Chapter 9

Scanning

9.1 Introduction

A scanner is a device that looks and works much like a xerox machine except that the copy is not a duplicate on a sheet of paper but rather a computer file. Scanner software can save the image in a variety of formats including as a Postscript file. The image can then be included as part of a book or scientific publication; many journals now can receive Postscript files as email and print directly from the file on their very high resolution printers. The image can also be imported into a drawing program and then edited and modified.

Low cost scanners cost only $200 or less, but still do a good job of scanning line art at 300 dpi, and are fair at scanning color photographs or images. Mass-market magazines employ high-resolution scanners (as much as 2000 dpi) to convert colored drawings, advertisements and photographs into a form suitable for layout and typesetting software.

9.2 Why-To or Reasons Scanning is Useful in Scientific Computing

1. To include graphs and diagroms from previously printed work in computer-generated journal articles and reports.

2. To use previously published images as a starting point from a new diagram or image. (A scanned image can be traced in a drawing program, combined with other FORTRAN or Matlab-generated graphs, or otherwise modified to suit present needs better than the original picture.)

3. To extract numerical information from previously published graphs ("data theft").

9.2.1 Recycling Graphs

To introduce a topic or provide background, a previously published graph is sometimes worth more than a dozen paragraphs of text. One then has a choice: regenerate the graph oneself, which may take a year of work, or scanning/xeroxing the image from a book, and asking permission of the copyright owner to reprint the picture.

Xeroxing is sufficient if your own work is to be typeset by the publisher. However, for a computer-generated manuscript, it is highly convenient to have the copy in a computer file, and that requires scanning.

9.2.2   Scan and Modify

Much scientific and engineering work consists of doing old work better. It is then very useful to publish a direct, graphical comparison between the old and the new. The most convenient way to make such an old-and-new graph is to scan the old picture in and then use a drawing program to combine this with your new computer-generated image. One could use a professional draftsman to combine the curves. However, scan-and-modify is almost certainly what the artist will do!

9.3   Scanner Types

Scanners come in several species:

1. Flatbed scanners

2. All-in-one printer/scanners

3. Slide/film scanners

4. Photo scanners

5. Drum scanners

6. Oversize scanners

7. Handheld scanners

8. Digital cameras

Flatbed scanners are the most common. They vary in price, but a decent scanner is only $100. Many home computers now have flatbed scanners attached, and many stores bundle computers with scanners.

Flatbed scanners usually employ a laser as the light source; xerox machines also employ laser light sources, and indeed the mechanics and electronics of the moving laser is very similar in both types of equipment. The reflected or transmitted light is captured by a CCD array just as in a digital camera. However, the scanner array does not need to be as dense or as sophisticated as in a camera because the camera must capture the entire image in a fraction of a second whereas the scanner array literally scans the image over a few seconds, capturing a tiny piece of the image at each instant.

Because there is great commonality between scanners, xerox machines and laser printers, so-called "all-in-ones" have become popular. These combine a flatbed scanner with a copier, printer and fax machine. The earliest all-in-ones printed only black-and-white, using a laser printer mechanism. (The laser did double-duty for both printer and scanner). Many new models combine a color ink-jet printer with a scanner; these can make color copies.

As scanners, the all-in-ones are usually less capable than stand-alone scanners. The great selling point of all-in-ones is that they are very compact, allowing several functions in a container which can fit into a crowded home office or computer study. Consequently, all-in-ones rarely allow the scanning of images bigger than 8.5" x 11" whereas

larger sizes are common in flatbed scanners, even cheap ones. However, many all-in-ones offer the option of an automatic document feeder (ADF).

Slide/film scanners are designed to scan *transparent* media such as slides or film in *small* sizes at very *high resolution.* Inexpensive slide/film scanners usually have much higher dots-per-inch ratings than even top-of-the-line flatbed scanners, but are limited to media that will fit in a holder that can accomodate only 35mm film or slides.

A recent variation is the dedicated photoscanner. A developed 3" x 5" scanner is an opaque media that can be scanned perfectly well on a flatbed scanner. However, the photo scanners are more compact than flatbeds because they are just big enough to accomodate a photo. Their optics has been designed so as to trade moderate resolution over a wide area (flatbed) for higher resolution over a small image size (photoscanners).

If one intends primarily to scan either slides or photos, a special-purpose scanner will give better results, and take up less space, than a general purpose flatbed scanner. However, a flatbed is more flexible in that it scan just about anything as explained below.

