# Quantifying and Processing HRTEM Images

# 30




D. B. Williams et al., *Transmission Electron Microscopy*

## CHAPTER PREVIEW

In this chapter we will equate our title with the use of the computer. We will simply use image processing to extract more information from a TEM image than we can obtain by eye. In the past the optical bench was also used for this purpose, but the number of optical benches is negligible compared to the number of computers now found in every TEM lab. Optical benches did allow us to form DPs which we could then modify to produce a processed image. This analog approach has now largely been replaced by its digital counterpart. The computer can be much cheaper than the optical bench and is far more flexible. The number of software packages which are designed for, or can easily be adapted to, TEM is also growing.

We can use image processing to produce a clearer view of the image, for example by subtracting unwanted background detail, correcting for noise or drift, or removing artifacts. The big warning, though, is that, when removing one artifact, you must be very careful not to introduce others.

Although it's nice to see information more clearly, the unique feature of the computer approach is that we can *quantify* the data in any image and then normalize these data. Now we can directly compare the quantified experimental image with computer-simulated images. Although throughout this chapter we will be concerned with HRTEM images, most of what we say can be transferred directly to the analysis of diffraction-contrast images.

The other general point is that the ideas we'll discuss are also applicable to images derived from different sources. Once the data are in the computer, i.e., in digital form, the source becomes unimportant as far as processing possibilities are concerned. Examples of "images" which might be obtained from the TEM include X-ray or EELS maps, STEM images, TEM images, and CBED or BSE patterns.

Most of our discussion will concern the use of computers. All you need to know is how best to get the data into the computer, how to process it, what to do with the data, and how to display the result!

# Quantifying and Processing HRTEM Images

# 30

## 30.1. WHAT IS IMAGE PROCESSING?

Image processing is essentially manipulating images. The topic arises in many fields, so we need to understand the words/jargon; we'll discuss the language of image processing as it is applied to TEM.

> The basic idea of image processing is changing images into numbers and manipulating the numbers.

Image processing is not only becoming more common, but is also finding new applications in many fields. Faster, more powerful computers and increased memory storage are making tasks possible which could not previously have been considered. As a result of this increasing user base, there are now many software packages available which can be used in microscopy; we listed some back in Section 1.5. These range from programs used widely in desktop publishing to those which have been custom designed for EM. The goals of image processing include that of quantitative microscopy. You must choose between the different packages, commercial and freeware, and match them to the computer available in your lab. One point to remember is that some very simple optical methods which don't rely on a computer can often be very helpful. The other point is that the eye is hard to beat.

There are many specialized books on this topic for the beginner or the expert; a selection is given in the references. The purpose of this chapter is to give a generalized overview. One problem in discussing this topic is that it is a very rapidly changing field. We will try to avoid specifics concerning particular programs but will mention these programs at the end of this chapter.

## 30.2. PROCESSING AND QUANTIFYING IMAGES

We process images primarily for two reasons:

- We may want to improve the appearance of an image, make it look sharper, more even in contrast, higher contrast, etc.
- We may want to quantify the information contained in the image.

Processing for improving the appearance of TEM images has been practiced for many years using such photographic techniques as "dodging," using "filters," selecting different emulsions, or varying the developer, etc. It is only recently that relatively powerful personal computers have become widely available, but the term "image processing" almost automatically implies the use of computers. Computer image processing will be the emphasis of our discussion. We have three requirements:

- We must be able to create a digital form of the image in the computer.
- We need appropriate software for processing the image.
- We need a computer which can perform the processing in an acceptable period of time with the required resolution.

Many comments here are similar to those we made in discussing the microscope itself. For example, you may have to work with the built-in system or the system that's already available in the lab. The difference is that some of the freeware programs are extremely powerful, so that all you need is the desktop computer. Many programs designed for desktop publishing are relatively inexpensive.

Thus, you can almost always find a way to extend your processing capabilities.

The motivation is that we need to obtain more information from an image than we can get by just looking at it. This principle applies to more than HRTEM; we are discussing it here because HRTEM is where at present it is most needed/used in TEM. However, any TEM, X-ray map, or energy-filtered image or DP may benefit from processing and quantification. We need to quantify the TEM parameters, in particular $C_s$. One unique aspect of image processing in the TEM is that we have a choice between on-line and off-line processing. In fact, we often use on-line processing (frame averaging and background subtraction on the video image) to see the image even though the image we record may be unprocessed.

## 30.3. A CAUTIONARY NOTE

For most of our discussion we will consider only processing techniques using computers. To a large extent we can simulate a TEM using the computer. As we saw in Chapter 29, we can model a crystal, insert apertures, define the electron beam, including its broadening in the specimen, and then calculate the image. What we do in image processing is start with the image, add apertures and special filters, and then create a new image, the processed image. This image is a *real* image. What we must be careful about is explaining just what processing procedures we have used, since these may affect the interpretation of the data. This reporting is particularly critical when the raw data (the "original" image) are not being reported at the same time.

> Always report how you have treated your image, so that the reader can compare your data with related data that may have been processed differently or not at all.

## 30.4. IMAGE INPUT

There are several methods you can use to put the TEM image on the computer. The choice depends in part on how much detail you want in your digitized image, but also depends on how much work you're prepared to do. In this discussion, we'll only consider images which you have looked at on a video monitor, a CRT, or the fluorescent screen. Your basic choices are:

- Transfer the image directly from the TEM to the computer.
- Record the image on film, then digitize it using a microdensitometer.
- Record the image on video tape.
- Record the image on film, then print it and use a flat-bed scanner.

There are many methods for creating a digital form of an image in your computer. The simplest is to use a slow-scan CCD camera, which we discussed in Chapter 7. The drawback to CCD cameras is that high-quality CCD chips are very expensive for 1k × 1k arrays and astronomically expensive for 2k × 2k arrays. Although such cameras may become routine add-ons for all TEMs in the near future, even when you do have such a camera you will probably also want to use film or video. With film, you can record a larger area than you can using a CCD; you should use a video-recorder for *in situ* studies when using a heating or straining holder.

We can transfer the image from a video tape or a video camera to the computer using a frame grabber. Frame-grabber boards are readily available for most computers. You can use a high-resolution scanner for photographs or negatives. At this time, scanners with a resolution of 1k × 1k cost about the same as a video camera with comparable resolution. The purist's approach is to use a microdensitometer to measure the intensity of the film point by point and read this directly into the computer. The advantage of the microdensitometer is that it is very precise and can achieve the highest resolution for a very large area. The main problem is that it is slow, being a serial-collection technique. If you use it to its best advantage, your image will require a large amount of computer memory, which in itself is not a problem, but manipulating such images will still be slow.

## 30.5. PROCESSING TECHNIQUES

### 30.5.A. Fourier Filtering and Reconstruction

The principle involved in filtering is that a mask is used to remove some information from an image in order to enhance or emphasize other information. As an extra complication we can process the image, e.g., Fourier transform an HRTEM image, then apply a mask and then reverse the processing.

We can vary the size of the apertures and the sharpness of their edges. This is not possible in a modern TEM with normal fixed-diameter objective apertures. A single

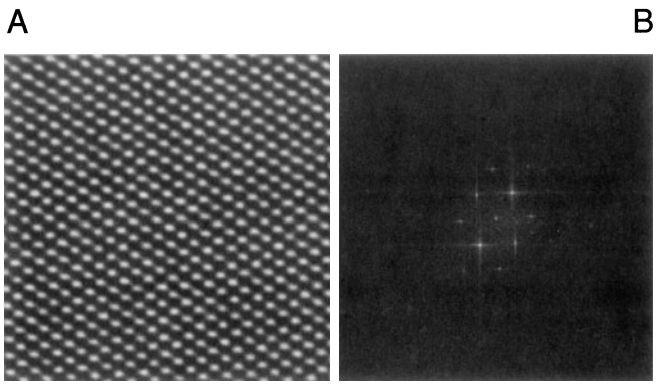


**Figure 30.1.** A square mask has been used to select the area shown in (A) from a much larger print of the image. The Fourier transform of this region is shown in (B) where you can see not only the spots in the 110 DP but also long streaks which run normal to the edges of the mask.

variable SAD aperture was used on some early TEMs, but it was triangular in shape and used three movable blades, so it was not much use as an objective aperture. You can best understand the procedure by an example. A square mask was used to select the region in Figure 30.1A from a much larger region of the HRTEM image, and its Fourier

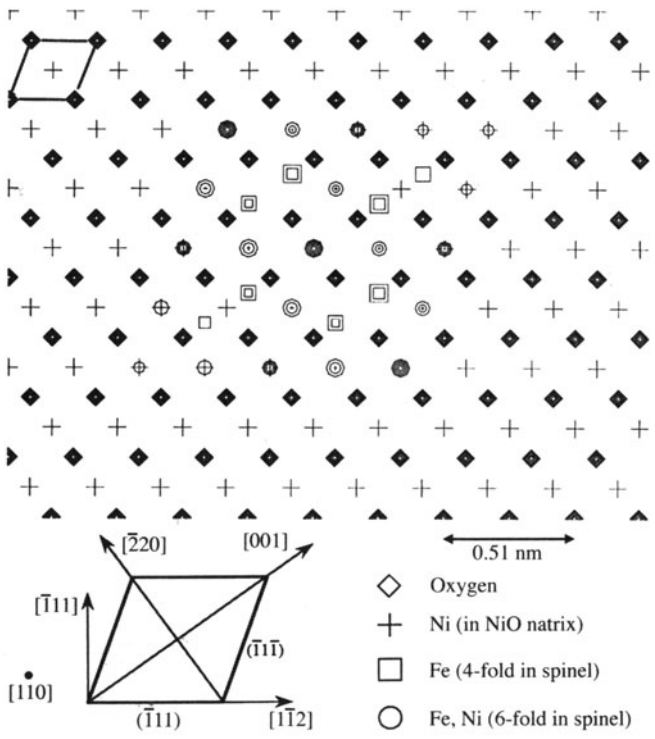


**Figure 30.2.** A model of an octahedron of spinel which is fully enclosed within a matrix of NiO. The rest of the specimen could then be modeled by adding extra layers of NiO above or below the defect layer.

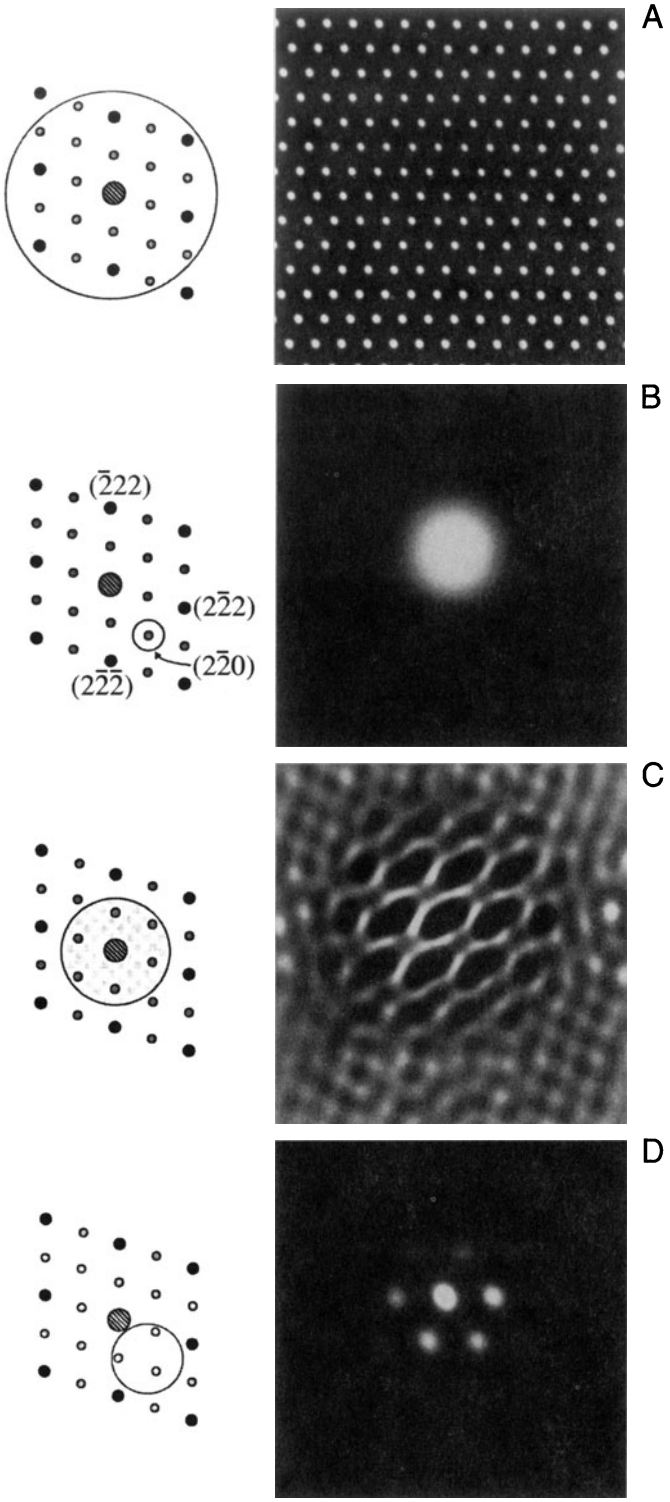


**Figure 30.3.** The DP from a specimen containing a particle such as that shown in Figure 30.2 is shown schematically in (A) together with the resulting lattice image. The other three pairs of diagrams illustrate how we can use the computer to produce different masks and thus generate different images, such as the DF image in (B). The image in (C) corresponds to the image shown in Figure 28.16, while the DF lattice image shown in (D) is analogous to that discussed in Figures 28.14 and 28.15.

transform (i.e., DP effectively from a few nanometers) is shown in Figure 30.1B.

What this technique does is to allow you to do microscopy in the computer. Your image becomes the specimen. You form the DP, then you can use apertures to select one or more beams to form the image; these apertures are the computer version of the objective aperture. Small apertures limit resolution as in the "real" TEM because the information about anything other than the perfect lattice is carried between the reciprocal lattice points. Figure 30.2 gives an illustration of how a model can be constructed of a particle in a matrix, which can be useful when simulating HRTEM and conventional BF/DF images in the computer, as shown in Figure 30.3. This was done using the Digital Micrograph package (Section 1.5).

## 30.5.B. Analyzing Diffractograms

In Chapter 28 we showed that the transfer function could be plotted out as shown schematically in, e.g., Figure 28.4. Another way of thinking about this plot is to imagine what would happen if you had a specimen which generated equally every possible value of **u**, i.e., every possible spatial frequency.

An amorphous film of Ge can provide just such a plot, but it is difficult to record the result because the scattered intensity is low.

> You will often see similar diffractograms obtained using a film of amorphous carbon. While such films are easier to make, they give little diffracted intensity for the range of **u** values between 6 and 8.5 $nm^{-1}$, which is important in HRTEM.

We thus record the image at high resolution, preferably directly using a slow-scan CCD camera, although digitizing the negative is fine. Then, by comparing your experimental plot of $I$ versus **u** with those calculated for different values of $\Delta f$ and $C_s$ you can determine the astigmatism, the defocus, $\Delta f$, and the value of $C_s$ (as we'll see below). It helps if you have a few particles of Au on the Ge film since the Au spots then give an internal calibration. Such a set of images and their corresponding diffractograms is shown in Figure 30.4. Notice that as the defocus of the objective lens increases, the number of rings increases but the rings become narrower. The contrast transfer gradually extends to larger values of **u**.

