atomic terrace is 1 Å the signal to background ratio is 1/26 in this case.

We have seen that even a small tilt between sample and scanner leads to a substantial slope in the images. This slope can be eliminated by a background subtraction. This is usually done by fitting a straight line to the data of each scan line and by displaying only the deviations of the data with respect to this fit, as shown in Fig. 7.2c, d.

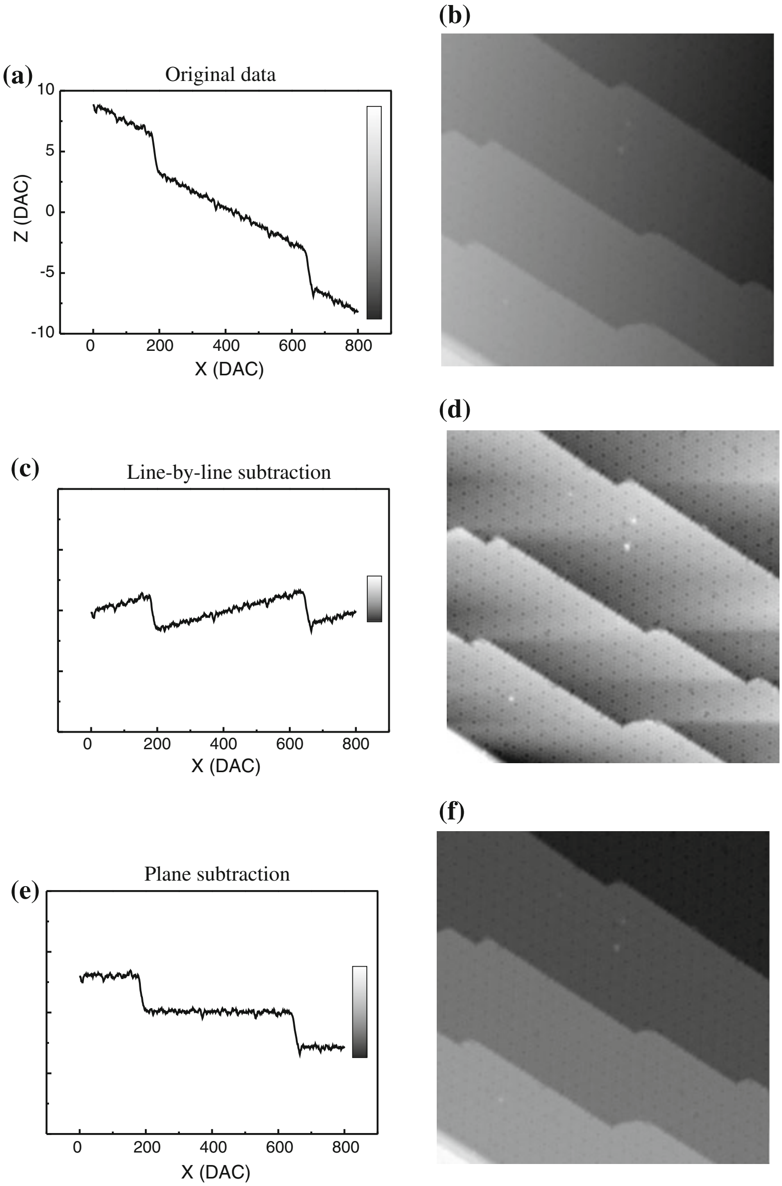

Fig. 7.2 STM data taken on a stepped Si(111) surface with the atomically resolved (7 × 7) reconstruction contained in the data. Comparison of different kinds of background subtraction for a single scan line (*left panel*) and a whole image (*right panel*). **a** and **b** Show the original data without background subtraction. In **c** and **d** a line-by-line background subtraction was applied. In **e** and **f** a plane subtraction relative to one of the terraces, between steps of a single atom height, was applied. The image size is 600 Å. In this image, the scanning slope corresponds to an angle of 0.7° between sample and scanner

This background subtraction increases the contrast in the image, but also leads to artifacts like the black shadows which can arise due to some higher parts of the scan line which pull the fitted line up. The next higher approximation is to use a fit to a quadratic function as background. This can also remove the part of the background that arises from the scanner bow in large scans. This scanner bow arises because the xy-motion induced by the tube scanner is approximately a motion on a sphere with a radius of the piezo tube length.