A drum scanner requires that the scanned object be thin and flexible so that it can be wrapped around a drum which rotates in front of the light source and light collector. Drum scanners have much higher resolution than flatbed scanners because they employ a different light source (halogen lamps) and different light collector (Photomultiplier Tubes, alias PMT). Such scanners are so expensive (many thousands of dollars and up) that it is unlikely that an individual scientist or even a large research group would ever own one.

Nevertheless, drum scanners are important to the scientific community. The reason is that publishers use them as the standard means of scanning documents for publication. These days, publishers try to bypass the scanner by collecting electronic image files from authors whenever possible. However, because of the lack of standardization of image files and so on, journal publishers are routinely forced to fall back on their drum scanner. This is particularly so for photographs and other images which are initially generated on a non-electronic media such as film; the publisher would much rather use his high resolution drum scanner for a photo than accept an electronic image that is the author's low resolution scan of the photograph.

In addition, it should be noted there are "service bureaus" in most large cities that serve the graphics design and publishing industries. An engineer who needs an ultra-high resolution scan can take his image to a service bureau and pay a small charge to have it drum-scanned and returned as an electronic file on a Zip disk or CD-ROM.

There are oversize scanners that typically have a long bar that can accomodate paper as much as two or three feet wide plus a feed mechanism to pull the paper through the scanner and past the imaging equipment. These scanners are sometimes found in large engineering and architecture establishments to allow the scanning of blueprints and architectural drawings, which often are of rather heroic size. These scanners also are too expensive to buy, but often not that hard to "borrow" by using a service bureau, or a centralized university facility.

Hand-held scanners received a lot of publicity a few years ago. The theory was that these little scanners could be slipped in a backpack and taken anywhere. Instead of xeroxing articles in the library, one could use a hand-held scanner to capture the image electronically.

Unfortunately, the head is a lot less steady than the mechanical track of a flatbed scanner, so the quality of hand-held scans of an 8.5" x 11" page is awful. Such scanners are still sold, but never became popular.

Nevertheless, small, portable scanners will likely make a comeback. First, scanner technology has improved and price has come down. Second, libraries in the near future

may have coin-operated scanners similar to the coin-operated xerox machines that they already have.

Digital cameras are not usually thought of as scanners, but sometimes are used that way. The cartographer David Rumsey has put his library of 6000+ historical maps online. Because the maps are very large, they cannot fit in a standard flatbed or drum scanner; because many are a couple of centuries old, they are too fragile to be pulled through an oversize blueprint scanner. His procedure is to spread the maps on the floor, illuminate them with standard photographic spotlights, and then capture the map using a high-quality digital camera mounted on a standard several feet about the maps.

9.4   Scanner Substitutes

9.4.1   Digital Cameras

As noted previously, digital cameras are often the best way to scan very large maps. Flatbed scanners have been used in the past to scan three-dimensional objects such as coins or leaves, etc., taking care not to scratch the glass surface. However, digital cameras are probably better for imaging three-dimensional objects. (The exception is very small objects that can be laid on a flatbed, but not easily photographed because a "close-up" would be outside the focal range of the camera.)

Digital cameras vary widely in cost and resolution. The latter is usually expressed in "megapixels" where a megapixel is a million pixels, or in other words a 1000 x 1000 bitmap. How many megapixels does one need?

A digital image should be printed at a resolution of 150 dpi to 300 dpi. At less than 150 dpi, the image looks fuzzy and pixelated to most observers. A picture at more than 300 dpi is indistinguishable from a print at 300 dpi even for the most keen-sighted observer. In general, a picture printed at about 200 dpi looks crystal clear. It follows that 200 dpi is a good resolution goal for digital cameras.

The number of megapixels required to reach this goal depends on two issues:

1. Final print size

2. Cropping

Table 9.1 shows the number of pixels needed at 200 dpi resolution for various print sizes. The lowest resolution digital cameras, costing \$35 to \$50, create images that match the VGA screen resolution, 640×480. For a passport photo, even the most humble of digital cameras would seem to be adequate, but this is not the whole story.

First, digital cameras (and film cameras) are bothered by imperfect color fidelity; people photographed with a VGA camera often look slightly seasick ("green cast") or drunk ("red cast"). More expensive cameras reproduce colors more faithfully than cheap cameras, if color fidelity is important.

Second, and more important, digital cameras allow *very easy cropping.* An ordinary photograph can be physically cropped by cutting away extraneous parts of the picture with scissors. Unfortunately, this renders the picture a non-standard size. The cropped picture will fall out of our scrapbook, and look wierd when viewed on the same page with a lot of standard size figures.

In contrast, digital pictures may be cropped with reckless abandon in Photoshop or similar programs. The image can then be reexpanded to a standard size.