*Determining astigmatism.* You can use such diffractograms to correct the astigmatism, since a perfectly stigmated image will give a DP with circular symmetry. As you can see in Figure 30.5, even a small amount of astigmatism can be detected by eye. The computer can readily measure and correct such a pattern, as we will see shortly. This set of diffractograms also shows that the computer can distinguish astigmatism and drift in the image while your eye can easily mistake one for the other. Drift produces a circular pattern but the higher spatial frequencies are lost in the direction of drift.

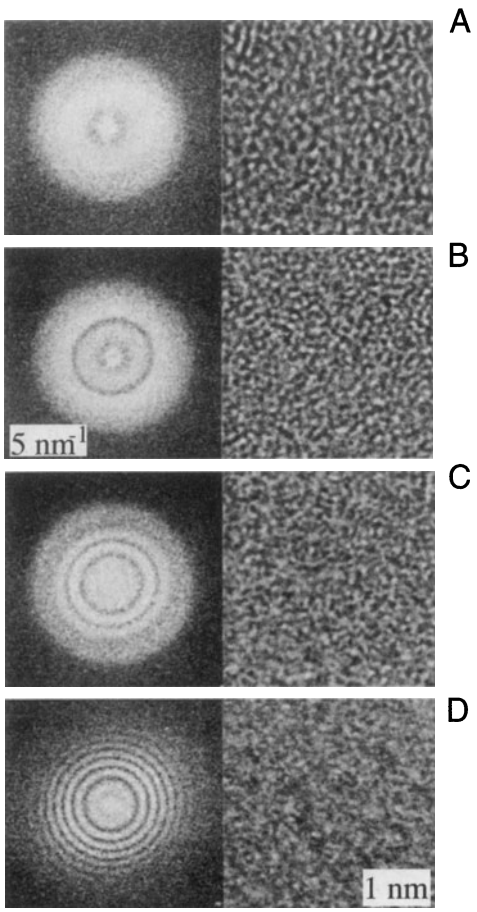


**Figure 30.4.** Four images of an amorphous Ge film and their corresponding diffractograms. $\Delta f$ has the following values: (A) 1 sch; (B) 1.87 sch; (C) 2.35 sch; (D) 3.87 sch. 1 sch = $-(C_s\lambda)^{1/2}$.

*Determining $\Delta f$ and $C_s$.* You can determine $\Delta f$ for any image by measuring the radii of the bright and dark rings in the diffractogram, since bright rings correspond to $\sin \chi(\mathbf{u}) = 1$ and dark rings correspond to $\sin \chi(\mathbf{u}) = 0$.

$$\sin \chi(\mathbf{u}) = 1 \quad \text{when} \chi(\mathbf{u}) = \frac{n\pi}{2} \text{ and } n \text{ is odd} \quad [30.1]$$

$$\sin \chi(\mathbf{u}) = 0 \quad \text{when} \chi(\mathbf{u}) = \frac{n\pi}{2} \text{ and } n \text{ is even} \quad [30.2]$$

Since $C_s$ will also influence the location of the rings, you need at least two rings. Krivanek (1976) has given a simple

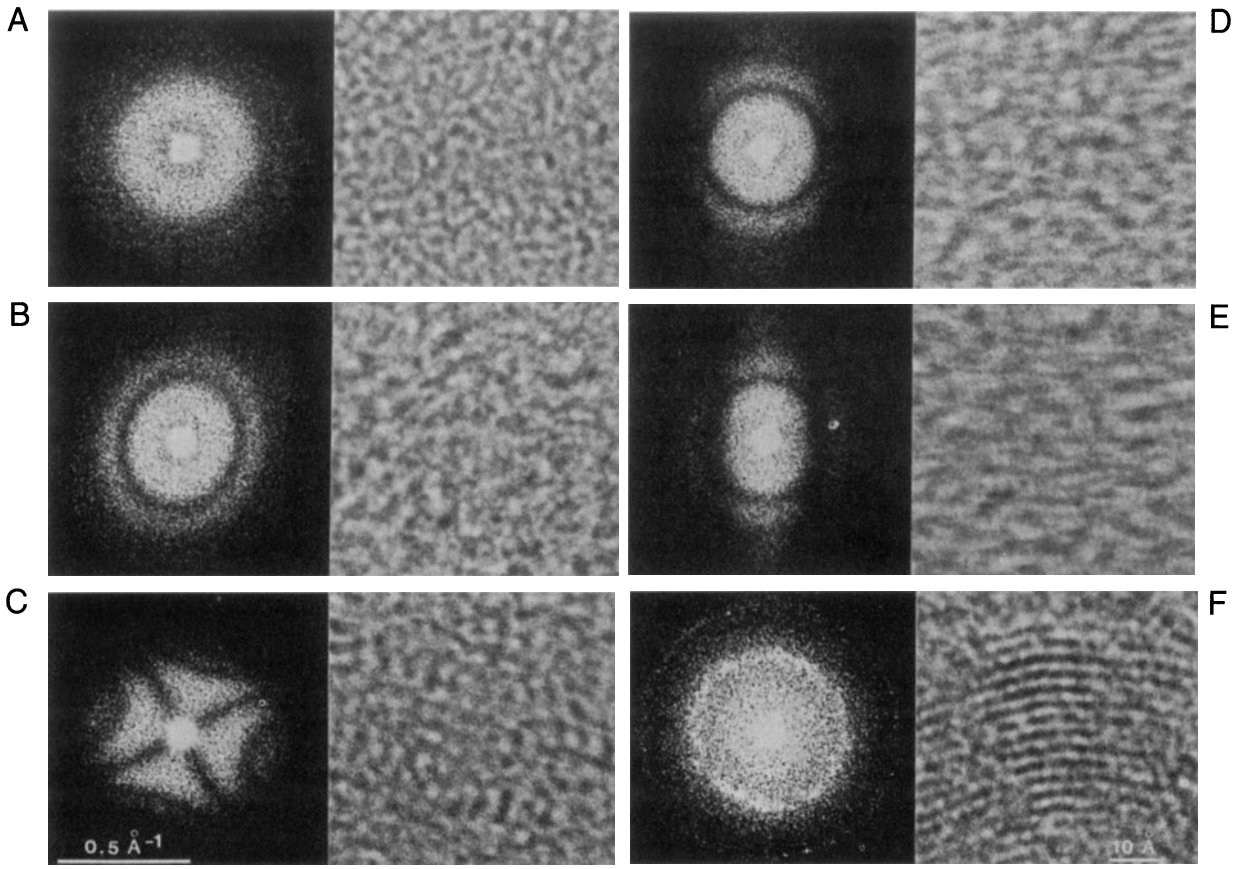


**Figure 30.5.** Six images of an amorphous carbon film and their corresponding diffractograms illustrating different misalignments of a 300-kV HRTEM: (A) well aligned and no drift; (B) some astigmatism ($C_a$ = 14 nm); (C) more astigmatism ($C_a$ = 80 nm); (D) no astigmatism but drifted 0.3 nm; (E) no astigmatism but drifted 0.5 nm; (F) well aligned and no drift showing graphite calibration fringes of 0.344-nm spacing. (B,D,E) $\Delta f = -2.24$ sch; (C) $\Delta f = 0$.

procedure for finding both $C_s$ and $\Delta f$. If we start with our definition of $\chi$

$$\chi(\mathbf{u}) = \pi \Delta f \lambda u^2 + \frac{1}{2}\pi C_s \lambda^3 u^4 \qquad [30.3]$$

then, inserting the values given in equations 30.1 and 30.2 leads to

$$\frac{n}{u^2} = C_s \lambda^3 u^2 + 2\Delta f \lambda \qquad [30.4]$$

All we now have to do is plot $nu^{-2}$ versus $u^2$ to obtain a straight line with slope $C_s\lambda^3$, and with an intercept on the $nu^{-2}$ axis of $2\Delta f\lambda$. Assign $n$ = 1 to the intensity maximum of the central bright ring, $n$ = 2 to the first dark ring, etc. The analysis can be trickier if you have used an underfocus condition or if you are very close to Scherzer defocus, but you will know when you have not found a straight line. You should find that your value of $C_s$ will be close to that given by the manufacturer! A rather neat result is that, if you plot $nu^{-2}$ versus $u^2$ for different diffractograms (i.e., different values of $\Delta f$), then the points corresponding to each particular value of $n$ will lie on a hyperbola, as shown in Figure 30.6A. You can use these hyperbolas to determine $C_s$ for any microscope and $\Delta f$ for any diffractogram (Krivanek 1976).

*Diffractograms and beam tilt.* Beam tilt is very difficult to correct by eye; even worse, it can cause the diffractogram to look astigmatic so you correct the astigmatism instead. In the image, as we saw earlier, beam tilt can improve the appearance but confuse the interpretation! The set of diffractograms shown in Figure 30.6B shows you how to overcome the problem. You have to compare diffractograms taken at different beam tilts to determine the zero-tilt condition. A pair of diffractograms taken at ± θ° tilt will only look the same (though rotated) if the beam had zero tilt at θ = 0°. In the example shown, the diffractograms above and below the horizontal line are similar, so $\theta_y$ was very close to zero for the central condition. However, the pairs of diffractograms on opposite sides of the vertical axis differ slightly, so the alignment of $\theta_x$ was not perfect.

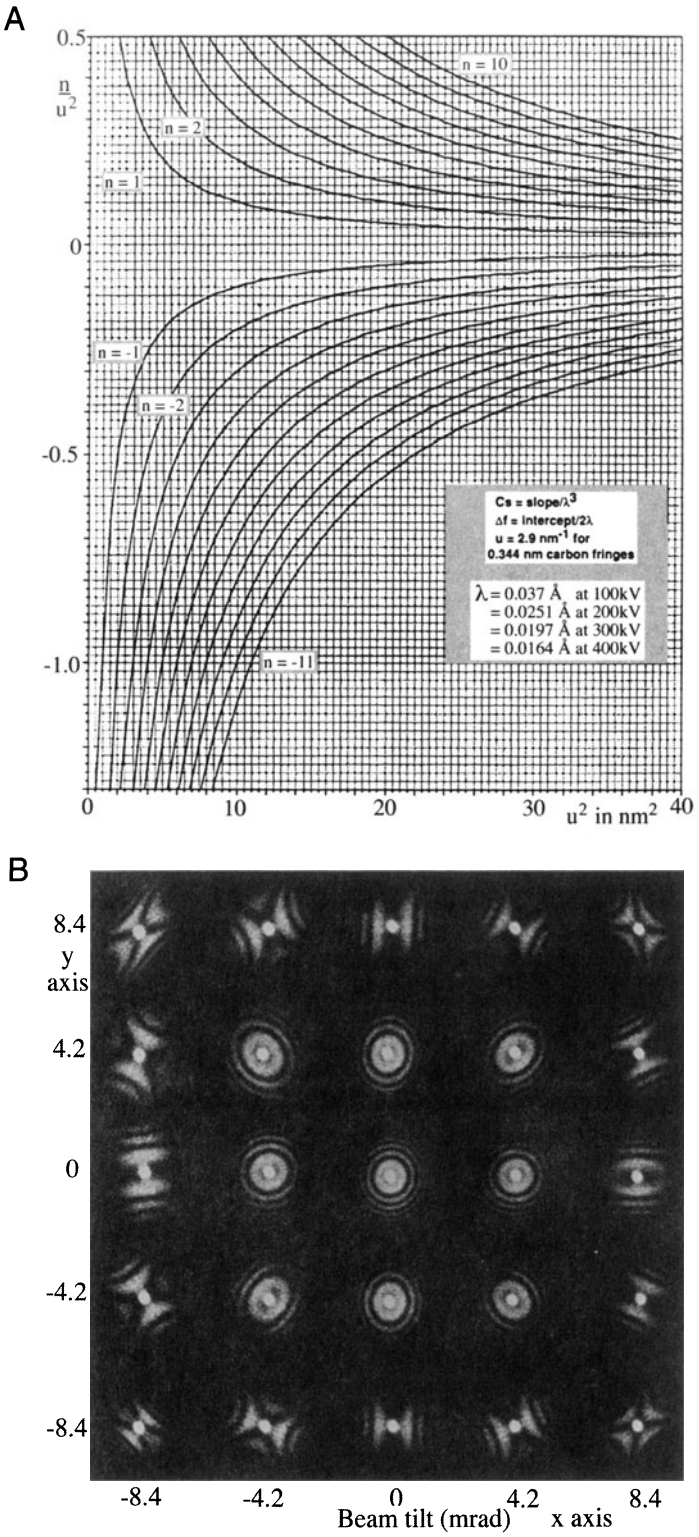


**Figure 30.6.** (A) Plot of $nu^{-2}$ versus $u^2$. The rings in any diffractogram correspond to a series of $n$ values which allow you to draw straight lines on this figure and thus determine the slope and the ordinate intercept, giving $C_s$ and $\Delta f$ respectively. (B) Set of diffractograms showing the effect of incident beam tilt.

## 30.5.C. Averaging Images and Other Techniques

If you have recorded a series of images using a video camera, for example, you can average them over several frames as your eye does automatically. The result of such a process is illustrated in Figure 30.7. Different methods can be used to average the images. The easiest approach appears to be as good as any and simply involves taking the unweighted average of your best images, i.e., in the video example, just average over a series of frames. If you know that the object you're studying has a certain symmetry, you can use that information to improve the image further. The article by Trus *et al.* (1992) will give you a start on this process. If you want to remove the blur due to motion of the image, then you will really need to delve much more into this subject.

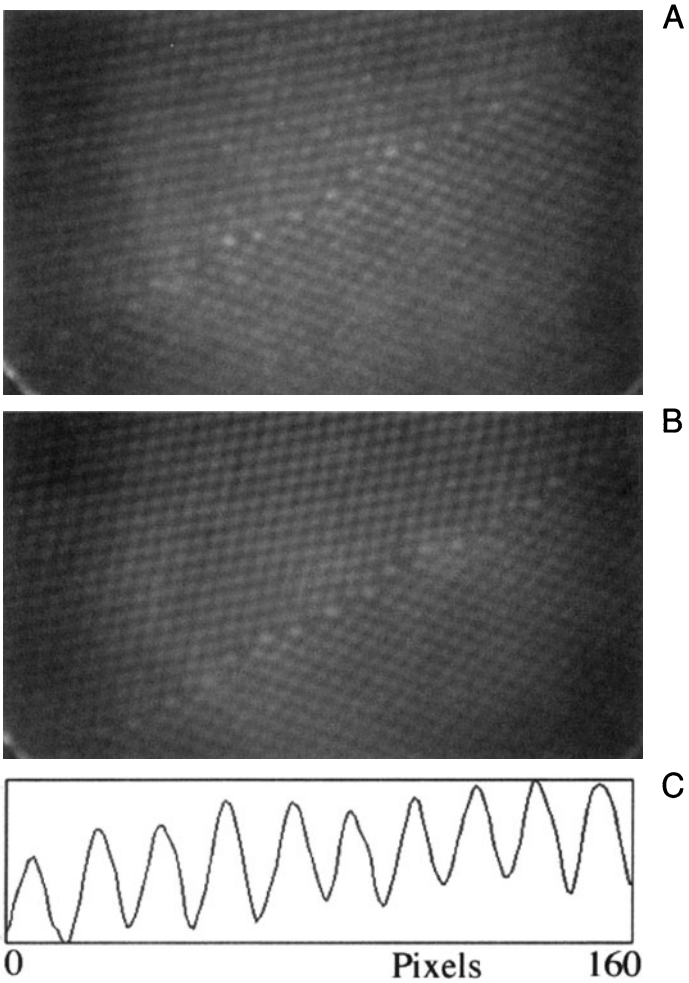


**Figure 30.7.** An example of the benefit of frame averaging to improve information from a video recording: (A) one frame, (B) 16 frames, (C) intensity profile along a (111) plane in (B).

