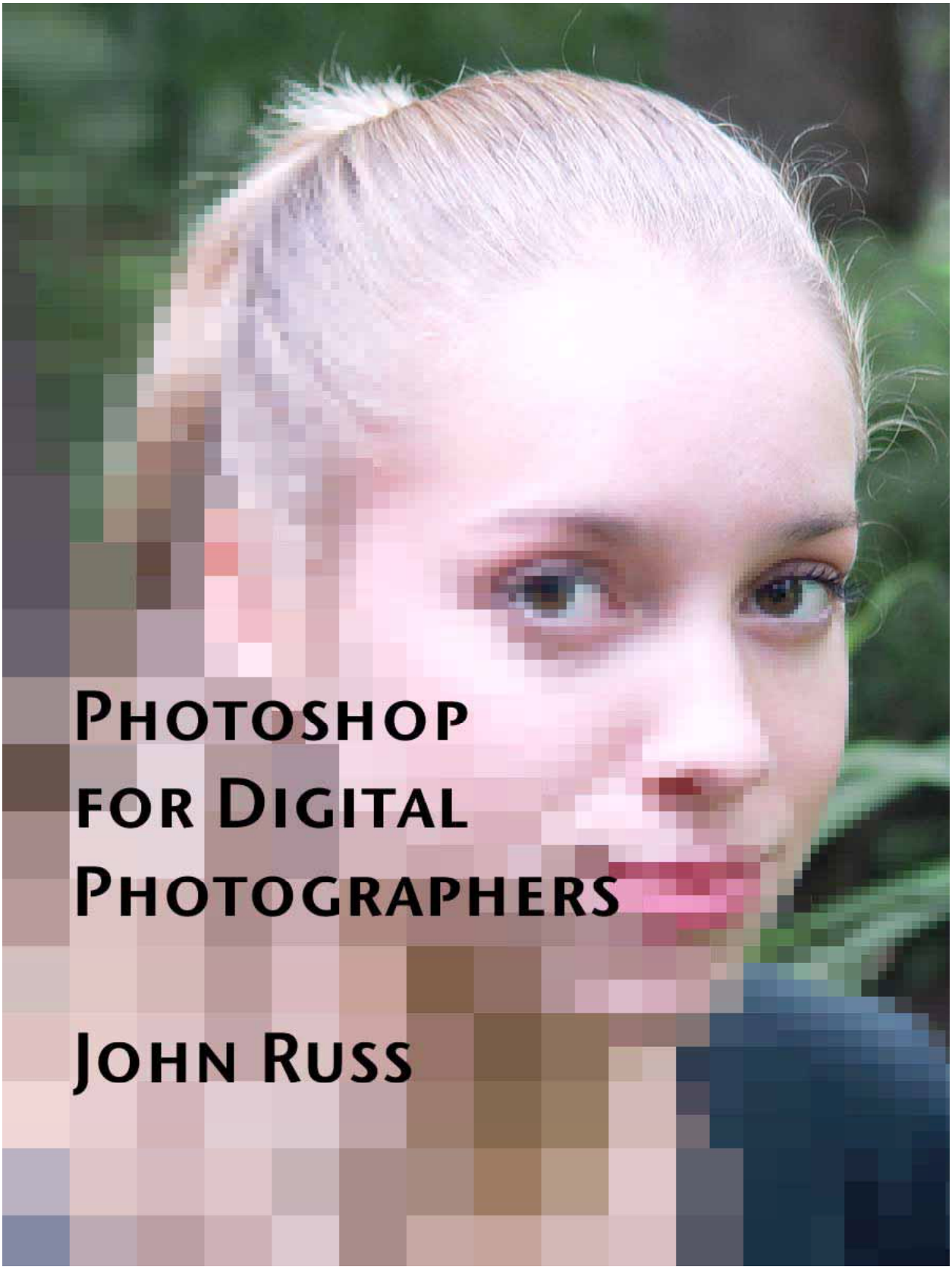


Photoshop for Digital Photographers

John Russ

Photoshop
For
Digital Photographers

John C. Russ

(author of The Image Processing Handbook
and Forensic Uses of Digital Imaging)

A Practical, Hands-on Guide
For Digital Photographers
Using Adobe Photoshop
To Adjust, Enhance and Print
Optimized Images

Photoshop for Digital Photographers

John C. Russ

Copyright © 2002. All Rights Reserved.

Contents

Chapter 1 - Digital cameras and typical image defects **6**
Resolution and image size 6
CCD vs. CMOS 9
Single chip color cameras 10
Other ways to get color 12
Random noise 13
The Gaussian blur and median filters 16
Color shifts, saturation and gamut 27
Other optical problems 29
Storage and compression 31

Chapter 2 - Adjusting Contrast and Color **38**
The human visual system 38
The image histogram 41
Color spaces 44
Hue, saturation and intensity 46
Converting Color to Grey Scale 49
Limits and Clipping 51
Gamma (mid grey adjustment) 54
Curves 56
Color adjustments 59
Selecting control points 63
Other control candidates 64
Color adjustments for the subject of interest 66
Matching colors within the image 68
True color 72

Chapter 3 - Combining multiple images **75**
Merging pixel values for optimum exposure 75
Loss of saturation 76
Blending range 77
Keeping brighter pixel values 79
Keeping darker pixel values 80
Averaging multiple exposures 82
Highlights 83
Averaging multiple images 85
Extending the focal depth 87
Aligning and adjusting multiple images 90
Anaglyph stereo images 91
Assembling a montage 93
Larger arrays 97
Foreground and background isolation 101

Chapter 4 - Photo Post-Processing **105**
Visual enhancement 105
Color sensitivity 107
Color resolution 110
Spatial resolution 112
Lighting gradients 113
Generating a background 115
Sharpening edges 118
Color and Noise 121
Laplacian and Difference-of Gaussians filters 123
Light and Dark edges 126
Other detail enhancement techniques 131
Edges 133
Reducing noise 137
Other types of noise 143
Periodic noise 148
Cropping, zooming and interpolation 153
Restricting operations to a selected region 156
Masks and Selections 163
Layers 166
Layer blending modes 170
Manual editing 174
Cloning, Healing and Patching 175
Red Eye 181
Dodging and burning 183
Localized Sharpen and Blur 184

Chapter 5 - Printing images **187**
CIE color space 187
ICC Profiles 190
Other working spaces 191
Out of Gamut 192
CMYK and RGB 193
How printers work 195
Printing “resolution” 197
Contours 200
Compressing dynamic range 203
Color 205
Ink jet printers 207
Preparing images for printing 211

Index **214**

Introduction

Adobe Photoshop is a powerful program with tools that address several very different audiences and their needs. One is the creation of commercial artwork, another is the preparation of such images for printing, and still a third is the processing of photographic images for enhancement or printing.

As there are many audiences for Photoshop, so are there many books about it. Some of these just present substitutes for the program manual, with exhaustive treatment of how to use the various tools and functions but little information on why. Others select some narrow (but often deep) field of use and present valuable but highly specific details on how to retouch photographs, how to adjust colors for prepress workflow, the use of digital imagery in scientific and forensic applications, and so on.

This book intends to reach the growing group of digital photographers, either serious amateurs or beginning professionals, who realize that after the camera has acquired the photo, more work is needed to produce the final perfected image and hardcopy. It assumes some familiarity with cameras, either film or digital, and the elements of taking a well composed picture. But it emphasizes the differences between film and digital cameras, which impose some limitations and present some opportunities to the digital photographer.

One of the strengths of Photoshop is its ability to incorporate additional capabilities provided by third-party plug-ins, and several of these are described in the text as appropriate. But wherever a procedure can be performed using only the native Photoshop functions, that method is described, even if a plug-in exists that can simplify or speed up the process.

Many of the procedures described in this text are straightforward and even obvious once understood, but they can't really be learned just be reading about them and examining the example photographs in the book. It is essential to apply these methods to your own pictures, in order to really master the methods. Don't be afraid to try things in order to see what various operations do to your photo. One of the great advantages of digital photography is, after all, the low cost of messing up and discarding a picture.

Happy photography….

Chapter 1
Digital cameras and typical image defects

The range of digital cameras available on the market is extremely broad. With prices ranging from under $100 to several thousand dollars, consumer, prosumer and professional models offer a wide range of performance. Most digital cameras are intended for general use, but some are targetted specifically at markets such as studio use, attachment to microscopes or other scientific apparatus, high speed photography, etc. Most digital cameras are self-contained, but some models must be tethered to a computer. The highest-end cameras typically accommodate interchangeable lenses, often the same ones used on high-end 35mm camera bodies (even though the chip sensor is smaller than the exposed film area, which changes the meaning the lens focal length in terms of wide angle or telephoto use).

Understanding the principal camera specifications will enable the user to better select the model that is most appropriate for the intended application, and to use it intelligently to obtain the best pictures. The most recognized specification, which correlates at least approximately with price, is the stated resolution of the chip (e.g., 3 million pixels). At least in rough terms it is true that the greater the number of light sensors present on the chip, the greater the image resolution and the larger the resulting prints can be made, but the meaning of resolution, and indeed the meaning of the often abused word “pixel” need more explanation and discussion.

The internet makes it relatively easy to collect specification data on various camera models, and to compare them. One site that maintains up-to-date and unbiased data on most models, with useful comparisons and reviews, is Digital Photography Review (http://www.dpreview.com/). The illustrations in this text were taken with a variety of different digital cameras, principally a Sony DSC F717, but also using a Nikon CoolPix 990, Polaroid DMCle, and several other older Sony, Fuji, Olympus and Canon models. The examples and procedures shown are applicable to all of them.

Resolution and image size

The first important thing to understand about digital cameras is that with the exception of a few professional models fromCanon and Kodak they do not have as much resolution of a 35mm film camera. A moderately high end digital cameras as this is written has about 5 million transistors on the chip, and stores an image of about 2500x1900 pixels, each with 256 values for red, green and blue light intensity. Film does much better. The spatial resolution of a 35mm transparency is about 4000x3000 pixels, corresponding to 12 million pixels. A good film scanner can capture that amount of information from the film, and represents another way to bring digital images into the camera for processing.

More important than the spatial resolution in many instances is the dynamic range. The 0..255 range of brightness values stored in a typical digital image is usually described as an “8 bit” image, because 2^8 = 256 and corresponds to the number of bits in a computer byte used to store the red, green and blue values. That is plenty of range to produce displays on a computer CRT for viewing, with no visible steps between values. It is also more than enough range to drive modern printers, as discussed in **Chapter 5**. But is not enough to capture images from scenes containing bright highlights and dark shadows.

Film can typically record several thousand distinct brightness values, which would correspond to 10-12 bits. Achieving that dynamic range (sometimes called the tonal resolution) with a digital camera requires cooling the chip to reduce thermal noise, and recording a dark frame to subtract from each image. Such techniques are used with digital cameras in specific scientific applications, such as microscopy and

astronomy, but are not available for cameras used for general photography. The “RAW” output from some high end cameras can exceed 8 bits, but does not approach the range of film.

Scanning film can capture a much greater dynamic range in the digital image than a digital camera. Most scanners produce at least 12 bits of data per channel (i.e., for red, green and blue). This is usually specified as the maximum density that can be read. A value of 3.0 for maximum density corresponds to a 1000:1 dynamic range, and many scanners produce 3.4 or 3.6. The density scale is logarithmic, so a value of 3.6 corresponds to nearly 4000:1 (12 bits).

Photoshop handles images with bit depth greater than 8 as 16 bit per channel images, using two bytes to hold the intensity data for each of the red, green and blue channels. This makes the image files larger, and has a small effect on the image processing speed. Also, some Photoshop functions, particularly the use of layers, are not possible with 16 bit images. There are work-arounds for these limitations, some involving third party plug-ins and some using 8 bit copies of images to create selections that are then transferred back to the original 16 bit image. These techniques will be illustrated throughout this book. Generally it is wise to perform all of the color adjustments and processing on these images in their original 16 bit form and only reduce the result to 8 bits per channel at the end, in order to produce a copy for internet transmission or printing.

Capturing an image with a higher bit depth than can be displayed, viewed or printed may seem unnecessary but it is not. With a good photographic negative, it is possible to make a print in the darkroom that brings out the details in the bright regions, or in the shadow areas, and so forth. It is also possible to use techniques such as the unsharp mask (discussed in **Chapter 4**) to show details in both the bright and dark areas in a single print. (In fact, the name of the commonly used - and often abused - unsharp mask filter in Photoshop originates in the performance of the same method in the photographic darkroom.)

But if the captured image has only a shallow dynamic range, it is not possible to perform these subsequent manipulations. Either the bright highlight areas have been clipped to white, or the dark shadow areas have been clipped to black, or both. Once the information is lost, it cannot be recovered. There is much greater premium placed on getting the exposure exactly right in the first place, and for many scenes this means selecting an exposure that keeps just the portions of the scene that are known to be of interest at the time.

Camera manufacturers all specify “resolution” in terms of the number of transistors on the chip, or in some cases (even worse) in terms of the number of pixels in the image stored in memory and transferred to the computer. First, to clarify, the word “pixel” has several inconsistent meanings. It may be used to count the number of transistors on the chip, or the number of display points on the computer CRT, but it will be used in this text only for the number of pixels in the stored image, which is not necessarily the same as either of the other values. The display can be “zoomed” or enlarged to show individual stored image pixels as quite large squares on the screen. And there is not necessarily any direct correspondence between the number of transistors present and the number of numeric intensity values extracted from the chip to be stored as an image. That relationship depends on the nature of the chip and how color information is obtained, and is discussed below.

Most cameras store more pixels than is justified by the actual image resolution. Resolution in optics has several working definitions, more or less equivalent. The ones that make the most direct sense for evaluating digital camera images have to do with the smallest size feature that can be distinguished and the smallest separation between two features that can be seen. This is complicated by the effect of

contrast. An object only slightly different in brightness or color from its surroundings must be larger to be seen than one that has high contrast. A black and white test pattern provides the maximum contrast and the best (highest) estimate of resolution. A typical pattern, as shown in **Figure 1**, has lines oriented horizontally, vertically, and at various angles. As they get closer together the lines merge together and the smallest line width and separation that can be seen gives the resolution.

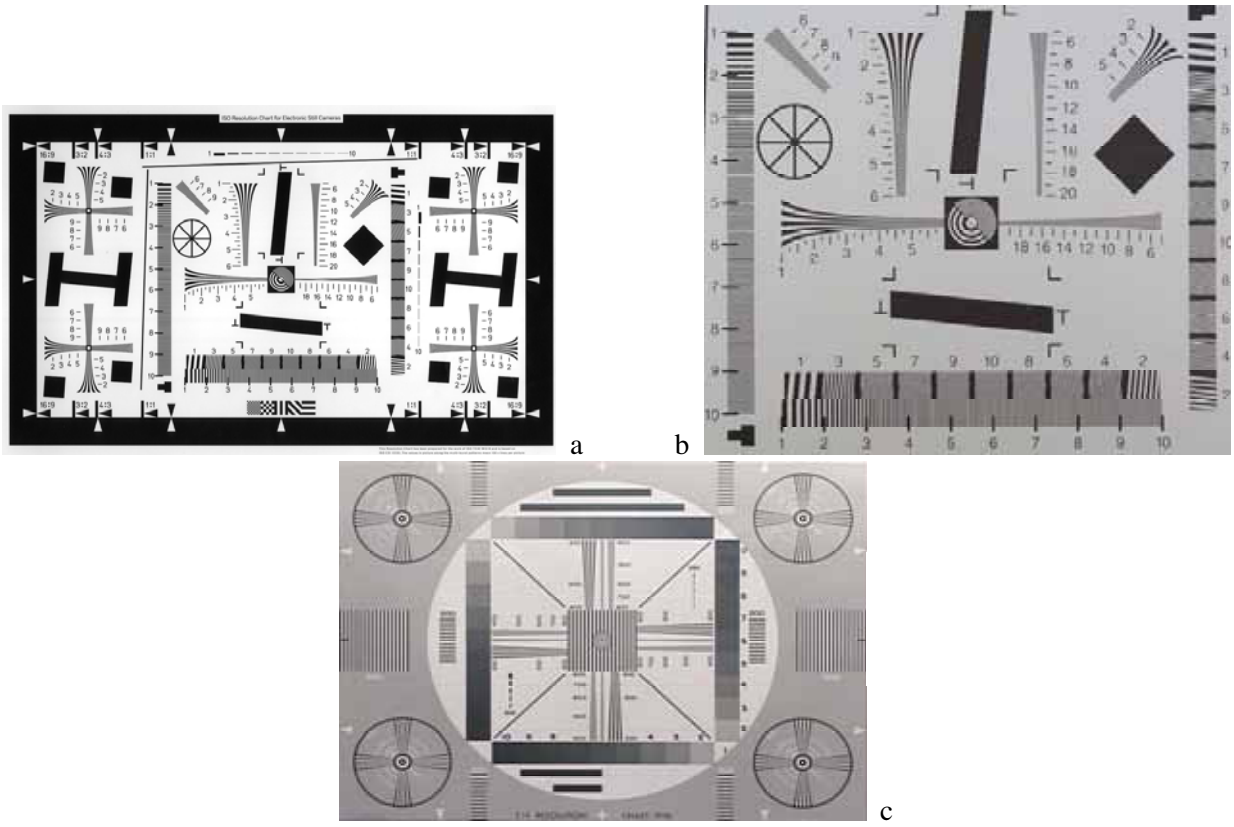


Figure 1. Examples of commonly used test patterns: a) ISO resolution chart (some areas with fine, unresolved lines appear as solid grey); b) detail of the central area of the chart at higher resolution; c) EIA test chart used for video cameras has ten grey scale steps as well as line patterns.

Since a digital image is not continuous but has pixels of finite size that average the intensity over some area, it is easy to see that narrow lines present a serious problem. If a black line (for instance) lies exactly on a column or row of pixels, it can be detected in the image even if it is much narrower than the pixel width, although it will appear grey rather than black. If the same line is shifted by one-half pixel, it will produce less contrast in each column and will be harder to detect. This would be true even if the pixels in fact represented a continuous sampling of the original image, but in practice the camera sensors are separated by spaces on the chip that provide electrical insulation, and perhaps contain additional electronics, and it is possible for the line to disappear altogether. That is why the line patterns are usually aligned at small angles in the test pattern. If a narrow line breaks up or produces a moiré pattern as it crosses rows or columns of pixels, it is too small to be resolved.

Typical measurements on digital camera images produce real resolution measurements expressed as "lines." This is the number of parallel lines in a test pattern that can be resolved, usually given as the number per height of the image. For a representative camera which stores images that are about 1900

pixels high, the typical measured resolution is about 1000-1300 lines. In other words, the actual resolution is about 50-70% of the value implied by the number of stored pixels (and, in many cases, by the number of transistors on the chip). The reasons for this discrepancy arise from the use of a single chip to get color images, and are discussed below.

It is also possible to measure resolution in other ways. For optical engineers, a common practice is to use the lens diffraction pattern or Fourier transform to locate the frequency at which the power spectrum becomes flat. Don't worry if that sentence is gibberish to you. Testing a number of images using that procedure, with some Photoshop plug-ins intended for a scientific market that perform the appropriate math, produced results that are the same as with the lines on the test charts. Most camera images are stored with many more pixels than justified by the actual resolution.

In photography, that is called empty magnification - enlarging an image past the point where detail can be seen, and where fine lines and edges just become fuzzy. For example, many scanner manufacturers specify as separate numbers the maximum optical resolution (more or less the real image resolution) and the maximum "interpolated" image size. The latter may be much greater than the former, and simply represents empty magnification. There is no additional information in such images; they merely take longer to process and more space to store.

CCD vs. CMOS

When a light photon strikes a silicon semiconductor it produces electrons. By collecting the electrons, it is possible to measure the amount of light that fell on the device. That is the basis for all of the solid state digital cameras, and many other devices including the majority of video cameras. Older video technology using vacuum tubes has been replaced in most applications by solid state chips because of their greater efficiency, ruggedness, immunity to environmental effects such as electromagnetic fields, and cost. There are two principle types of chips being fabricated, CMOS and CCD. Each consists of an array of virtually identical transistors formed by lithographic processes and covering an area that may vary from a few millimeters up to more than 12 mm, in which there may be hundreds of thousands (for video) or millions (for digital cameras) of transistors, each one used to measure the light intensity at that position.

An increasing fraction of inexpensive consumer cameras are using CMOS (complementary metal oxide on silicon) process chips as light detectors. These are also used in many inexpensive surveillance and conferencing video cameras, as well as toys, cell-phone cameras, etc. The advantage that they offer is one of cost, compared to CCD (charge coupled device) chips. CMOS chips are made using the same technology as memory chips, although they require somewhat tighter quality control and cleanliness. The devices essentially *are* memory chips, and each location in the array can be read out individually by addressing it (part of the address locates the column and part of it locates the row, so each location has a unique address).

The ability for selective readout is not present for the more costly CCD devices, but may sometimes be useful for reading out just those sampled locations needed to form a preview image, or for exposure metering, or to capture high speed images of a selected area, but few digital cameras take advantage of that possibility. On the CMOS chip, each light-collecting transistor has two or three (depending on the circuit design) additional transistors associated with it, which clear and control the device to start acquiring an image, and amplify the output signal for transmission to the digitization circuitry. It is the ability to place these control and signal amplification devices on the same chip with the light sensor that makes the CMOS chip low in cost, because little additional circuitry is needed to construct a camera.

The “camera on a chip” design is used in many toys, cell phones, and other devices. The lower power requirements of CMOS than CCD devices also simplifies needs for battery power in these applications.

In comparison, the CCD requires a more complex fabrication procedure that typically is limited to smaller size wafers containing fewer devices. External control and amplification devices are needed to control the chip, which is read out by shifting the collected charge row by row through the transistors until it can be amplified and measured. This process means that the entire image array is always read out in order to obtain image data, rather than being able to access any particular location as desired.

The CMOS devices are inexpensive, but have many important limitations for serious imaging. The presence of additional transistors means that a significant portion of the chip surface area is not light sensitive, so much of the light from the scene is lost. This “fill factor” can be improved by placing individual lenses over each light sensor to collect light from a larger area that covers the associated circuitry, but that adds cost to the chip and introduces another source of variation from one detector to another.

That variation is due in large part to the individual circuits that control and amplify the signal from each light detector. Pixel-to-pixel variation in the output signal from a uniform scene, called fixed pattern noise, is a characteristic of CMOS detectors that can be eliminated by post processing, by collecting a suitable background image (e.g. a photograph of a uniform grey card) that is then divided, pixel by pixel, into the measured image, but this is difficult or inappropriate for most applications of a digital camera.

Another limitation of CMOS cameras is the small volume of the transistors used as light detectors (this plagues some of the smaller chip sizes of CCD devices also). The dynamic range of the image, discussed above in terms of the number of brightness levels that can be distinguished, depends on two factors: the maximum signal that each detector can accommodate (the “full well” signal) and the amount of random noise present, which shows up as random speckle superimposed on the image. CMOS devices are poorer than CCDs in both categories, with the result that they typically have small dynamic ranges (sometimes less than 8 bits) and comparatively large amounts of noise. The small well size can be alleviated in some applications by binning (combining the signal from multiple detectors) but of course this reduces the spatial resolution of the image.

For all these reasons, but primarily the lower amounts of noise in the image, CCD chips are used for practically all higher end cameras, and are expected to continue that dominance well into the future. CMOS chips will continue to dominate the low end market and account for an increasing percentage of the number and total value of camera chips.

Single chip color cameras

The majority of digital still cameras use a single chip to collect light. Unlike many high end prosumer video cameras, which use prism optics to send the red, green and blue light from the image to separate CCD chips and then combine the electronic signals from those chips to get the composite color signal, most digital still cameras use a single chip with a pattern of colored filters in front of the light-sensing transistors that allow some of them to record light in different parts of the color spectrum.

The advantage of the three-chip design is, of course, that of higher spatial resolution. Each chip collects a full resolution image whereas in the single chip design only a fraction of the transistors see each portion of the color range. But in addition to the greater cost of the three-chip design, the complexity of

the optics required limits the ability to use short focal length lenses and absorbs a significant fraction of the incident light.

In the single-chip design, the most widely used arrangement of color filters is the Bayer pattern shown in **Figure 2**. Half of the transistors are filtered to detect green light, while one fourth receive blue or red light, respectively. There are other filter arrangements used, some with other colors, such as CMGY (cyan, magenta, green and yellow), and other spatial arrangements. Fuji uses a layout of transistors that is not a rectilinear array but rather a diamond pattern, and experimental designs with a hexagonal arrangement of detectors have been tried. In all cases, the choice of a filter pattern is closely tied to the method used for interpolating to determine the colors present, and it is only meaningful to examine or compare the overall results in test images.

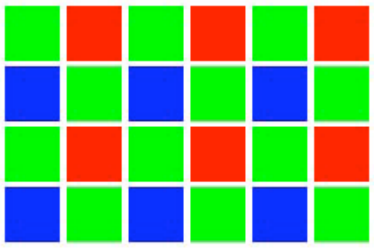


Figure 2. Bayer filter pattern

With all of these schemes, the major drawback is that the resolution of the image is reduced compared to the number of transistors present. To determine the amount of red light that fell at a location where there was no red detector, for instance, it is necessary to interpolate from the nearest places where red detectors were located. This interpolation cannot simply be done along each horizontal row of detectors as they are read out, for instance by using passive electronic circuits. It is necessary to perform the interpolation using information from adjacent rows as well, so at least some primitive computation is required.

Actually, there are some very advanced algorithms for performing this interpolation, called "demosaicking." The interpolation is generally done in a color space other than RGB, and may involve significant computation. Most of the algorithms are covered by patents, and belong to one or another of the various camera manufacturers. Since the user will generally have no ability to control or even discover the particular algorithm used, they are not discussed in detail here. But if you are really interested in learning more about them, there is a rich technical literature on the subject (a good place to start is "Demosaicking methods for the Bayer Color Array," by R. Ramanath, W.E. Snyder, G.L. Bilbro and W.A. Sander, which appeared in the Journal of Electronic Imaging, vol. 11, no. 3, pp. 306-315, July 2002; your local university library should have a copy, or you can download a pdf file from the web at <http://www4.ncsu.edu:8030/~rramana/Research/demosaicking-JEI-02.pdf>)

Regardless of how the interpolation is performed, color artefacts can appear along the boundaries between colored regions. Usually these are not objectionable in real photographs because the colors are not fully saturated and the boundaries are not perfectly sharp anyway, but in some cases they may be visible. The same hybrid color median described below in **Figure 14** generally cleans up such boundaries well. More disturbing than boundary artefacts is the presence of a moiré pattern that superimposes color bands on regions of the image that have a texture with about the same spacing as the color filter spacing. **Figure 50** in **Chapter 5** shows an example, along with a way to remove the problem from the image.

In most cases, the primary effect of the color filters used in a single chip camera is to reduce the image resolution. The camera manufacturer will generally emphasize the number of transistors present in the camera and/or the number of pixels in the stored image. Often these two numbers are nearly the same, but some camera makers (especially Fuji) typically store images with considerably more pixels than the number of transistors. Clearly these additional pixels represent empty magnification, but even if the number of stored pixels is the same as the number of transistors the image does not actually have that much resolution. As noted above, the real resolution is limited by the interpolation and results in images that could be stored with from 50 to 70% as many pixels across the width and height of the image.

Other ways to get color

The use of a color mosaic filter on a single chip is not the only way to get color information. Some cameras used for scientific applications acquire three separate images using a single chip, using a separate external filter for each exposure. That has the advantage of capturing images with the full resolution inherent in the number of transistors on the chip, and also allows adjusting the exposure for each filter to capture low noise, well balanced images. The three exposures correspond to red, green and blue, and are combined in the camera in exactly the same way as the Photoshop Merge Channels command. In fact, it is possible to capture three separate images using a monochrome (black and white) camera with three filters and then merge them in Photoshop to obtain a color image. Of course, the problem with this method is that it is very time consuming and only applicable to static, unchanging scenes. This might be acceptable in some studio work, but not for pictures involving people.

Another approach to constructing a color camera uses three separate chips, each with an array of photo-sensitive detectors. The incoming image is split by prisms into red, green and blue light, each of which falls onto one of the detectors. The output signals from the chips are then combined electronically to produce the color result. This method is often used in high end video cameras, but the resolution of color images is not that high (as compared to modern digital still cameras), and the chips used are relatively small. The cost of three-chip cameras is also increased because of the optics and the need to maintain alignment of the three chips. Scaling this arrangement up to handle the larger, higher resolution chips used for digital still cameras is difficult because angled light paths through the prisms absorb more light and produce color casts or shading across the image (usually from top to bottom). This problem, and the physical size of the prisms, limits this method to relatively long focal length (telephoto) lenses. In addition, the absorption of light in the optics makes the three-chip method quite inefficient, requiring lots of light intensity to obtain good images.

A very new method for getting color images with a single chip and without sacrificing resolution has been developed by Foveon Corp. At each location on the chip, three transistors are constructed. These are deposited one on top of the other, as shown in **Figure 3**. The topmost transistor captures the blue light, with green and red light penetrating through it to greater depth in the silicon. The next transistor measures the green light. The long wavelength red light penetrates the farthest in the silicon, and is measured by the deepest transistor. The color responses aren't that exact, but they are known for the dimensions of the device and allow processing the output from the transistors to produce a color signal containing the conventional R, G, B information.

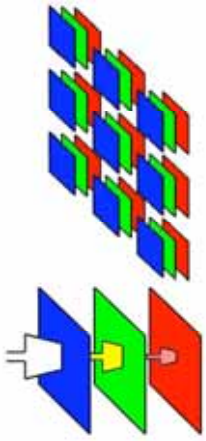


Figure 3. Foveon sensor (schematic).

The Foveon approach has great potential, and is used in the SD9 camera marketed by Sigma, but at present there are many problems still to overcome. The greatest of these is that the fabrication process for depositing multiple transistors in layers is only applicable to CMOS devices, which as noted before are much noisier (and have other problems) as compared to the CCD chips used in most high end cameras. For the red channel, the absorption due to penetration through the two upper transistors produces significant additional noise. The shallow depth of the transistors and the need for support transistors which occupy space on the chip results in small transistors with a limited dynamic range,

Random noise

Single chip cameras with mosaic color filters are by far the most common type used in digital still cameras. In the most common design (Bayer pattern), one fourth of the transistors are filtered to detect blue light. The use of half of the transistors to detect green and the remaining fourth to detect red is a relatively crude approximation to the wavelength response of human vision, which is most sensitive to green light and falls off at both ends. Silicon is not very sensitive to blue light, and the amplification of the signal from the blue transistors creates noise in the signal. In almost all cases, the random ("speckle") noise in the blue channel of a digitized image is much greater than that in the green channel. Red is usually intermediate.

In the example shown in **Figure 4**, a long exposure (1 second) and high ISO setting combine to produce considerable speckle noise. This is least evident in the luminance channel, with the blue channel showing the most noise. Also note that the color balance is off, with a cyan cast to the grey image.

Because human vision is also not very sensitive to blue light, the presence of this noise may not be noticeable unless the photograph is processed in the computer, to adjust color balance or enhance the visibility of edges and detail. In that case, the noise is often magnified and becomes visually distracting.

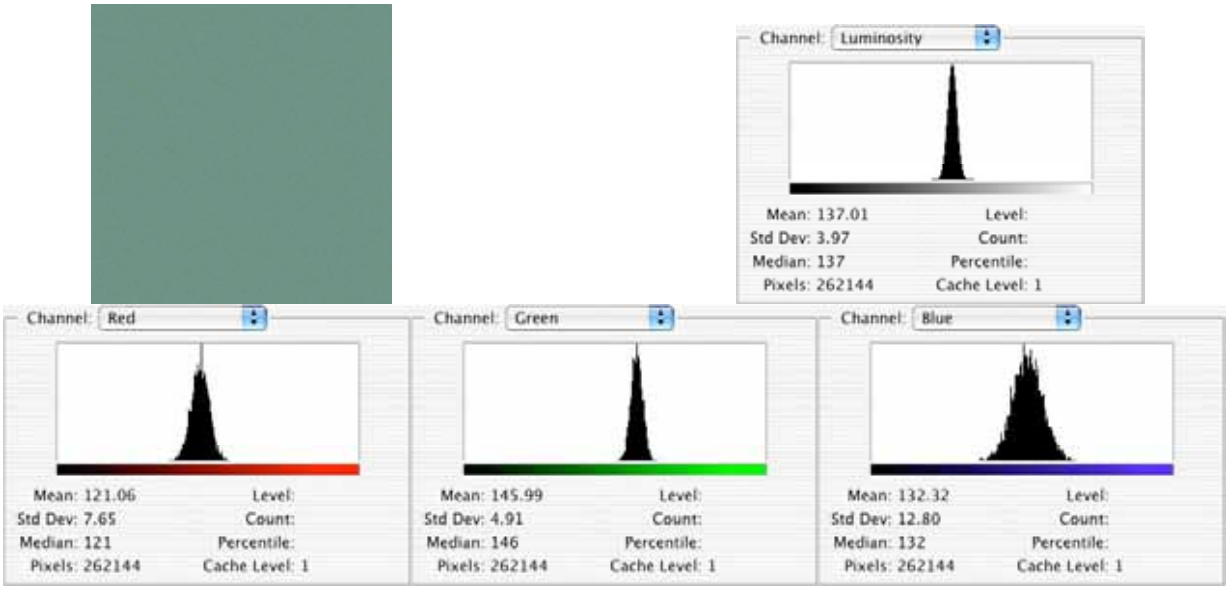


Figure 4. Fragment of an image of a uniform grey test card, with histograms of the luminance and the red, green and blue channels.

It is helpful to have a basic understanding of the noise and its multiple sources. There are two fundamentally different types of noise present. One, called multiplicative noise, produces a variation in pixel brightness that is proportional to the brightness. In other words, the noise source is multiplied by the signal - hence the name - and there is more noise in the bright areas than in the dark regions. But most cameras output an image that mimics the behavior of film, and film responds to light logarithmically rather than linearly. (Note that this applies to the final output image; depending on how the conversion from the inherently linear output of the chip is performed, the raw image that some cameras make accessible may or may not be converted to this behavior!) The result is that this type of noise, which arises in the creation of charge in the chip and its subsequent amplification, ends up having the same contribution across the full range of brightness in the image. Consequently, the width of the peaks in the histogram are the same for dark, medium, and bright values.

The width of the peaks (the amount of noise) generally increases as the amount of signal amplification increases, which corresponds to the ISO setting of the camera. This value typically lies in the range from 50 to 800, and is intended to reflect the equivalent speed of traditional film under similar exposure conditions. Just as high speed film tends to be grainier than low speed film, so a high gain in the digital camera produces more random pixel variations (“noise”) in the recorded image.

The second basic type of noise is additive noise. It arises from thermal noise and transfer losses in the chip, and increases with the length of exposure. Cooling the chip can reduce this source of noise, but while that is routinely done in many scientific and astronomical applications it is not practical for most digital camera applications. If the exposure time exceeds about 1/10 second (and especially if it becomes several seconds) this can be an important, even dominant noise source. Additive noise contributes a constant amount of variation to both bright and dark pixels, but because of the logarithmic nature of the output image this shows up as broader peaks in the dark part of the histogram and narrower ones in the bright part.

That characteristic is particularly objectionable, since human vision detects brightness changes based on the relative difference in intensity, not absolute differences. In other words, a change from a pixel of brightness 20 to one of 25 is a change of 5 parts out of 20, or 25%, but a change of the same magnitude from a pixel with brightness 220 to 225 is only 5 parts in 220 or about 2%. The latter is just at the threshold of visibility, while the former is glaringly obvious. Additive noise produces more pixel variation in the dark parts of the image, which is exactly where it is most visually apparent.

Some better cameras reduce the effect of additive noise by recording a "dark image" for the same length of time as the real exposure but with the shutter closed, and then subtract this from the image. This method is effective and can, of course, be carried out manually if the camera does not perform the function automatically. This also reduces some other sources of image defects, particularly fixed pattern noise that results from variability in the output from different transistors (especially a problem in CMOS cameras).

Most camera images contain at least some of both additive and multiplicative noise. This results in peak widths in the histogram that are at least slightly broader for dark regions that for light ones. **Figure 5** shows an image that is useful for comparison of several noise removal techniques (image courtesy of photographer Michael Almond, one of many on his site at www.michaelalmond.com). Taken with a Minolta DiMage 7i with ISO setting of 100, shutter speed of 1/125 second, and F-stop 8, it has random speckle noise that is visible in the uniform regions, as well as a variety of sharply defined but irregular edges corresponding to the various mountain ridges, which vary in contrast.

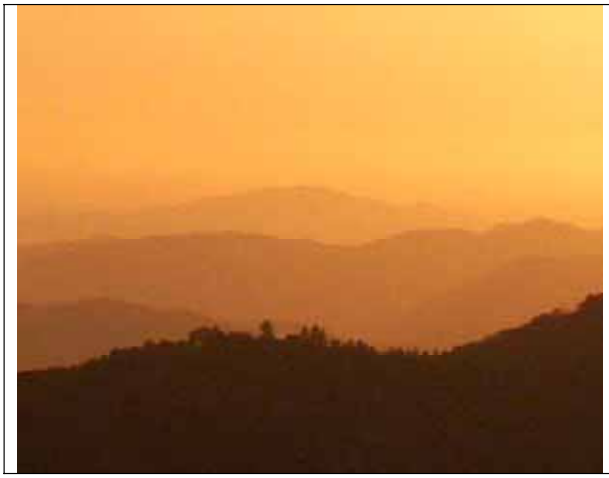


Figure 5. Sunset scene with atmospheric glow and mountain ridges: a) original image; b) selected region used in following examples.

A histogram of three small regions in the light, medium and dark areas of the image is shown in **Figure 6**. The peak widths vary, indicating that the brightness variation is a combination of additive and multiplicative noise. The change in the width of the noise peaks is small enough that it is practical to use the same techniques for reducing noise in both the bright and dark areas. This is usually the case, although in extreme cases where this is not true, selections can be used to apply different methods to regions with different noise characteristics.

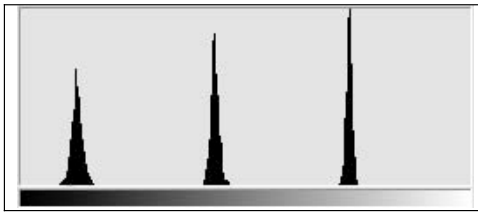


Figure 6. Histogram of three areas in the dark, medium and light regions of **Figure 5b** showing the variation in peak widths.

The Gaussian blur and median filters

The best solution to the noise problem is to take pictures with lots of light, which reduces several of the sources of random noise. But obviously this isn't always possible, and sometimes it is important to find ways to process a noisy image to reduce the visibility of random speckle noise.

Two basic tools are available for noise reduction. Gaussian smoothing averages each pixel with its neighbors. Median filtering ranks the pixels in a small neighborhood in order and replaces the central pixel with the middle value in the ranked list. Both methods are applied to each pixel to produce a new image that shows less noise.

These are very basic and important techniques in image processing, which use quite different methods, and it to make the processes clear it is helpful to consider a very simplified example. In **Figure 7**, an array of pixels is shown both as grey scale and integer values. The Gaussian smoothing method can be illustrated in a simplified way by using the Photoshop Custom filter. The array of weight values shown in the figure is multiplied by the individual pixel values, the results are added, and the sum divided by the total of the weights, to arrive at a new value for the pixel in the center of the array. In practice, larger arrays of weights are used that better fit the shape of a Gaussian or bell curve.

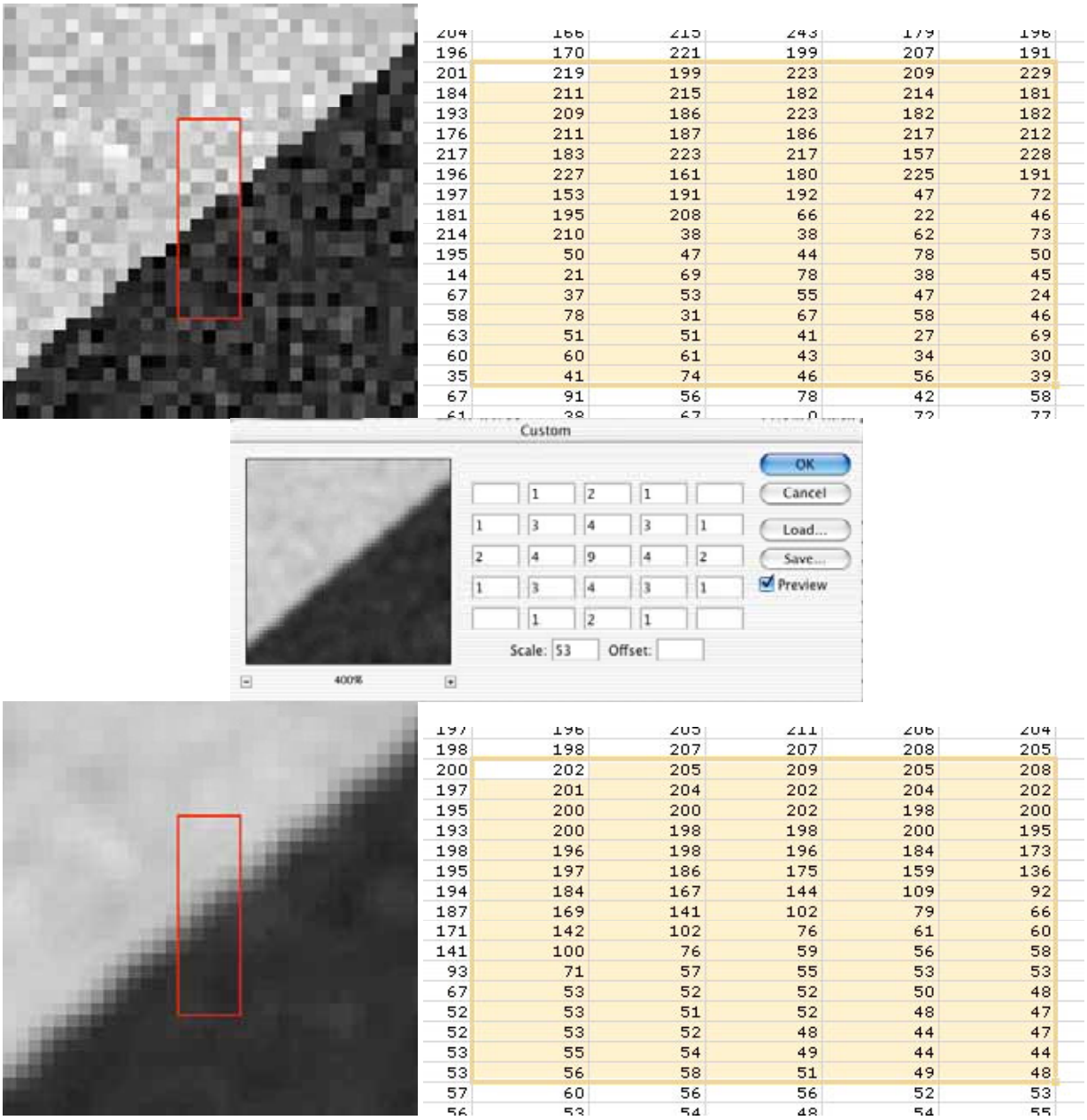


Figure 7. Enlarged fragment of an image with random noise and a step or boundary: a) grey scale representation of pixels; b) numeric values (0 = black, 255 = white); c) Custom filter approximating a Gaussian smooth; d, e) result of applying filter to original image.

Weighted averaging using numeric values that approximate the shape of the Gaussian curve can be shown mathematically to produce the least blurring of detail for a given amount of noise reduction. The controlling parameter is the radius of the filter (in technical terms, the radius is the "standard deviation" of the Gaussian, the point at which the value falls to 67% of the maximum). The filter actually includes pixels in a neighborhood with a radius more than 6 times as large as the radius value. **Figure 8** shows two sets of Gaussian weights with different radii, and hence different amounts of noise reduction. For the random noise present in most digital camera images, radius values in the range from about 0.5 to about 2 pixels are generally useful. Note that the Photoshop Custom filter is not large enough to apply

most Gaussian filters, and is restricted to integer values; the built-in Gaussian Blur filter works correctly and takes advantage of the fact that the Gaussian filter is "separable," meaning that it can be implemented much more quickly and efficiently by applying sequential vertical and horizontal blurs.

0.5 pixel

	0.0020	0.0060	0.0020	
0.0020	0.0432	0.1172	0.0432	0.0020
0.0060	0.1172	0.3184	0.1172	0.0060
0.0020	0.0432	0.1172	0.0432	0.0020
	0.0020	0.0060	0.0020	

2 pixel

			0.0001	0.0001	0.0002	0.0003	0.0003	0.0004	0.0003	0.0003	0.0002	0.0001	0.0001			
		0.0001	0.0002	0.0003	0.0005	0.0007	0.0009	0.0009	0.0009	0.0007	0.0005	0.0003	0.0002	0.0001		
	0.0001	0.0002	0.0004	0.0008	0.0012	0.0016	0.0020	0.0021	0.0020	0.0016	0.0012	0.0008	0.0004	0.0002	0.0001	
0.0001	0.0002	0.0004	0.0009	0.0015	0.0024	0.0032	0.0039	0.0042	0.0039	0.0032	0.0024	0.0015	0.0009	0.0004	0.0002	0.0001
0.0001	0.0003	0.0008	0.0015	0.0027	0.0042	0.0057	0.0069	0.0073	0.0069	0.0057	0.0042	0.0027	0.0015	0.0008	0.0003	0.0001
0.0002	0.0005	0.0012	0.0024	0.0042	0.0065	0.0088	0.0106	0.0113	0.0106	0.0088	0.0065	0.0042	0.0024	0.0012	0.0005	0.0002
0.0003	0.0007	0.0016	0.0032	0.0057	0.0088	0.0121	0.0146	0.0155	0.0146	0.0121	0.0088	0.0057	0.0032	0.0016	0.0007	0.0003
0.0003	0.0009	0.0020	0.0039	0.0069	0.0106	0.0146	0.0176	0.0187	0.0176	0.0146	0.0106	0.0069	0.0039	0.0020	0.0009	0.0003
0.0004	0.0009	0.0021	0.0042	0.0073	0.0113	0.0155	0.0187	0.0199	0.0187	0.0155	0.0113	0.0073	0.0042	0.0021	0.0009	0.0004
0.0003	0.0009	0.0020	0.0039	0.0069	0.0106	0.0146	0.0176	0.0187	0.0176	0.0146	0.0106	0.0069	0.0039	0.0020	0.0009	0.0003
0.0003	0.0007	0.0016	0.0032	0.0057	0.0088	0.0121	0.0146	0.0155	0.0146	0.0121	0.0088	0.0057	0.0032	0.0016	0.0007	0.0003
0.0002	0.0005	0.0012	0.0024	0.0042	0.0065	0.0088	0.0106	0.0113	0.0106	0.0088	0.0065	0.0042	0.0024	0.0012	0.0005	0.0002
0.0001	0.0003	0.0008	0.0015	0.0027	0.0042	0.0057	0.0069	0.0073	0.0069	0.0057	0.0042	0.0027	0.0015	0.0008	0.0003	0.0001
0.0001	0.0002	0.0004	0.0009	0.0015	0.0024	0.0032	0.0039	0.0042	0.0039	0.0032	0.0024	0.0015	0.0009	0.0004	0.0002	0.0001
	0.0001	0.0002	0.0004	0.0008	0.0012	0.0016	0.0020	0.0021	0.0020	0.0016	0.0012	0.0008	0.0004	0.0002	0.0001	
		0.0001	0.0002	0.0003	0.0005	0.0007	0.0009	0.0009	0.0009	0.0007	0.0005	0.0003	0.0002	0.0001		
			0.0001	0.0001	0.0002	0.0003	0.0003	0.0004	0.0003	0.0003	0.0002	0.0001	0.0001			

Figure 8. Two Gaussian filters: a) 0.5 pixel radius; b) 2.0 pixel radius. The weight values are shown as decimals summing up to a total of 1.0; in many cases they are approximated by a set of small integers and the result is divided by the sum of the weights.

The same averaging procedure is carried out for every pixel in the image, to produce the result shown. Since every pixel in the neighborhood is involved in forming the total, no values are discarded. Because the noise randomly increases and decreases the brightness of some pixels, the weighted average of the pixels in the neighborhood has less variation that the original and so the random speckle noise is reduced. This process is called a "linear procedure."

The simple Gaussian smooth (selected in Photoshop as Filter->Blur->Gaussian Blur) is effective at reducing the speckle noise in uniform regions, but as the undesirable side effect of blurring sharp boundaries by averaging the pixel values across the boundary and introducing new, intermediate brightnesses for pixels along the boundary (**Figure 9**).

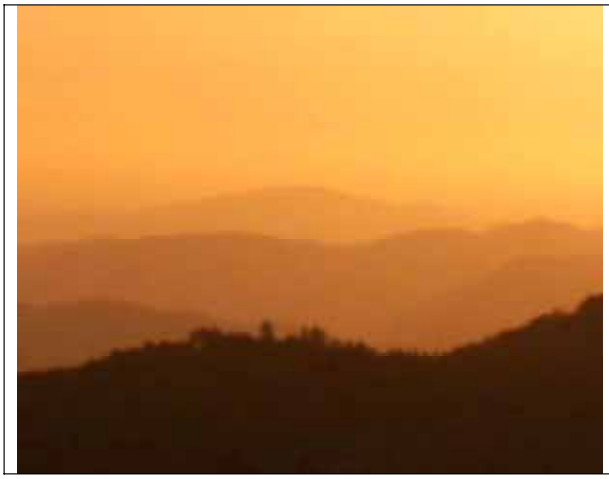


Figure 9. Gaussian smoothing (radius = 1.5 pixels) applied to test image.

Median filtering, by comparison, is a “nonlinear procedure.” It discards values from some pixels in the image. The pixel values in the neighborhood are ranked into brightness order, and the middle value in the list (statisticians call this the “median”) is assigned to the central pixel. Again, this process is carried out for all of the pixels, to produce a new resulting image. By discarding the extreme values, the amount of scatter (the noise) is reduced. One of the characteristics of the median is that edges are not blurred, but retain their full original contrast, because there is no averaging of the original pixel values. This is usually very beneficial in maintaining the sharp appearance of images while reducing the random noise present. However, the fine details of the shape of the boundary are altered (lines become straighter, corners and protrusions are rounded off) as shown in **Figure 10**.

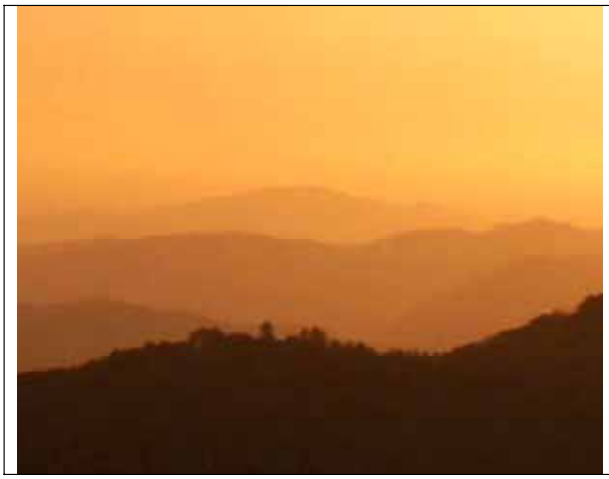


Figure 10. Median filter (radius = 2 pixels) applied to test image.

The radius adjustment for the median filter in Photoshop controls the size of the neighborhood of pixels that are ranked. A radius of one pixel corresponds to the central pixel and the 8 pixels that touch it, while a radius of two pixels corresponds to a 5x5 square. As the neighborhood size increases, the use of a square neighborhood rather than a circular one introduces some subtle artefacts into the resulting image. Third-party plugins such as the Reindeer Graphics Fovea and Optipix packages include medians that use round neighborhoods to overcome this.

In some special cases, it is desirable to have a median filter that employs a neighborhood that is not round but has a specific shape that can remove linear features such as scratches or power lines. Scratches on film that is digitized, or scan line noise from video images, typically has an horizontal orientation. Power lines may be horizontal, or have any other alignment and may not all be the same, in which case a selection of a region may be used to isolate ones with a particular directionality. Using a median filter with a user-defined neighborhood that runs across the lines or scratches to be removed can effectively remove them, as shown in **Figure 11**. This is quicker and requires less interaction than the Photoshop healing brush or patch tools.

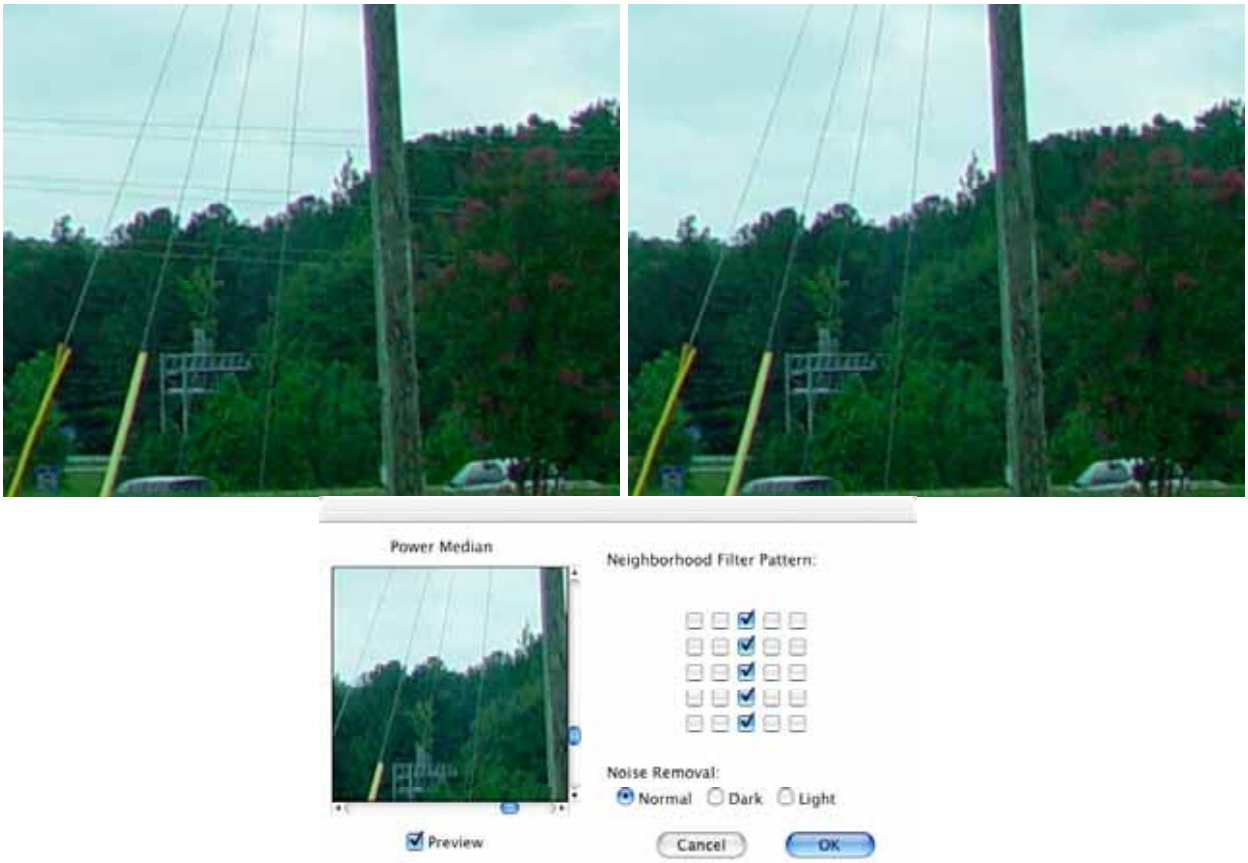


Figure 11. Image fragment containing power lines (a) and their removal by a median filter (b) using a neighborhood perpendicular to their orientation (c), defined using a Reindeer Graphics plugin. Notice that the guy wires have a different orientation and are not removed by this application of the filter.

Median filters also do not shift edges, but they do round corners and blunt sharp points. That problem can be eliminated by using a modification of the conventional median - called a hybrid median - that ranks several subsets of the pixels in the neighborhood, and then ranks their medians, to produce a final result. The hybrid median produces a result (shown in **Figure 12**) that better preserves the irregularity of the boundaries and presence of narrow lines in the original image. This method is not provided in Photoshop, but the Reindeer Graphics Optipix and Fovea packages offer it as a plug-in selection.

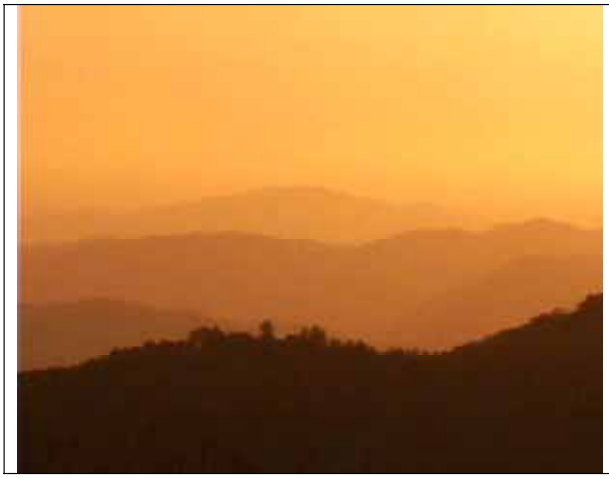
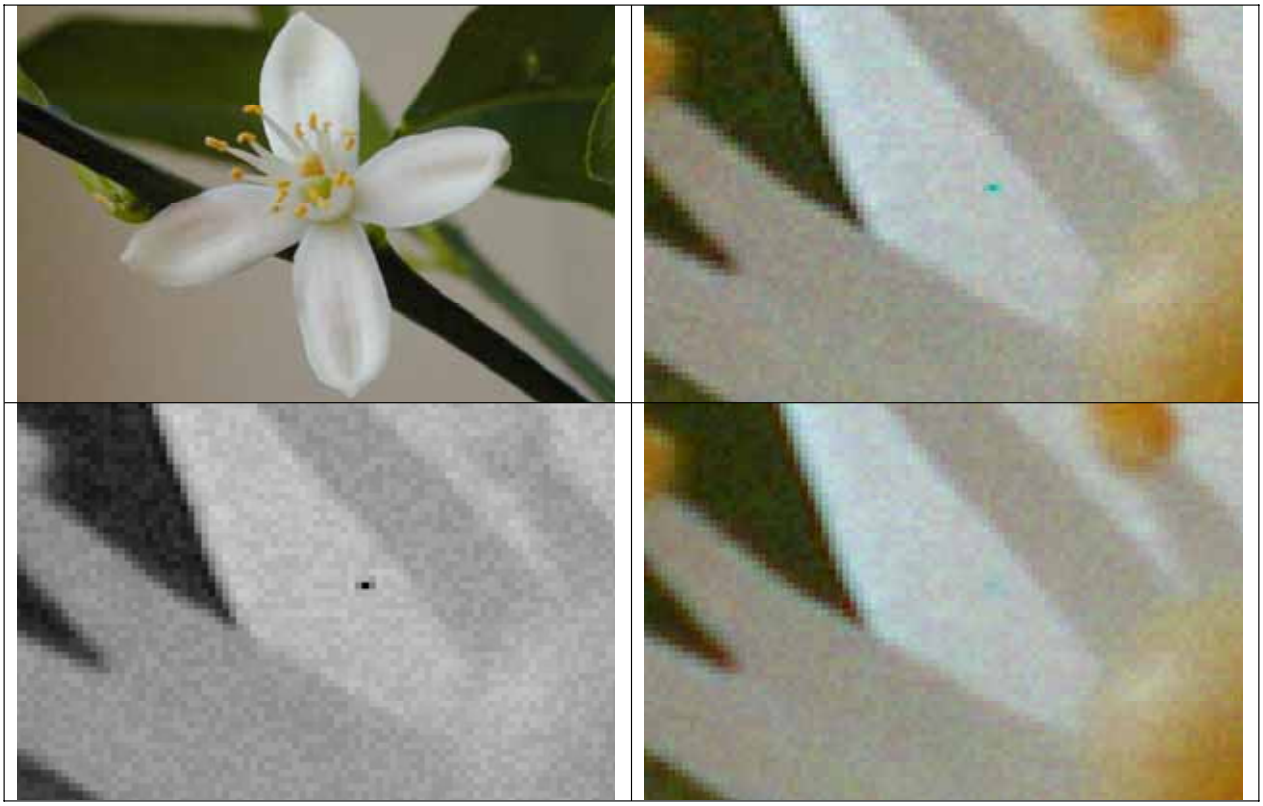


Figure 12. Hybrid median filter (radius = 2 pixels) applied to image.

The median filter is also very good at removing another source of noise in digital images, which results from dead or locked transistors. In most production chips with millions of transistors, there are a few that are either inoperative (produce no output) or shorted out (produce always the maximum output). This produces a characteristic type of noise (called “shot noise”). In a monochrome image, pixels would be black or white. In a color image, depending on the filter pattern and interpolation scheme used, it generally results in one or several pixels with a strong primary color cast. For example, a locked (maximum output) red-filtered transistor would produce a red spot from 1 to 3 pixels across, while a dead (zero output) red-filtered transistor would result in a similar sized spot with a strong cyan color. **Figure 13** shows an example.



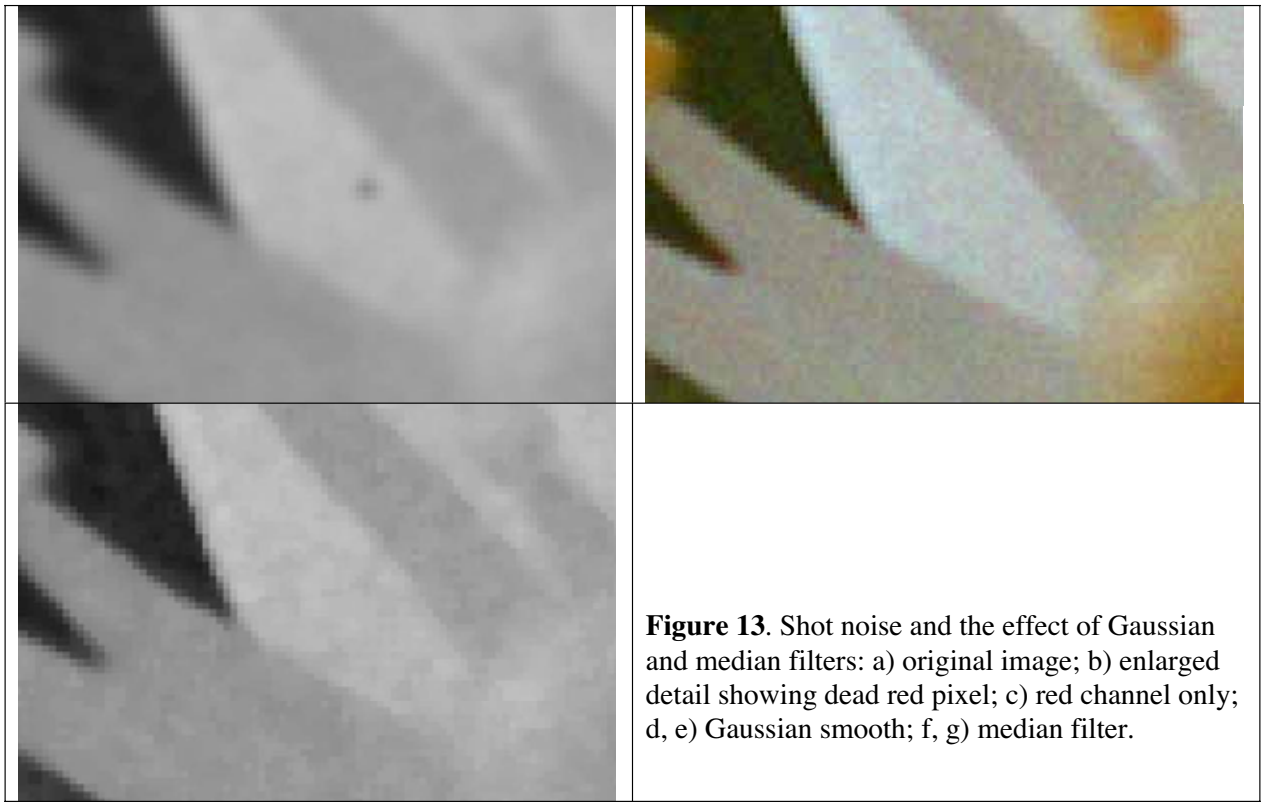


Figure 13. Shot noise and the effect of Gaussian and median filters: a) original image; b) enlarged detail showing dead red pixel; c) red channel only; d, e) Gaussian smooth; f, g) median filter.