This goes double for scientific graphics. There is no standard size; a graph can expanded to fill an entire column width in Latex without even bothering to alter the size in Photoshop.

Furthermore, there is no such thing as "pretty background scenery" in a scientific image; it is desirable to crop away everything except the experimental apparatus, or whatever is the focus of the picture.

Thus, it is a little tricky to use a digital camera as a scanner substitute. A picture that may be satisfactory as a 3" x 5" image may be cropped to 2" x 3" and then published as a 3.2" x 5" — with visible pixelation and streaking.

In estimating resolution needs, one must allow for cropping and also estimate the final *published* size of the cropped image.

The conventional wisdom is that a six-megapixel digital camera is needed to match the resolution of 35 mm film. Cameras in this range are now available, but very expensive (thousands of dollars) in 2002. However, prices are dropping rapidly.

Digital cameras have one significant drawback compared to film; they are relatively slow at capturing images. Film speed is rated from AS100 to AS1000; by using the latter and a very short exposure, a 35 mm camera can freeze very rapid motion without blurring. Digital cameras have a speed roughly equivalent to AS100. Consequently, sports photographers are forced to use film for many applications even though they can afford six megapixel digital cameras and love the freedom to bypass the delays of development, and go straight to digital editing.

9.4.2 Scanning Substitute: PDF files

The American Meteorological Society has archived all articles in its journals, going back to the founding of each journal, as PDF files. These can be opened in Photoshop and individual figures extracted and modified.

1. First, open a PDF file such as **KeyseretalBAMS01.pdf**. A little window will appear with the legend "Generic PDF Parser".

2. Select the desired page number by clicking on the arrows and watching the page marker change. Click OK. A window will appear that allows one to select the numbers of pixels per inch, etc. Click "continue". The selected page will then be rasterized.

3. Erase unwanted bits of text and figures to leave only the desired figure. The easiest way is probably to simply crop the Photoship image around the desired picture.

4. Add text to the figure as desired, and whatever modifications or adornments as seems appropriate.

5. Save the image in Photoshop format. Save it again as an EPS file if that format is needed.

Table 9.1: Pixel Requirements of Digital Pictures

Print Size	Pixels at 200 dpi	Total pixels	Print name
1.375" x 1.75"	275 x 350	0.10 megapixels	passport photo
3" x 5"	600 x 1000	0.60 megapixels	standard snapshot
5" x 7"	1000 x 1400	1.40 megapixels	medium framed picture
9" x 12"	1800 x 2400	4.30 megapixels	oversized

When all this is done, one obtains the same result as an annotated scan, but without actually using a scanner

9.4.3   Scanning by Video Camera Frame-Grabbing

This is effective. However, it is limited to low-resolution images (640 x 480 pixels) on older video cameras. Some of the newer digital camcorders have higher resolution, but even the best are inferior to a good digital camera.

9.4.4   Scanning by Photo CD

With this option, film is developed and returned to the photographer as a CD-ROM. The scanning is effectively done with a pretty good drum scanner. The CD contains several image sizes for each photo. The largest size image is 2048 x 3072 pixels.

9.5   Copyright and All That

Many images, including most pictures in a scientific journal or a book, are protected by copyright. Some images are now protected by a "digital watermark" which is encoded in the pictured; decoding software can recover the watermark for a scanned image and so prove that the image was copied in violation of copyright laws. There are several ways to circumvent copyright.

1. Use old images for which copyright has expired. For example, images created in the nineteenth and early twentieth century are now in the public domain and can be freely copied. Some books contain collections of such images; these books of "clip art" explicitly state that these images are not protected by copyright.

2. Trace the scanned image and then delete the scan.

3. Ask permission of the copyright owner.

9.5.1   Copyright-Expired Images

American and international copyright laws contain time limits. For example, until recently, copyright protection expired 56 years after the original publication. Changes in the copyright law (about 20 years ago) altered the status of publications after that date: copyright now extends fifty years from the death of the author. All graphs published before 1942 are now in the "public domain", which means that copyright law no longer protects them and these images can be freely copied and republished.

There are subtleties. For example, da Vinci painted the *Mona Lisa* in the early sixteenth century and it is therefore not directly protected by copyright. However, the Louvre Museum owns the physical copy of the painting and does not allow unauthorized reproduction. One could copy the image from a book, but technically, the book is protected by copyright. In practice, since the image has been published in many books, one can safely scan it as long as one crops away all information that might indicate the book where it was published.

Government authors of scientific publications are required by law to place them in the public domain, and publishers have been forced to respect this. Such publications can be a useful source of royalty-free, copyright-free scanned images. But there is still a moral and cultural obligation to credit the original source in your caption.