If you use a TV-rate video, you'll almost certainly use background subtraction as a matter of routine. For example, you can record an image of the honeycomb pattern of the YAG detector, store it, and then automatically subtract it from all subsequent images in real time.

You may find it useful to add artificial color (pseudo-color) to your TEM images. Although it is often assumed that this is just done to make the images even more appealing to the nonmicroscopist (or nonscientist), there is actually a valid reason for the practice. Our eyes are much more sensitive to small variations in color than they are to small variations in gray level. You might therefore find color useful if you have a wide range of gray levels and want the viewer to be able to "see" some subtle variations. Similarly, you can use color to emphasize a particular gray level in an image. However, you have to be very careful in your choice of look-up table (LUT), the table which relates each gray level to a particular color. To get a feel for the dangers, play with Photoshop™ and your favorite TEM image. In the TEM, all of our apertures have relatively sharp edges. In the computer, you have the possibility of using multiple apertures, apertures with different shapes, and apertures with diffuse edges. Apertures with diffuse edges will help eliminate the streaking which will otherwise be present. You can also use the computer to do "unsharp masking," which is not the same as simply using a diffuse mask. The technique comes from the photographic process whereby we first print and image out of focus onto film, thus making a complementary image, except where there is fine detail present in the original image; in digital processing this is called Laplacian filtering. Many more examples are given in Russ's two books (Russ 1990, 1995).

## 30.5.D. Kernels

A kernel is simply an array of numbers which we can use to perform operations on a digital image. If we have the $3 \times 3$ kernel, $K$ (we can have $5 \times 5$, $7 \times 7$, etc. but the computation time becomes too long, especially for real-time situations)

$$K = \begin{matrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{matrix} \qquad K_o = \begin{matrix} A & B & C \\ D & E & F \\ G & H & I \end{matrix}$$

we can apply it to every $3 \times 3$ group of pixels in our image, eg., $K_o$, and put the result in a new digital image. If we call our new $3 \times 3$ image $K_i$, then

$$K_i = \begin{matrix} A' & B' & C' \\ D' & E' & F' \\ G' & H' & I' \end{matrix}$$

The new image will have, for example, E′ = 8E - A - D - G - B - H - C - F - I. This kernel then gives us a digital Laplacian (an approximation to the second linear derivative, $\nabla^2$). What this kernel is doing is subtracting the brightness value of each neighboring pixel from the center pixel. If the area is a uniform gray, it will become white, so changes in contrast will be exaggerated. We can design a wide range of kernel operators. For example, the edge enhancer kernel has the effect of digitally differentiating the image. (We'll see a related digital-processing procedure applied to spectra in Chapters 35 and 39.) The Sobel and Kirsch operators are examples of such edge detectors; each can be thought of as the sum of several kernel operators. We can also use binary morphological operators which make binary features become larger or smaller. All of these operations can be carried out in any standard image processing package. In general, you should be very careful when using such techniques in TEM; their value is in displaying data which might otherwise be missed, rather than helping you quantify an image.

# 30.6. APPLICATIONS

This section will give you a taste of how image processing is being used now. It is just part of a rapidly growing list, so we are not going to be detailed or inclusive. We can separate the applications into two groups:

- Noise reduction or improving the signal/noise ratio.
- Quantifying images.

Of course, the first topic is included in the second.

## 30.6.A. Beam-Sensitive Materials

Low-dose microscopy necessarily implies that the signal-to-noise ratio will not be large; if it is large, the dose could have been smaller. This problem has been extensively addressed in biological EM and led to Klug's Nobel Prize for "Development of crystallographic electron microscopy and the structural elucidation of biologically important nucleic acid–protein complexes" in 1982 (see, e.g., Erickson and Klug 1971). In materials science we have tended to accept "beam damage" as a fact of life, but this attitude will not be acceptable for future quantitative HRTEM. Most modern microscopes will allow you to perform all your alignments on one area and then translate the beam a pre-

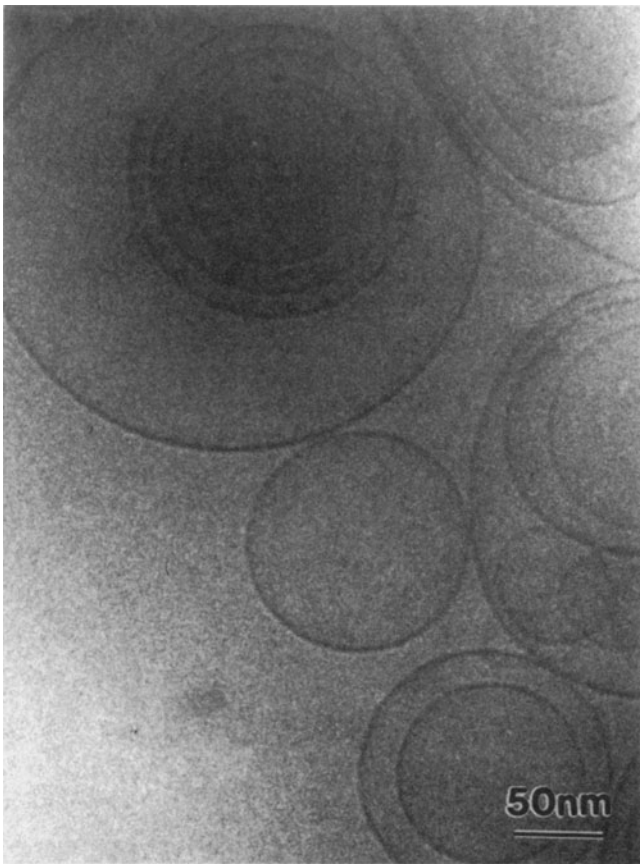


**Figure 30.8.** The image is from a highly beam-sensitive solution of surfactants in water. The solution has been frozen by plunging a film into liquid ethane and then transferring it to the TEM. The large circles are the surfactants that have aggregated to form vesicles; the concentration of the surfactants in the solution is just right for them to form lamellar structures. Texture starts to appear in the image as soon as the beam interacts with the specimen.

determined distance in a predetermined direction before recording the image of a pristine area. Clearly, the CCD camera will not only let you see your image without waiting to develop the plates, but you can take a series of images for noise reduction purposes and/or assess whether the imaging conditions were what you had intended. The image shown in Figure 30.8 illustrates the possibilities. If you read the review by van Heel *et al.* (1992), you will get some idea of how far you can already go in this field.

## 30.6.B. Periodic Images

In discussing quantitative analysis, we have already noted how we can use the computer to identify similar features and combine them in order to reduce the noise. This technique has many possible variations. Again, biological applications are leading the way with 3D crystallographic reconstruction (Downing 1992, Dorset 1995) and even correcting for distortions in the specimen (Saxton 1992).

## 30.6.C. Correcting Drift

Although drift is not as limiting on new machines, many older TEMs are still in use. Drift can be corrected now if the rate and direction of movement are constant. The computer can calculate the relative translation of two images and change the current in the image translation coils appropriately (which avoids moving the specimen). The difficulty is that the drift may not be linear. When implemented, such routines will be particularly valuable for frame averaging using a video camera. Then there will also be many applications for diffraction-contrast imaging, as well as for microanalysis.

## 30.6.D. Reconstructing the Phase

Although we are studying phase contrast, the image intensity doesn't directly give us phase information. Kirkland *et al.* (1982) have shown that the phase can be reconstructed by processing a defocus series. In their approach they use an iterative nonlinear image-processing technique to reconstruct the complex electron-transmission function. The technique was demonstrated using images of $CuCl_{16}PC$ (hexadecachlorophthalocyanine copper).

Five images from the experimental defocus series are shown in Figures 30.9a–e, together with the reconstructed transmission function plotted both as a real and imaginary part and then as an amplitude and phase part Figures 30.9f–i. The projected structure of the known unit cell is also shown Figure 30.9j. The phase image contains most of the structural information: it corresponds to the projected potential while the amplitude image contains features due to inelastic scattering. Notice in particular that we can now identify the benzene ring. This is one of the few examples of full phase reconstructions published. You must record such a series of defocus images if you want to do quantitative HRTEM.

## 30.6.E. Diffraction Patterns

We have generally ignored the intensities in DPs because they are so strongly influenced by dynamical scattering. However, if the specimen is very thin, we can use the intensity in the SAD pattern to carry out electron crystallography in the same way as in classical X-ray crystallography. As you can appreciate from Figure 30.10, particularly if the unit cell is large and the specimen examined is thin, there is a great deal of information in the SAD pattern but

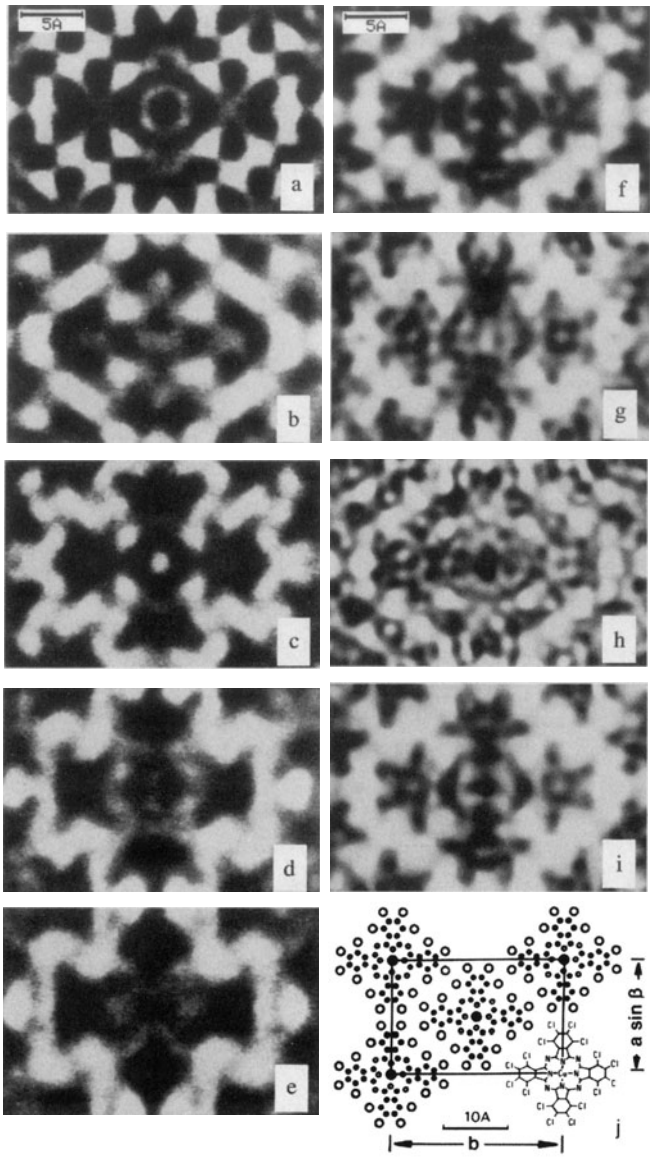


**Figure 30.9.** (a–e) Images from an experimental defocus series of $CuCl_{16}PC$; the reconstructed transmission function plotted as both a real and imaginary part (f,g) and then as an amplitude and phase part (h,i); (j) the projected structure of the crystal.

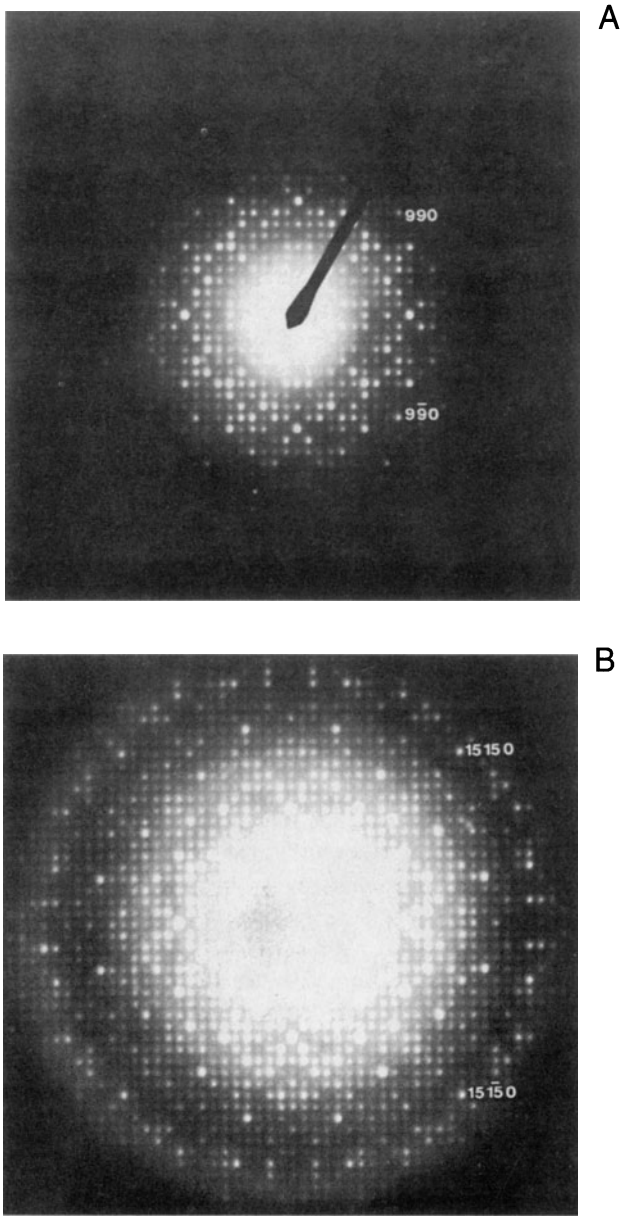


**Figure 30.10.** SAD patterns from $K_2O.7Nb_2O_5$ recorded using two different exposures. More than one exposure is needed to get all the information in the pattern. The space group in P4bm with $a = b = 2.75$ nm. The (15, 15, 0) reflections correspond to spacing of 0.13 nm.

you can't get it all in one exposure. Hovmöller's group (see, e.g., Zou 1995 and the ELD program in Section 1.5) have provided a routine for analyzing such patterns and getting structure-factor information. For example:

- Several patterns are recorded using exposure times of 0.5 s to >15 s.
- The patterns are digitized directly from the negatives using a CCD camera and a light box for backlighting.
- The intensity of each film is calibrated using a calibration strip with 20 equal exposure steps.
- The intensity is measured for all the spots and the processing begins.

This digitization process is particularly demanding, because each reflection typically covers an area of < 0.5 mm diameter on the photographic film. You will need to be able

to index three strong, but clear, reflections. The computer can then perform a series of functions:

- Optimize the location of these points using a center-of-gravity approach, locate the origin, and index the rest of the pattern.
- Extract the intensities of each peak, taking care not to be misled by any shape effects of the specimen.
- Use reflections which are present on two successive negatives (since the intensities are now in digital form) to calibrate films recorded with different exposures and thus develop a very large dynamic range.

A cooled slow-scan CCD camera will give you a large dynamic range and better linearity than its room-temperature counterpart and should simplify this type of analysis. There are other complications in using electrons rather than X-rays for this kind of crystallography. While the Ewald sphere is still curved, as with X-rays, electrons can easily damage your specimen. However, the technique clearly has potential! Like all TEM techniques it can be applied to much smaller regions of the specimen than is possible for X-ray beams. We can also use the symmetry present in the SAD patterns. Because the specimen is very thin, this technique could be described as "kinematical" crystallography and complements the "dynamical" electron crystallography that we described for CBED patterns from thicker specimens in Chapter 21. The process of extracting intensities from DPs can now be carried out using the ELD software package (Section 1.5).

