A Primer on Image Histograms and Curves

Contents

Introduction -- *"How Do I Know When to Stop?"*

If you're new to using a scanner, or perhaps even not so new, the process of getting an acceptable image as well as what constitutes a final image can at first be a total mystery -- it sure was for me. A paucity of useful books on this subject only aggravates one's attempts to address one's ignorance. No matter how good appearing the scans one gets, unless one develops an understanding of the underlying concepts and principles of operation, one is left with the nagging suspicion that the final scan could have been better.

Another reason for such a measure is that if you have a high-end scanner and want that last 2% of performance -- which differentiates the scans of an excellent scanner from an average one -- you need to be able to quantify just how good each image is. Otherwise how else does one justify the premium paid for such a scanner?

"I Can Tell My Scanner Is Good Because the Results Look Great."

You see, that was the nub of the problem. Once one acquires a minimal competence in adjusting images one is able to adjust images scanned under a wide range of scanner settings and get acceptable results on the monitor. I would make several scans of the same image, each with different settings, and after tone and color adjustment end up with images of apparently equivalent quality. Which one to save? Obviously you want to save the best one, but they all look the same. What is needed is an objective measure of scan quality that eliminates variables such as monitor performance and the subjective judgments of tone and color adjustment.

Because I tended to use my scanner intermittently, I accumulated a pile of notes over the years. In an effort to organize and consolidate these notes, mostly for my own convenience, I decided to put them into Web pages. Hopefully these also may be beneficial to others. Basically these pages outline, more or less, the evolution of a personal approach to scanning.

To be honest, this is not simple material, but when I reflect on how many hours I've wasted making scans, I realize that I would have saved most of those hours if I had spent just a few of them learning this material beforehand.

The Context for This Tutorial

Often a properly scanned image will look much better than an image scanned carelessly at twice the resolution. This tutorial emphasizes a quality of data approach to film scanning over quantity (resolution). And as a practical result, it advocates a two-step, professional-type of workflow for image production:

1. In the scanning step, one scans for maximal data quality, i.e. capture image data as accurately possible. The resultant image constitutes the final scan.

2. In the image production step, one or more final images are derived from the final scan in an editing program and are modified through tools such as filters, cropping, sharpening, etc. What distinguishes these final images is that they are modified to address different needs by abstracting from the final scan. A final scan with high data quality as a resource has more potential for meeting more needs.

Most amateurs and beginners attempt to accomplish all this in a single step and as a result must re-scan the image as different needs arise. If the final scan contains all capturable data, re-scanning is superfluous. This is the main theme of this tutorial. Histograms and curves are merely the tools that are used as aids in extracting every bit of data and thereby implement this approach.

Goals of a Scanning Procedure -- First Principles

To me, getting a quality scan should almost be a mechanical process. I don't mean that one works like a mindless robot; only that at each step one is able to read and interpret the diagnostic tools available and from them know exactly what to do next. Part of this requires a good scanner with software with the necessary features. The other important component is, of course, a competent operator who knows what to do.

The abiding idea the operator follows is that in scanning the name of the game is accurate and comprehensive data capture while maintaining the relationships among the pixels. If this sounds a bit dry and mechanical, I would argue that's how it should be. Having good original scans as a resource gives you more data to work with when you exercise your creative impulses later in the image-editing program. That's where you alter and make abstractions of your original scans. A good final scan may not look snappy and may not grab your attention, but it will stand-up to your manipulations in the image-processing program so that the final product does look snappy and grabs attention.

Scanner Sensitivity -- The Limiting Resource

The typical consumer film scanner reads each pixel as an analog signal from an array of charge-coupled-device (CCD) sensors, which sweep longitudinally across the film plane. An analog to digital converter translates it to an integer from 0 to 255 (in 8-bit mode) or 0 to 4095 (in 12-bit mode). This is done for each of the red, green, and blue color channels. These readings are passed to the scanner software, e.g. NikonScan, which uses a user-specified function to transform the tone and color to an output number, also from 0 to 255, which is stored in the resultant image file. Unfortunately the CCDs in non-professional scanners have a very limited sensitivity range and are often incapable of capturing all of the tonal values in an image. In addition, because they operate at the limits of their capabilities, they are susceptible to electrical noise artifacts. To minimize these artifacts and make optimal use of the limited range, the user is forced to define how the image is to be scanned by assigning weights to areas of the image.

The 256 tones are just sufficient to form a visually continuous gradient. The problem is that once the scan is made one is working with discrete (integer) values. With each adjustment of contrast or brightness, some of those values are aggregated. After several adjustments, an area containing a smooth tonal gradation may degenerate into visibly distinct areas.

LS-2000 Process Schematic

Next Page

Contents

Image Histograms

How does one judge a scanned picture? More basically, what are the criteria by which one judges the quality of a scan? Conceptually simple factors such as focus and frame alignment are immediately obvious from viewing the monitor and are easily adjustable through simple controls or manual adjustment. Other factors such as color and tonal balance are much more difficult to evaluate by looking at the scanned image on the monitor. A full-frame scan on a scanner with a default resolution of 1350 dots per inch will generate approximately 6 million bytes. At a resolution of 2700 dpi, the file will contain over 24 million bytes. Obviously no one is intellectually capable of following such a mass of points individually. At the other extreme, the person who can consistently obtain quality scans by just previewing images on the monitor is exceptionally talented indeed.

An image histogram is a way of making quantitative sense of this mass of data.

A brief course on histograms

One of the characteristics of an image that we want to analyze is the lightness or darkness of its pixels, or luminosity. An image of tone steps is a good introduction to understanding image histograms because similar tones are grouped together, making the correspondence between pixel count of each tone (the area of each tone) and bar height obvious:

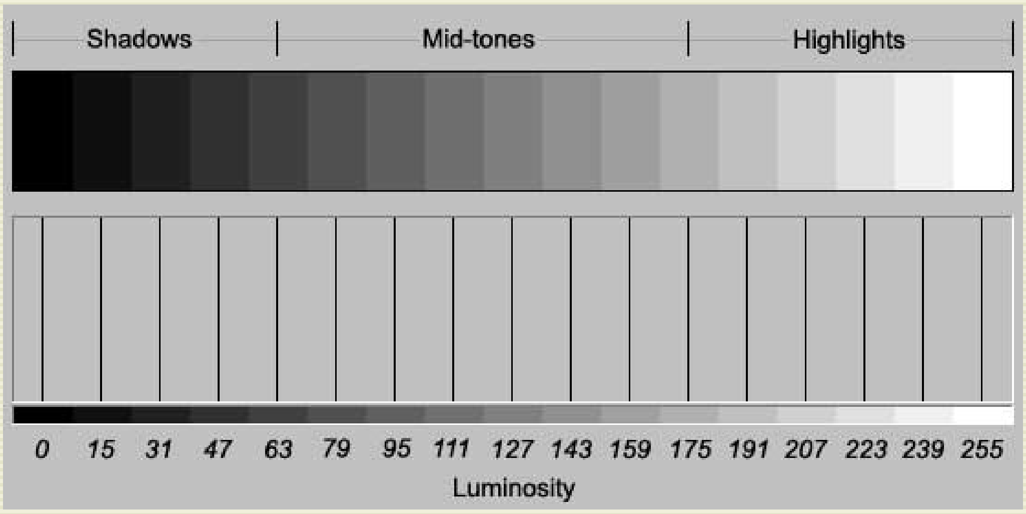
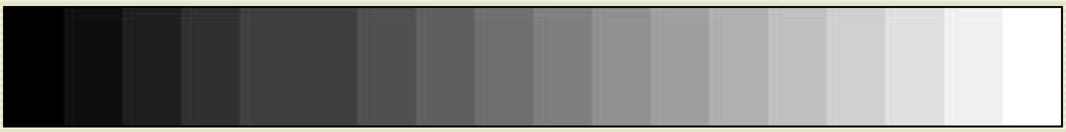


fig. 1a

Luminosity values range from 0 (black) to 255 (white) and are whole positive numbers; there can be no 22.5, -10.0, 156.19, etc. In fig. 1a each of the 17 steps contains 2048 pixels, so the height of each bar underneath represents a count of 2048. This histogram has a fixed-height vertical axis, which is set equal to the tone with the greatest frequency (the statistical mode). The heights of the other bars are then set relative to the tone with the greatest frequency. Since the steps in this example are equal in size -- have the same number of pixels -- the bars are equal in height.

In this image, the fifth step is doubled in size to demonstrate that histogram bar heights are relative to the sample with the greatest frequency and are not absolute:



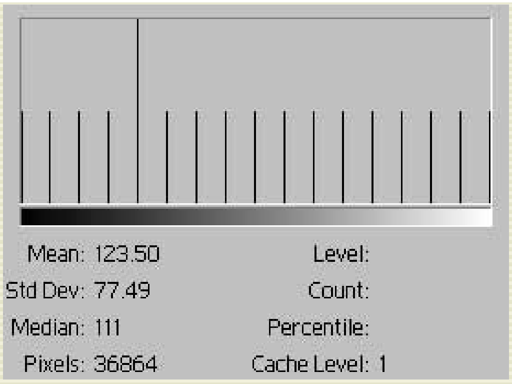
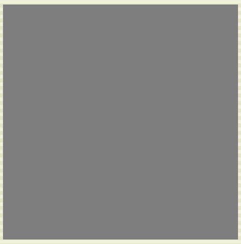


fig. 1b

Here's something much simpler:

Single tone, no contrast



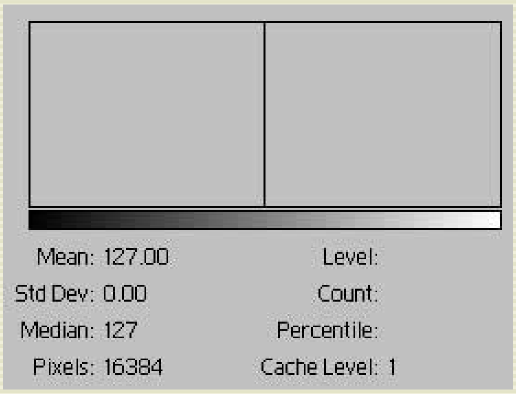
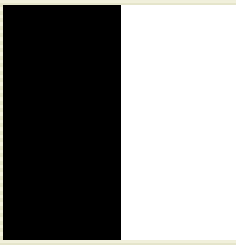


fig. 2

This is as simple as it gets, a one color-one bar histogram. The RGB levels of the color are (127, 127, 127). The bar has a height of 16,384 -- all the pixels in the image. As a result the mean and median are 127. The bar's height could also be read as 100%, a percent scale, since that's the only color. With every pixel having the same tone, the standard deviation is zero.