9.5.2   Scan, Trace and Delete the Original

Drawing programs such as Adobe Illustrator support graphical layers. One can copy the scanned image into layer one, trace the image in layer two, and then delete layer one. Some specialized programs, such as Adobe Streamline, are designed specifically to do high quality automatic traces; Illustrator has a lower quality automatic trace. Even the best automatic trace usually requires some manual clean-up, but auto-tracing can greatly improve the usefulness of scanning.

Copyright law is not infringed as long as the trace is modified so that the final image is different from the original. (How different? Not very, but ask your lawyer!) The law states that the *expression* of an idea can be protected by copyright, but not the *idea itself*. Thus, one can always draw (or auto-trace and modify) a graph of the partial sums of the piecewise linear (or "sawtooth" function) even though such graphs have published many times before. The *idea* of a partial sum graph cannot be copyrighted. However, a pixel-by-pixel copy of someone else's published partial sum graph would be a violation of copyright law.

9.5.3   Asking Permission

Since scientists want maximum publicity for their own work, obtaining permission to use a copyrighted image is easy. One simply sends a form letter to the author.

The subtlety is that most journals and professional societies require the book or article author to transfer copyright to the publisher. However, to avoid annoying the author, publishers usually grant only conditional permission: the permission takes effect only when one has also obtained the consent of the author if still living. So, one must write to *both* the author and publisher. If the article has multiple authors, one must write to *all* the authors unless one authors signs on behalf of all the authors.

Only once have I been asked to pay a fee to reuse an image. A textbook publisher asked $200 to use an image of the partial sum of a Fourier series for the piecewise linear function. I regenerated the graph using Matlab, and paid nothing.

9.6   How-To

A scan is composed of mandatory steps plus others that are optional. The steps that must be done for every scan are the following.

1. Place the image to be scanned on top of the glass sheet and drop the cover over it, as far as the cover will go without forcing. This step is identical to making an ordinary xerox.

2. Open the scanning software, such as Ofoto which is standard for Apple scanners, and click the button to begin the scan.

3. Save the image as a computer file of the chosen file type: Postscript, Encapsulated Postscript (EPS), Pict [Macintosh only], TIFF, etc.

The scanner *hardware* almost never requires manipulation by the operator except to turn it on and off and to put the transparency adapter in place. Almost all options are implemented by clicking in the scanner program.

These options include the following:

- Scan as line art, grayscale, or color photograph.
  If an image is drawn only in black lines, dashes, and curves without grayscale shading or color, then it is a good idea to click the "line art" option on the scanning software. There are two reasons. First, line art is stored as one-bit-per-pixel. Grayscale is 8-bit and color as 24-bit, which implies that at a given number of dots-per-inch, a grayscale file will be eight times as large as a line art scan of the same image. Second, line art images, if scanned as grayscale, will often be rather muddy-looking as the scanner represents the background as pale shades of gray.

  Some scanner software, such as Ofoto, provides the option to *scan* line art at 8-bit resolution even though the result will be *stored* and *printed* at one bit resolution with the "threshold" option. The reason that the in-memory grayscale scan is useful is that one can observe what areas are treated as black and which areas as white by the default threshold criterion. If the scan target is a xerox, for example, the paper is likely sprinkled with a lot of tiny random dark spots. If these little gray noise blotches are being interpreted as black, then altering the threshold so that only really black pixels are printed as black may effectively eliminate a lot of noise from the scan.

- Dithering & Halftones

  Most scanning software allows several 1-bit options. The best choice for line graphs is usually "threshold", which interprets all sufficiently dark pixels as black, and all pixels lighter than the threshold as white. (The threshold can be adjusted in the "Tone Controls" menu of Ofoto software, and similarly in other scanning software).

  Another option is "dithering". This tries to correct for the "jaggies". The best option is to scan at higher resolution, but if this isn't possible, dithering may be helpful.

  "Halftone" represents grayscale images by little black dots of various sizes and densities, similar to the way photographs are printed in newspapers. If the image is really grayscale, then one should scan it as a grayscale image. The little black dots of a halftone image will likely "bleed" when printed in ways even more unpleasant and uncontrollable than when grayscale images as printed. (To put it another way, scanning software and laser printers do a better job with grayscale images when they know the images are grayscale.)

- Printer & Monitor Calibration

  Scanner software will optimize the image file so that it will print as well as possible on a given model of printer. (The software manual describes how to do calibrations for your particular printer; the software has been built-in calibrations for many common models of printer). This calibration is particularly helpful for *color*, less so for grayscale, and almost irrelevent for line art.