Applying a Gaussian smooth to the red channel in this image spreads the colored spot out across a wider area and reduces its magnitude, but does not eliminate it. Applying a median filter to the red color channel discards the extreme values and replaces them with a value from the nearby pixels in the neighborhood, which produces a much superior result as shown in the figure.

In this case, applying a median to the individual color channels makes sense because the defect being removed arises from the color filtering used in the camera and is present in the channels. But for the case of random speckle, a median filter should not be applied to individual color channels. Doing so selects different pixels in the neighborhood in each channel, and so combines different values of red, green and blue. That can produce new combinations of colors that produce a greater visual distraction and impression of noise than present in the original image.

The most common way of median filtering a color image is to use the pixel brightness values (either the average of R, G and B, or the L channel in the Lab color space representation of the image) to rank the pixels in the neighborhood and choose the median, and then use the color values from the selected pixel for the central pixel. When the principal type of random noise present is a variation in brightness, that is a reasonable method and gives good results. But in many color images, particularly those that result from the interpolation of color values from a color-filtered single chip camera, the noise is commonly greater in the color data than in the brightness values.

How can the idea of a median filter - selecting the pixel within a neighborhood that is the “middle” value - be applied to color images in which each pixel has three numeric values? The answer to that lies in vector math. Probably the simplest way to understand the process is to visualize a color space (this works in any color space: RGB, Lab, HSI, etc.) with each pixel in the neighborhood plotted as a point. It is possible to calculate the distance from each point to all of the others. The point whose sum of

distances to its neighbors is least is the vector or color median, and its values are selected for the central pixel in the neighborhood. The color median does a very good job of reducing both random speckle noise and shot noise in color images, as shown in **Figure 14**. The Reindeer Graphics Optipix filters combine the color median with the hybrid median to preserve corners and lines.

Figure 14. Enlarged portion of an image showing both speckle and shot noise: a) original; b) after application of a hybrid color median.

Another approach uses the Gaussian smoothing method but restricts its application to areas of similar pixel values, so that averaging across edges does not occur, and consequently the edges are not blurred. One way to accomplish this is by using a selection. Drawing selections by hand is rarely practical, but fortunately there are ways to generate them automatically. **Figure 15** shows the most widely used approach, which is available at numerous web sites to be downloaded as an action, an automation plugin, or even a droplet.

The basic procedure is to duplicate the image and use the edge tracing routine in Photoshop (Filter -> Stylize -> Find Edges). There are many other edge-locating routines available as plugins, and for this purpose any of them can be used. They produce a result as shown in **Figure 15a**. Copying this image and pasting it into the quick mask channel for the original image produces the result shown in **Figure 15b**. The mask covers the edges and prevents the smoothing operation from affecting the pixels there (**Figure 15c**). Similar use of masks to limit processing is discussed at length in **Chapter 4**, where it is most commonly applied in the opposite way: to limit sharpening to edges.

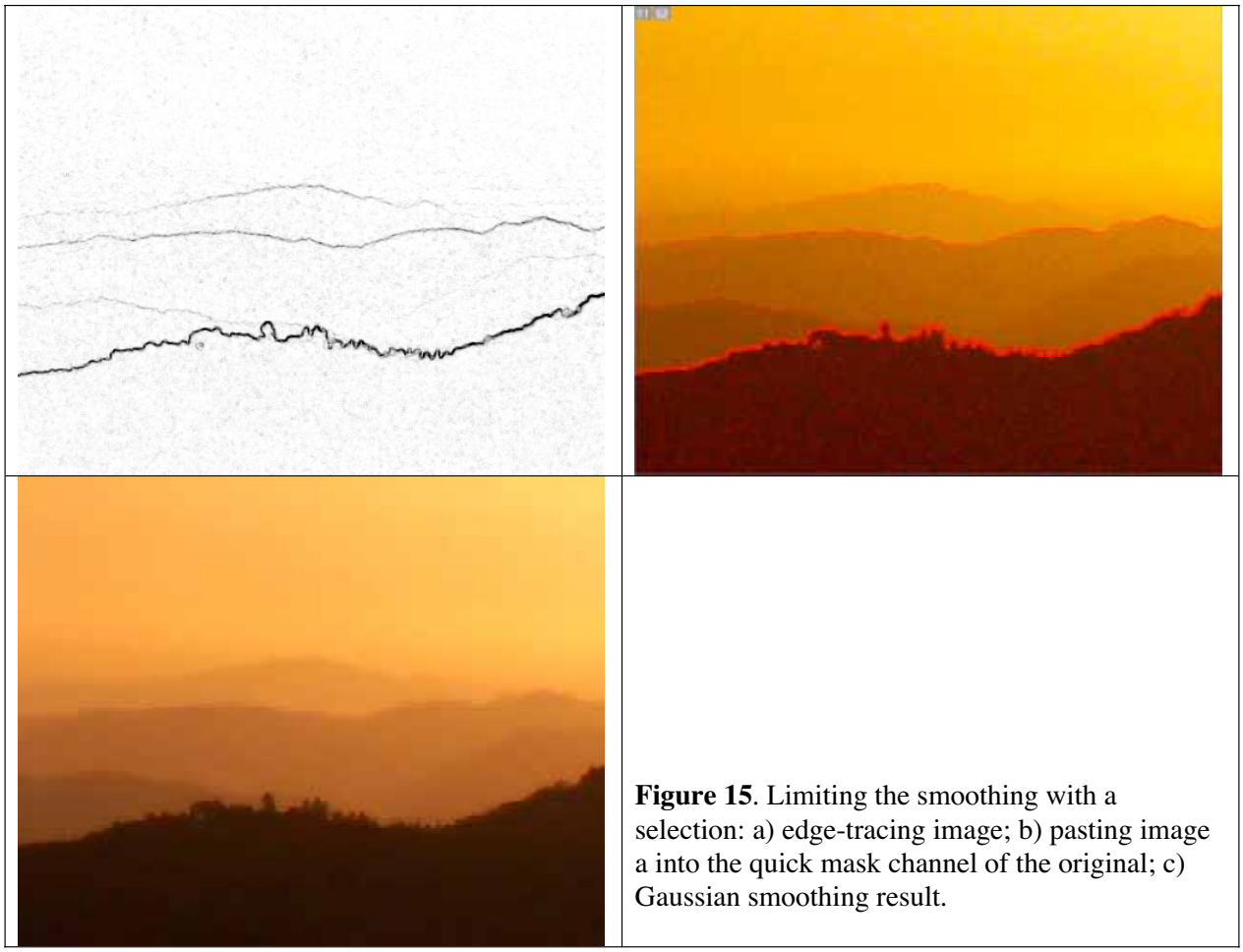


Figure 15. Limiting the smoothing with a selection: a) edge-tracing image; b) pasting image a into the quick mask channel of the original; c) Gaussian smoothing result.

Unfortunately, the contrast in the edge image must be adjusted before using it as a selection in order to protect the fainter edges. This is done using the Image -> Adjust -> Levels controls, but is inherently a trial-and-error process with no preview of the final result. Also, the selection protects some pixels along the sides of the more prominent edges so that they are not smoothed, and still appear to be noisy.

Better results can be obtained using a thresholded Gaussian blur filter. Photoshop includes a threshold slider on the unsharp mask filter (discussed in **Chapter 4**), but not on the Gaussian blur filter. However, plugins such as those from Reindeer Graphics provide that additional functionality, calling it a Conditional Smooth. The radius slider selects the size of the neighborhood, but the threshold slider excludes from the averaging calculation those pixels that lie within that region but whose difference from the central pixel exceeds the threshold value. **Figure 16a** shows the effect of a large threshold (20 brightness steps), which preserves the more prominent edges. A smaller threshold (6 brightness steps) also preserves the subtle edges of distant ridges while still reducing the random speckle noise within uniform regions. (**Figure 16b**).

Figure 16. Conditional or Thresholded smoothing: a) threshold = 20; b) threshold = 6.

Most noise reduction using Gaussian blurs is performed on the individual red, green and blue channels, typically using the same filter radius on all three. That is convenient but not optimum. It was pointed out before that there is typically more noise in the blue channel than the green, for example, so using a larger radius on the blue channel would better reduce the noise, while using a smaller radius on the green channel would better preserve the sharpness of boundaries.

Several standard methods recommended for processing digital camera images follow this logic. They apply a Gaussian smooth, or in a few cases a median filter, to the noisy blue and perhaps also to the red channel. The green channel is not blurred, and in many cases is actually sharpened slightly using an unsharp mask filter as discussed in **Chapter 4**. The result is an image in which the visual sharpness of edges is enhanced without increasing the visibility of the random speckle noise. The downside of this approach is that performing different manipulations on the individual color channels can alter the proportions of the red, green and blue, resulting in the introduction of new colors that are visually objectionable, as shown in **Figure 17**.

Figure 17. Processing in RGB space: a) original noisy image fragment; b) red and blue channels smoothed, green channel sharpened.

Superior results can be obtained by working in a different color space, such as Lab or HSI. Converting the image from RGB to Lab space in Photoshop, performing the sharpening operation on the L (lightness or luminance) channel, while smoothing the noise in the a (red-green) and b (blue-yellow) channels produces better results as shown in **Figure 18**. It also retains fine detail better as shown in **Figure 19**, although it can introduce some color shifts that can corrected as discussed below.

Figure 18. Processing the same image from **Figure 17** in Lab space.

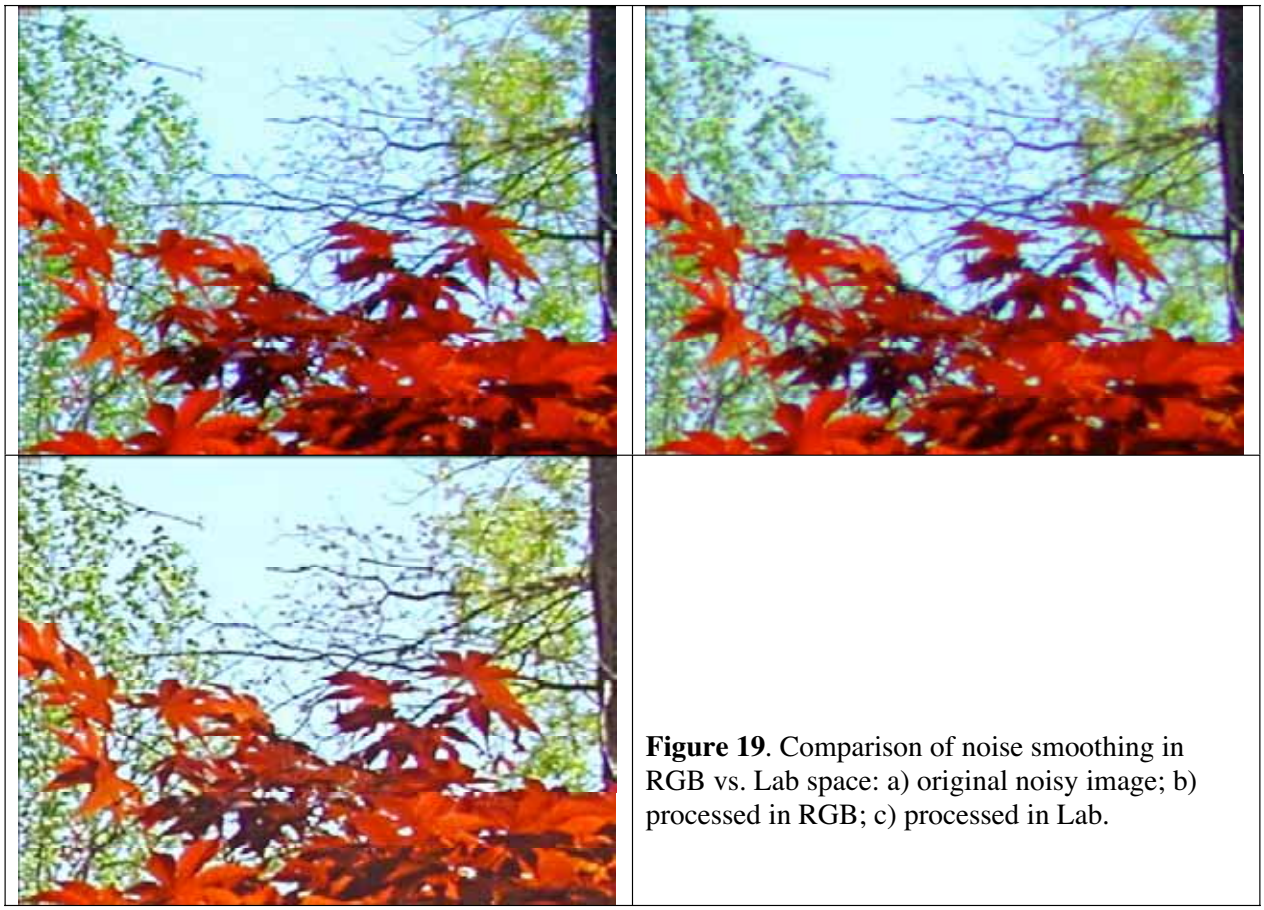


Figure 19. Comparison of noise smoothing in RGB vs. Lab space: a) original noisy image; b) processed in RGB; c) processed in Lab.

Color shifts, saturation and gamut

All cameras, digital or film, have difficulty in recording "true" color. The response of the detector (chip or chemistry) to different wavelengths of light produces a result that represents the original scene in a different way than human vision, which itself is notoriously poor at judging accurate color information. In all cases, the most important uncontrolled variable is the lighting, which is never neutral but has some color distribution itself that is compounded with the colors and surface characteristics of objects in the scene to produce the color data. So the problem that arises is to know about the lighting and how it alters the perceived or recorded colors.

The transistors on the chip in a digital camera are filtered for specific colors, often RGB, but the output of the chip is usually YCC (Y is luminance, the two C or chrominance channels are the red-green and blue-cyan signals as discussed above). The conversion of this signal to recorded RGB values depends on the "white point," or in other words what balance point of these signals is neutral in color. Some cameras have a few built-in white points that attempt to handle common situations such as sunlight, open shade, tungsten indoor lighting, etc. Many of the high end cameras allow you to perform a white point correction by pointing the camera at a neutral target and pressing a button. The former method assumes that your scene conforms to one of the standard types, and the latter requires that you actually have a white or grey target in your scene. Given a suitable white point correction, most cameras do a pretty good job of making neutral colors grey or close to it. Further color corrections can be handled in Photoshop as discussed in **Chapter 2**.

More problematic is the gamut of the camera. This is the range of colors that can be distinguished, and no digital camera covers the full range of colors visible to the human eye. The gamut is still typically greater than that of the printers that will be asked to produce hard copy, as discussed in **Chapter 5**, but typically does not include the most saturated colors. In order to prevent the recorded images from appearing undersaturated and "flat," many cameras boost the colors - particularly the red - for greater visual effect and more "zip." To be sure, some films do this as well, and photographers are accustomed to selecting color films for their ability (some would say bias) to capture and/or enhance particular colors.

The goal in recording a picture is rarely to produce a spectrometrically accurate recording of the colors present in the scene. A different type of device (a spectrophotometer) can be used for that purpose. Cameras are used to record pleasing and plausible representations of a scene, not to render true color. Remembering that distinction is important to obtaining satisfactory results. We are typically concerned with the ability of the camera image to distinguish small changes or differences in color within an image, or rarely between one image and another. For that purpose the color gamut of the camera, and the 8 bits (256 values) available for red, green and blue intensity, are usually adequate, although with little to spare. In some cases more than 8 bits are necessary (particularly to distinguish values in the darker regions of scenes, such as within shadows). In some cases highly saturated regions of color will appear uniform and flat, losing shading and texture, because the camera gamut is insufficient.

Closely related to color gamut and correct setting of the white point for color conversion, is the determination of exposure. Many cameras use the chip itself to sample the image to calculate the best exposure, using various algorithms to center weight or sample various areas. These algorithms are good for a surprisingly wide range of "typical" images, as indeed they are for film cameras. But full adjustability and spot metering are still important for serious work. The need to set proper exposure is greater for digital cameras than for film cameras because the dynamic range of the chip is less than that of film, as noted above. It isn't possible to recover as much detail from the under- or over-exposed regions of an image in a digital camera as can be done in the darkroom, and this is particularly true if the film development can be varied in addition to selecting papers and print exposures.

The balancing advantage of digital cameras is the live preview and instant review that they afford. Examination of the image itself can be used to determine whether the desired detail is present in highlights and shadows, and if not the exposure can be adjusted and (hopefully) another picture recorded. Methods for capturing multiple images with different exposures and combining them are discussed in **Chapter 3**.

Color problems can also appear as a side effect of the optics in the camera. Chromatic aberration in lenses produces color fringes at sharp edges because the different wavelengths of light cannot be focused to exactly the same plane. This is true for any lenses, but generally those with the best corrections (through the use of multiple lens elements made from glasses with different refractive indices) are fixed focal length lenses available for professional 35 mm. cameras. These can be used with some professional digital cameras that use camera backs compatible with those same lenses.

But most digital cameras come with their own lenses, which are not interchangeable and usually have fairly large zoom ranges, from wide angle to telephoto. This places severe demands on the optical designer, and in many cases some chromatic aberrations can be seen as color fringing at extreme wide angle or zoom settings with wide open apertures, typically in the corners of the image (farthest from the optical axis). The problem can be sees as purple fringes along edges, particularly ones oriented

circumferentially in the image. The current state of the lens design art is pretty good, however, and except for the inexpensive range of cameras this is not likely to produce an serious problems, particularly since the viewer usually does not examine the corners of images as closely as the center.

Fringing can also occur at the edges of images if the light does not fall perpendicularly onto the chip. The transistors and filters do not lie in a perfect plane, but have some depth. Light striking the devices at an angle can cast shadows or scatter to a different detector, producing color shadows and fringes. This is rare, because most optics are designed to direct the light perpendicularly onto the detector. Furthermore, this is not a new requirement. Film also has depth, with the color sensitive layers at different positions through the film thickness, so non-perpendicular incidence also produces color fringes with film cameras.

Other optical problems

Several types of optical problems are common with digital cameras. Inexpensive consumer cameras with fixed focal length lenses are often prone to distortion and vignetting. The use of higher quality zoom lenses with adjustable focal length in intermediate and high end cameras does not always alleviate this problem, because at extreme wide angle or telephoto settings both distortion and vignetting may occur. Even with professional grade cameras that accept interchangeable lenses from 35 mm cameras, improper matching of the lens to the chip, or non-perpendicular illumination of the chip, can cause problems. Added to that is the fact that the chips are smaller than the original 35mm frame, and so the viewfinder does not correspond to the captured image, and the lens' stated focal length is not directly meaningful.

Distortion in the camera optics causes a stretching or compression of the image, typically at the corners or edges. The two most common types of distortion are pincushion or barrel distortion. If the scene contains lines known to be straight, or if other images taken with the same lens (at the same zoom setting) have previously been used to calculate the lens parameters, then third-party Photoshop add-ons such as Andromeda's LensDoc can be used to remove the distortion.

Image distortion can also occur due to the use of extreme wide angle lenses, or foreshortening due to perspective. This is not an optical defect, and is inherent in the geometry of viewing a projection of the scene, but it can also be corrected using software as shown in **Chapter 3** (for examples, see **Figures 26** and **27**).

Many camera lenses produce vignetting, or darkening of the image toward the edges and corners. This may arise from the greater thickness of glass that the light rays reaching the corners of the image pass through, or because the individual lens elements are not quite large enough to pass as many rays to those areas. It occurs in film cameras, too, but when images are projected from slides or printed onto paper, human vision is not very sensitive to gradual changes in brightness and ignores the dropoff in brightness. The more limited dynamic range of digital cameras, and the desire to perform post-processing on the images, makes vignetting more of a concern for digital cameras. Vignetting is often accompanied by geometrical distortions and by loss of sharpness at the edges of the scene. **Figure 20** shows an example.

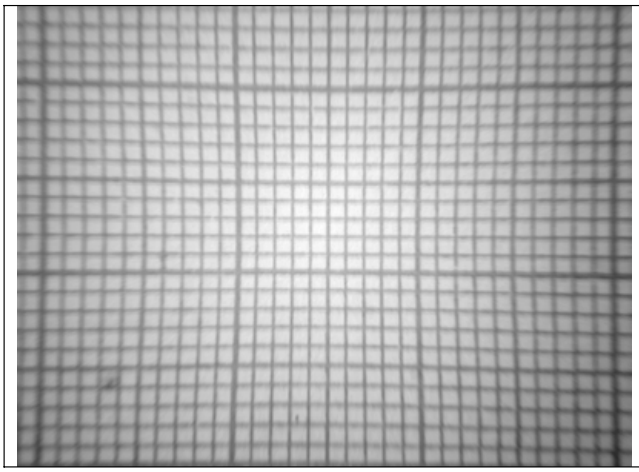


Figure 20. Photo (with a consumer-quality digital camera) of graph paper showing pincushion distortion, with vignetting and loss of sharpness at the edges.

In some studio or copystand situations, vignetting (and other shading such as may result from nonuniform illumination) can be corrected by recording an image of a uniform grey card with the same camera position and setup. Place this image in a layer on top of the original, invert its contrast, and adjust the opacity slider until the shading in the original image is cancelled out. **Figure 21** shows the results for the vignetting example in **Figure 20**. For scenes in which the direct recording of the shading or vignetting is not practical, it is sometimes possible to use the gradient tool to create an image that represents the shading and can be used instead.

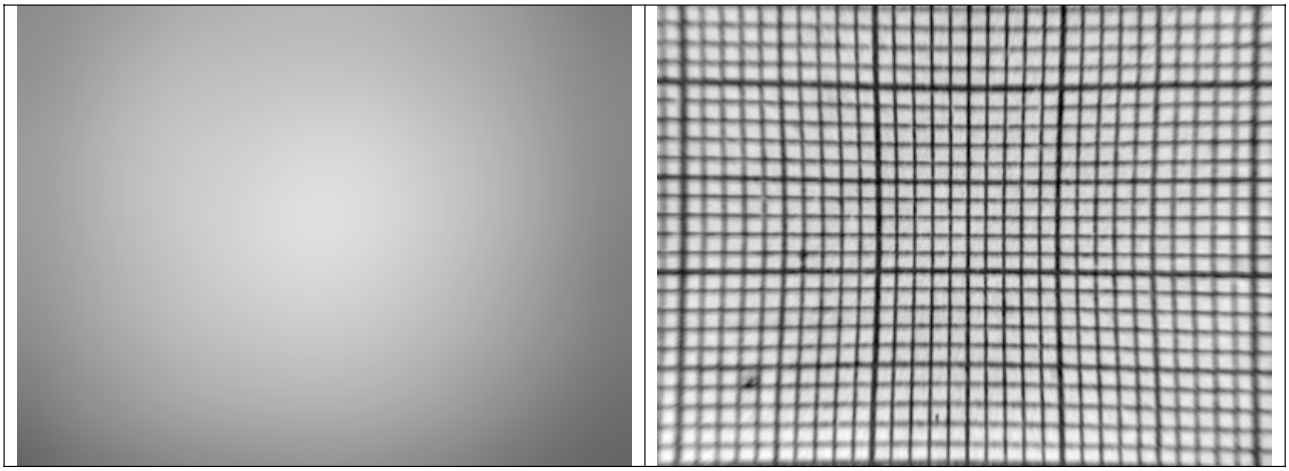


Figure 21. Correcting the vignetting in **Figure 20**: a) shading recorded by photographing a grey test card; b) result of overlaying the inverse of a onto the original and adjusting the opacity to about 50%. Notice that the distortion and out-of-focus problems remain.

Another problem that is essentially optical in nature is the viewfinder present in many inexpensive cameras. These are obviously much less costly than an LCD preview screen, and also do not require battery power, but create difficulties in knowing just what the captured image will show. Just as for 35 mm cameras of 50 years ago, before the widespread use of TTL (through the lens) viewfinders, the inexpensive optical viewfinder does not see exactly the same area as the lens, and it also views it from a different angle so that the agreement between the viewfinder image and the image recorded by the lens and chip becomes progressively worse for closeups. Even some LCD preview screens can be misleading

because of the limited resolution, difficulty in viewing them at an angle, and the fact that some display slightly less than the full image area.

Storage and compression

With the exception of a few camera models that directly transfer their images to a host computer, most digital cameras include some type of storage for the images. The possibilities include floppy disks, writable CDs, hard disks, and various configurations of flash memory (compact flash, memory stick, etc.). Most of these are removable, so that in principle an unlimited number of images can be acquired and stored.

In fact, for most models of digital cameras, battery life is more of a limitation than storage space. With LCD preview displays, motorized zoom lenses, built-in flash, and of course plenty of electronics, many cameras go through AA cells like locusts through a wheat field. Some cameras have high capacity rechargeable batteries, but even then carrying a spare fully charged battery as well as some spare memory is a good idea.

Flash memory, which requires no power to retain the data written into it, is the most common storage medium in digital cameras. Miniature hard disks can provide larger storage capacity and a lower cost per megabyte, but have higher initial costs, higher power requirements, and less tolerance for physical abuse. Because flash memory is relatively expensive, many cameras come with only a minimal amount. But adding the largest capacity (at this writing, from 128 to 512 MB depending on configuration) available is a worthwhile investment.

Because memory is expensive, and finite, the number of images that can be recorded is limited. For a typical high-end consumer camera, with “5 Megapixel resolution,” a full color image recorded without any compression would require 15 MB (5 each for red, green and blue). As noted before, some of the pixel values are interpolated and that isn’t the actual resolution of the image, but it is the size of the array that needs to be stored. It doesn’t take very many 15 MB files to fill up the storage on a compact flash or memory stick.

Consequently, practically all cameras provide image compression to fit more images into a limited amount of memory, generally using the widely standardized JPEG (Joint Photographers Expert Group) algorithms. A detailed understanding of how they work is not necessary here. In general terms, the method converts the image data from RGB to a YCC color space, and the color values are averaged to reduce the resolution of the color information by a factor of 2. The process then divides the image up into 8x8 pixel blocks, performs a discrete cosine transform on each block, quantizes the resulting values (rounds them off to varying degrees of precision), and encodes the results. The important compression step that reduces the file size is the quantization of the values. The “quality” of the JPEG compression is directly related to the size of the steps used in the quantization.

Other compression methods, using wavelets and fractals, use different strategies but all are driven by the same need - to store more images in limited memory - and use as a measure of success the same criteria - the ability to present to the viewer an image in which features remain recognizable and artefacts are generally not visually evident. Specifically, the goal of image compression is to keep those aspects of the image that seem to be important to the human vision system and its interpretation of images, and to discard as much as possible of the things that are not essential to that understanding. Obviously, this works differently for scenes of very familiar objects, where only a few clues are needed for recognition, than for unfamiliar scenes.

Compression factors of 3-10X are typically accompanied by modest loss of visually significant information from the image. Higher factors, from 10 to 50X, are generally not so satisfactory. To see what effects various amounts of compression have on an image, start with one that has been saved without compression (or, if your camera does not offer that possibility, with the highest quality - least compression setting). Save copies of this with Photoshop using the JPEG format with various quality settings, and note how the file sizes drop as the quality setting is reduced. Then open these compressed images, copy each one, paste it in a layer on top of the original, and set the layer blending mode to difference. **Figure 22** shows an example. Using a low quality setting, the image file size was reduced by a factor of 15. However, the position and color of edges and fine detail have been altered significantly.

Figure 22. Original uncompressed image (a) and the difference between the original and the JPEG compressed version (b), shown here with expanded contrast.

Another tell-tale sign that compression has been performed is the loss of spatial resolution in the color channels. **Figure 23** shows the compressed image from **Figure 22** converted to Lab color mode. The L (lightness) channel contains some visually discernible artefacts, such as the loss of defined texture in the center of the daisy and some of the petals on the rose. But the a (red-green) and b (blue-yellow) channels show just how much color information has been discarded. The color resolution was first reduced by half to compress the amount of data, and the results have been quantized so that each 8x8 pixel block has very few color values remaining, with large discontinuous steps from one block to the next. It is amazing that the human eye does not easily see this artefact in the color presentation of the image, being tolerant (as discussed above) of considerable color bleeding across brightness boundaries.

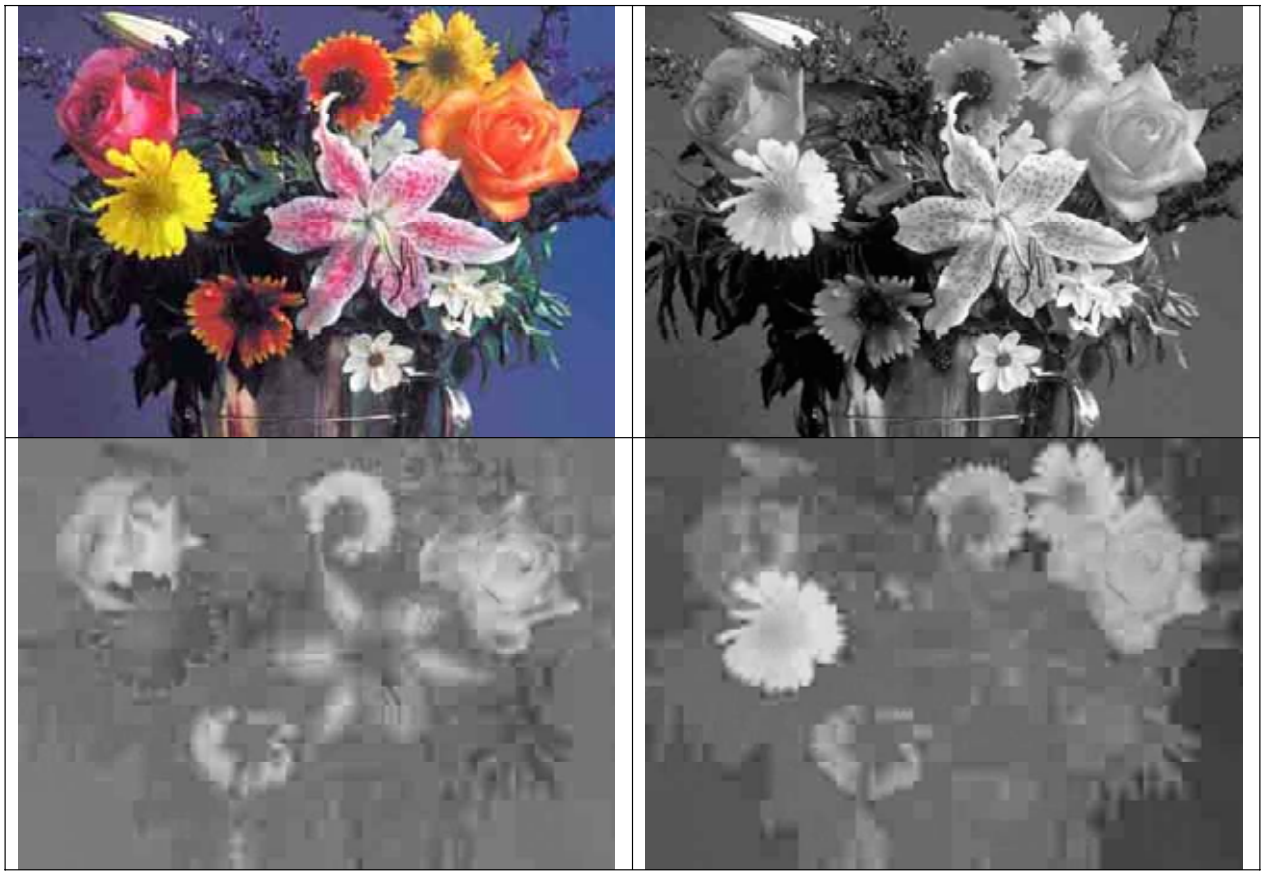


Figure 23. The compressed image from **Figure 22**: a) color display; b) L channel; c) a channel; d) b channel.

These same strategies - reducing the resolution of color information, discarding detail in portions of an image where there is a lot of detail, and allowing shifts in colors and in the positions of edges or other detail, are used by all of the compression techniques.

The most visually distracting artefact from JPEG compression is usually the “blockiness” that can result from the 8x8 blocks that are used to process the original values. Abrupt discontinuities between pixels in one block and their neighbors in the adjacent block attract attention because they line up along vertical and horizontal lines. There is no way to recover the information discarded in the compression process, but the visual distraction that the block artefacts cause can be alleviated by processing. **Figure 24** shows the result.

In this example, the original image was compressed from 1 MB to 40K (a factor of 25), and has the typical nasty block pattern. The image was converted from RGB to Lab space and a duplicate layer generated. A image was created consisting of two-pixel-wide vertical stripes, with a spacing of 8 pixels. This was aligned with the vertical edges of the block pattern and used as a selection mask (by pasting it into the Quickmask channel). The pixels within the selection, namely those at the left and right edges of each block, were then smoothed using a linear Gaussian smooth (the weights of 0.1, 0.4, 0.75, 1.0, 0.75, 0.4, 0.1 are shown in the figure). Because of the selection, this smoothing used all of the pixels in the image but applied the results only to those along the block boundaries. The same procedure was then used to smooth the horizontal block boundaries. The final result is an image in which the block

boundaries are well hidden and the appearance of the image greatly improved. Obviously, this process can be automated as either an action or plug-in for routine use.

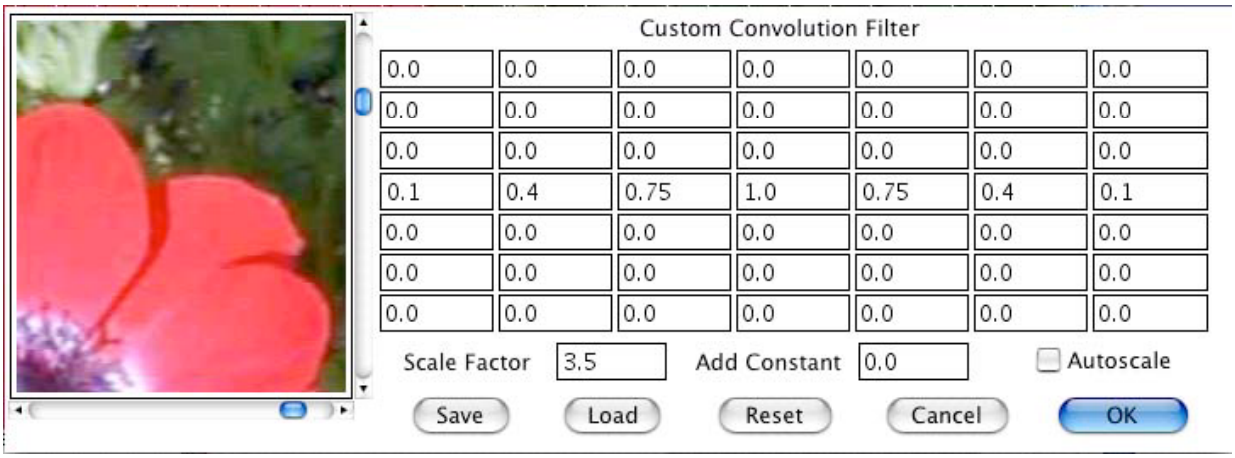

Custom Convolution Filter
0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.4 0.75 1.0 0.75 0.4 0.1
0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.0 0.0 0.0 0.0 0.0 0.0 0.0
Scale Factor 3.5
Add Constant 0.0
Autoscale
Save
Load
Reset
Cancel
OK

Figure 24. Removal of JPEG block boundary artefacts: a) original JPEG compressed image (enlarged); b) selection of pixels adjacent to vertical boundaries; c) custom filter applying a linear Gaussian smooth across the vertical boundaries; d) result of smoothing vertical and horizontal block boundaries.

Images can be transferred to the computer or directly to the printer either by removing the storage media and inserting that into the computer or printer, or by using the interface built into most cameras with a suitable cable. Most cameras now use USB (Universal Serial Bus) interfaces and are recognized by modern computer operating systems are either hard disks containing files that can be transferred, or directly as digital cameras. Some cameras use SCSI or Firewire interfaces, but these are not so widely available on the current crop of desktop computers.

If a digital image is destined for transfer to a printer (either directly or through a computer) without processing or modification, for printing without significant enlargement, then a modest amount of compression usually won't be noticed. When processing in the computer is used to adjust color balance, reduce noise, sharpen edges, or otherwise manipulate the image, the artefacts introduced by the JPEG process become more pronounced and visible. Enlarging the image generally makes the problems more obvious, as well.

Consequently, it may be acceptable to use a medium quality setting in the camera for "snapshot" purposes, but for serious work the least compression/ highest quality setting possible should be used.

And if your camera provides for an uncompressed image storage (usually RAW or TIFF), then that should be selected for really important shots or for ones that you anticipate will be processed significantly in the computer.

Uncompressed images not only take up much more storage space in the limited memory of the camera, they also take a lot longer to store. In most cameras, dedicated chips are used to perform the JPEG compression in a very short time, and the smaller file size can then be written to the comparatively slow flash memory more quickly than a larger, uncompressed image file.

The time required to write the acquired image to memory is an important consideration with digital cameras because it affects the "cycle time" between pictures. Many high end cameras have enough fast internal memory to hold a few full-resolution images, so they can quickly acquire a few photos to bracket exposure or focus settings. But once that memory is full, the images have to be written out to the much slower flash memory (or other, even slower storage), and all the user can do is wait.

There are other delays and waits with digital cameras, as compared to film cameras. One is the "startup" time, from switching on to being ready to take a picture. A second, probably the most annoying, is the delay lag between pressing the "shutter release" button and the actual acquisition of the picture. Many cameras reduce that somewhat by allowing a partial press to set the exposure and focus, but even so a delay of several tenths of a second while the electronics do their work (clearing the chip, which has been used to acquire the preview image and calculate exposure, etc.) feels like a long time to a photographer used to the immediacy of a mechanical shutter, and forces learning to anticipate the moment to be captured. If the camera then acquires a second "dark frame" to reduce noise, as discussed above, there is a further delay before the result can be seen and another photo recorded.

Whether or not these drawbacks of digital cameras outweigh their many advantages is a personal decision that varies with the user's experience and the type of images being taken. But the trend is certainly toward greater and greater use of digital techniques for acquiring, adjusting, enhancing and printing photographs.

Chapter 2
Adjusting Contrast and Color

All software packages, ranging from low-end utilities that may come bundled with cameras or scanners to high-end general purpose programs such as Adobe Photoshop, contain some tools for adjusting image contrast. As will be discussed below, these may be applied to the brightness information only, or in various color spaces to also adjust the color of the pixels. The goal in all cases is to increase the visibility of detail in the images, and to make the images more visually appealing (sometimes this is described as having more "pop" or "zip"). It is not, in general, to produce faithful and true representations of the color or brightness information in the original scene. That topic is very important in some photographic applications (such as producing illustrations for catalogs) but involves the image display and printing process at least as much as the image acquisition. It is taken up separately in **Chapter 5**.

The human visual system

Because the adjustment processes discussed here are specifically intended to produce results that are well suited for the human visual system, it is important to note a few characteristics of our eyes and brains and the effects they have on what we see, or think we see. First and most basic of all, we do not have any sense of absolute color or brightness. Almost everything in the human vision system is comparative. We compare the brightness or size or shape of features in the same scene, two at a time, almost literally by mentally shifting them so that they lie next to each other. If the two regions or objects were initially separated, or had different surroundings, that slows the process down and often biases the result. We compare images to objects or scenes in memory the same way, and that process is notoriously imprecise. In our minds, pictures are not stored as arrays of pixels, as the camera does, but in a highly abstract representation that concentrates on a few details that have attracted our attention and ignores much of the rest.

Figure 1 shows several three-dimensional objects. Some are different views of the same object, but some show a mirror image shape. The more different the orientations of the objects are, the longer it takes us to decide whether they are the same or different - the "turning over in the mind" process takes more time. When we compare things to remembered objects or scenes, not only does it take longer but (as one might expect) the accuracy suffers. Rarely do we store in memory those details of a scene that turn out to later be of importance for comparisons, and even that information is only stored as comparisons to still other remembrances. No wonder eyewitness descriptions of crime scenes and individuals are so unreliable.

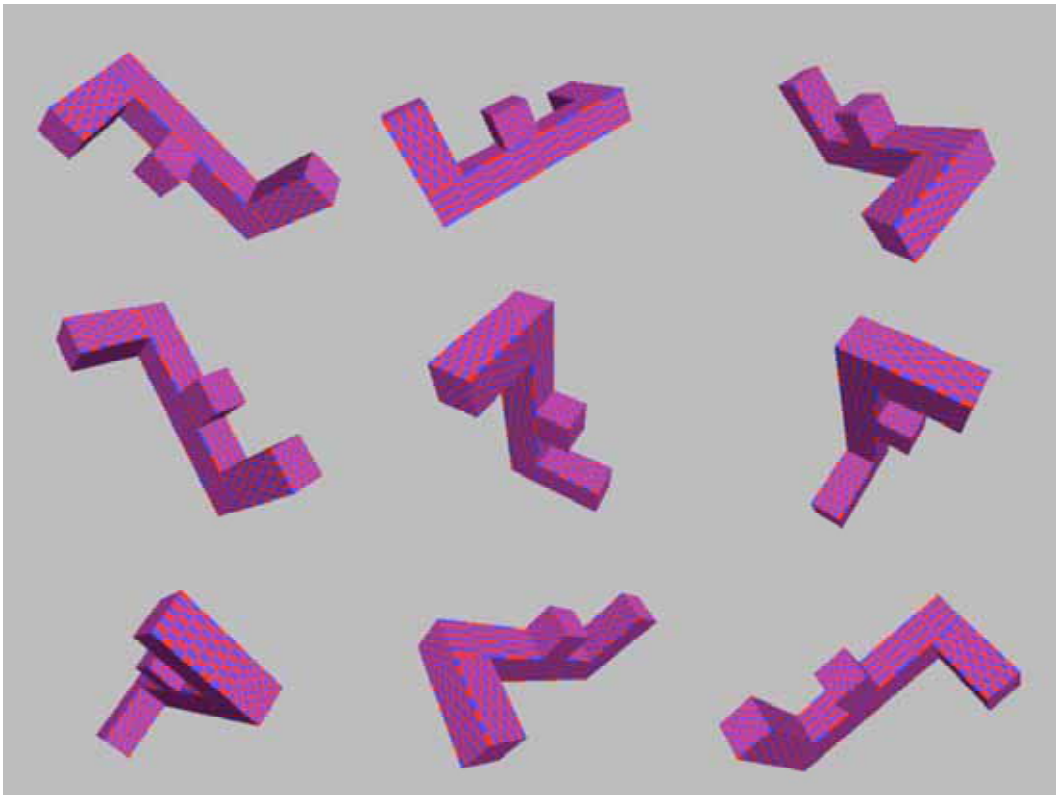


Figure 1. Several views of a three dimensional object. Which of the representations are identical and which are mirror images? The time needed to decide is proportional to the difference in orientation, indicating that we literally "turn the objects over" in our minds to compare them.

It is only in the central portion of the visual field (the fovea) where we have significant color sensitivity (or good spatial resolution, for that matter) because of the very high density of color-sensitive cone cells. We “look at” the features in a scene that our brains find interesting, which consists of rotating our eyes so that the chosen part of the image falls onto the fovea. It is the act of directing each eye toward the same feature and the muscle movements in the eye sockets from which we decide whether one object is closer or farther away than another.

The region surrounding the fovea (peripheral vision) is good for detecting motion, and has good low light level sensitivity because it has primarily rod cells which detect light but not color (in fact, we can see things better by averting our vision and looking “next to” regions of interest when it is very dark). The surroundings may also play an important role in color perception by providing information about the scene illumination. There is also a “hole” in the scene recorded by our eyes, the blind spot where the optic nerve connects to the retina, but our mind fills that in so that we are normally not aware of it.

Vision adapts in a short time to vastly different levels of illumination, and not just by changing the size of the aperture in the eye’s pupil. We are largely unaware of the automatic adaptation, except when the scene brightness changes so abruptly that the vision system cannot keep up with it. As the level of

illumination changes, so usually does the color of the light. This alters all of the colors reaching the eye, but our mind compensates for this and, again, we don't notice it.

We do not "measure" size or angle visually any more than we do brightness or color. Because we live in a world that has a lot of horizontal and vertical structures, both natural and man made, we compare things to those orientations easily. People can judge deviations from vertical of about one degree, and from horizontal of a few degrees. But at other angles our judgment is poor. Without markings on a clock dial, it is difficult to judge the time to the nearest minute (6 degrees). And the presence of other alignments in the image biases our judgment as well. **Figure 2** shows an example of diagonal lines that are in fact exactly parallel, but the don't appear to be so because of the horizontal and vertical lines around them. This process, called inhibition, makes nearby objects or details that are different in size, shape, or orientation appear to be more different.

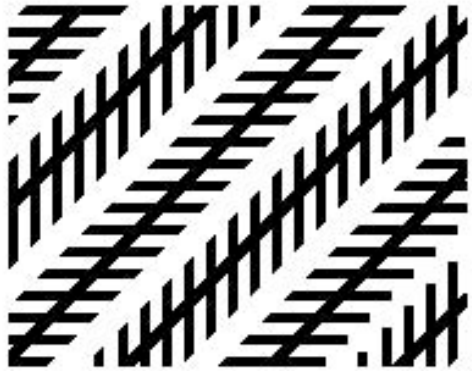


Figure 2. The diagonal lines are parallel, but the crossing lines cause them to appear the diverge due to grouping and inhibition.

Inhibition works on brightness and color values, too. A common visual illusion, shown in **Figure 3**, has two identical grey squares, one surrounded by lighter and one by darker values. People consistently judge that the one with the lighter surroundings is darker than the other. The same thing happens with color. In **Figure 4**, the change in the surroundings makes the color of the test bars appear to vary in an opposite way, in terms of the hue or position around the color wheel. Tests also show that the apparent size of objects is a function of their color and brightness.

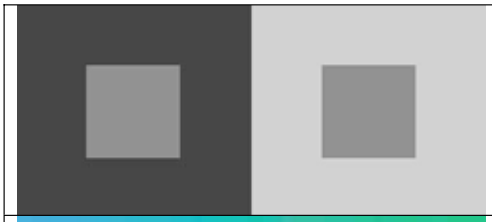


Figure 3. The two central grey patches are identical in brightness, but inhibition due to the surrounding regions makes them appear to be different.

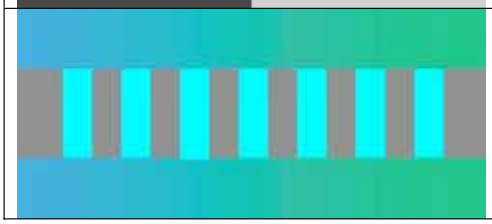


Figure 4. Inhibition due to comparison with surroundings also affects color perception. The cyan bars are all exactly the same.

What consequences does this have for image contrast adjustments? First, it means that the contrast adjustments will depend on the actual contents of the image, that there is no “universal” solution. In fact, there may be different adjustments appropriate for different parts of an image, for example if the left side and the right side for example have different color or brightness ranges (using selections to differently process regions in the image is discussed in **Chapter 4**). That means, in turn, that the adjustment will be user-dependent and what looks best to each person will depend on which features in the image they judge to be most important and therefore which should be given the best visual enhancement. Attention will generally focus on features near the center of the image, rather than ones near the corners and edges.

There are a few common threads that will organize the many possible adjustments and give some structure to the process. First, while the visual contrast of the computer display depends in significant measure on the ambient light levels, there is almost never any reason to use less than the full range of the display for the picture information. It is true that the dynamic range, or gamut, of the computer display is greater than the gamut of the printer, and it may be difficult to reproduce in hard copy the full range of colors from the screen display, but that problem is deferred to **Chapter 5**. If the original picture is of low contrast and does not cover the full range of values that could be recorded by the camera, we will certainly want to expand the contrast by setting dark and light limits for the display.

The image histogram

Setting limits on the darkest and lightest values in the image brings us to consideration of the image histogram. All programs, and even many of the higher end cameras, display a graphical plot of the number of pixels (or, equivalently, the area of the image) that has each possible value of brightness. For a typical 8 bit image this is 256 brightness levels, so we have graphs like those in **Figure 5**. Selecting Image -> Histogram in Photoshop displays these graphs.

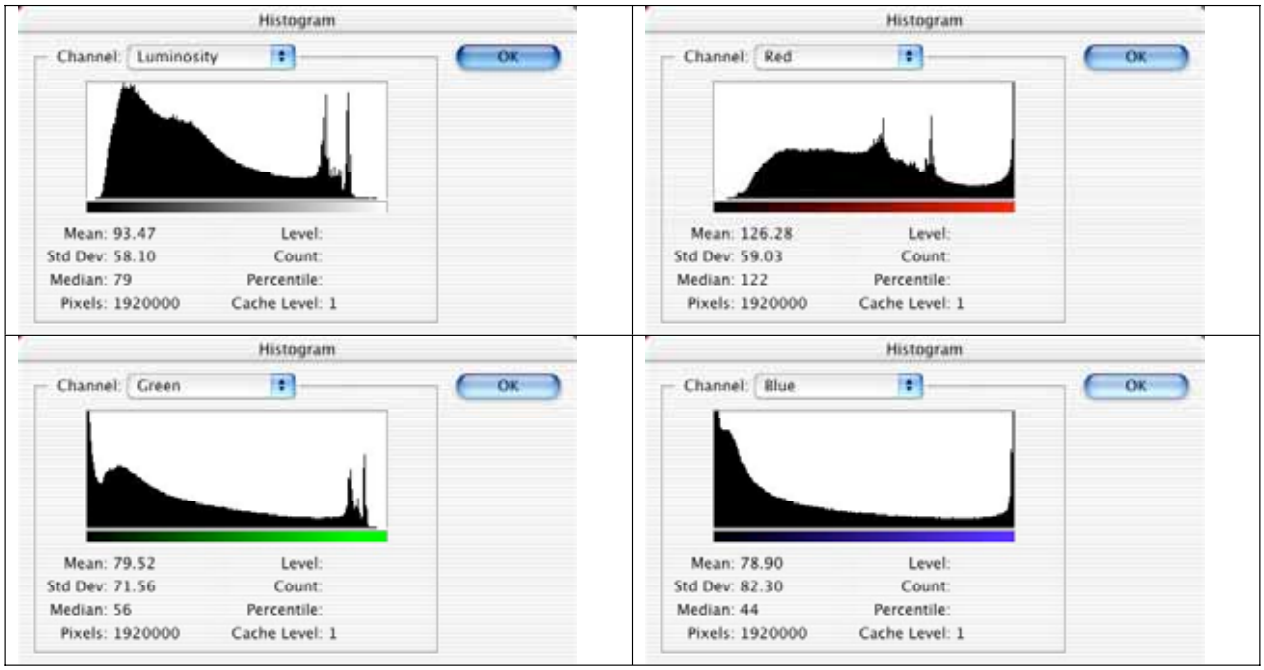


Figure 5. A color image (a) with its histograms: b) the intensity or luminosity channel; c) the red channel; d) the green channel; e) the blue channel.

Plots can be shown for the red, green and blue channels that store the actual information, but at present we just want the overall brightness, which is usually calculated as the average of R, G and B. If the intent is to more closely mimic the actual color response of the human eye, a weighted average of the general form Intensity = 0.25 * Red + 0.65 * Green + 0.1 * Blue may be used. The actual numbers vary somewhat with overall brightness, and probably also from individual to individual (e.g., with age) but that isn't important for our purposes here.

Many of the illustrations in this chapter will use the image histogram as a tool for making various adjustment settings, and for judging the contrast in the original image. The other graph that will be used is the transfer function, which Photoshop calls Adjustment Curves. This is a plot relating the stored brightness value (either for intensity or for one specific color) on the horizontal axis to the value that is displayed on the computer screen and, later on, sent to the printer, on the vertical axis.

The transfer function may be implemented in various programs either by modifying the stored contents of the image, or by a "lookup table" that tells the display what value to use for each of the possible stored values. Most computer display controllers allow writing a lookup table of 256 values for red, green and blue that modifies the display without altering the stored pixel values. Of course, this alters the entire display and not just one image in its window, so Photoshop instead implements the transfer function internally and substitutes the modified values before the data are sent to the display. This allows making adjustments to the display hardware to achieve color matching, discussed in **Chapter 5**.

Photoshop also has the ability to apply curves and other contrast modifications as an “Adjustment Layer,” which alters the display and can be edited as desired, but leaves the underlying data unchanged. It is always desirable to use an adjustment layer to make color and contrast adjustments, but this can only be done for 8 bit per channel images, and in many cases it is best to carry out the image processing steps in 16 bit per channel mode, only reducing the image to 8 bits for printing at the very end. This is obviously true if the image was initially acquired from a camera or scanner that produced more than 8 bits, in order to preserve the original information. It is also advantageous when the image was originally only 8 bits, because many of the operations compress a portion of the data; in order to preserve precision the image can be converted to 16 bits per channel for the intermediate processing steps.

The dialogs for creating adjustment layers for levels and curves are identical to those for the same functions that alter the stored data. The latter are selected as Image -> Adjustments -> Levels (or Curves, etc.), and the latter as Layer -> New Adjustment Layer -> Levels (or Curves, etc.).

If the image has low overall contrast, the histogram of brightness values will not cover the full range from black to white (0 to 255). Setting limits at the ends of the actual histogram (**Figure 6**) is one way to instruct the display hardware to substitute values using a transfer function like that shown in the figure. Sometimes the adjustment is instead made by increasing the gain (the slope of the transfer function) and the brightness (the point at which the curve starts to rise), but this is usually more difficult to control (the Adjustments -> Brightness/Contrast dialog is probably the least useful means for trying to control image appearance). When limits are set on the histogram, the displayed values then vary linearly from black for whatever the actual darkest pixel is to white for whatever the actual brightest pixel is.

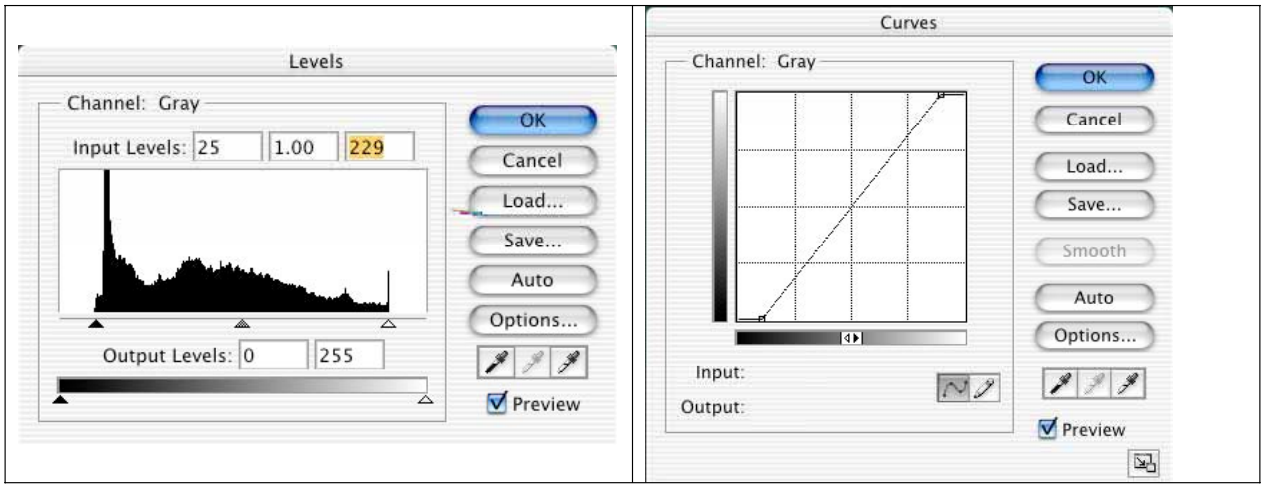


Figure 6. Adjusting the bright and dark limits to linearly expand image contrast: a) original image; b) resulting image; c) image histogram as shown in the Image -> Adjustments -> Levels dialog, with bright and dark limit markers positioned at the ends of the histogram; d) the transfer function shown in the Image -> Adjustments -> Curves dialog, with the same end points.

Color spaces

For a color image, this adjustment should be made only to the brightness and must not alter the colors (unless you also want to change them), so typically the adjustment is carried out not in the RGB space of the stored values (and, for that matter, the space used by the display, and by many cameras and scanners), but in a hue-saturation-intensity space like HSL, HSB or L*a*b. The casual user of image processing software may wish to skip this discussion, but understanding the different ways that color can be represented does help to understand and master the adjustment and processing tools.

The RGB color space is mathematically very simple, with three perpendicular axes and a cubic shape. **Figure 7** shows how combinations of red, green and blue specify a color. Points along the diagonal of the cube, with equal amounts of all three colors, are neutral grey and range from black (R=0, G=0, B=0, the bottom corner of the cube) to white (R=255, G=255, B=255). Since the complementary colors to red, green and blue are cyan, magenta and yellow, respectively, this same cubic space can be used to describe the CMY color space, also shown in the figure.

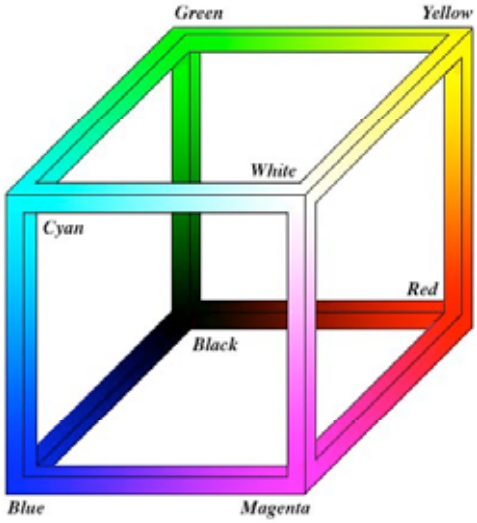


Figure 7. The Red-Green-Blue color cube, showing the major perpendicular axes and the addition of the primary RGB colors to produce cyan, yellow and magenta. The diagonal from the black corner to the white corner has equal components of all three colors to represent grey values.

As indicated in **Figure 8**, RGB space is an additive space. For example, in a typical CRT display the addition of various intensities from those colored phosphors to the black background produces the various colors. The greater the color intensities, the brighter the combined result. CMY (cyan-magenta-yellow) is a subtractive space. For example, in typical printing the white paper provides a background from which the various inks reduce the complementary color (e.g., magenta ink reduces the amount of green light reflected). The maximum amount of ink produces the darkest color, although because it is difficult to get a true black with only these three inks most color printers also use a black ink (the K in CMYK). Printing is covered in **Chapter 5**.

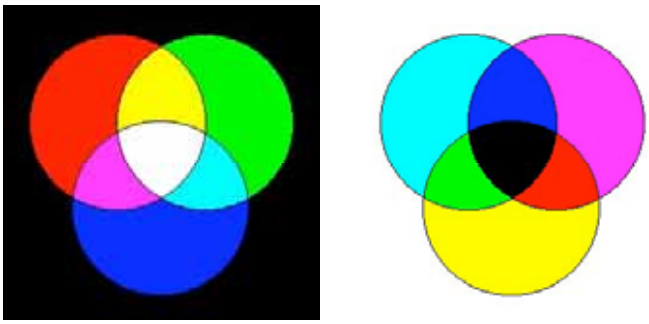


Figure 8. RGB colors add to a black background to produce white (a). CMY colors subtract from a white background to produce black (b).

Most people do not instinctively see colors as mixtures of red, green and blue. Looking at an orange-pink rose does not translate into the amount these three primary colors, but they are all present. There are certainly a few people, experts in prepress image preparation, who know that C=10, M=35, Y=45 is a skin tone, but it isn't something that most photographers care (or need) to learn.

Hue, saturation and intensity

The natural way that artists (and children) relate to color is based on the color wheel shown in **Figure 9**. Colors vary from red through orange, yellow, green, cyan, blue, and magenta with angle. The technical name for this angle is hue. The radius is the amount of color present, for instance the difference between pink and red or sky blue and ocean blue. In the part of the country where I live, this is the difference between “Carolina Blue” and “Duke Blue” - very important during basketball season. Called saturation, this radius ranges from zero for grey scale values to maximum for a fully saturated or “pure” color.

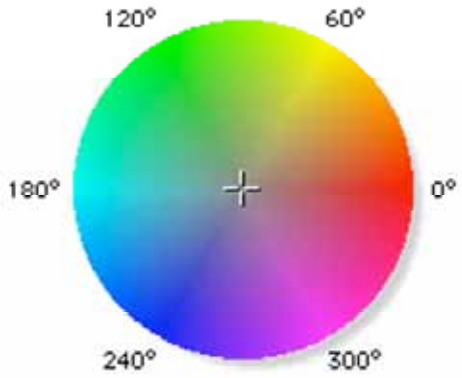


Figure 9. The “color wheel” in which angle represents hue and radius represents saturation.

The hue and saturation define the color information, but third parameter is needed along with them to describe brightness or intensity. It is measured perpendicular to the plane in which the color wheel lies. Depending on the details of the color space this may be called luminance (HSL) or value (HSV) or brightness (HSB), but collectively this book will refer to these color models as hue-saturation-intensity (HSI). The space may be visualized as a cylinder, cone, or bi-cone. The latter, shown in **Figure 10**, is probably the most meaningful geometry. It has a central axis that is the grey scale axis, running from black at one vertex to white at the other. The tapering of the space toward both ends corresponds to the fact that very dark values cannot have very much color added to them (and hence have a small radius or saturation) while very bright values are similarly limited. This is because a fully white spot on a display monitor is produced by turning all three RGB phosphors on fully, and any introduction of color must be accomplished by reducing one or more of the three phosphors, which reduces the intensity.

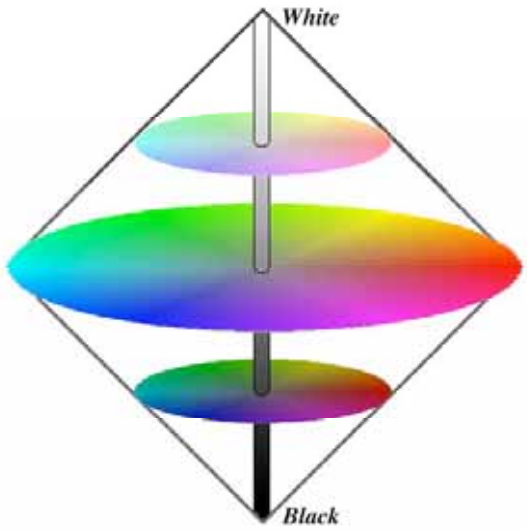


Figure 10. The bi-conic Hue-Saturation-Intensity color space.

The bi-conic HSI space is a useful way to visualize color. Grey scale or neutral values without color lie along the central axis, ranging from black to white. Saturation adds color to the grey values, and can be greatest for mid-range brightness values. The hue, which corresponds to the “color” that people recognize, is the angular direction, which is traditionally measured with a zero angle corresponding to red. The relationship between this space and the RGB cube can be seen by aligning the axis of the bi-cone with the diagonal of the cube. The “black” and “white” corners of the cube are the vertices of the cones. Of course, the two spaces have slightly different shapes. This distortion is not important, because the relative position of one color to another stays the same (it is brighter, or greener, or more saturated, etc.) and it is only the relative colors and not absolute values that interest us here.

It is possible to adjust images based on the hue, saturation and intensity values, although that rarely produces optimum results. The example in **Figure 11** shows an original image that is rather flat in appearance, with colors that are not highly saturated and poor contrast. The red color is somewhat to the blue or purple side of true red on the color wheel. Using the Image -> Adjustments -> Hue/Saturation dialog, the sliders were adjusted so that the red in the flag became exactly red (hue of zero in the bright portions of the stripes), the saturation was increased and the overall lightness was increased slightly.

Figure 11. Color adjustments: a) original image; b) adjustments to hue, saturation and lightness; c) adjustment curves applied to the RGB channels.

Note in the process of adjusting the hue and saturation for the flag, that the colors of the dark foliage and the light sky have been somewhat altered, the foliage becoming a bit too yellow and the sky a bit too cyan, as a consequence of the global hue adjustment. As shown in the figure, it is possible to make a better correction, which does not alter the background colors but still increases the contrast for the flag. This was accomplished using the adjustment curves for each channel. It should be noted, however, that the highly saturated red and green colors in this image as it appears on the computer monitor cannot be printed, as they exceed the gamut possible with a CMYK printer. Gamut problems are discussed in **Chapter 5**.