Two tones, high contrast



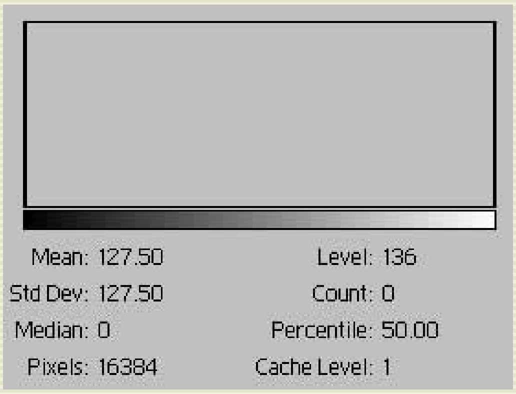
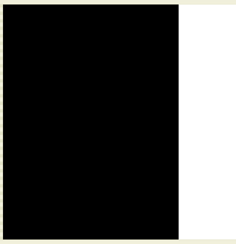


fig. 3a

This image is 50% black and 50% white. The bar at the extreme left represents the frequency of black and has a height of 8192; the bar at the extreme right represents the frequency of white and has a height of 8192. The average tone is the same as in fig. 1a, but a gray tone of this value is not present in the image. The extreme difference in tones is indicated in the high standard deviation -- a pixel is either black or white. The standard deviation is useful as a numerical proxy for the contrast of the entire image. The median is incorrect and should be 127.5.

Two tones, high contrast



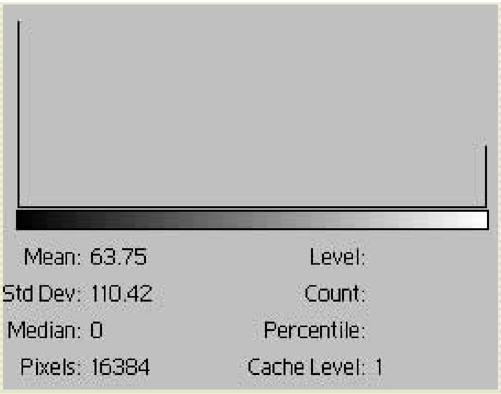
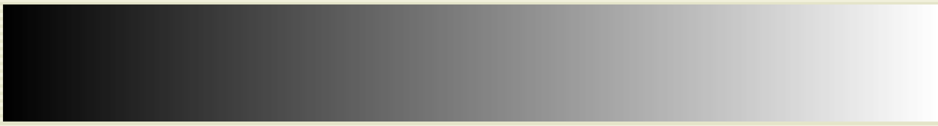


fig. 3b

In fig. 3b we've changed the proportions of black to white pixels to 3 to 1. That is 12,288 pixels are black and 4096 are white. As a result, the black bar is 3 times higher than the white bar. The histogram frame has been removed to make this clearer.



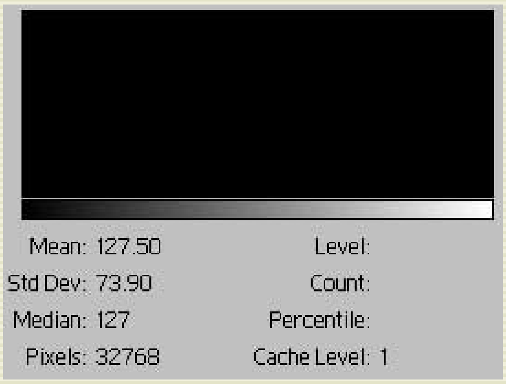


fig. 4

Fig. 4 is a scale of 256 gray tones that appears continuous. The histogram is solid with bars because the continuous gradation has no missing tones.

An image histogram consists of 1 to 256 tonal values, or bars

The height of each bar represents the count of pixels having that tonal value (0 to 255)

Previous Page Next Page

Contents

Color Channels and Luminosity

The great majority of Photoshop histograms used in this tutorial display *luminosity*, or brightness. In the RGB color system, the predominant color system of non-professional scanning, this channel exists as a composite value, or in the current idiom, a "virtual" channel calculated from the levels of red, green, and blue.

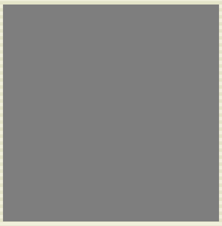
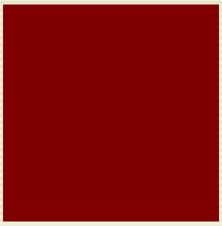
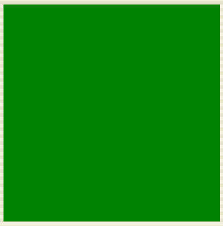
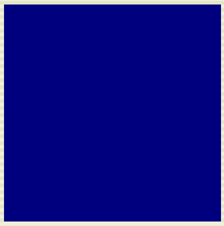
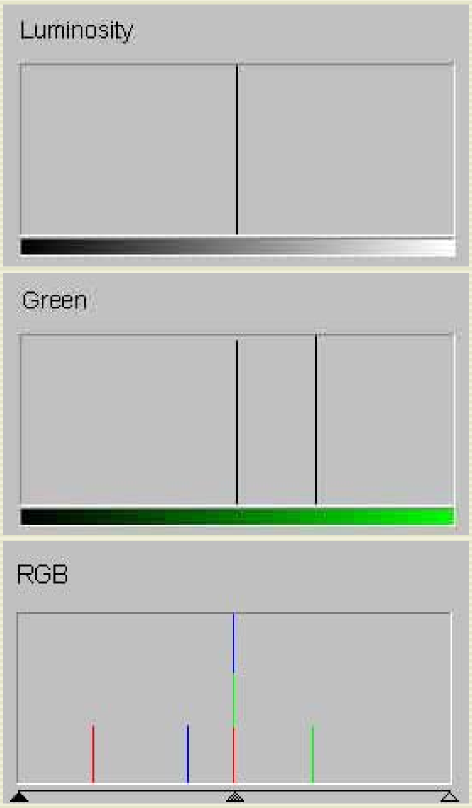
	Gray is a composite of			
<127, 127, 127> Luminosity: 127	is the sum of	red <127, 0, 0> + Luminosity: 38 +	green <0, 127, 0> + Luminosity: 75 +	blue <0, 0, 127> Luminosity: 14

fig. 5.

Note the weights for luminosity assigned to the three colors: 0.30 : 0.59 : 0.11. Green is weighted 5.4 times that of blue! Human vision is much more sensitive to green and the weightings reflect the relative sensitivities. This unequal effect on luminosity should be kept in mind when one adjusts the separate color channels.

A Bar in the Luminosity Channel May Count Pixels of Different Colors

fig. 6



The left half of fig. 6 is the same gray as in fig. 5. The color in the right half, also a composite of RGB, was contrived such that it also has a luminosity of 127. As a result the luminosity histogram for the image results in all the image's pixels being observed in the same bar. The histogram for the green channel, however, shows two bars. The center bar is associated with the gray in fig. 6 and the bar to the right of it is the green component of the right half of fig. 6. The RGB histogram is a sum of the three underlying channels: red, green, and blue. The three components of the green half appear as the smaller bars in the rightmost histogram for RGB. The center bar reflects the 3 components of the gray half of fig. 6.

Why Luminosity?

To the human eye, brightness rather than hue is the quality that is more important in perceiving images. This is merely confirmation of what those who produce images, whether in the darkroom or in the computer, instinctively do in practice: correct brightness before color balance. Ideally one would use the luminosity histogram to apply and assess changes in brightness or contrast. The problem is that luminosity is not available in most scanning programs and is available with some inconvenience in Photoshop. (This is no longer the case since the release of version 8.)

The RGB histogram is a cumulative enumeration of its underlying red, green, and blue components, and unlike the luminosity histogram, does not summarize a visually apparent quality. That is one can perceive redness, greenness, blueness, brightness (luminosity), but not "rgbness". You'll use the RGB histogram in performing the scan because it clearly identifies the "outliers," or values that book-end the tonal scale of the image. Once the image has been acquired, however, we're more interested in analyzing and modifying what's between those ends -- how tones are distributed throughout the image. The luminosity histogram is the tool to do this. In practice, as long as the image is large enough and the black-white point settings have been applied, I've found that the RGB histogram often, but not always, approximates the luminosity histogram in indicating the location of peaks and end-points.

Previous Page Next Page

Contents

Image Histograms and Pixel Arrangement

Keep in mind that a histogram simply gives a summary of the image's tonal values. It tells us nothing about how the individual pixels are arranged.

To demonstrate this we'll contrive an image consisting of a visually continuous tonal gradient. A gradient has tones that are spatially related because the image elements are quantitatively ordered. That is similar tones are spatially contiguous.

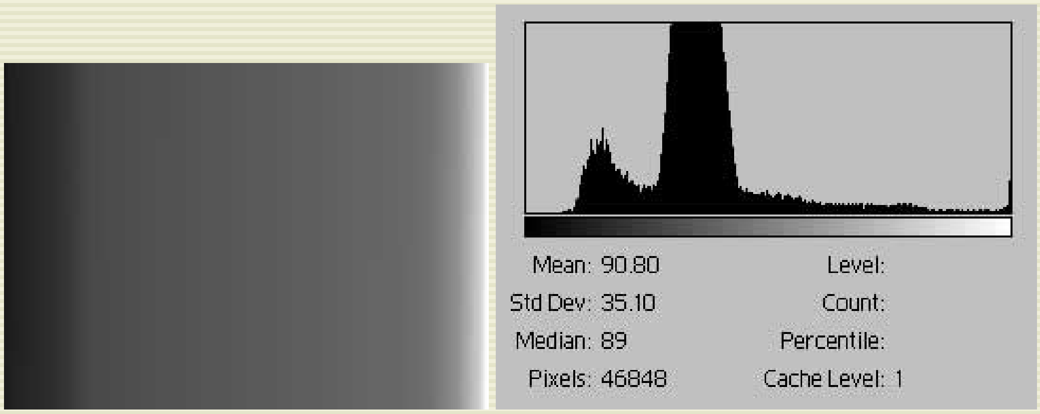


fig. 6

To the right of fig. 6 is its histogram. In this image the tonal value of a pixel is absolutely correlated with its location. Given a location we can predict what the pixel will be.