Another kind of background subtraction is not taking each line individually into account, but the whole matrix of measured data as one entity. Here the obvious approaches are to fit a plane or square function (paraboloid) to the data for background subtraction. Another approach is that the user can define points in an image which are known to belong to one specific height (for instance one atomic terrace). The background subtraction is then performed relative to this user-defined plane. An example of this background subtraction relative to a user defined plane is shown in Fig. 7.2f. The different methods of background subtraction each have their advantages and disadvantages. The advantage of the (user-defined) plane subtraction is that locations of the same height on the surface are displayed by the same gray level. The advantage of line-by-line subtraction is that the contrast is higher and the small height corrugations due to the atomic structure of the Si atoms are more easily visible. As another variant the whole contrast range from black to white can be used for one atomic terrace, leaving however all lower terraces black and all higher ones white. This is also called clipping. If you see larger areas in an image either white or black, the real data are outside the contrast range and are clipped to black or white.

Apart from the gray scale images considered so far, it is, of course, also possible to use color in the image representation. In the false color representation, the 8-bit gray scale palette is replaced by a color palette. The most popular one is the fire palette ranging from black via red and yellow to white. In Fig. 7.3a a gray scale representation (subtracted line-by-line) of a stepped Si(7×7) surface is used, while in Fig. 7.3b a false color representation with the fire palette is used. In Fig. 7.3c a plane subtracted representation of the same image is shown in gray scale and false color representation using a palette with several colors is shown in Fig. 7.3d. Here the palette was chosen such that each terrace has a specific color. In Fig. 7.3e a 3D image representation of the same image is shown. Here techniques like rendering and ray tracing are used to give a plastic impression of an actual three dimensional landscape of the measured data. While such images look like the real morphology of a landscape it must be kept in mind that the z-scale in SPM images is almost always quite exaggerated relative to the lateral scale. For the example in Fig. 7.3e, the z-scale in the image is only 12 Å, while the image size is 600 Å. Going one step further a fly-by movie through the atomic or nano canyons at the surface can be generated. With all these different kinds of image representations it should not be forgotten that they are only different representations of the same initial data matrix. The appropriate image representation should always be chosen for the respective purpose. An elaborated image representation with a lot of colors may be well suited to impress laypeople but may obscure the visibility of important details. Therefore, a simple gray scale representation is often sufficient to convey the scientific information.