In broadcast television, where attempting to control the actual colors is important (if not always successful), still another variant of HSI space is used. The CIE (Commission Internationale de L'Éclairage) color plane is not a circle, but instead has a shape in which colors from various phosphors add along straight lines. There is also a modification of that which makes equal distances represent equal ability to discriminate color differences. It is not necessary for the photographer to be familiar with these technical issues.

But while the HSI bi-cone is a convenient space to represent perceptual color, it is a difficult one mathematically. The angle jumps from 360 to zero degrees at red, so it isn't immediately apparent that 10 degrees (a slightly orange red) and 350 degrees (a slightly magenta red) are actually very close

colors. Certainly the average of these two values should not be (10 + 350) / 2 = 180 degrees (blue-green). Also, the variation of maximum possible saturation with intensity adds complexity.

One of the common modifications of this space that simplifies the math is to consider the space as a sphere with the north-south axis corresponding to the grey scale intensity axis. This fits the bi-cone and the cube well enough. Instead of using polar coordinates (hue and saturation) the L*a*b model places conventional perpendicular axes at the center of the sphere, with one axis running from green to red and the other from blue to yellow as shown in **Figure 12**. This slightly distorts the color wheel, which in HSI space has cyan opposite red, but stretching a portion of the wheel is a minor detail since the order of the colors stays the same. Many programs including Photoshop use this model for internal calculations involving color.

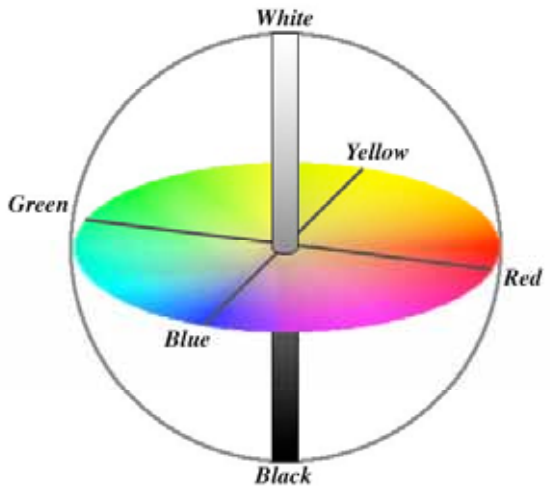


Figure 12. Spherical L*a*b color space. The a and b axes are red-green and blue-yellow.

Converting Color to Grey Scale

The various color models have all described an intensity axis running from black to white through various greys with no color. Converting a color image to grey scale may be required for printing or display, or to process the grey scale image to use as a mask (discussed in **Chapter 4**). There are many possible ways to combine the color values.

For display on the computer monitor, averaging the red, green and blue together gives a pure grey value with no visible color tint, but the resulting grey scale image is very different in contrast from what the human eye sees, or what a monochrome film camera would record from the same scene. Because the human eye is quite sensitive to green, less so to red, and even less to blue, a weighting something like the equation given above is often used:

Intensity = 0.25 * Red + 0.65 * Green + 0.1 * Blue

But there is no reason to limit the conversion to any particular equation, if the purpose is to produce a grey scale representation with visually interesting yet realistic contrast. The image in **Figure 13** shows a typical problem situation. The background is mostly green, and the flower blue. Simply converting the image to grey scale using the built-in Photoshop Image -> Mode -> Grey scale menu selection is identical to converting the image to Lab color space and keeping just the lightness channel. As shown in the example, this doesn't produce a very interesting image: the flower petals can hardly be distinguished from the leaves.

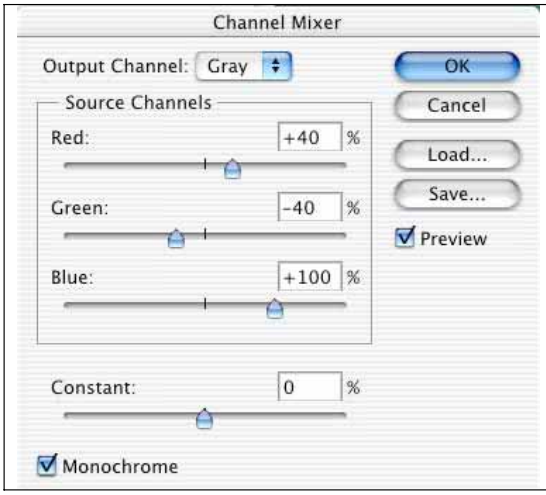


Figure 13. Converting color to grey scale: a) original; b) L channel from Lab; c) average (reducing saturation to zero in RGB space); d) channel mixer result; e) channel mixer dialog.

Using the Image -> Adjustments -> Hue & Saturation function and reducing the saturation to zero produces a slightly better result, since there is now some contrast between the flower and the leaves. This indicates that the RGB color space in which this operation takes place weights the green less and the blue more, which is consistent with the simple averaging of red, green and blue.

The most versatile method of combining the color values is to use the Layer -> New Adjustment Layer -> Channel Mixer routine. This allows the amount of each channel to be set arbitrarily to produce a result with the desired contrast. In the example using a negative amount of green (subtracting the green value from the sum) and a positive one for blue produces a grey scale image in which the flowers have definite contrast with the leaves. Other combinations would reverse the contrast.

Limits and Clipping

The simplest modification that can be made to an image with low contrast is to create a transfer function that starts at the lowest brightness value actually present in the picture and ends at the highest, so that the image contrast is stretched out uniformly (linearly) over the dynamic range of the display. But in most images there are highlights and shadows that really do not contain any important or perceptible detail.

Specular reflections from surfaces, for example, are simply bright points of light that tell our visual system something about the curvature of the surface and about the light source, but have no internal detail. Dark shadow areas also give shape to the structures in front of them, but we don't expect to see inside them. Furthermore, many cameras have a few "dead" or "locked" transistors, as discussed in **Chapter 1**, that have either maximum or zero output, and these create small points (not necessarily limited to a single stored pixel) of bright or dark values in the picture.

It makes sense to set the bright and dark limits for the display to clip these extreme bright or dark values. When clipping occurs, the original extreme bright and dark pixels are simply set to the maximum or minimum display values. This can be done manually, using the histogram as a guide, or in many cases an automatic method that clips a certain percentage of the pixels in the image to white or black can be

used. The percentage may vary from as little as 0.05% to as much as 0.25%, depending on personal taste and the nature of the image. Different values may also be used for the bright and dark ends of the histogram. To specify the amount of clipping, click on the Options button in either the Adjustments->Levels or Adjustments->Curves dialogs (**Figure 14**). The Auto button (or the Adjustments -> Auto Levels menu selection) performs this operation using whatever clipping percentages have been established.

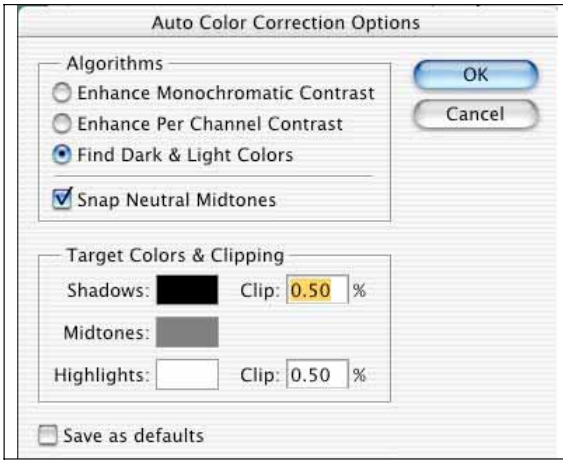


Figure 14. Setting the clipping levels in the Options dialog.

It is important to understand that setting limits and creating a transfer function is something that must be done on the intensity values in the image, not the individual stored red, green and blue values. If the limits are set on the individual color channels, the result is to alter the ratios of the different primary colors. This in turn changes the hues in the image, and typically adds color to neutral grey regions. The latter effect is particularly a problem, since maintaining neutral tones in features and areas that should not have color is one of the most important steps in the color correction discussed below.

In terms of the color space discussion above, the process of setting limits is to construct a histogram of the pixel brightness (or luminance, or intensity) values, to set the limits on that histogram either manually or automatically, and then to apply that transformation to the individual pixels by converting their values from the stored R, G, B space to an HSI space (Lab color is often used), adjust the intensity values, and then convert back to R, G, B so that the computer display can properly show the image. This is very different from the common mistake of setting limits to expand the range of contrast on the individual R, G, B channels, which can produce serious color shifts when the histograms for the individual channels are different. Instead, the procedure of working on the Luminance or intensity channel alters the brightness of the pixels but retains the original color information as shown in **Figure 15**. The process is imperfect because human vision interprets colors somewhat differently in bright and dark images or regions of images, but is the best general approach before addressing the more advanced techniques shown below.

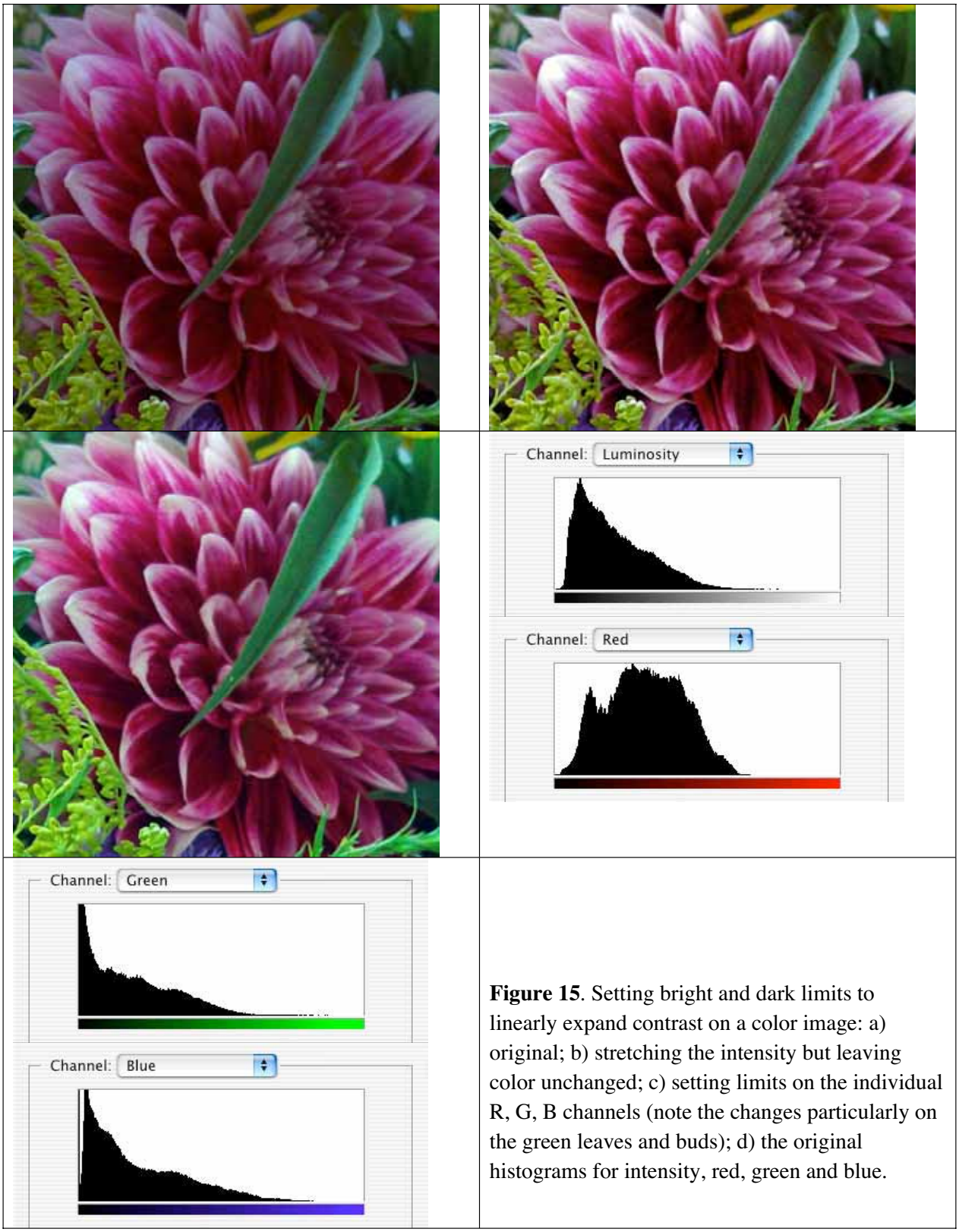


Figure 15. Setting bright and dark limits to linearly expand contrast on a color image: a) original; b) stretching the intensity but leaving color unchanged; c) setting limits on the individual R, G, B channels (note the changes particularly on the green leaves and buds); d) the original histograms for intensity, red, green and blue.

In **Figure 15**, the contrast expansion obtained by setting the limits on each color channel individually (selecting "Enhance Per Channel Contrast" in the Options dialog for Adjustments -> Curves or Adjustments -> Levels, and the method used if Adjustments -> Auto Levels is selected) produces color shifts and blows out some of the white areas. Setting the limits instead on the intensity channel (select "Enhance Monochromatic Contrast" in the same dialog) does a better job of preserving the color information.

Some cameras perform this type of limit setting, stretching and clipping the data internally and automatically. Most digital cameras store images with 8 bits each for red, green and blue. Many of the higher-end models actually have analog-to-digital converters with more than 8 bits of precision. This allows the built-in camera software to find the actual dark and bright limits in the image and scale the output values between those points, which produces an image whose histogram covers the full range of the 8 bits used for storage. However, if you are not using automatic exposure settings, or want to clip bright or dark regions, further expansion of contrast in the computer may still be appropriate.

Gamma (mid grey adjustment)

The stretching shown in the preceding example was linear, but there is no particular reason for this restriction. The simplest departure from linearity is the introduction of a gamma adjustment. For film photographers, this term is familiar and describes the choice of film or printing papers of various "hardness" values, which is the slope of the density vs. exposure response curve. The meaning is the same for digital image processing, but the appearance of the curves is different.

This arises because photographic materials (like the human visual system) respond to light with a logarithmic response. For the non-mathematically inclined, this means that the change in response (density for film, ability to detect a brightness difference for vision) is proportional to the ratio of intensities. Human vision can detect, under optimum viewing conditions, a change in intensity of a few percent. In the dark portions of a scene that is a smaller absolute change than in a bright region. Changing gamma shows up in the transfer function plot as a family of curves, as indicated in **Figure 16.**

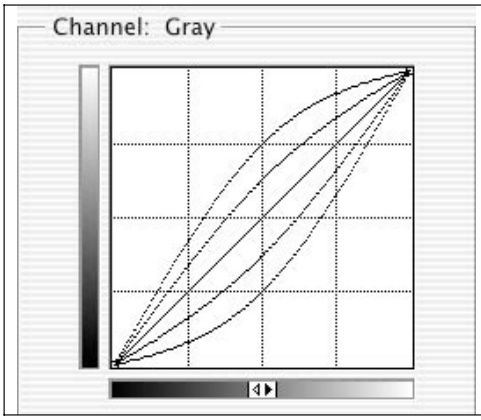


Figure 16. Transfer functions showing various gamma settings. The straight line is gamma = 1.0; moving the midpoint down expands the range for bright values and corresponds to gamma > 1, and vice versa.

In Photoshop, the gamma setting can be made by moving the mid-grey marker on the histogram display. As shown in **Figure 17**, shifting the marker toward darker values compresses the range available for them while expanding it for the light values and produces gamma values greater than 1. Conversely, shifting the marker toward lighter values expands the range for showing the darker pixels (making small differences more visible) and corresponds to gamma values less than 1. Automatic setting of gamma can sometimes be done well by positioning the mid-grey slider to a brightness value such that 50% of the pixels are brighter and 50% darker. The Photoshop Histogram display shows this Median brightness value.

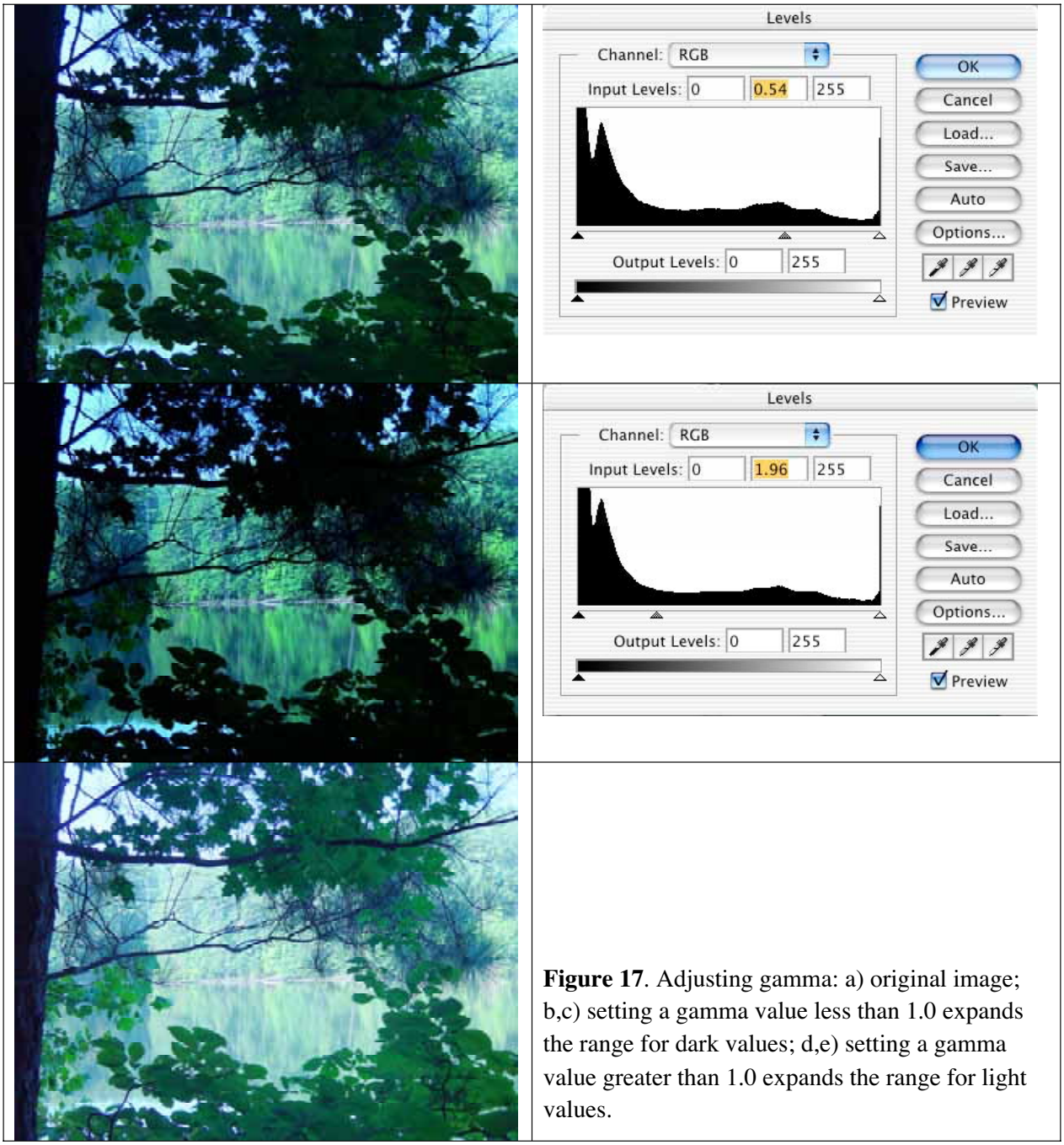
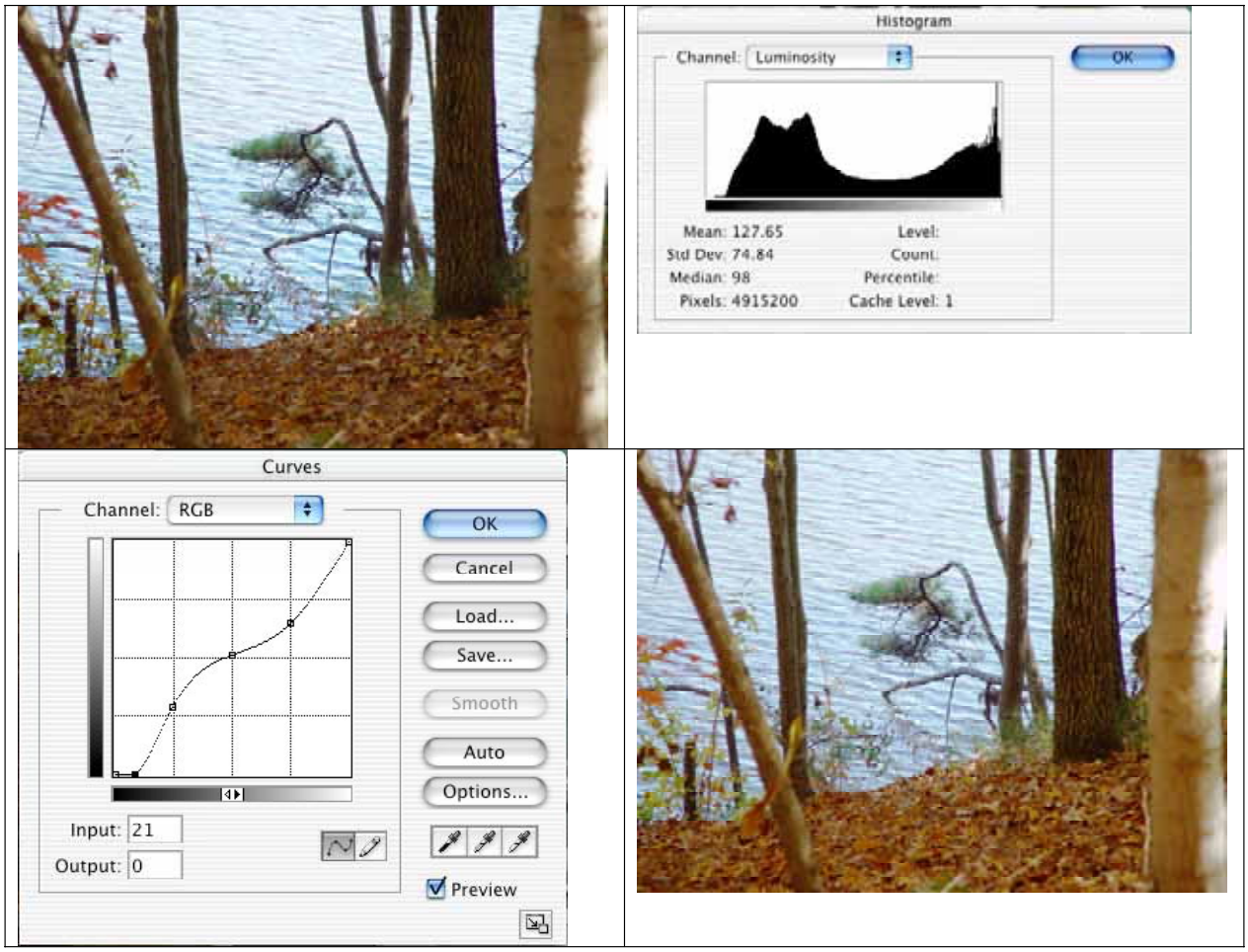


Figure 17. Adjusting gamma: a) original image; b,c) setting a gamma value less than 1.0 expands the range for dark values; d,e) setting a gamma value greater than 1.0 expands the range for light values.

As for setting limits, gamma adjustments should be made on the intensity channel rather than the individual red, green and blue channels in order to avoid altering the colors in the image.

Curves

Adjusting gamma by positioning the mid-grey slider allows adding contrast to the dark or light tones in an image, but even more control is possible by manipulating the transfer function curve directly. The goal is to increase contrast in whichever range of tones requires it (usually those tones that predominate in the image, which means that they produce a peak in the histogram). This is accomplished by compressing the contrast for other tonal ranges. The transfer function has steep sections where the contrast is being expanded and flat sections where it is being compressed, as shown in **Figure 18**.



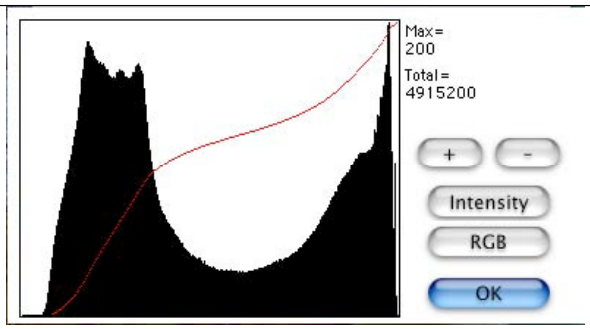


Figure 18. Adjusting curves to expand dark and light areas: a) original image; b) histogram, showing peaks at both ends; c) adjustment curve; d) resulting image; e) detail showing increased contrast range for both bright and dark regions (top: original; bottom: expanded contrast); f) histogram with superimposed cumulative (integral) curve in red.

Note in this example that control points have been used to shape the transfer function curve, placed at about the 25% and 75% points on the brightness scale (often called the "quartertones") along with the 50% points (the "midtone"), with another point at about 10% to clip off the dark range that has few pixels, as evident from the histogram. It is also useful to observe that the shape of the transfer function is similar to that of the histogram when it is plotted in a slightly different way. Instead of the conventional plot showing the number of pixels as a function of brightness, the cumulative histogram shows the fraction of the image that is darker than each brightness level (mathematically, for those so inclined, this is the sum or integral of the conventional histogram).

Because of this relationship between the shape of the cumulative histogram and the transfer function, it is sometimes useful to use the histogram to perform the operation automatically. This is called histogram equalization, because it tries to shift brightness values up or down to make equal areas of the image have each possible brightness level. While histogram equalization rarely produces the optimum setting, it is often close enough that it gives a useful guide to making manual adjustments.

Figure 19 shows the opposite situation, an image whose histogram has a broad, single peak near the center of the brightness range, but still has some pixels at the ends. We want to increase the contrast for the central region but not by simply clipping the ends of the histogram by setting limits that erase all detail in the bright and dark regions by setting those pixels to maximum and minimum brightness. In this case, the transfer function is adjusted to bend the other way, making the ends flatter and the center steeper.

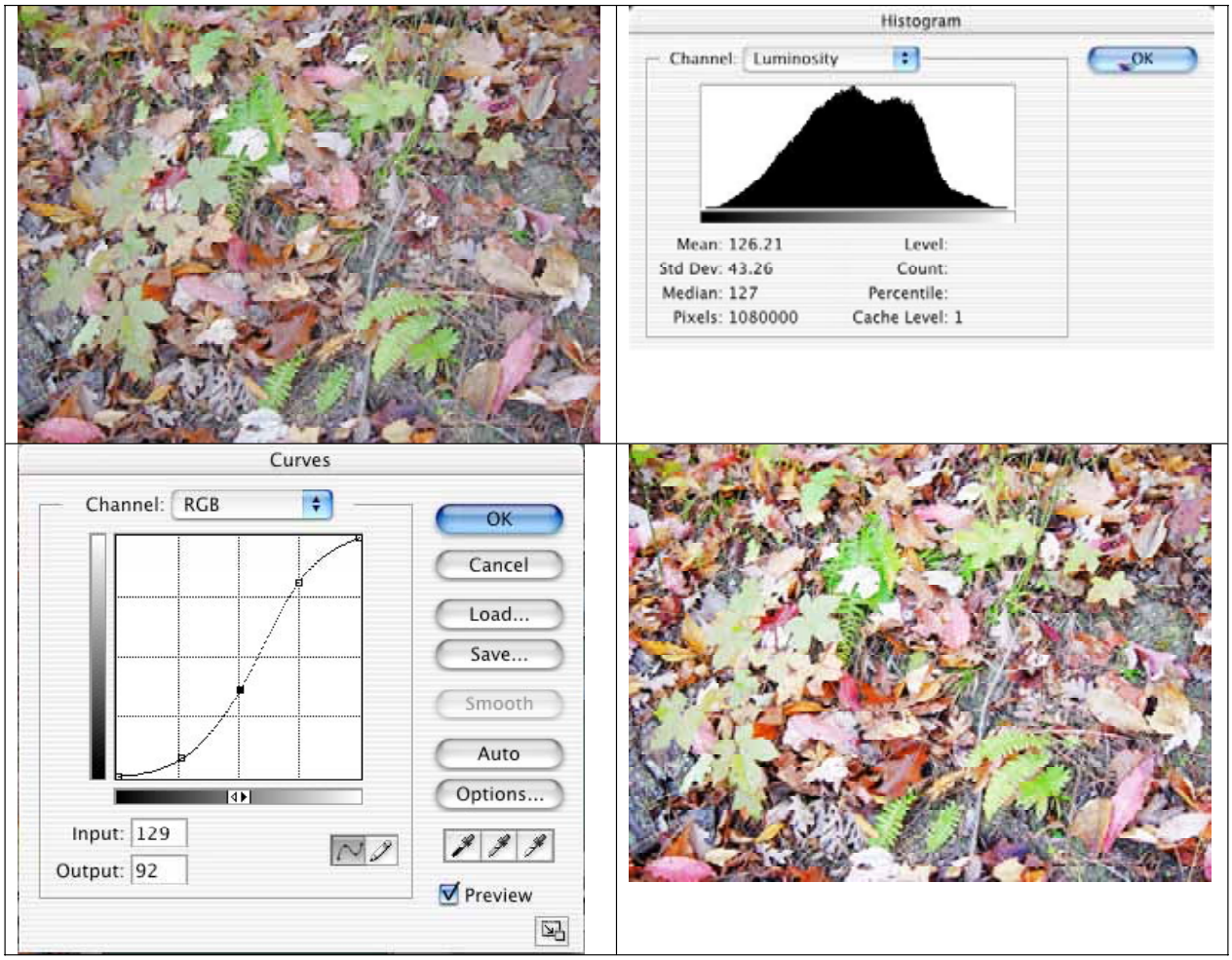


Figure 19. Adjusting curves to expand the midrange: a) original image; b) histogram showing a central peak with tails to both ends; c) adjustment curve; d) resulting image.

However, it is also evident in **Figure 19** that the contrast expansion performed in the intensity channel has made more evident a problem with the color balance in the image. Compare this to **Figure 29** to see that much better results can be obtained by applying adjustment curves to the individual color channels.

Figure 20 shows another example. The original image is somewhat too dark, as is very evident from the histogram. Taking good pictures of black dogs and cats is always difficult because there isn't much in the way of a grey scale. Black fur is a really good absorber of light, and all you get is nearly uniform black with a few really bright highlights. Simply moving the bright limit to the topmost point in the intensity histogram creates a background in which the green grass is much too bright and too saturated. This can be corrected, as shown, by adjusting curves for the individual color channels. So the next topic to be taken up must be making adjustments to the transfer curves for each color channel separately, in order to correct color problems.

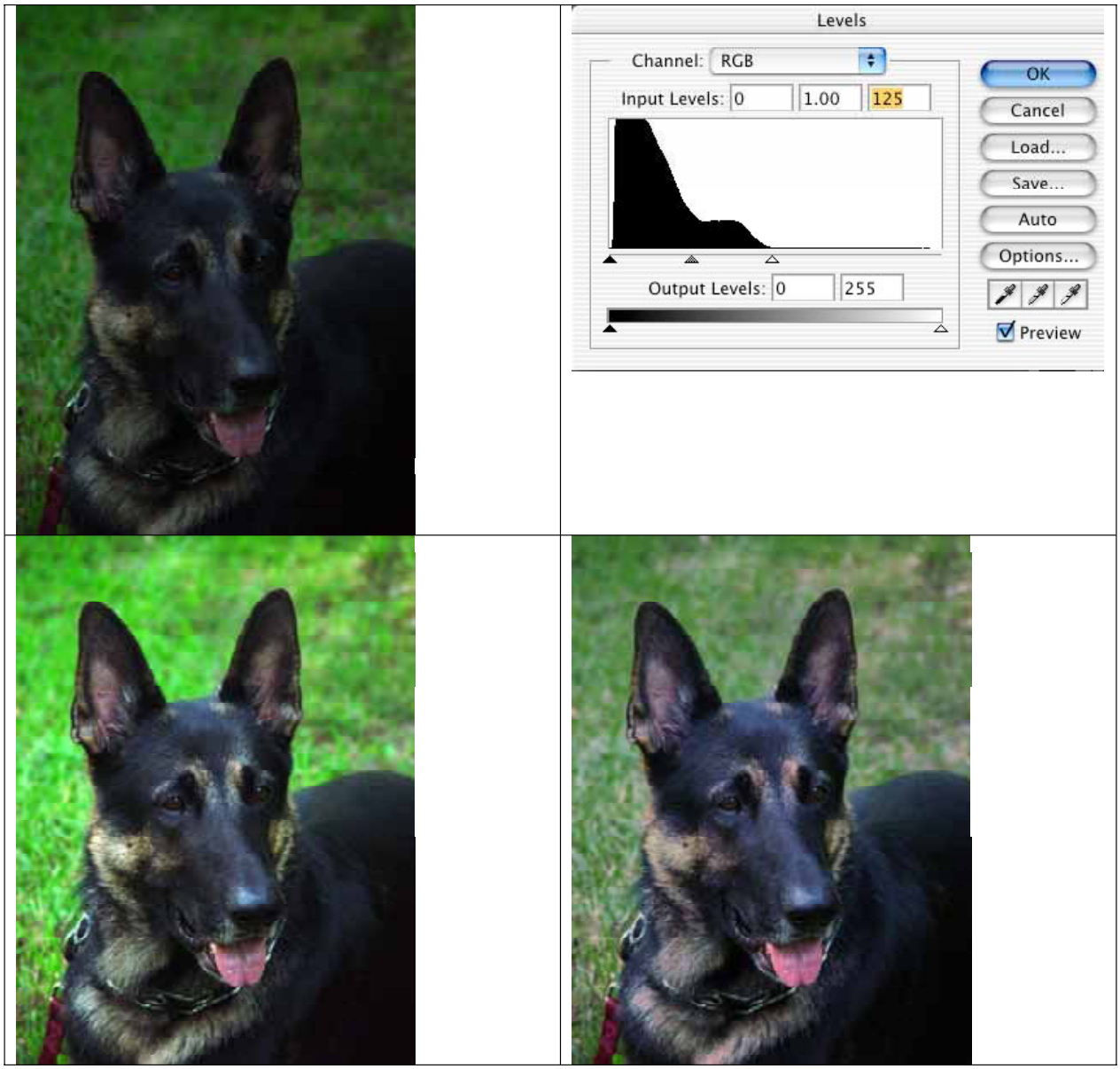


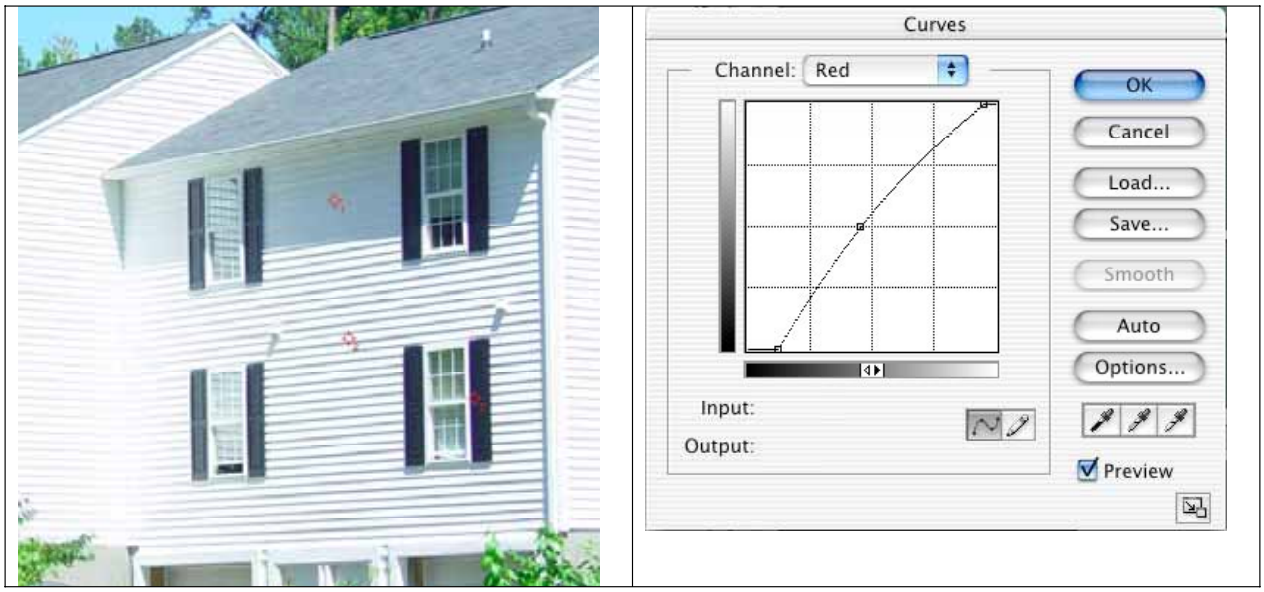
Figure 20. Color shift from contrast expansion: a) original image; b) histogram with limits; c) resulting image (note green background); d) contrast expansion without false green background resulting from curve adjustments in individual color channels.

Color adjustments

The curve adjustment approach shown above for intensity is, of course, also applicable to the individual red, green and blue channels. And the same basic principles apply: a steep curve will increase contrast in one portion of the intensity range, while a shallow curve can reduce it in another. The problem is that people just don't see color intuitively as a mixture of red, green and blue.

It is very important to keep in mind what the goal of the contrast adjustment on color images should be. We are not trying to get “true colors.” In fact ,as we will consider shortly, that is a much more difficult challenge that involves the illumination, the nature of the colored surfaces reflecting that illumination, and the color filtering used in the camera. Instead, what we want is to increase the visual contrast in the scene and the ability to discriminate colors *without presenting the viewer with obviously implausible colors* that would disturb the comprehension of the image.

Making the curves adjustment while monitoring the colors of a few selected points in the image is one good way to accomplish that goal. The very simplest situation arises when the image contains neutral grey tones that represent a dark grey, medium grey and light grey. In **Figure 21**, three monitoring points have been established in Photoshop on such regions. This was done by selecting the “Color sampler tool” with a 5x5 pixel average, and clicking on appropriate points in the image. Adjusting the red, green and blue channel curves is done by moving the control points that approximately correspond to those brightness levels up or down while watching the displayed RGB color values at the selected control points, so that the amount of red, green and blue in each of them is equal, thus giving a neutral grey color.



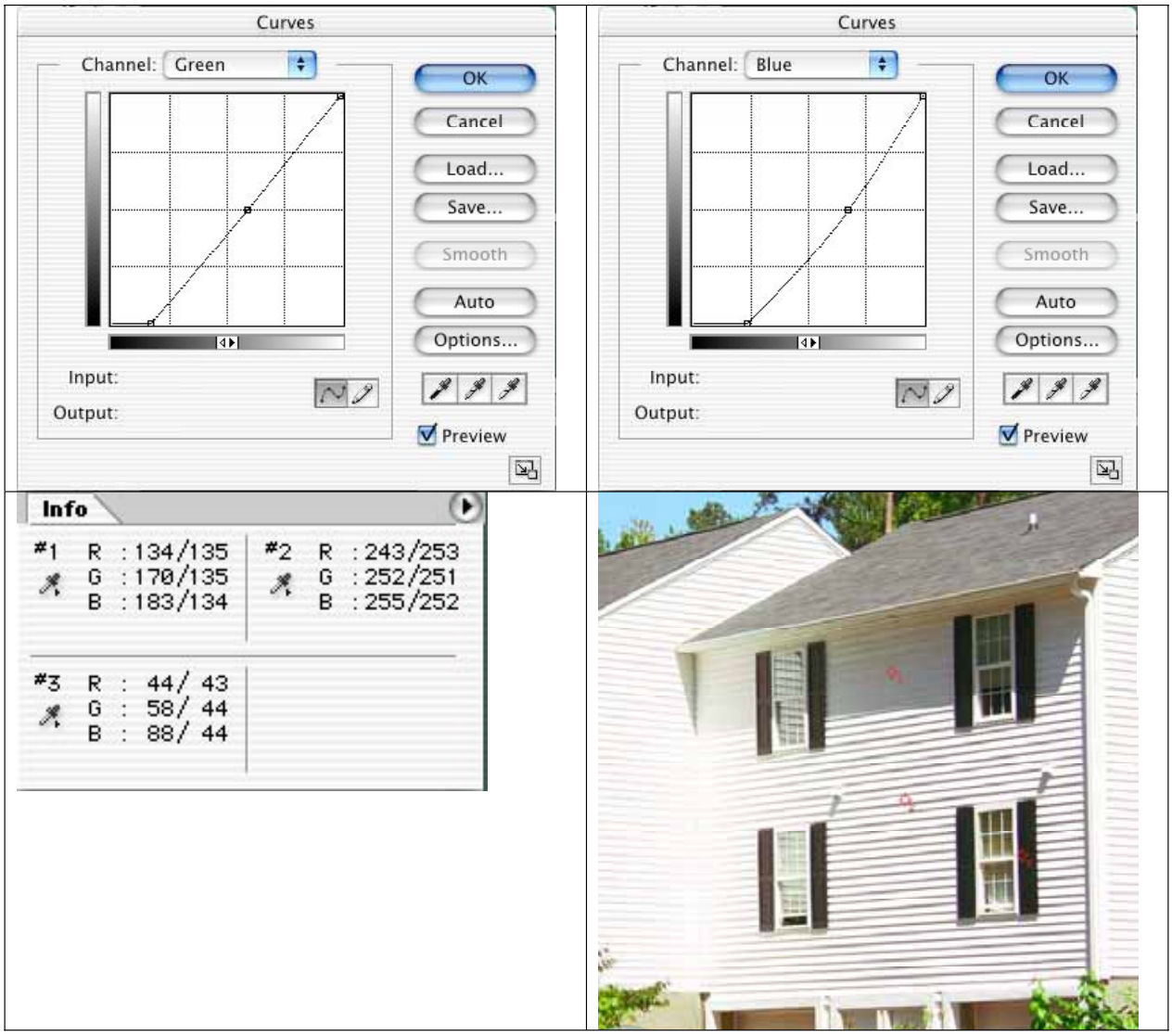


Figure 21. Using neutral colors to adjust color channel curves: a) original image with three monitor points set; b, c, d) adjustment curves for the red, green and blue channels; e) Info display showing the original and adjusted RGB values at each location; f) the final adjusted image.

A useful short cut to this manipulation is to use the three eyedroppers in the dialog to click on points in the image that represent bright, medium grey, and dark neutral tones (the same three points marked on the images were used). The program then performs the adjustment for you, making the red, green and blue values equal at the selected locations, and displaying the resulting curves.

Photoshop also has “automatic” tools that try to accomplish the same result as that shown. In the Adjustments -> Levels and Adjustments -> Curves dialogs, clicking on the “Options button” opens another dialog window (shown in **Figure 14**, above) in which you can select “Find Dark and Light

Colors" and "Snap Neutral Midtones." (This is also the dialog mentioned above in which you select the extent of clipping at the bright and dark ends of the histogram, for setting levels). Working with three points that it finds - bright, mid-grey, and dark - the program adjusts the red, green and blue curves attempting to make the neutral colors into true greys.

The use of three neutral points (dark, medium grey and bright) is often useful as a starting point for further manual manipulation, since it does shift the curves in the right direction. But with only a single mid-grey point it is unable to achieve the much more complete correction that the use of multiple points accomplishes.

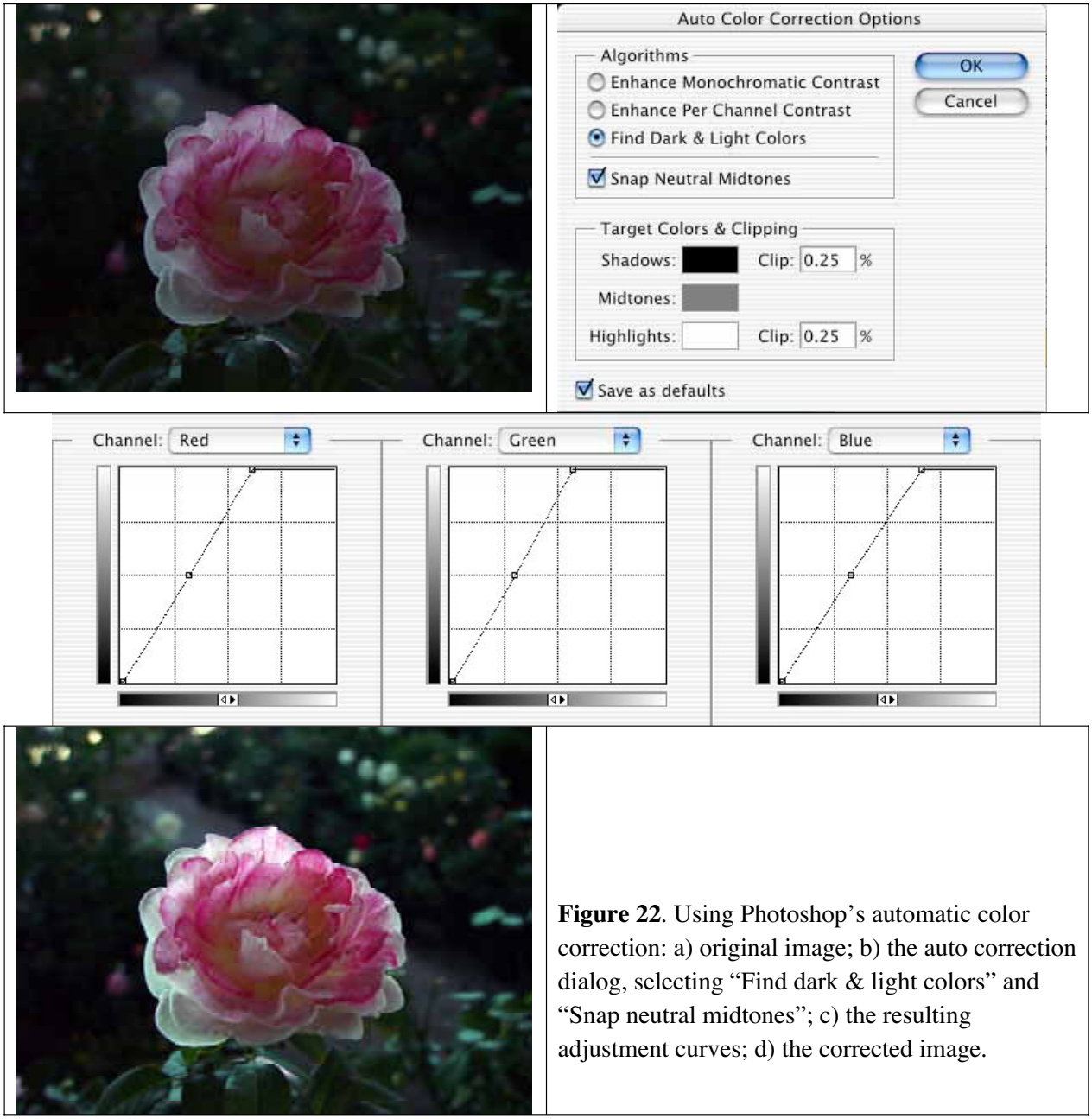


Figure 22. Using Photoshop's automatic color correction: a) original image; b) the auto correction dialog, selecting "Find dark & light colors" and "Snap neutral midtones"; c) the resulting adjustment curves; d) the corrected image.

Sometimes the simple, automatic method is enough. The flower shown in **Figure 22** has white areas in the petals as well as a pink (between magenta and red in hue, not fully saturated) tone. Using the automatic dialog to allow clipping for just a few pixels (there are a few bright and dark spots in the background that can be clipped without harm), the program sets the maximum brightness point for the red, green and blue values such that the brightest petals are exactly white (and the darkest shadows exactly black). The result is quite good, and no further adjustment of midtones is needed. This is exactly the same correction that some cameras allow by selecting a point in the image as defining "white." This automatic white point correction drags all other colors and shades along with it, which is better than nothing but not always an adequate substitute for curve adjustment.

Selecting control points

Not all apparently neutral colors really are, of course. An elderly lady's grey hair may have a tinge of blue in it, while the black hair of a younger girl is likely to contain some red. With the exception of dyed hair, such as platinum blonde, blueish grey, or henna, no natural hair color will ever shift toward the blue or magenta, however. When it is appropriate, such adjustments can be made easily by allowing the corresponding color to be a few percent greater than the others in the adjusted mixture.

Watching the numbers at the control points is important, because it can be visually very difficult to determine just what is a neutral grey color. One of the characteristics of human vision, mentioned above, is that of inhibition. When placed next to a bright region, a feature will appear darker than when it is next to a darker region. Similarly, the tone of a neutral region will shift away from the surroundings. If a medium grey tree trunk is surrounded by green leaves, it will appear to take on a reddish cast. The same shade of grey surrounded by orange leaves in the fall would appear bluish.

Sometimes control regions are selected that are not really perfectly neutral, but the image will visually appear more realistic if they are set to be so. For example, a white shirt or blouse will often have a color cast to it because of the color of the illumination (or because it isn't really white material, but actually has some slight color because of laundering - or lack of it), but our minds "know" that the true color should be white, and by using this as the bright control point for setting the color balance the image will have a better appearance.

Deep shadow areas may seem to be convenient as dark grey control points, but don't usually work so well. First, they probably really do contain some color, but are just too dark for it to be easily seen. Second, increasing the contrast so that they really do become black is probably a better strategy, and as points approach true black there is no color.

Ideally, control points at about 25, 50 and 75% neutral grey (quartertones) would be optimum. This can't always be achieved, but it gives you an idea of what to look for first in the scene. Sometimes you can find them all on the same feature, just selecting points that have different illumination (such as different areas on a white house or grey pavement where there may be shadows).

Other control candidates

Adjustment of the color channel curves using neutral grey points is most conveniently done in terms of R, G, B values because the goal is to make them equal, or nearly so. When colored regions must be used as control points it is generally better to monitor the color values in terms of hue, saturation and intensity.

Photoshop offers a variety of color spaces for reporting color, including CMYK for the pre-press print professionals. The hue-saturation-intensity model used to display color information in the Info dialog is HSB, a color space shaped like a cylinder. The bottom end of the cylinder is black, and increasing the brightness is accompanied by an increasing amount of saturation until the top face of the cylinder is the familiar color wheel (**Figure 23**). The hue value is reported as an angle starting at red (0 degrees) and increasing through yellow (60°), green (120°), cyan (180°), blue (240°) and magenta (300°) back to red. Saturation (the radius in the cylinder) and brightness (position up the axis from the black bottom) are both reported as percentages from 0 to 100%. White is located at the center of the top face of the cylinder. This is one of the standard color spaces for selecting colors on Macintosh computers. With very little practice and examination of colors in representative images, this way of describing color will become very familiar and automatic.

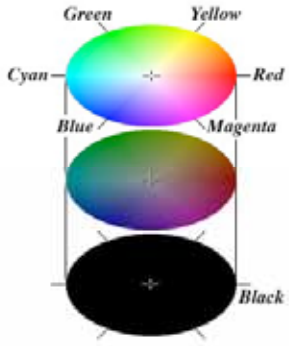


Figure 23. HSB (Hue-Saturation-Brightness) space shown as a cylinder with black at the bottom and fully bright colors at the top.

In many photos, a convenient bright control point will be the sky. Blue sky without clouds has a hue between blue and cyan, typically about 200°, a rather low saturation (30-50%, decreasing toward the horizon) and in many cases can be set to be the brightest part of the image (brightness of 90%). When clouds are present they complicate things because while they are nearly neutral grey, they can vary over a wide range in brightness. The tops of white puffy summer clouds are a good white control point, but the dark bottoms of thunderclouds may be used as a 25% bright (i.e., dark grey) neutral point, and thin cumulus may have a significant amount of blue.

The color of water is typically a little bluer than the sky (by about 10-15 degrees in hue), somewhat more saturation (again decreasing toward the horizon), and brightness values of about 70-80%. You do have to be careful about reflections from water, which may contribute other colors, and about the shadows or whitecaps on waves. Shadows increase the saturation and reduce the brightness, by 10-15%

each. Whitecaps may be used in some cases as a white control point, but usually do have a slight blue or blue-green cast. At least, they should never have a hue between yellow and magenta, unless the water is seriously polluted. Water in a swimming pool often shows the color of the paint on the pool bottom, which can be anything but is often a greenish blue or bluish green.

The numbers given in the preceding paragraphs are useful as starting points but are not gospel. If slightly different values look better to you, then by all means use them (and note them down for future use). Remember always that human vision does not measure color, but compares it to the immediate surroundings. Consequently colors will appear visually different depending on their context - the sky over a forest scene will look different than one over a cityscape, even if the numerical values are the same, so perhaps they should be set differently with the appearance taking priority over the numbers.

Human vision consists of much more than the eyes as cameras. The information sent from the retina to the optical cortex and on to the conscious levels of thought is highly selective (we generally see what we are looking for). And culture plays an important role, as well. The sky and sea colors mentioned above look like sky and sea to me, and hopefully to you (see **Figure 24**). But in the Odyssey, Homer consistently describes them quite differently. The figure shows his “bronze sky over a wine dark sea.” Are we supposed to believe that nature has altered the actual wavelengths of reflected and refracted light over the past few millenia? Or is there clues in the facts that Homer was blind, the ancient Greek language did not even have a word meaning “blue,” and it is common in poetry for words to describe an emotional response rather than the literal actuality?

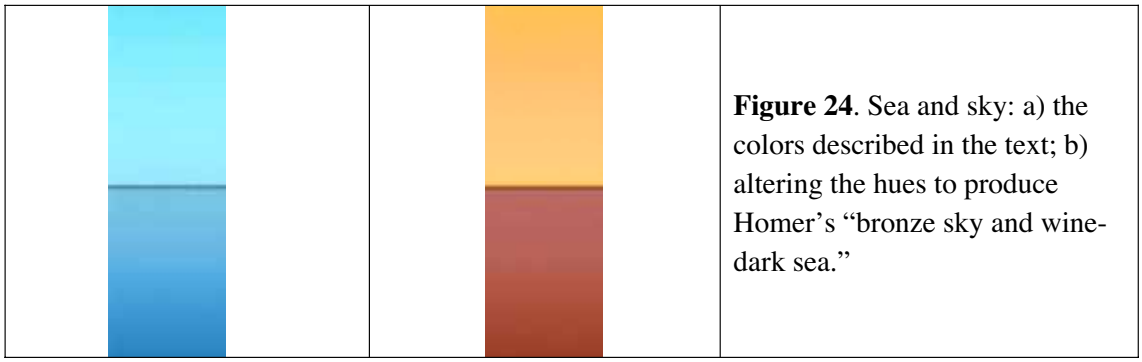


Figure 24. Sea and sky: a) the colors described in the text; b) altering the hues to produce Homer’s “bronze sky and wine-dark sea.”

Figure 24. Sea and sky: a) the colors described in the text; b) altering the hues to produce Homer’s “bronze sky and wine-dark sea.”

Also, human vision manages to find more contrast in areas when we concentrate on them. So looking into a dark region or a bright one will reveal more detail, and that will produce a visual increase in the contrast there. Of course, this isn’t accompanied by any change in the numeric values present. So there is always the question of whether an artificial increase in contrast should be imposed in order to draw the eye to a region, or whether some inherent features of interest will attract attention anyway.

Foliage varies so much on color, from yellow-green to blue-green, that leaves do not usually make very reliable control colors. Fortunately, however, human vision is not very sensitive to changes in color in

this part of the spectrum. We can detect much smaller variations in the blue and red parts of the color wheel than in the green region, so there is somewhat more latitude in adjusting green colors without offending the viewer.

Colors of clothing are rarely useful choices because it is difficult to know what the real color is, or how it should appear in the lighting used for the picture. And skin tones are very tricky indeed. The range of hue, saturation and brightness values on human skin vary not only from race to race, but considerably within each race. The pinkish color of Nordic blonde skin and the olive-toned hues of a Mediterranean brunette give only a hint of the variability encountered. Most persons of one race are not very familiar with (yet alone able to describe) the range of skin tones that occur in members of other races. In addition to the problem of having to guess at the skin color itself, the values recorded by the camera will be influenced by shadows, the nature of the light source, and the fact that some of the light illuminating the skin has been colored by reflection from clothing.

Color adjustments for the subject of interest

Few images represent scenes in which everything is of equal interest. Usually there is a foreground subject and some background to provide context. It is much more important to adjust the color for the foreground than to preserve all of the background colors or even detail. In **Figure 25**, the camera has captured an image that shows the background detail behind the bride and groom, but interest is centered on them. While the image histogram shows good range and balance, it is clear visually that the bride's dress is so bright that folds and detail are lost, and the skin tones are not well balanced.

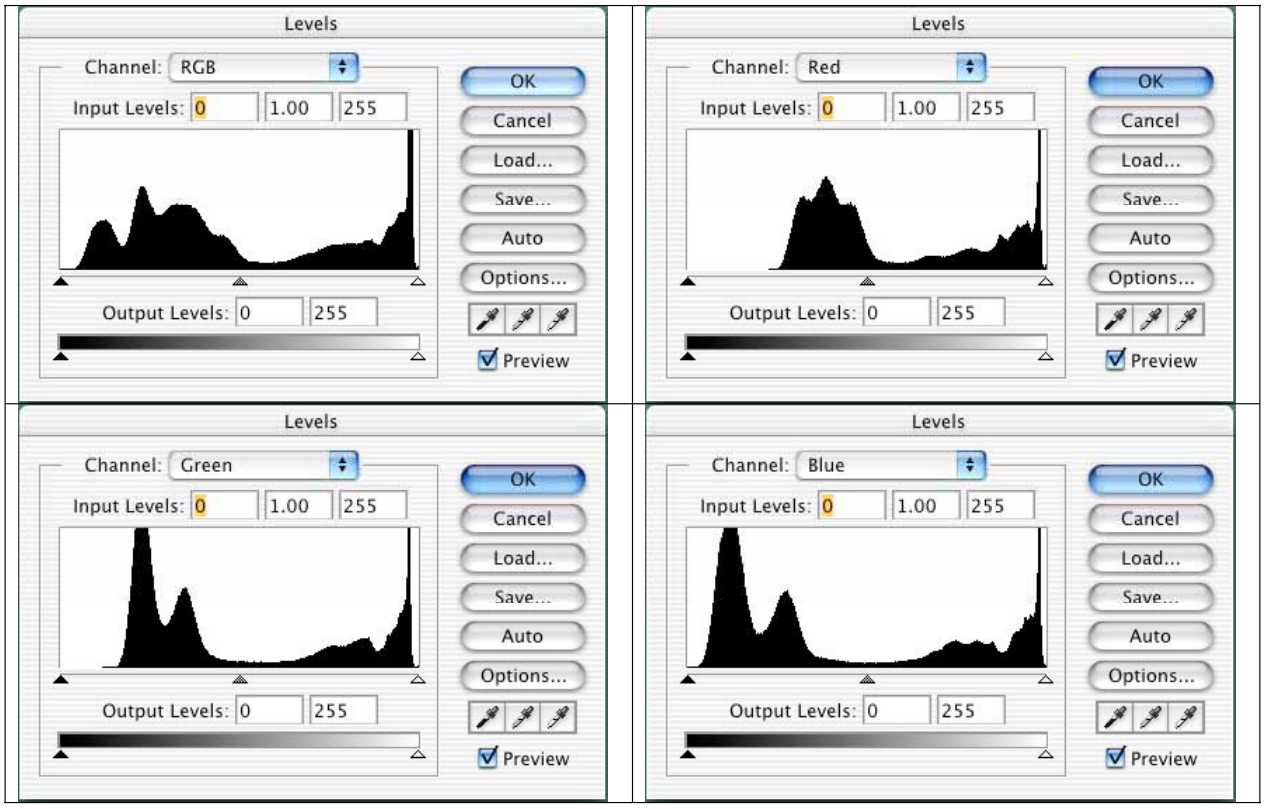


Figure 25. Wedding picture (a) with the histograms for each channel (b-e).

Several improvements can be made to this image. First, it needs to be cropped. The uninteresting background takes up too much of the picture area. The Photoshop cropping tool was used to select the are containing the bride and groom. This left a part of the lamp hanging over their heads, which was removed using the patch tool discussed in **Chapter 4**.

Curves for the red, green and blue channel adjustments were made as shown in **Figure 26**. First, the light and dark eyedroppers were double-clicked and the maximum and minimum brightness values for adjustment were set to 95% and 5%, respectively. The eyedroppers were then used to select points in the shadow area on the groom's tuxedo and a bright area on the bride's gown. The result was a neutral dark grey for the tuxedo and a neutral near-white for the gown, achieved by positioning the end points of the adjustment curves. The midtone and quartertone points were then adjusted to increase contrast in the bright areas of the gown (showing the folds that had been hidden before) and to balance the skin tones. As a side effect of the adjustments, the background becomes very dark, and serves to effectively highlight the people, which is the proper focal point for the image.

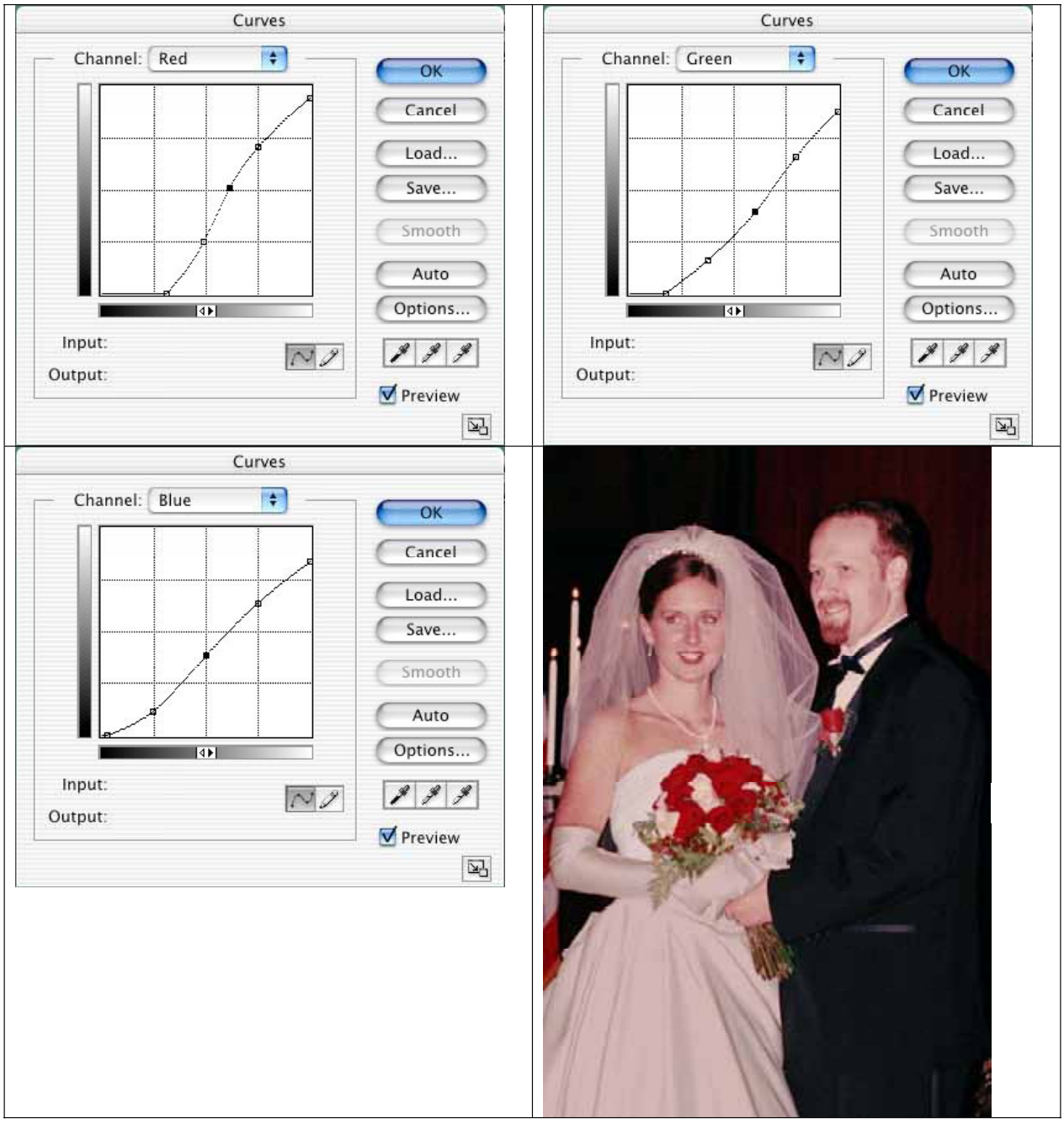


Figure 26. Adjustment curves applied to **Figure 25**, and the resulting image.