Next, with some of the tools, we're going to push the pixels around randomly, making sure as we move each pixel that its tone is unchanged:

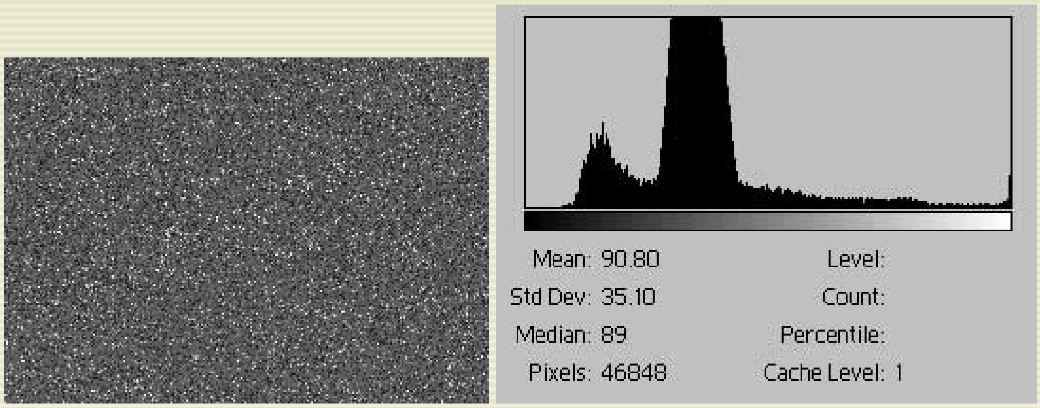


fig. 7

As you can see the histogram is exactly the same as the gradient histogram, but in fig. 7 the tonal value of a pixel is uncorrelated with its location.

Now, let's push the pixels around a bit more creatively:

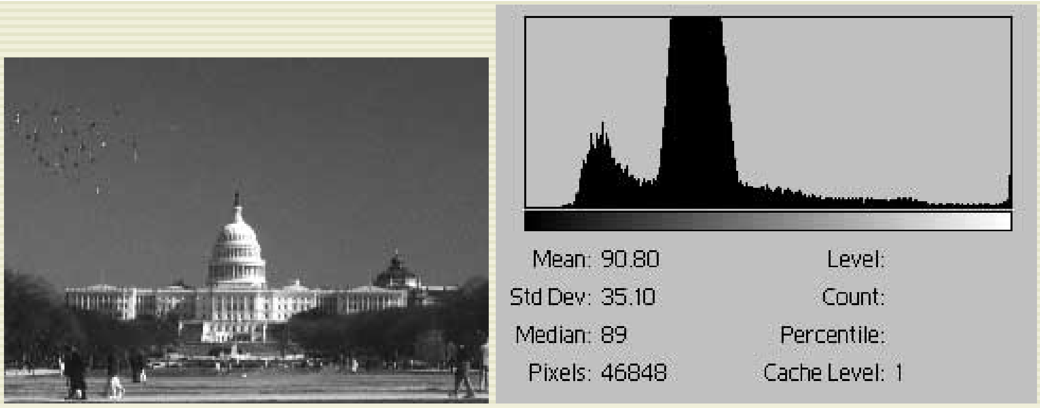


fig. 8

Again the histogram is exactly the same. Here the tonal value of a pixel has some correlation with its location. Given a location we can predict with good probability what its value will be.

The three pictures have exactly the same pixels and as a result their histograms are exactly similar. How the pixels were arranged does not affect the histograms. As far as the histogram function is concerned, it's looking at the same picture in each of the three cases. Our visual sensibilities attach significance and order to an image.

A bar in an image histogram aggregates observations from the entire image

Previous Page Next Page

Contents

Local Histograms

Another important diagnostic tool is what I call the local histogram, as contrasted to the global histogram, which summarizes the entire image. A local histogram samples an area of particular interest lying within the image, a spatial subset.

Returning to our tonal gradient, we'll construct a rectangular selection (dotted line), which defines the area to be measured by the histogram function:

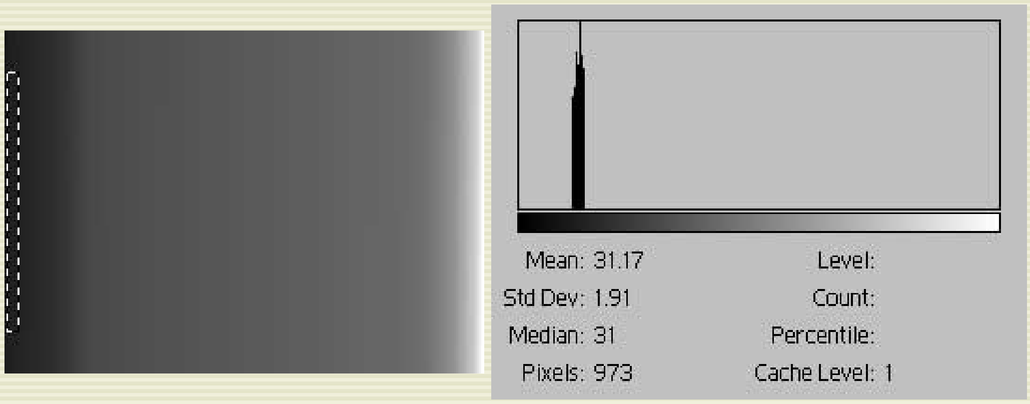


fig. 9

The histogram shows that the values are within an extremely narrow range of values. In contrast to the global histogram, this local histogram tells us we are examining a nearly solid tone. Even without seeing the selection, one could guess the shape, location, and size of the selection.

Returning to our second image (fig. 7), where we've taken the first image and randomly moved the pixels, we'll make another rectangular selection:

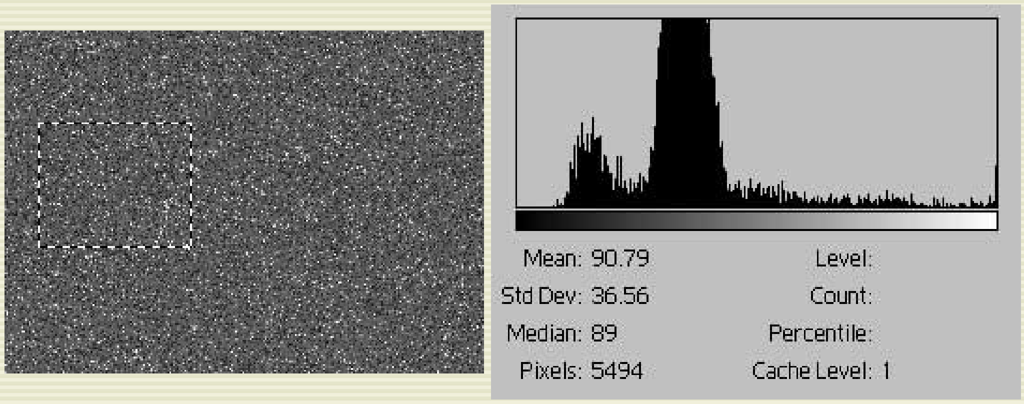


fig. 10

This local histogram is almost exactly the same as the global histogram. This makes sense because the pixels are arranged randomly, so a large enough sample is representative of the distribution of pixels of the entire image. In this image there is no privileged view. A privileged view is a local view that offers information about where it is in the global view. For this image, any view should have the same histogram as the global as well as any other local view.

For the third picture (fig. 8), instead of a rectangular selection, we'll draw the selection around the US Capitol Building, the subject of the picture.

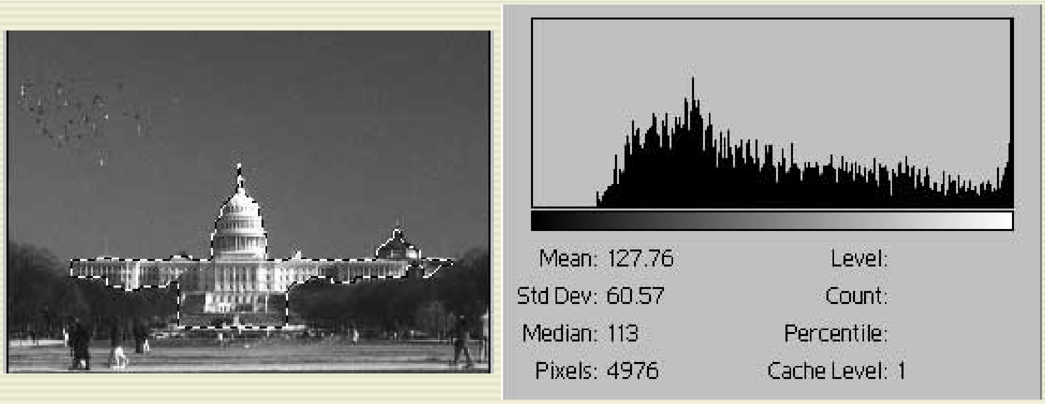
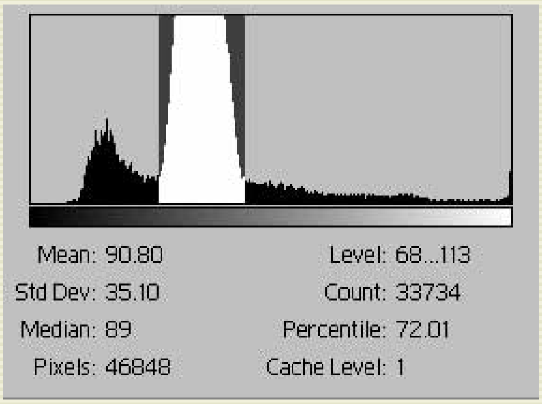


fig. 11

This histogram shows a fairly well distributed range of tonal values from the mid-tones to the highlight tones. Some of the pixels have gone to complete white. This is not unexpected since the capitol is made of white limestone. The global histogram is darker, the mean value is 91, and has a much lower standard deviation of tone. A change in contrast will have a greater effect on the subject area than the rest of the image.

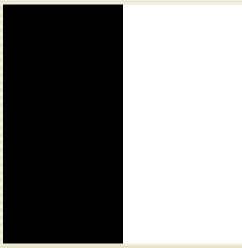


Much of the area of the image (about 72%) is of the sky. Only a few hundred or so of the subject's pixels are in this hump (white area).

There is a feedback mechanism between local histograms and the global histogram. In general, the local histogram will be associated with the area of the image containing the subject. When it comes to adjusting the image, as you make changes to one, global or local, you'll want to check the effect changes have on the other.

Image Detail Size Affects Apparent Contrast

It's obvious that fig. 12 is maximally contrasty (the standard deviation is 127.5). Locally, on the pixel level, however, it would have zero contrast. Fig. 13 appears to be a monochromatic gray with no contrast at all. Yet it is composed of alternating black and white pixels, exactly the same pixel elements as fig. 12 but arranged differently. Therefore, according to the histogram, it also has maximum contrast. Note that fig. 13 has only bars at the extremes. Locally, on the pixel level and in contrast to fig. 12, it would have maximum contrast. So 12 and 13 differ in their local contrast. Of course, this is a most extreme case, but pointillistic effects often lead to subjective judgments counter to the histogram's purely objective display. Why is this detail important? Curves might have radically different effects depending on local contrast. If fig. 13 were instead mid-tone gray pixels, it would behave differently in response to a tone curve.