  The scanner software will generate a standard chart to be printed, and then scanned. By comparing the scan of the calibration chart with its exact specifications, the software can "tweak" the scan so that the printed image resembles the scanned image as closely as possible.

  The monitor can be similarly calibrated. Both printer and monitor calibrations are stored as files, to be applied as needed.

- Manual rotation.
  The scanner software will try to identify the top edge of the page and then rotate

the image is right side up. Of course, if one placed the paper absolutely level, this automatic rotation would be unnecessary. However, it is difficult to place the paper perfectly flush against the top of the glass sheet; even if the placement is perfect, the act of closing the cover may twist the paper a couple of degrees. So, the scanner tries to compensate. Often, automatic rotation fails either because the image to be scanned is twisted relative to the top of the sheet — common when scanning a xeroxed page, for example, or because the contrast between the top of the page and empty space is so weak that the software cannot detect the top edge. The software allows one to rotate the image manually when automatic rotation fails. Usually, only a rotation of one degree, either way, is sufficient.

- Cropping
When a xerox of a journal article is scanned, the scanner will scan the entire 8.5" x 11" page. If the goal is to copy a graph that occupies only a fraction of the page, then the extraneous material needs to be deleted. The cropping option allows one to move a box around the target image, using the mouse and cursor, and then cut away everything else.

  Cropping is easy but important because it tremendously reduces the size of the image file

- Size Transformations
Scanning software allows one to change the size of an image. This is rarely needed because Latex, for example, can scale images when it prints them. (The image size can also be changed in a program like Photoshop or Illustrator).

- Photoshop filters
Photoshop is a program for transforming bit-map images, which are the only kind generated by a scan. Photoshop offers a wide variety of "filters" to distort the shapes, change colors, and do almost anything one wishes to an image. The computer shelves at a local bookstore will offer a large collection of books devoted entirely to Photoshop.

- Contrast

  Scanning software usually has a built-in filter to manipulate the contrast of an image.

- Lighten/Darken, Balance and Exposure

  This lighten/darken filter allows one to manipulate what artists call the "key" of an image. A very dark image, composed almost entirely of dark grayscale, can be lightened to use a wider range of grays.

  Underexposure or overexposure of a photographic image will artificially change the key of an image. Therefore, Ofoto has a "balance" option which corrects for under/over exposure by automatically adjusting the image to a mid-range key. If this is inadequate, one can manually apply the lighten/darken filter.

  "Key" correction can be done more precisely in Photoshop. However, if one is going to scan a whole bunch of very dark or very light images, it may be useful to adjust the default on the scanner to automatically apply a lighten/darken filter, if the scanner software allows this.

- Resolution
Usually specified in dots-per-inch (dpi). The default for line art is 300 dpi (which is the resolution of most laser printers). The default for photographs is usually 150

dpi because color images yield very large files when scanned at high resolution. However, the software allows one to vary the resolution within wide limits.

9.7   Choice of Scanning Resolution

There is a tradeoff in the user-chosen resolution of a scan: high resolution yields very large file sizes. It is silly to scan at 300 dpi if the image is intended for a Web site where it will be displayed at 72 dpi. What choices are good for different situations?

The usual rules-of-thumb are the following:

Rule-of-Thumb 1 (Scanning Resolution) *If an image is to be used at its physical size at a desired resolution of X dpi (dots-per-inch), then the scanner resolution should be set, in the scanning software, to scan at 2X dpi. In other words, scan at twice the desired resolution.*

This rule is most significant for making Web sites. A computer screen has a resolution of only 72 dpi, so it is sufficient to scan images for the Web at 150 dpi. The exception is if the Web image is going to displayed at a larger size than the original. If a tiny picture only 3 cm on a side is to be displayed as 6 cm square screen image, then the original should be scanned at 300 dpi.

This rule isn't too useful for published images because journal typesetters have a resolution of 2000 dpi, far beyond the capabilities of most scanners in a scientific or engineering facility. Even a low cost laser printer has a resolution of 300 dpi; the better models are 600 dpi. This is roughly the same as the maximum resolution of low and medium cost scanners.

For published images, the alternative rule is the following:

Rule-of-Thumb 2 (Scanning Resolution for Print Publication) *Scan at the print's highest resolution, especially for line art [black-and-white images].*

Although not strictly a part of "visualization", scanners are often used to convert printed *text* into electronic form by first scanning the text and then deciphering using Optical Character Recognition (OCR) software.

Rule-of-Thumb 3 (Scanning Resolution for OCR) *Sullivan(1995), pg.   47, states that the best procedure is to scan at 300 dpi, black-and-white, which is the resolution that most OCR software is optimized for.*

Grayscale just slows down the OCR software and may degrade quality a little.