Matching colors within the image

In some cases, the issue is not so much what color a feature should have as making sure that all of the similar objects in the scene have the same color. **Figure 27** shows an example in which the pumpkins in the shadow of the building have a much more reddish tone than the yellow-orange pumpkins in full sun. This is a typical example of the color shifts that cameras - both digital and film - record when the

lighting changes. Because the pumpkins in the shade are also darker than those in the sun, it is relatively straightforward to adjust this image so that the hues are consistent. In the example, four control points were set. One pair is on a sunlit pumpkin and one pair on a similar location on a pumpkin in shadow. Within each control point one has been set to display the HSB values while the other shows the RGB data. The latter may be useful in developing experience with the relationship between these different color representations.

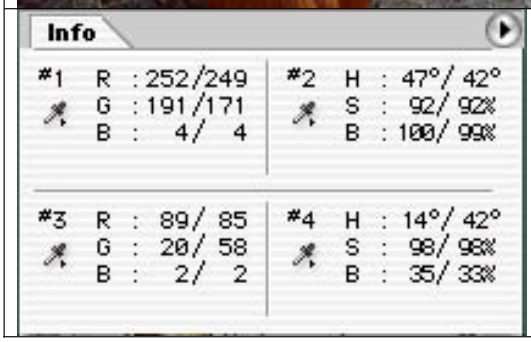


Figure 27. Color correction used to make similar objects the same color: a) original image; b) corrected image with control points (1 through 4) marked; c) Info display showing the RGB and HSB color values for the original and adjusted image.

Using the hue values in the two different regions as a guide, the curves were adjusted so that the two regions match. The Photoshop Info display shows the original color values followed by the modified ones. There are still differences in brightness and saturation, of course. When the original and corrected images are viewed side-by-side the redder colors in the shadow area are quite apparent. But unless we were using a measurement tool on the original image, we probably would not have noticed the difference visually.

The fact is that human vision makes a pretty good automatic compensation for shadows and for changes in the absolute color of objects with illumination. Since we know that the pumpkins in shadow are really identical to those in sunlight, we don't allow ourselves to see any color difference between them. So while technically this image can be adjusted to have uniform hues in the sunlit and shady regions, it is not so obvious that such an adjustment needs to be made. Sometimes it is good enough to depend on the peculiarities of the human vision system to perform contrast enhancement or color balancing.

This is especially true in color rich scenes like the one in **Figure 28**. The same shadows are present as in the previous example, but there are so many different colors present that adjustment may not really be necessary to comprehension of the image or to pretty accurate comparisons of the various colors. In fact, the white areas (the signs on the baskets, shirts on several people) all indicate that far from being a neutral color, the reds are much too low and the blues too high. Using the same methods shown above for neutral colors, it is easy to adjust this image to correct the neutral tones and to set the bright limit for the display to the top of the histogram. The result is a much improved image from the standpoint of color saturation and contrast, but whether it actually makes any difference in recognizing the vegetables is an open question.

Figure 28. Adjusting neutral white in a color-rich image: a) original; b) expanding contrast linearly by setting limits to the ends of the histogram; c) image with color channel adjustment curves used to set control points (marked 1 and 2) to neutral white; d) Info display showing the RGB and HSI color values for the control points in images a, b and c.

The situation can be difficult if the image has no obvious known colors or neutral colors. In that case the goal is often to present the viewer with an much contrast in brightness and color as can be achieved without making any of the individual colors implausible. In the example of **Figure 29**, the camera has

recorded colors using an outdoor white point setting but with open sky as the light source, not direct sunlight. In addition to the image being generally flat due to the gentle lighting that produces few shadows (even though the histogram does cover the full dynamic range from dark to light), the colors are visibly too magenta on the red leaves (we expect the hue to lie on the orange side of red, with no blue cast).

With few other known reference points to act as a guide, the usual strategy is to make the adjustment curves into an “S” shape that increase contrast in the mid range. By doing it channel by channel it is possible to avoid false colors. Since green is the predominant color, it was done first. Then the red channel was expanded even more than the green, while examining the tan and brown colors. Tan and brown have an orange hue, with low saturation and low brightness, respectively. Finally, the blue was reduced enough to keep the leaves red or orange red, rather than magenta. Throughout the process, colors were monitored in HSB at a number of points on the image.

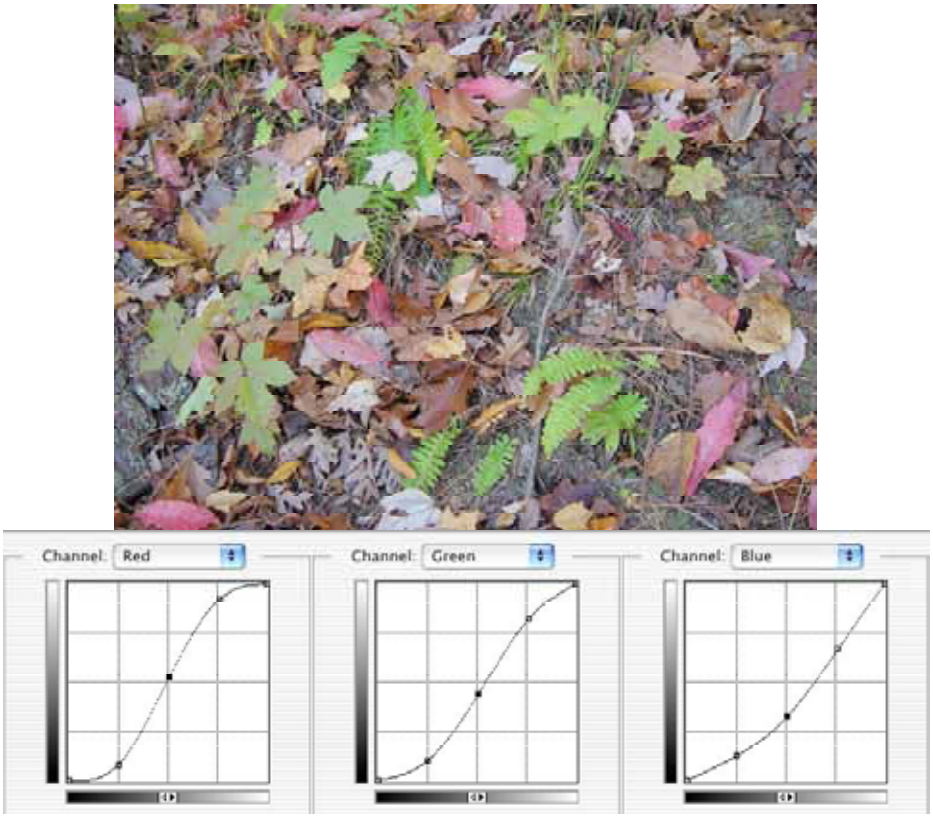


Figure 29. Adjusting colors arbitrarily: a) original image; b) adjustment curves; c) final result (compare this to **Figure 19**).

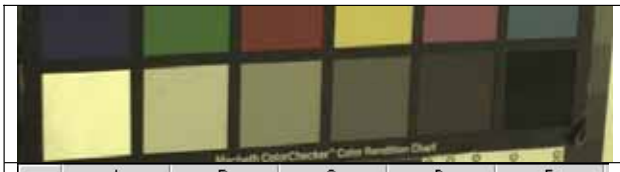
True color

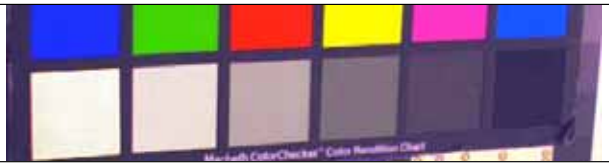
In some applications, such as photography of clothing for catalogs, it is very important to have the printed colors match the original objects. That is mostly a matter of properly adjusting the camera white point for the color temperature of the lighting, performing a calibration of the monitor, and using the correct printer software so that the match is achieved. This is covered in **Chapter 5** in the context of printing hardcopy of images. It does not involve actually measuring the colors in the image.

The color temperature of the light can be controlled to a good extent in studio situations, and in uniformly illuminated outdoor settings, but varies considerably in scenes with shadows, and even more so in interior scenes with mixtures of incandescent and fluorescent lighting, extended or diffuse light sources and point source lighting, light reflecting from colored walls, etc. Because the camera chip has colored filters that pass ranges of wavelengths to the transistors, there is literally an infinite number of combinations of actual color spectra from objects that will produce identical red, green, blue outputs from the camera. Digital cameras are not spectrophotometers and cannot be used to "measure" color.

But it is possible to adjust the acquired images to produce the same visual impression as the original scene. This method is both simpler and more complicated than the curve adjustment methods discussed above. The math is a little bit more complicated, but the computer can handle that. The correction procedure is actually simpler than manipulating curves. The basis of the method is that the color filters in the camera have wavelength cutoffs that overlap somewhat (as do the response functions of the long, medium and short wavelength cones in the human eye). That means that some wavelengths of light are detected by both the red and green, or by the green and blue sensors. Also, the filters don't cut off entirely so there is even some red light detected by the blue sensor and vice versa.

These overlaps can be measured. The standard procedure is to use a target with known color swatches, such as the widely used GretagMacbeth chart. This is photographed under the same lighting conditions as the objects (obviously, this works best in a studio setting). Film photographers have commonly taken a similar reference picture at the start of each roll, to aid in controlling developing and printing, but the process is somewhat simpler with digital image processing. The RGB intensities are measured in the color swatch areas corresponding to visually true red, green and blue, to serve as reference points for subsequent correction of all colors that are composed of those three primaries. In principle it is possible to do this with other colors, but the math and procedure are both simplest using RGB.





	A	B	C	D	E
1	**Color Correction using Tristimulus values**				
2					
3	Measure the RGB intensities from three target areas that are				
4	to be corrected to pure red, green and blue. Enter the intensities in				
5	the matrix below (example data from Macbeth image using the				
6	**IP*Measure Global -> Brightness** function on each region).				
7	Measured intensity:		Red	Green	Blue
8	Area :	Red	90.223	45.865	35.556
9		Green	59.933	86.941	40.296
10		Blue	38.088	36.635	43.028
11					
12	Intermediate normalized intensity matrix (above values converted				
13	to a percentage of 255)				
14			0.35381569	0.17986275	0.13943529
15			0.23503137	0.3409451	0.15802353
16			0.14936471	0.14366667	0.16873725
17	Inverse matrix calculated by Excel (MINVERSE function)				
18	Enter these 9 values into the **IP*Color -> Tristimulus** dialog box				
19	Channel:		Red	Green	Blue
20	Result:	Red	4.9455668	-1.465084	-2.7146869
21		Green	-2.2799016	5.52035567	-3.285862
22		Blue	-2.4366123	-3.4032749	11.1270463

Figure 30. Tristimulus color correction: a) original image of a GretagMacbeth color chart; b) calculation described in the text (screen shot from Excel); c) corrected image.

The set of nine measured values - the red, green and blue intensities measured on the red, green and blue swatches - form a 3x3 matrix of numbers. This matrix of numbers must be inverted. The example shown in **Figure 30**, illustrates carrying this operation out in an Excel spreadsheet using the MINVERSE command. The result is a set of nine correction factors, called a tristimulus correction. Usually the three factors along the diagonal of the matrix are positive (the contribution of red light to the red sensor, green light to the green one, and blue to the blue) while the rest are negative (to remove the contribution of one color light from the other sensors). The procedure is to calculate for every pixel in the image a new set of R'G'B' colors from those that were measured.

$$R' = \alpha_{RR} * R + \alpha_{RG} * G + \alpha_{RB} * B$$
$$G' = \alpha_{GR} * R + \alpha_{GG} * G + \alpha_{GB} * B$$
$$B' = \alpha_{BR} * R + \alpha_{BG} * G + \alpha_{BB} * B$$

As shown in **Figure 30**, when this tristimulus procedure is carried out it corrects the red, green and blue color swatches as well as the cyan, magenta, yellow, and grey scale swatches. Since this procedure requires performing the matrix inversion and also accessing the three color channel values for every

pixel in the image, carrying out the calculation, and writing the new values back, it is easiest to perform it using a Photoshop plug-in. The Fovea Pro 3.0 Tristimulus correction plug-in (www.ReindeerGraphics.com) was used in the example.

Chapter 3
Combining multiple images

Combining several pictures, either taken with different exposure settings, different focus settings, or different point of view, is difficult for the film photographer, but can offer some unique advantages with digital images. This chapter shows a variety of applications and methods. The first topic deals with ways to extend the dynamic range of digital camera images.

The film photographer always faces the challenge of adjusting the exposure and lighting of a scene to balance the brightness and contrast of all parts of the resulting image at the same time, in order to record them on film. One of the major advantages of digital cameras is the ability to capture multiple images with different focus, exposure and lighting conditions, or overlapping points of view, or over a period of time, and then combine them to produce an optimum result.

There are several ways that this can be done, which are illustrated in the following sections. The tools used are averaging, blending arbitrary percentages, keeping the brighter or darker pixels values at each location, and blending the exposures based on the local contrast in each image. By choosing from this short list of techniques, a wide variety of difficult imaging situations can be accommodated to use multiple images acquired with different lighting or exposure conditions to produce a significantly enhanced result.

Merging pixel values for optimum exposure

One limitation of digital cameras is that they generally have a significantly smaller dynamic range than film. For instance, the most common 8-bit detectors record 256 brightness levels (and as discussed elsewhere, resolve fewer distinct levels than that), while color film can typically distinguish 3-4000 levels. But even film has difficulty in providing adequate contrast in scenes such as the example shown in **Figure 1**. The sunset colors in the sky wash out when the foreground detail is well exposed, and the foreground and the backlit trees are too dark when the sky is well exposed.

Figure 1. Two original exposures (images courtesy of Jim Reswick) taken with a tripod mounted camera and a difference of three stops in exposure.

One very labor-intensive way to combine these two images (taken with a tripod-mounted camera) is to manually select the sky region, and copy those pixels from the good exposure of the sunset sky to paste

onto the image with the good exposure of the foreground. Even if this can be accomplished successfully, it usually creates problems along the region boundary where fine detail is either lost or shows a halo effect.

This is exactly the same approach and result – and requires a similar amount of effort – as creating a mask for dodging and burning while printing a photographic negative in the darkroom. That is necessary because while the dynamic range of film is very high, that of photographic prints is not. Compressing the entire image into a print by using soft paper means the loss of local contrast (and with it detail).

A much better (and automatic) method adds the two digital images together, combining the two recorded values for each pixel according to the local contrast present in each image (in other words, the blending factor varies from pixel to pixel, in order to produce a result in which the final local contrast is optimized). **Figure 2** shows the resulting blended image.

Figure 2. Blended result, which can then be further manipulated using the contrast controls discussed in Chapter 2.

Loss of saturation

The preceding example started with images taken with a difference of three stops of exposure. Attempting to extend this technique to differences of 5 or 6 stops often introduces another problem, loss of color saturation. Very dark areas are so underexposed they have very little color information, and at the same time the overexposed areas are blown out and also have little color. Pixels with color values at the extreme light and dark corners of color space contain only a small amount of color (that is, little saturation). Combining them using the blending method shown above produces values of intermediate brightness, but with low contrast and lack of saturation.

Of course, the traditional method for dealing with situation is to use a mask to copy just the properly exposed pixels from one image and paste them into the second. Rather than trying to draw the mask by hand, it is often useful to convert the image to CMYK and use the wand tool to select regions with low saturation. Feathering this selection slightly also helps.

It is interesting to note another problem that may arise in blending. If te two exposures were acquired by adjusting the aperture, this also changes the optical depth of field, with the result that while far-away

features remain in sharp focus in the brighter image, the nearer details do not. When the two images are combined, those details acquire halos and ghost edges where the sharply focused and blurred edges are added together. When acquiring multiple images with a digital camera, it is better to adjust the shutter speed and keep the optical settings fixed.

Blending range

The problem of losing saturation when blending extreme exposure combinations is basic, and arises because of the nature of color space. This puts a practical limit to just how different the exposures can be, and reminds us that even with the ability of the computer to manipulate and combine digital images, there must be information there to start with if we are to recover and use it. Film has an extremely wide latitude, typically 3-4 thousand brightness levels can be resolved. The digital camera is typically limited to 256 brightness levels, and usually does not actually resolve that many. In this regard, the digital camera image is more like the photographic print than like film. Photographic prints are also quite limited in dynamic range, and every film photographer has faced the problem of trying to capture detail in the dark and light areas of a print, when that detail is quite evident on the film. Blending together multiple exposures from a digital camera produces a resulting image that has a greater dynamic range than a single exposure, but the problem of compressing this extended range into a single viewable (or printable) photo remains.

The four example images shown in **Figure 3** are each slightly more than two stops different, and cover the range from much too light to much too dark. In the dark shadow areas of the underexposed image, there is no visually discernible detail at all. Nor is there in the blown out bright areas of the overexposed image.

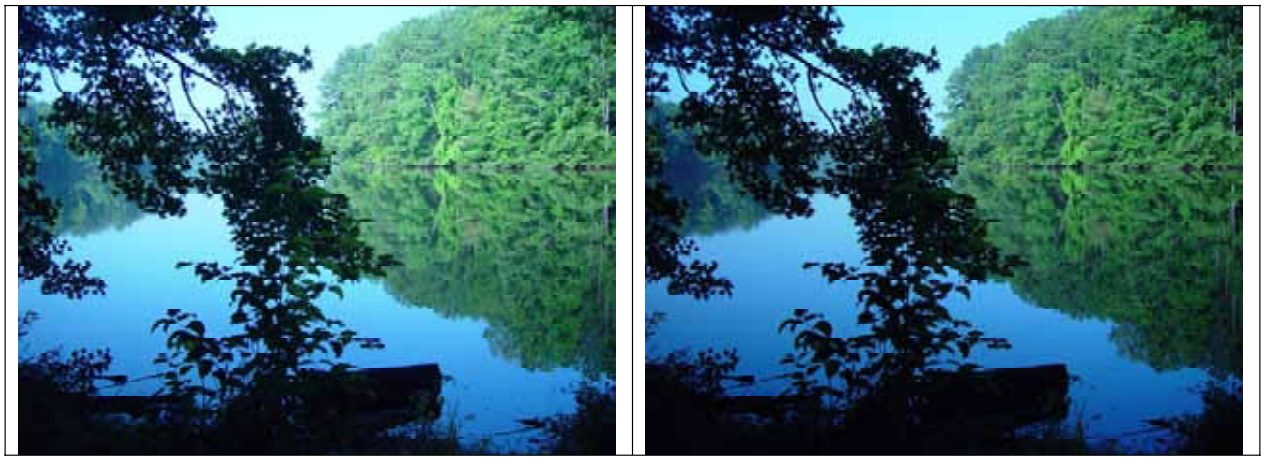


Figure 3. Original exposures used to illustrate blending.

Blending the extreme images, which are about 7 stops different in exposure, produces a resulting image that is clearly undersaturated (**Figure 4a**). The brightness histogram ranges over the full 255 values, so it is not a lack of contrast but rather one of color saturation or richness. Blending images that are only about 4 stops different produces a result that is much better (**Figure 4b**). This is because the pixel values in the shadow areas of the dark image and the bright areas of the overexposed image still have some color information that can be used to intelligently interpolate between them.

Figure 4. Blending the images from **Figure 3**: a) over 7 stops; b) over 4 stops.

The best results are generally obtained when blending images over about 2-4 stops, which extends the dynamic range without seriously compromising saturation. As always in photography, it helps to have good quality images to begin with.

Keeping brighter pixel values

Blending two exposures works best when the individual images have some detail in every location. It works poorly when the images have such extreme exposures that areas are too bright and hence washed or too dark and hence without information. The maximum saturation of colors occurs at intermediate

brightnesses. When pixels are very dark or very light, there cannot be much color in them. For example, the only way to increase the amount of red color in a white pixel is to reduce the intensity of green and blue, which darkens the pixel's intensity.

Another useful way to combine images, that works in these situations, is to keep the lighter pixel value at each location. Consider as an example the pictures in **Figure 5**. The first one uses flash to record the interior of a room, but the view out through the window is too dark. The second, taken without the flash, records the view out through the window but the interior is dark. In principle, it is possible to adjust the shutter speed to give a good outdoor exposure and the aperture to work with the flash intensity to give a good interior exposure, but in practice this can be difficult to set up.

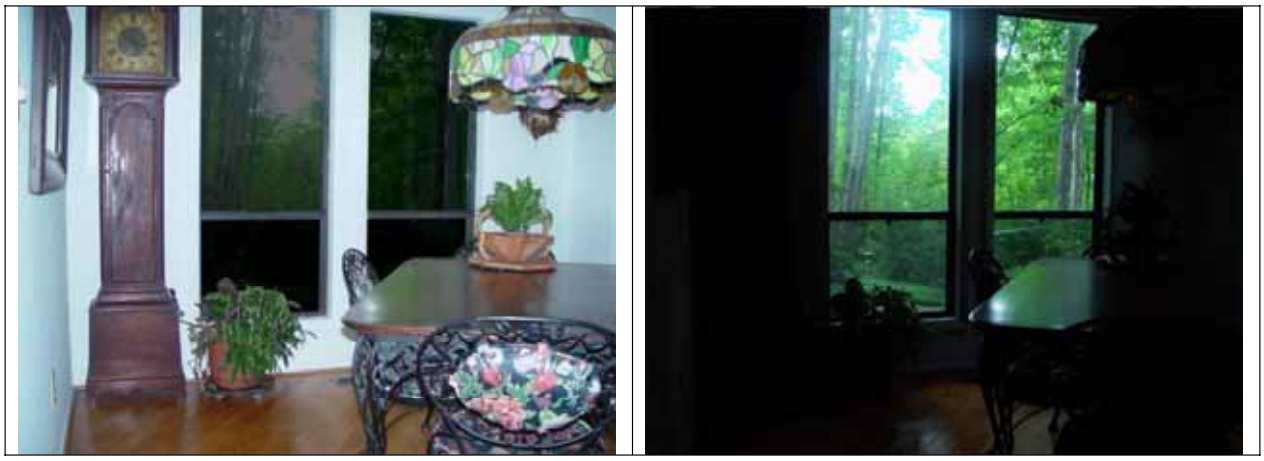


Figure 5. Two exposures taken with a tripod-mounted camera, one using flash to illuminate the interior of a room and one longer exposure to capture the exterior scene.

Combining these two images, pixel by pixel, and keeping at each location the values for the lighter pixel effectively merges together the well exposed bright interior and the well exposed outside view (**Figure 6**). This automatic technique is far easier (and more accurate) than manually trying to outline selections around the bright regions in the exterior image, cutting out those pixels, and pasting them into the interior image, but produces the identical result.

Figure 6. Combined result from the images in **Figure 5**, showing well-exposed pixels in both the interior and exterior.

Keeping darker pixel values

Photographing objects, especially in closeup or macro mode, often creates difficulties with specular reflections and highlights. Sometimes these can be controlled by diffuse lighting (umbrellas or diffusers) but these also tend to reduce the "pop" in the image that results from clearly and sharply defined edges and detail. A different approach is shown in the example below.

The object in **Figure 7** is a crystalline-glazed pot, which has many fine crystals on the surface, each of which reflects light. A picture with single-point lighting shows a large patch of specular reflections. By adjusting exposure these could be blown out to white, producing a region with no detail. The alternative is to reduce the dynamic range available for the rest of the image. Capturing a second image with illumination moved to another location moves the reflection but otherwise has the same problem.

Figure 7. Two images taken with a fixed camera position but changing the location of the light source.

Since these two images were captured with the camera on a tripod, they are in registration with each other. Hence it is straightforward to combine them, keeping whichever value is darker at each pixel location. The result (**Figure 8**) eliminates both of the glare spots, and hence gives more range to the colors for the rest of the image, while keeping the sharp detail and interior contrast in the surface finish. Obviously, this technique can be readily extended to more than 2 light sources.

Figure 8. Combined result from the images in **Figure 7**, eliminating the glare spots.

If a specific highlight is desired in the image, particularly one from a backlight, it is easy to blend in a small percentage of that as well. In the example shown in **Figure 9**, 15% of the backlit image has been added to brighten the top of the pot. This addition can be performed in Photoshop by placing the images into layers and adjusting the opacity of the top layer, or by using the Image->Calculation function.

Figure 9. An additional backlit image (a) and the result of blending 15% of it into **Figure 8** to provide a highlight and reduce shadows (b).

Averaging multiple exposures

The situation is quite different when the object being photographed is black. Getting good snappy pictures of dark or black objects is one of the most difficult challenges most photographers face. With no colors and little reflected light, the textures and shape are revealed by small specular reflections from points on the object. Typically, any placement of illumination that shows some of these will hide others, and as in the example above large areas of specular reflection are blown out and affect the overall contrast of the picture.

Figure 10. Several images of a black piece of pottery with variations in surface finish, taken with different locations for the light source.

In the example shown in **Figure 10**, four images were captured using different locations for the illuminating light (a flood shining into an umbrella). None of the pictures is bad. Each one shows some of the surface texture, but has deep shadows in other areas. It is possible to get a better result by combining them. In this case, we do not want to eliminate the specular reflections by keeping the darker pixel because we need them to reveal the object's surface texture and shape. Instead, the four images are averaged together into a 16 bit space, from which an optimum 8 bit image (**Figure 11**) is obtained by setting the limits to the minimum and maximum luminance values. This is the same procedure used to combine multiple images for noise reduction, discussed below, and it has that effect here but it is less important than the blending of the images to control the reflections. The operation was carried out using the Average Frames plug-in in Reindeer Graphics' Optipix package (www.ReindeerGraphics.com).

Figure 11. Result from averaging the multiple exposures in **Figure 10.**

Note in the composite that all of the specular reflections are still present, and that they help significantly in giving shape to the pot (especially to its handle) but are limited in brightness and size. At the same time, the texture is also revealed in all areas (including the interior) and there are no completely shadowed regions.

Highlights

Another situation in which highlights are important is the always-difficult task of photographing gemstones. But in this case we must combine images that have been acquired with quite different types of lighting. A top-lighted image of the stones gives nice highlights from the facets, which help to show the shapes of the cut stones. But lit from the top, the colors of the stones are typically dark and muddy. Lighting them from below with a light box provides clear and true colors, but does not produce good shape definition or highlight the facets. **Figure 12** shows the original photos.

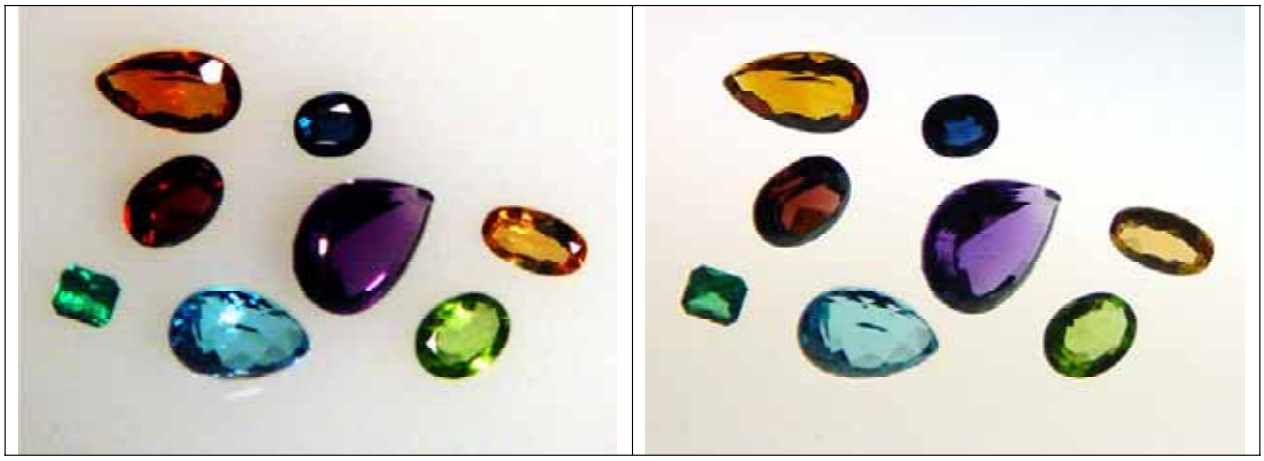


Figure 12. Original images of gemstones, with top and bottom illumination.

The best solution for this case is to combine the two images keeping whichever pixel is brighter (**Figure 13**). This preserves the true colors of the stones but superimposes the highlights, which give the resulting image more character and impact.

Figure 13. Result of combining the images in **Figure 12**.

Of course, this same approach works in a lot of other cases. A common problem is taking a good picture of the Christmas tree. If the room lights are on or you use a flash, the details of the tree and ornaments are recorded but not the tree lights. Take a second exposure with the room lights off (or the flash off) to capture the lights, and then combine them with the original keeping the brighter pixel values. The lights will turn on in the picture. When you do this, it is important to make sure the dark image is well focused. Since the aperture will probably be small to get adequate depth of field, this means that the shutter speed will be long. As discussed in **Chapter 1**, digital cameras have noise problems with long exposures, so a noise canceling step of recording and subtracting a dark frame should be used if your camera supports it. And, of course, you need to use a tripod both for the long exposure and to make sure the image of the lights is properly aligned with the one of the tree, but all of these multiple exposure operations require that.

Averaging multiple images

Multiple images taken with the same exposure settings can also be combined for an improved result. In a single 8 bit per channel image, the presence of random speckle noise, discussed in **Chapter 1**, can reduce the amount of real information significantly. This noise arises from many sources, including the generation of electrons by photons, the collection of those electrons, the amplification of the signal, and the measurement or digitization process. The random variation can be seen as the width of the histogram peak for an image of a nominally uniform region. As shown in **Figure 14**, an image of a photographic test wedge shows a series of peaks whose spacing can be used to plot the linearity of the camera detector and electronics, while the width of each peak measures the noise, summed from all of the various sources.

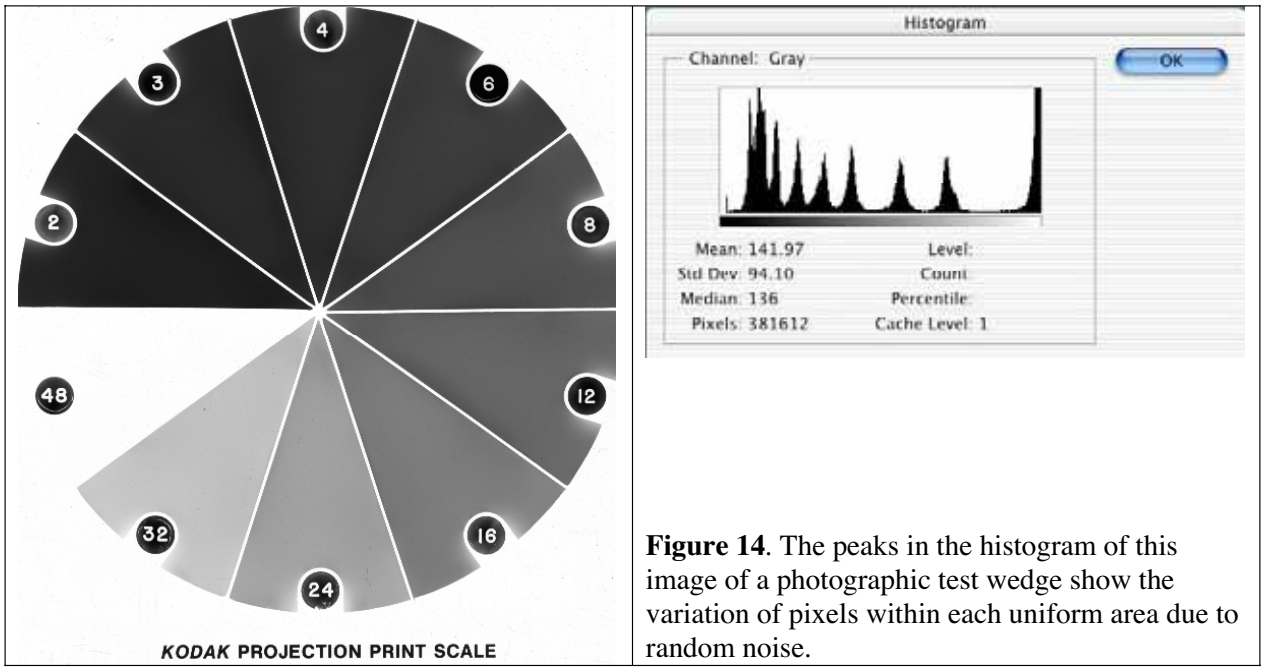


Figure 14. The peaks in the histogram of this image of a photographic test wedge show the variation of pixels within each uniform area due to random noise.

In the example shown, which was acquired with a scientific grade cooled camera, the peaks have a width (measured by the standard deviation) of 2.5-3.5 shades of grey, out of a total range of 256. This means that of the nominal 256 shades of grey recorded, the actual working grey scale resolution is less than 100. It also means that a change in brightness must be larger than the noise level to assure detection. Typically, the noise level gets worse and the number of grey levels that can be discriminated drops as the amount of gain used to amplify the signal is increased (e.g., taking pictures with reduced lighting and increasing the effective ASA rating for the detector).

The noise level in the image can make it difficult to distinguish true pixel values that are close together. The visual appearance of the speckle noise can be increased by the use of adjustments such as levels or curves. There is an effective way to reduce the proportion of the image that is noise, by adding together a sequence of images. These are usually acquired by mounting the camera on a tripod, and of course only apply to a scene that does not change for the length of time needed to capture the images. Some digital cameras have a burst mode that will automatically acquire and save a few images. Others can acquire a short movie clip, albeit with reduced resolution. Or a video camera might be used and the individual frames in the movie averaged together.

In all of these cases, the signal that corresponds to the scene adds up in direct proportion to the number of frames added together. If a single image has a dynamic range of 8 bits (256 brightness values), then adding together as few as four frames will produce an image with 10 bits (1024 values). But on the average the random speckle noise will cancel out in multiple images, and so the signal to noise ratio improves in proportion to the square root of the number of images. In other words, adding together 4 images doubles the signal to noise, adding together 16 would quadruple it.

Of course, the adding together must take place in a 16 bit image space, which can hold the sum of more than 32000 frames. Once the images have been added together, a new 8 bit image can be extracted by finding the brightest and darkest ends of the image histogram and rescaling the pixel values between those limits. Unlike the use of limit markers in the Adjustments -> Levels function, this does not produce a histogram with missing values, because the higher precision of the 16 bit space typically contains pixels at all values in the range. In many cases it may be desirable to perform other processing or pixel value adjustments on the image in the 16 bit space, before converting in back to 8 bits for printing.

It is important to understand that this process is not the same as using a longer exposure time to capture more light in a single image. If the result did not overflow the dynamic range of the detector (a significant risk considering the small well size of some of the tiny chips used in consumer cameras, as discussed in **Chapter 1**), the additional exposure would increase the number of electrons to be collected. That would reduce the contribution made to the noise by the statistical processes by which a photon is absorbed to create free electrons. But it would not affect the other sources of noise in amplification, readout and digitization. And in most cases, increasing the exposure will cause many pixels to exceed their maximum capacity, resulting in blown-out highlights.

Averaging multiple exposures does not produce these problems. It can be accomplished in Photoshop by using a plug-in such as the image averaging routine in Optipix (Reindeer Graphics.com). **Figure 15** shows an example of the improvement in noise level and ability to adjust contrast in an image. The original images were acquired in deep dusk. With the longest exposure setting on the camera (1 second at f 4.5) each image was still underexposed, and the camera's preview screen was virtually black so that focusing was done by forcing the setting to infinity. Increasing the contrast on a single frame by setting the limits in Adjustments -> Levels reveals the presence of speckle noise. Averaging together 6 frames provides an image with much lower noise as well as pixel values that cover the full 0 to 255 range.

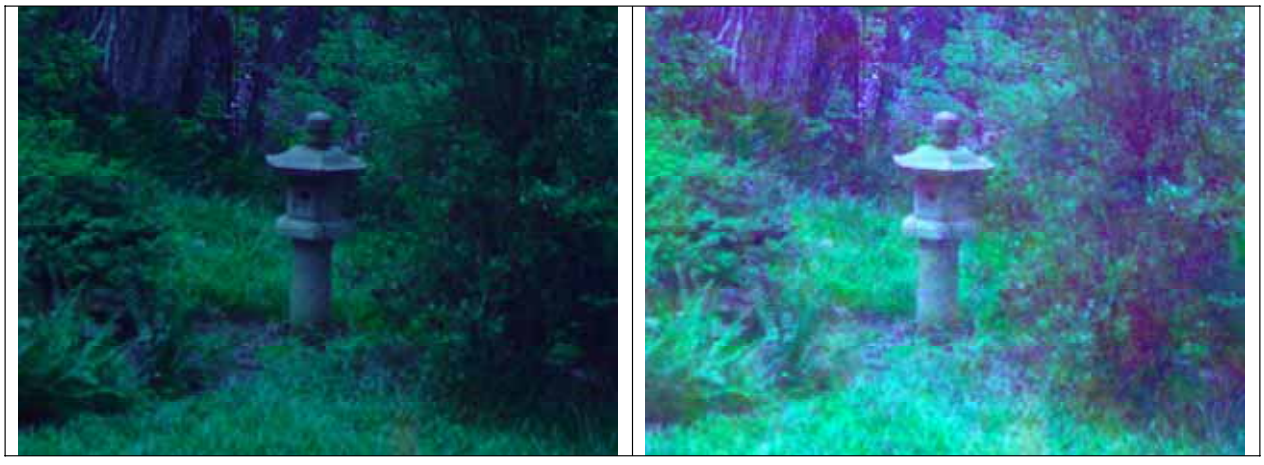


Figure 15. Averaging together multiple exposures: a) single frame; b) with levels adjusted; c) average of 6 frames, automatically scaled to full contrast range.

Extending the focal depth

With all optics, there is a tradeoff between aperture and depth of field. Smaller number for the lens f-stop (a wider lens opening) admits more light, but reduces the distance range over which that light is sharply focused on the film or chip. Even with a fully adjustable camera in which lens aperture and shutter speed can be set independently, and in a studio setting with plenty of light available, the depth of field of the lens may not be sufficient to produce a sharply focused image of the entire scene. Of course, sometimes it is desirable to let parts of the scene go out of focus. But the interest here is in the situation where the achievable depth of field, either because of available light, or the need to use a high shutter speed, or the inability to control the camera adjustments, or just the characteristics of the lens, is less than needed to get the depth of field needed to produce a sharp focus where it is desired.

Sometimes this can be solved with a relatively simple cut-and-paste with two images. For example, if a model's face and hand (e.g., holding the product to be advertised) must both be sharply focused, then two pictures can be taken. If the hand can be cut from one and pasted into the second, with a blending area between the two sharply focused regions that is ignored because it is not in sharp focus, then a good result can be obtained. This is the way such problems were typically handled in the photographic darkroom with film negatives, and it is even easier to do it with digital images in the computer.

Figure 16 shows an example with two well separated foreground and background objects. Photographs with each in focus were acquired with identical lighting and exposure. A region drawn around the vase and several pixels larger, with a feathered edges, was used to cut the sharply focused image of the vase from its picture and paste it onto the other, covering up the out-of-focus image. The result looks quite satisfactory, although a careful observer would note that the reflection of the vase on the table surface is not in focus, as it should be. It can be difficult to collect a series of images in which the size of features has not been altered by the change in focus (or at least the region around the in-focus feature can be made large enough to cover the out-of-focus image), and if the features are not well separated, the cutting and pasting becomes much more difficult.

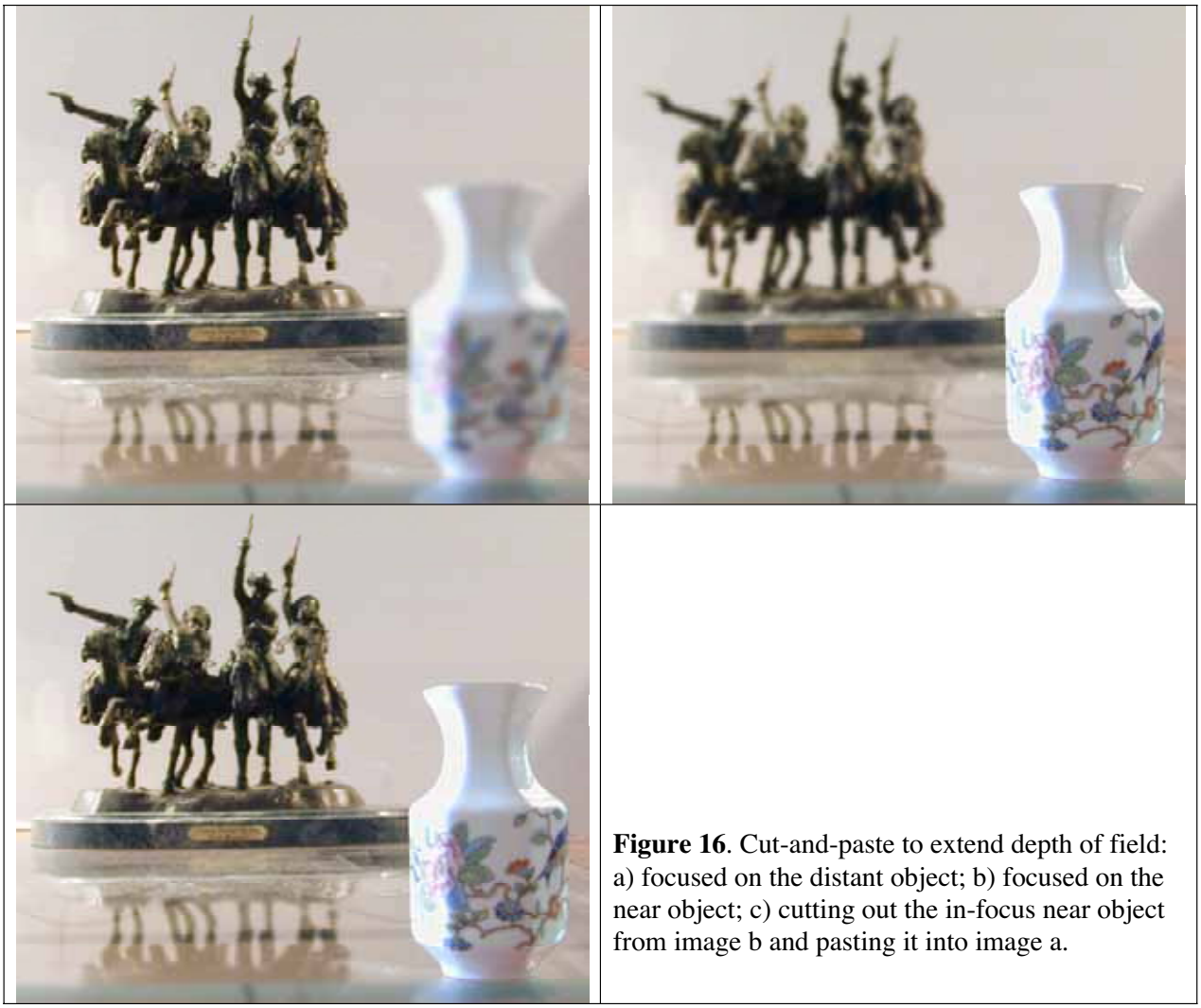


Figure 16. Cut-and-paste to extend depth of field: a) focused on the distant object; b) focused on the near object; c) cutting out the in-focus near object from image b and pasting it into image a.

There is another way to handle this, without the manual outlining requirement. If two (or more) images are collected with different (but overlapping) focal depths, it is possible to combine them algorithmically to get a result that keeps the best-focused pixels from each. It is important that the images all have the same exposure and lighting, and that they are at the same magnification. The latter can be a problem. Most camera lenses change the size of the image as the focus adjustment is made. If this is the case, then pasting the sequence into layers allows them to be aligned (if the camera has been moved slightly on the tripod as the adjustments are made) and their scale adjusted.

Adjusting the layer opacity so the active layer and the background layer can both be seen, choosing Select -> All, and then Edit -> Transform -> Scale allows typing in the scale factor. Typically the size change from one focus adjustment to the next will be small, so the H or W setting might be 99.7% or 100.3%. Enter either one and click on the link icon to maintain the aspect ratio and the other will be set to the same value. It is easier to enter the numbers until the images match than to try to drag edges of the region frame. Once all of the layers are the same size, they can be merged by examining the pixel values at each location and keeping the one that has the sharpest focus. The example in **Figure 17** was created using the Best Focus plug-in from Fovea Pro (ReindeerGraphics.com).

Figure 17. Merging multiple layers to produce an extended focus result: a..f) individual images taken with different focus settings; g) merged result with both leaves and flower in focus.

Sometimes, there is another way that a similar result can be achieved. If the multiple images in the preceding figure are simply averaged together (in a 16 bit space, as described above), the image appears visually to be blurred everywhere. But in fact, all of the well-focused feature edges are present in that averaged image, just hard to see because of the overlying out-of-focus values. A custom sharpening filter like the one shown in **Figure 18** can increase the contrast for the edges (it uses the method described as a difference of Gaussians in **Chapter 4**) and produce a very pleasing result.

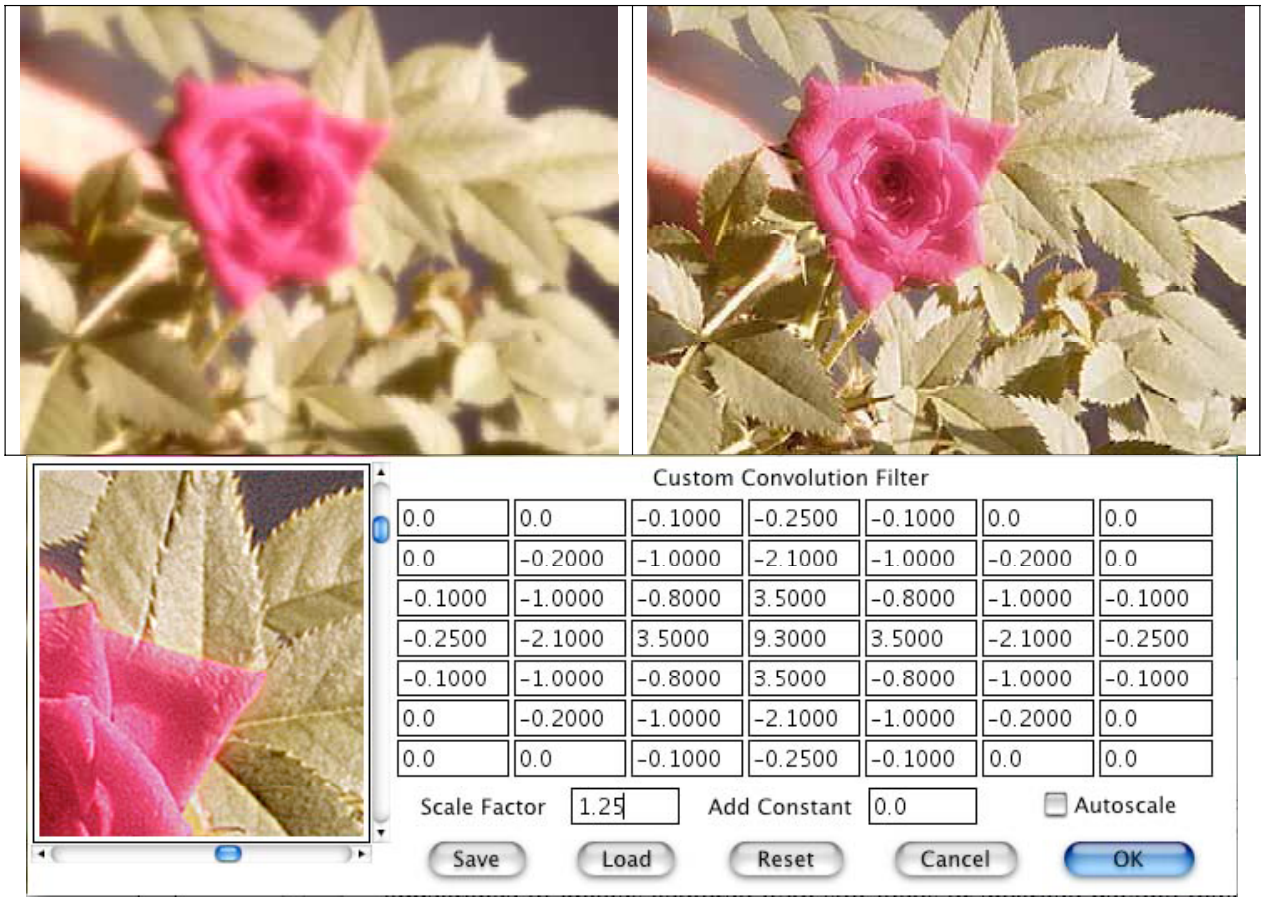


Figure 18. Sharpening the average: a) sum of the six individual images from Figure 21; b) sharpened result; c) custom filter (shown in the Custom dialog from the Optipix plugins, ReindeerGraphics.com).

Instead of adjusting the camera lens focus, it is often easier (and makes unnecessary the scaling of image size) to leave the camera lens unchanged and to move the camera closer to the subject, or vice versa. this is typically the way the procedure is carried out in macro photography, by lowering the camera on a copy stand.

Aligning and adjusting multiple images

In the preceding example, the multiple images were placed in layers and manually aligned and scaled. In general, most of the uses of multiple images require this alignment. If it cannot be achieved by mounting the camera on a rigid tripod, and perhaps using timed shutter release to eliminate any motion of the photographers hand, then layers provide the best environment for manually making the necessary adjustments. Varying the opacity of layers and/or turning them off altogether by clicking on the eye icon for each layer in the Layers window provides flexibility for viewing. Use the Select->All function, choose the layer to be moved, and depress the command/control key while using the mouse to drag or the arrow keys to nudge a layer. The Edit->Transform options also provide tools for either manually or numerically altering one layer to fit well with others.

Sometimes it is advantageous to use a numeric entry to shift the layer rather than simple visual judgment of the alignment. If there are well defined reference points (usually called fiducial marks) in the images,

then their positions in each layer (measured with the cursor) can be used as the basis for aligning one image with respect to another. A single point can define the simple X, Y shifts required if the images are not displaced by rotation. Two points can define both shift and rotation, assuming that the images are not stretched or distorted. Three points are needed if the images may have been stretched, and more are required if the stretching may be nonuniform.

In most cases, the user must identify the fiducial points. If these are marked in a contrasting color, it may facilitate the manual alignment method with the layers. If the coordinates of the points are recorded, then it may be attractive to use a plug-in (such as the one in Fovea Pro, ReindeerGraphics.com) to perform the calculations and apply the transformation for rotation, translation and uniform stretching to the image.

The Fovea Pro plugins can also perform translation alignment without user-selected fiducial marks, by a process called cross correlation. This may be visualized in terms of layers by imagining the process of sliding one image over the other, to all possible positions. The pixel values that align with each possible shift are multiplied together and added. The result is a maximum total when the images are in alignment. The advantages of the plug-ins are that they handle 16 bit per channel images, which cannot be placed into layers, and they are capable of interpolating with sub-pixel accuracy, whereas the shifting of one layer with respect to another uses in whole pixel increments.

These methods may be useful in some few situations for digital photographers, but are primarily of interest to quite different applications such as aligning medical images taken with MRI, CT and PET scans, or for lining up remote sensing images taken at different times and from different planes or satellites. The digital photographer can generally handle the requirements for multiple image alignment (when the simple solution of using a good tripod doesn't solve the problem) by interactive use of layers, or by using the image stitching routines discussed below.

For many situations in which multiple images are to be combined, a second concern in addition to their proper alignment is the adjustment of brightness, contrast and color so that when the images are viewed together the seam where they meet is invisible. Sometimes this can be done by selecting a few common points using the color sampler tool and adjusting the levels or curves until they match. If the problem is simply a slightly different exposure between the images, levels adjustment is often adequate. If the color changes result from altered illumination, the problem is more difficult because different colors are affected differently. In that case, monitoring several different colors and making adjustments to the curves is usually required. Of course, if a color target is included for each lighting condition, a tristimulus correction as discussed in **Chapter 2** can be applied.

Anaglyph stereo images

A rather specific application that requires two images of the same subject, but not taken from the same camera position, is creating images for stereo viewing. A periodically recurring enthusiasm for taking stereo photographs has led to the development of film cameras that simultaneously expose two images through lenses about as far apart as the human eyes (roughly 7 cm). Developed and viewed through a device that allows each eye to see the appropriate image, these are perceived as a full depth live view of the original scene.

There are no stereo cameras using digital sensors, but it is fairly easy to take two sequential pictures with a digital camera, shifting the camera sideways between shots. If the scene subject is quite far away, a simple shift works fine, but if the subject is close to the camera, then in addition to shifting the camera

sideways it must be turned slightly so that the subject stays at the center of the field of view. This is exactly analogous to the fact that when viewing a nearby subject the human eyes rotate in the sockets ("vergence") to point towards the subject. The typical vergence angle for stereo viewing is 5 to 10 degrees. Sometimes it is possible to keep the camera in one position and rotate the subject by this amount for a second picture.

It is not always easy to reposition the camera so that the subject remains centered, but the use of the same approach described above will work nicely for final alignment. Place the images into layers and shift one so that the center of attention in the images is lined up. These images can then be printed out or even displayed on the CRT for side-by-side stereo viewing. People who work extensively with such images have usually learned how to view them without optical aids, either by allowing their eyes to point straight ahead while still focusing on a nearby image plane, or by crossing their eyes. Most viewers will require some type of viewing device to satisfactorily see side-by-side stereo.

Another very useful viewing method uses colored glasses. The two images must be grey scale for this mode (the side-by-side method works fine with color). The left eye image is placed in the red channel and the right eye image in the green and blue channels of an RGB image. Once the images have been aligned, this is easy to accomplish by duplicating the right eye image and then using the Merge Channels function in the Channels window. Glasses with colored lenses are often given away in advertising promotions, or can be saved from a stereo movie showing (which uses the same method). They can also be purchased (see web sites such as <www.2spi.com/catalog/stereo-3D/stereo_red.shtml> or <taggart.glg.msu.edu/stereo/3dview.htm>).

As shown in **Figure 19**, the two images can be merged into color channels with different horizontal positions. It is important that the vertical alignment be correct, as people have a very difficult time shifting one eye up or down with respect to the other. But a horizontal shift only affects the amount of eye vergence, which is interpreted as the distance to the subject. Specifically, shifting the left and right eye images laterally will make the subject appear to move behind or in front of the plane of the paper or screen.

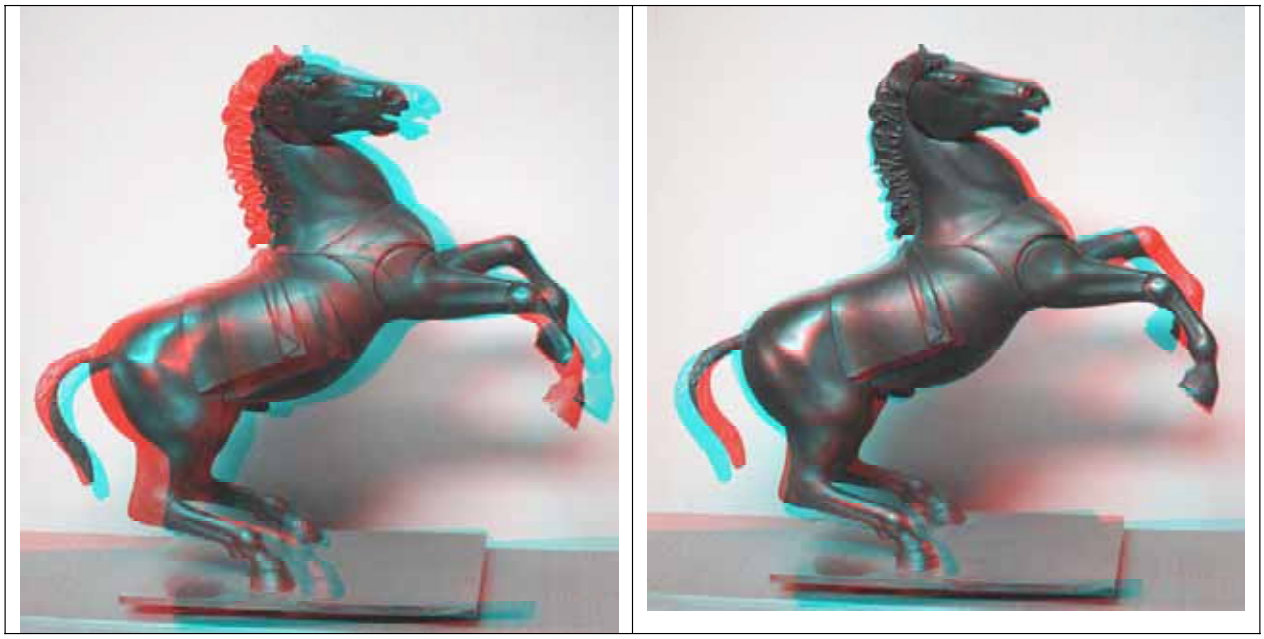


Figure 19. Stereo pairs: a) side-by-side views of a statue; b, c) red-cyan composite of the views. These should be viewed with red/green or red/blue glasses with the red filter on the left eye.

Assembling a montage

Most digital cameras are equipped with lenses with modest optical parameters. Super-wide-angle lenses, panoramic lenses, tiltable or drop-front lenses for distortion control, just aren't generally available. Fortunately, software is available to handle many of the situations that required these optical tools on film cameras (and they were only available for expensive high-end film cameras, too).

On of the simplest problems is obtaining a panoramic shot. Whether you are at the top of a Empire State building or overlooking the Grand Canyon, the camera cannot directly capture the wide expanse of the scene that you assemble in your mind by turning your head. But placing the camera on a tripod and turning it so that a series of slightly overlapping pictures is photographed captures the raw material needed to address this need.

Using Copy and Paste, it is easy in Photoshop to place two images in layers that can be shifted for alignment. The problem is that in most real cases the simple pasting approach will not work. Using normal blending mode and 50% opacity for the topmost layer, it may be possible to shift one layer using either the arrow keys or the mouse (with the Command/Control key depressed) in order to align features in the overlapped portion of the images. Sometimes it is preferable to set the opacity to 100% and the blending mode to difference, which shows the absolute difference between the images in the overlap region.

But as shown in **Figure 20**, aligning one point does not correctly align others. The images are not rectilinear but represent projections of a three-dimensional scene onto the image plane. In is possible in theory to use the various Edit -> Transform tools (scale, skew, rotate, perspective, etc.) to interactively manipulate the images so that many common points are aligned. However, the process is very difficult

and clumsy because it is the boundaries of the selected area that must be dragged to align interior points, and moving one moves all of them. Also, it should be noted that the distortions introduced by this procedure may be visually acceptable but are essentially arbitrary and do not correctly represent the scene.

Figure 20. Attempt to align two side-by-side images in layers. The single arrow marks a point exactly aligned in the overlap region, and the pairs of arrows show the offsets of other locations.

Simple side-by-side pasting does work correctly (and is much easier to implement) when the subject is flat and the camera can be shifted laterally to acquire multiple images while maintaining a perpendicular view at a uniform distance. In the example of **Figure 21**, multiple exposures (or "tiles") were acquired to build a mosaic of an antique map (the same procedure is commonly used to obtain high resolution records of paintings and other art works). The camera was tripod-mounted, and a square and ruler were used to make sure the macro lens was always at the same distance from the map, and pointed perpendicular to it. Nevertheless, slight errors that alter the scale by about 1% or rotate the image by about 1 degree are virtually impossible to avoid unless a complicated and expensive scanning framework is used.

This approach of shifting the camera over the object to be recorded (or vice versa in some cases such as a copystand or a microscope) should ideally produce rectilinear images that can be aligned by simple rigid shifting and rotation of the individual tiles. In fact, some microscope systems with highly precise, motorized stages, acquire the tiles automatically and assemble them based simply on the distance the stage has been commanded to move, assuming perfect hardware precision and no need to consider rotation, optical focus, etc.

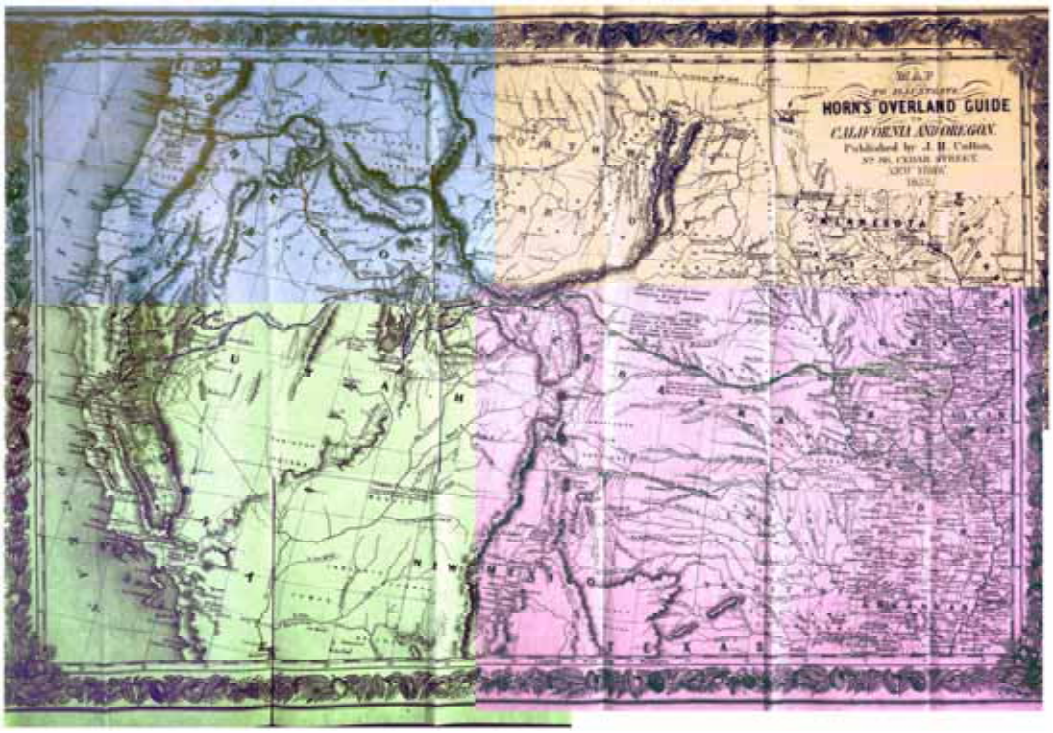


Figure 21. Four tiles arranged to show their coverage of the original map. Each tile has had its hue shifted to make it easier to distinguish from the others.

These assumptions become more doubtful as magnification is increased, and for very high magnification imaging such as in the electron microscope or atomic force microscope it is necessary to shift and rotate the tiles for best alignment, using computer algorithms such as cross-correlation to find the best alignment of the tiles. Even with perfect mechanical devices there would still be a problem if the edges of each tile had slightly different magnification than the center, which is typically the case. If the surfaces aren't flat but have elevation variations that also affect the local image magnification, the complications grow quickly. Constructing large mosaics of images from a microscope, whether it is done with digital camera images or traditional film, is very difficult and rarely produces a perfect result. That application and the methods used to address it are beyond the scope of this text.