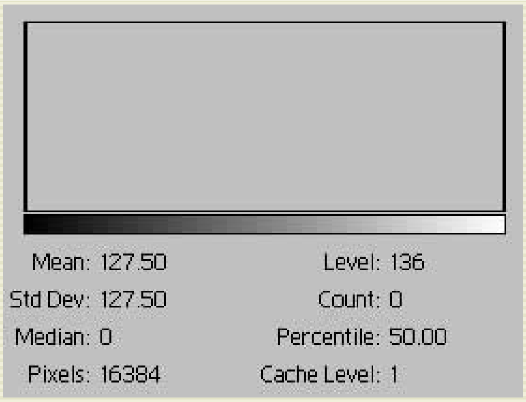
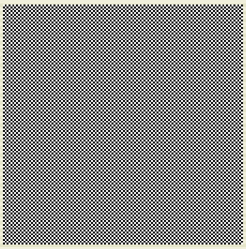


fig. 12



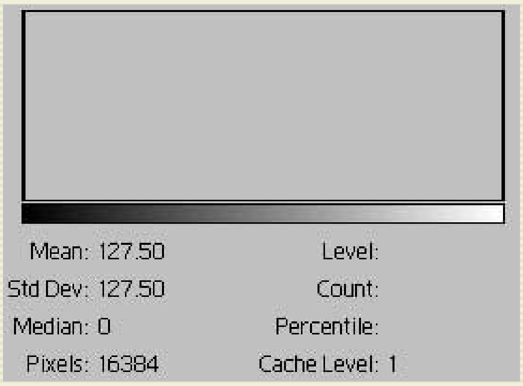


fig. 13

(The half-tone gray in fig. 13, can be converted into a real gray by using the contrast slide control. Decreasing contrast will shift the black and white pixels towards each other in tone.)

A local histogram used with the image histogram enables the scanner operator to vary the allocation of the available tonal range between the subject and the image as a whole

Previous Page Next Page

Contents

Using Histograms as a Scanner Tool

It's a Matter of Correct Exposure

If you're serious about photography you take great care in setting your camera's exposure. A scanner is also a camera, and a possible weak link in your production chain unless similar care is taken in setting its exposure. As an exposure meter is used to monitor a camera's exposure, the histogram is the primary tool used to monitor a scanner's exposure. That's why histograms are important.

Histograms are used in scanning to show how the values of the pixels that have been captured are distributed over the scanner's tonal sensitivity range, just as a camera's meter indicates how well the image matches the film's tonal range. In other words, they tell you how efficiently the scanner is being used. Without them it would be extremely difficult to achieve technically optimal scans.

- Histograms provide a graphical display of the state of the image, summarizing the vast quantity of data contained in a typical scan
- It is nearly impossible to make the final adjustments through visual feedback alone. Monitors are incapable of displaying the subtle differences in the shadow areas which are easily discerned in a histogram

These principles also apply to digital cameras which display histograms.

These two images illustrate the limitations of the monitor as a scan diagnostic tool. The picture on the right has more shadow detail than the one on the left. To view these differences, however, one must increase monitor brightness and contrast to the maximum -- hardly a convenient way of comparing image quality. Why should one worry about detail not visible on the monitor? For a Web application, for example, you may not need such detail, but in another application you may need that detail if some manipulation of the tones or colors make these tonal areas more visible. Or some other output medium that you may want to use, presently or in the future, may be capable of displaying these distinctions.

How the Scanning Process Relates to Image Histograms

The typical scanner has a density range less than that attainable on film. Any tones denser than that perceivable by the scanner become black, also known as clipping.

The width of the scanner histogram's horizontal axis--from the left (black) to right margin (white)--represents the scanner's density range. Each of the bars, of course, represents the pixel count for the tone value of its position.

This particular histogram shows that approximately two-thirds of the scanned image's pixels--the dark gray area--are in the darker half of possible tonal values.

The Photoshop version of the histogram for the image is functionally similar:

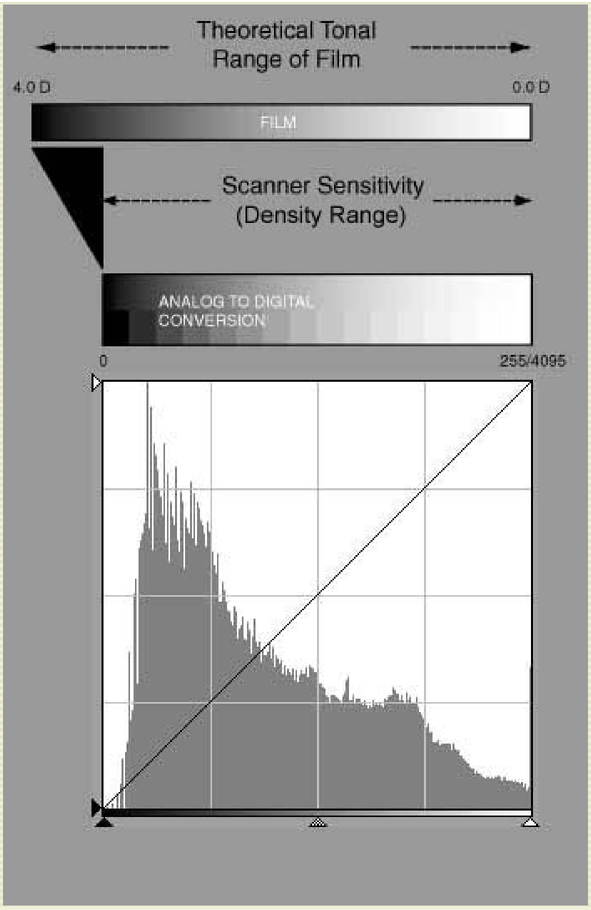
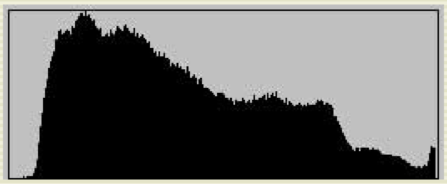


fig. 1



The scanner histogram, at left, is for the preview image, which is based on a small sample of pixels. The Photoshop histogram is for the scanned image.

[Comparison of Photoshop and NikonScan Histograms]

Image Density Range and Scanner Density Range

One of the primary reasons for performing a preview scan is to obtain an image histogram. Yet I would guess that because of simple ignorance the vast majority of non-professional scanner users simply ignore this extremely useful device. I would also guess the same applies to users of graphics editing programs such as Photoshop. This section gives an idea of how to begin to use histograms in scanning.

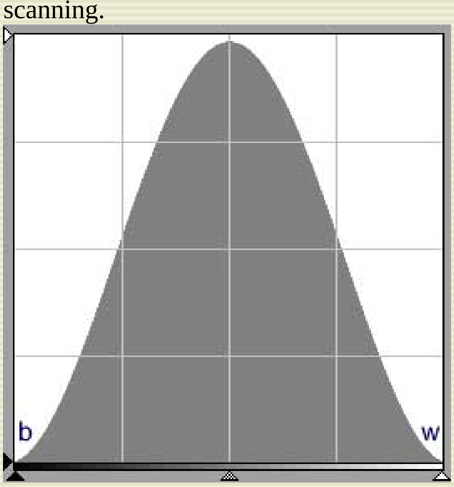


fig. 2

If you get fig. 2 as an initial preview histogram, this is about as good as it gets. The image's tonal range neatly matches the scanner's density range without having to apply curves adjustments. The left of the histogram (b) abuts the base of the scanner's density range and the right of the histogram (w) abuts the high-end of the range. The scanner's sensitivity is used at maximum efficiency: none of the image is outside the scanner's sensitivity range and samples will be obtained for all 256 tonal values.

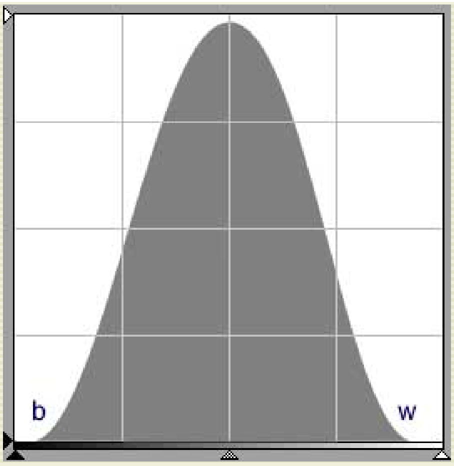


fig. 3

Fig. 3 is a histogram where both b and w are in the left-most and right-most quarters of the histogram but not touching the edges. The scanner's sensitivity is used efficiently.
Examples: Average density range, average contrast

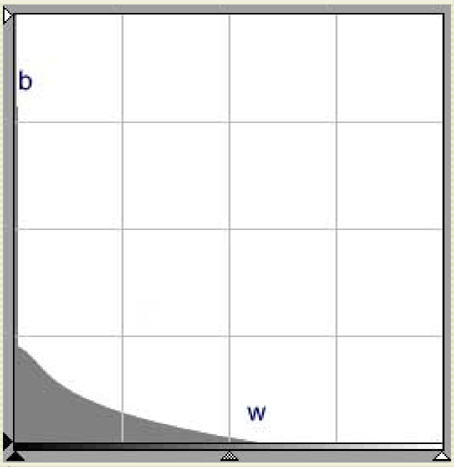


fig. 4a

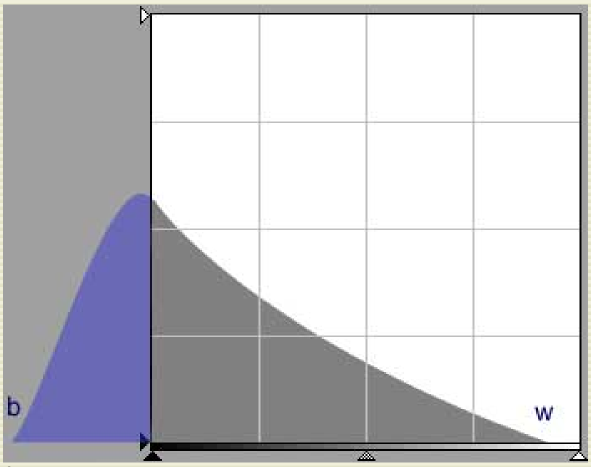


fig. 4b

Fig. 4a is a histogram, most commonly encountered in scanning transparencies, where some of the film's tones are beyond the scanner's sensitivity. Fig. 4b, a notional histogram implied by fig. 4a, depicts the part of histogram (blue) falling outside the scanner's density range. The tones in the blue area are imaged as black.
Example: Wide density range, average contrast

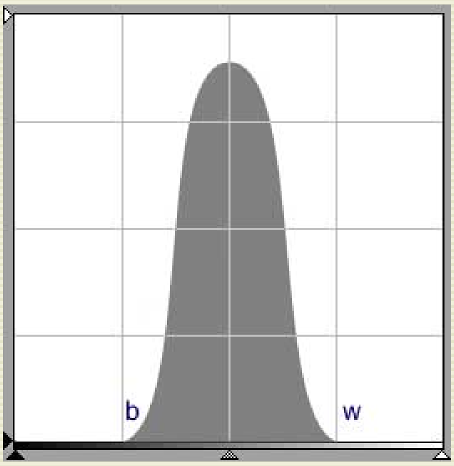


fig. 5

Fig. 5 is a histogram with endpoints close or near the quarter tone markers. This is an uncommon occurrence. Much of the scanner's density range is unused.
Example: Narrow density range, very low contrast

Evaluating the Preview Histogram

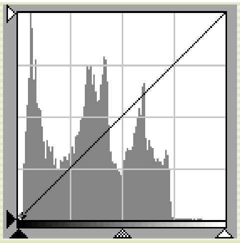


fig. 6

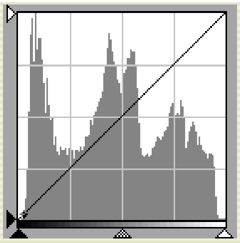


fig. 7

Figs. 6 and 7 are histograms for scans of the same image without curve adjustment and illustrate making use of the available density range. Even if the scan represented in fig. 6 had resulted in an image visually closer to the desired final image, the image represented by fig. 7 is preferred because it spans the range of available tone values and as a result contains more information than 6. With more information, the user has greater latitude for subsequent manipulations of the image. One could produce the image in 6 from 7 but not the reverse. Note that although figs. 6 and 7 are for the same image, fig. 7 occupies a greater area.