> The structure deduced from the HRTEM approach should generate the experimental SAD pattern, so it should be possible to use these diffraction data to further refine the structure.

If we can use the quantitative information available in DPs, we could combine this information with our experimental and simulated HRTEM images. The quantitative analysis of the DP is known as structure-factor-modulus restoration or reconstruction (Tang *et al.* 1995). One limitation of this approach is that the specimen be sufficiently thin that diffraction is kinematical. Of course, this requirement is necessarily similar to the HRTEM requirement of the WPOA.

## 30.6.F. Tilted-Beam Series

Having gone to great trouble to remove any beam tilt, we will now mention how beam tilt can be used to extend the resolution of your microscope! The basic idea goes back to the tilted-beam lattice-fringe imaging we discussed in Section 27.2. Now you use a computer to combine information in different tilted-beam images. The method proposed by Kirkland *et al.* (1995) assumes that you know when the beam tilt is zero. You tilt the beam through different angles in well-defined directions so that you transfer information in overlapping regions of reciprocal space, as shown in Figure 30.11A; you also need the on-axis image, as shown in the tableau in Figure 30.11B. Since it is important that

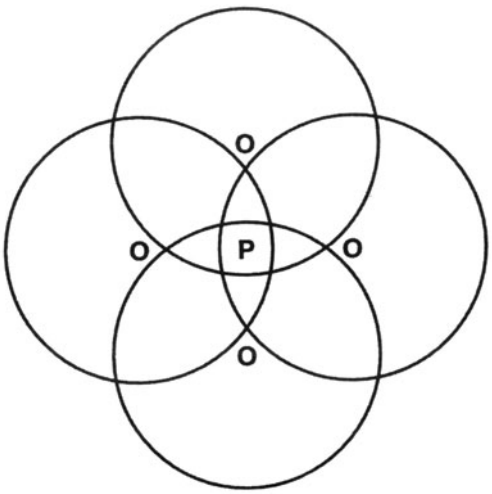


**Figure 30.11.** A method for extending the resolution of your TEM. Set the beam tilt to zero, then tilt the beam through different angles. (A) The four regions of Fourier space are shown by the four circles; O is each position of the tilted beam, P is the optic axis, and PO corresponds to the angle of tilt. (B) The five images used in the restoration arranged according to the beam tilt used in (A) with the on-axis image at the center.

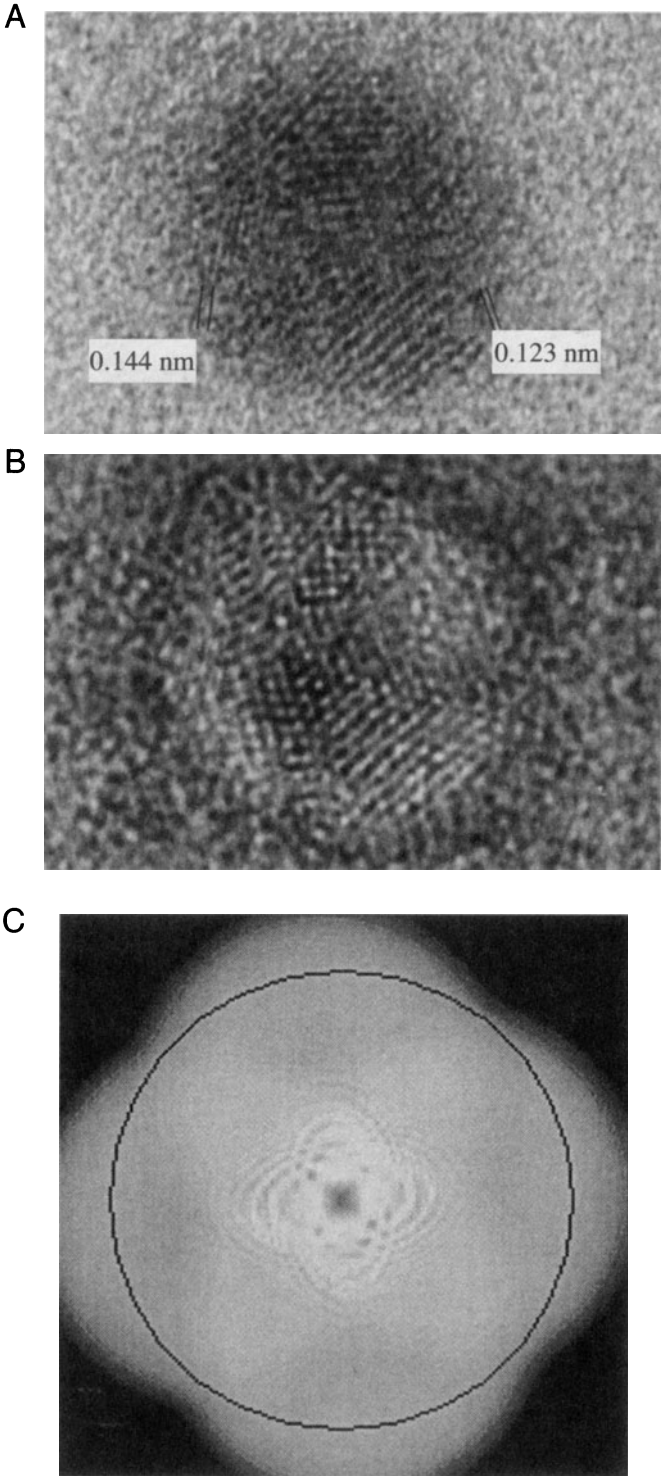


**Figure 30.12.** The restored image of a gold particle on amorphous Ge: (A) the amplitude (modulus) image showing 0.123-nm fringes, (B) the corrsponding phase image, (C) the transfer function after restoration, plotted in two dimensions. The circle corresponds to 0.125-nm detail successfully transferred to the image. For such thin specimens, atom-dense positions have a reduced modulus [black in (A)] and an increased phase [white in (B)].

the same area is imaged, a sixth (on-axis) image is recorded and correlated with the first on-axis image to check for drift and specimen degradation. You now need to restore the modulus and phase to create a higher-resolution image.

Kirkland's paper is a beautiful demonstration of the care needed in image processing. Even aligning the images is not trivial. However, the resulting restoration shown in Figure 30.12 demonstrates the potential of the technique: detail is present in the image at a resolution of 0.123 nm using a 400-kV microscope.

## 30.7. AUTOMATED ALIGNMENT

In the not-too-distant future, all TEMs will have automatic beam alignment, astigmatism correction, and readout for $\Delta f$. What makes this possible is the diffractogram analysis, a slow-scan CCD camera to digitize the image, and computer control of all the microscope functions. By microscope functions we mean all lens currents, deflector currents, specimen drive, and aperture drives. The slow-scan camera is needed because the computer needs to make measurements on more than one ring in the diffractogram. So you will not actually have to sit in front of the microscope once you have loaded the specimen.

> The big advantage in remote control will not be just that you can sit in Bethlehem and operate a microscope in California but that you can locate the microscope in its own controlled environment with no one opening the door to check if it is working (it was) or entering the room and thus changing the heat-load. Also, if your specimen is not ideal or the microscope breaks down, you won't have to go for a walk on the beach in California but can continue word processing in Pennsylvania.

The procedure has been extensively discussed and implemented by Saxton and Koch (1982), Krivanek and Mooney (1993), Koster *et al.* (1988), Koster and de Ruijter (1992), and others. Your role is to select a suitable region of the specimen close to the area of interest; the area of interest should ideally only be examined at low magnification. At present, you will make the initial alignment manually and then turn the process over to the computer. The computer will then adjust the astigmatism and correct the beam tilt independently and quickly.

Figure 30.13 shows how well and quickly this procedure can now be done. The different diffractograms in each tableau correspond to incremental changes in the

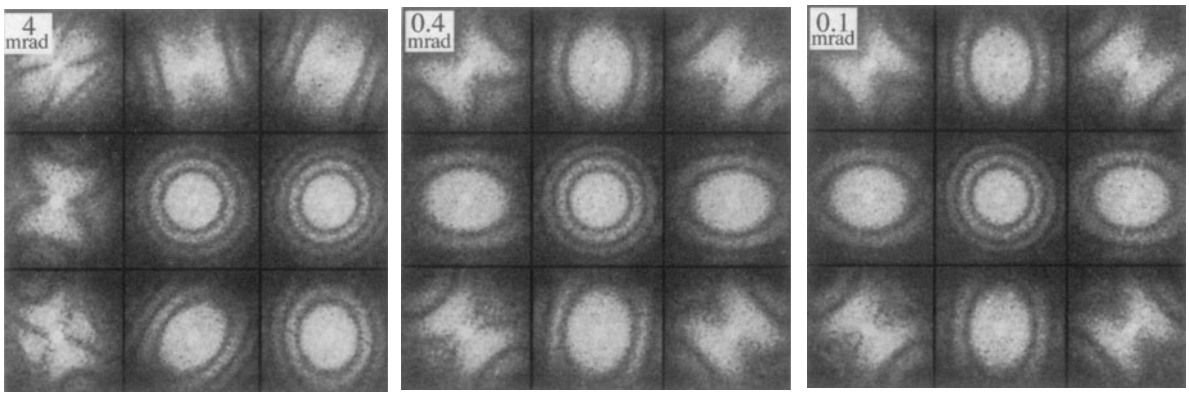


**Figure 30.13.** Using the computer to correct the beam tilt: The different diffractograms in each tableau correspond to incremental changes in the beam tilt of 6 mrad in the $x$ and $y$ directions away from the initial beam tilt in the central diffractogram. The computer initially determined a misalignment of 4 mrad, then corrected this to 0.4 mrad and finally to 0.1 mrad. Note the central diffractogram is almost unchanged, emphasizing the need for computer-controlled tilting to give correct alignment.

beam tilt of 6 mrad in the $x$ and $y$ directions. The computer showed that the initial tilt error was 4 mrad, which was reduced to 0.4 mrad after one pass and < 0.1 mrad after the second pass. Each pass only took 28 s! The astigmatism shown in Figure 30.14 was initially 53 nm. It was reduced to 3 nm after one pass and to < 1 nm after the second pass. For this correction each pass took only 8 s. Even an experienced operator can't match this speed or accuracy for either correction, and both corrections are now *quantitative*.

The defocus value is then found by calibrating the image with minimum contrast at $\Delta f_{MC}$. The value of $\Delta f_{sch}$ can then be found when the image contrast is a maximum. Although the method described here uses the diffractogram, a corresponding approach can be followed by analyzing variations in the contrast of the image. This technique has been described by Saxton *et al.* (1983) and uses a method of cross-correlating pairs of images recorded at each focus setting of the microscope. The reason for cross-correlating images is to remove the effects of electron shot noise; variations due to the photographic emulsion are avoided by using the slow-scan camera.

## 30.8. QUANTITATIVE METHODS OF IMAGE ANALYSIS

In the next six sections we will go through several particular illustrations of image processing in HRTEM. Our discussion will draw heavily on the work of a few pioneers in this field; we will also emphasize that, although this subject is still in its infancy, it is developing rapidly. The main cause for the delay in its application in materials science has been the lack of affordable fast computers and the feeling that everyone must write their own image processing program; the latter is not true and is certainly not recom-

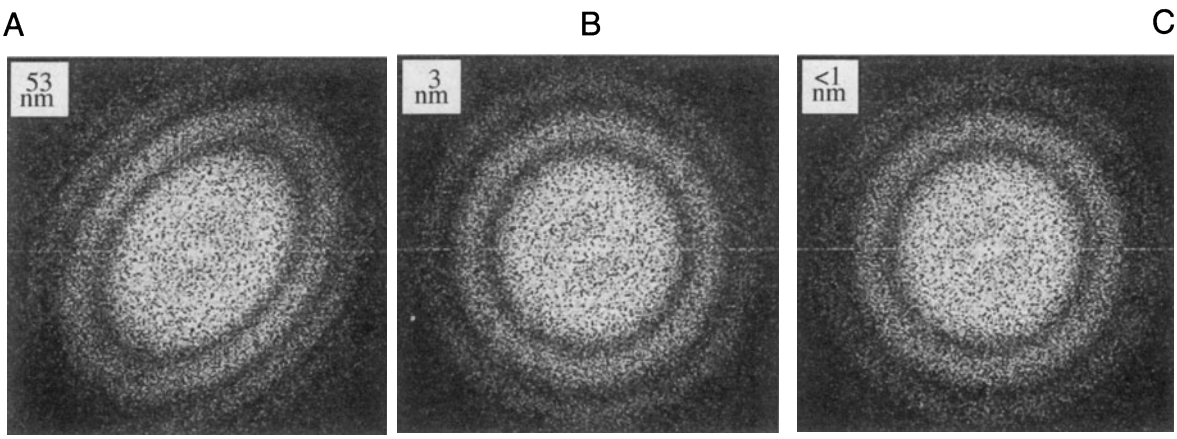


**Figure 30.14.** Diffractograms showing the astigmatism corrections made by the computer following a similar procedure to that shown in Figure 30.13. The final diffractogram shows that the HRTEM is now very well stigmated.

mended. At this time we can summarize the situation as follows:

- Quantitative analysis is difficult, often tedious, and invariably time-consuming.
- You have to understand the basic ideas of image theory.
- Your analysis is only as good as your image and your image is only as good as your specimen.

We gave the necessary information on how to obtain the software in Section 1.5.

## 30.9. PATTERN RECOGNITION IN HRTEM

The most obvious feature of the majority of HRTEM images is that we see patterns of white, gray, and black dots or other shapes. If the pattern is perfect everywhere, your specimen is probably a single crystal with no defects, no thickness variations, no variation in atomic composition, and no use. If it is not perfect, then we can use pattern recognition to quantify the variations.

> The principle of pattern recognition is to take or make a template, move it across your image, and measure how closely the image resembles your template.

Clearly you need a computer for this! Your template needs to match the magnification and rotation of the pattern you are examining. Then you need a method to say how close your match is, i.e., you need to know your "goodness of fit." We will go through some basics here, but strongly recommend that you consult the list of original papers given at the end of this chapter when you are ready to apply this technique.

We can illustrate the approach as shown in Figure 30.15 following Paciornik *et al.* (1996). The large rectangle represents your digitized image and could be 1k × 1k; remember, the numbers indicate pixels. The small rectangle represents your template. This template might be a small area of the pattern or a simulated image, in which case it might be a 128 × 128 pixel template. If the template is taken from your image, then you have already got the

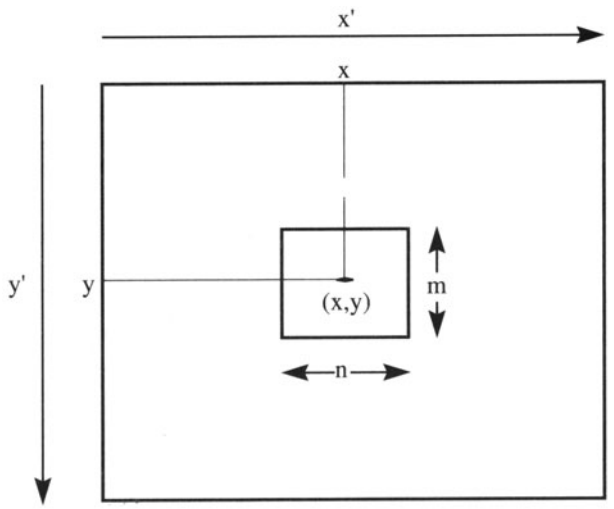


**Figure 30.15.** The large rectangle represents the digitized image, size ($x' \times y'$); the small rectangle, ($m \times n$) pixels, represents the template used in the cross-correlation calculation. The small rectangle is moved to different ($x$,$y$) positions during the process.