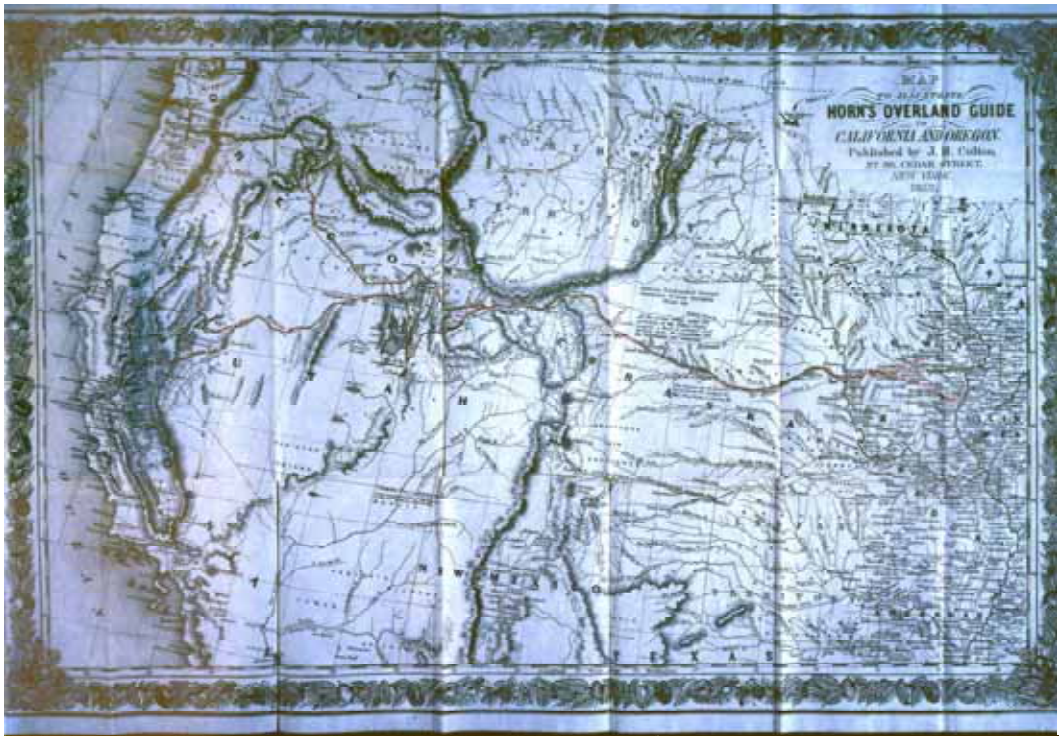


Figure 22. Assembled map image.

The best alignment obtained by rigid shifting of each tile does not bring all of the various points along the common boundaries in **Figure 21** into registration. Tedious manual adjustment of the tiles using the Photoshop Edit-> Transform routines did produce an acceptable map (**Figure 22**), in which the offsets at the tile boundaries and distortions of the map are no worse than those produced by the folds in the original map.

It is sometimes suggested that in the overlap region where two (or more) tiles have pixel information, that blending them together can produce an improved result. The gradient blending procedure varies the fraction of each layer linearly across the overlap region. This is done by selecting the top layer and choosing Layer -> Add Layer Mask -> Reveal All. Then click on the mask in the Layers window and use the gradient tool to draw a gradient on the image which defines the blending of the two layers. The result is shown in **Figure 23**. In most cases, however, this additional effort does not produce an improved visual result. Any slight difference in the alignment of the pixels in the two layers is seen as a blurred final image, which degrades the overall appearance. It is generally better to make the top layer fully opaque and to have a sharp cutoff, even if this results in a few localized misalignments along the boundary.

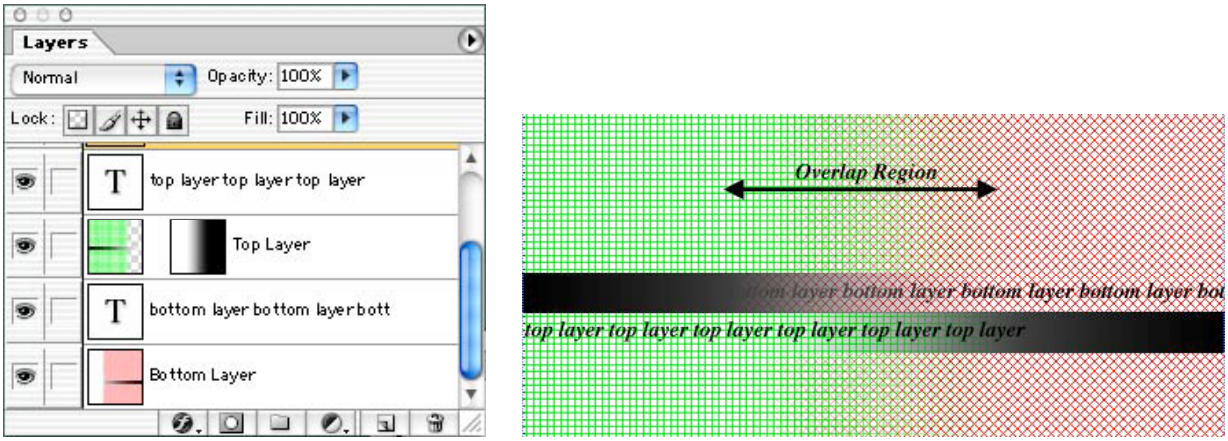


Figure 23. Blending layers in an overlap region: a) Layers window; b) resulting image.

Larger arrays

The original set of separate images included not just the four shown in **Figure 21**, but a 3x3 array with much greater overlaps. In fact, these made the job much harder rather than simpler, and only the four images were ultimately used. **Figure 24** shows the reason schematically. Arranging tiles in a row or column requires a match only along one common edge. In a 2x2 square, each tile must match with others along two edges. In a 3x3 square, the constraints grow rapidly. All of the eight outside tiles overlap along two or three edges, and the central tile is overlapped along all four, plus probably significant overlaps in the four corners. Manipulating and distorting one image to match another, which produces effects that alter the goodness-of-fit with a third and fourth, propagates changes through the set of images that make an overall perfect fit very difficult to achieve.

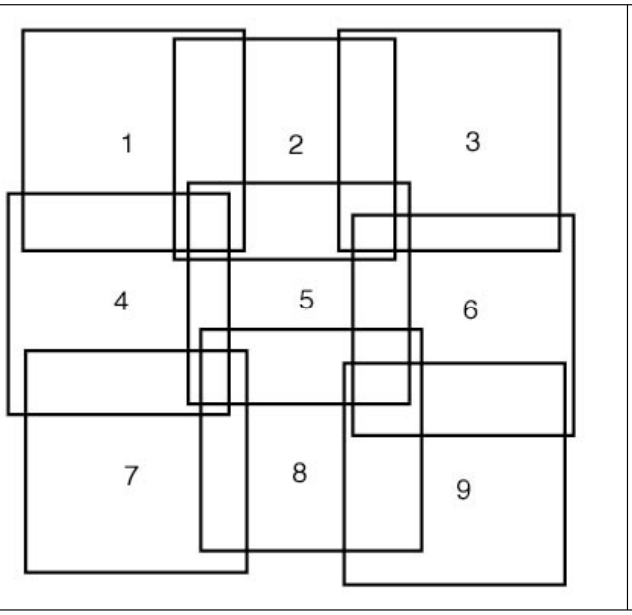


Figure 24. Diagram of a 3x3 mosaic.

Programs that perform the matching and fitting with varying degrees of automation may use either manual marking of common points in the various tiles, automatic matching of points, or a combination

of the two. A guide to the wide variety of these tools on the market can be found at www.panoguide.com. The kinds of distortion that can be applied to the images are typically constrained to match the known behavior of lenses. As for the manual alignment case, the situation is simplest for images that construct a horizontal panorama or vertical column. **Figure 25** shows a typical panorama constructed from a half dozen digital images, each acquired by simply panning the camera on a tripod. A combination of software and hardware (such as an automated panning head for the tripod – see www.kaidan.com) can simplify the construction of a panorama. Notice that each image is distorted into the shape of a barrel to align features along the common edges. There are some minor distortions within the images as well, but as long as the vertical lines remain vertical the overall appearance is acceptable to the viewer. Some stitching programs can specify rectilinear, cylindrical or spherical projection as a way to constrain the bending of features in scenes.

Figure 25. Panorama view of the New York skyline assembled from six digital photos (Quickstitch, Enroute Imaging).

The three vertical tile images in **Figure 26** show considerable barrel distortion (the edges of the tower are curved inwards) as well as perspective foreshortening of the campus bell tower. The distortion is an optical effect and can be corrected. One method is to have calibration data for specific lenses that map the known distortions, while a second approach is to find features that should be straight lines and deduce the lens distortion from them. The latter method was used in the example.

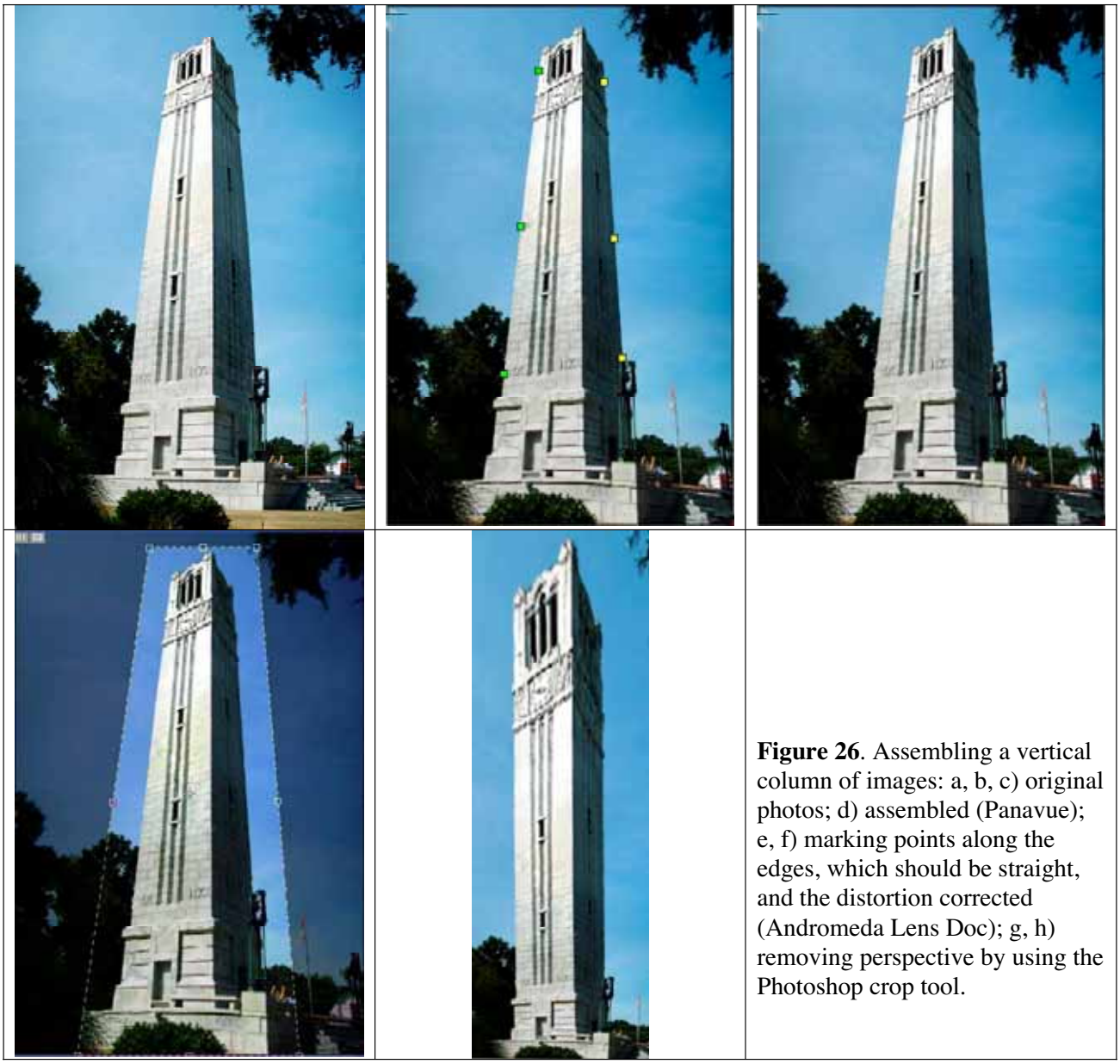


Figure 26. Assembling a vertical column of images: a, b, c) original photos; d) assembled (Panavue); e, f) marking points along the edges, which should be straight, and the distortion corrected (Andromeda Lens Doc); g, h) removing perspective by using the Photoshop crop tool.

After the sides of the tower have been straightened, the perspective distortion of the entire structure remains. This can also be corrected, by using the crop tool in Photoshop and dragging the corners to mark a region outlining to the tower. The final result is a tower with straight vertical sides, but the image is unnatural looking and the top is artificially enlarged compared to the bottom. The problem is that people see things in perspective and do not correctly judge the proportions of the tower when it is shown without foreshortening and perspective.

Figure 27. Perspective correction: a, b) building; c, d) laptop computer.

This type of perspective correction can be useful in other situations, as shown in **Figure 27**. The method is only appropriate when the surface that is foreshortened in the recorded view is planar. Notice in the building image that the side of the building no longer appears perpendicular to the front, and that the sharpness of the image deteriorates at the far end of the building because of the greater enlargement of that region. When the crop region is positioned around the portion of the scene to be rectified, you may enlarge the region by depressing the command/control key and drag the edges to proportionately increase the coverage of the scene.

Assembling multiple images, even without accurate correction of lens curvature or perspective, can be an effective tool to create images with enough pixels and resolution to justify making large hardcopy prints. As pointed out in **Chapter 1**, the resolution of the current crop of digital cameras is much poorer than 35 mm film cameras. As discussed in **Chapter 5**, an image taken with a 5 megapixel chip has less than 2000 pixels of real resolution across its width. Printed at about 200 pixels per inch, which is somewhat less than the human eye can resolve but not seriously blurry in appearance, this corresponds to about an 8 x 10 image. Enlarging the same scene to poster size will result in a pixelated or blurred appearance except when viewed from far away. Combining an array of multiple frames can overcome this limitation.

By mounting the camera on a tripod and panning to collect a series of images that are assembled into a montage, as shown in **Figure 28**, it is also possible to obtain an image that would otherwise call for a very short focal length, wide angle lens. In many real life situations it is not possible to back up far

enough from a building or other subject of interest to capture it all with the modest wide angle lenses provided on most digital cameras. This is not an accident. As pointed out in **Chapter 1**, the requirement that light fall perpendicularly onto the chip makes the design of very wide angle lenses much more difficult.

Figure 28. Automatic assembly of a series of images into a montage (Quickstitch, Enroute Imaging).

Foreground and background isolation

Careful examination of **Figure 28** reveals that the person in the foreground moved slightly between the exposure of the two tiles in which he appears. This has created a distorted image of the person that is acceptable in this instance, but highlights a potential problem with montage construction. It is important

to acquire all of the images in a sufficiently short time that the various portions of the scene still belong together. The motion of people, clouds, vehicles, etc., or the change in lighting conditions, can alter or damage the appearance of the resulting montage.

Multiple images of the same scene can be used to isolate or remove people (or other moving objects) from the foreground. The example in **Figure 29** shows two photos of the entrance to a building taken at different times. Because it is a busy location, getting a photo without people present is impractical. But in a series of such pictures each part of the building front is captured eventually. It would be possible by manual cutting and pasting to assemble a montage of fragments of the building that had no people present, but it is more instructive to see how this can be done automatically using various Photoshop tools.

In this case, the two images were combined as layers using the Difference blending mode. The resulting image is nearly black (no difference) where there are no people, but much greater wherever someone or something is present. This difference image can be thresholded (Image -> Adjustments -> Threshold) to select the pixels where there is a large difference, and that black and white image will become a mask to blend the images together. As shown in the figure, the mask was blurred to produce feathering where pieces of different images will be blended together. The resulting final image eliminates all of the people and produces a “clean” image of the building.

Figure 29. Removing foreground objects: a, b) two views of the same scene at different times; c) blending images a and b using Difference mode; d) thresholding image c to create a blending mask; e) the foreground objects selected by the mask; f) the scene without the foreground objects.

Similar logic can be used to isolate the foreground objects. This is particularly useful in surveillance work. Typically, an image showing a person in a surveillance scene is quite noisy and often has poor lighting. Enhancing the view of the person alters the surrounding scene in unrealistic and visually bothersome ways. Comparing the surveillance scene to another view of the same scene taken shortly before, without the person of interest, allows creating a mask of just the changed area (in fact, some surveillance systems perform this comparison continually and only save the images when someone or something has moved in the field of view). Expanding the contrast in just the masked region while leaving the rest of the scene unchanged improves the visibility of the subject, as shown in **Figure 30**.

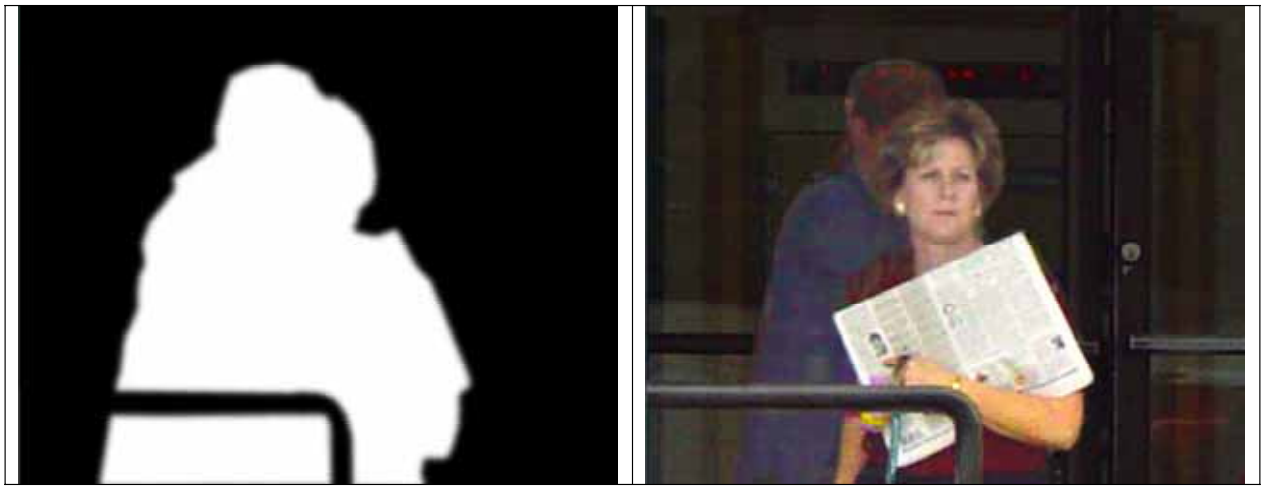


Figure 30. Enhancing the foreground object: a) static background scene; b) subject in scene; c) mask of pixels that have changed from static background (with feathered edges); d) histogram equalization of changed pixels highlights the foreground subject.

Chapter 4
Photo Post-Processing

After the image has been acquired, and global contrast and color adjustments made, it may be necessary to apply various post-processing or editing steps. Some of these are inherently manual and interactive, such as touching up blemishes, removing dust marks, etc. Many of them can be performed algorithmically using routines built into Photoshop or added to it as Filter plug-ins.

There are a variety of powerful methods for correcting image problems such as scratches or speckle noise, and for enhancing the visibility of detail. Since the goal of these operations is to improve the visual appearance of images, it is very useful to understand the human visual system's response to detail, contrast and color. It may also be interesting to know what the underlying algorithms for some of the more useful methods are, and they are explained in the following sections but the reader can skip the technical explanations in favor of examples if that is preferred.

Both manual and automatic operations benefit from the use of regions to limit the area of application, and layers to blend the processed results into the original. The selection of regions (manually or based on color criteria) and the layer blending options need not be confusing if a few fundamentals are understood.

Visual enhancement

It has been pointed out in earlier chapters that the human visual system has characteristics unlike a camera, and employs its own image processing methods that extract information from scenes to pass on to conscious layers of the mind. In order to design techniques for improving the visual appearance of images, it is helpful to understand some of those characteristics and to take advantage of them. Human vision is capable of operating over an extreme range of absolute brightness levels, from a sunny day at the beach to a starlit night (although in the bottom portion of this range we lose color sensitivity and see only in shades of grey).

But we cannot resolve very many shades of brightness in a single scene. A change of a few percent in brightness, occurring over a short distance, is visibly detectable. A much larger change that occurs gradually is not: consider the lighting in a typical room - the shading from locations near a lamp or window to a dark corner is not noticed unless we concentrate on it, and even then does not interfere significantly with our ability to judge contrast and color. In **Figure 1**, the area of wall shadowed by the roof varies in brightness by a factor of 2 - some of it is darker than the roof and some lighter - but we see it as a continuous wall and "know" that it is the same color. The horizontal lines where the siding overlaps are in some cases only changes of 5% in brightness, but because they are very narrow and because we "expect" them to continue from other locations in the image, we see them easily.

Figure 1. Changes in brightness that occur gradually are largely ignored, those that are abrupt are highly visible.

Abrupt changes in brightness are highly visible, and gradual ones less so, because of the way our retinas work. Within the first few layers of neurons that collect the signals from the light-sensing rods and cones, are comparators that look for differences between one point (not quite the same thing as a "pixel" on the screen, but enough like it for now) and its immediate surroundings. This is called inhibition, and it means that the higher levels of the visual cortex and brain only get signals where a local change is present. Similar comparisons over time help us to sense motion. There are no circuits that compare over long distances, so the gradual changes are not detected.

The example in **Figure 2** is a familiar one in many textbooks on human sensory perception. The vertical bands of grey are absolutely uniform in brightness, but most people see them as having variation - they get darker next to the lighter neighboring band and brighter next to the dark one. That is inhibition at work. Since the visual system is so attuned to finding brightness steps and interpreting them as the edges or boundaries of objects or detail, one common technique in image processing is to increase the contrast at these locations.

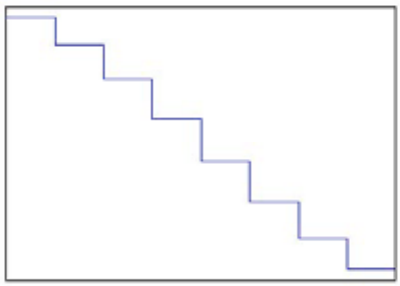
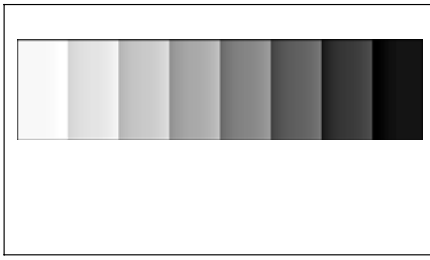


Figure 2. Grey scale bands are perfectly uniform but the human eye increases contrast at the steps.



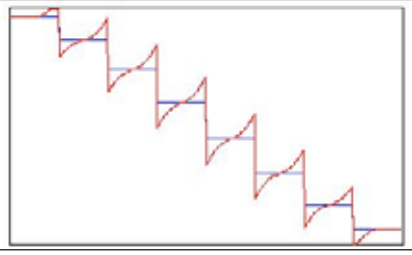


Figure 3. Applying an unsharp mask to **Figure 2** adds contrast at the steps and increases their visibility.

Figure 3 shows the result of applying an "unsharp mask" to the bands from **Figure 2**. The unsharp mask, discussed in more detail later in this chapter, is one of the most used (and abused) processing tools. It is designed to increase the contrast at steps, just like the human eye does. The gradual variations within each band are largely ignored, but the sharp ones at the steps make the edges appear even more visible, and makes images look "crisper."

Color sensitivity

Color is very important in human vision, and the ability to distinguish and recognize features. But, like brightness, it is only judged by comparison to other colors. There is, by the way, a set of Photoshop plug-ins (Insight from Colorfield Digital Media, www.colorfield.com) that simulate the effects of various kinds of color blindness on images, which may be very useful in determining how they will appear to persons (mostly male) who suffer from those handicaps.

There are three kinds of cones in the human eye, packed densely into the fovea, the central spot that we use to "look at" features of interest in a scene. They are often called "red," "green," and "blue" sensitive but in reality each type covers a broad and overlapping range of wavelengths as shown in **Figure 4**. We are most sensitive to green light, least to blue. The rod cells have lower spatial density but are highly sensitive to a broad range of wavelengths (but not to red, which is why we can use red flashlights at night without destroying our sensitive night vision that uses the cones).

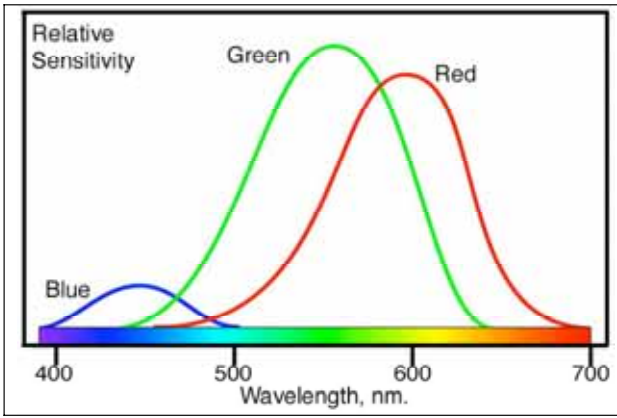


Figure 4. Relative sensitivity of the red (long wavelength), green (medium wavelength) and blue (short wavelength) cones in the human eye.

While our sensitivity to green light intensity is very high, we are not very sensitive to small color changes in the green part of the spectrum. It is in the transition range green-yellow-orange-red (where two of the cones have sensitivities that overlap extensively) that changes in color are most strongly perceived, followed by the red-violet-blue range produced by excitation of the red and blue cones but not the green. This means that the greatest color contrast in images will arise when those colors are present. Since it is color differences that are most important, and we are rarely concerned with or able to

accurately judge actual colors, altering the color balance in images (as discussed in **Chapter 2**) can increase visual contrast by shifting the hues slightly.

Of course, this can also be done locally in the image. In the example of **Figure 5**, the colored autumn leaves are distinguished from the green leafy background and from each other. It is easy to increase this difference and with it the visibility of the leaves by using a selection based on color (selections are discussed later in this chapter) to limit changes to pixels in the yellow-to-red range, and then modify the adjustment curves to increase the contrast and saturation of those colors. The blue water and sky, and the green leafy areas of the trees, are not affected by this change. The result is to place more exciting colors into the image and to use them to draw the viewers attention to the central region where they appear.

Figure 5. Changing leaf colors: a) original image; b) applying adjustment curves only to the red-orange-yellow pixels.

One of the very interesting things about color, which lies at the heart of image compression algorithms and broadcast television, is that our visual spatial resolution for color is not as good as it is for brightness. Intentionally degrading the spatial resolution of the color values reduces the amount of data that must be stored or transmitted, and in most cases does not affect our visual impression of the scene. But it goes back much farther than that - children don't always color inside the lines, and many artists have also intentionally allowed colors to extend across boundaries (**Figure 6**). As long as there are lines or intensity steps present in the image to define the actual boundaries of objects, this color bleeding doesn't bother us visually. Examine your television image closely and you will see the same thing - the colors bleed across boundaries without affecting comprehension at all.

Figure 6. A child's drawing. Colors extend beyond the black lines that locate edges.

It might be imagined from the preceding discussion of the inhibition theory of vision that this approach to edge sharpening as a means of visual enhancement could also be successfully applied by manipulating the colors near boundaries. It doesn't work that way. As shown in **Figure 7**, increasing the change in brightness at each step with an unsharp mask in the intensity channel, while leaving color information unchanged, does make the boundaries appear sharper. Doing the same thing in the hue channel, while leaving the intensity the same, merely introduces visible color shifts that are confusing and not helpful.

Figure 7. Increasing edge steps by local processing: a) original image; b) processing just the intensity channel; c) processing just the hue channel.

Another way to demonstrate how little we rely on changes in color to define feature boundaries in images is shown in **Figure 8**. The original image is blurred so that all edge and detail definition is lost. The resulting color splotches are applied to the original image by using "color" blending layers mode, which replaces the colors in the original image but keeps the original brightness values. The color shifts and bleeding of colors from one flower to another are evident, but they don't have any affect on the ability to decide where each flower's petals end.

Figure 8. Blurring colors: a) original image; b) heavily blurred (a Gaussian blur with 40 pixel radius); c) applying the colors from image b to image a.

Color resolution

In a typical scene, because of the eye's ability to distinguish a brightness change of a few percent, only some twenty to thirty shades of brightness can actually be resolved. This may seem like a very small number, when the typical digital camera stores images with 256 shades of brightness, and photographic film can resolve several thousand shades (but photographic prints can not - the typical print is not much better than vision requires, and some cases such as the early Polaroid black-and-white prints had less than 15 distinguishable grey values, but were still considered very useful). But apparently this limited intensity resolution is enough for us to interpret what we see, even in scenes that cover enormous ranges of brightness from glare to shadow because the values are compared locally, and concentrating our view on one area allows our vision to rapidly adjust its sensitivity to the actual range of values present there, ignoring other brighter or darker regions.

Using an approximate value of 30 shades of brightness resolution, if the argument is made that this should mean the ability to resolve 30 * 30 * 30 = 27,000 colors (thirty shades each of red, green, and blue) then the prediction is that a very large number of colors should be distinguishable. They are not. In typical viewing situations, the actual number is a few thousand. This information is also important for image compression, because it means that the actual colors in a scene can be greatly reduced for storage.

The 256 colors displayable on early generations of colored computer displays is clearly inadequate for realistic viewing. A display with thousands of colors (5 bits each or 32 intensity levels for red, green and

blue) should by that argument be enough - and it would be if that number of colors could be distributed properly across the gamut of color space. But the linear intensity steps in the display hardware produce too great a change in brightness for relatively dark values, and these can be visually perceived. So "photorealistic" computer displays typically use 256 shades each for red, green and blue, and human vision does not detect any of the steps that are present from one shade to the next. The color wheel in **Figure 9** has about 4000 separate color blocks, not all of which can be distinguished visually.

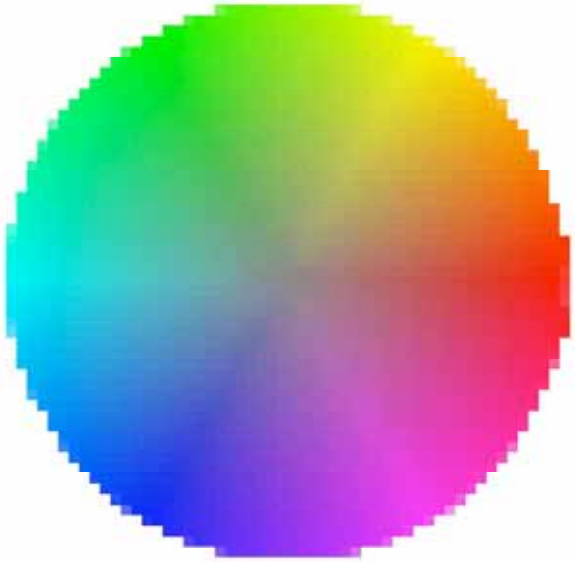


Figure 9. The color wheel shown as an array of color blocks.

Actually, a more limited display with thousands or even hundreds of discrete colors can be made to appear better by introducing dithering, in which groups of display pixels in uniform areas are given different colors so that at normal viewing distances they are visually averaged to create the impression of intermediate colors, and the sharp steps from one color to another are broken up so they are not visually noticeable. This technique will be discussed further in the context of printing, in **Chapter 5**.

In a typical scene or photograph, the presence of local contrast - the juxtaposition of regions with different brightness or color values - attracts the eye. Edges, lines and steps are visually interesting, and it is there that we focus our attention. Cartoons and sketches work because we look at the lines, recognize them as corresponding to boundaries that occur in real images, and accept them as a useful shortcut to extracting the very parts of the scene that our eyes would find and transmit to the brain.

What this means for image processing and enhancement is that anything that increasing the visibility of contrast, lines, edges, steps or texture will improve the image (or at least make it more interesting). Similarly, the suppression of contrast, lines, edge and steps or texture that are not part of the interesting part of the image (based on the photographer's intent) will make them less visible. Uniform or gradually varying (e.g., blurred) regions, and dark regions, do not attract attention. This can be used to reduce or remove defects and suppress background, while making the subject of the photograph dominate the scene.

The camera specifications discussed in **Chapter 1** note that a typical digital camera records images with 8 bits each of red, green and blue values (the usual 0..255). That does not mean there is that much color or intensity resolution. The histogram of a region in an image that is all the same color or brightness

does not show that all of the values are identical. The width of the peak in that histogram is a measure of the actual resolution, and as pointed out in **Chapter 1** that depends on the noise characteristics of the camera and electronics, and may not be the same for bright and dark regions, or in the different color channels. In general, the recorded image will not have more than a few tens of thousands of meaningfully different color values recorded. By no coincidence at all, this is similar to the capabilities of the display and the requirements of the human eye.

Spatial resolution

Based on the density of cones in the fovea, and the optics of the human eye, it is possible to resolve distances of about 0.1 mm at close viewing distances. That means that two points that far apart can be distinguished. It does not mean that smaller features are invisible, provided they have enough contrast as compared to their surroundings. A star in the night sky covers a very tiny angular arc, but because of its brightness can still be seen. But two stars close together can only be separated with a telescope.

Computer video monitors do not have a dense enough array of phosphor dots to meet this criterion - a typical value is 0.25 to 0.3 mm for the dot pitch. And sure enough, if you look closely, you can see the individual colored dots. But at a normal viewing distance this is adequate. It corresponds to the typical values for displays of 72 to about 90 display pixels per inch that are used on computer displays. If you have a typical medium sized monitor, say one that is 17 inches across and able to display 1280x1024 pixels, that is 75 pixels per inch (=1280/17).

If your camera takes 5 Megapixel images, 2560x1920 pixels in size, that image can't be seen in its entirety on your display. You can zoom it up to 1:1 and scroll around the image to see parts of it at a time, or you can shrink it to "half size" and display only one fourth of the stored pixels - every other row and column. The values from in-between points aren't shown because there is no place on the display to put them. Fine lines may disappear, or even more disconcertingly may be seen in some places and not others, depending on whether they lie on even or odd addresses. At other degrees of reduction, other artefacts and aliasing will occur.

Of course, the situation is much worse for hardcopy printing. That is discussed separately in **Chapter 5**. But as a simple sanity check, a good 35 mm film negative can resolve 4-5000 points across its width, and a good photographic print can resolve about 1000. So digital cameras are not capable of recording images that preserve as fine a detail as film cameras. In practical terms, that means it is necessary to use close-up pictures to record fine detail, or perhaps build a mosaic from multiple images as described in **Chapter 3**.

Another consequence is that as resolution is limited, fine detail in the image is lost. Since it is the detail, lines and texture that are important, the inability of the camera to capture it may cause regions that appear visually to be interesting in the original scene to lose that interest in the photograph. In **Figure 10** the individual hairs on the dogs coat are not resolved, and this becomes more severe as the number of pixels in the image is reduced. As the texture is lost, so is much of the information we use to judge the shape of the dog. The few highlights that are present can be enhanced in visibility to restore some of the flow of the hair, but the scale is much coarser than the individual hairs and the appearance is not very natural.

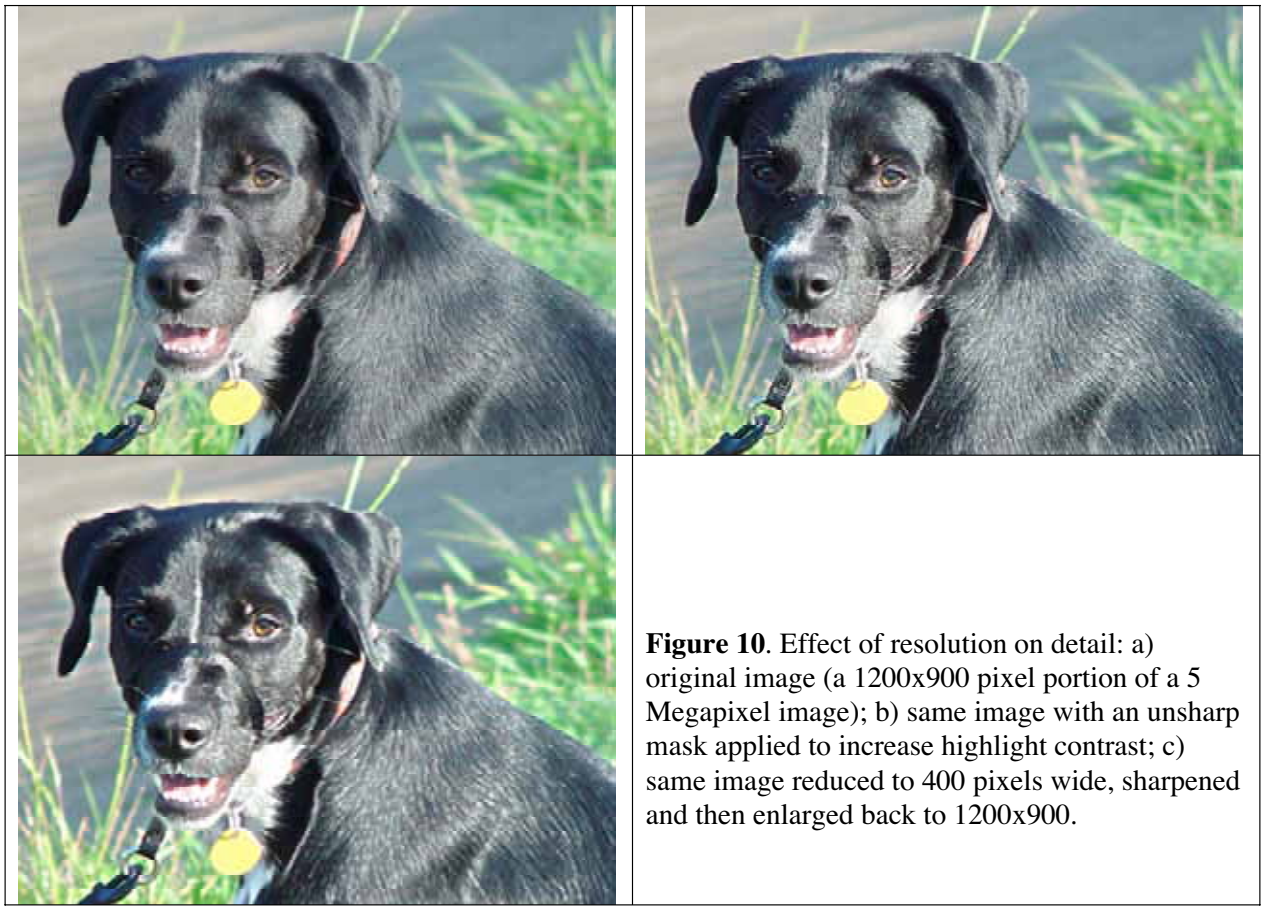


Figure 10. Effect of resolution on detail: a) original image (a 1200x900 pixel portion of a 5 Megapixel image); b) same image with an unsharp mask applied to increase highlight contrast; c) same image reduced to 400 pixels wide, sharpened and then enlarged back to 1200x900.

Since the prospects for having digital cameras that can approach the approximately 150 million pixel resolution of film in a typical studio camera or of the human eye are not very good, the detail that we would be able to observe directly viewing the scene or in a photograph from a 4x5 film back just will not be captured in the digital image (and the poorer dynamic range of the digital cameras is an even more serious limitation in scenes with large contrast ranges). It is by the effective use of computer processing to enhance local detail and suppress unwanted defects that we can make the digital images useful and acceptable as realistic and interesting representations of the scenes to a human visual observer.

Lighting gradients

Shading or brightness gradients in images may be visually ignored because, as noted above, the human eye doesn't particularly notice gradual changes in brightness. But we do tend to overlook the darker areas of photos, so one reason to correct for shading is to give all parts of the scene the same average brightness. When we look at the dark region of an scene we can adapt our vision to bring out the detail. But the dynamic range of a digital photograph is not great enough to retain the subtle details in extreme light and dark areas. Correcting for shading or brightness gradients can allow enough dynamic range for enough local contrast to make detail visible everywhere.

In a studio setting, the placement of lights controls the levels of illumination, and it is possible to arrange them so that all of the interesting parts of the scene have adequate light intensity. It may even be

desirable to have the background areas darker, to focus attention onto the foreground. If it isn't practical to place all of the lights needed to illuminate the subject perfectly, acquiring multiple images with different lighting and combining them as discussed in **Chapter 3** may provide an effective solution.

Particularly in macro photography, it can be very difficult to arrange uniform lighting. Ring lights around the lens, or an array of floods on both sides, can solve the problem in many situations. But sometimes, even with the best practical effort, the recorded image will have a brightness gradient. Also, optics can contribute to the problem. In many situations a mismatch between lens, apertures and chip can produce vignetting, in which less light reaches the edges and corners of the chip and the image is darker away from the center. Many of the less expensive consumer cameras use extremely low cost lenses that often suffer from vignetting problems, especially if limited capability optics are asked to provide a zoom capability (these lenses also have out-of-focus and distortion problems away from the center of the image).

For whatever reason vignetting or side-to-side brightness gradients may arise, if they can't be eliminated before the image is taken it may be necessary to correct for them afterwards. The most straightforward way to accomplish this is to take a second image with the same lighting and camera settings, but without the subject being photographed. In place of the subject, a uniform grey card is used. The image of a standard "18% grey" test card will record the lighting variations on the subject and can be used to level the image brightness and remove vignetting or gradients.

The name "18% grey" may be misleading - because of the logarithmic response of the eye, film cameras, and many digital cameras, this card will actually be seen as a medium grey and should ideally correspond to stored values halfway between white and black (in other words, about 128 on the 0..255 brightness scale). Since it is a neutral grey, it should have equal intensities for red, green and blue. Some copy stands incorporate a grey card in the base, so simply removing the macro object being photographed permits recording the grey background image.

Figure 11 shows an example of rather severe nonuniform lighting (one of the floodlights on the copy stand was turned off). Removing the necklace and capturing a second image with the same camera settings produced a background picture that recorded the texture in the cloth used as a background. To eliminate this from the background shading image, a large Gaussian blur was applied. This image was then divided into the original and the results automatically scaled to the full range of the display (using an autoscaling general image math function in the Fovea Pro plug-ins, www.ReindeerGraphics.com). If the background variation was smaller, no more than about 10% of the full dynamic range of the image, subtraction using the built-in Photoshop Image->Calculations function would have been a sufficient approximation, but the proper way to remove the gradient is to use the ratio of the original image to the background.

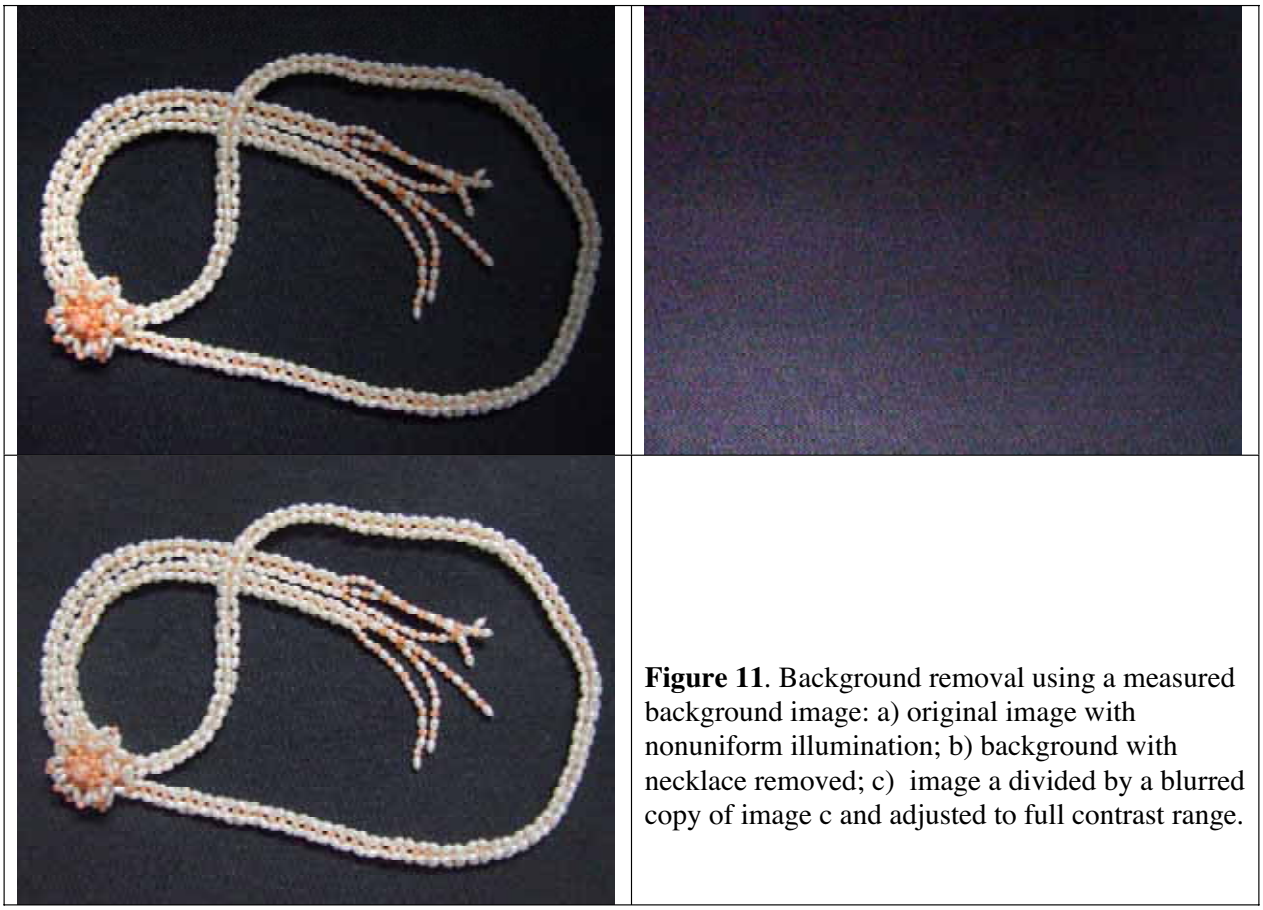


Figure 11. Background removal using a measured background image: a) original image with nonuniform illumination; b) background with necklace removed; c) image a divided by a blurred copy of image c and adjusted to full contrast range.

(It sometimes occurs to people who dimly recall some of their high school algebra that there might be an alternative way to accomplish the division by using the Photoshop Image->Adjust->Invert command and then the multiply option in either the Image->Calculations dialog or in the layer blending mode. They are thinking that dividing by X is the same thing as multiplying by the inverse of X. But, alas, that doesn't apply here. Inverting an image doesn't calculate the inverse in the sense of 1/X but rather inverts the image contrast by subtracting the pixel values from 255. Sorry about that...)

Because the background image has a brightness range that covers about 25% of the full 0..255 dynamic range of the original image, the resulting image has lost some of its brightness resolution, but by expanding the remaining 75% to cover the full range (using the Adjustments -> Levels function) the image is still quite satisfactory. If the background variation becomes too large, for instance 50% of the dynamic range, it creates problems because the foreground objects have too little contrast remaining, and some regions may be clipped to full white or black.

Generating a background

Obviously, the grey card method is a lot harder to implement in an outdoor scene. It may be possible to locate features in various parts of the image that should have the same brightness and color (for instance, the shirts on team members in a group picture taken with lighting from the sides that makes the center darker). If several such locations can be found, they can be used to construct a background image.

In the example of **Figure 12**, the brightness values on the shirts at the left and right sides of the image were measured and used with the manual gradient tool to construct a ramp. This was done in grey so that the colors in the original image would not be altered. Then the background ramp was divided into the original image, pixel by pixel, to remove the shading. This reduces the dynamic range of the image, but after increasing the contrast top cover the full display range using the Adjustments->Levels function, the result is an image that does not show the shading.

The built-in Photoshop calculations routine offers subtraction but not division. For a relatively small amount of shading, as in this picture, subtraction of the gradient produces a nearly identical result and can be used satisfactorily. When the extent of shading is greater, it becomes important to construct the new image correctly as the ratio of the original image to the shading. Plug-ins such as The Image Processing Tool Kit and Fovea Pro (www.ReindeerGraphics.com) offer ratioing with optional automatic scaling of the result.

Figure 12. Removing shading from an image: a) original image; b) linear ramp constructed with the manual gradient tool, using brightness values measured on shirts at the left and right sides; c) image a divided by image b (pixel by pixel), with levels adjustment to increase contrast.

More complex backgrounds (using polynomials instead of linear ramps) can be constructed automatically by selecting multiple regions in the image and using the Background Fitting and Background Removal plug-ins in Fovea Pro (www.ReindeerGraphics.com). **Figure 13** shows using the Select -> Color Range dialog to select all of the yellow areas on the boys shirts in the original image of **Figure 12a**. The automatic fitting of a background function to these pixel values and its removal produces the corrected image shown.

Figure 13. automatic shading correction: a) selecting the yellow shirts with the Select -> Color Range function; b) removal of the background fit to the selected pixels.

Of course, sometimes a brightness gradient is desirable to focus attention on a foreground object and away from the background. This is easy to accomplish. Create a duplicate layer and select the gradient tool. Click on the center of interest (usually right between the eyes for a person or animal) for the white point of a radial gradient, and extend the gradient to the farthest corner of the image for the black point. Set the layer blending mode to multiply and adjust the opacity as desired to control the strength of the vignetting. Just as we use division to remove shading, we use multiplication to apply it. **Figure 14** shows an example.

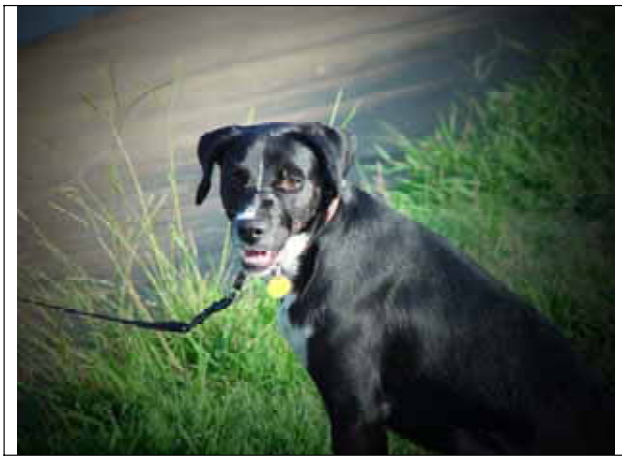


Figure 14. Creating vignetting: a) original picture; b) radial brightness gradient layer; c) resulting image.

Sharpening edges

The discussion above of human vision emphasized the importance of detecting edges and boundaries. This primary step is sometimes called the formation of a primary sketch from the image, highlighting the relationship between sketches and image understanding. The first few layers of neurons, right in the retina of the eye, use comparison between pixels and their neighborhood to find the edges, boundaries and detail that convey important information about the scene to higher levels of the visual cortex. Processing the stored digital image to help the eye find and extract this detail is one way that we can make images (or parts of them) more visually interpretable and more interesting. This is sometimes called "sharpening" the image, and sometimes is described as "edge enhancement."

The Unsharp Mask filter in Photoshop is the most widely used tool for accomplishing this purpose, but it is not the only one, nor always the most appropriate. By way of understanding what it does, and why it has such a counter-intuitive name, it may help to recall the original use of the method, in the photographic darkroom. Early film and print materials had a more limited dynamic range than most modern types, which made it difficult to capture the full range of contrast in prints, or the fine detail in both bright and dark regions of the image. This is still sometimes an issue since photographic prints (and computer hard copy, as discussed in **Chapter 5**) still have a narrower dynamic range and less tonal resolution than film or many digital camera images.

In order to use more of that range and resolution for the details in the image, at the sacrifice of some of the global contrast between bright and dark regions, film photographers could produce an "unsharp mask" by printing their image onto a piece of film material, at 1:1 magnification but slightly out of focus (hence "unsharp"). When developed, this film was a positive print, dense in regions that the negative was thin and vice versa. Sandwiching this "mask" with the negative and printing through both produced a final result in which the overall contrast between bright dark areas was greatly reduced, but in the regions around edges and detail the out-of-focus mask allowed the fine detail on the original negative to print through. The result was an image that used more of the print's tonal range for the detail and less for the overall contrast.

It is easy to demonstrate this same process with a digital image, as shown in **Figure 15**. A duplicate of the image is blurred using the Filter->Blur->Gaussian Blur function. The radius is adjusted to be just large enough to remove all of the detail that is important, typically a radius value of a few pixels. This blurred image still has the overall contrast between light and dark regions, but the fine detail and edges have been blurred. Subtracting this image from the original accomplishes the same thing as sandwiching

the negative and unsharp mask together and printing through both of them. The result is the difference between the two images, which is just the fine detail and edges that were removed by the blurring. This detail is then added back to the original image to highlight those edges and that detail, which increases the visual sharpness of the resulting image.

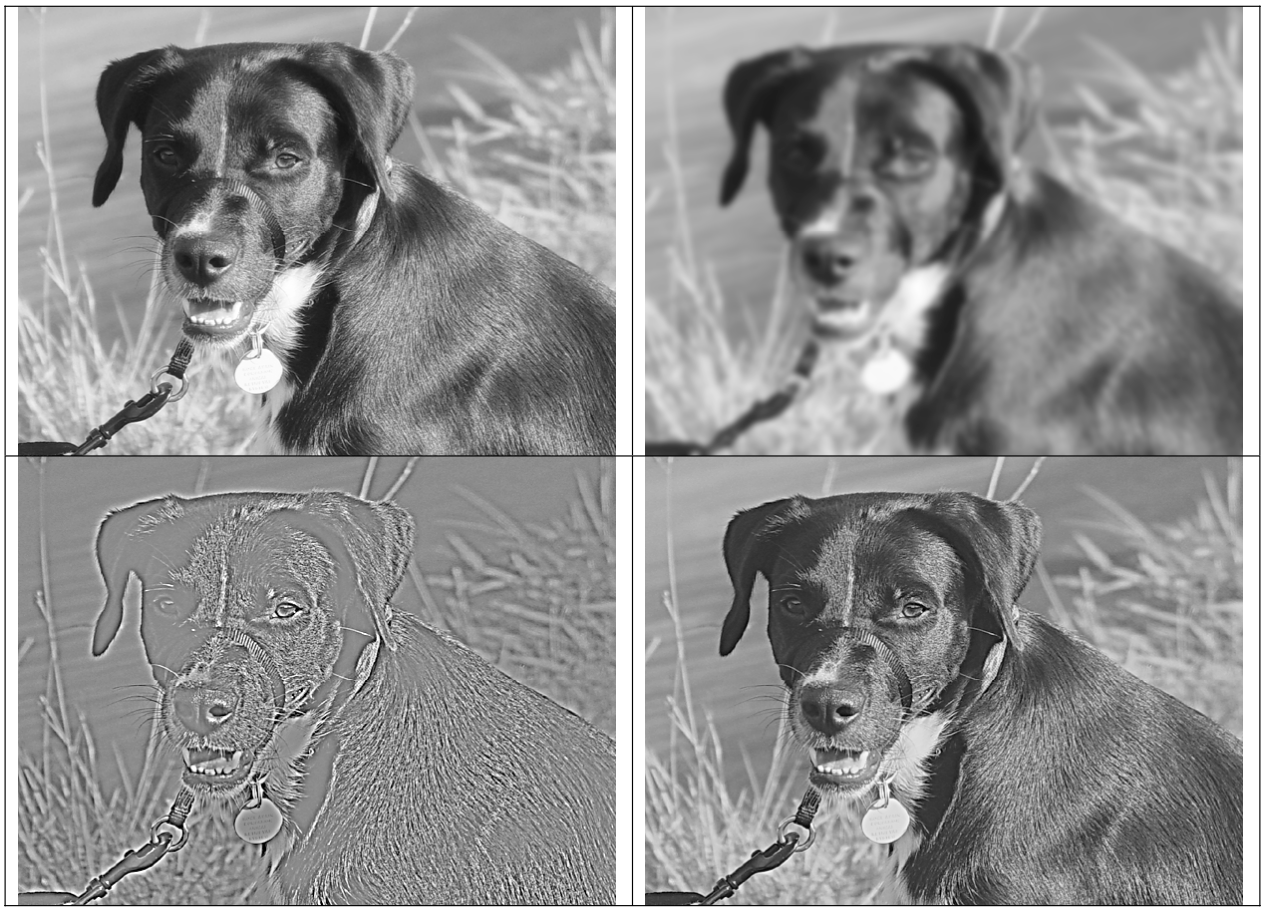


Figure 15. The Unsharp Mask process: a) original image; b) blurred duplicate; c) image a minus image b; d) image a plus a percentage of image c.

The key variables in this process are the amount of the blur, and the amount of the detail image to add to the original. The Photoshop Unsharp Mask filter has sliders for both of those controls (**Figure 16**). The third control allows selection of a threshold value for how much difference there must be between a pixel and its neighbors to be considered detail and kept after the subtraction process. Increasing the value will add sharpness to only the major contrast steps and not the places in the image where the detail has less contrast. Understanding what the sliders do in terms of the process will help the user to manipulate them to a successful result, but the criterion for selecting the “best” settings for any particular image is a visual one. If the detail that is of interest to the viewer is enhanced, then the image will have a better visual appearance and the operation is beneficial.

Figure 16. The Photoshop Unsharp Mask dialog, showing controls for the amount of the edge image to add to the original, the amount of blur to be applied, and a threshold for which edge detail to keep.

It is important not to make the blur in the unsharp mask too large. That increases the depth of the shadows or halos around edges, which increases their visual impact, but if the shadow from one edge or detail covers another, lesser edge it hides it and reduces the overall sense of detail and sharpness. **Figures 17** and **18** show examples. In **Figure 17** the spacing of the rock layers is fairly uniform and so a good setting of the unsharp mask improves all of the detail, while a larger setting degrades the image. In **Figure 18** the detail occurs on many different size scales, and the best blur setting is the one that corresponds to the finest detail that is to be kept.

Figure 17. Varying the amount of blur: a) original image; b) good setting (2.5 pixels); c) too large a blur (9 pixels).

Figure 18. Varying the amount of blur: a) original image; b) good setting (2 pixels); c) too large a blur (10 pixels).

Color and Noise

Applying an unsharp mask to a color image by processing each of the red, green and blue channels creates problems. The boost in edge contrast is a function of the pixel values on each side of the boundary, and except for neutral grey colors this is different in each color channel. The consequence is that the amount of change in each color is different, and this results in their combination producing a new color, or hue, that was not present in the image originally. Highlighting red areas with orange or purple, for example, creates a distraction for the viewer.

Noise in the image is also a problem. One common type of noise is simple random speckle superimposed on the image. As discussed in **Chapter 1**, it arises from statistical variations in collecting

the electrons in the chip, amplifying them to produce a signal, and measuring the voltage produced, and is unavoidable. Lower light levels (requiring a higher effective ASA rating and more amplification) increase the amount of this type of noise. The unsharp mask responds to a difference between a pixel and its surroundings, but that is exactly what this type of noise is. Consequently, the visibility of random noise is increased by an unsharp mask, as illustrated in **Figure 19**.

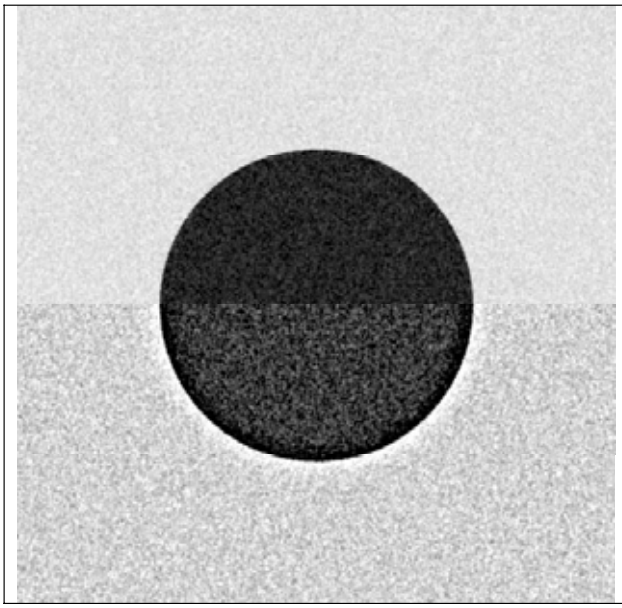


Figure 19. The visibility of the random speckle noise superimposed on this simple image of a circle is significantly increased by the application of an unsharp mask to the bottom half of the image.

But in a color image, the noise in each channel is largely independent of the others. There is usually the least noise in the green channel, because the most common sensor design (the Bayer pattern discussed in **Chapter 1**) uses more detectors for green and thus generates more signal. The blue channel typically has the most noise, because the silicon-based chips are not very sensitive to blue light (neither is human vision). Because the noise fluctuations in each color channel are different, at one pixel location the red value may be high and the green low, while at another the green may be high and the blue low. Each of these differences is amplified by the unsharp mask. When the variations in the different color channels are viewed together, they produce new colors, not present in the image originally. This colored "snow" is very distracting and degrades the image considerably.

As shown in **Figure 20**, the unsharp mask works much better when applied not to the individual color channels, but to the intensity values only. This can be done in Photoshop by converting the image from RGB to L*a*b color space (Image -> Mode -> Lab), selecting just the Lightness channel, and performing the processing on that channel before converting the image back to RGB. This leaves the color values unchanged, so the random speckle noise in the various channels is still present but does not increase or give rise to new, distracting colors in the processing.

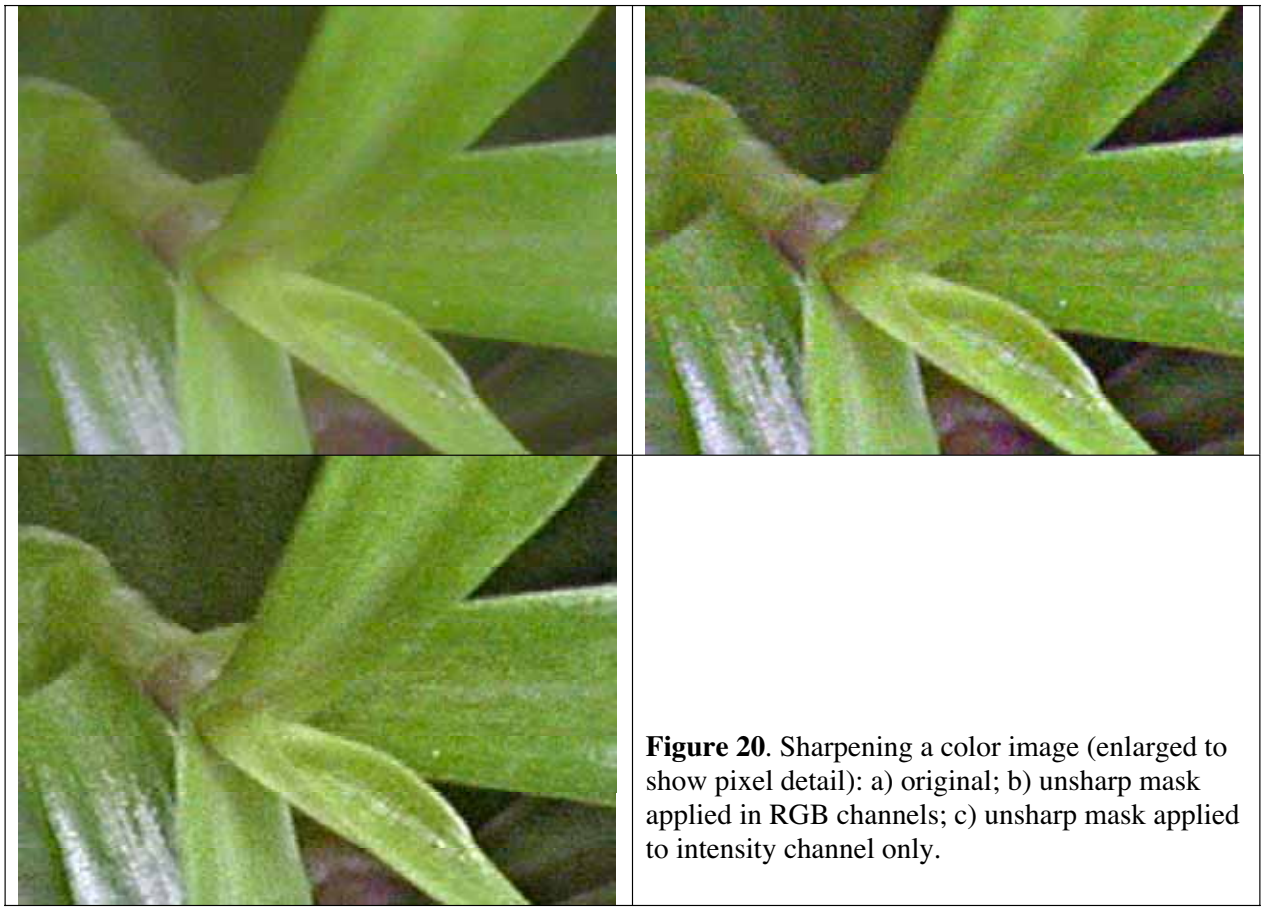


Figure 20. Sharpening a color image (enlarged to show pixel detail): a) original; b) unsharp mask applied in RGB channels; c) unsharp mask applied to intensity channel only.