Previous Page Next Page

Contents

Postscan Histograms

Histograms displayed by scanner programs are formed by sampling the image preview. A histogram of the final image, obtainable in an image processing program such as Photoshop, gives a more accurate representation of the image's characteristics. It is useful when it is necessary to examine the histogram in detail.

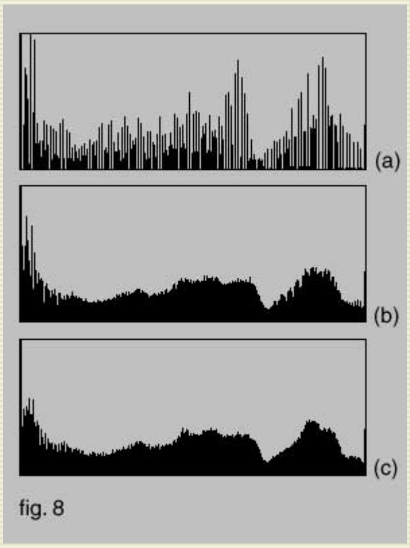


fig. 8

Note the vertical bar at the left margin, implying a high number of tones that could not be imaged. In addition, the relatively smooth contours of distributions 8b and 8c degenerate near the left tails. This is because this area is near the limits of the scanner's density range, where pixel values are more likely to be distorted by noise and thus exhibit greater randomness. This does not necessarily mean that the scan will be unacceptable. It does demonstrate the effectiveness of the histogram in indicating potential problems.

Fig. 8 shows a progression in image quality for 3 scans of the same image. Fig. 8a is for an image which has been tonally adjusted by using curves. Note the development of drop-outs. Some tones have combined with adjacent tones to form tall spikes and leave gaps indicating their absence. Fig. 8b is for the same image, but tonal adjustment has been achieved by using analog gain rather than curves. Fig. 8c is the same as 8b but in addition was scanned in 12-bit mode. As a result it exhibits the smoothest contour of the three scans, although in practice 8b and 8c would be visually identical in appearance. Both 8b and 8c efficiently use the tonal space since all values are used.

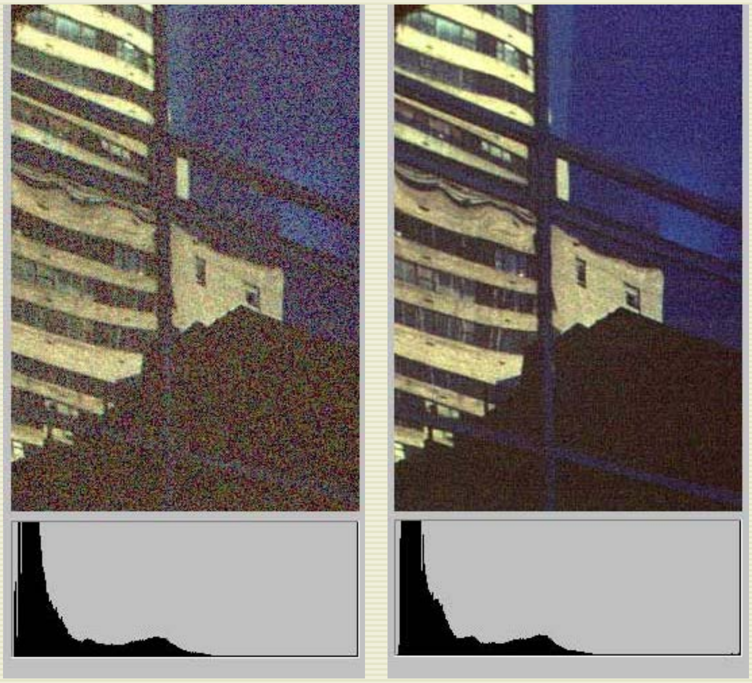


fig. 9a

Using Multi-Sampling

The effectiveness of multi-sampling to suppress noise relies on the premise that errors in reading pixel values are non-systematic, i.e. truly random. Thus they can be eliminated by scanning the image several times and taking an average value for each pixel. Implementing this in software requires making several copies of the image as a source for the final image. Some scanners do this in hardware by taking up to 16 readings at each position as the CCD sensor array is stepped over the image. This has the advantage of maintaining registration and performing the task in one pass.

Fig. 9 shows the effect of using of multi-sampling in reducing noise near the limits of the scanner's sensitivity. The scan for the histogram for 9a was performed without multi-sampling and shows an erratic contour at its left edge, showing the effects of noise. Fig. 9b, performed with 16x sampling, has a clearly defined contour at its left edge. (Details shown above histograms.)

fig. 10

Fig. 10a, the histogram for the picture above, has three salient features. The dark foreground appears as the largest peak on the left. The buildings are represented as the mid-tone peak. An essentially featureless sky in the source transparency was forced into white (the peak along the right margin) so that the building features and foreground tones could occupy more of the tonal space. It is important to emphasize that the foreground is not black and contains detail.

Fig. 10b is a hypothetical histogram for the same image scanned on a scanner with a more limited density range. Because many of the picture's tones are beyond the scanner's ability to image them they become black (black line against left margin).

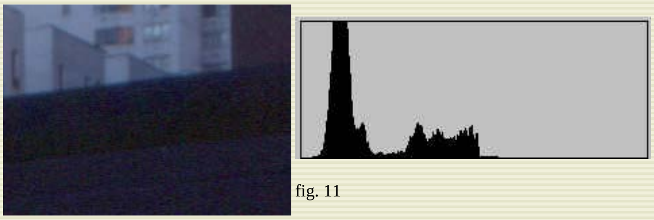


fig. 11

In reality, the typical behavior for scanners with a limited density range is shown in the detail in fig. 11. Non-imageable areas become noise, which register as non-black tones instead. Figs. 10 and 11 emphasize the importance of the area near the black end-point in a scanner histogram.

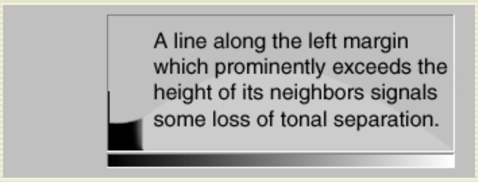
A Histogram for Negative Film is the Reverse of a Histogram for Transparency Film

The examples I have used so far have been from *transparencies*. For *negative* film, of course, the reverse applies--it is the area near the right end, the highlight end-point, that is more important. This is where the pixel values for the densest areas on film will appear. Fortunately for amateur photographers, negative film tends to be more scanner friendly than transparency film because the densest tones are not as dense as the densest tones on transparency film, thus avoiding the limitations of the scanner's density range. Kodachrome, a transparency film, with its nearly opaque dark areas, for example, can be particularly challenging for most scanners. To say it is the *bête noire* of scanners seems particularly apt. For the sake of simplicity, this tutorial assumes the use of transparency film.

Look for Bars of Maximum Height along Either Margin

A bar of maximum height at either margin (b, w) may signal a loss of tonal separation.

The cause may be a tone curve that has thrown values into the extremes; tones beyond the capability of the scanner to image; or truly characteristic of an unusual image.



Indicators of a loss of tonal separation -- merged bars, such as fig. 8a, bars along the margins, discontinuities, or rough contours in larger images -- imply a loss of information in the image.

In preparing an image for production, some loss of information is inevitable, even desirable if the image's main elements are to be emphasized. The importance is in knowing how to control information loss as a means of enhancing the production image.

Previous Page Next Page

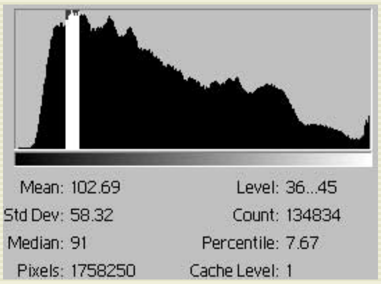
Contents

Interpreting Area under the Curve

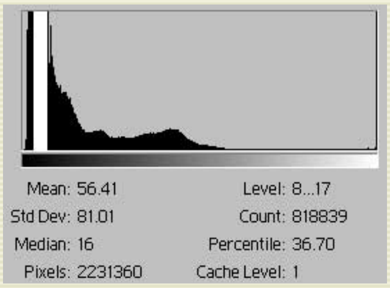
An important visual cue is the area under the curve. A histogram occupying a small area within its frame implies that a greater proportion of pixels have been imaged within a narrow band of tones. As a result, tonal separation may be diminished. This may be entirely consistent with the image, but if in addition it displays a bar of maximum height adjacent to the left or right margin, a problem is likely to have occurred.

A typical continuous tone image will exhibit a relatively large area such as in the top histogram in fig. 12. In the bottom histogram in fig. 12, the relatively small area indicates a concentration in tones at or near the most frequent value. If two images are scanned at the same resolutions and have approximately the same crop area sizes, why are there differences in the absolute areas? After all, the histograms must account for an equal number of pixels and thus it seems they should have the same areas. As I mentioned earlier, the histogram is drawn within a frame of constant dimensions: there is a fixed vertical distance available to represent bar heights. The tonal value with the greatest frequency, the statistical mode, is set to this height no matter how many values are represented by this bar. The histogram's other bars are scaled relative to the mode. Thus a histogram with a small area indicates that the concentration at or near the mode has a much greater weight in representing the image's pixels. Another way to look at it would be to say that the vertical axes are scaled differently. If one were to label the two graphs, the bottom graph would have a greater number at its vertical maximum than the top graph. When the scaling difference is taken into account, the area under the curves would be the same.