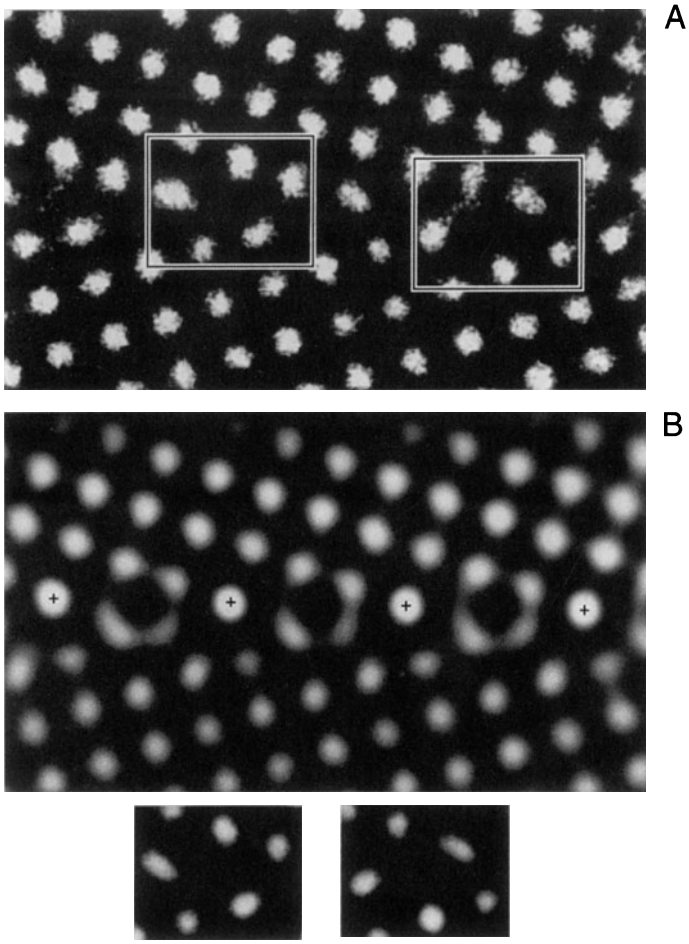


**Figure 30.16.** Analysis of a small region of a $\Sigma = 5$ tilt boundary in $TiO_2$. The two small boxed regions in (A) are only present at the boundary; these are used as the templates for the cross-correlation method. (B) The cross-correlated image. The small rectangles at the bottom of the figure are the low-noise averaged images of the GB templates.

right magnification and rotation. If not, you have to set these first and we will return to this problem shortly.

This is a real-space approach.

The process is best understood by an example. Figure 30.16A shows an HRTEM image of a small region of a $\Sigma = 5$ tilt boundary in $TiO_2$. The two small boxed regions appear only at the boundary and are selected as templates. The matching process has then been carried out and the new image is shown in Figure 30.16B. Having found all the regions which match the template, we could then take the average of these to produce low-noise images of the grain boundary templates. The final step is the comparison of these templates with models of the grain boundary structure. There are two important points to remember:

- When you average images, you implicitly assume that they are all the same.
- Don't forget our discussion in Chapter 27 of interface grooving and the problems associated with interfacial segregation.

## 30.10. PARAMETERIZING THE IMAGE USING QUANTITEM

In general, the thickness or chemistry will vary as you cross the specimen, i.e., the projected potential varies across the specimen. This means that one template will only match a small area, so you have to use many templates. These templates could, in principle, be totally empirical, but to be quantitative you must derive them from image simulations. This approach has been described for two special cases by Kisielowski *et al.* (1995) and Ourmazd *et al.* (1990).

### 30.10.A. The Example of a Specimen with Uniform Composition

In the QUANTITEM approach developed by Ourmazd, the results of Chapter 29 are summarized by a general equation linking the intensity and all the imaging ($S_i$) and materials ($P$) parameters,

$$I(x, y) = F(P(x, y), S_i) \qquad [30.5]$$

This equation just tells us that the intensity depends on the imaging conditions and on the specimen. For a particular set of imaging conditions, $S_i$ will be known (more or less) and we'll call it $S_i^0$. Then we can write that

$$I(x, y) = F\left(P(x, y), S_i^0\right) = F^0(P(x, y)) \qquad [30.6]$$

The basis of this approach is quite straightforward:

- Define the function $F^0$ for each image that you may obtain.
- Then construct a set of templates for your matching process.

Providing you stay within one extinction band, $F^0$ will be directly related to the projected potential of the specimen. Ourmazd gives a helpful simple analogy for this process, as illustrated in Figures 30.17A and B. The function $F^0$ describes the path of a swinging pendulum as it varies with time. Each value of $F^0$ corresponds to a snapshot of the pendulum, so if you plot $F^0$ you can "see" the path of the pendulum. The velocity of the pendulum is related to the density of points along the path. So it should be possible to plot out the function $F^0$ from a single lattice image even if you don't know the microscope parameters used to form the image.

Yes, there are limitations and conditions and we'll discuss them later. All we need now is a method for representing each image by a snapshot of the pendulum: we have to *parameterize* the image. This process is the key to the technique. Manipulating and quantifying, in principle, thousands of images, each requiring 4 Mbyte of memory, is not a fast process, even if you do have that much memory. If we could characterize each image by a few numbers (a vector or parameter), the comparison process could be much faster.

We separate the image into unit cells and digitize these to give many templates, which are $n$ pixels by $m$ pixels as shown in Figures 30.17C–E. If we define $N$ to be $n \times m$, then we have $N$ numbers for the $N$ pixels, where each number represents a gray level. Now the $N$ numbers are regarded as the $N$ components of an $N$-dimensional vector. (The math is not complicated, but don't try to visualize this vector.) So now all the information in each unit cell is represented by a vector in $N$-dimensional space. The function $F^0$ describes the path of these $N$-dimensional vectors as the projected potential changes.

The next step is to define a reference frame for these vectors. Three basis vectors are derived from the experimental image. Ourmazd *et al.* (1990) argue that three basis vectors will be sufficient, as we can show in the following way. We will be using a low-index zone axis for any HRTEM analysis. Then we have three types of images:

- The background, $\mathbf{R}^B$, due to the direct beam, O.
- A single-period image, $\mathbf{R}^S$, due to the interference between O and the strongest reflections, $G_i$.

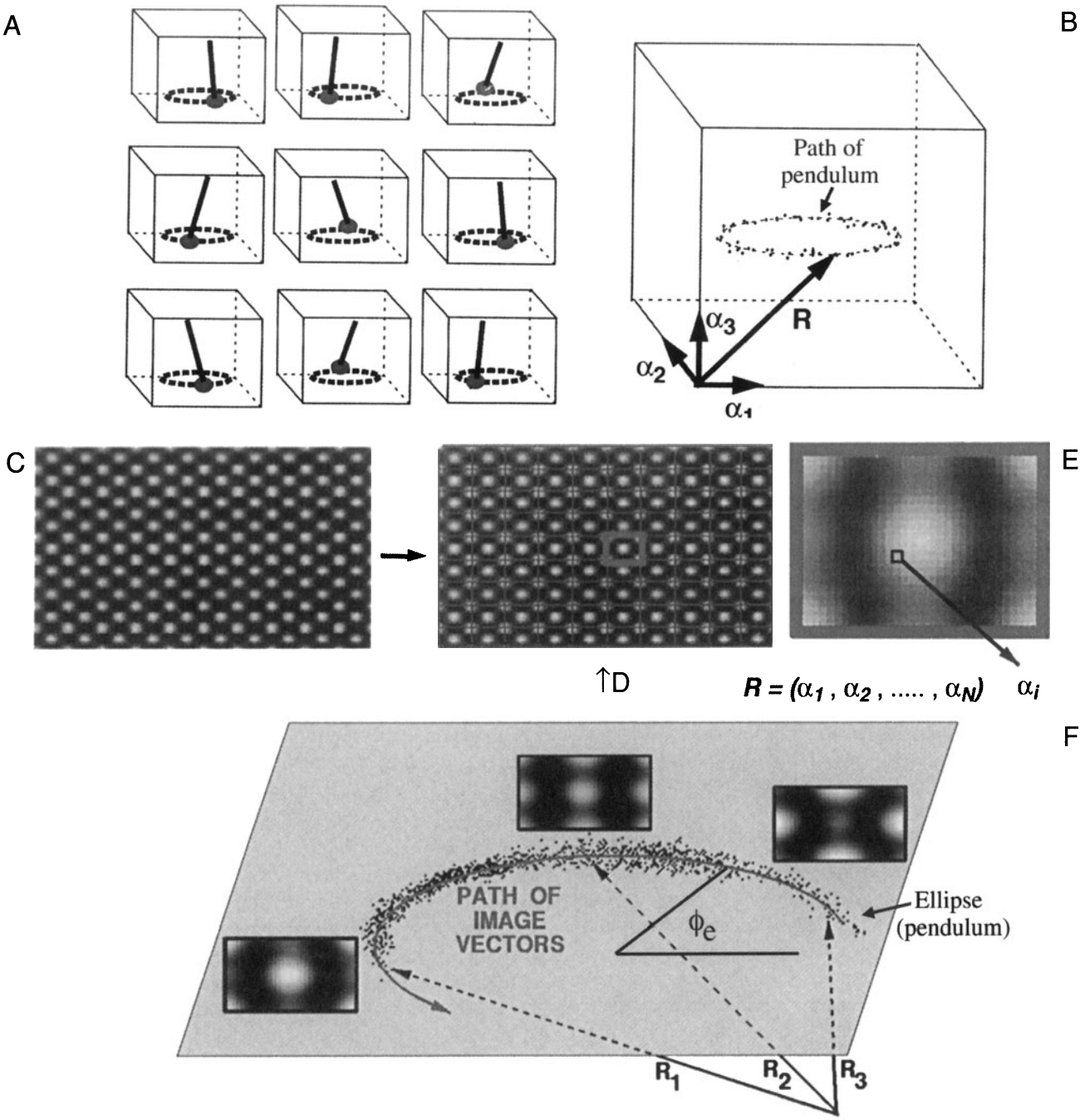


**Figure 30.17.** The principle of vector parameterization used in QUANTITEM. (A) shows the different pendulum positions and (B) shows the path of the pendulum. Each HRTEM image is represented by a single vector R which has *N* dimensions; (C–E) The image is separated into unit cells and digitized to give ($n \times m$) pixel templates; (F) Three vector-parameterized image ($\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$) of a wedge-shaped specimen of Si at different thicknesses.

- A double-period image, $\mathbf{R}^D$, due to the interference between these strong $G_i$ reflections.

Each of these **R** terms is a vector which represents an image. Any image we can form must be a combination of these three types of image, so a general image, G, can be written as

$$\mathbf{R}^G = a_G\mathbf{R}^B + b_G\mathbf{R}^S + c_G\mathbf{R}^D \quad [30.7]$$

Each of the basis vectors (images) can be expressed in the same manner

$$\mathbf{R}_i^T = a_i\mathbf{R}^B + b_i\mathbf{R}^S + c_i\mathbf{R}^D \quad [30.8]$$

giving three vectors for $i = 1$, 2, and 3.

We can turn these equations around to define any vector $\mathbf{R}^G$ in terms of the basis vectors

$$\mathbf{R}^G = \alpha_G \mathbf{R}_1^T + \beta_G \mathbf{R}_2^T + \gamma_G \mathbf{R}_3^T \qquad [30.9]$$

which is what we wanted to show.

Ourmazd *et al.* point out that this treatment gives three important results:

- The vector notation allows us to parameterize the lattice image.
- Projecting the vectors onto planes and/or paths aids noise reduction.
- Any noise which remains can be quantified.

The result of vector-parameterizing an experimental image of a wedge-shaped specimen of Si is shown in Figure 30.17F.

## 30.10.B. Calibrating the Path of R

In order to relate any image to the projected potential, we have to calibrate the curve showing the path of $\mathbf{R}^G$. This is where the image simulation comes in. We start with the vector-parameterized analysis of a series of simulated images as shown in Figure 30.17F. Each point on the curve corresponds to a unit-cell image, and thus to a vector $\mathbf{R}^G$. The ellipse has been fitted empirically, and the thickness of the cell has been increased by 0.38 nm for successive calculations. The points are closer together in some parts of the plot because, as we saw in Chapter 29, some characteristic images appear for a wider range of thicknesses. Now we have a way to quantify this "experimental" observation. What the ellipse does is to allow us to parameterize the path in terms of the phase angle of the ellipse $\phi_e$, shown in Figure 30.17F. Thinking back to the pendulum analogy, the path parameters are the image version of the coordinates for the harmonic oscillator.

Now the $\phi_e$ curve parameters can be obtained from a series of images. We can change the material and in each case examine three other variables:

- The orientation of the specimen (i.e., the zone axis).
- The defocus, $\Delta f$, of the objective lens.
- The specimen thickness.

The remarkable result is that when we plot $\phi_e$ versus the thickness, normalized by the extinction distance, we obtain a straight line. The explanation for this result is related to the fact that only a small number of Bloch waves usually contribute to the image, as we saw in Chapter 29. In materials such as YBCO, this is not the case, because too many Bloch waves are important and the curve is not a straight line.

## 30.10.C. Noise Analysis

If noise moves the vector off the ellipse, we can analyze the noise. If it moves the vector exactly along the ellipse, we can't analyze the noise, but that is quite unlikely since the noise would then be accurately mimicking a change in projected potential. So this method should reduce the noise by a factor of $\sqrt{N}$, which for a 10 pixel × 10 pixel cell is a factor of 10!

The analysis given by Ourmazd *et al.* then shows that, in the case where only two Bloch waves are excited, the image intensity, $I$, can be expressed as

$$I = B + S + D \qquad [30.10]$$

where $B$, $S$, and $D$ are the contributions from the background, single interaction, and double interaction, as we defined them above. The point ($B$, $S$, $D$) does indeed describe an ellipse which lies on a plane independent of $\Delta f$.

The value of this approach can be appreciated if you look at the examples shown in Figures 30.18A and B. In the first example, the technique has been used to provide a map of the roughness of the Si surface. The experimental image looks really uniform until you analyze it using this method, when you can discern the roughness at the ~0.5-nm level, as in Figure 30.18B.

As you know from earlier discussions, changes in chemistry produce effects which are similar to changes in thickness, because they change the projected potential. In terms of the present analysis the effects are different: composition changes cause changes in the ellipse and in $\xi$ (notice that there is no subscript, since this $\xi$ is a many-beam value).