Filtering the intensity channel only is a very general approach - almost all image processing operations should be applied to the intensity channel of a color image, not to the individual RGB channels. The exception is the reduction of noise, discussed below, which should be performed on each channel since that is where the noise arises and since it is largely independent in each channel.

Laplacian and Difference-of Gaussians filters

The unsharp mask should be understood as one member of a family of operations that compare pixel values to find differences. On the very smallest scale, comparison of each pixel to its immediate neighbors (called a Laplacian operation in the image processing literature) is carried out with a filter as shown in **Figure 21**. The Photoshop “custom” filter is a powerful tool for understanding how filters work and experimenting with various rules for combining a pixel with its neighbors.

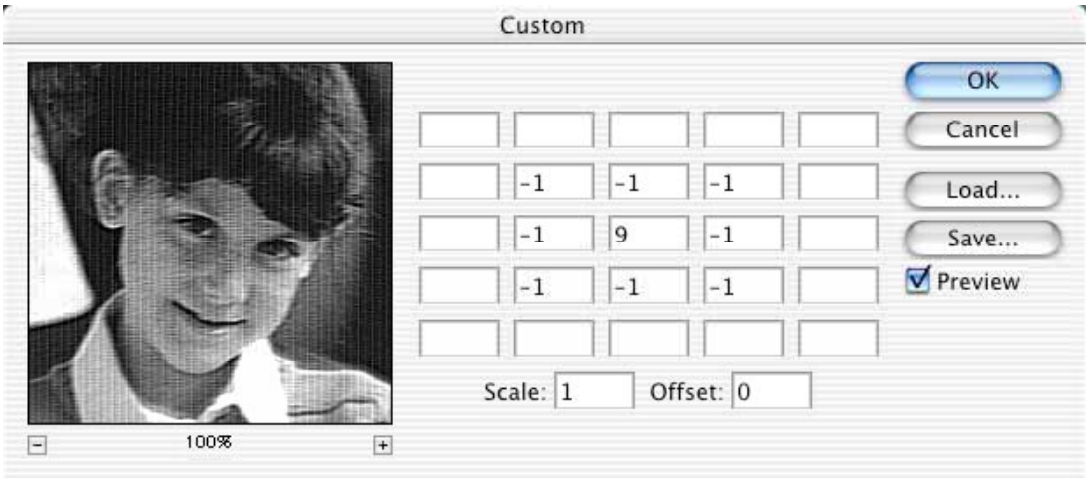


Figure 21. The Photoshop "custom" filter with weights entered that apply a classic Laplacian filter.

In the example, the process is to multiply each pixel value by 9 and then subtract each of the 8 neighbor pixels. In a region where all of the values are nearly the same, this leaves the original value nearly unchanged. But if there is a difference in brightness in any direction it produces a larger result with the effect that a pixel brighter than its neighbors becomes brighter still, and vice versa. That is the same effect produced by the unsharp mask, and it is easy to imagine the ring of "-1" values in the kernel of weights as the blurred mask that is subtracted from the sharp original. The Laplacian is a very good noise finder because it finds differences between immediate neighbors.

In contrast to the tiny comparison shown in the Laplacian, in the human eye there is a hierarchy of comparisons that find differences between a central region and its surroundings on many different scales. The filter that performs the same type of operation in the computer is a generalized version of the unsharp mask. Instead of subtracting a blurred copy from the original, we can make two blurred copies, one blurred just enough to smooth the pixel noise and the other enough to smooth out the important detail, and subtract one from the other. Because the blurring is accomplished with a Gaussian blur, this method is called the "Difference of Gaussians" (sometimes abbreviated DoG) filter.

The advantage of the DoG method when applied to a noisy image is shown in **Figure 22**. By comparison to the example in **Figure 19**, the increase in edge contrast is not accompanied by an increase in visibility of the noise. The radius for the Gaussian blur used to reduce noise is typically in the range of 0.5 to 2 pixels, just enough to reduce the speckle, while the radius for the Gaussian blur used for the unsharp mask is typically several times greater. Of course, this assumes that the scale of the detail is larger than the scale of the noise.

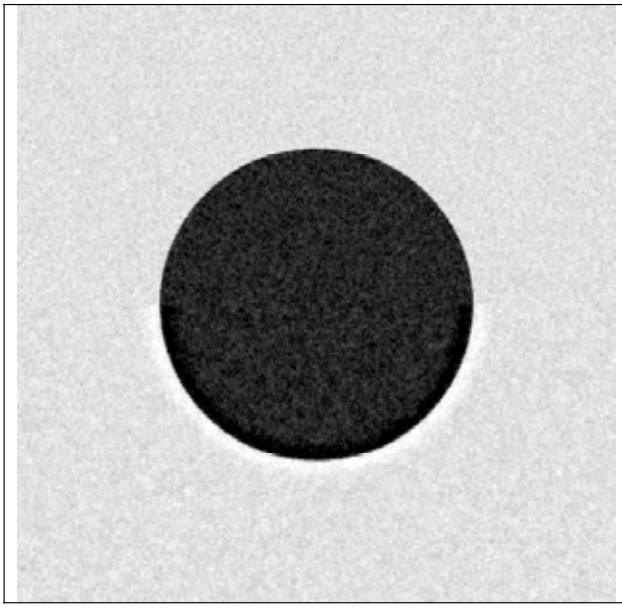


Figure 22. The random speckle noise superimposed on this simple image of a circle is less visible than the example of **Figure 19**.

The difference of Gaussians is a more general form of the unsharp mask, which includes two smoothing adjustments. One increases the edge contrast, as in the conventional routine, while the other reduces the noise. It would be equivalent to perform the smoothing first, separately, and then use a conventional unsharp mask filter except that it is difficult to know how much noise reduction is necessary until the sharpening process makes it more visually evident. **Figure 23** shows a plug-in (part of the Optipix plug-in set from Reindeer Graphics) that implements this method. Setting the radius for noise reduction to zero produces the conventional unsharp mask result, which in this implementation is automatically applied to the intensity channel of a color image.

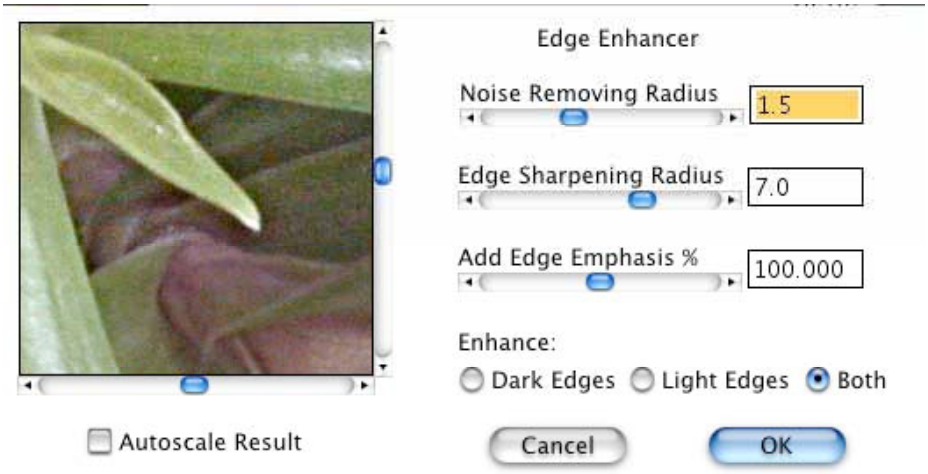


Figure 23. Edge sharpening using a difference-of-Gaussians method. The sliders allow interactive control over the radius of the Gaussians used to remove noise and to enhance detail.

The other approach is to reduce the noise first. Since the noise in an RGB image is actually found in the individual color channels, and is mostly independent and has different characteristics in each channel, it is possible to apply a Gaussian smooth directly to each channel. In the example shown in **Figure 24**, a

different amount of smoothing was used in each channel (a radius of about 2 pixels in the noisier blue channel, a radius less than 1 pixel in the green channel, and an intermediate value in red). Then the conventional unsharp mask was applied to the intensity channel as described above. This produces the best result, but requires the most effort. And since the increase in noise due to the unsharp mask is hard to judge beforehand, a few trial-and-error iterations were needed to get the final result shown, reducing the noise without sacrificing detail sharpness.

Figure 24. Sharpening the color image from **Figure 20** with noise reduction: a) difference-of-Gaussians enhancement (unsharp mask plus smoothing) applied to the intensity channel; b) smoothing each color channel separately, then applying an unsharp mask in the intensity channel.

Light and Dark edges

The unsharp mask (and related methods like the difference of Gaussians) highlight edges and detail by increasing the change in brightness close to each step. The standard version of the technique, shown in the preceding examples, adds a bright halo along the bright edge of the step and a dark halo along the dark edge. Sometimes there are advantaged in just using one or the other, rather than both. **Figure 25** shows the results, applied to the same set of steps used in **Figures 2** and **3**.

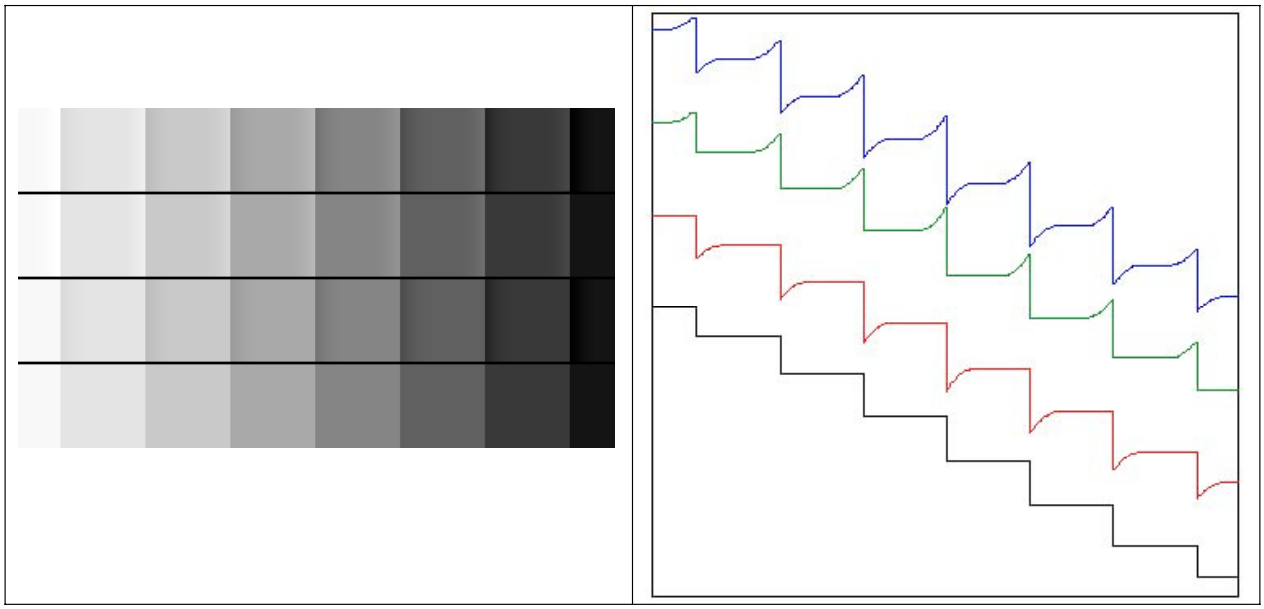


Figure 25. A set of uniform brightness steps, with an Unsharp mask used to apply both edges, light edges only, and dark edges only.

For real images, this has the advantage of reducing interference between steps or detail and the haloes from other, nearby steps. Also, it can be used to make features stand out better from background. To do this in Photoshop, duplicate the layer holding the image, apply the conventional unsharp mask, and then set the layer blending mode to “Darken” or “Lighten.” Of course, this only works for 8 bit per channel images (which can be put into layers), but the plug-in shown in **Figure 23**, which allows direct selection of dark or light edges, also functions on 16 bit per channel pictures.

The general rules for deciding which edge halos should be used are: 1) light edges don’t show up well on light background tones, and vice versa; and 2) the halo should lie on the background rather than the foreground, to help the feature stand out without altering its brightness values. Sometimes these rules are in conflict and it is necessary to try several combinations (both edges, light edges only, dark edges only) to decide which is best. And, as will be discussed below, different regions of the image may call for different answers.

In all of the examples shown to compare bright and dark edges, the width of the halos (radius of the blur function) and the amount of the effect added have been set larger than ideal, in order to make it easier to see and compare the effects. **Figures 26** and **27** illustrate the first rule. The light halos create visual artefacts (for instance next to the shutters) on the light colored house, and the dark halos merely interfere with other highlights on the dark dog. So we would use the dark edges on the light house and the light edges on the dark dog for best results.

Figure 26 Comparison (using the image from **Figure 1**): a) enlarged detail of original image; b) light edges only; c) dark edges only.

Figure 27. Comparison (using the image from **Figure 10**): a) enlarged detail of original image; b) light edges only; c) dark edges only.

Figures 28 and 29 illustrate the second rule. The light halos are visually distracting and unnatural around the model's eye, but the dark edges simply serve to increase the local visual contrast and bring out the existing detail. The leaf is darker than the background, so adding a light halo to the image helps the foreground object to stand out.

Figure 28. Comparison: a) enlarged detail of original image (using the image from **Figure 30**); b) light edges only; c) dark edges only.

Figure 29. Comparison: a) enlarged detail of original image (using the image from **Figure 35**); b) light edges only; c) dark edges only.

The typical problem for users of the unsharp mask filter is adding too much additional local contrast. Many of the preceding examples have used settings for the radius of the Gaussian blur and the amount of the effect to add to the original that are large, in order to make the nature of the effect quite evident in the example images. Adding some local contrast can make a bland image into a good one. Adding too much can make a good one into a disaster. Seeing the effect on the computer screen and watching details "pop" out of the background can easily mislead the practitioner into the "some is good, more must be better" mindset that produces caricatures instead of photos. It is hard to dispassionately review the results of processing to detect this flaw, and it is ultimately up to the individual user to decide where the limit lies. And it should be noted that what looks best on the computer screen will not usually be the same as what looks best in print, because the printing process compresses contrast and blurs detail somewhat as discussed in **Chapter 5**.

Figure 30 shows an example. The original picture has had some minor retouching of blemishes and darkening of the background, as discussed in following sections, but is otherwise unmodified. A small amount of added local contrast using an unsharp mask (actually a difference of Gaussians) improves the vitality and sharpness of the photo. Increasing the radius and percentage settings by one third clearly blows out the highlights, creates artificial shadows, and hardens the image in an unfortunate way.

Figure 30. Taking the unsharp mask too far: a) original image; b) modest sharpening with good results; c) abuse of the technique.

Other detail enhancement techniques

While it is the most widely used technique for processing of photographic images, the unsharp mask (and its various relatives discussed above) is certainly not the only tool available. One of its problems is that by calculating the difference between a pixel and its neighborhood, the results depend on the absolute brightness of the image. The difference between a pixel value of 30 and an average neighbor value of 20 is 10, or 50% of the background level. The same difference between a pixel value of 230 and a neighbor value of 220 is less than 5% of the background level. Human perception can detect changes of a few percent. The first case would represent a highly visible detail, while the second would be barely perceptible.

Another, related problem is that the unsharp mask serves only to increase local contrast. Sometimes it is necessary to reduce the overall image contrast so that local detail can be discerned, particularly in the very dark or very bright regions of images. There are other techniques that fare better in extracting detail in these situations.

Figure 31: Recovering detail: a) original image (courtesy Eastman Kodak Co.); b) Unsharp mask; c) Adaptive local equalization; d) Variance equalization.

Figure 31 shows an image in which the shadow areas inside the streetcar and inside the building behind it hide the detail present. An unsharp mask pumps up the contrast but does little to reveal that detail. Two other statistical techniques fare much better. The adaptive equalization method is the more widely used of the two (there is a free Photoshop plug-in available from www.ReindeerGraphics.com). It constructs the histogram for a small neighborhood around each pixel to perform the histogram

equalization function described in **Chapter 2**. Effectively, if a pixel is brighter than its neighborhood it makes it brighter still, and if it is darker, it makes it darker still. This increases the local contrast.

The variance equalization technique (part of Fovea Pro from Reindeer Graphics) also performs this shift, but by an amount the makes the statistical variance - a measure of the differences between all of the pixels in the neighborhood - equal across the image. This limits the effect of random speckle noise in the image somewhat better than the comparison of just the central pixel to the neighborhood, although all of these so-called “high pass” filters that emphasize short-range variations in brightness are inherently sensitive to noise and tend to increase its visibility as well. Both of these methods bring out the details in the nearly uniform shadow areas of the original image.

Of course, it is still possible to use an unsharp mask to increase the sharpness of the detail after one of these equalization techniques has been used to recover it from the original, as shown in **Figure 32**.

Figure 32: Applying an unsharp mask filter to **Figure 31d.**

Another situation arises when edges and region boundaries in an image are broad. This may be due to problems such as focus blur but is often inherent in the shape and shading of objects. For some purposes, such as reducing the number of colors needed to print the picture, to eliminate fine detail that will not print anyway while increasing the definition of the remaining structures, or to simplify it for graphics arts representation, it may be desirable to make the change in brightness at the boundary very abrupt and to eliminate the very fine details in order to increase the visibility of the major edges.

Figure 33 shows an example. The median filter is a widely used tool that has its primary function in eliminating isolated noise pixels and fine scratches. It compares the values of pixels in an adjustable neighborhood (for this purpose we make that neighborhood small, with a radius of 1 or 2 pixels) and finds the median value, the one with the same number of brighter and darker neighbors. This pixel’s value is then transferred to the central location. When this process is repeated for every pixel in the image, it rejects pixels that are very bright or dark compared to their surroundings. It also tends to sharpen up edges in the sense that a gradual change in brightness becomes more abrupt as values on either side of the boundary take on the values of their similar neighbors.

Figure 33. Applying a median filter: a) original image; b) one application of a median filter; c) twelve repetitions of the median filter; d) unsharp mask applied to image c.

One of the nice properties of the median filter is that it can be repeated without shifting or distorting edges. In the example, repeating the application of a small median filter a dozen times produces an image with well defined boundaries, and smooth tones in which fine detail has been replaced by gradually varying blocks of color, resulting in an almost painted quality to the original photo. This visual impression can be further enhanced by applying an unsharp mask after the median filters. This technique is sometimes called posterization, but should not be confused with the Photoshop Posterize function which simply reduces the number of bits used to represent brightness values.

Edges

Human vision is highly adapted to looking at surfaces under various lighting situations, and interpreting the resulting image. A common trick for seeing subtle surface detail is to shine a light on it at a low angle, so that even small roughness or protrusions are brightly lit and cast shadows behind them. Digital images can be treated as though changes in brightness represent changes in height, and the same technique applied to see the detail. **Figure 34** shows the method, usually called “embossing” because it created the appearance of raised relief on surfaces. It is particularly useful for examining macro images of surfaces for defects or forensic evidence.

Figure 34. Embossing: a) original image; b) embossed result; c) custom embossing filter; d) dedicated embossing dialog.

A Photoshop custom filter to produce embossing uses positive numbers on one side and the central pixel, and negative ones on the opposite, to produce the bright and dark edges that are interpreted as relief. The offset value adds a medium grey so that both light and dark edges can be seen. The direction of the shadows (indeed, whether the result is interpreted visually as an indentation or a bump on the surface) is simply controlled by shifting the positive and negative values. A dedicated filter to perform the embossing operation is also provided in Photoshop.

For color images, it is necessary (as for most image processing filters) to work in a hue-saturation-intensity space and apply the filter only to the intensity, leaving the color alone. In **Figure 35** this was done by converting the image to Lab color space (Image -> Mode -> Lab Color), selecting the Lightness channel, and processing it before converting back to RGB Color. The colors have been altered somewhat in brightness by the operation, but are still useful for distinguishing structure. If the operation is applied directly to the RGB image, the individual red, green and blue channels are processed, and the overall color information is lost.

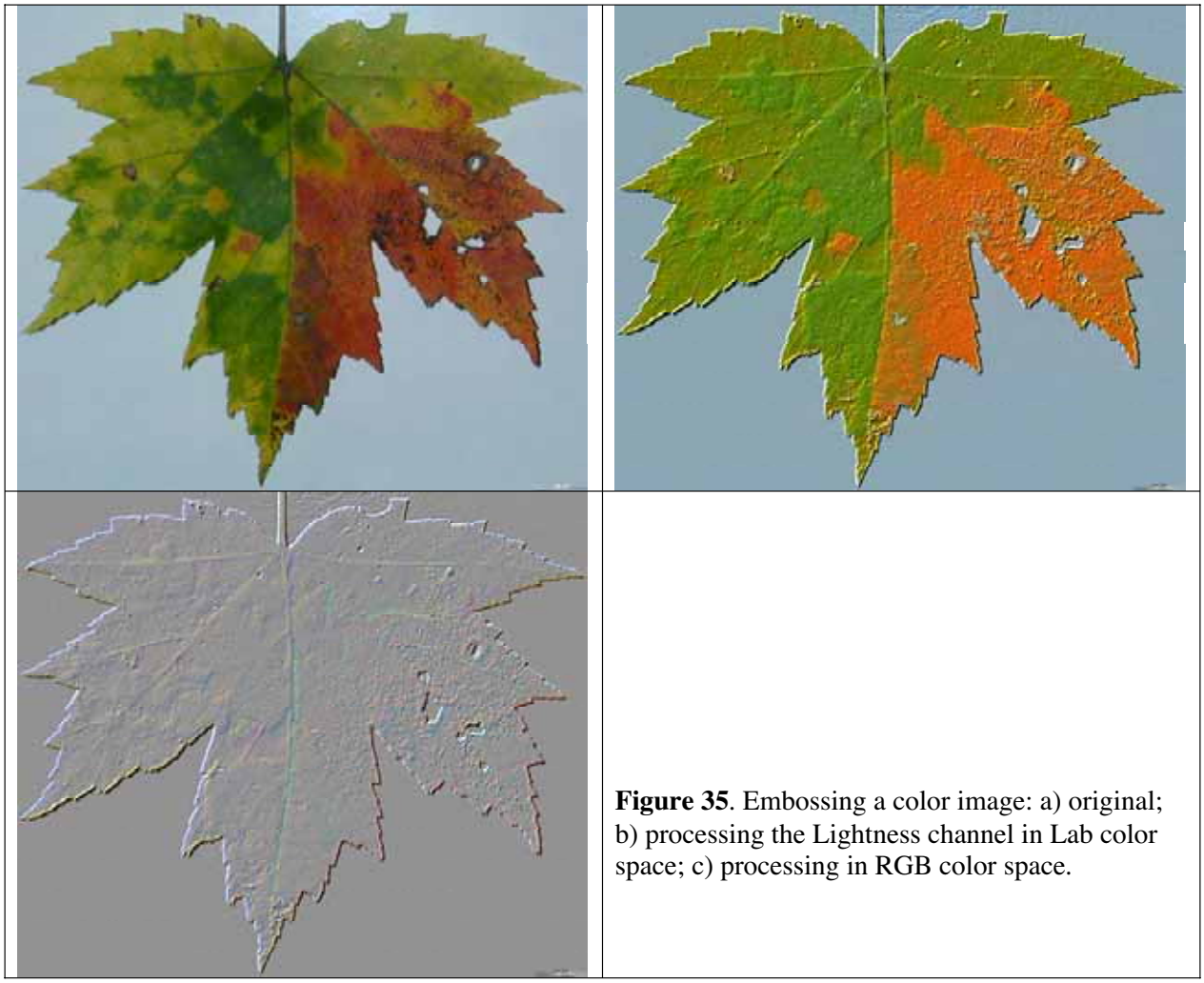


Figure 35. Embossing a color image: a) original; b) processing the Lightness channel in Lab color space; c) processing in RGB color space.

It is also possible to extract the edges from an image entirely. There are several quite different algorithms used for this purpose (edge finding is a major field of endeavor in the image processing

world), with subtle differences between them but a common purpose. **Figure 36** shows two representative approaches. The examples were processed in RGB color space, which is why the boundaries are colored but the interior regions are not. If the same operation is performed on only the intensity channel (for example, the L channel after converting to Lab space), the boundaries are dark and the colors of regions are retained as shown in **Figure 37**.

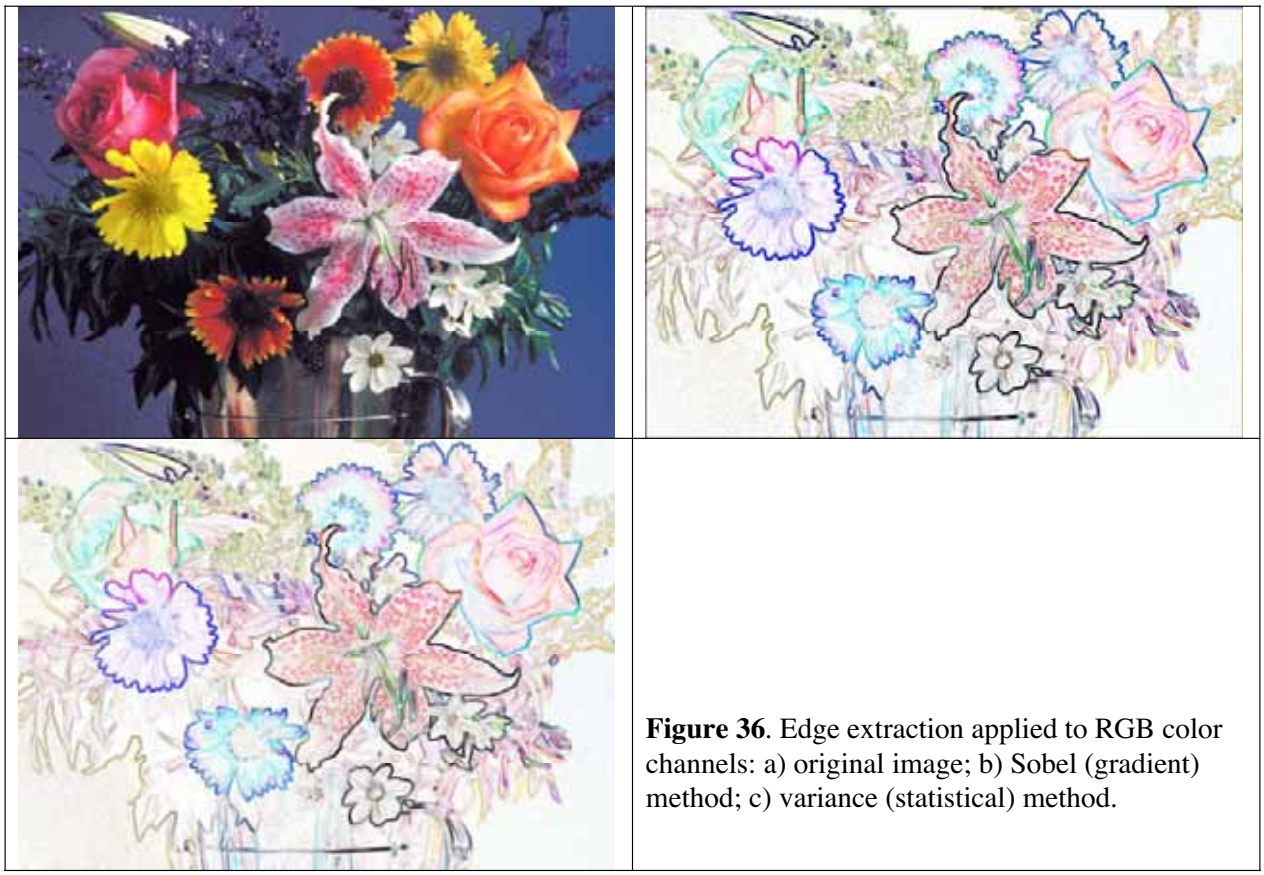


Figure 36. Edge extraction applied to RGB color channels: a) original image; b) Sobel (gradient) method; c) variance (statistical) method.

Figure 37. Edge extraction applied to the L channel in Lab color space: a) Sobel (gradient) method; b) variance (statistical) method.

The Sobel gradient approach, which is probably the most widely used and seems to be the method used by Photoshop's "Find Edges" routine, calculates the derivative of brightness in the horizontal and vertical directions using a filter very similar to the embossing filter shown above. These two derivatives are then combined as vectors to determine the magnitude of the local change in brightness, and the resulting value highlights the edges with lines. Since this is a very localized operator, it generally highlights the finest detail of the three approaches, but it responds differently to edges in bright and dark regions of the image.

The variance technique calculates the statistical variance of the pixel values in a small neighborhood. In uniform regions this is a very small value, but it rises quickly when a texture or edge is encountered. This method (performed using a plug-in from the Reindeer Graphics Fovea Pro suite) finds both major and minor edges well but responds to textural details as well as edges.

There are also methods to thin the resulting edge lines down to single pixel width (essentially converting the original image to a sketch), and by combining that sketch with the original or modified colors from the image to create line drawings and cartoons from photographs (**Figure 38**). Since it is not the intent of this book to wander too far into the graphics arts field, as opposed to the enhancement of digital photographs, this will merely serve to indicate that there is a rich field of processing options that start with photos and end up with something that may or may not be "art."

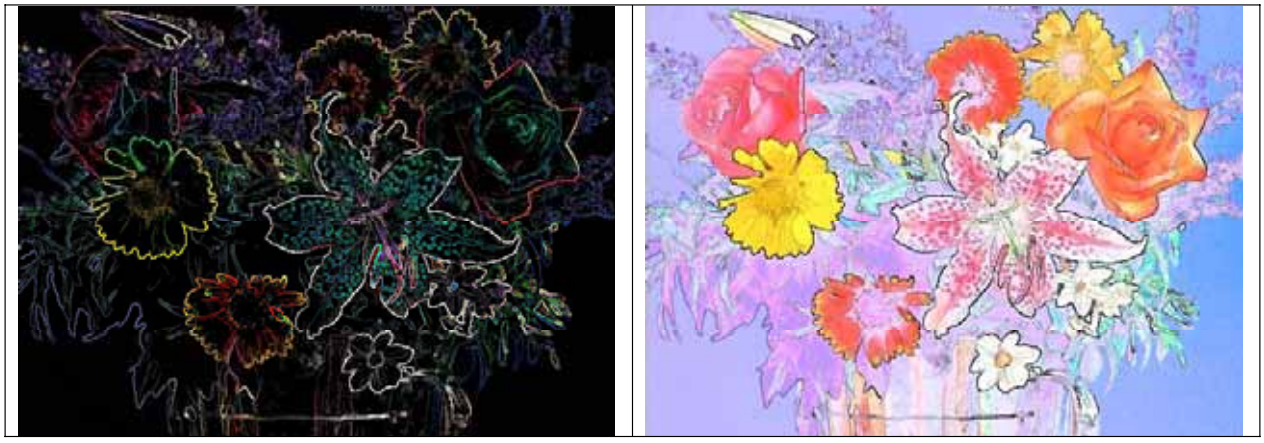


Figure 38. Thinning the edge lines from **Figure 35b** to single pixel width, and mixing them with the color information from the original image.

Reducing noise

In the field of signal processing, of which digital image processing is a part, the word "noise" has a generic meaning. Noise is that part of the signal (or image) that isn't wanted and doesn't represent the original input (the scene). There are many types and sources of noise, and different techniques are required to remove or reduce their effects. Typically image noise is expected to refer to corruption of the true picture by imperfections in the process of acquiring, transmitting or storing the data, but sometimes this meaning is broadened to include defects in the scene itself (for instance, removing falling raindrops) or in a previous, intermediate representation of the scene (for instance, scratches or dust on a photographic negative that is being scanned into the computer).

In digital photography, the storage and transmission of the image data is digital and not normally subject to corruption by noise. Either the signal survives the storage and transmission process and is identical to the original, or it is lost and may or may not be recovered, but it isn't partially altered or degraded. But the detector that converts the incident light to electrons, subsequent electronic amplification, and conversion to digital measurement (discussed in **Chapter 1**) are definitely subject to various kinds of problems and limitations, and can produce noise in the resulting image.

Chapter 1 discussed random speckle noise in acquired digital images, and illustrated several means for alleviating the appearance of the noise. Because the presence of noise makes it more difficult to sharpen or otherwise post-process images, a further discussion of noise reduction methods is included here.

Probably the most common type of noise appears as a random variation in brightness. Ideally a picture of a uniformly illuminated, uniformly colored region should have identical values for all of the pixels. In practice, it does not. The histogram of the image shows a peak with finite width, and the wider the peak is the greater the pixel value variations, which show up visually as a random speckle superimposed on the image. The peaks are usually Gaussian in shape, simply because of the statistical law known as the "central limit theorem." This states that a Gaussian variation results when there are a great many independent sources of variation, regardless of the nature of each one. And in the detection, amplification and digitization of the image data there are plenty of opportunities for variations to occur.

In **Chapter 3**, the averaging together of multiple images was used to increase the dynamic range of the data and allow better control over contrast and color adjustments. This is also a way to reduce noise. Generally, when the sources of noise are multiple and random, increasing the number of samples that are averaged together allows the signal to build up while the noise fluctuations tend, over the long run, to cancel out. Technically, the signal-to-noise ratio increases as the square root of the number of separate images that are combined. But except in a few cases such as using the successive frames from a video, it is only practical to average together a relatively small number of images.

If temporal averaging cannot be used as a noise reduction method, the next candidate is spatial averaging. Most of the structures and objects in images are larger than the size of the pixels. Within each region of the image, we would expect one pixel to be pretty much like its neighbors, except for the random speckle that arises as noise. A Gaussian blur, as discussed in **Chapter 1**, replaces each pixel with a weighted average of the pixels in the immediate neighborhood.

The controlling parameter is the radius of the blur function, usually expressed as the standard deviation of the Gaussian. It isn't necessary to know anything about statistics or math to use this. The radius control in Photoshop controls the amount of the blur, and the total diameter of the neighborhood of pixels involved in the calculation is about 6 times the blur radius. For example, the custom filter shown in **Figure 39** corresponds to a Gaussian blur with a radius (standard deviation) of 1 pixel.

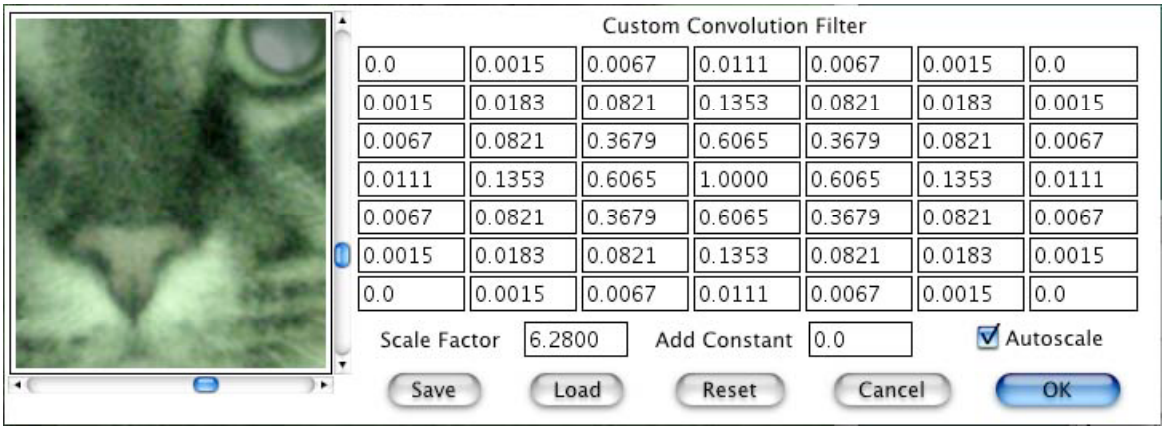
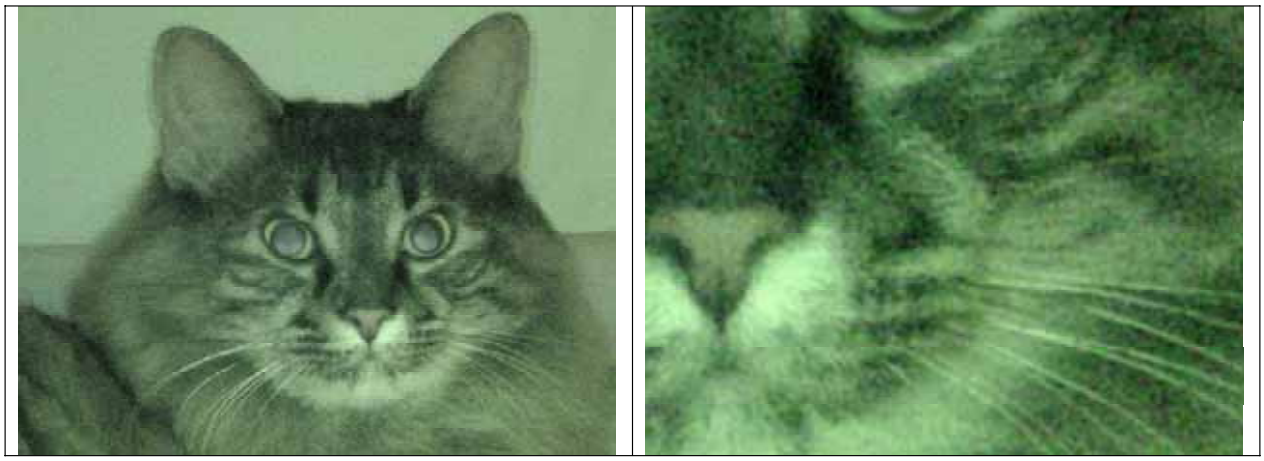


Figure 39. Custom filter that performs a Gaussian blur with standard deviation of 1.0 pixels. Instead of the built-in Photoshop custom dialog, this is the one in Reindeer Graphics Fovea Pro, which allows the entry of floating point numbers, works on 16 bit per channel images, and has optional automatic scaling of the results.

In practice it is not necessary to use custom filters for performing this operation. The Gaussian blur routine allows direct control of the radius and generates the kernel weights as required. In fact, it isn't necessary to use a full array of weights, because the Gaussian smooth is separable, meaning that a purely vertical smoothing followed by a separate purely horizontal smoothing produces the same result. For noise reduction purposes, values from about 0.3 to about 2.0 pixels are typically useful. One of the properties of the Gaussian is that it provides the maximum amount of noise reduction for any given amount of blur.

Typically, in digital photography, the greatest amount of random speckle noise is found when high effective ASA ratings and dim lighting are used, often with relatively long exposures (a lot of electronic gain applied to a signal produced by few photons). **Figure 40** shows an example, actually an infrared photograph taken by available light in a darkened room, by removing the infrared filter in front of the chip. The breadth of the histogram peak for just the background pixels above the cat's head indicates that there is a very large amount of variation present in the pixel values.



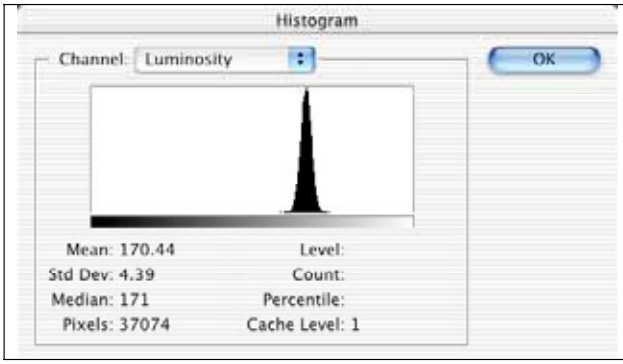


Figure 40. Random speckle noise in a monochrome image: a) original; b) enlarged detail; c) histogram of the wall area over the cat's head.

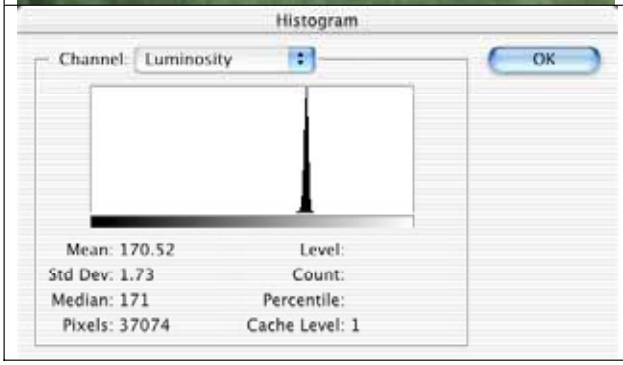


Figure 41. Gaussian smoothing applied to the image in **Figure 40**: a) entire image; b) enlarged detail; c) histogram for comparison to **Figure 40c**.

Applying a Gaussian blur to this image with a radius of 2 pixels reduces the breadth of that peak, as shown in **Figure 41**. But it also blurs the fine detail. Any lines, whiskers or other markings in the original image that are as small as the scale of the noise - in other words the size of pixels - will be affected by the blurring as much as the noise is, and this may cause the entire image to appear to be out of focus. It is true that applying an unsharp mask to the smoothed image can increase the contrast for fine lines and detail, but it does not remove the effects of the blur. The operation described previously as a difference-of-Gaussians makes it possible to increase local contrast without having the noise in the image increase in visibility and dominate the result, but it does not remove blur (**Figure 42**).

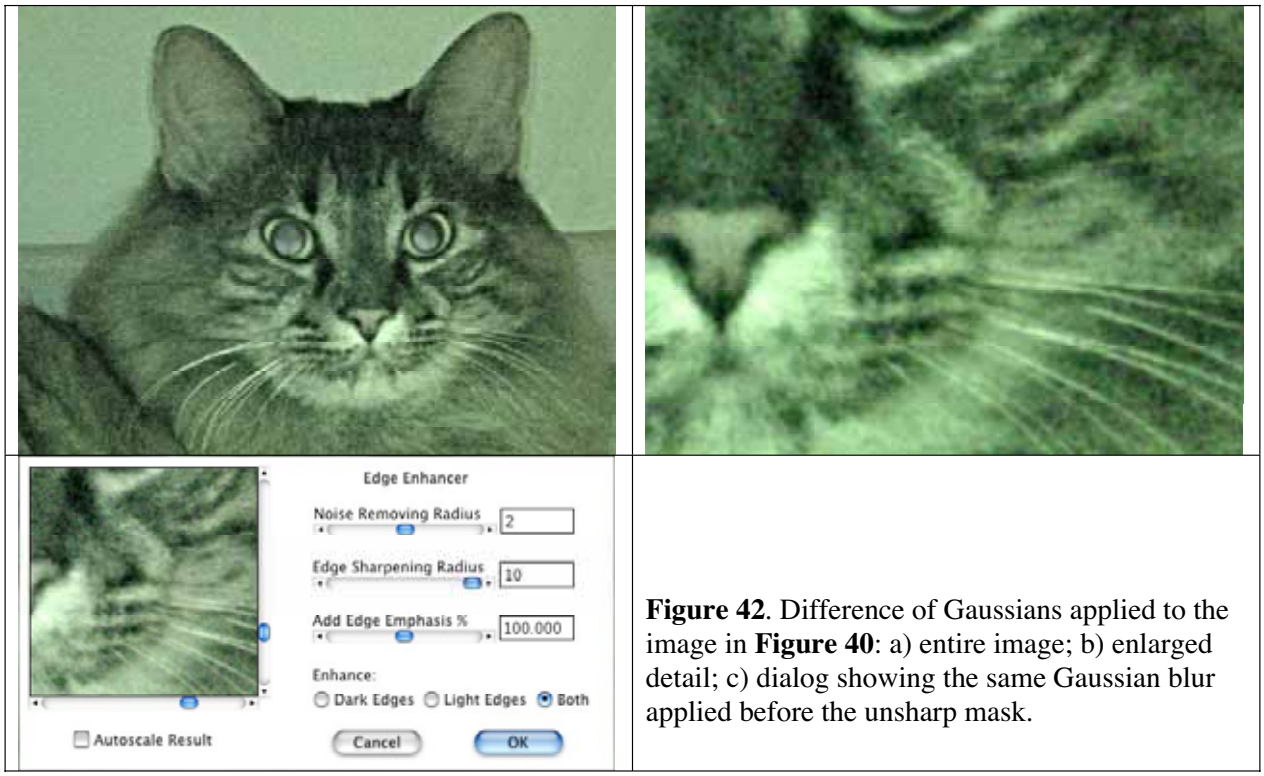


Figure 42. Difference of Gaussians applied to the image in **Figure 40**: a) entire image; b) enlarged detail; c) dialog showing the same Gaussian blur applied before the unsharp mask.

There is another technique for noise reduction that is best suited to the removal of scratches and shot noise, discussed in **Chapter 1** and again below, that can also be applied in some cases to speckle noise. The median filter ranks the pixel values in a small neighborhood into order based on their brightness and replaces the central pixel with the median value, from the center of the ranked list. Unlike a Gaussian blur, it does not reduce the contrast at steps or boundaries, and does not shift edges. **Figure 43** shows that when applied to an image with speckle noise it reduces the breadth of the histogram peak and the visible speckle nearly as much as the Gaussian blur. However, the appearance of the image is altered visually because of the steps in brightness between regions in which the pixels have been made very similar by the filtering process.

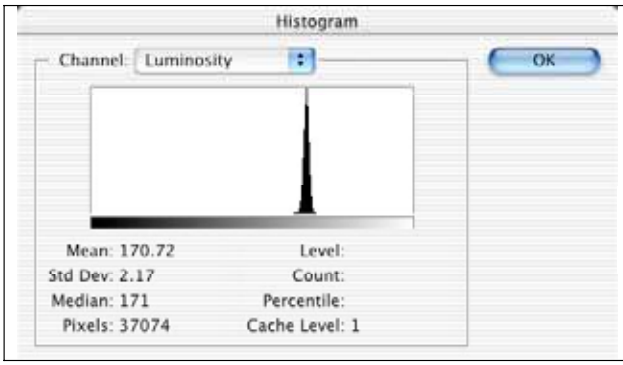


Figure 43. Median filter (diameter = 5 pixels) applied to the image in **Figure 41**: a) entire image; b) enlarged detail; c) histogram for comparison to **Figure 40c** and **41c**.

Noise in color images is largely independent in the red, green and blue channels. **Figure 44** shows an enlarged portion of a color image taken in dim light with high ASA gain and a 1 second exposure. Such long exposures produce more thermal noise in the chip; some cameras compensate for this by acquiring a separate image with no incident light, to subtract from the image, but this was not done in this case. Both the Gaussian blur and the median filter applied to the intensity values reduce the visible speckle somewhat in this image.

Figure 44. Application of noise reduction filters to a color image: a) original; b) Gaussian blur; c) median filter.

It is instructive to examine the individual noise channels in the original image (**Figure 45**). The green channel shows the highest spatial resolution and the least speckle noise, as we would expect from the fact that more sensors are used to collect the green light. The blue and red channels are noisier, and also show less spatial resolution. Note that in the green channel the speckle variation is from one pixel to the next, but in the red and blue channels the speckle covers several pixels indicating this lower resolution.

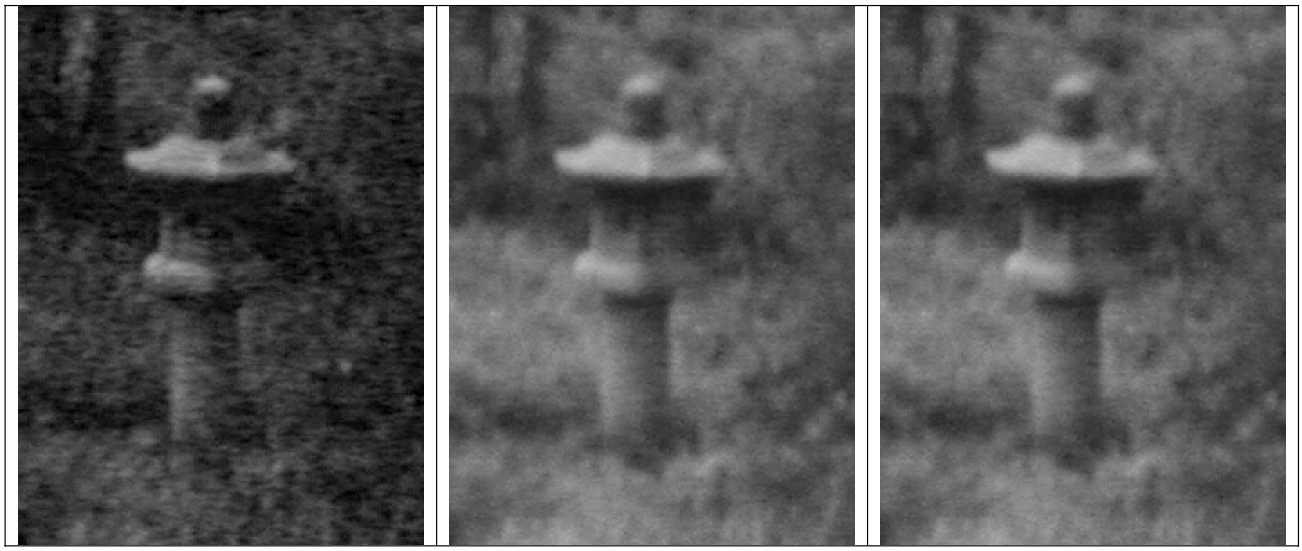


Figure 45. Individual color channels in **Figure 44a**: a) red; b) green; c) blue.

The different noise characteristics of the individual color channels suggests that different amounts of blur or even different noise reduction methods might be applied to each channel. That approach can be effective, but care must be taken to avoid producing color shifts or introducing new colors into the image, particularly at boundaries. **Figure 46** shows an example in which different Gaussian filters were applied to the red and blue channels and a median to the green channel.

Figure 46. Processing the individual color channels separately (compare to **Figure 44**).

Other types of noise

The random fluctuations in the processes of producing electrons from photons, collecting and amplifying the electrons, and measuring the resulting voltage, produce a variation in pixel brightness that has been characterized above as speckle, and has a more-or-less Gaussian distribution. There are some other sources of noise that have very different sources and appearance. One is the “shot” noise, also called salt-and-pepper noise, because of the black or white pixels that result in a monochrome

image. Most chips have at least a few dead or locked pixels (defects that set the voltage from a transistor on the chip to either zero or maximum output) out of the several million present, and except for some very high end devices used by NASA this is accepted as a practical limit of manufacturing quality control.

Because the readout of the chip is analog, and particularly for a single-chip camera in which the color information is obtained by interpolation, a single dead or locked transistor does not translate directly into a black or white pixel in the resulting image. More typically, it produces a brightly colored spot that may be one or several pixels across, as shown in Figure 45. Using a Gaussian blur on an image with this type of defect does not work well, because it averages the extreme value into the image, spreading the problem out to affect several surrounding pixels. A median filter is much better because it eliminates the extreme value, replacing it with one from a nearby pixel.

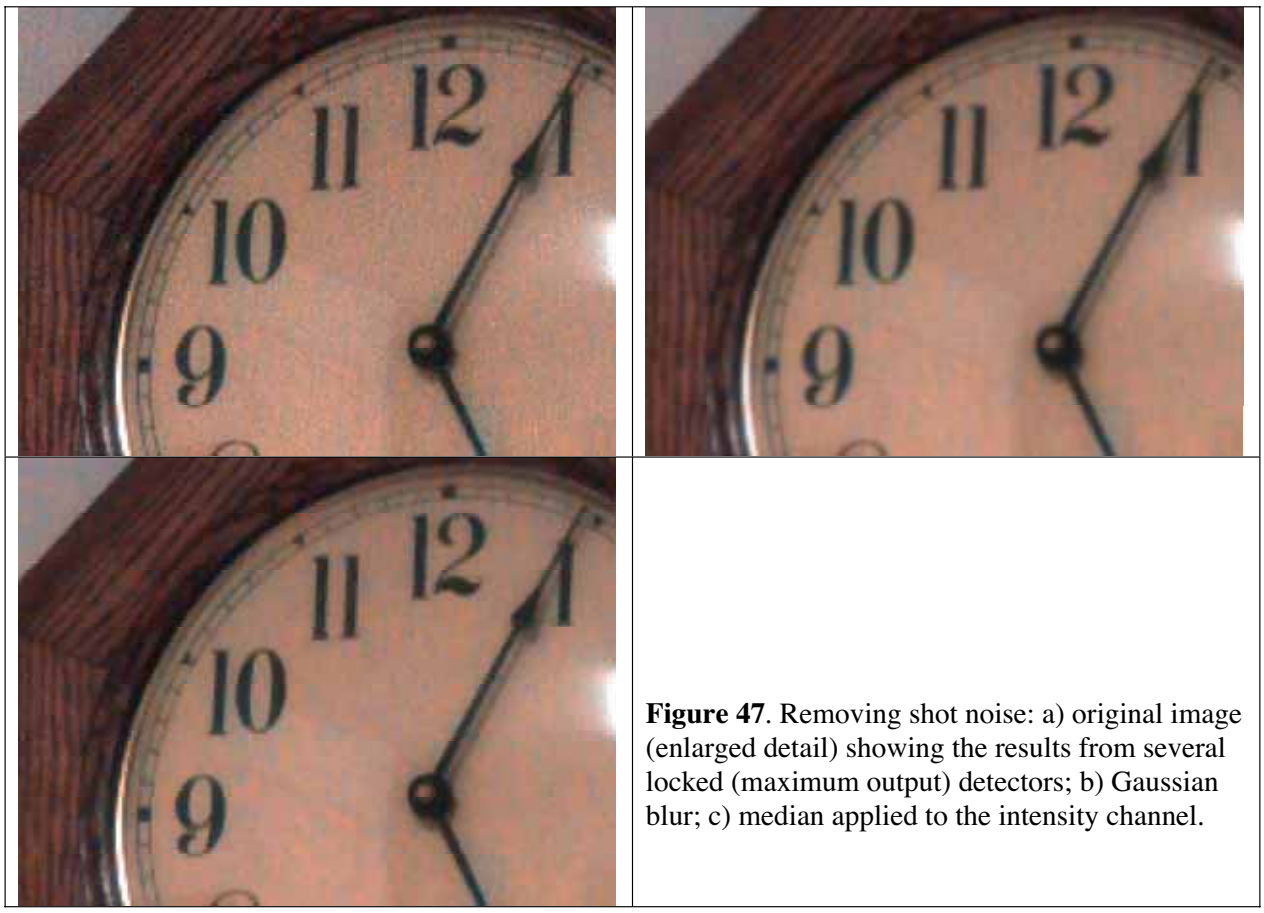


Figure 47. Removing shot noise: a) original image (enlarged detail) showing the results from several locked (maximum output) detectors; b) Gaussian blur; c) median applied to the intensity channel.

Because the defect typically affects just one color detector, the operation can also be performed on the individual color channels, as shown in **Figure 48**. This result is better than operating on just the intensity channel, and is another example of the advantages of using knowledge of the source of an image defect to choose the best way to ameliorate its effects.

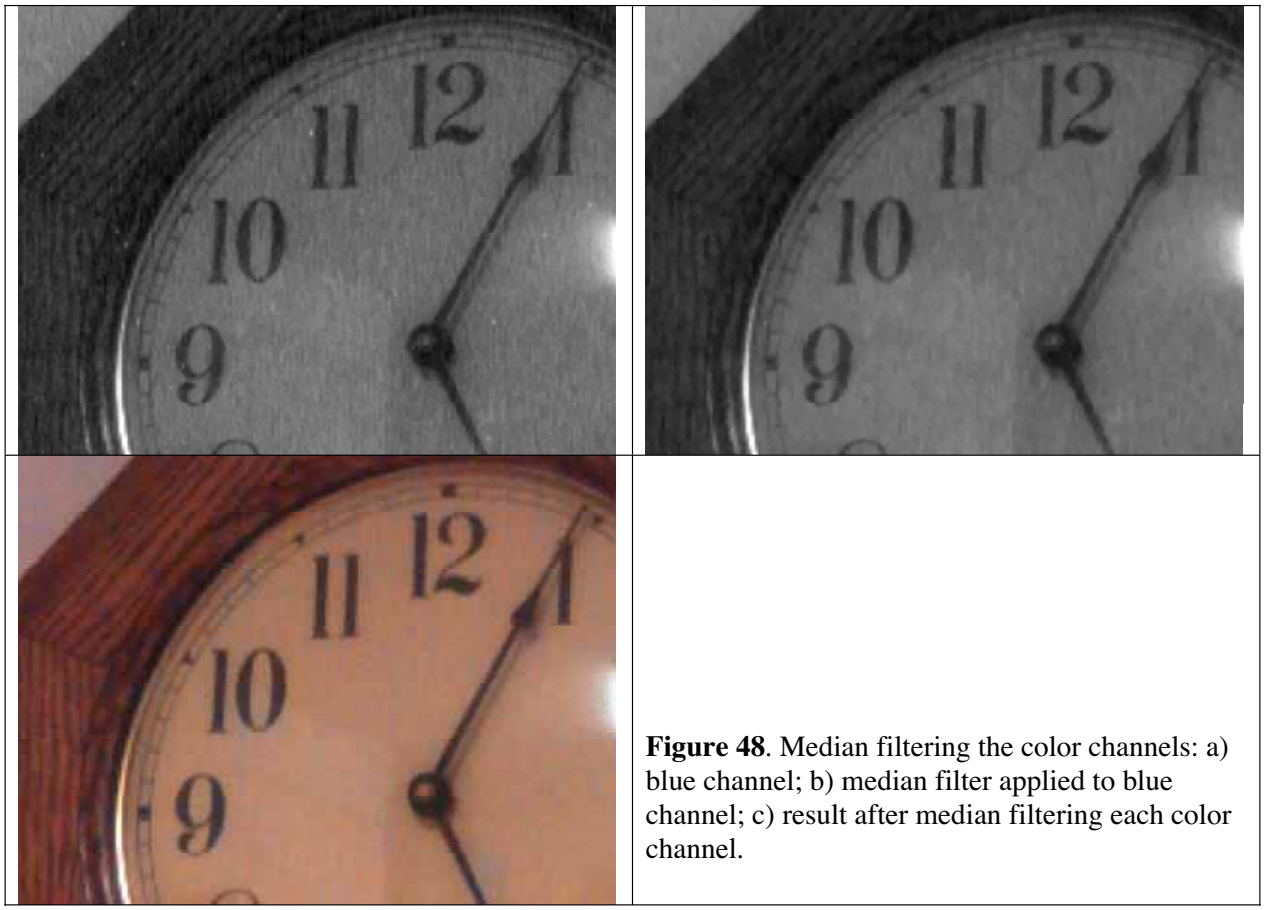
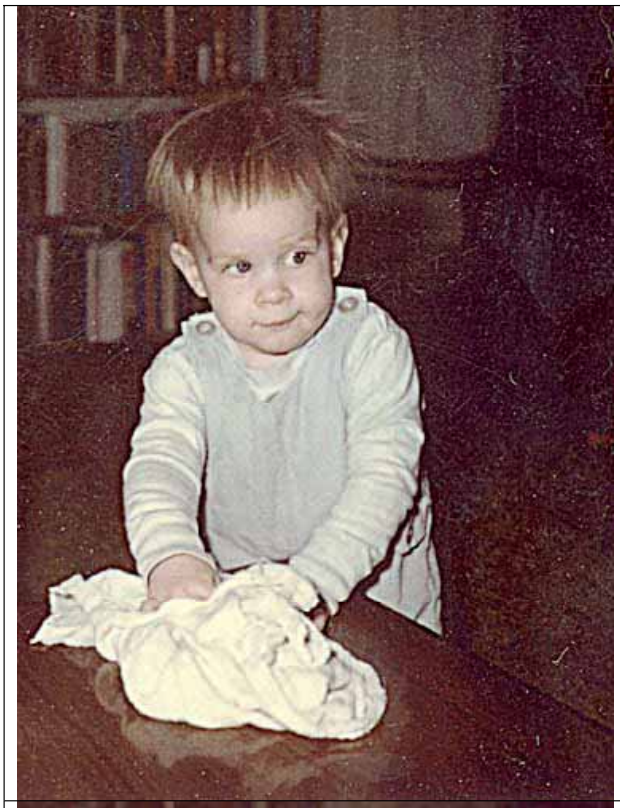


Figure 48. Median filtering the color channels: a) blue channel; b) median filter applied to blue channel; c) result after median filtering each color channel.

The median filter is also very good at removing fine lines or points such as scratches and dust from images. These most commonly occur due to the handling of film or prints, and so are more common in scanned images than ones directly acquired from a digital camera. However, scan line noise in video cameras can also be treated using the same technique if it is necessary to deal with single frames taken from videotape. So can lines such as utility wires in scenes, as discussed in **Chapter 1** in terms of median filters with adjustable neighborhoods, and in this chapter's section on image retouching. **Chapter 1** also shows examples of median filters applied to color values, as contrasted to working on the individual color channels in an image.

Unlike the shot noise that arises in digital cameras, scratches and dust usually show up as white on the image, and so applying the procedure in the intensity channel is appropriate. The radius used for the median filter should be wide enough that the damaged pixels will never cover more than half of the neighborhood (or the median value would be taken from the defect rather than the surroundings). Because the median filter does alter the appearance of steps and detail, it is usually best to apply it only in a region around the defects, as shown in **Figure 49**. The region in this example was created using the Select->Color Range process, but of course manual drawing methods can also be used. These types of defects can also be handled by manual retouching, as discussed below.



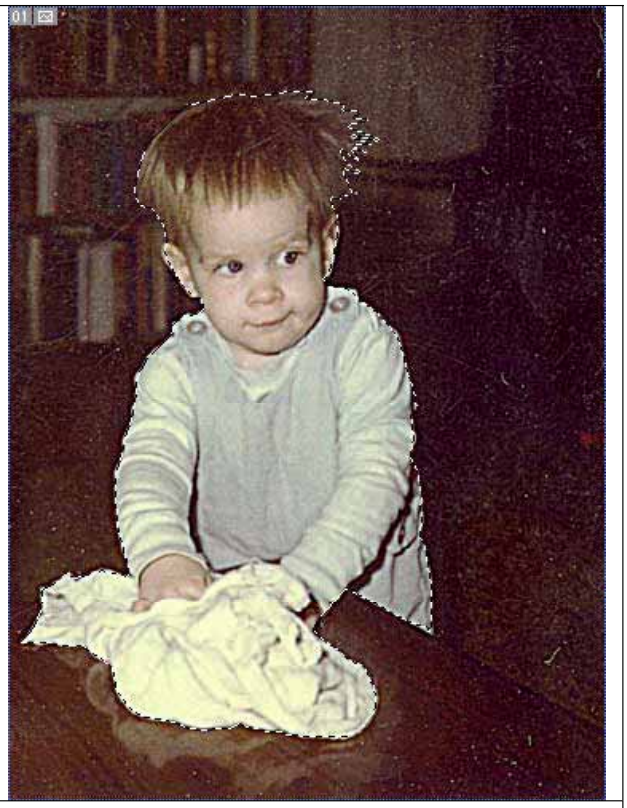
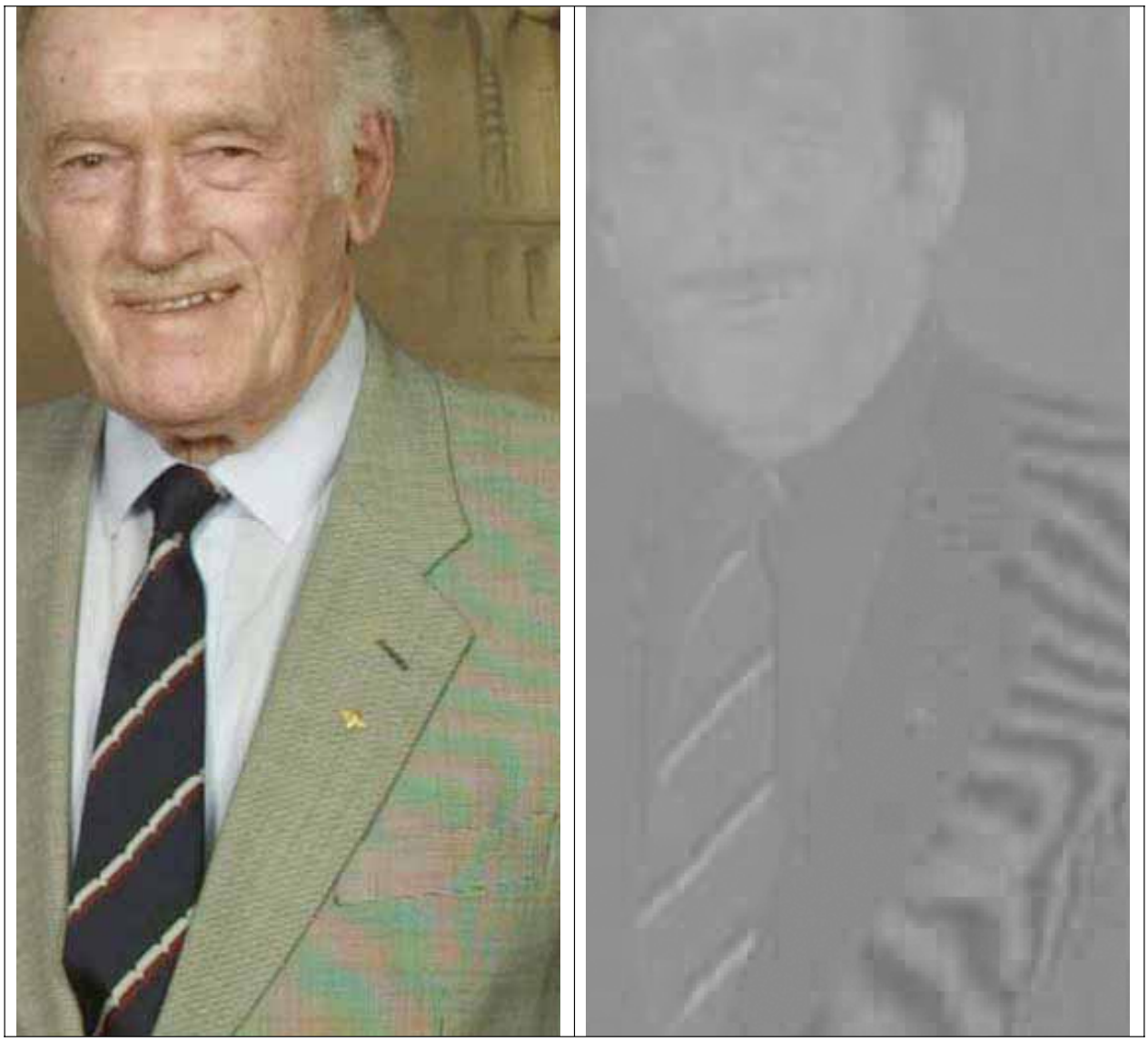


Figure 49. Eliminating dust with a median filter: a) scanned-in image from a photographic print; b) region selected for processing; c) result after median filter.