This fortuitous combination of scaling about the mode and a fixed-frame histogram "normalizes" the histograms. In other words, this enables us to compare the tonal separations of two different images, independent of image size, by visually comparing the areas occupied by their histograms.



An average image will occupy a significant portion of the frame as in the histogram at left.



A concentration of tones, and thus a potential lack of tonal separation, is indicated by a histogram which occupies a relatively small area, left (white plus black). If the concentration, white for emphasis, occurs close to the left extreme, a problem with density range may be indicated. In this example, an area around the most frequent value (8 to 17) contains 37% of the image's pixels compared to 8% in the histogram above it.

fig. 12

The evaluation of histogram areas is more useful in comparing successive scans of the same image.

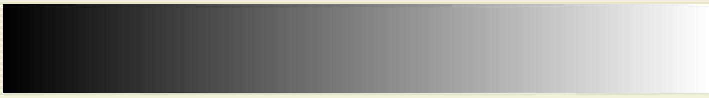


fig. 13

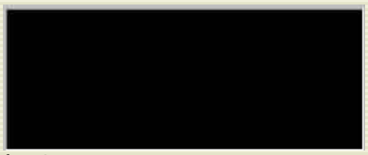


fig. 13a

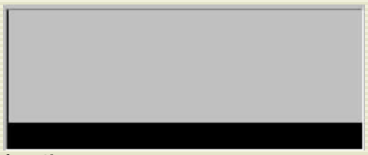


fig. 13b

Figs. 13a and 13b are histograms for scans of the same hypothetical image, fig. 13. Fig. 13a shows that the scan of fig. 13 has resulted in an image with an exactly even dispersion of tones. The histogram occupies the entire frame. Fig. 13b, done with notionally incorrect settings or an inferior scanner, has a much smaller area because many of the scanned pixels have ended up in the bar adjacent to the left margin. Again, note that 13a and 13b account for the same number of pixels but have different areas because the vertical axes are scaled differently.

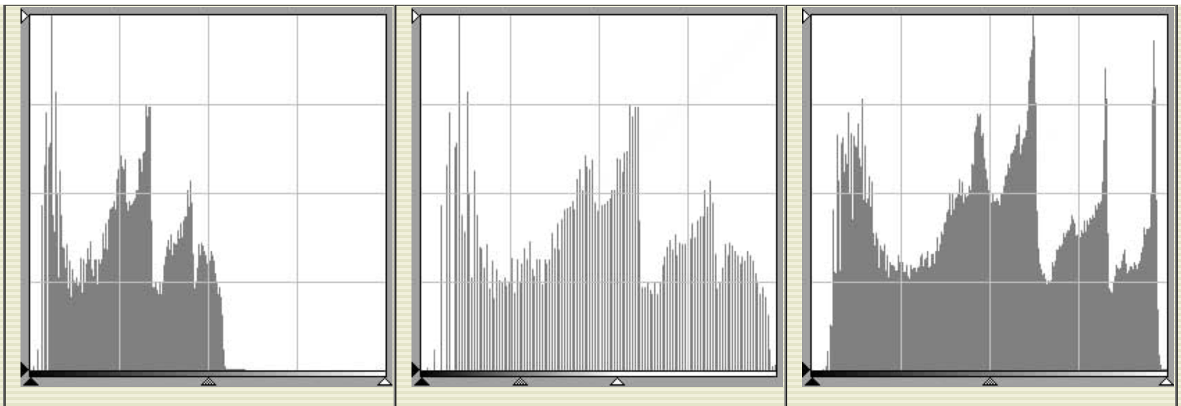


fig. 14a fig. 14b fig. 14c

Figs. 14 demonstrate the effects of two alternative methods of attaining a full range (black to white) of tones in an image. As previewed by the scanner, fig. 14a indicates an image with a narrow tonal range. By applying curves, fig. 14b is made from 14a by spreading the graph apart, much like an accordion, until it spans the full width of the graph. (Why this is done and how curves work is covered in the next section of this tutorial.) Gaps are left to fill the values vacated by the shifted bars. Note that since histogram 14b is 14a with its bars shifted -- gaps between its tones -- the area under the curve is the same. In contrast, fig. 14c achieves a full range of tones by adjusting the exposure through analog gain. It has a far greater area, about 80%, than fig. 14b and will have smoother tonal transitions. This is not to say that the image for 14b will be "bad," it depends on how much quality one needs to extract from the image.

The difference in areas in figs. 13 result from a concentration of tones. The difference in areas in figs. 14 result from gaps in data.

Everything else being equal for the same image, the scan with a histogram with greater area will contain more information and thus tend to exhibit superior tonal separation.

There's No Substitute for Practical Experience

I want to emphasize that there's no substitute for scanning on your own and analyzing your own scans and histograms. Start with images with various film types, lighting contrasts, and take notes. It's essential to be able to mentally associate the characteristics of an image with the resultant histogram, something that comes easily only with practice and experience.

What to Look for in The Histogram of A Well-Scanned Image of A Continuous Tone Picture

To summarize this section, this is what you look for in the histogram of a well-scanned *continuous* tone image:

- Absence of large gaps between values
  Usually caused by the application of curves. Minimized by additional bit-depth.
- Image makes use of available scanner density range
  Adjusted by analog gain.
- Smooth transitions from value to value
  Improved by bit-depth and multi-sampling (primarily in the shadow areas).

The final scan is performed when the scanner user cannot improve further on these criteria. The final scan serves as the basis of the production image(s), and therefore only coincidentally appears as ultimately intended. It is intended to obviate the need for rescans because it contains all of the readable data from the image.

Previous Page

Next Page Tone Curves *or* Next Page Setting Exposure -- An Interactive Demonstration

Contents

The Scanning Process, Exposure, and Tone Curves

You modify the brightness and/or contrast of an image by specifying a tone curve.

An Overview of Specifying Scanner Exposure

Before we examine curves, however, I would be remiss in pursuing our objective of getting quality scans if I didn't put tone curves into the context of setting overall scanner exposure.

The myriad controls that the Nikon LS, for example, has to set exposure -- tone curves, analog gain, brightness and contrast sliders, eye droppers, prescan, 'cont', etc. -- can at first be confusing and intimidating. How, when, and why do you use them? You can make sense of these controls if you conceptualize the scanning process as occurring in what I call the *pre-scan*, *scan*, and *post-scan* stages, **and** if you consider that a scanner operates in a world of continuous (the analog domain) and discrete numbers (the digital domain).

Pre-Scan

The operator sets the operating conditions (luminance of the LEDs, film frame rotation, film flatness, even cleaning the film, etc.) so that the scanner maximizes data capture from the image. Insofar as this tutorial is concerned, however, this means matching the tonal range of the image to the scanner's tonal sensitivity range. You want to illuminate the image such that the range from the darkest to the lightest meaningful values that you want to capture match the extremes of the scanner's sensitivity for dark to light tones -- setting the black and white points. Of course you can do this in post-scan using curves, but How a Tone Curve Works shows the cost of using a curve: the histogram is expanded causing gaps; data are lost. To do this on the Nikon LS the 'prescan' is performed to get the scanner's guess of the black and white setting. Then the user may use the analog gain controls to move the base of the histogram curve to meet the white or black point. If it is possible for the operator to set the black and white points in this stage, the scanner will be able to obtain samples from all 256 intervals and all will be meaningful. Operations in this stage are manifestly analog operations. (If you're technically minded, this stage forms a continuous function that the scanner samples. You define the domain (the meaningful tones) of this function such that it maps precisely into the range of the scanner's optical sensitivity. This will result in a sample for every possible tonal value.) Scanner Exposure -- Getting It Right

Scan

The scanner takes readings from the film and converts (quantizes) them to digital values. User specified settings that have effect are bits/color channel (8, 12, or 14) and scanning resolution.

At this point the image's tones have been digitized into a range of discrete numbers.

Post-Scan

The scanning program applies the tone and color curves that the user has specified to the set of numbers that the scanner read from the film. Applying adjustments at this stage has the advantage of greater flexibility, but since discrete numbers are being manipulated, the cost may be distortions in the data. Areas with increased contrast will develop gaps and possible posterization. Areas with decreased contrast will develop loss of tonal separation. Besides tone curves, eye droppers, sliders, and 'auto' function in this stage. These are all digital operations and may be applied outside of the scanning session without affecting the quality of the scan.

If a more extensive tone control tool repertoire is required, an alternative is to specify the tone curves in an image editor such as Photoshop, save them, and load them back into the scanning program. If this approach is anticipated, it may be more convenient to access the scanner through the image editor's Twain interface.

It is after the post-scan stage that further adjustments are made (e.g. analog gain, curves), and another scan is performed.

To be more specific, scanner exposure controls are set in the following order:

1. Run autoexposure to get tentative black and white point setting (analog adjustment)
2. Use analog gain (if necessary) to get a more precise black and white point setting (analog adjustment)
3. Make the digital setting of the black and white points
4. Specify tone curves (digital adjustment)

These 4 steps set the black and white points, set the color balance, and allocate tonal range between the subject and background.

By putting adjustments up-front, in the pre-scan, you don't have to apply them in the back-end, in the post-scan, where distortions may occur. Books have been written about analog to digital conversation processes, so to keep it simple, I'll summarize this section with one rule:

Perform as many exposure adjustments as possible in the pre-scan and as few as possible in the post-scan

Setting the Black and White Points

What You're Trying to Do with A Scanner Tone Curve -- The Objectives

Now let's return to tone curves. The scanner displays the histogram for the input image, and by specifying a tone curve you modify those tones so that they take on the values you desire. To determine the effect of the tone curve on a tonal value project its value vertically from the input axis until it intersects the tone curve. At the point of intersection move left horizontally until it intercepts the vertical axis. The new values are reflected in the histogram for the output image and give the user feedback.
How a Tone Curve Works

If a histogram bar is shifted to the right, raising it in value, the pixels in the bar are lightened. If it is shifted to the left, lowering it in value, the pixels in the bar are darkened.

In general you want to specify a tone curve that

- Minimizes breaks in tonal continuity Discussion
- Emphasizes the subject elements by allocating them more of the available tonal range. This comes at the expense of the tonal range of less important image elements. Case study

If you've performed the scanning procedure properly, all global settings to tone and color will have been performed during the scan. Think of the final scanned image as a resource for current and future applications, not just the immediate present. If all of these global adjustments are incorporated in the final scan, you won't have to do rescans and subsequent uses of the image will merely require small local tonal changes.