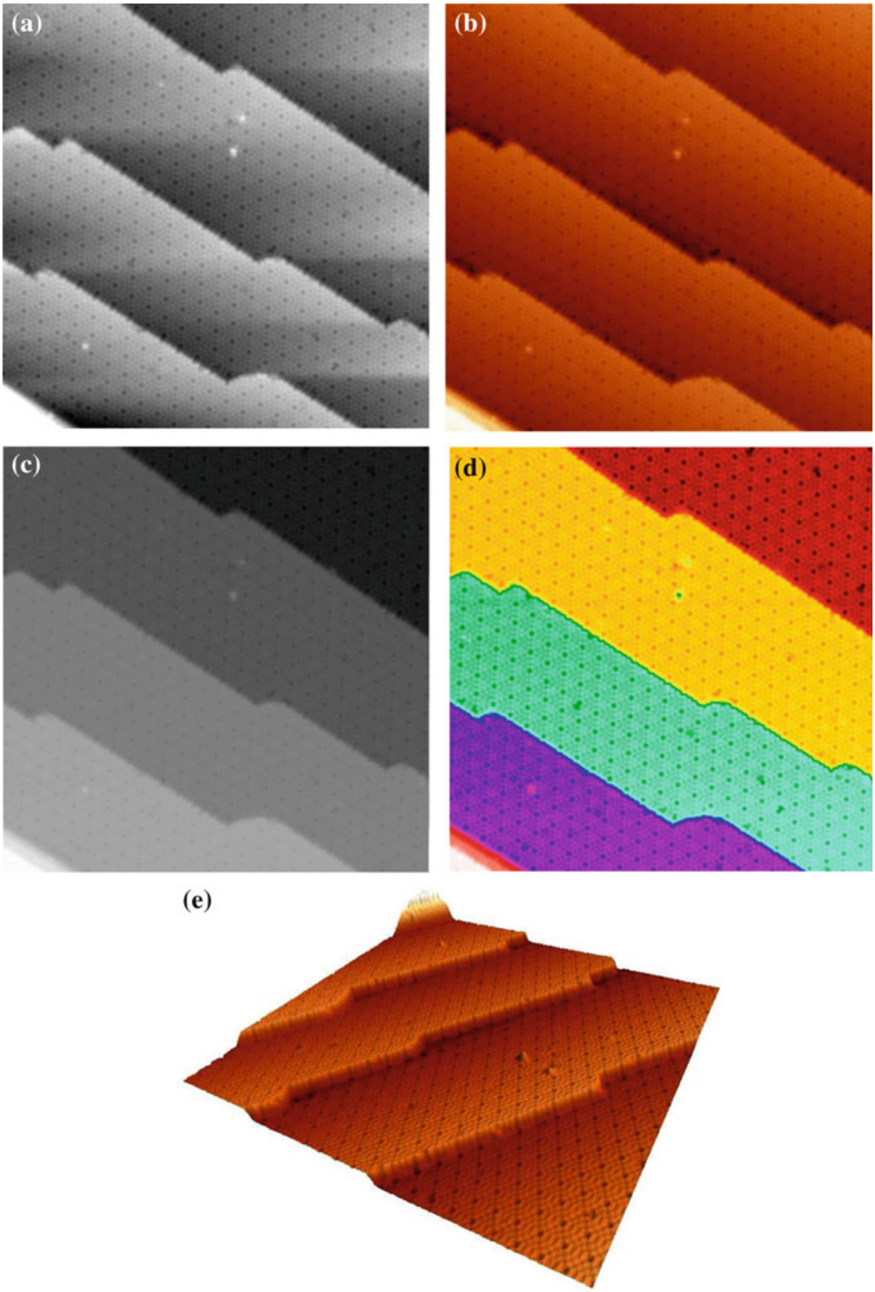

Fig. 7.3 STM image of a Si(111)-7 × 7 surface shown in different representations. Line-by-line background subtraction using **a** a *gray scale* palette and **b** a color palette. Plane background subtraction on one terrace **c** with *gray scale* palette and **d** a color palette with different colors for each atomic terrace. **e** Three dimensional representation of the same image

7.2 Image Processing

The application of image processing filters has two purposes. First, to enhance the image representation contrast above that possible with simple background subtraction and, second, to remove image artifacts such as high-frequency noise, noise pixels or noise lines. These are often eliminated by simple matrix filters. These filters consist of a sum of products of nearby pixel values with elements of a weighting matrix.

Matrix or convolution filters are used (a) to remove noise from the images, (b) to sharpen (high-pass), or (c) to smoothen (low pass) the images. The following algorithm describes the 3 × 3 convolution of image pixels. The measured value of an image pixel in the image matrix $z(x, y)$ is replaced by a modified value $z'(x, y)$

$$z'(x, y) = \frac{\sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} W_{(i-x+2, j-y+2)} z(i, j)}{\sum_{i=1}^{3} \sum_{j=1}^{3} |W(i, j)|}. \tag{7.1}$$

Depending on the properties of the matrix W high-pass, low-pass and other kinds of filters can be realized.

Another very simple and effective filter is the median filter. It removes speckle noise in the images, i.e. pixels which have, a very different gray value than the neighboring pixels. The advantage of this filter is that it does not lead to a pronounced blurring of sharp edges in the image, as other averaging filters do. For a median-filtered pixel consider the 8 pixels surrounding one pixel plus the center (original) pixel (9 pixels) and take as the new (gray) value for the center pixel the median of these nine pixels. The median is not the mean of the 9 pixels but the 5th highest value (i.e. the middle value, which is 68 in the example in Fig. 7.4a). The same procedure is applied to all pixels in the image. Median filtering is robust with respect to outlier pixels which would influence the mean considerably but not the median. In Fig. 7.4b, an image with white noise pixels is shown and Fig. 7.4c shows the image after median filtering.

Another frequently applied method for filtering SPM images is Fourier filtering. However, this kind of filtering is often not very useful for "improving" images. From the 2D Fourier transform of an image some parts considered to be noise are cut out and a reverse transformation is performed. With this procedure the image information in the respective frequency range is removed also. The emphasis in Fourier filtering is on enhancing the periodic part of the image, while in SPM often the defects and deviations from a periodic ideal lattice are interesting. Strong Fourier filtering can highlight the periodic part so strongly that atoms are "produced" by Fourier filtering and defect sites are "filled" by atoms.