Another rather specific defect that arises with single-chip color cameras is moiré patterns that arise from interference between the spacing of the color-filtered transistors on the chip and the spacing of some regular pattern in the scene. Often, the problem is patterns in clothing, as shown in **Figure 50.** The alternating light and dark lines in the jacket weave have a spacing that is close to the spacing of the color filters on the transistors, and so a beat pattern shows up in the image in which red or green predominates in certain areas.

This type of defect is usually fairly easy to correct after converting the image from RGB to Lab color space. In the a channel (red-green) the moiré bands are very distinct. Drawing a region around the banded area and filling it with the average pixel value in the region eliminates the bands. The resulting color image retains all of the detail as the original but the visible moiré banding is eliminated. Any slight color shift in the affected region is visually undetectable.



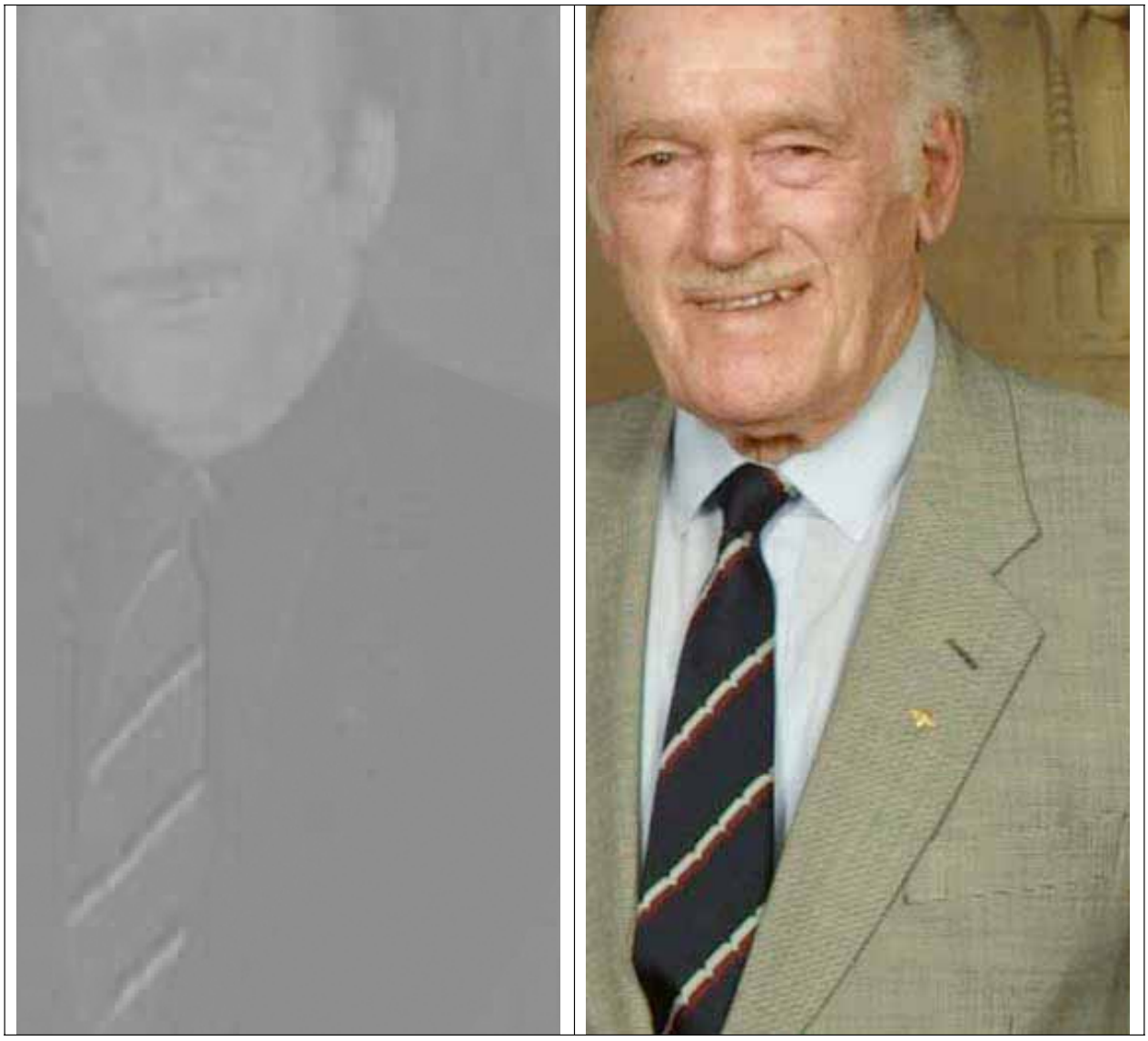


Figure 50. Eliminating moiré banding: a) original image; b) “a” channel in Lab color representation; c) filling a region around the bands with the average pixel value; d) resulting color image without banding.

Periodic noise

A special category of image “noise” is the periodic noise that can arise from electrical interference (including interference generated by fluorescent lights when television cameras are used), or vibration. This usually shows up as a set of intensity bands that overlay the entire image. Similar problems arise in forensic applications, such as trying to obtain a clear image of a fingerprint from a fabric that has a dominating regular pattern.

Finally, scanning in pictures that have been halftone printed in books or magazines shows the regular pattern of the halftone dots, which significantly degrade the quality of the image (and betray its origins). Many scanners have “descreen” functions built into their software that can sometimes eliminate this

pattern, provided the user can specify the spacing of the dot screen. Another trick that is commonly recommended in newsgroups are to rotate the original on the scanner to find a magic angle that will minimize the ability of the linear detector in the scanner to resolve the dot pattern. But what definitely does not work satisfactorily is to apply a Gaussian blur or any other "low pass" filter (for example, by scanning the image at one size and then forcing Photoshop to resample the data and change it to another size - the interpolation process is also a low pass filter) to remove the dots. That blurs the important image detail as well and does not solve the underlying problem.

There is a good way to remove periodic noise, whether it arises from a halftone pattern, electrical interference, vibration, or any other source. It involves working with the image in a different representation, one that is probably unfamiliar to most Photoshop users and photographers, although the roots of the underlying math go back to basic optical principles. And the math can be intimidating as well, although it really isn't necessary to deal with it just to use a plug-in based on the methods. The presentation that follows shows the fundamentals of the Fourier space filtering technique by example, without resorting to the math. But there are dozens of "serious" image processing texts (including John Russ, "The Image Processing Handbook," 4th edition, CRC Press, 2002) that can supply the underlying theory and details for anyone interested.

Any set of regularly spaced lines in an image can be described by their orientation and spacing. Fourier's theorem states that any complex signal - such as an image - can be made up by adding together enough sets of regularly spaced sets of lines. When an image is transformed from the "spatial" domain - the world of pixels - to the "Fourier" domain, the contrast or amplitude of each of the sets of lines that it takes to build up the original image is represented by one point in the Fourier space representation. The predominant sets of periodic lines, which in this context represent the noise or defects that we want to eliminate from the image, show up as dark points in the Fourier space representation. This is also commonly called the power spectrum, because the magnitude of each point is a measure of how much that set of lines - with its particular spacing and orientation - contribute to the final image.

Figure 51 shows an example. The "image" consists of three sets of regularly spaced lines, with different spacings and orientations. The Fourier space representation of this image consists of just three points (each one plotted twice because of the convention used to plot the data). Each point corresponds to one set of lines, as shown by the colored arrows. The direction from the center of the image to the point is perpendicular to the orientation of the lines, and the distance is the frequency (the reciprocal of the spacing - high frequencies means small spacing and vice versa).

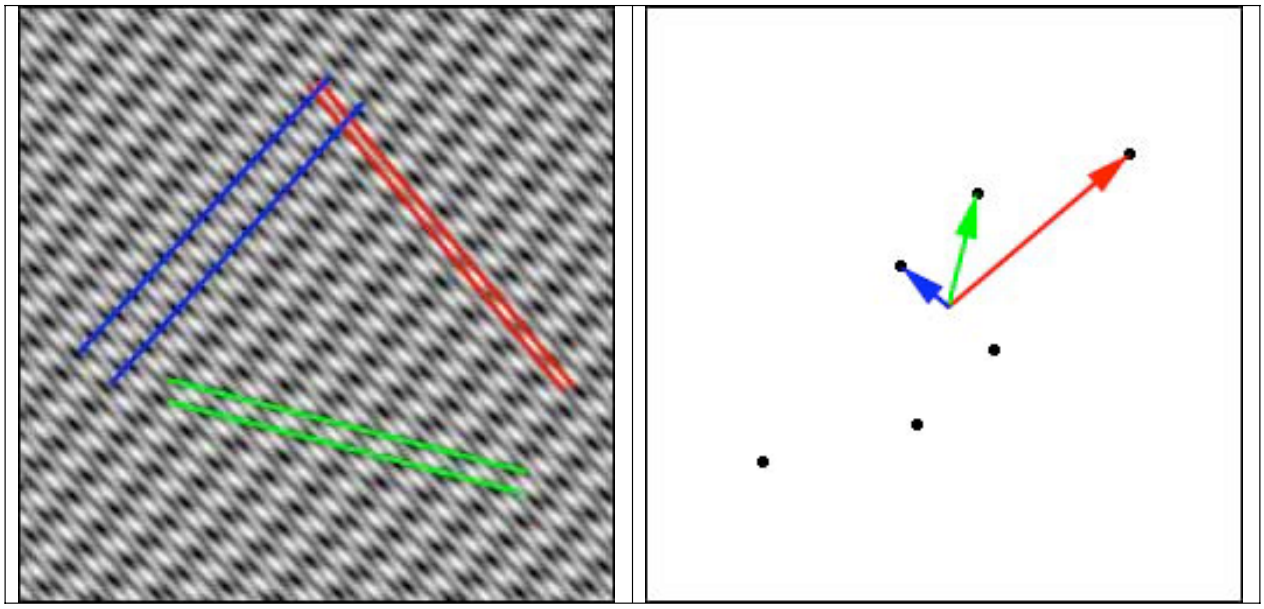


Figure 51. The relationship between the pixel or spatial domain image and its Fourier or frequency representation: a) three sets of regularly spaced lines, marked by red, green and blue overlays; b) the Fourier transform of the image, showing three "spikes" that correspond to the three sets of lines as shown by the colored overlay vectors.

That may not seem very helpful or intuitive, but it allows the periodic noise to be isolated from the rest of the image. If the same set of lines is superimposed on a real image, as shown in **Figure 52**, the same "spikes" appear in the Fourier power spectrum. Removing just those spikes and leaving all of the other data alone makes it possible to transform back from the Fourier to the pixel space representation and recover the image without the lines. This removal of spikes that represent sets of regularly spaced lines in the image is the way that periodic noise can be efficiently removed from images. Notice that the lines are imperfectly removed at the edges of the picture. The usual solution to this problem, which is inherent in the method is to copy the image into a larger window and keep only the central portion.

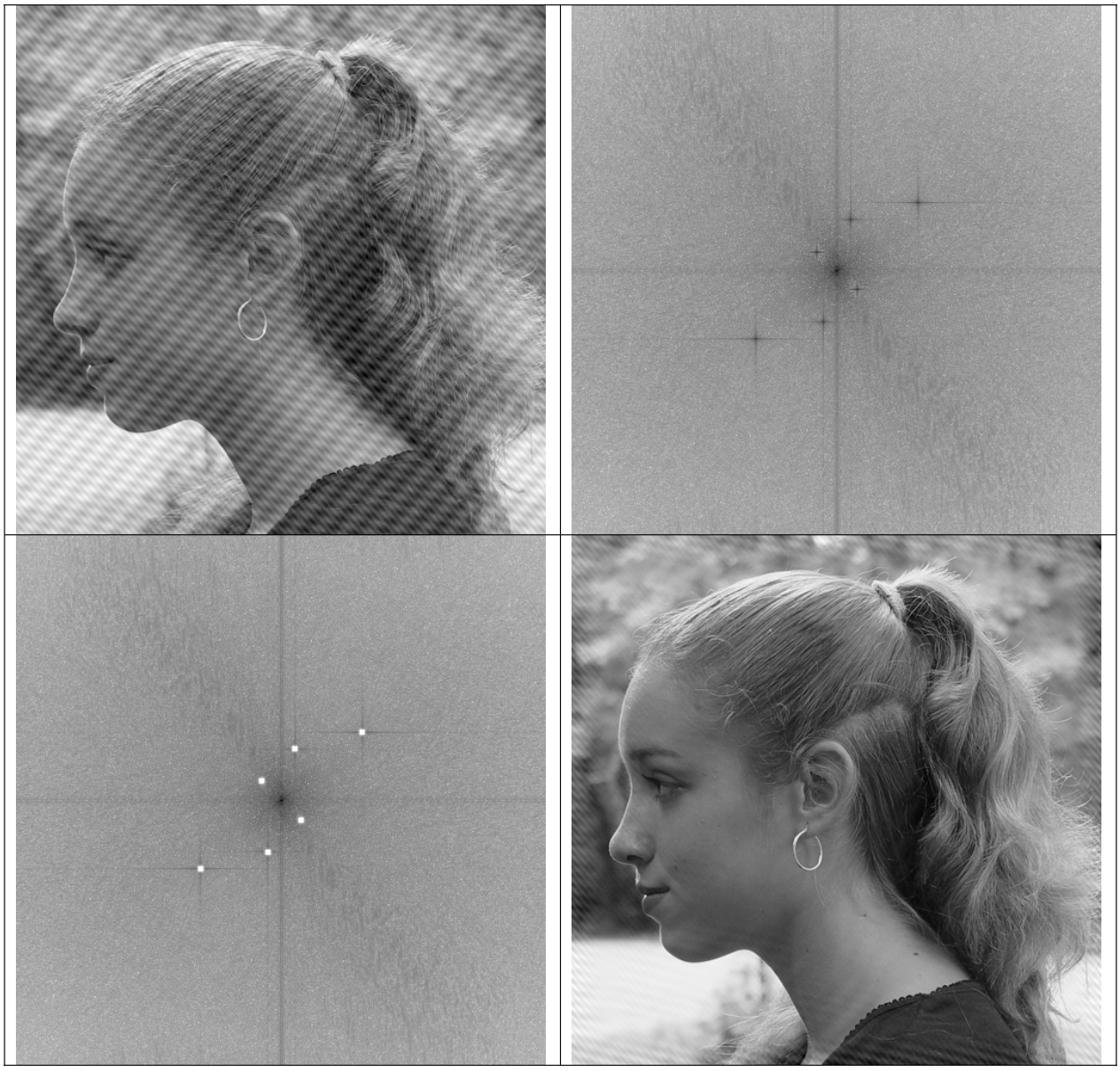


Figure 52. Removal of periodic noise by removing the spikes in Fourier space: a) image with corrupting lines; b) the Fourier power spectrum showing the spikes representing the lines superimposed on other information that represents the image; c) removal of the spikes; d) transforming the data back to the spatial domain removes the lines from the image.

The same procedure can be applied to images with more complicated periodic noise, such as the pattern of dots used in offset printing. **Figure 53** shows a picture scanned in from a newspaper. The dot pattern or screen is made up of quite a few sets of lines, which appear in the Fourier power spectrum with regular spacings and an orientation that matches the angle of the screen. Removing them as shown, either manually or automatically, allows the reconstruction of an intact image, with even subtle details like the cat's whiskers well defined. By comparison, blurring the image to eliminate the dot pattern also blurs the fine details and produces a very much poorer result.

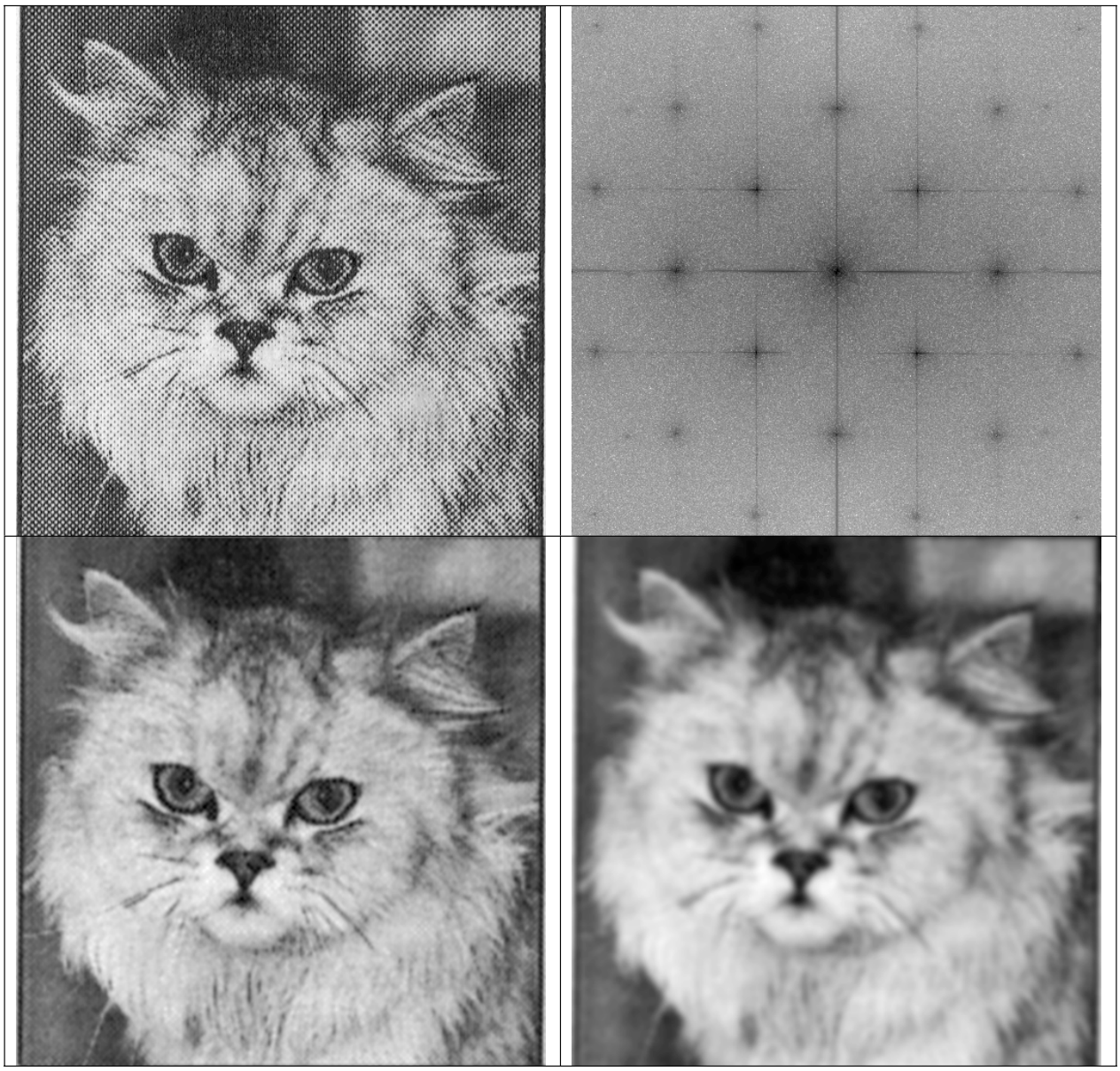


Figure 53. Removal of a halftone pattern: a) original image scanned from a newspaper; b) Fourier power spectrum showing the spikes that correspond to the halftone screen pattern; c) image obtained by removing the spikes and retransforming the image; d) result of blurring the original image to eliminate the dot pattern.

The preceding three figures showed monochrome images. For colored images the noise pattern is usually different in each channel. For example, some single chip digital cameras have patterns in the red, green and blue channels that result from the arrangement of each set of filters on the transistors. Halftone printed color images use screens for the cyan, magenta, yellow and black inks that are oriented at different angles. This produces different sets of spikes for each color, and the best solution is to convert the image to CMYK color space and process each channel separately. That was the procedure used in **Figure 54**.

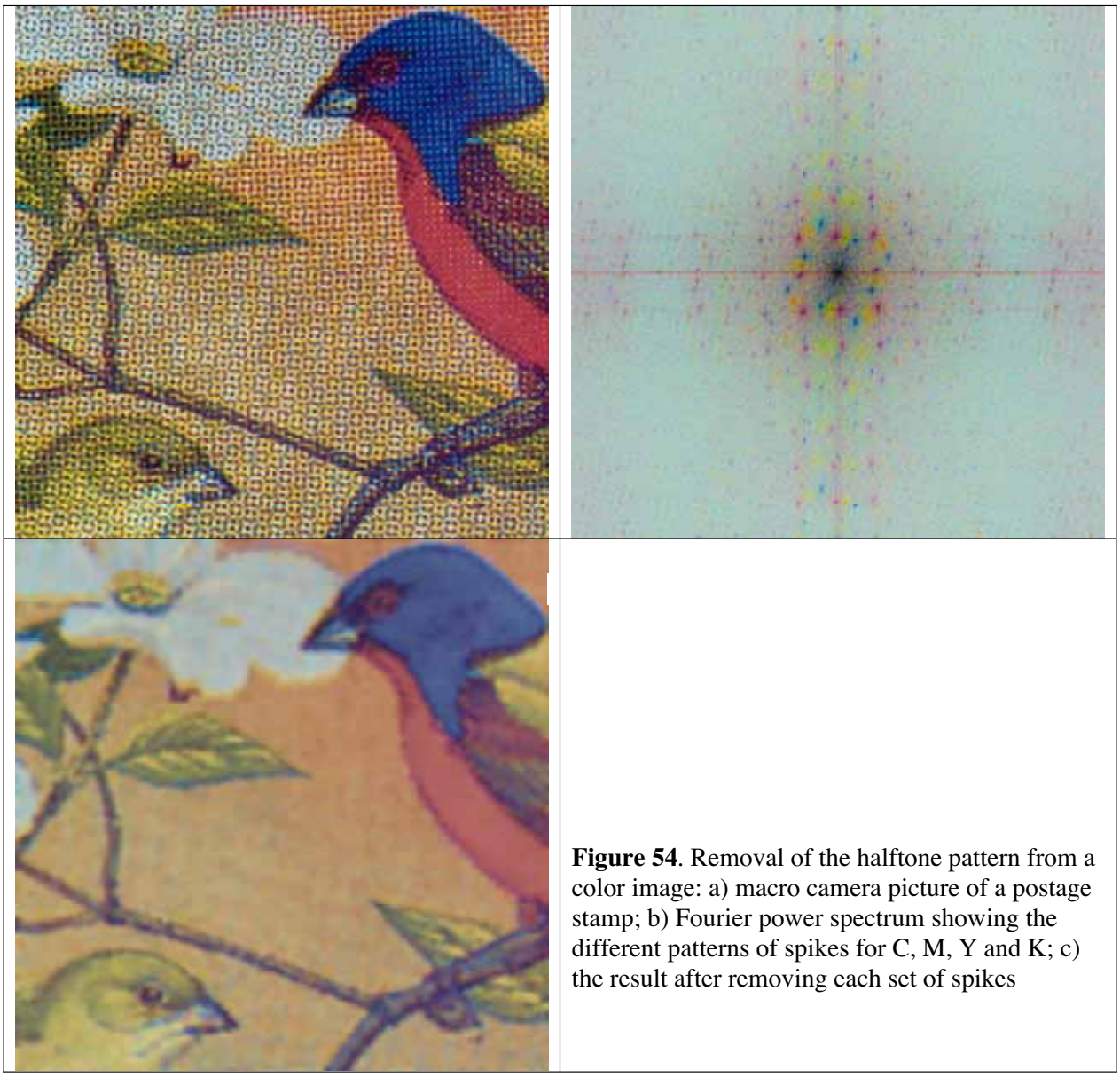


Figure 54. Removal of the halftone pattern from a color image: a) macro camera picture of a postage stamp; b) Fourier power spectrum showing the different patterns of spikes for C, M, Y and K; c) the result after removing each set of spikes

Filtering in frequency space requires a little practice, but no particular knowledge of with the mathematics involved. The examples shown here were performed using the Fovea Pro Photoshop plug-ins from Reindeer Graphics (www.ReindeerGraphics.com). Obviously this is an answer to a specific problem, which most digital photographers will not encounter in routine work, but which there is no other good way to solve when it does arise. Fourier space processing is also the key to removing other image defects not dealt with in this text, such as blur due to out-of-focus optics, or due to motion, and the plug-ins include tools for that purpose as well.

Cropping, zooming and interpolation

In the conventional photographic darkroom, the procedure of selecting the portion of the negative to print is so commonplace as to require no mention. The film negative, even from a 35mm frame, has high

enough resolution (limited by the grain structure of the silver halide particles or colored dyes in the film) that enlargement of even quite small regions can usually be accomplished without serious degradation of the resulting print.

Digital photographers haven't that luxury. The resolution of a digital camera image is much more limited. The 35mm film negative can resolve much finer detail than the best current digital cameras, such as a 5 Megapixel camera with a stored picture width of 2500+ pixels (and a real resolution, as discussed in **Chapter 1**, of perhaps 50-65% that value). Consequently, cropping and enlarging digital images is much more problematic.

Ideally, of course, the photographer would carefully frame the subject and the entire captured image area would be useful. In reality, one of the first rules of photography - especially when working with wildlife or children - is to grab the picture at the moment something interesting is seen, not to wait for the perfect moment that may never arrive. That means that some selection of the area to be kept or printed will certainly be needed.

It seems easy enough. The rectangular selection tool in Photoshop can be used to select a region of interest, and then Image->Crop will discard the surroundings. If this is too small, you can select Image->Size and enter in the dimensions that you want. you should check "Constrain Proportions" to prevent stretching the image in one direction or the other. The Resample Image selection allows you to choose several different ways to create the information that will be needed to fill in the gaps between the original pixels as the image is enlarged, and there are also third-party plug-ins that add other methods.

The criterion for the selection of an interpolation method should be the visual quality of the result, and this depends to some degree on the nature of the image and the preferences of the observer. The important thing to bear in mind is that when an image is enlarged, the original pixel values are spread apart. There is no information in the original image about what lies "between" the pixels so all of the interpolation methods use the values in the local neighborhood to create new pixels in the gaps. Varying degrees of mathematical effort are used to perform that calculation.

The discussion here involves the enlargement of images, which is the usual need. However, the exact same interpolation methods are used when the image is reduced in size, stretched or rotated. In all cases, each new pixel value is derived from the old ones by finding the corresponding location in the old image and using the pixels that surround that point.

The simplest method is to convert the coordinates of every pixel in the new image to a location in the original, see which pixel that point lies closest to, and use the values of that pixel for the new one. This "nearest neighbor" method has the advantage that it does not alter any of the brightness or color values in the picture, but that is the only advantage. The image appearance produced by nearest neighbor interpolation is blocky and visually distracting.

Bilinear and bicubic interpolation, both built into Photoshop, fit either a straight line or a smooth curve through the neighborhood. Bilinear uses just the four pixels that surround the new location, while bicubic uses a 4x4 neighborhood with 16 pixels. In practically every case, bicubic does the best job of producing a visually acceptable result.

Interpolation using larger neighborhoods allows fitting more complicated math functions such as spline fits (using a plug-in from Shortcut software, www.shortcut.nl) and sinc functions. In some cases these offer slight improvements in the perceived sharpness of boundaries, but in most cases the differences are

very, very small. In the case of boundaries with a large amount of contrast, the sinc function in particular can cause ringing, or overshoots of brightness and darkness on both sides of the edge.

Another technique, called fractal interpolation using a Photoshop plug-in called Genuine Fractals (www.Altamira-Group.com), manufacturers new detail to fill in the gaps. This is totally spurious, but in certain cases successfully fools the eye into thinking the image has more resolution than is actually present. More often, it makes the image look noisier than the other methods.

Figure 55 compares these different approaches using a representative situation. The picture of the heron was taken with an equivalent 105 mm lens, the longest optical zoom setting on the camera (digital zoom is just the same as interpolation in the computer, but with less control over the region or the method). This photo was taken before I tried to get closer for a better one, during which time the bird flew away. So the challenge was to crop and enlarge the central portion of the original to show the bird.

Figure 55: Cropping and enlargement: a) original image; b) cropped and enlarged to show region of interest.

Figure 56 shows enlarged details of the bird using the various interpolation techniques. Except for nearest neighbor, the differences between the different methods are very subtle. There is little to choose between spline fitting and bicubic interpolation. The important thing is to understand that any of these techniques cannot produce the same kind of enlargement results that film is capable of, and that enlargement cannot create meaningful detail finer than the resolution of the original picture.

Figure 56: Comparison of interpolation methods: a) nearest neighbor; b) bilinear; c) bicubic; d) fractal; e) spline.

Restricting operations to a selected region

Several of the examples shown in this and previous chapters have been applied to portions of an image rather than the entire pixel array. This is important for improving the appearance of photographs, since it enables selectively improving color balance, sharpness or other characteristics of each region or object in a picture. There are several ways to create, view and edit selections; some are better suited for some purposes than others so it is worthwhile learning about them.

Selections of pixels do not have to be contiguous, but can include a few pixels in one part of the image and many in another area, etc. To see exactly what the selection includes, it is often important to enlarge the image to at least 100% and scroll around the image to see the various parts of a selection. When making selections, it may be important to be able to see the entire image (which usually means that it must be reduced to less than 100% size on the display) and also the local details. Create a second display of the image by selecting Window -> Documents -> New Window and adjust the zoom settings and window sizes to fit your display and needs. The selection marquee lines will only appear on whichever of the two representations is active, but changes to the image will show up in both windows. Of course, there is no such thing as having too large a display area to work on.

Manual selection can be performed using the Lasso tool (**Figure 57**). Drawing a line around the intended region on an image is painstakingly slow, but sometimes the best or only way to achieve the desired selectivity. Holding down the shift key to add another region to the current one, or the option/alt key to remove a region from the current selection, allows you to adjust a selection region, including ones created using other more efficient methods, so the Lasso is often used as a final step to tweak a selection slightly.

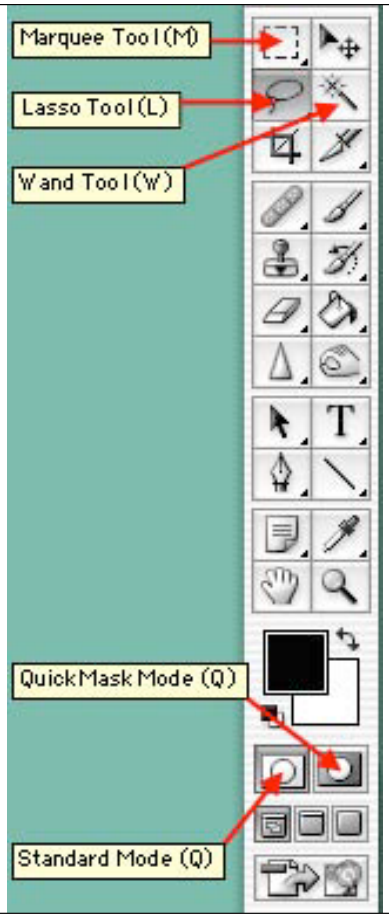


Figure 57: Selection tools: The Lasso, Marquee and Wand tools are shown, as are the buttons to select how the selection is displayed and edited.

The Lasso has three forms: the completely freehand version that draws as long as you hold the mouse button down, the Polygonal version that draws the region boundary in straight line segments between points clicked with the mouse, and the Magnetic version that tries to follow boundaries as you move the mouse close to them. The latter method rarely gives exactly the desired result without considerable

subsequent editing, and is primarily useful in graphics arts applications where the contrast is more consistent and boundaries better defined than in photographs.

Two other sets of tools are also shown in **Figure 57**. The Marquee tool is used to draw rectangles and ellipses on the image, and while it is very useful for cropping images it is rarely appropriate for selecting regions for correction or processing, because natural objects rarely have just ideal geometric shapes. However, rectangles and ellipses can be effective in creating masks for vignetting or framing.

The Wand tool is closely related to the color range selection described below (many Photoshop algorithms have several different ways to access them, each using a different user interface to access the same program functionality). When the user clicks on a point in the image using this tool, all pixels whose color values are the same as that point, within the tolerance setting on the toolbar, and which are connected to it by a continuous path of similar pixels, are selected. This can be extremely useful for selecting a particular object for processing, although some tweaking of the tolerance may be required.

Once a selection has been made with the wand tool (or any of the others, but this procedure is most often used after wand tool use), the menu selection Select -> Similar will find all other pixels in the image that have the same range of pixel values, but were not part of the original selection, and will also make them part of the selection.

The tolerance value for the wand applies to all color channels (R, G, B or C, M, Y, K). However, it is possible to be more selective about the color range in each channel. The selection mask applies to the entire image, but the wand tool that adds or subtracts pixels from it can be used in any individual color channel. By choosing one color channel (e.g., Red) in the Channels window, setting an appropriate tolerance value, and clicking with the wand (with the Shift or Alt/Option key depressed to add or subtract from the existing selection), pixels can be added or removed based on the range of values for just that one color. This may seem tedious and complicated, and it does require you to keep track of which channel is being displayed (another good reason to keep two copies of the image on view), but it can still be more efficient than completely manual outlining with the lasso tool.

In the example shown in **Figure 58**, selecting the boat is best performed with the wand because a global color range selection would also include the life vests. But in the red channel it can be seen that the boat is very similar to the muddy water in brightness, and in the blue channel it is similar to the rocks. Selection using several clicks on different parts of the boat (fore and aft, high on the side and down in the shadow area) in these two color channels can select all of the boat but get part of the surroundings as well. switching the image mode to Lab and using the Alt/Option click on the unwanted regions in the a (red-green) channel eliminates those pixels and leaves a good selection of the boat hull. It is then easy to use the Image -> Adjustments -> Hue & Saturation dialog to change the hue of the boat, from red to turquoise.

Figure 58: Selection using multiple channels: a) original picture with final selection region shown; b) red channel; c) blue channel; d) “a” channel (red-green) from Lab space; e) changing the boat’s color.

In the example, the selection process as carried out with the wand tool did not include the lettering or stripe on the boat. One way to add those pixels to the selection would be to use the Lasso tool with the shift key depressed to encircle the lettering. The way that was actually used, which also had the benefit of smoothing the boundaries of the region, was to use the Select -> Modify -> Expand function to increase the size of the selection region in all directions by 3 pixels, which filled in the letters, and then Select -> Modify -> Contract to shrink the selection region by the same three pixels. This returned the outer boundaries to their original location but the filled-in holes and small irregularities were no longer present.

This use of expansion (usually called dilation in the technical image processing field) and contraction (called erosion) is widely useful. Doing the expansion first fills in small voids and heals breaks in the selection. Reversing the order and doing the contraction first erases small regions, for instance where noise has been selected. The distance value in pixels for the two operations is usually the same in order to maintain the location of the outer boundaries, and large enough to eliminate the imperfections.

For most purposes, the manual selection of regions involves a lot of clicking, some undo-ing of steps and trying again, and a modest level of frustrating use of time. In many situations, a more efficient result can be achieved with the Select -> Color Range dialog. **Figure 59** shows an example. Clicking on a few points in the sky using the "+" eyedropper adds to the selection region all of the pixels whose colors match those at the chosen points, within the tolerance set by the "fuzziness" slider. This is not quite the same meaning as the wand tolerance, but both specify a tolerance range that is the same in all color channels. The wand tolerance sets a range of colors for pixels that are selected. The color range fuzziness controls the degree to which a pixel is selected based on how different it is from the chosen points. The idea that a pixel can be partially selected involves the values in the image's alpha channel and is discussed below.

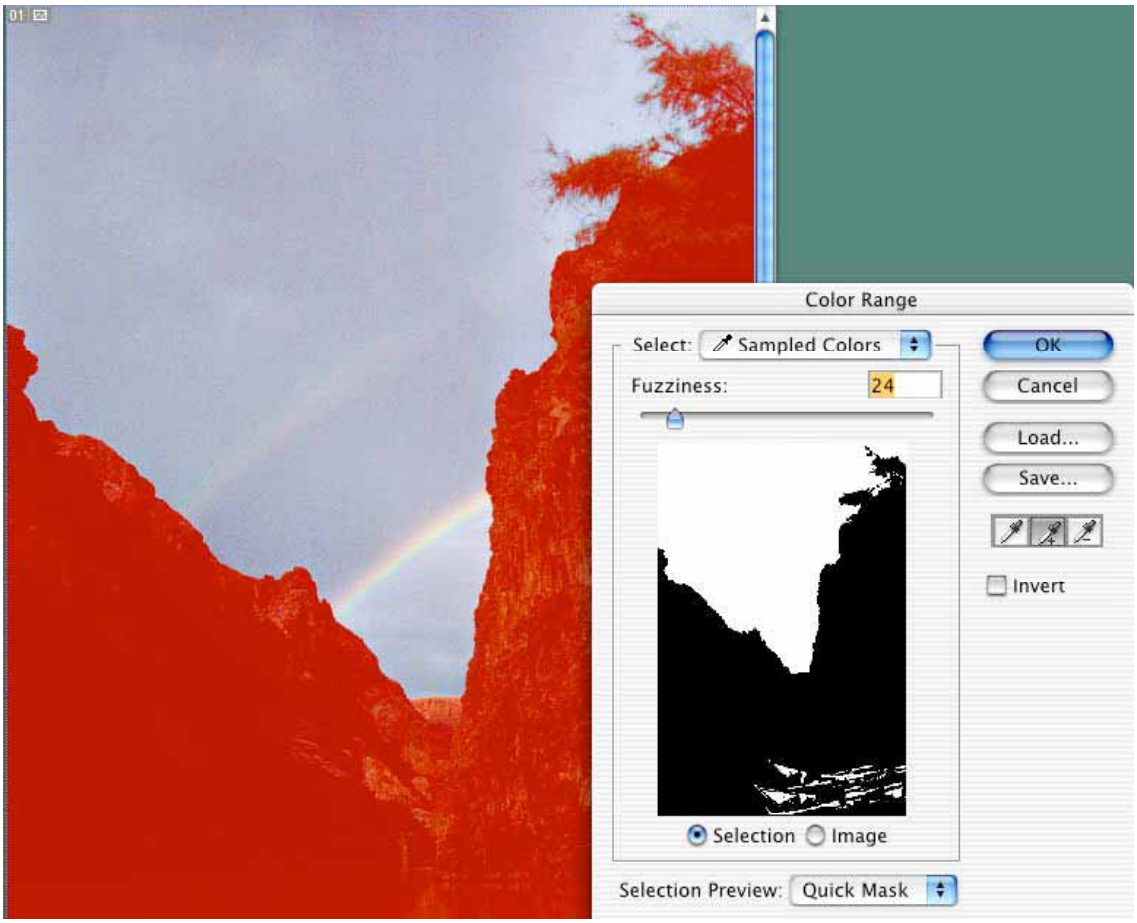


Figure 59. Color range selection: a) original image; b) final result after altering the color balance of the sky; c) selection dialog.

Unlike the wand method, the color range technique also includes some regions in the boats at the bottom of the picture (the wand can also be set to automatically include all pixels, not just those connected to the marked points). Exclusion of these unwanted regions is easy, by drawing a rectangular marquee around them with the alt/option key depressed.

Manipulating the color adjustment for the sky to increase the visibility of the first- and second-order rainbows and Alexander's band can then be done without altering anything gin the rest of the picture. In this case, increasing the saturation of the colors also increased the visibility of random speckle in the region, and a Gaussian blue was applied to reduce it. This was also confined by the region to the sky, where there are no abrupt edges or fine detail, and so does not affect the sharpness of any of the visually important foreground detail.

Color range selection is also commonly used in "blue screen" applications where a person or object is photographed against a background that has a uniform color not present in the object. That makes it easy to draw the region around the object, so that it can be cut out and pasted onto another image. You probably see that same method used every evening when your local weather forecaster appears in front of the daily weather maps.

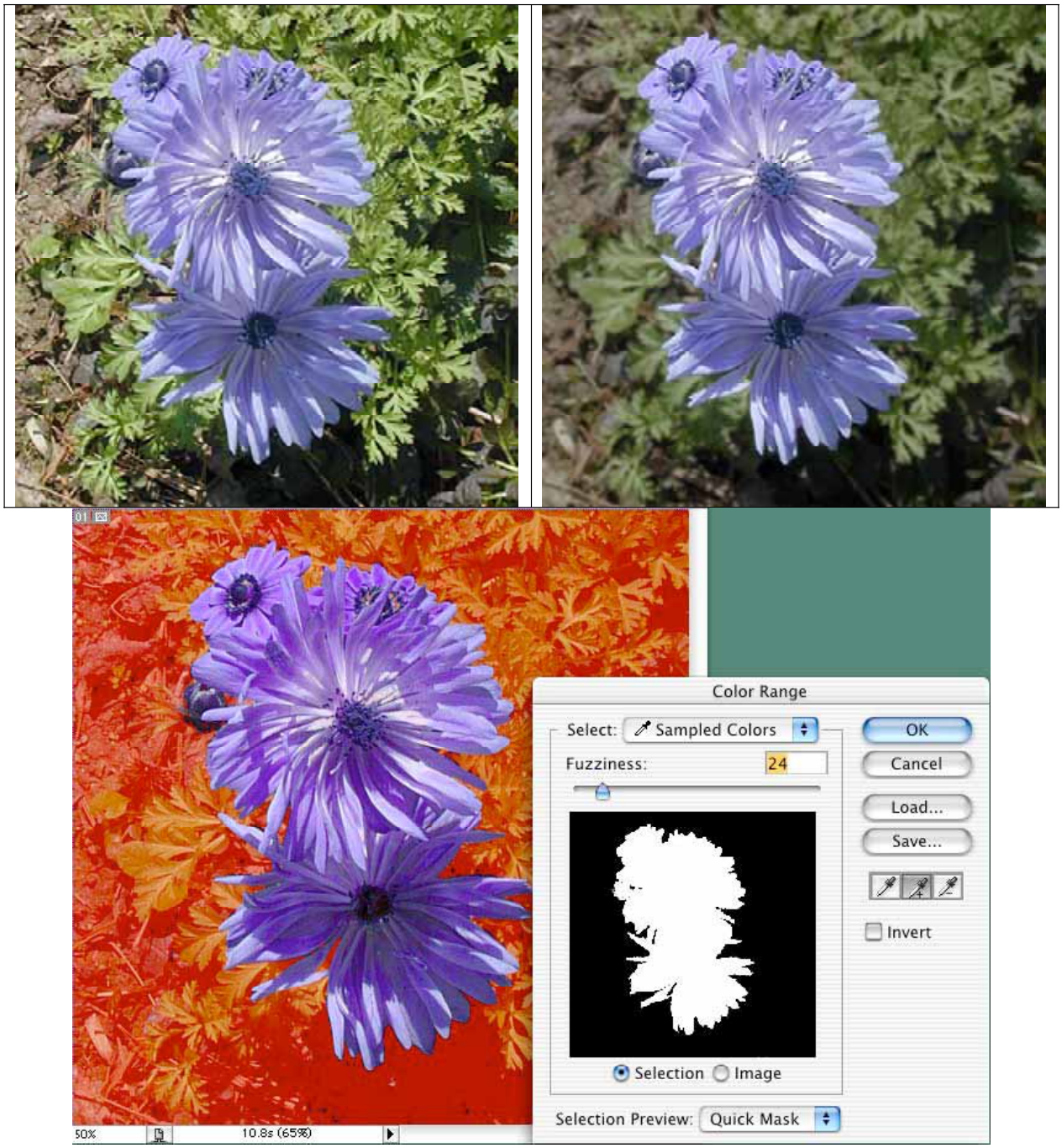


Figure 60. Inverting the selection: a) original image with blue flower on variable background; b) blurring the background and reducing the saturation and intensity; c) selecting the flower, after which Select -> Inverse changes the selection to the background.

In blue screen applications, it is the uniformity of the background that simplifies region selection. Once the background has been selected, then inverting the selection (Select -> Inverse) changes the sense of the selection so that pixels that were previously selected are not, and vice versa. There are many situations in which selecting the pixels you are actually interested in processing is harder than selecting

everything else and then inverting the selection. **Figure 60** shows an example. The background around the flower is partially green and partially brown, with a great range of brightness from leaf to shadow. The flower, on the other hand, is a fairly uniform shade of blue. The color range selection of the flower permits inverting the selection to all of the pixels that are not part of the flower. This background can then be subjected to a Gaussian blur, and the Image -> Adjustments -> Hue & Saturation function used to reduce saturation and intensity to make the flower stand out more from the background.

Masks and Selections

In the preceding examples, pixels have either been included in the region or excluded from it, so that processing or color changes can either affect the pixel or not. However, it is possible to exert much more control than that. The amount by which each pixel can be affected is actually adjustable in 256 steps, from entirely to not at all. The amount is controlled by the picture's alpha channel, which has an 8 bit range and can be viewed as an image in which black = 0 means the pixel is completely protected and not affected by processing and white = 255 means the pixel is 100% effected by processing.

Instead of the selection display that has been used above, in which the marquee of black and white pixels (also called the "marching ants") outlines the regions, another mode of display is available. It was used above in Figure 57b. Called the "Quick Mask" mode, it shows the alpha channel as a colored overlay on top of the image. The greater the amount of color at each pixel, the more that pixel is "protected" from processing. Clicking on icons in the Tools palette shown in Figure 55 allows switching between the standard mode of display in which the regions are outlined, and the quick mask mode in which the colored overlay appears (to change the color of the overlay, double-click on the quick mask icon).

Adjusting the mask value for pixels has many uses. One of the most common is to avoid an abrupt change in values for pixels at the edge of a region. For example, when a region is used to limit the action of the clone tool, any abrupt discontinuities at the edge of the region attract visual attention. Smoothing the edge over several pixels hides these changes as shown in **Figure 61**. This edge smoothing, called feathering, can be accomplished in several ways. One is the Select -> Feather menu function, which operates when the standard display mode of regions is in use. Note that the standard display cannot show whether a pixel is "partially" selected. The marquee line is drawn at the 50% contour line around the selection edge. Pixels just inside the line are more than 50% but less than 100% selected, while those outside the line are less than 50% but more than 0% selected.

Another way to feather the selection boundary requires displaying it as a quick mask. In that display mode, the alpha channel contents are treated by Photoshop just like an 8 bit grey scale image, and all of the processing routines and plug-ins can be applied. Performing a Gaussian blur on the quick mask blurs the edges and produces the same result as feathering the selection boundary in standard display mode, except that the result can be seen in the quickmask by the variation in the intensity of the colored overlay.

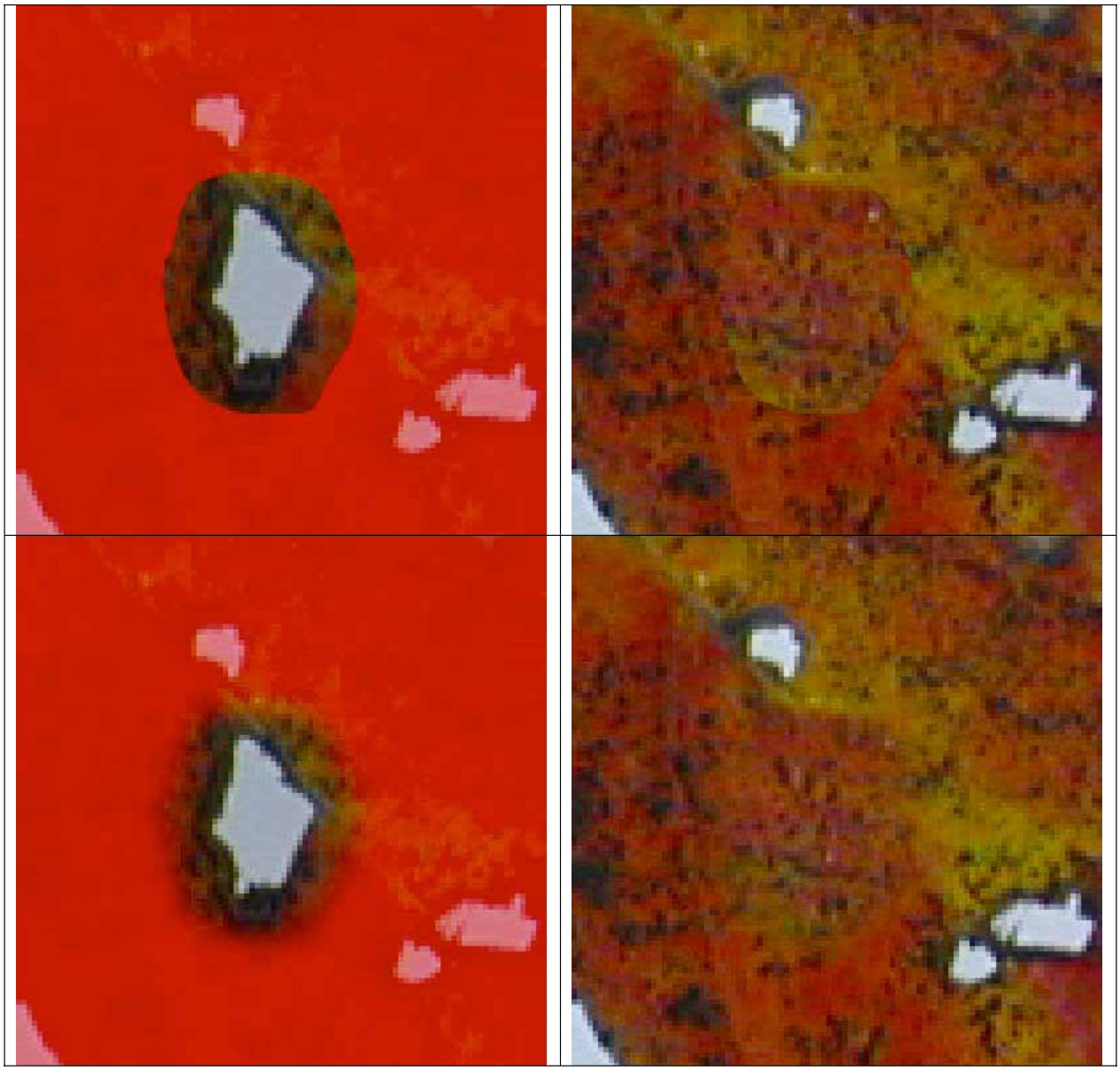


Figure 61. Feathering a region boundary: a) hard edge; b) resulting visible boundary of cloned pixels; b) feathered edge; d) resulting hidden boundary of cloned pixels.

Other processing tools can also be applied to the quickmask, or a grey scale image can be pasted into the quickmask. This offers some very useful possibilities for creating a mask that controls the blending of images. **Figure 62** shows an example. While the unsharp mask is a commonly used technique for improving the visual appearance of images by increasing the local contrast at edges and detail, it can produce an objectionable increase in the visibility of noise in smooth areas in the image. In the Photoshop unsharp mask dialog, it is possible to set a threshold to only apply the algorithm to pixels whose difference from their local surroundings is large, but there is another way to gain even more specificity and control.

In the example, two duplicates were made from the original image. One was sharpened using an unsharp mask (for the purposes of this demonstration, it was over-sharpened; the edge contrast is too exaggerated particularly around the eyes and mouth, and noise is evident on smooth areas such as the model's forehead and cheek). The second copy was converted to grey scale and the edges outlined using the Find Edges routine. Then the three images were combined by selecting the original and choosing Image -> Apply Image from the menu. The dialog shown in the figure indicates the procedure. The pixel values from the sharpened image are blended into the original image using the mask image value to control the amount. Because the original mask image is black at the edges, it is inverted in the dialog. The result is that the sharpened values are used but only at the major edges, and only in proportion to the magnitude of the edge. Smooth areas in the original image are not modified.

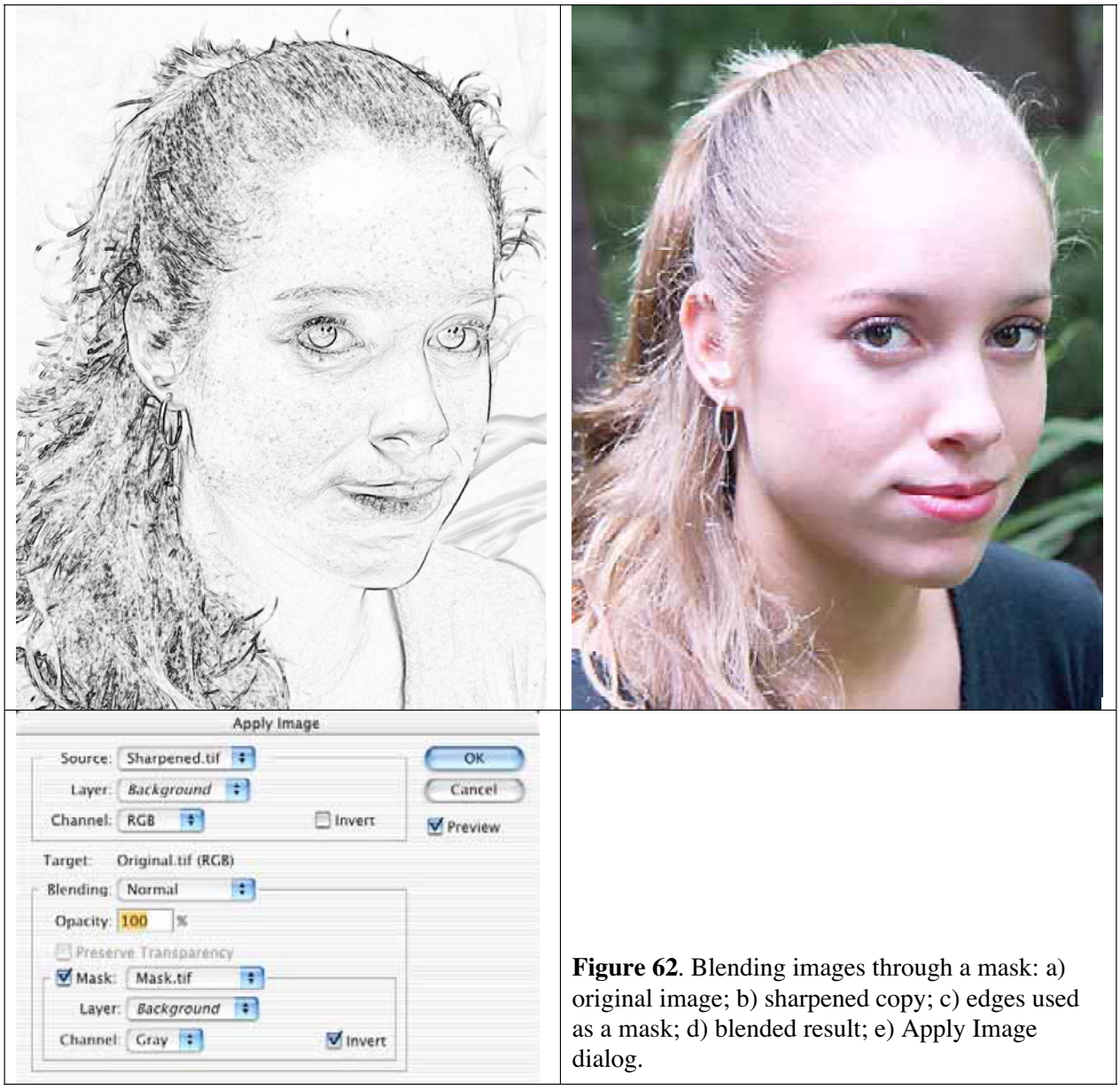


Figure 62. Blending images through a mask: a) original image; b) sharpened copy; c) edges used as a mask; d) blended result; e) Apply Image dialog.

This is a simple and commonly used procedure for controlling the unsharp mask, and it can be extended to other situations, including controlling the use of the median filter to eliminate lines and scratches, etc. Any mask image can be employed in this way, and tools such as setting levels or adjustment curves, smoothing, etc., can be applied to the mask to control the relative contribution of the original and processed images.

As is usual in Photoshop, there is another way to carry out this process. By placing the images in layers, and using the mask image to control the blending, the exact same result can be achieved. That has the advantage that you can modify the mask or the processed image and see the final effects immediately. However, the order of the images in the layers is critical and it is easy to get confused. Until you are comfortable with the technique, performing the operation with the Apply Image function has the advantage of clarity and simplicity.

Only 8 bit per channel images can be combined using the Apply Image function, or placed in layers, so unless you use plug-ins such as Optipix from Reindeer Graphics this blending technique does not apply to 16 bit per channel images. However, other uses of selections to control processing or color balance in specific regions are still useful with 16 bit per channel images. The problem is how to create the selection. The wand and color selection tools cannot be applied to 16 bit images (nor can any of the drawing tools), and there is no quickmask display mode. It is possible, however, to have an 8 bit alpha channel for a 16 bit image.

The best procedure for creating the alpha channel is to duplicate the 16 bit image and convert the copy to 8 bits. Use this to create the selection in any of the ways described above. Then choose Select -> Save Selection from the menu. This saves the selection as an 8 bit alpha channel (the same as that seen in the quick mask mode) in the current image. Then bring the 16 bit image to the front and choose Select -> Load Selection, choose the image and alpha channel you have just saved, and the selection region will be displayed in standard mode on the image. Although only the selection outline is shown (as noted above, the line is drawn at the 50% values in the alpha channel), the usual meaning of the mask values is still in effect. Processing of the 16 bit image can now be used and will affect only the selected region. There can be multiple channels saved in an image, so you create different selections and save them with individual names, and load each one as required to process portions of your image.

Layers

Selections, layers and masks can be confusing in Photoshop because there are so many different options and many different ways to accomplish equivalent techniques to arrive at the same result. Part of this confusion also stems from the fact that the layers modes in Photoshop are intended to address several quite different purposes. One is the compositing of elements to create a graphic arts layout. Since this includes the placing of text onto other layers, it does have some applications to working with photographs, and will be discussed below. A second, very different purpose is to apply corrections to images, including color balancing, sharpening, and other processing, either of entire images or portions of images. A third is to combine multiple images, for example copies that have been processed differently. But it is also a way to align and combine arrays of images that form a mosaic. The following sections will concentrate on those operations and layer modes that are most useful for working with photographs.

In their simplest form, layers can be imagined as a stack of photographic negatives. Viewing through the stack when placed on a light table produces a visual combination of them. In Photoshop, with the viewing mode is set to “normal,” the opacity control can be adjusted for each layer (except the bottom or background layer) to control the extent to which that layer affects the light passing through. At 100% opacity, nothing behind the layer can be seen at all. At 50% the layer is averaged with the image coming from layers behind it.

Opacity of a layer is quite different from the transparency introduced above, which varies within a layer. Individual pixels can be made entirely or partially transparent regardless of their color or density, and regardless of the layer’s overall opacity. This transparency is controlled by a mask, also known as a selection mask or an alpha channel. It is in effect another 8 bit grey scale image that has the same dimensions as the image or layer. For each pixel in the image, the corresponding pixel value in the mask sets the transparency. Since this mask can also be viewed (and edited and processed) as an image, it is useful to know that the 0..255 scale of values (0 = black, 255 = white) controls the transparency. Black

makes the pixel completely transparent, while white makes it fully opaque. But if the overall layer transparency is less than 100%, it will still be possible to partially see through the opaque pixels.

When one image, or a portion of an image, is selected and copied to the clipboard, it can be pasted onto a layer in another image. This is the classic way that Photoshop enables people to place one person's head on somebody else's body. In that case, it is probably easier to select the head than the body, but in most situations there is always the question of which image to perform the selection on. In the example of Figure 61, the original canyon picture was exposed to obtain good contrast and adequate brightness on the rock walls, which are in shadow. The consequence of this choice is that the sky visible at the end of the canyon is overexposed.

Pasting the canyon picture into a layer on top of a sky picture makes it possible to select the overexposed sky, cut it out, and so reveal the sky behind. Selecting the bright area where the sky should be is easily accomplished using the wand tool. In **Figure 63**, two different sky images were used, one with clouds and one without. In both cases the background sky layer was shifted so that the prettiest part of the sky showed up in the gap (choose Select -> All, click on the proper layer in the Layers window, depress the Control/Command key, and use the mouse or arrow keys to move the layer). The ability to assemble an image from fragments of others provides a great deal of flexibility. Turning layers on or off by clicking the eye icon in the layers window makes it easy to compare various combinations.

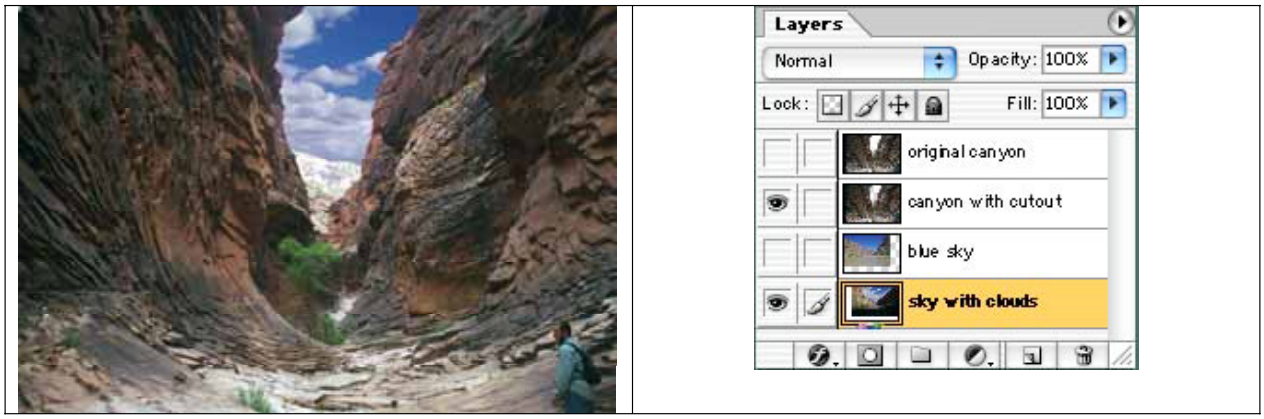


Figure 63. Assembling a scene: a) original canyon photo; b) image of clear blue sky; c) image of sky with clouds; d) composite of a and b; e) composite of a and c; f) layers window.

Certainly the simplest type of layer is one containing text. This is a different type of layer than the preceding example, since it does not consist of pixels forming an image. Instead, the information on the font, style, color and size selected, and the text that has been typed in, are used to draw ideal characters on top of the layers beneath. All of the rest of the layer is transparent. This layer can be edited (to change the text, font, color, etc.) and shifted relative to the other layers, but none of the conventional processing operations can be applied to it unless it is first converted to a normal pixel layer (in which case the text is no longer editable). Besides the ability to edit the contents of a type layer, the other advantage is that the text is no aliased and will always display or print at the highest resolution of the screen or printer.