Drawing programs like Adobe Illustrator allow one to import scanned images. An imported bitmap or EPS image may have some undesirable features such as very large file size and extraneous graphical elements. Often the simplest way to keep the essential parts of the image while eliminating the undesirable features is to trace the scanned image.

Of course, manually tracing a complicated image can take a long time. Illustrator and its ilk therefore have an "auto-trace" option which lets the software do the tracing. Unfortunately, auto-trace is far from flawless. First, software in the year 2002 isn't capable of making the subtle judgments that a human can: should two adjacent lines, perhaps separated only by scanner error, be joined into one or left with a pixel-size gap in between? Second, such decisions about what to join and what to smooth are problem-dependent, and even very sophisticated software will be unable to make all

the decisions correctly unless the software is somehow educated in the specifics of the problem.

Nevertheless, auto-trace can take much of the work out of tracing a scanned image if one is willing to do some manual clean-up afterwards.

Rule-of-Thumb 4 (Scanning for Auto-tracing) *Sullivan (1995), pg. 47, states that "For autotracing purposes, only one thing counts — the highest possible resolution."*

On pg. 80, he suggests that a good way to cope with a low resolution scanner (300 dpi) is to scan as a grayscale image. This is less subject to the jaggies than a black-and-white image.

9.7.1   Resolution and Resizing

When one is saddled with a printer which has *lower* resolution than the scanner, one useful trick with a small image is to rescale it to larger size. For example, suppose the scanner has a 600 dpi resolution but the printer is limited to 300 dpi. If the image is doubled in size, and the scanning software is order to prepare the image for printing at 300 dpi, the result will use the same number of print dots as if the image were printed at its original size, but with 600 dpi resolution.

The restriction of this trick is to print the image at N times its original size at a resolution of X dpi, the scanner must be able to resolve NX dpi.

9.7.2   Service Bureaus

Most large cities and many universities have "service bureaus" or "university printshop" operations where very high resolution scanners are available. Industrial-strength scanners have a resolution of at least 2000 dpi. If very high resolution is needed, these specialized shops will provide it (for a fee).

For photographs, the simplest way to high resolution is to use a purpose-built photograph/slide scanner.

9.8   Scanning Transparencies

Transparencies cannot be scanned in the same way as images in a book or on a sheet of white paper. The reason is that the green scanning light *reflects* off a sheet of white paper, but it goes *through* a transparency.

The good way to scan transparencies is to use a scanner which has a transparency adapter. Such adapters are widely available, but add several hundred dollars (or more) to the cost of the scanner. So, many scanners don't have them.

9.8.1   White-Paper-Trick

The sneaky way to scan transparencies on an adapter-less scanner is to insert a sheet of white paper above the transparency. The scanning beam will then pass through the transparency and bounce off the white paper back to the imaging head — but only where not blocked off by the pigment on the transparency. This trick works well for transparencies which are colored only in black. We have no experience in copying colored transparencies this way.
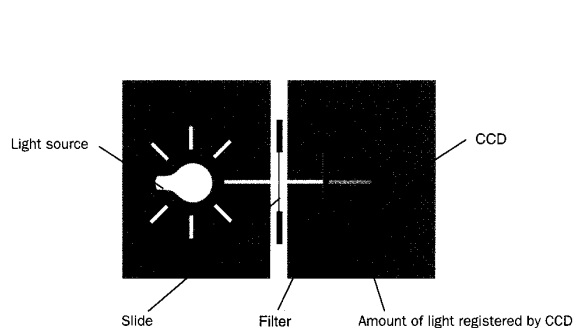
FIG 42 The basic principle of a slide scanner. The light passes through the slide and then a filter. After that, the remaining light is registered by CCD photodiodes.
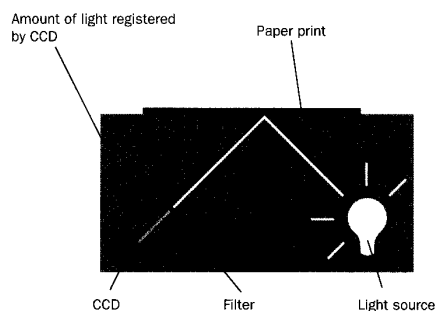
FIG 43 The basic principle of a scanner for paper prints. Light is reflected from the print and passes through a filter. Then the remaining light is registered by the CCD photodiodes.

Figure 9.1: A comparison of scanning for opaque and for transparent media.