The method is more limited in respect to change in composition but can be used if the thickness and roughness are known, i.e., if you can measure the roughness elsewhere on your specimen (using a known reference cell), and infer it for the area you want to analyze. (Warning lights should be flashing.) The approach is as follows:

- Use QUANTITEM to measure the advance in $\phi_e$ at your target cell relative to your reference cell.
- Subtract $\Delta\phi_e$, which is due to a thickness change.
- Then the rest of the change in $\phi_e$ must be due to changes in $\xi$. If you know how $\xi$ varies for dif-

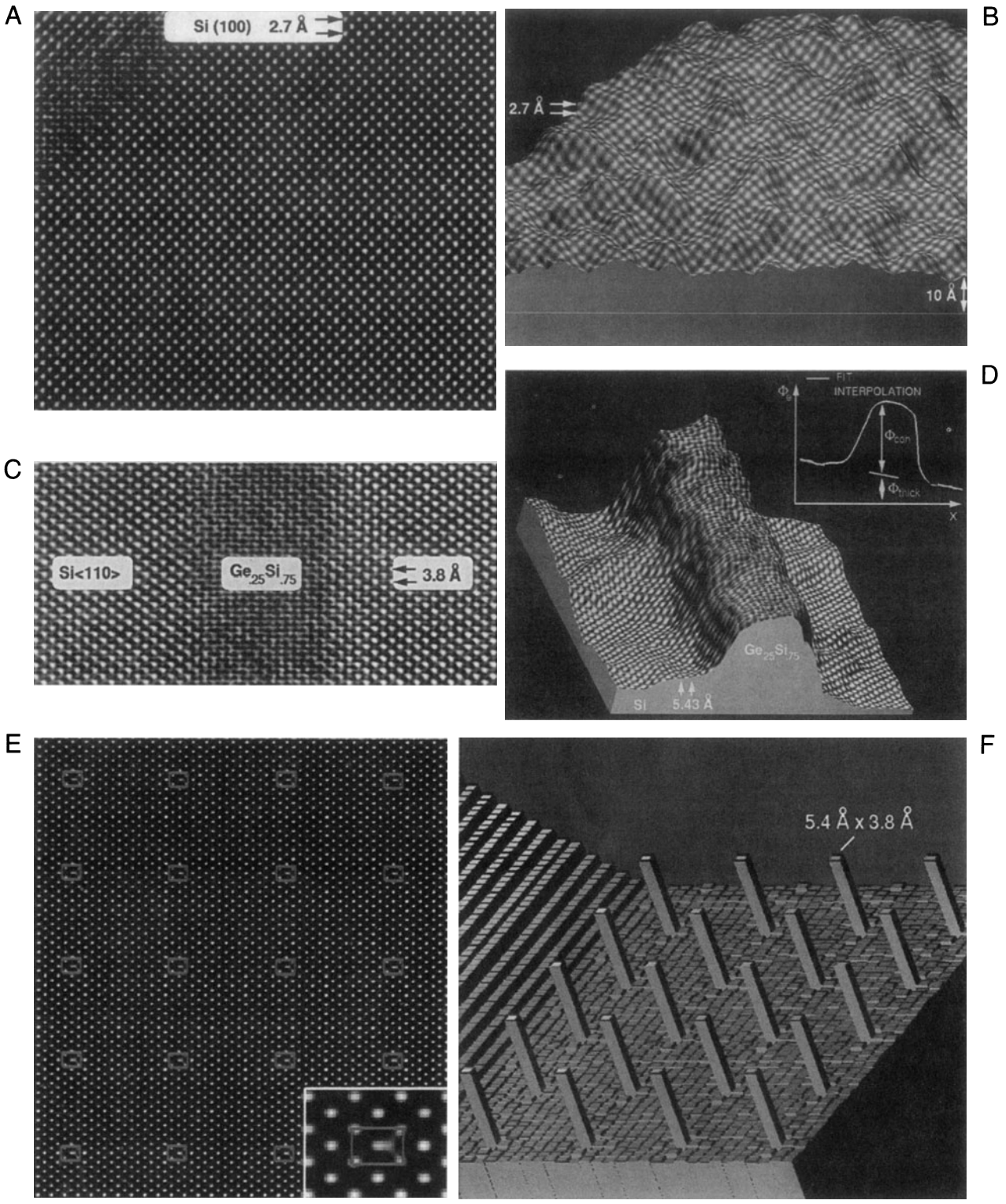


**Figure 30.18.** Examples of the application of QUANTITEM: on the left is the conventional HRTEM image, on the right is the QUANTITEM image. (A,B) Mapping of the roughness of the Si surface covered by $SiO_2$; (C,D) a layer of $Ge_xSi_{1-x}$ in a matrix of Si, the inset shows the plot of $\phi_e$ versus $x$; (E,F) analyzing a simulated image of columns of Ge (a δ-function in concentration) in Si.

ferent compositions, you have determined the local composition.

The example shown in Figures 30.18C and D is a near-perfect application for the technique, since the elements in the alloy $Ge_xSi_{1-x}$ are randomly located on the lattice sites. The slope tells us how abruptly the composition varies.

You can test the potential resolution of the technique, its sensitivity to the alignment of the beam, bending of the specimen, and beam divergence in the usual way by creating model structures, simulating the images, and then analyzing them. Figures 30.18E and F show that the potential resolution is superb, but beam tilt can cause 10% errors in thickness measurement. The conclusion is clear: as always, you will only get the best results if the specimen is ideal and both the microscope and the specimen were perfectly aligned. Note, however, that the technique has not yet been successfully applied to a wide range of materials, but it is complementary in many ways to STEM Z-contrast (see Figure 22.15).

## 30.11. QUANTITATIVE CHEMICAL LATTICE IMAGING

This technique uses the approach described in Section 30.10, but can only be applied to materials where we have chemically sensitive reflections, which we discussed in Chapter 17. We will use these reflections in Chapter 31 to produce chemically sensitive DF images. In HRTEM, the chemically sensitive reflections not only contribute to the overall image but they will generally have a different dependence on thickness, too.

This effect is shown in Figures 30.19A and B for AlAs and GaAs, which have identical structures. The 002 reflection is allowed for both, but is stronger for AlAs since $F$, the structure factor, is proportional to $f_{III} - f_V$. You can see that, under the conditions chosen for this comparison, the intensity of the 022 reflections is also very different for the different thicknesses. Figure 30.19C shows the sort of image we can analyze with this approach. We want to know how abruptly the composition changes at the interface.

In this example, the ideal GaAs and $Al_{0.4}Ga_{0.6}As$ unit-cell images are characterized by the two vectors, $\mathbf{R}_{GaAs}$ and $\mathbf{R}_{AlGaAs}$, following the approach described in Section 30.10. This was done in this case by simulating the cells, dividing them into 30 × 30 pixel arrays (so $N = 900$), and then plotting **R**. The information content is contained in $\theta_C$. As before, we can directly assess the noise in such an image. So how is the direction of **R** dependent on composition?

The technique is explained by Figure 30.19D. The three known simulated templates each produce a vector $\mathbf{R}^t$. Although the vector for the intermediate composition does not lie in the plane, it can be projected onto this plane to give a unique vector for certain ranges of thickness. Since this is a complex procedure, you'll find the flow chart shown in Figure 30.20 helpful.

- In Figure 30.20, the experimental image (A) is digitized; the image contained approximately 25 × 25 unit cells and used a 514 × 480 frame buffer.
- Next, the image must be separated into individual cells (B).
- The pair of templates shown (C) is then used to calculate the angular positions of the **R** vectors for all the unit cells. Such templates can be calculated or taken from known areas of the specimen.
- These **R** vectors are characterized in terms of where they cut through a plane (D) (see Figure 30.19D also).

The maximum chemical difference determines how far apart the two principal distributions can be (E). Since the image is now fully parameterized, we can do the statistics and finally invert the angular data to give the compositions (K).

This technique has enormous potential, but you must also remember that it is susceptible to all the drawbacks inherent in HRTEM. The big advance is that now you can put numbers on those effects. The technique is material-specific, but if you know your material, you can combine image simulation and this processing method to examine what will be the limiting factors for your material. You can construct a test image like that shown back in Figure 30.18E. If your specimen is ideal, you could, in principle, easily detect a column of Al in a mainly GaAs matrix without any "spreading" due to the electron beam.

## 30.12. METHODS OF MEASURING FIT

There are two methods presently used to obtain a measurement of the goodness of fit, namely:

- Cross-correlation.
- Least-squares refinement (Section 30.13).

In this section, we'll use the cross-correlation method to compare an $n \times m$ pixel template (see Section 30.9) with every possible $n \times m$ rectangle in the image. The computer

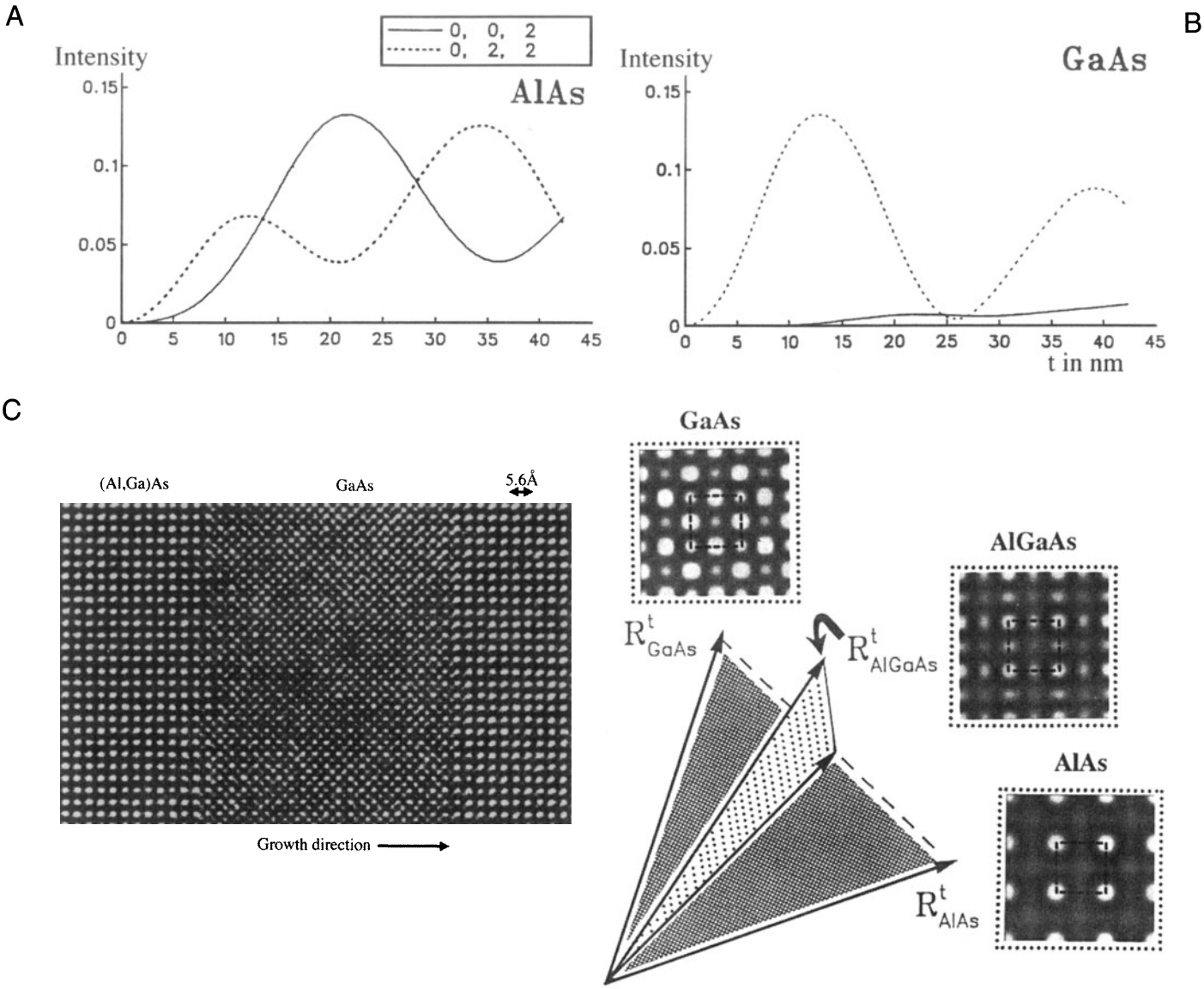


**Figure 30.19.** (A,B) Variation in intensity of the (002) and (022) beams along [100] in AlAs and GaAs (400 keV); (C) chemical lattice image of a layer of GaAs between two layers of Al$x$Ga$_{1-x}$As ($x$ = 0.4); (D) templates simulated for different values of $x$ each produce a vector $\mathbf{R}^t$.

moves the template across the image one pixel column at a time, then shifts down one pixel row, and repeats the exercise. The cross-correlation function (CCF) gives the goodness of fit or a "measure of similarity" between the template and each $n \times m$ image

$$CCF(x, y) = \frac{\sum_{x'}\sum_{y'}\left[i(x', y') - \langle i(x', y')\rangle\right]\cdot\left[t(x' - x, y' - y) - \langle t\rangle\right]}{\sqrt{\left\{\sum_{x'}\sum_{y'}\left[i(x', y') - \langle i(x, y)\rangle\right]^2 \sum_{x'}\sum_{y'}\left[t(x' - x, y' - y) - \langle t\rangle\right]^2\right\}}} \quad [30.11]$$

In this equation x varies from 0 to $x_{max}$, $y$ varies from 0 to $y_{max}$.

- $i(x', y')$ represents the image.
- $t(x', y')$ represents the template.
- $\langle t \rangle$ is the average value of the pixels in $t(x', y')$; it is computed just once.
- $\langle i(x, y) \rangle$ is the average of $i(x', y')$ in the region coincident with the current location of $t$.

The summations are taken over the coordinates common to both $i$ and $t$. The origin of the image is at its top left corner and the origin of the template is at its center. In this equa-

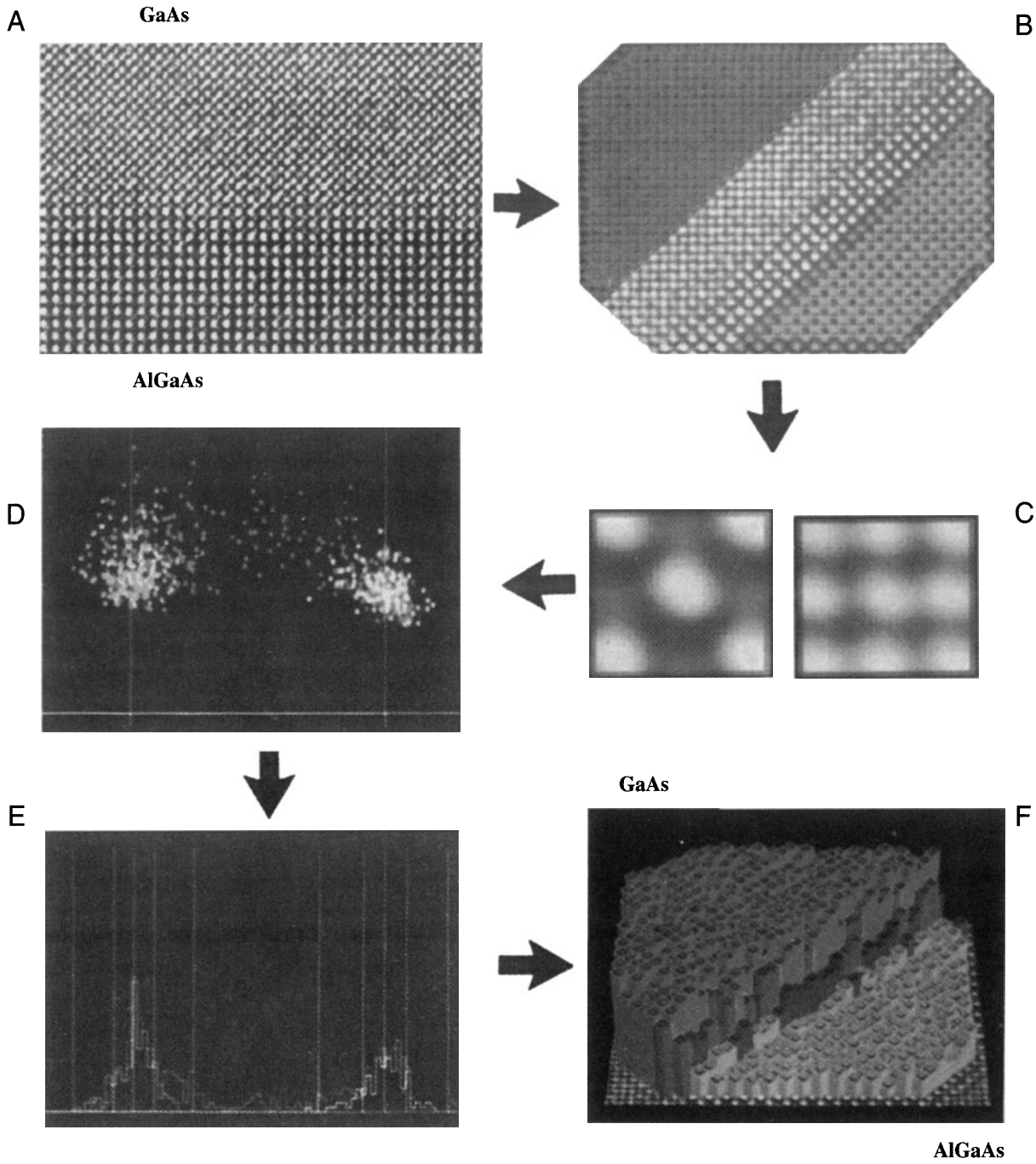


**Figure 30.20.** A flow chart summarizing the pattern recognition procedure.

tion the denominator is a normalization factor, so that the CCF will not depend on any difference in the intensity scale between the template and the image.