One useful application of Fourier analysis for SPM images is the identification of a long-range periodic corrugation signal in the image which may be hidden by noise in the original image. Another application of a Fourier transform is to compare quantitatively two different periodicities which are present in one image, for instance

(a)

63	68	74
55	255	27
69	70	66

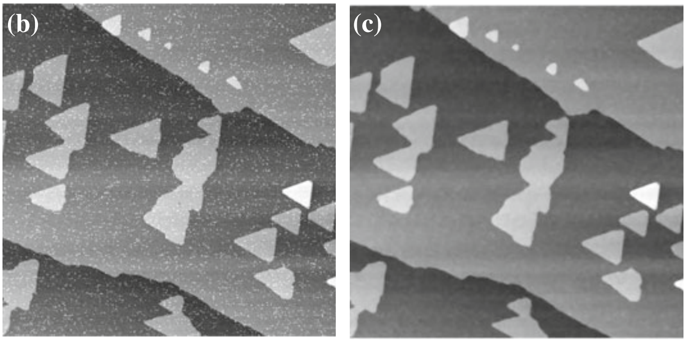

Fig. 7.4 **a** Example of the median filter showing gray values in a matrix of 8 pixels around a center pixel. When applying the median filter, the value of the center pixel is replaced by the fifth highest value (68 in the example). Thus the outlier value of 255 is replaced by the more reasonable value of 68. **b** STM image of triangular Si islands on Si(111) with speckle noise. **c** After median filtering this noise is removed

the atomic lattice and an additional periodic long-range modulation, as for instance a Moiré pattern.

It is important to mention in detail in presentations and publications which kind of image processing algorithms have been applied to the original data.

7.3 Data Analysis

There are a whole range of image analysis procedures which are often very specific to the problem under study. For instance, if in studies of epitaxial growth, island populations are analyzed, questions arise like: What is the island density per area? Also other questions about the distribution of the volume, the width, or the height of islands can be answered using SPM data. In principle, all questions related to the morphology of the surface can be answered, since the complete surface morphology is measured. Such analysis tasks can be performed more or less automatically. However, such data analysis procedures are very specific to the problem considered and we will not discuss them further here.

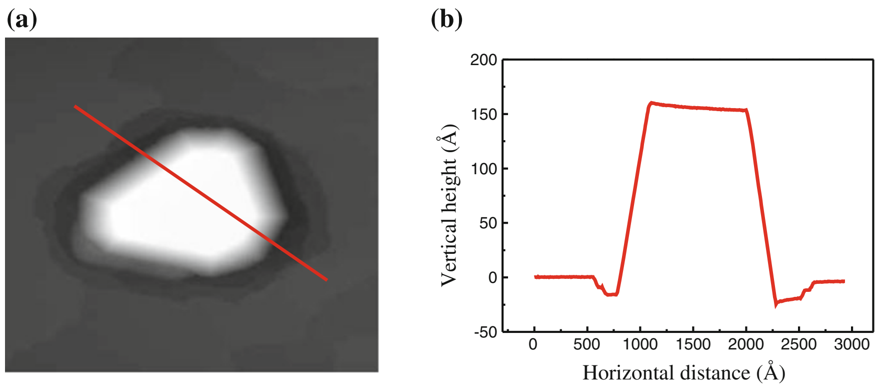

Fig. 7.5 **a** Gray scale STM image of a 3D Ge island. **b** Line scan across this island

A simple and general procedure for data analysis is the line scan. By interactive mouse clicking, a line is defined in an image on the computer screen and the height levels along this line (sometimes averaged over a certain width perpendicular to this line) are displayed and can be used for high-accuracy measurements of topographic heights as shown in Fig. 7.5, or horizontal spacings of features (atoms). Also the slopes of facets of surface features such as islands can be determined.

A second example of data analysis is the measurement of the roughness of a surface. The usual quantity characterizing the roughness of a surface is the RMS roughness defined as the standard deviation of the heights $h(x, y)$

$$\sigma = \sqrt{\langle (h(x,y) - \bar{h})^2 \rangle} = \sqrt{\frac{\sum_{x=1}^{L} \sum_{y=1}^{W} (h(x,y) - \bar{h})^2}{LW}}, \tag{7.2}$$

with L and W being the length and width of the image (number of pixels), and $\bar{h}$ the average height. A necessary requirement for a correct determination of the roughness is a good background subtraction of the scan slope.

7.4 Summary

- Data representation is the task to map the measured heights (DAC values) to gray levels in an image in an optimal way.
- Line-by-line background subtraction and plane background subtraction are commonly used.
- Matrix filters can be used to sharpen, or smooth the images, or to remove outlier pixels.
- In order to measure heights, width, or slopes of topographic features line scans can be used as data analysis tool.