9.8.2 Newton Rings

When a transparency is scanned on a normal flatbed planner, the interaction of partial reflections between the glass platen of the scanner and the transparency film can create patterns of concentric rings on the scanned image which are called "Newton rings". These are not inevitable, and indeed can often be eliminated simply by repositioning the transparency film and rescanning. (The rings tend to occupy only a small portion of the image.)

Transparency scanners therefore *eliminate* the glass sheet so as to remove any possibility of Newton rings.

9.9 Scanning Photographs

A photograph which has been printed or developed onto paper is an opaque object, and can therefore be scanned like any other sheet of paper. A photographic slide is transparent, and therefore should be scanned on a flatbed printer using the white-paper-backing trick like other transparent media.

The best option with slides is to use a purpose-built slide scanner.

One great advantage of purpose-built slide scanners is that they have *very high resolution*, typically at least 1800 dpi. They also come with software specially designed to handle photographs, photographic negatives and so on.

"They offer superior software algorithms for capturing the subtle differences between different transparency types and color negatives. Because film companies such as Kodak and Fuji manufacture their films differently for consumer and professional use, this can be a significant feature. Your sunsets look great when shooting Kodachrome because "amateur" films like Kodachrome and Fujichrome are oversaturated to give 'truer' reds and greens. Ironically, Kodak's and Fuji's professional E6 films are nearly indistinguishable — they don't want to unduly influence a professional photographer's results. Dedicated slide scanners have to know how to distinguish between these two film types. As a result, most transparency scanners come with a fairly complete list of different film types that the scanner software supports and thus can compensate for when scanning." (Sullivan, 1995, pg. 37).

9.10   Scanning Material Objects

Sullivan (1995) notes that a scanner will often give good images of objects such as a dime or a bottle of white-out. It is helpful to put a sheet of white paper or Bristol board on top of the object to give a better image. It is usually necessary to use Photoshop to clean up the background. Still, this might be a good way to put images of small pieces of computer equipment or experimental apparatus into an image.

However, digital cameras as noted earlier are usually better unless the object is *small*. A closeup so that the object fills most of the picture may then be beyond the minimum focal distance of the camera.

9.11   Tone Curves, Gamma and All That

Definition 19 (Tone Curve (of a Scanner)) *This is a curve that gives the scanned brightness (output) as a function of the intensity of the image being scanned (input).*

By convention, the horizontal axis of the plot of a tone curve is *input* and the vertical axis is for output intensity.

Ideally, the default tone curve of a scanner is linear, but this is never exactly true. However, to understand the adjustments that can be made by software, it is helpful to pretend the default tone curve is linear.

Scanning software has the ability to increase or decrease contrast so as to reveal details in deeply-shadowed areas, for example, that would be lost if scanned using the scanner's default settings. However, there are some subleties.

First, a few scanners and many image-processing programs allow for the *linear* controls of "contrast" and "lightness". Sullivan detests linear corrections to the tone curve because they usually lead to loss of data.

Definition 20 (Contrast Adjustment (of a Scanner)) *The angle of the tone curve is increased or decreased.*

When contrast is increased, a range of dark on the image is output as black and a range of light tones is also output as white (right middle panel in Fig. 9.2). Information is lost because a whole range of intensity is mapped to a single tone.

When contrast is decreased, the output range is truncated. A low-contrast image does not use the blackest blacks or the lightest whites; we have the approximate equation

$$\text{low-contrast} \quad \leftrightarrow \quad \text{all-grey, black-free \& white-free output}$$

Definition 21 (Lighten/Darken (of an Image)) *The tone curve is moved up or down without a change in shape or slope. (An artist would refer to this as raising or lowering the "key" of the image.)*

Lightening adds a constant amount to each tone in the output. A range of dark tones is thus truncated from the image:

$$\text{lightened} \quad \leftrightarrow \quad \text{grey \& white, black-free output}$$