We can rewrite equation 30.11 as the dot product of two vectors **t** and **i** which gives us the $n \times m$ component of the template

$$CCF(x, y) = \cos(\theta) = \frac{\mathbf{t} \cdot \mathbf{i}}{|\mathbf{t}||\mathbf{i}|} \qquad [30.12]$$

Now we can plot the CCF as a map of our image and then examine it to deduce where there is a particularly good match. Since the CCF value varies from 0 to 1, we can plot

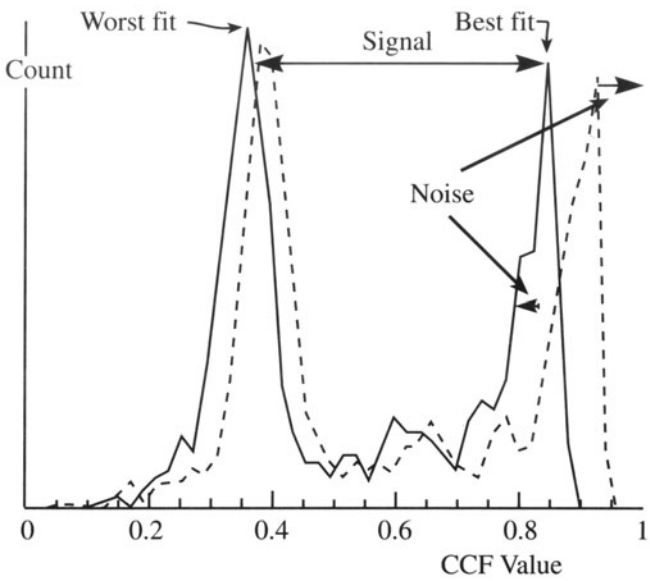


**Figure 30.21.** A plot of how often a particular CCF value occurs. The two peaks in the continuous curve are the best fit and the worst fit: their separation gives a measure of the discrimination signal; the width of the peaks gives a measure of the noise and hence a signal-to-noise ratio. The plot can be redrawn after repeating the process (dashed line) to estimate the improvement in signal-to-noise ratio.

out the number of times each particular CCF value occurs, as illustrated in Figure 30.21. The two peaks in this curve correspond to the best fit and the worst fit, so the distance between them gives a measure of the "discrimination signal." From the width of the peaks we have a measure of the noise, and hence a signal-to-noise ratio. The regions of good fit can be combined to produce a better template, and the process repeated, giving the dashed line. A second measure of the noise is then given by how far the good peak differs from unity. A particularly nice feature of this approach is that the procedure is available as a plug-in module for Digital Micrograph (see Section 1.5). Your template could alternatively be a simulated image and the process repeated for a series of different thicknesses and/or defocus values. When you want to learn more about correlation techniques, see the article by Frank (1980).

## 30.13. QUANTITATIVE COMPARISON OF SIMULATED AND EXPERIMENTAL HRTEM IMAGES

If we want to compare simulated and experiment images quantitatively, we really should modify our usual approaches to both simulation and experiment (King and Campbell 1993 and 1994). When doing the simulation, most programs automatically adjust the gray scale for each image so that darkest is 1 and brightest is 0 (or vice versa). This means that two simulated images might appear similar even though you would hardly see the pattern in one if both appeared on the same negative. In a similar way, we usually print an image to be as clear as possible using the full range of contrast of the photographic paper.

We need methods for normalizing these procedures if we want to make quantitative comparisons. The solution for the simulation is simple. For the experimentalist, it means recording extra data while you're at the microscope. After recording the image, you record another image with the specimen removed. You then use this image to scale your lattice image such that you correct for variations in intensity across the field of view and the nonlinearity of the response from the photographic film. Figure 30.22 illustrates the experimental transmittance for Kodak SO-163 film, 400-keV electrons, plotted against the digital value on a CCD array. Of course, you must process both images at the same time. This is called the "flat-field" correction; a slow-scan CCD camera would simplify this procedure at the cost of reducing the area you examine.

> Don't forget that since HRTEM uses higher voltages, the perfect image will only be recorded from an area of the specimen that has only seen the beam while you recorded the image! So you should always use low-dose techniques for quantitative imaging.

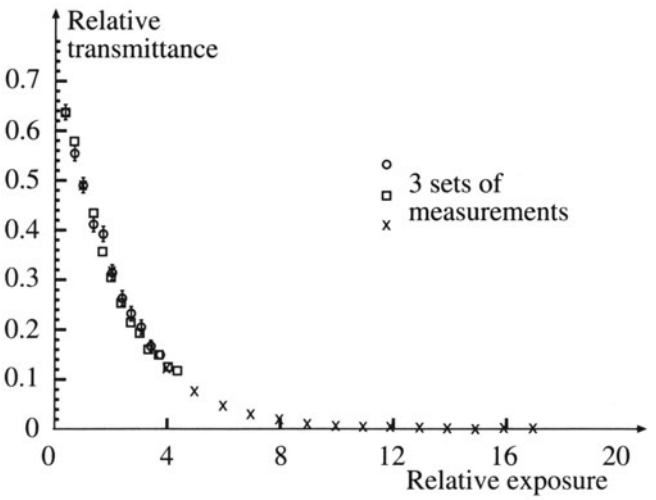


**Figure 30.22.** Plot of transmittance versus relative exposure measured using a CCD camera to digitize images from Kodak SO-163 film (the symbols indicate three different microscopes).

When you analyze the image, you'll find out if the area you photographed was correctly aligned. Since your image only takes two seconds or so to record, you may risk several exposures using this technique.

You are now comparing numbers, so you can use a least-squares fit where the residual $f_i(x)$ is defined as

$$f_i(x) = \frac{\left[f_i^{obs} - f_i^{calc}(x)\right]}{W_i} \qquad [30.13]$$

and your task is to minimize $f_i(x)$. The difference between the intensity in the experimentally observed $i$th pixel and its calculated value would be zero if everything had been scaled correctly, the imaging conditions ($\Delta f$, $C_s$, etc.) were correct, and you have the right structure.

Let's say $W_i$ is the image which represents the error bars for pixel $i$. Then we can write that

$$W_i = \min\left[\sum_{i=1}^{N} f_i(x)^2\right] \qquad [30.14]$$

This equation defines the nonlinear least-squares problem. We use $x$ to summarize a set of parameters ($\Delta f$, $C_s$, the model, etc.); $N$ is the number of pixels in the image. Fortunately, this analysis is now routine statistics. You'll need a computer program to tell you how good the first guess was, make an improvement, and continue until it meets our specific criterion for matching [King and Campbell used MINIPACK-1 (Moré 1977, Moré *et al.* 1980)].

In their demonstrations of this approach to analyze a [001] tilt grain boundary in Nb, King and Campbell (1993, 1994) varied four parameters: thickness, defocus, $x$-tilt, and $y$-tilt. The steps were as follows:

- They first optimized the electron-optical parameters using a 64 × 64 pixel image, giving $N$ = 4096 and an image computational cell of 3.303 nm by 3.303 nm. Using the EMS program (Section 1.5), the optimization took 20 iterations and 80 multislice calculations.
- Next, they had to optimize the structure of the grain boundary. This process required defining 84 atomic positions in a unit cell of 4.16 nm × 1.04 nm and a 512 × 128 (= 65,536) pixel image. Now the optimization required 16 iterations and 1300 multislice calculations.

These numbers are instructive. First they tell you that this computation can be done, which wasn't obvious. Secondly, they tell you that this is a computer-intensive process; that you could have guessed!

You'll need to take enormous care in this type of analysis:

- Align the simulated cell with the experimental cell and measure the unit cell in pixels.
- Choose a number of cells and relate them by the translation vector parallel to the rows of the image array.
- Calculate the standard-deviation images.
- Rotate the unit cell and repeat the exercise several times.

The orientation which gives the smallest standard deviation is your alignment. You must now adjust the magnification of the experimental image to fit the simulation, in a similar way to what you did for rotation. Next, you have to match the origins of both cells; the procedure is the same as we just described, but translating the unit cell, not rotating it. For a bicrystal, you now repeat this exercise for the other grain and then for the grain boundary. You can improve the fit further if you take account of a constant background contribution which probably arises due to the amorphous layer on both surfaces. Comparing experimental and calculated images quantitatively, we define $f_i^{obs}$ as the intensity value of the $i$th pixel in the experimental image and $f_i^{calc}$ as the corresponding value in the simulated image. We then calculate the residual $f_i(x)$ as follows

$$f_i(x) = \frac{\left(f_i^{obs}(x) - \left(f_i^{calc}(x) + b^{fit}\right)\right)}{W_i} \qquad [30.15]$$

where $b^{fit}$ is included as a free parameter in the optimization procedure. King and Campbell's calculations showed that $W_i$ could be expressed as

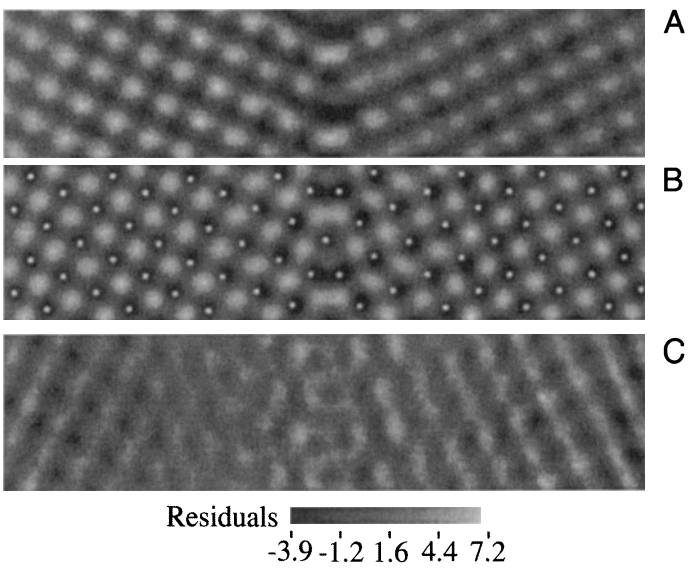


**Figure 30.23.** (A) Experimental image, (B) best-fit simulation, and (C) normalized residuals of a Σ = 5 symmetric tilt boundary in Nb.

$$W_i = \sigma_i^{obs} + 0.05 f_i^{obs} \quad [30.16]$$

where $\sigma_i^{obs}$ is the standard deviation of the $i$th pixel. Examples of the experimental, best fit, and normalized residuals are shown in Figure 30.23 for images from a $\Sigma = 5$, (310), [001] GB in Nb. The values for the thickness and $\Delta f$ show how consistent this technique can be, especially since the images in Figures 30.23A and B were from *opposite sides of the grain boundary*; (C) was like (B), but for a different defocus value.

## 30.14. A FOURIER TECHNIQUE FOR QUANTITATIVE ANALYSIS

Möbus *et al.* (1993) have proposed using what is referred to as an adaptive Fourier-filtering technique. The HRTEM image is digitized in the usual manner and then a special spatial-frequency filter is applied. This type of mask is most promising for analyzing regions which contain defects.

> An adaptive filter is one where the shape of the filter, or mask, is adapted to fit the shape of the "image" it's filtering.

So the idea is that the computer automatically optimizes the mask to maximize the separation of the signal and the noise. This approach has not been widely practiced in TEM but clearly holds enormous promise. By varying the mask, this approach can prevent the analysis of a defect layer being dominated by the bulk information. Since the approach is quite straightforward signal processing, we will just illustrate an example found in the analysis of a simulated $\Sigma = 5$ grain boundary with an extra period along the boundary. To test the analysis, white noise was added to a calculated image to give the image shown in Figure 30.24A. The power spectrum (computer-generated DP) of the micrograph is shown in Figure 30.24B. The adaptive filter and the filtered image are shown in Figure 30.24C and D. The important feature of the adaptive filter is that it was created as such because the computer detected the doubling of the periodicity, which is *only* present in the grain boundary. Secondly, the mask consists of elongated openings which we know we need when analyzing the grain boundary because of the shape effect.

## 30.15. REAL OR RECIPROCAL SPACE?

In principle we could equally well compare two images in reciprocal space rather than real space. However, while the Fourier transforms can generally be carried out much faster, the real-space approach has several advantages:

- Fourier analysis separates local information into sine and cosine functions which are delocalized. When we reassemble the real-space image, higher parts of the frequency spectrum will be lost which will degrade the resolution.

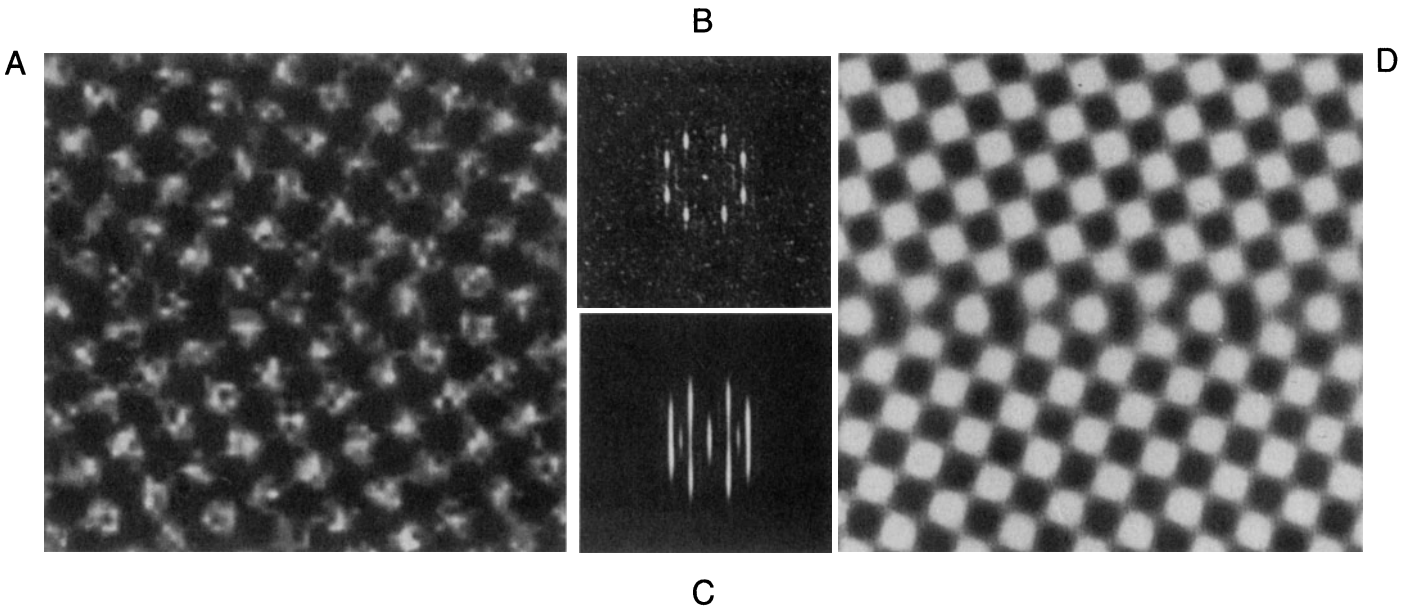


**Figure 30.24.** (A) White noise added to a calculated image of a $\Sigma = 5$ grain boundary; (B) The power spectrum of (A); (C) the adaptive filter; (D) the filtered image.