Properties of Tone Curves

Tone curves (Tone Curve Characteristics) have several important characteristics and restrictions:

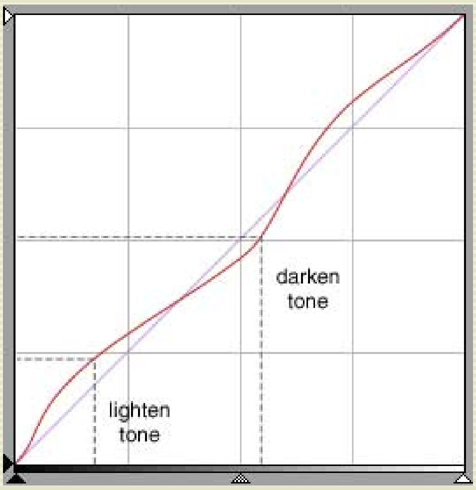
- The horizontal and vertical axes are both numbered 0 to 255, taking whole values only
- The end-points are always the lower-left (0, 0) and the upper-right (255, 255) corners of the chart. This means the overall angle, or prevailing slope, of the curve is 45º
- The horizontal axis represents the tones of the source (input) image.
- The vertical axis represents the output image
- The tone curve is continuous -- a tone curve with gaps or breaks cannot be specified

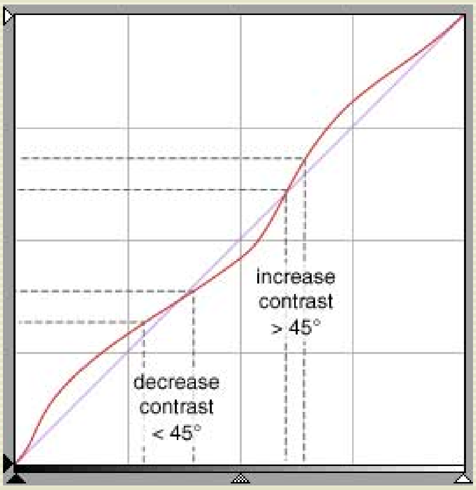
The curve, defined by the user, shifts the bars (i.e. counts of pixels) in the histogram to their desired positions. If in the process of shifting the bars two or more bars have the same value, they are then aggregated into one bar. For example, if the curve causes several bars to fall into the range 32.5 inclusive to 33.5 exclusive, then all of those bars are aggregated into the value 33. Obviously this represents a loss of information to be avoided.

Remember how emphatic I was about maximizing the use of the tonal space in the previous section? The user is rewarded for achieving this with greater flexibility in applying curves because the tones are more precisely defined.

Previous Page Next Page

Modifying Brightness and Contrast





The properties of the tone curve (red, above) imply

- *A tone is darkened (lowered in value) if the input tone intersects the curve below the 45º diagonal*
- *A tone is lightened (increased in value) if the input tone intersects the curve above the 45º diagonal*
- *The contrast between two tones is increased if the interval between the two tones is increased*
- *The contrast between two tones is decreased if the interval between the two tones is decreased*

Here is what we do

- *To lighten an image, grab the tone curve at the center and drag it slightly upwards above the 45º diagonal. Illustration*
- *To darken an image, grab the tone curve at the center and drag it slightly downwards below the 45º diagonal. Illustration*
- *To increase contrast, grab the tonal curve at a darker tone interval and drag it slightly down. Then grab the tonal curve at a lighter tone interval and drag it slightly up. This forms a characteristic "S" curve. You form the "S" between the two points between which the contrast is to be increased. Illustration*
- *To decrease contrast, grab the tonal curve at a darker tone interval and drag it slightly up. Then grab the tonal curve at a lighter tone interval and drag it slightly down. This forms a reverse "S" curve. Illustration*

The end-points are absolute, 0 and 255, so the contrast between black and white cannot be increased. (In this tutorial, I imply two types of contrast: local and global. Local for contrast between two tones. For an entire image's contrast, since the end-points are fixed, I look at the histogram's standard deviation as an indicator.) We can, however, increase the contrast between two points if one is not lying on an extreme. Increasing the slope of the curve* between the 2 points to more than 45º has the effect of spreading an input tones over a wider interval of the output line.

When I say slight, I mean slight. To increase contrast in one range of tones means to decrease contrast in another. If you go too far, tones in the increased contrast area will start to lose their continuity and posterize (fewer tones are distributed into a larger number space), and tones in the decreased area will begin to merge (more tones are squeezed into a few values). We're restricted to working within a very small interval of 256 numbers--a zero-sum game. We can perform adjustments on an image, not miracles. Even the most expert scan cannot compensate for a lack of tonal range in the source image.

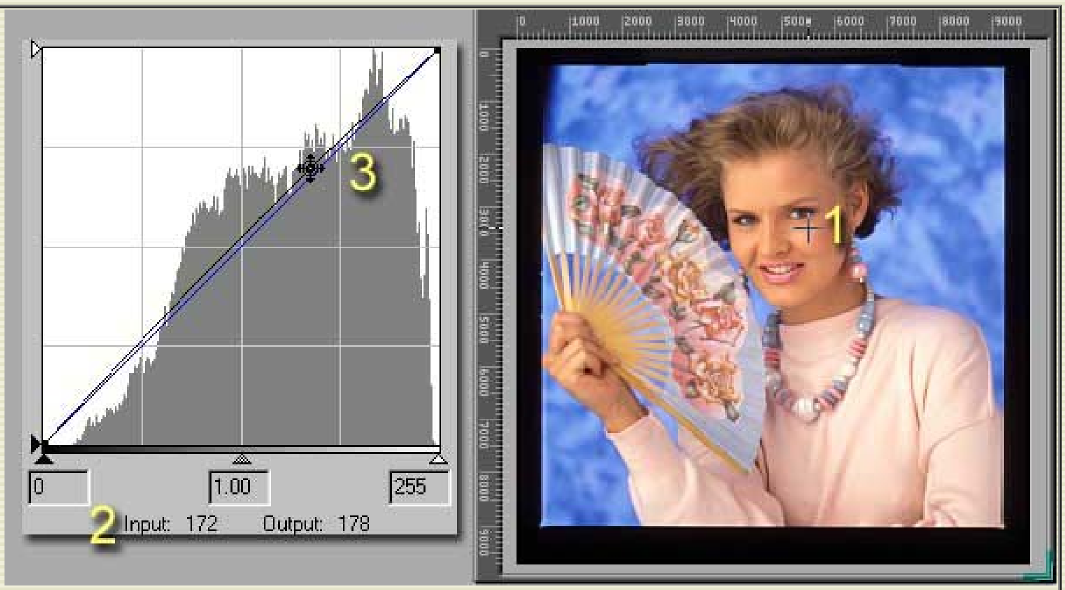
Keep in mind that each pixel in a color image is a composite of red, green, and blue values. When an RGB curve is applied to an image, the scanner software or image editing program applies the curve to the underlying components separately.

The effect of tone curve slope

A note on setting the B/W points
Comments on tone curves
Some examples of curves
An interactive demonstration of curves (requires Java enabled browser)

Identifying Tonal Areas as Grab Points

Of course, when we modify a curve we don't grab just any point, we manipulate the curve with specific tonal areas in mind.



1. Place the cursor (the cross-hairs) on a point within the area of interest.
2. The value of the pixel is displayed in the input field (172) and a knot at that value appears on the curve.
3. If, for example, you want to lighten that area, you would grab the curve at that point of the knot and drag it upwards until it's at the level of the desired output value.
   Setting gamma

Changes in contrast requires that two points be identified. An increase in contrast is achieved by moving the two points so that the slope, or angle, of a line drawn through them is increased. A decrease in contrast is achieved by moving the two points so that the slope, or angle, of a line drawn through them is decreased.

With practice, knowing where to grab the tone curve becomes almost a natural reflex.

Ameliorating the Loss of Tonal Separation

There are three methods to address the distortions caused by the application of tone curves:

1. Scan at double the resolution, apply curves, and down sample to the lower resolution This method relies on the down sampling process to average out errors. This is expensive in storage and CPU resources.
2. Scan at a greater bit-depth apply curves and down sample to 8-bits This is the most common method employed to minimize the effects of tone shifts due to curves. The Nikon LS-2000 has a provision to scan at 12-bits. This means that instead of a tonal range of 256 tones, there are 4096, a 16 times finer tonal resolution. The cost is a doubling of file size. Extra bit scanning should be considered in the critical scanning of color negatives: A Brief Note on Scanning Color Negative Film.
3. Do nothing Some applications simply do not require very high image quality. Tonal distortions would not be visible or of little concern to the end-user. This is particularly true of images destined for Web page display, which benefit more from an adroit application of an unsharp mask.

That's the theory and received wisdom. In practice, it's rare that methods 1 and 2, although resulting in better looking histograms, ever result in visually improved images. The one possible exception arises when scanning a transparency with important shadow detail or a transparency that is underexposed. If the source image is properly exposed and won't undergo extensive tonal modification, 8+-bit scanning is a bit of a waste. In any case, if you're determined to scan at greater bit depths be sure to use multi-sampling. It makes little sense to attempt to obtain the benefits of the *precision* of 8+-bit depth scanning without the increased *accuracy* of multi-sampling.

Tone Control Sliders

You might wonder what the role of the sliders might be in the method I have outlined. Simply speaking, none. I hate to sound preachy and dogmatic, but if there's anything I feel strongly about, it's that one should not use the sliders to adjust contrast, brightness, or color balance in creating a final scan. This is not especially because in NikonScan these adjustments are not reflected in the histogram, thus giving you no objective feedback, but because the effects of the sliders are so crude and thus work against the objectives of proper tonal correction. This may be demonstrated by analyzing the sliders as tone curves.

Brightness slider control demonstration
Contrast slider control demonstration

Naturally, this remonstrance applies equally to using the tone control sliders in imaging editing programs such as Photoshop.

Previous Page Next Page

---

* So closely are tone curves associated with the concept of slope that some refer to them as gamma curves--functions used in photography which relate exposure to emulsion density.

Contents

Color Corrections Using Curves

The proper means of performing image-wide color modifications is through curves -- not by using the sliders.

Modifying colors with curves is similar to modifying tones. When we speak of modifying tones it is in terms of lightening or darkening. For colors, on the other hand, we speak in terms of addition or subtraction. In addition, a modification of a color component -- red, green, or blue -- can also be expressed as an inverse adjustment of its color complement:

- Adding/subtracting red is equivalent to subtracting/adding cyan
- Adding/subtracting green is equivalent to subtracting/adding magenta
- Adding/subtracting blue is equivalent to subtracting/adding yellow

For example, if your image has a cyan cast you correct it by applying a curve that adds red, its complement. Or to say that an image lacks yellow is equivalent to saying that blue must be subtracted.

The common practice in finding the correct color setting is to evaluate each correction quickly because the human visual system rapidly accustoms itself to erroneous settings.