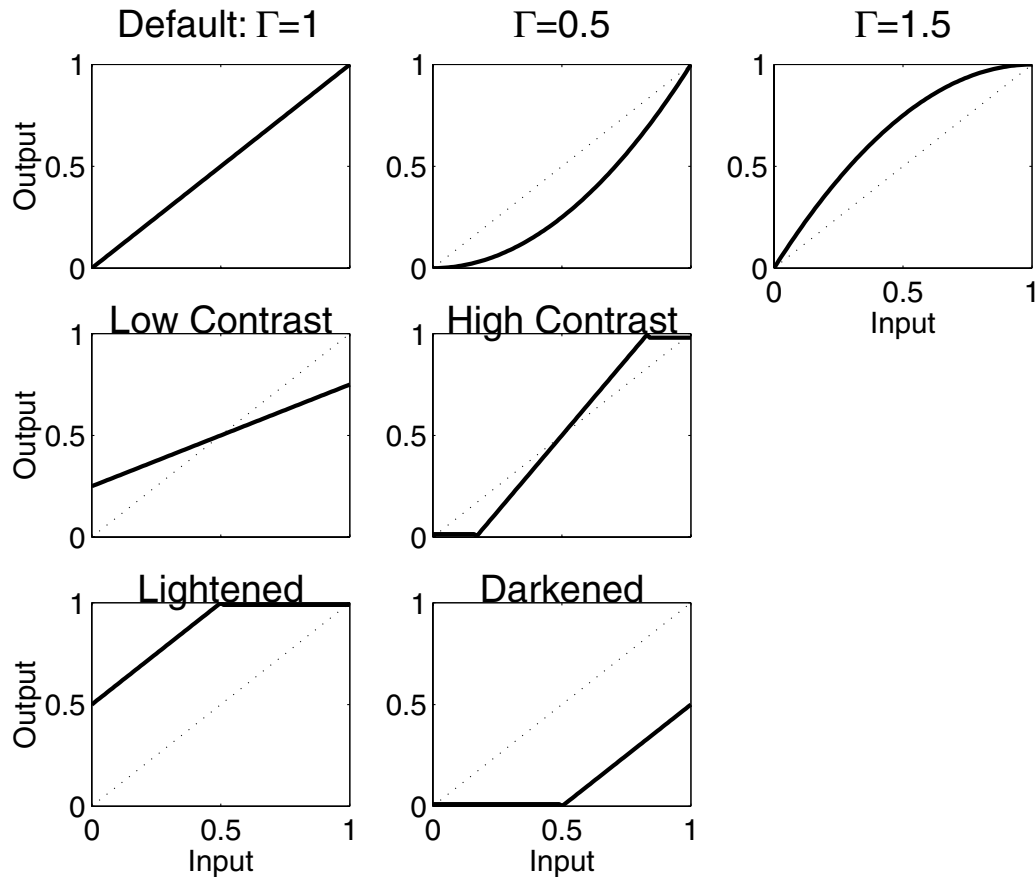$$\text{darkened} \quad \leftrightarrow \quad \text{black \& grey , white-free output}$$

Figure 9.2: Tone curves for an image, with and without the use of various filters.

Contrast uses the full output range of the scanner, but there is loss of data because part of the input range, either the black end or the white end, is ignored. Lightening and darkening use the full input range, but part of the output range is ignored.

In contrast, scanner software allows easy manipulations of the tone curve which force it to be *nonlinear*, but preserve the totality of both the input and output range.

The simplest nonlinear tone curve manipulation is to vary the parameter called "gamma".

Definition 22 (Gamma) *"Gamma" is a numerical parameter which specifies the brightness of midtones relative to default. Thus, Gamma=1.0 is the default. Gamma = 0.75 improves resolution in light areas; Gamma = 1.5 improves resolution in dark areas. The ends of the tone curve (blackest-input-to-blackest-output, and similarly at the white end) are not altered by manipulating Gamma.*

Sullivan is a big fan of Gamma manipulation because the full input and output ranges are preserved. When Gamma is set equal to 0.75, the midtone input is mapped to a darker tone on output so that the tone curve becomes a parabola dragged down below the default line curve. Thus, the brighter half range of tones in the image is mapped to say, two-thirds of the range of tones in the output. Thus, contrast is increased for the lighter tones.

9.12 Summary

1. Scanners come in an enormous range of prices, but good inexpensive flatbed scanners are now widely available in stores and in university computer labs.

2. Transparencies are best scanned in a scanner with a transparency adapter which eliminates the glass sheet on which the paper rests when scanning paper. However, an ordinary flatbed-scanner-with-glass can usually be employed successfully, too, with the aid of one or more of the following tricks:
   (i) Insert blank white paper on top of the transparency.
   (ii) Reposition the transparency if "Newton rings" appear.

3. Photographs on paper can be scanned like other images on paper; slides can be scanned as transparencies, but the best option is to use a purpose-built slide scanner, which has very high resolution and software that can correct for the idiosyncrasies of various kinds of film.

4. Scanning can be performed without a scanner by using the following alternatives:
   (i). Video camera (with a computer frame-grabber board).
   (ii). Photo CD
   (iii) Extraction from PDF files
   (iv) Digital camera

5. An excellent reference on scanning is the following short monograph: *Make Your Scanner a Great Design & Production Tool* by Michael J. Sullivan, Northlight Press, Cincinnati, 160 pp. (1995).