- We want to maintain information on the absolute value of the intensities.
- Real-space methods are visually more intuitive for most of us. We can easily see what we have removed in the process.
- The real-space approach allows us to choose any values of $n$ and $m$ in defining our templates. Fourier space prefers aspect ratios given by $2^n$.

## 30.16. THE OPTICAL BENCH

Although not widely used now, the optical bench is still a useful instructional tool. A typical experimental set-up is shown in Figure 30.25. The laser provides a coherent source of illumination representing the electron beam. The negative acts as the specimen. If it contains a set of lattice fringes, these act as a diffraction grating and give rise to a row of spots on the screen placed at the back focal plane of the "objective" lens. The lens is thus performing an optical Fourier transform of the photograph. If you move the screen to the image plane, the fringes reappear. You can make different masks and place them at the back focal plane, or even create an "adaptive filter" by exposing a photographic film and using this as the template for your mask. These masks correspond to our objective apertures. Students will find it instructive to transform their instructor or another suitable photograph, examine the frequency spectrum, and investigate the resulting spatial effect of different masks. The detail in the image is quickly lost as you remove the high spatial frequencies. This corresponds to inserting a smaller aperture in the back focal plane of the objective lens, as illustrated in Figure 30.26. So Figure 30.26D is effectively a BF image: you lose a lot of information in such images!

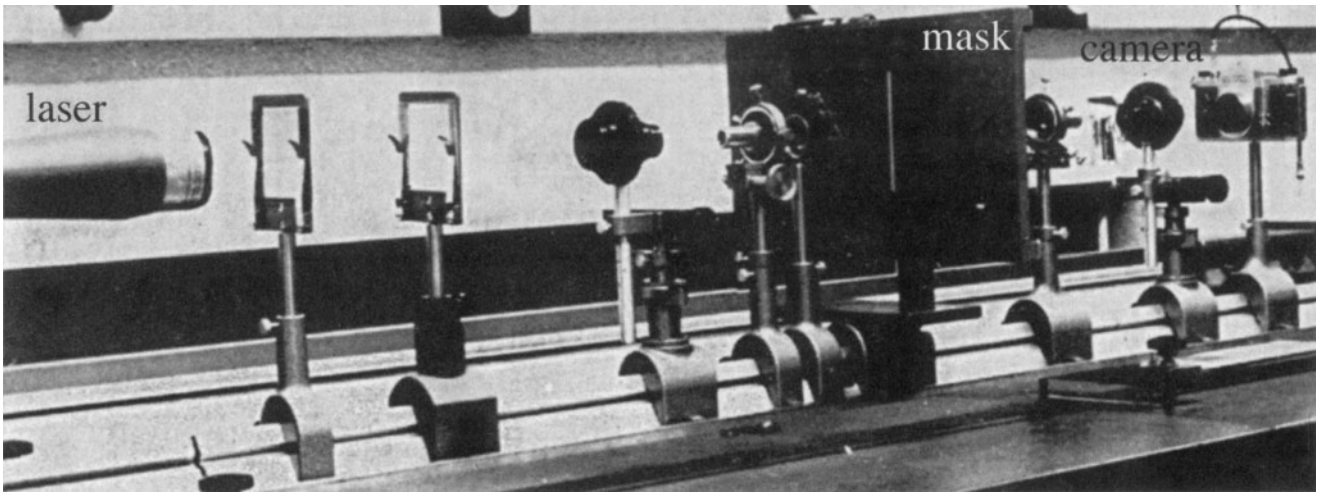


**Figure 30.25.** A typical experimental set-up for an optical bench, with the mask in the back focal plane.

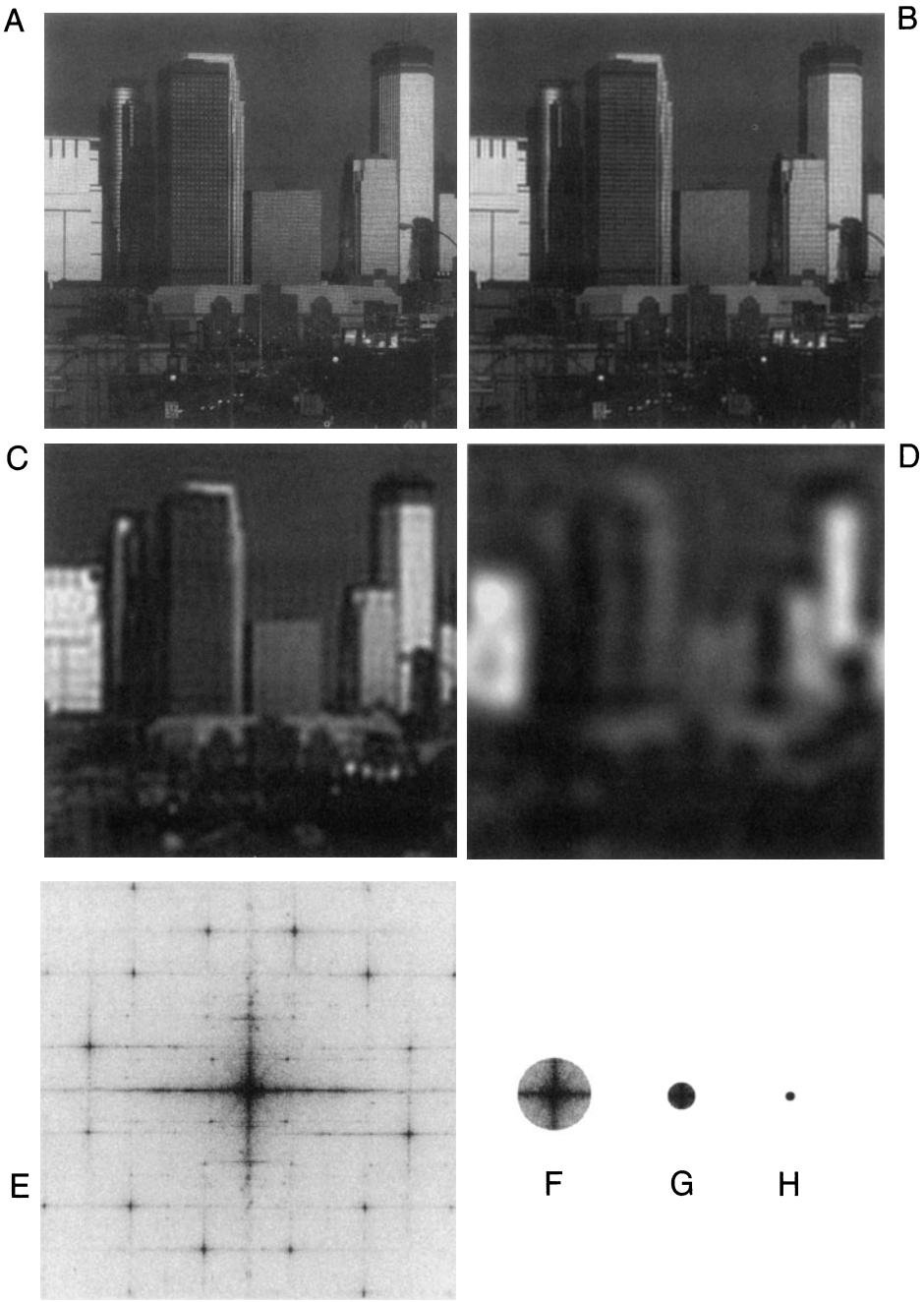


**Figure 30.26.** The effect of mask (aperture) size on a nonperiodic image of the Minneapolis skyline viewed from near the Guthrie Theater. (A–D) Reducing the aperture size, as indicated in the corresponding optical transforms (diffractograms) (E–H) reduces the image detail. The streaks in (E) arise from the edges of the photographs.

## CHAPTER SUMMARY

We have been doing image processing for many years; it's called "dodging" in the photographic darkroom. You can even do this with a special enlarger. However, we have done very little quantitative imaging in materials science. The points you should remember when starting in the field are the following:

- Quantitative comparison of simulated and experimental images depends on both the simulation program and your experimental parameters.
- If you are going to use reciprocal-space techniques for quantitative analysis, you should let the computer design the optimum mask as part of this process; usually, it will not be a circular mask, especially if you are studying interfaces!
- The potential for image restoration is not limited by the signal mixing due to $C_s$ and $\Delta f$. You can unscramble those effects. Ultimately, the limit is set by the signal-to-noise ratio in your image.
- You will notice the repeated use of the word "potential," where we don't mean $V(\mathbf{r})$! In many ways this chapter is a guide to the future of HRTEM and TEM in general. Some of the features won't be commonly available or optimized until the manufacturers realize their importance to the user.

There is always the possibility of removing information which is important. For example, Fresnel fringes often should be there! Beware of making reality match your simulation, rather than the reverse. In the same vein, we draw your attention to the conclusion of Hÿtch and Stobbs (1994), who found that they could only match their experimental and simulated images if they used a value for the specimen thickness which they knew was wrong! Their study emphasizes that, wherever possible, you should obtain independent measurements of the characteristics of your specimen and your machine. Remember the double-headed rhino in Figure 1.7; don't publish artifacts, even well-processed ones.

In this chapter we have discussed several different techniques used for processing TEM images. Several software packages are widely used by the TEM community and have been listed in Section 1.5. In its earliest application, image processing in TEM was almost exclusively applied to HRTEM images. This is no longer the case. Remember: always start with the best possible data. You can't always obtain a perfect image because your specimen might be beam-sensitive, or coated with oxide, and you need to be aware of these limitations when processing or quantifying the image. This chapter has given you a hint of what is possible and where the subject is developing. We recommend that you obtain the software and start experimenting.

## REFERENCES

### General References


Harrel, B. *et al.* (1995) *Using Photoshop for Macintosh,* Que, Indianapolis. A detailed description of the possibilities using Photoshop; this volume should also serve as a warning for all microscopists. Ask the question: has the image you are examining been processed, and if so, how?

Hawkes, P.W., Ed. (1980) *Computer Processing of Electron Microscope Images,* Springer-Verlag, New York. This collection of articles is for advanced students.

Hawkes, P.W., Ed. (1992) *Signal and Image Processing in Microscopy and Microanalysis,* Scanning Microscopy Supplement **6**, SEM, Inc. AMF, O'Hare, Illinois. This volume and its 1988 and 1996 companions are required reading for this subject.

Hawkes, P.W., Ottensmeyer, F.P., Saxton, W.O., and Rosenfeld, A. Eds. (1988) *Image and Signal Processing for Electron Microscopy,* Scanning Microscopy Supplement **2**, SEM, Inc. AMF, O'Hare, Illinois.

Hawkes, P.W., Saxton, W.O., and Frank, J., Eds. (1996) *Image Processing,* Scanning Microscopy Supplement, SEM, Inc. AMF, O'Hare, Illinois.

Russ, J.C. (1990) *Computer-Assisted Microscopy,* Plenum Press, New York. Chapter 3 is particularly relevant. Subsequent chapters give detailed analyses of the topics introduced here.

Russ, J.C. (1995) *The Image Processing Handbook,* 2nd edition, CRC Press, Boca Raton, Florida. A beautiful, comprehensive, and an essential component of any EM lab (or home).


### Specific References


Dorset, D.L. (1995) *Structural Electron Crystallography,* Plenum Press, New York.

Downing, K.H. (1992) Scanning Microscopy Supplement **6**, ibid., p. 405.

Erickson, H.P. and Klug, A. (1971) *Phil. Trans. Roy. Soc. Lond.* **B261**, 105.

Frank, J. (1980) in *Computer Processing of Electron Microscope Images,* p. 187.

Hÿtch, M.J. and Stobbs, W.M. (1994) *Ultramicroscopy* **53**, 191.

King, W.E. and Campbell, G.H. (1993) *Ultramicroscopy* **51**, 128.

King, W.E. and Campbell, G.H. (1994) *Ultramicroscopy* **56**, 46.

Kirkland, E.J., Siegel, B.M., Uyeda, N., and Fujiyoshi, Y. (1982) *Ultramicroscopy* **9**, 65.

Kirkland, A.I., Saxton, W.O., Chau, K.-L., Tsuno, K., and Kawasaki, M. (1995) *Ultramicroscopy* **57**, 355.

Kisielowski, C., Schwander, P., Baumann, F.H., Seibt, M., Kim, Y., and Ourmazd, A. (1995) *Ultramicroscopy* **58**, 131.

Koster, A.J., van den Bos, A., and van der Mast, K.D. (1988) Scanning Microscopy Supplement **2**, p. 83.

Koster, A.J. and de Ruijter, W.J. (1992) *Ultramicroscopy* **40**, 89.

Krivanek, O.L. (1976) *Optik* **45**, 97.

Krivanek, O.L. and Mooney, P.E. (1993) *Ultramicroscopy* **49**, 95.

Möbus, G., Necker, G., and Rühle, M. (1993) *Ultramicroscopy* **49**, 46.

Moré, J.J. (1977) in *Lectures Notes in Mathematics* (Ed. G.A. Watson), p. 630, Springer, Berlin.

Moré, J.J., Garbow, B.S., and Hillstrom, K.E. (1980) User Guide for MINIPACK-1.

Ourmazd, A., Baumann, F.H., Bode, M., and Kim, Y. (1990) *Ultramicroscopy* **34**, 237.

Paciornik, S., Kilaas, R., Turner, J., and Dahmen, U. (1996) *Ultramicroscopy* **62**, 15.

Saxton, W.O. (1992) Scanning Microscopy Supplement **6**, p. 405.

Saxton, W.O. and Koch, T.L. (1982) *J. Microsc.* **127**, 69.

Saxton, W.O., Smith, D.J., and Erasmus, S.J.J. (1983) *J. Microsc.* **130**,187.

Tang, D., Jansen, J., Zandbergen, H.W., and Schenk, H. (1995) *Acta Cryst.* **A51**, 188.

Trus, B.L., Unser, M., Pun, T., and Stevens, A.C. (1992) *Scanning Microscopy Supplement* **6**, p. 441.

van Heel, M., Winkler, H., Orlora, E., and Schatz, M. (1992) *Scanning Microscopy Supplement* **6**, p. 23.

Zou, X. (1995) *Electron Crystallography of Inorganic Structures,* Chemical Communications, Stockholm University, Stockholm.