Example of application of color correction curve function

Post Color Correction

After a color change is applied, a compensating adjustment of brightness may be required if the image's tonality is to be preserved, something you may be familiar with if you've developed color prints in the darkroom. If the color correction curve was shifted upwards, which has the side effect of increasing brightness, the compensating RGB curve would be downwards, and vice versa. Image tonality, or luminosity, is especially sensitive to changes in the green channel, which has a far greater weighting than blue or red in calculating luminosity (see Color Channels and Luminosity).

Use Analog Gain to Control Exposure; Use Curves to Correct Tone and Color Balance

Analog gain is for adjusting scanner exposure, not for controlling tone or color balance. Although analog gain affects color balance, it only does so as a side effect to setting exposure, its sole function. The proper order of scanner operation is to set for optimal exposure, if necessary, through analog gain (a pre-scan operation). Optimal exposure only coincidentally results in correct balance. Then the image's tone and color balance are achieved through curves (a post-scan operation). In addition, because it operates linearly like a slider, analog gain is a generally inadequate and cumbersome tool for controlling tone and color balance. If you value your sanity, don't try to accomplish both with analog gain. Keep the tasks conceptually separate and address each with the appropriate tool.

Previous Page Next Page

Contents

Scanning Examples

Now we'll return to the preview histograms we looked at earlier and apply our knowledge of image exposure and tone curves.

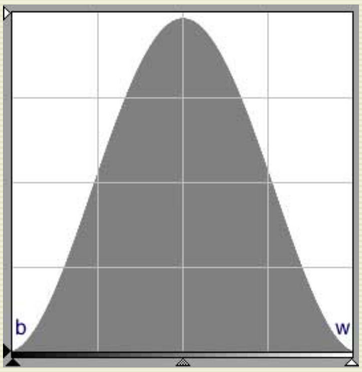


fig. 1

If you get a preview like this, just scan, no corrections are needed. The tonal range of the image (b to w) exactly matches the scanner's sensitivity (black to white triangle); the histogram's contour is smooth and continuous; there is no indication that tones have been clipped; and the mode is well away from the margins.

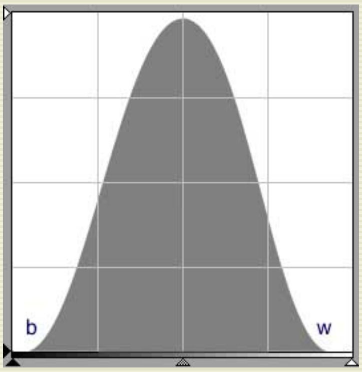


fig. 2

Fig. 2 is close enough. Just set the black and white points and scan.

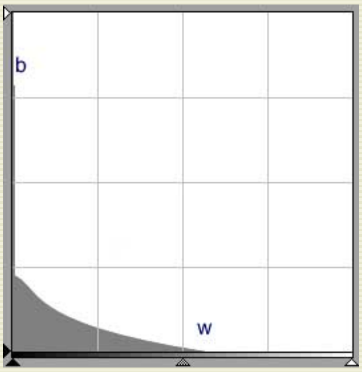


fig. 3a

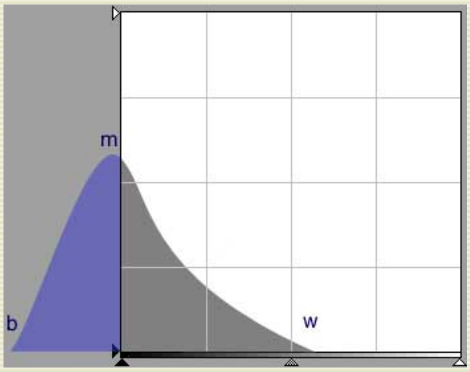


fig. 3b

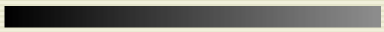


fig. 3c

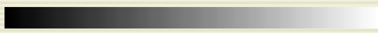


fig. 3d

For this image, some of the tones, as represented by the blue area in fig. 3b, are beyond the capabilities of the scanner. Scanned without adjustment, the blue area will appear as a long bar or cluster of bars adjacent to the vertical axis, as in fig. 3a. Fig. 3c is the tonal scale of the image prior to setting the white point (the black point is set already). Fig. 3d is the tonal scale after setting the white point through software.

Analog Gain

Adjusting the analog gain merely means changing the luminance of the scanner's light source. Usually this means an increase so that sufficient light can pass through denser tones for the sensors to measure. You can conceptualize the effect of increasing analog gain on the image histogram as follows:

Any pixel at point b is opaque black so increasing the luminance has no effect (no light passes through). Thus its value, or position on the graph, is unchanged.

The lightest tone, w, is lightened the most so its value and thus its position is shifted right the most. If it is to be a true white, analog gain is adjusted until it meets the right margin, as in fig. 3e.

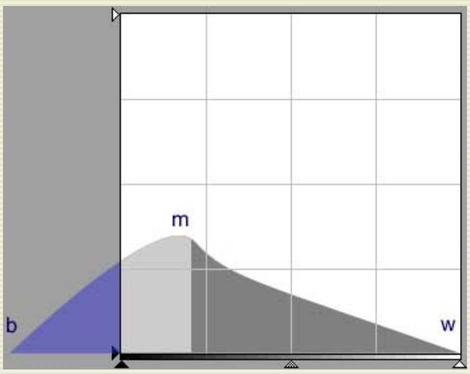


fig. 3e

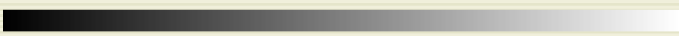


Image Processin...
John R Jensen

Radiographic Image Analysis, 3e
Kathy McQuillen Ma...

Digital Image Processing
Rafael C. Gonzalez...

Algorithms for Image Processing and ...
J. R. Parker

Privacy Information

fig. 3f

Tones in between (m) are shifted right proportionately.

To minimize this, we apply analog gain (i.e. increase the luminance of the LEDs) until the right edge of the histogram (w) touches the right margin, assuming that the tone represented by w is indeed white. This is also called setting the white point through hardware. Increasing analog gain has the effect of elongating the tonal scale (the distance from b to w) seen by the scanner. As a result more of the image's tones (light gray, fig. 3e) are shifted into the scanner's sensitivity range. Note that the tonal loss (blue area) has decreased and the tones properly imaged (gray areas) have increased. Without analog gain, a final scan with the distribution in fig. 3b will result in an image containing approximately 140 of the 256 available tones, whereas fig. 3e will contain all 256, as shown by fig. 3f. Compare tonal scale 3d with 3f. Figs. 3d and 3f are drawn with their widths proportionate to the number of tones in their images.

For an interactive simulation of the exposure adjustment process: Interactive Exposure Demo (requires Java enabled browser)

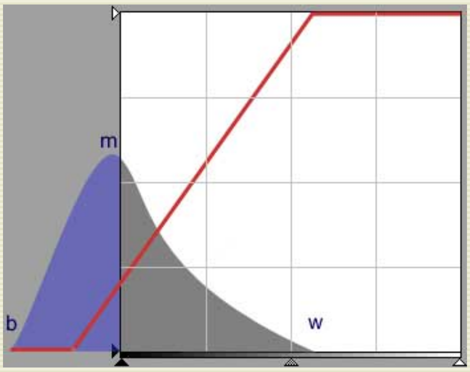


fig. 3g

Adjusting analog gain implicitly applies a linear curve to the image in the analog domain. (Thus far, the curves we've discussed function in the digital domain.) In this example, setting analog gain to transform distribution 3b to 3e, we notionally apply the curve shown in 3g (red). An analog gain curve resembles that which is often used in setting the black-white points (characterized by the horizontal end segments), precisely our intent here. In this case, however, the effect of the curve extends to transform values outside the digital domain of the scanner (the super-imposed graph frame). (The inclusion of the graph frame is meant to help illustrate that scanning is a process of moving an image defined in the analog domain into the scanner's digital domain.) Note that the curve is notional and isn't displayable by the scanning program because this is a pre-scan adjustment.

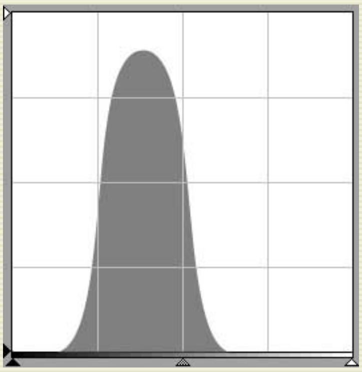


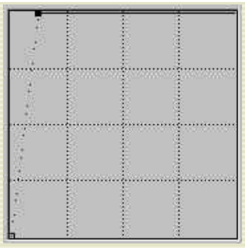
fig. 4

The problem with the histogram in fig. 4, often encountered with low-contrast films, is that it indicates a small range of tones. This is most often encountered in scanning color negative film. The approach to scanning low-contrast images is covered in A Brief Note on Scanning Color Negative Film

Temporary Curves

In addition to being used to form a final image, curves may also be used as a temporary expedient in evaluating an image. Returning to the earlier example, which was given to illustrate the limitations of using the monitor to evaluate the scan, suppose we want to make the differences visible:

In this case we can apply a curve with an 8:1 slope in the input interval 0 to 30 to both images to make the differences in the shadow areas more visible.



With a good scanner and careful control of exposure, images with full tonal ranges and shadow detail may be obtained, even from high contrast films, such as this one on Kodachrome 64:

 228 KB image

Previous Page Next Page

Contents

Summary

The purpose of this tutorial was to suggest that the process of using a scanner is greatly simplified if the user approaches it as an exercise in maximizing the quality of data capture. It defined the objectives of this method, the criteria for measuring these objectives, the methods for measurement, and the limiting factors imposed by the hardware and software. I believe this far more systematic method will result in less frustration and waste of time than the largely intuitive and haphazard approach that most scanner users employ.

Specifically, some of the key elements are that

- Most non-professional scanners have limited tonal range sensitivity and we are constrained to making the best of this
- Histograms display the tonal characteristics of an image and provide an objective measure of scan quality. Some of the attributes to be evaluated are the area, contour, and continuity
- The task of the user is to identify the range of meaningful tones of the image and match them to the scanner's capabilities. This maximizes data capture
- If image-wide changes are necessary, apply tone curves in the scanning step so they will not have to be applied in subsequent uses.

Hopefully, the information in this tutorial will also enable the consumer to evaluate scanners and scanner software more knowledgeably and thus provide vendors further motivation to improve their products. The user should ask herself or himself whether a prospective product has the features that enable the user to employ the principles outlined here.

Not only are histograms and tone curves powerful tools, but they are absolutely necessary in achieving quality scans. By comprehending and putting into practice what we have covered, I hope to make scanning a more productive, if not enjoyable, experience.

Previous Page

High Frequence Trading

Cost Cutting in Business

E Procurement System