Another type of layer that is neither type nor image is an adjustment layer. All of the adjustments shown above - levels, curves, hue and saturation, etc. - can be applied to a layer without permanently altering its contents. Selecting Layer -> New Adjustment Layer -> Curves (or any of the other choices in the menu) creates an adjustment layer just above it in the stack of layers. The dialog shown is the same as for the normal image adjustment routine, and so is the visual result. But because the effect has been created as a layer, it can be edited at any time, or turned on or off by clicking on the eye icon for the adjustment layer, or turned partially on or off by adjusting the opacity slider for the layer. Several different adjustment layers with different settings can be created and used separately, or several different types of adjustments (e.g., curves and hue/saturation) can be used in combination. Furthermore, each adjustment layer can have an associated mask (alpha channel) that defines which pixels in the image it is applied to (**Figure 64**).

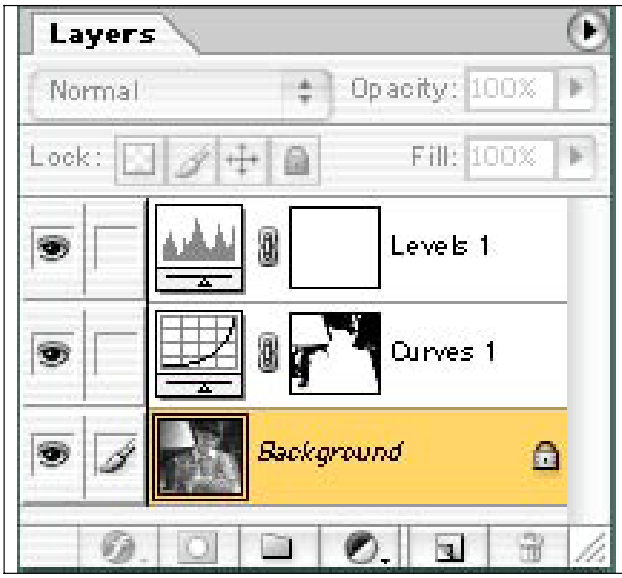


Figure 64. Adjustment layers. The Levels layer applies to the entire image but the Curves layer has a mask.

This provides a very flexible set of tools for manipulating the appearance of each region of an image and combining the results. As seen in some of the preceding examples, it is also easy to keep several copies of the original image that have been processed in different ways in layers for quick comparison, or for combining together. It also allows keeping the original image and all of the modifications together and yet individually accessible. On the other hand, a file with many layers can become fairly large, and is in a format that other programs cannot read.

If you haven't mastered the individual tools for controlling image contrast and brightness and the ways in which images can be combined, the potential complexity of the layers can be confusing. Start out with simple things and as they become familiar, add more.

Layer blending modes

One of the simplest uses of layers is to keep copies of the image at various stages of processing. Clicking on the eye icon in the layers window turns on or off the visibility of each layer. This makes it very quick and easy to visually compare the appearance of the image before and after each step. If multiple images of the same scene have been acquired (for instance, using different colored filters, or different focus settings), they can also be compared in this way. Copying one image and pasting it into a layer on top of the other is a great help in making sure they are perfectly aligned, either by clicking the top layer on and off to observe the shift, or by viewing both layers at once.

Leaving the layer blending option set to normal and adjusting the opacity of the top layer to about 50% shows the average of the two layers, and if they are misaligned this is very apparent. Choosing Select -> All from the menu and shifting the top layer by holding down the Control/Command key and either dragging with the mouse or using the arrow keys allows you to adjust the alignment while seeing both images. Nudging with the keys is best for fine control, but works only in steps of whole pixels. For subpixel shifts, the Optipix (ReindeerGraphics.com) nudge plug-ins allow fractional steps in X or Y.

Figure 65. Detecting changes: a, b) two images of the same scene at different times; c) 50% opacity, normal mode (identical to adding the images together); d) difference mode; e) darker mode; f) subtraction (note that even the hour hand has moved enough to be visible).

Instead of viewing the two layers with normal mode and 50% opacity, it may also be helpful to change the layer blending mode to “difference” and set the opacity of the top layer to 100% (in all of the other

discussions of layer blending the opacity is 100%, unless otherwise specifically noted). This mode shows the absolute value of the difference between the two layer images, so if they are exactly identical and aligned the difference image will be black. In the example of **Figure 65**, difference mode is used to show the changes that have taken place in the scene in the time interval between the two pictures.

Adjusting the opacity in normal mode is identical to adding together various fractions of the two layer images. Difference mode shows the minute hands and the pendulum, which is not in the same position. Setting the layer blending more to "Darker" shows both minute hands. Subtraction is not provided for in the blending modes, but can be accomplished using the Apply Image procedure described above.

The primary use of layers for photographic processing is to combine the results of processing with the previous image. We have seen several examples of this in the various processing techniques described earlier. For example, the result of an unsharp mask can be blended into the original using normal mode and opacity control, or just the light or dark edges can be kept by selecting a blending mode of lighter or darker, respectively.

Many of the layer blending modes seem to have no particular application to photographic processing, but to be intended for various graphic arts effects. They are described (albeit briefly) in the Photoshop User Guide, and the user is also encouraged to try them and observe the results as a way to determine when or if they might be useful in certain situations.

The most useful blending modes for photographic images are illustrated in **Figure 66**. The background layer is the original image, and the upper layer that is blended with it shows the edges of the features, produced by using the Photoshop "Find Edges" filter. These edges can be used to modify the original image in several different ways using the various blending modes. Some of these modes produce results that look more like paintings than photographs.

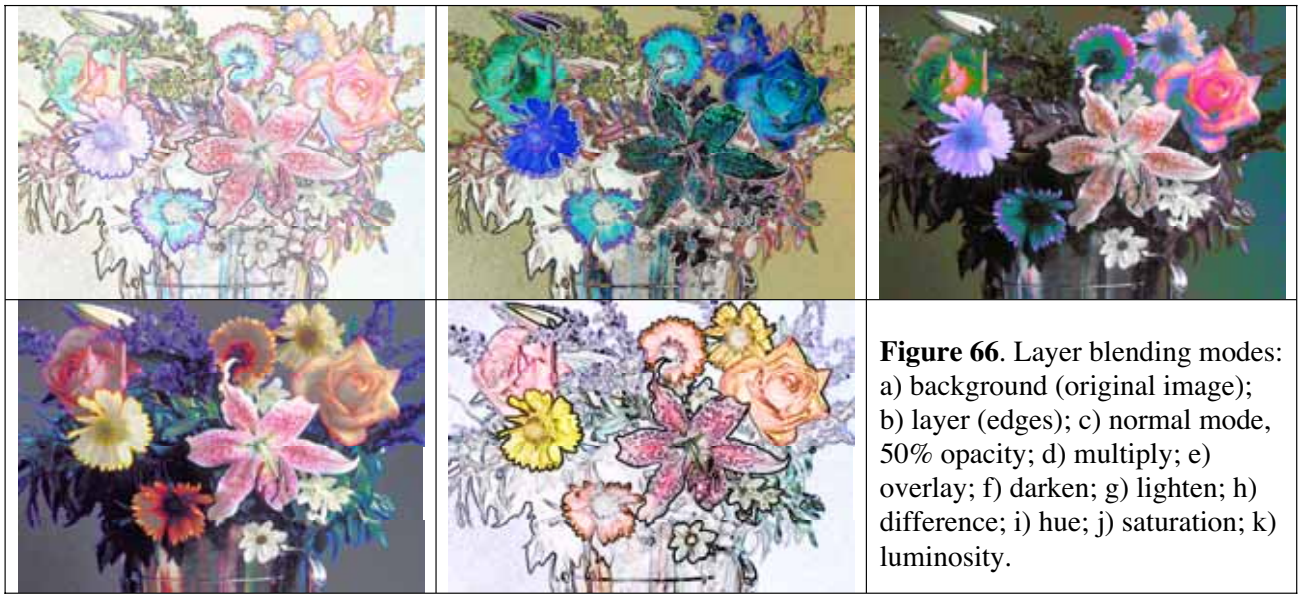


Figure 66. Layer blending modes: a) background (original image); b) layer (edges); c) normal mode, 50% opacity; d) multiply; e) overlay; f) darken; g) lighten; h) difference; i) hue; j) saturation; k) luminosity.

Multiply darkens the image because each pixel in the background is multiplied by the value in the layer, using the convention that white=255 is a multiplier of 1.0 while dark=0 is a multiplier of zero. Hence the edges, which are the non-white values in the layers, are darkened more than the uniform regions. Overlay also performs multiplication, but instead of using the intensity in the layer as the multiplier the individual channels are mixed together, which generally brightens the color values in uniform regions while darkening the edges. Darken and lighten keep either the background or layer pixel values depending on the intensity value. Difference shows the absolute difference of the color values (calculated channel by channel). Hue, saturation and luminosity use the intensity value from each pixel in the layer image to replace the hue, saturation or intensity value for the background image, while keeping the other values. For instance, the saturation mode keeps the hue and intensity from the background but uses the intensity in the layer image as the saturation for the resulting pixel color.

Layers can also be combined using a mask, as was done above (**Figure 62**) using separate images and the Apply Image function. In the example of **Figure 67**, the aerial photo has the typical problem for such pictures of a bluish atmospheric haze, which increases with distance. Two duplicate layers were separately adjusted to give good color balance and contrast for the far distant terrain and for the nearer terrain. As shown, these have very different appearances. The far image was also given a slight Gaussian blur to reduce speckle noise that became more evident as a result of the increase in contrast. In order to blend them together, the top layer selected and the Layer -> Add Layer Mask -> Reveal All menu item chosen. Then, by clicking on the icon for the mask in the layers window, a vertical linear gradient was drawn that blended a different fraction of the two images as a function of vertical position.

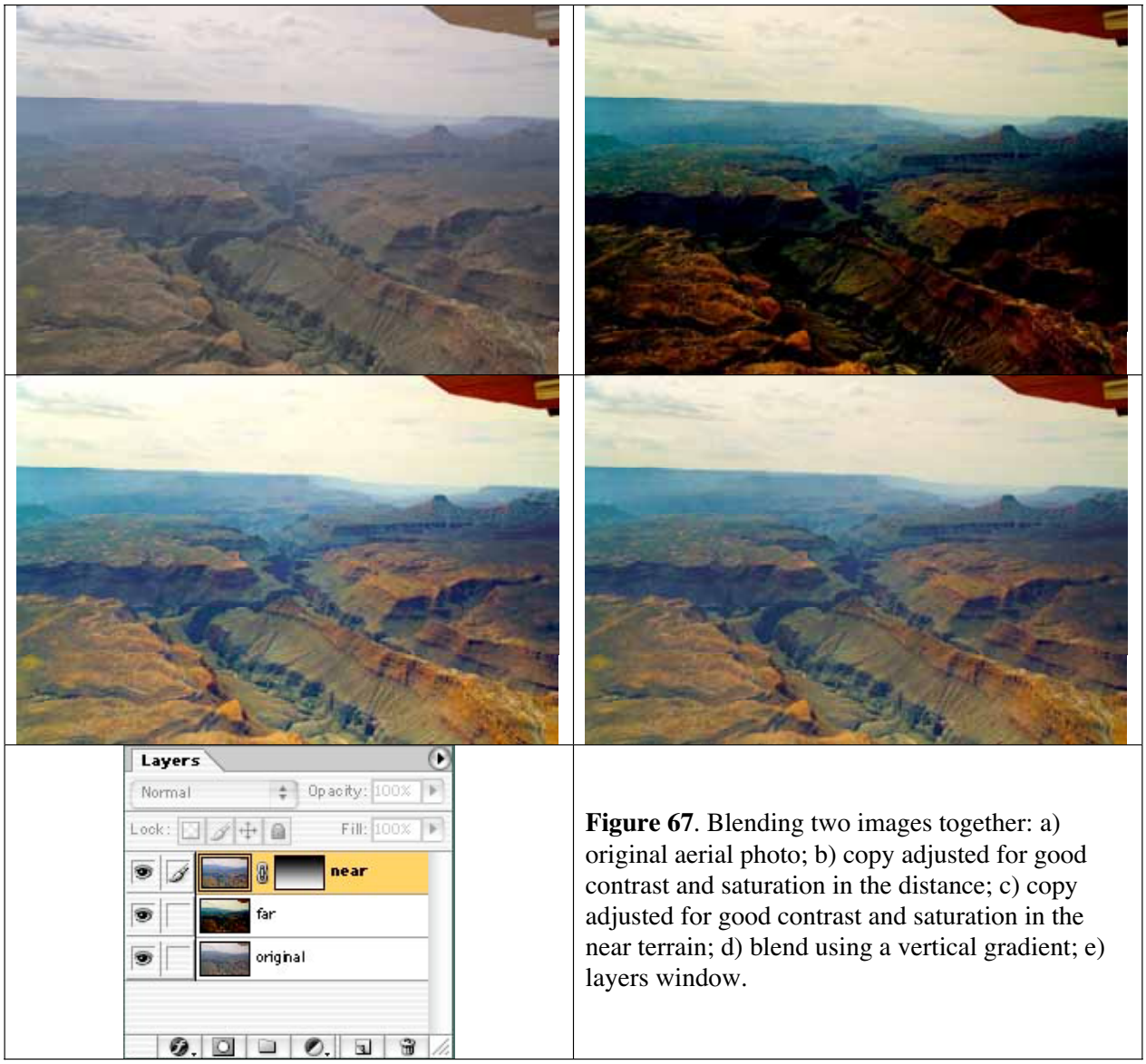


Figure 67. Blending two images together: a) original aerial photo; b) copy adjusted for good contrast and saturation in the distance; c) copy adjusted for good contrast and saturation in the near terrain; d) blend using a vertical gradient; e) layers window.

Manual editing

Every photographer encounters situations in which some retouching of images is required. Generally this is much easier to do with digital images in the computer than with chemicals and negatives in the darkroom. And while there are certainly some practitioners of this art with the skill of a renaissance painter who can create entire believable perfect images using the graphic arts tools in programs like Photoshop, it is not our intention here to do more than minimal necessary cleaning up of images.

In most cases the defects to be removed are actually present in the scene, although in a few instances they may arise from hardware limitations. Few of the things we photograph are perfect. Models have skin blemishes which may not be entirely covered by makeup. In a photograph, clothing may show a tear or stain that was not noticed beforehand or could not practically be corrected. A hastily grabbed action shot may have something unwanted in the background. Utility wires may intrude on an otherwise

idyllic scene. Or, if friendships or politics change over the years, it may seem necessary to remove someone from a group photograph. All of these situations call for retouching of the image.

The first rule to follow in attempting retouching is to always keep a backup copy of the original image, and to save the results of your work in stages as each part of the job is accomplished. This will save a lot of handwringing and grief later on. Secondly, spend a little time getting familiar with your tools. Some people feel quite comfortable using a mouse, or even a touch pad or trackball, to manipulate the paintbrush tool or the cursor used to select regions, while many others prefer a pressure-sensitive drawing tablet and stylus. Whatever you use, practice is needed to apply the tools so that your hands follow the thoughts your mind creates.

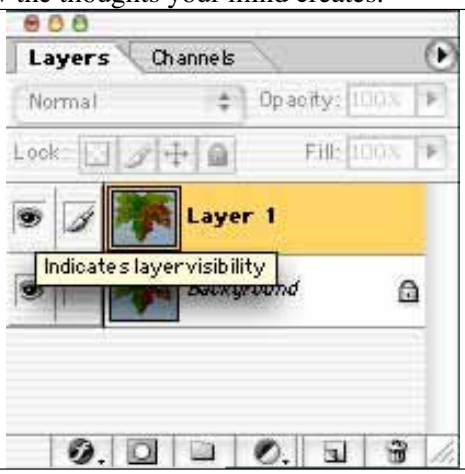


Figure 68. Clicking on the eye icon in the Layers palette turns the visibility of a layer on or off.

Most of the Photoshop tools discussed in this section can only be applied to 8 bit per channel images. Given that limitation, it makes sense to make extensive use of Layers (which are also enabled only for 8 bit images). Before working on a particular defect, create a duplicate layer and apply the tools to that copy of the image. Turning the layer visibility on and off (click on the eye icon in the layers palette as shown in **Figure 68**) makes it easy to compare the original image to the edited one, to be sure you are accomplishing what you wanted to do without altering things that should be left as they were.

The most basic editing tool is the paintbrush, which can be used to apply any selected color to the pixels in the image. For retouching of photographs this is almost always useless because areas of uniform color stand out as being quite artificial. The paintbrush tool is appropriate for graphic arts work, not photograph retouching.

Cloning, Healing and Patching

Much more useful is the “Clone” tool. This is also a paintbrush, whose size you can select, that copies values from one location in the image and places them in another. You would normally select a size that is no larger, and usually somewhat smaller than the defect you wish to cover up, and make sure that the brush has a very soft edge that will gradually blend values into the surroundings (**Figure 69**). An abrupt edge between the edited region and the original surroundings is very easy for the eye to detect, and should be avoided.

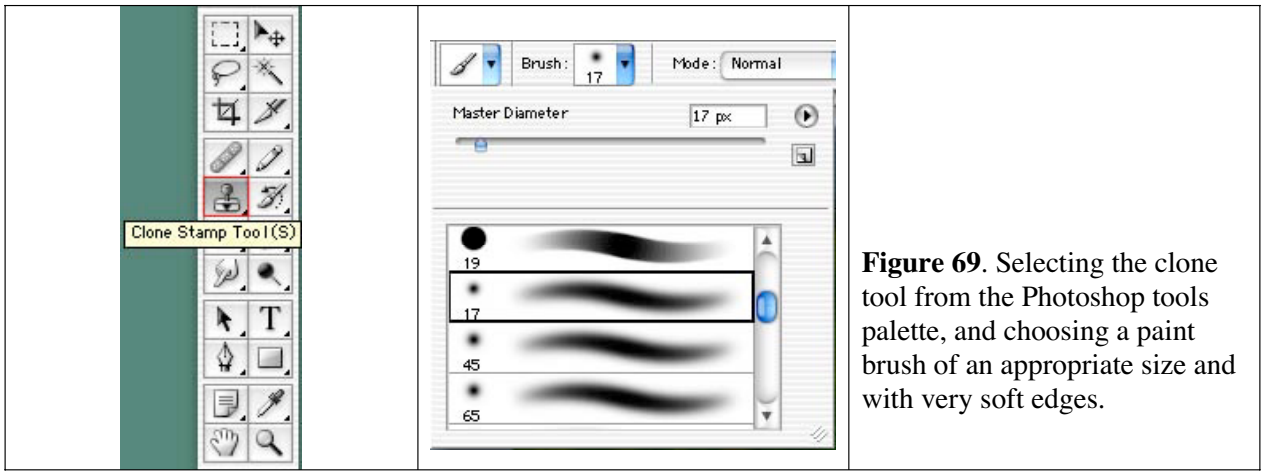
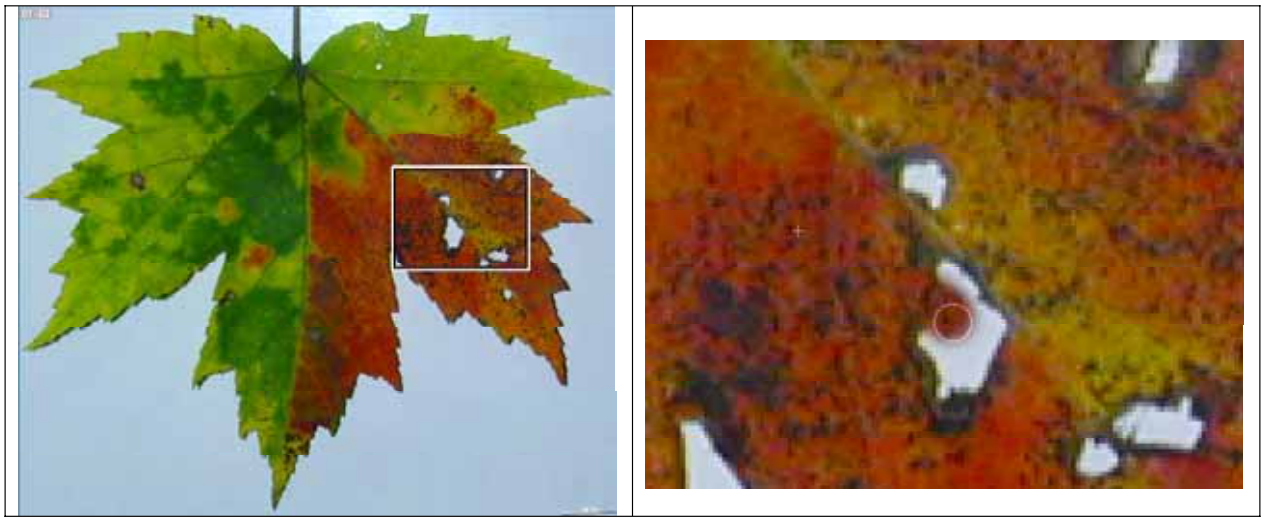


Figure 69. Selecting the clone tool from the Photoshop tools palette, and choosing a paint brush of an appropriate size and with very soft edges.

To use the clone tool you must first select the location on the image where the pixel values are to be copied from. Hold down the option/alt key and click on this point. Then when you click on the defect, an additional cursor will appear on the image to show where the clone tool is picking up the pixel values to apply to the selected defect area. The pickup point will move as you move the cursor, so that when you have painted over the defect, it will be an exact copy of the cloned location.

That is fine for removing small defects, such as covering up a blemish on a model's skin or perhaps removing a utility wire from a picture. But if the colors in the cloned region are even slightly different from those surrounding the region the pixels are copied to, the result may appear incorrect to the viewer. Also, the exact repeat of detail from the cloned region is likely to be visually recognizable, and being unnatural will reveal the editing (**Figure 70**).



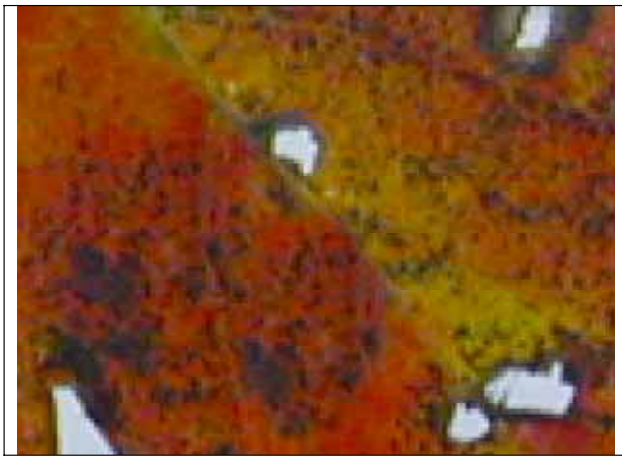


Figure 70. Leaf image showing holes: a) original showing enlarged region; b) enlarged region with clone tool being applied (circle indicates brush size, cursor mark indicates source location); c) clone tool result (notice exact texture repetition).

In Photoshop 7, another pair of tools was introduced (**Figure 71**). The "Healing brush" copies the texture from one location to another, but keeps the average color values that were present in the pixels being edited. Unlike the clone tool, the pickup point does not move when using the healing brush. Instead, the texture in a region about a dozen pixels wide is used. The brush size should be larger than the defect. Painting a new texture over incorrect colors will not heal a defect, but for narrow lines and scratches the tool samples the adjacent colors and the combination of the neighborhood color values and the texture from the pickup point produces a convincing result, as shown in **Figure 72**.

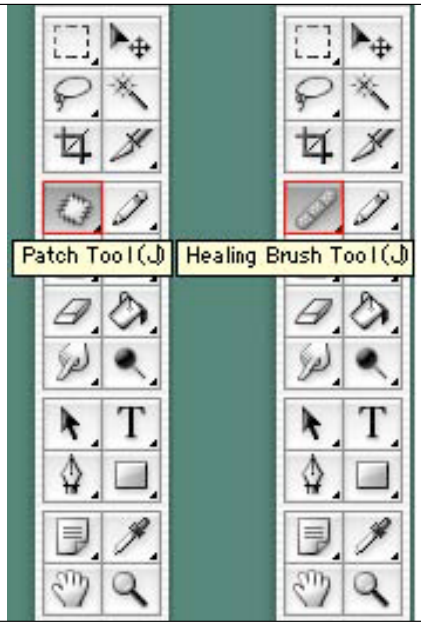


Figure 71. Selecting the Healing Brush and the Patch Tool from the Photoshop tools palette.

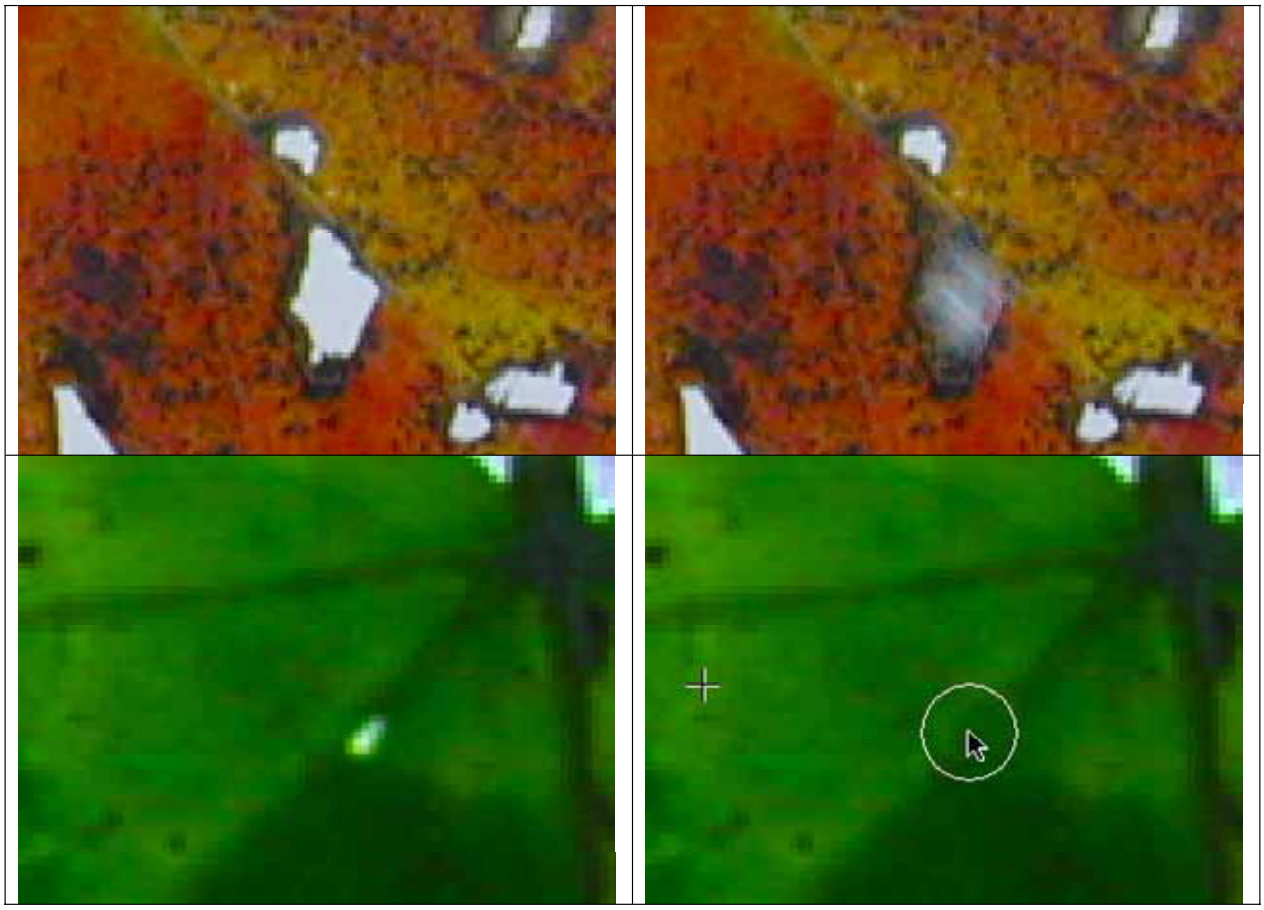


Figure 72. Editing with the healing brush: a) original image showing hole defect; b) the healing brush copies texture but not color - the brush size is too small for the defect; c) another area of the image showing a smaller defect; c) successful healing using a brush size larger than the defect width (brush size and pickup point are shown).

The use of texture from elsewhere in the same image is very important. In many instances people have tried to paste in parts of one image onto another (putting celebrity bodies on other heads is a common example), and the results are usually very easy to detect. Even if they get the sizes right, and have images in which the lighting and shadows are consistent, the different image texture provides a key clue. The texture results from the characteristics of the camera and the electronic gain needed for the lighting used, as well as the actual texture in the scene. Applying a texture that actually comes from elsewhere in the image makes the forgery much more difficult to detect.

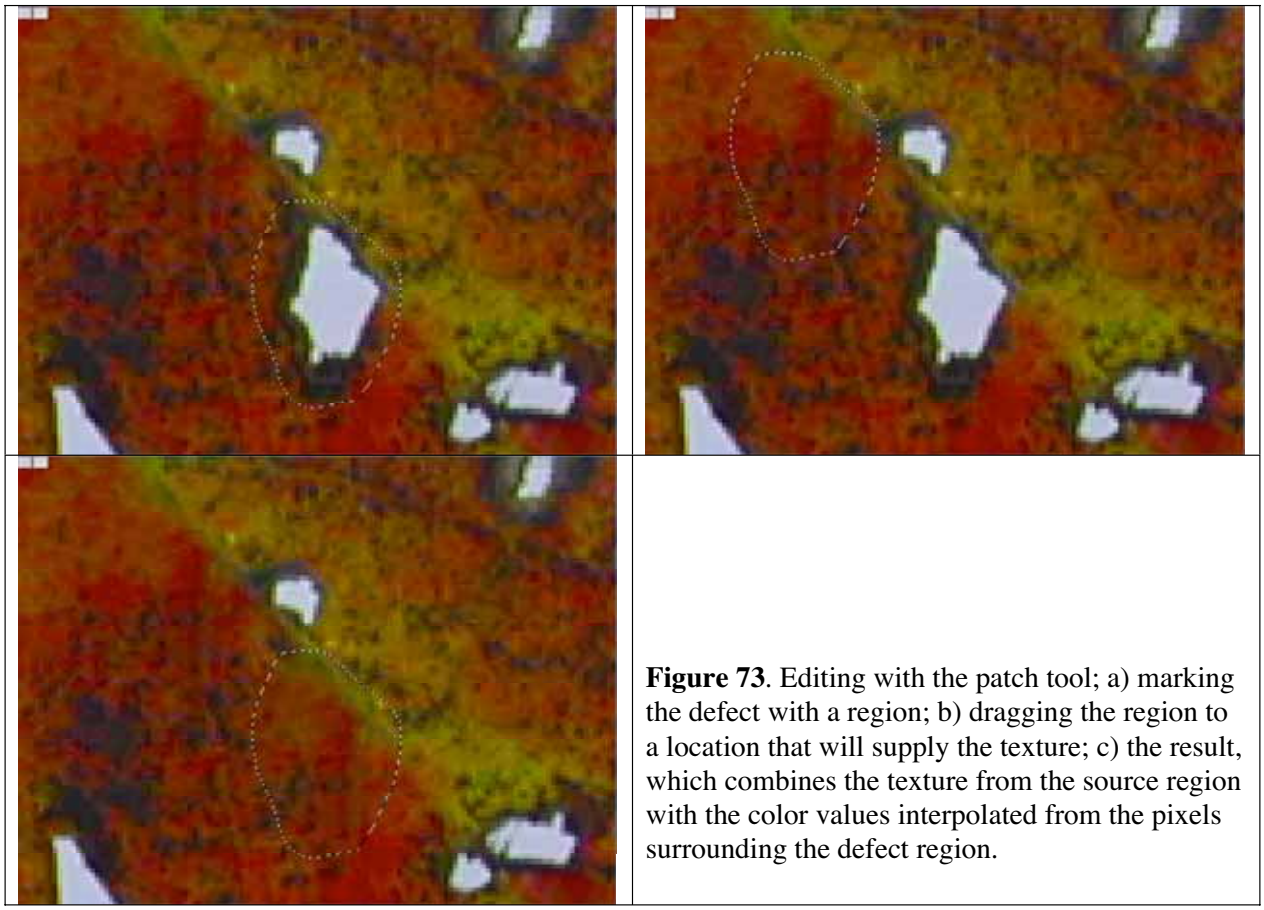


Figure 73. Editing with the patch tool; a) marking the defect with a region; b) dragging the region to a location that will supply the texture; c) the result, which combines the texture from the source region with the color values interpolated from the pixels surrounding the defect region.

The idea behind the healing brush, combining texture from one region with color values from another, is extended to larger area defects using the "Patch tool." With this approach you use a region selection, larger than the defect, to bring in the texture from another area. As shown in **Figure 73** this blends in a realistic texture from one area with the color values that surround the defect region. Doing this for all of the holes in the leaf took only a few minutes, and removes all of the defects without leaving any visible evidence, as shown in **Figure 74**.

Figure 74. Filling the holes in the leaf: a) original; b) final result.

Figure 75. Removing braces: a) original; b) after filling the regions.

Filling in regions from their periphery is a valuable tool in many instances. In the example shown in **Figure 75**, the girl's picture shows braces on her teeth. For cosmetic reasons, it may be desirable to remove them in the photograph. Drawing a region around the braces using the freehand lasso tool, and then filling the region using the same techniques as above, removes the braces and replaces those pixels

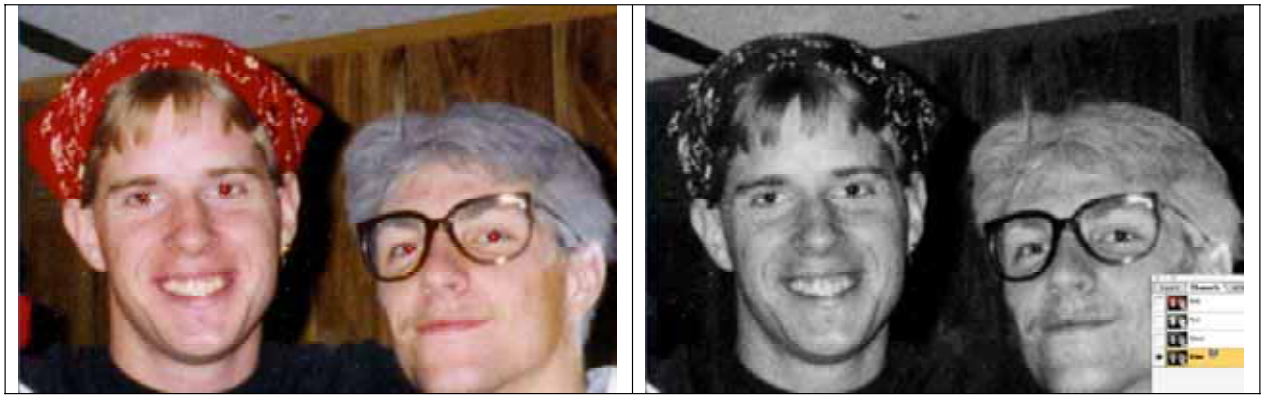
with values consistent with the nearby teeth. The same method works very well for removing utility lines, tears or stains on clothing, etc.

Even when simply painting in with the clone tool is being used, it is wise to draw a bounding region around the defect to prevent accidental straying onto the surroundings. When a region is used, in order to prevent an abrupt edge or discontinuity from appearing at the boundary, it is wise to feather the boundary. This means that edited changes will be gradually reduced in magnitude as they are applied to pixels near the boundary. From the menu, choose, Select->Feather and choose a width for the feathered region that depends on the scale (dots per inch) of the picture and the size of important details, so that the transition occurs over a reasonable distance. In the leaf example above, a feathering of five pixels was used.

The tools in Photoshop can be used for extensive and detailed retouching of images. For photographs from digital cameras, the need for this type of editing should only arise when there are blemishes on the original subject or scene. But for other related applications such as scanning in photographic prints or negatives, particularly old ones, retouching may be a very important requirement. This is a major subject in its own right, and the interested reader is referred to books such as “Photoshop Restoration and Retouching” by Katrin Eismann and Steve Simmons where the procedures and possibilities are explored in detail.

Red Eye

Particularly in snapshots, photographs of eyes exhibit a different problem - red eye. Reflection of light, particularly from an on-camera flash unit, from the retina in the eyes of people and animals produces a bright, colored spot in the pupil of the eye. The color of the spot depends on the absorption of light in the eye: people have red-eye while cats generally have green-eye. Either way, it is a problem easily correctable by editing of the stored image. (The other way to correct it is to prevent the light reflection in the first place. An off-camera flash won’t send the reflected beam straight back to the camera lens, and even many snapshot cameras have a “red-eye reduction” setting that flashes the light twice, once before the actual photo to cause the iris of the eye to stop down so not as much light will be reflected.)



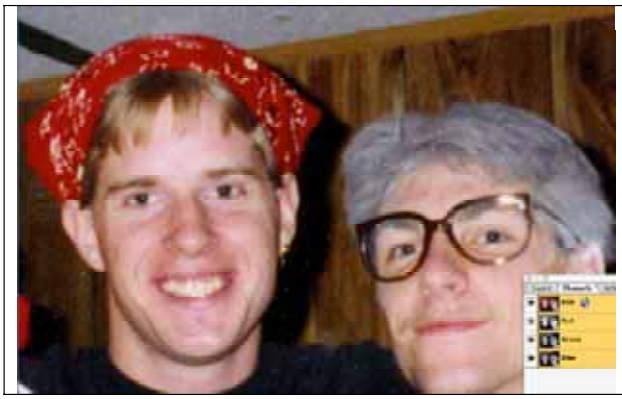


Figure 76. Red eye removal: a) original image (an enlarged fragment of a scanned snapshot), with selection regions around the iris and pupil of each eye; b) the blue channel, from which the pixel values are copied; c) the result of pasting the values from the blue channel into the red channel.

In the example of **Figure 76**, the red eyes in the image scanned in from a snapshot print are quite evident. Certainly it is possible to use the paintbrush to paint out the red, but there is a much simpler and quicker procedure. Use the elliptical selection tool to draw regions of interest around the pupils of the eyes. Feather these selections slightly (a width of 2 pixels was used in the examples). Then in the channels window select the blue channel and choose Edit->Copy. This copies the information in the blue channel, where there is no strong reflection from the eye. Next select the red channel and choose Edit->Paste to insert the pixel values that were copied. Clicking on RGB then shows the entire color image, from which the red-eye reflection has been removed. This method preserves highlights that result from reflection of light from the front surface of the eye, which are usually desirable.

For the example in **Figure 77**, a different approach was used. To select the yellow region in need of edition, the wand tool was used. Then the Adjustments -> Hue & Saturation function was selected to reduce the saturation in the region, and also to drop the intensity somewhat). This allows the original detail in the area, which was obscured by the strong yellow reflection, to be seen.

Figure 77. Yellow-eye removal for a cat picture: a) original image; b) final result; c) selection of the yellow regions; d) reduction of saturation and intensity.

Dodging and burning

Chapter 2 dealt with adjustments to contrast that can be made globally by adjustment of the Levels and Curves for the entire image. Those changes can also be restricted to regions. A region can be defined by manually drawing it on the image (using the rectangular or circular Marquee tools, or the freehand or polygonal lasso tool), or by using the Select -> Color Range menu entry to define a set of colors to be edited. It is also possible to use one of the color channels to create a selection by thresholding and pasting it into the Quickmask. These various methods were discussed fully in a preceding section of this chapter. Usually it is wise when modifying the contrast or color in a selected area to use feathering of the selection to smooth the transition between the modified and unmodified regions, to prevent a visible boundary from distracting the viewer.

There is another way to carry out local contrast and brightness changes for editing purposes. It is interactive and manual, but allows selective reduction of glare spots, or brightening of dark shadows, etc. Similar to the editing performed in the previous section using the clone tool, it employs a brush, and the size of the brush and smoothing of the brush edges can be adjusted by the user. The Dodging and Burning tools are selected from the tools palette (**Figure 78**).

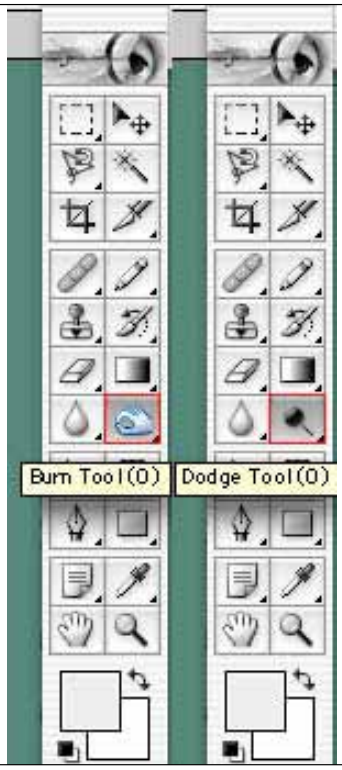


Figure 78. Selecting the Dodging and Burning tools from the Photoshop Tools palette.

The names come from common darkroom practice. Dodging is accomplished by partially blocking the light from the enlarger from falling onto selected regions of the paper, so that the exposure of the paper is reduced and the image is lightened. Burning is the reverse - light is blocked to the rest of the image while additional exposure of the selected region is carried out, producing a darker result. There is also a sponge tool that mimics the effect of sponging developer onto the print, to increase or decrease the local saturation of colors.

Figure 79. Parameter selection in the tool bars for the Dodging and Burning tools.

Using the brushes on the image, either by moving them back and forth over the intended region, or repeatedly clicking the mouse button on a region, applies the effect to either lighten or darken the selected pixels. Some control over the action of the brush is possible by entering the amount of exposure to be affected (larger numbers make a greater difference) and by selecting whether dark shadow areas, mid-tones, or bright highlights are to be altered. These adjustments (**Figure 79**) are not very precise or selective, and the tools will modify all of the pixels in the region to some extent.

Localized Sharpen and Blur

Just as dodging and burning are localized “paint brush” implementations of contrast adjustments that can be applied to either the entire image or to selected regions of the image, so too there are localized versions of the Gaussian blur and unsharp mask filters. These are selected from the tools palette (**Figure**

80) and have adjustments in the tool bar (**Figure 81**). Like the dodging and burning brush tools, the sharpening and blurring brushes have effects that are cumulative with moving across a region or repeated clicking on a spot. The tools can be applied to the intensity, or to various color channels. There is no adjustment of the smoothing or sharpening parameters.

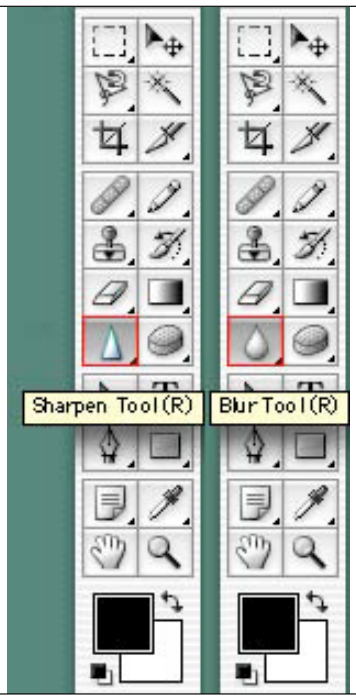


Figure 80. Selecting the Blur and Sharpen tools from the Photoshop Tools palette.

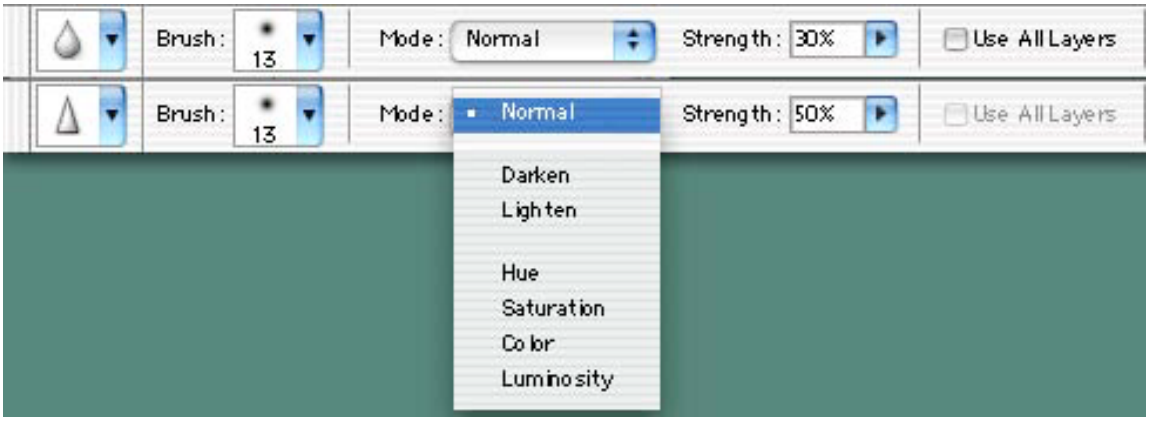


Figure 81. Parameter selection in the tool bars for the Blur and Sharpen tools.

While these tools can be effective in particular circumstances, it is usually better to apply the regular Gaussian blur filter or unsharp mask filter to a feathered region for localized blurring or sharpening. If this is done after duplicating a layer, then selective transparency of the layer can be used to blend in any desired amount of the original and processed image, which gives more control over the final result.

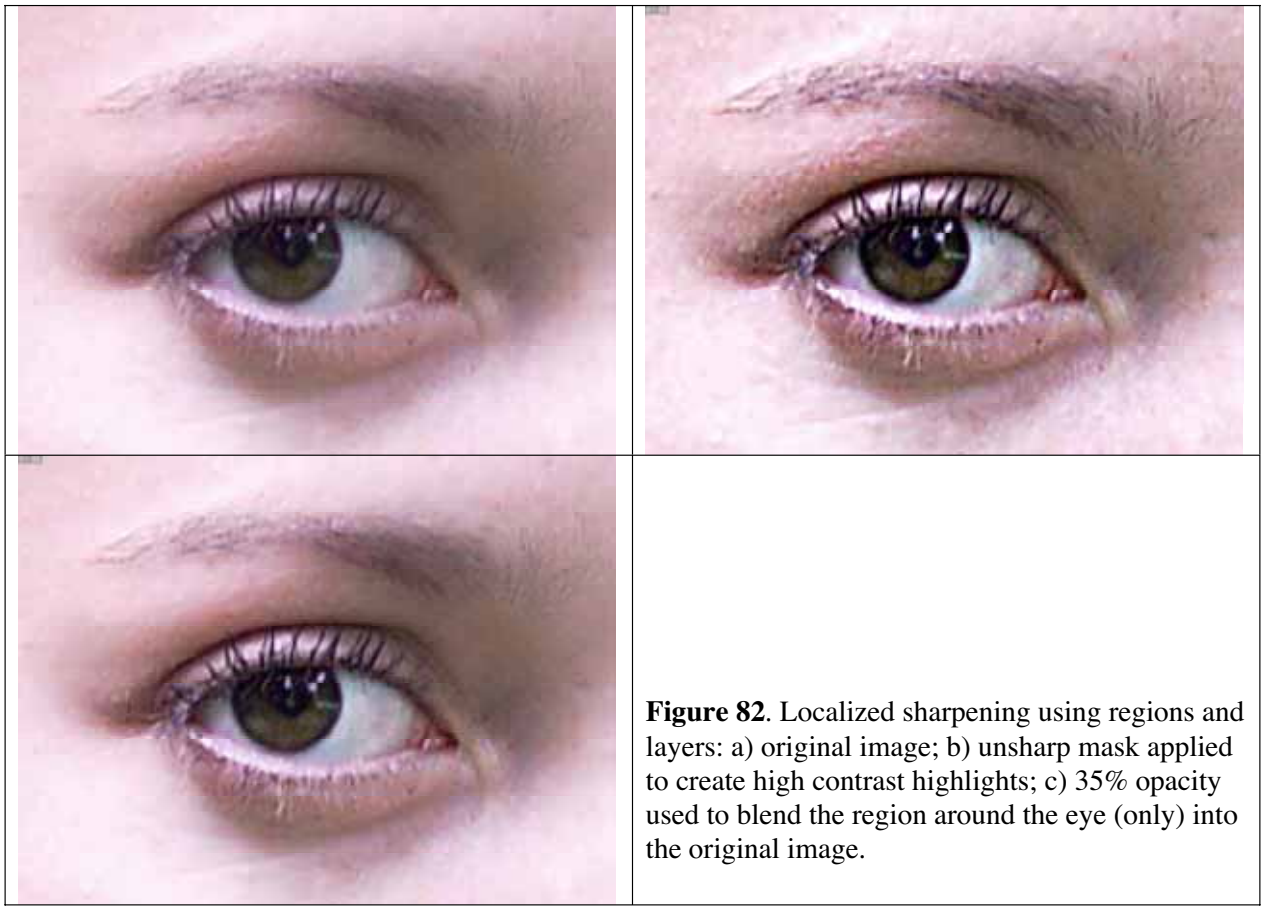


Figure 82. Localized sharpening using regions and layers: a) original image; b) unsharp mask applied to create high contrast highlights; c) 35% opacity used to blend the region around the eye (only) into the original image.

Figure 82 shows an enlarged portion of an image to which sharpening has been applied, and the result of adjusting the opacity of the layer and restricting the region using a feathered region. The sharpened image shows much more localized contrast, but this also reveals many superficial marks on the skin which are not desired in the final image. Drawing a freehand region around the eye and feathering it to blend the selected area into the surroundings, allows erasing the top layer except for the eye. Then adjusting the layer opacity to about 35% allows the sharpened detail in the eye to be effectively averaged with the original image so that the highlights complement the image rather than dominating it.

Chapter 5
Printing images

If we only intended to view our digital photographs on our own computer monitor, the color space in which the adjustments and optimizations were performed would not matter. The advantage of using a consistent color space to hold the corrected copy of the image, and depend on device-specific ICC profiles to translate those color values to the modified values required to get the same result from other devices, especially printers, is key to being able to making the widest possible use of images.

Monitors are output devices, too. The ability to get consistent color on different monitors is, on the one hand, easier than dealing with printers. The emission of light from phosphors covers a fairly broad range of colors (the "gamut") and the phosphors are well characterized, so that the relationships between colors are well known. User adjustments of brightness and contrast can alter this, and so can aging of the tube. As discussed below, LCD screens present something of a problem, because of the effect of viewing angle on perceived color. But in most cases monitors are easier to calibrate than printers, which have many variables including the inks and papers that can be used.

On the other hand, displaying images on a web site that is accessed via the internet does present a daunting challenge to obtaining consistent color: most web browsers do not pay any attention to the original color space for an image, and just display the values with whatever settings the user's monitor may have. Obviously, this leads to enormous variations in the appearance of images on different computer screens. The variation is not helped by the fact that the native screen display of Windows computers, which represent the majority of all systems, has a gamma that is different from the native display of Macintosh computers, which represent the majority of systems on which serious image processing is performed. Putting up helpful grey scales or providing information about image calibration on a web site seems to be a waste of time, as few visitors will pay any attention to them but will judge the images as they appear without questioning whether they look the same as on the original machine.

At least with paper printouts, everyone is looking at the same ink and paper. But, of course, not under the same lighting conditions, so consistency is still an issue.

CIE color space

To sensibly discuss ICC curves, color space conversions, and - above all - the gamut of output devices (especially monitors and printers), it is advantageous to introduce the CIE color space. ICC stands for International Color Consortium, an industry group pioneered by Apple Computer and now universally accepted as standard. CIE stands for Commission Internationale de l'Eclairage, the French name for the International Committee on Illumination that established the quantitative basis for color perception in 1931. The CIE space is one in which accurate colors are measured unambiguously, and through which most color conversions take place from one space to another.

The usual presentation of the space is shown in **Figure 1**, but this is in fact just the color plane through two of the three axes (called x and y, but roughly corresponding to the red-cyan direction and the magenta-green directions). The third axis is the luminance, and it is perpendicular to the color plane, so this space, like Lab and HSI spaces, separates the color information from the brightness. (There are actually several CIE color spaces, the original 1931 model called CIEXYZ, variants such as CIELUV and CIELAB, and modifications called "perceptually uniform" CIE in which the ability to distinguish colors visually is represented by the same distance everywhere in the space. These distinctions are important, but not for the discussion that follows here.)

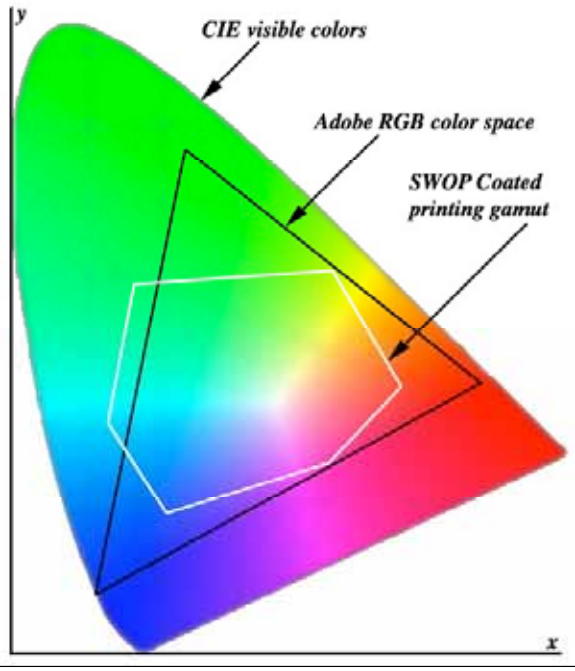


Figure 1. The color plane in CIE space. Adobe RGB space and the SWOP printing gamut (shown for coated paper) do not cover the entire range of visible colors.

It is easy to see the CIE "horseshoe" shaped space as a distortion of the simple color wheel introduced earlier. The reason for the distortion is that in this space, two colors can be linearly combined to make a third. In other words, if you know the colors of two phosphors - say the red and green ones - on your color monitor, then the line between those colors on the CIE color chart identifies all of the colors that can be generated by mixing those two phosphor outputs. With a third point corresponding to the blue phosphor, this produces a triangle inside which are all of the colors that the monitor can display.

The triangle for a typical monitor does not fill the entire CIE horseshoe, which represents all of the colors visible to a normal human. So there are colors that we can see that the monitor cannot generate. The triangle representing the RGB space covered by a display device like a computer monitor is called the "gamut" of the device, meaning that it shows the range of possible colors.

Printers have gamuts, too. Subtractive color generally has a much harder time producing saturated colors than emissive devices like a monitor. This can be seen by the smaller gamut produced by the standard CMYK inks used in offset printing. The color gamut depends on the paper as well, but even with coated paper the ability to produce saturated reds and blues is limited. Interestingly, the printer gamut is slightly greater than the monitor for cyan to green colors, but it turns out that in the green portion of the curve people have less color discrimination anyway, and this doesn't matter much.

The printer space is also more limited in the direction of the third axis. Most printers produce white simply by not depositing ink on paper. There are two problems with this. First, the paper can't be as bright in reflected light as a monitor can by emitting light. Second, there is a minimum amount of ink that can be deposited to produce the lightest possible grey value, and for many printers that step from no ink at all (white) to the lightest possible grey is a very visible one. The all-white areas of the print stand out visually and are distracting, so the solution is to not let the image go all the way to white, but to set the brightest possible output value to that lightest shade of grey, to prevent "blowing out" the whites.

Similarly, at the dark end, no combination of inks on paper can be as dark as a black monitor, because the inked paper still reflects some light. And when the ink nearly covers the paper entirely to produce its darkest value, there is an abrupt and visually distracting jump from the darkest grey value to solid "black." In addition, dot gain means that when the ink lands on the paper it usually spreads out to a larger area than the drop size, producing an increasing amount of darkness for a given amount of ink at the dark end of the space. Printers say that dark areas get "plugged" by too much ink. As at the bright end, the solution is to set the darkest possible output value to the darkest shade of distinguishable grey that does not cause plugging.

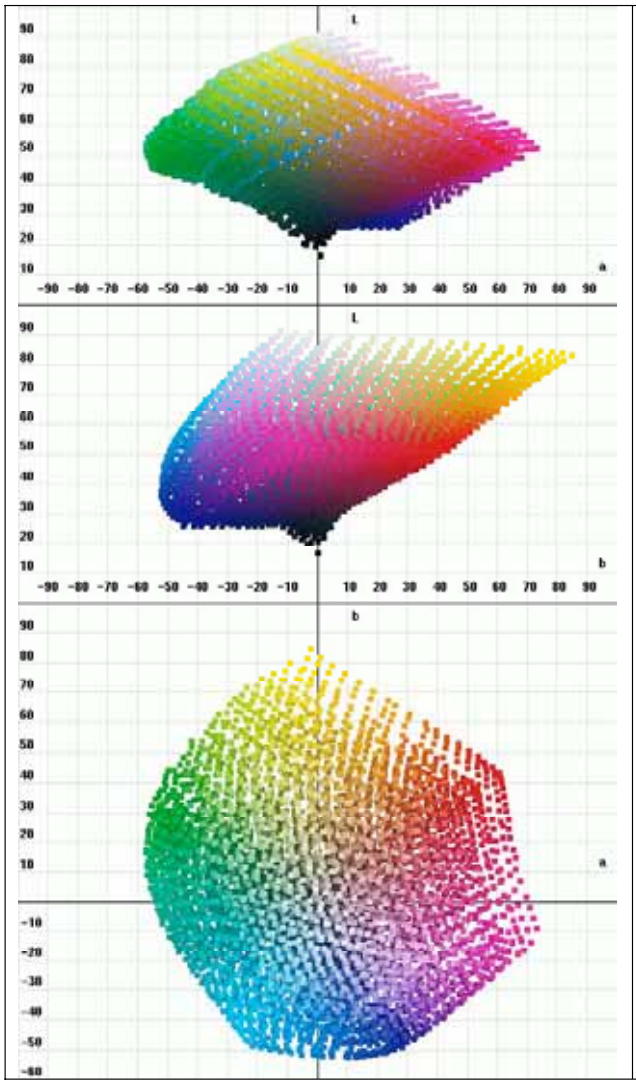


Figure 2. Three-dimensional gamut for a six-color CcMmYK ink jet printer, displayed as cross sections through L*a*b color space.

The entire gamut of a printer (or any other device) can be mapped out by using a series of color patches that vary in hue, saturation and brightness as shown in **Figure 2**. The result of the reduced range of brightness values is a further reduction of the gamut of the printer's color space. So it is typically necessary to decide how to handle images that may fit inside the gamut of colors that are stored internally and may look fine on the computer monitor, but exceed the gamut of the printer. That is where the ICC profiles come in.

ICC Profiles

The idea of having one standard color space for the image, and a series of ICC profiles that reproduce those colors on various output devices (especially monitors and printers) so that the appearance of the colors is the same is a very attractive one. Many hardware manufacturers have adopted this approach, and many scanners, printers and monitors have default ICC curves that do a pretty good job of handling this color conversion without user attention. Of course, it depends on keeping the device in calibration. You can't change printer inks or let the cartridge become dirty, or alter the contrast and brightness settings of the display, etc., without changing the ICC curves.

To produce a really accurate ICC profile for a particular device takes some hardware and effort. There are colorimeters that can attach to a monitor face and read the actual color produced, and software that produces a series of colors to be tested, that generate ICC profiles. Likewise, for printers, you can print a standard test document and then use a device to read the actual color of the various swatches under controlled lighting conditions. If you are in the clothing business and the match between screen images (e.g., shown on your catalog web site), print images (e.g., in your mail catalog), and the actual articles has to be very good indeed to satisfy customers, than you will want to invest in those devices.

But for most photographers it is acceptable to use the default profiles bundled with the various devices, or to carry out simple calibration routines using your visual comparison of images to make the necessary judgment. Apple has a display calibrator as part of the operating system (click on calibrate in the monitor color control panel), and Adobe supplies a similar Adobe Gamma calibrator for Windows computers. The steps are to set the gamma value and white point (often described as the color temperature), and to specify the colors of the R, G and B phosphors. Note that the profile for a CRT monitor may change over time as the phosphors age, so that the original information provided by the manufacturer may require tweaking.

The idea of an ICC profile to convert color values from the working RGB space to the output device space works well for CRT monitors, but not for LCD screens. That's a problem, because LCDs have a lot of other advantages (space, power consumption, long term consistency). A CRT produces color by combining the outputs from the three phosphors, and the linear combinations represented in CIE color space correctly calculate the resulting colors, including the all-important white point or neutral value. In the LCD screen, there are two independent sources of color, which interact nonlinearly. The backlight produces the "white" light (which may not be all that uniform across the screen), and the liquid crystals introduce color by absorption. Furthermore, the amount of that absorption varies for each color with viewing angle, so moving your head around in front of the screen will cause changes in the colors that are viewed. If you really need accurate color previews on your monitor, you will have to stick with CRTs for the present.

The other problem with displays, either CRT or LCD, is that the stored profiles that accompany the image may be ignored. For instance, most web browsers (the Macintosh version of Internet Explorer is an exception) ignore the profiles and simply use the raw RGB values. That means that whatever care you take in adjusting your image for proper color viewing may not reach the final viewer.

Profiles for acquisition devices such as scanners and digital cameras are not always provided, nor necessarily accurate. Instead, you can have Photoshop assign a default color space to the data (typically sRGB or AppleRGB. These have smaller gamuts than the Adobe RGB that is recommended as the best working space, so all of the color data from the acquired image can be accommodated. The color values in an acquired image depend on many things, some of which can't be controlled. For example, the color

temperature of the light and the white point set in the camera affect color. So does the background color of paper in a scanner. Once the image is in the computer, adjusting the colors as described in earlier chapters allows the user to correct for any problems.

The point of having the image in a working color space, with ICC profiles to convert the data for output to monitors and printers, is to allow making those corrections once, to deal with the acquisition variables. Then the corrected image can be sent to the various output devices without further modification.

There is a lot more to creating profiles and selecting working color spaces. The professional who needs accurate color and is willing to invest the time in mastering the subject, including creating calibration profiles for specific combinations of printers and inks, should read two books that cover the subject in detail. Both are from Peachpit Press, Berkeley, CA:

Michael Kieran (2002) "Photoshop Color Correction"
David Blatner and Bruce Fraser (2002) "Real World Adobe Photoshop"

Other working spaces

The DCF ("Design rule for Camera Files") standard used by most manufacturers of digital still cameras specifies sRGB as the standard color working space. This space was originally developed (by Microsoft and Hewlett Packard) for the display of images on CRT's, particularly when used for image delivery over the internet. The sRGB color space has a rather small gamut, much less than the ability of the camera to capture color information. This becomes particularly important when trying to print images as hard copy using the most recent generation of ink jet printers, which use multiple inks to increase their gamut.

For that reason, Epson has proposed the PIM ("Print Image Matching") technique, which most camera manufacturers have adopted. It takes advantage of the fact that most cameras store images in a JPEG format in which the color information is not encoded as RGB, but as YC_RC_B. This is a television-based standard in which Y is the luminance (brightness). The chrominance (color) signals are the red-green and blue-yellow signals, very similar to the a and b channels in the Lab space described in **Chapter 2**. This encoding allows for a greater gamut than sRGB, but normally when the JPEG image is read by the computer the values are either clipped or scaled to fit into the smaller space. If the sRGB color space has been selected as a working color space, this will result in a loss of saturation, which is why the larger AdobeRGB space was suggested above.

The PIM method adds information at the beginning of the image file that records the camera imaging conditions (e.g., settings for white point, gamma values, overall image brightness, whether landscape or macro settings were used, etc.). This information can be used to scale the color values back up into the color space of the printer gamut, rather than limiting them to the smaller sRGB gamut.

At present, while many cameras include the PIM header information in their JPEG files, only some Epson printers utilize it. But if your intent is to print directly from the image stored by the camera to a printer without any intermediate processing of the image in Photoshop, the technique does offer improvements in print quality. This is true whether the camera is connected directly to the printer, or the file is first transferred to the computer and later sent to the printer, so long as the image is not processed and re-saved in the computer. It is particularly useful if the built-in camera software includes the capability to specify print quantity and orientation instructions (the so-called DPOF or Digital Print Order Format, developed by Kodak, Fuji and other manufacturers).

Non-PIM-aware printers that can accept images directly from the camera (including executing DPOF instructions), and computer software such as Photoshop that read the camera JPEG files, ignore the PIM information in the file header. The PIM method is not an ICC profile, and does not substitute for that more general approach, which aims to make the printer output match the screen display of the image. For one thing, the PIM information does not include the characteristics of the printer, inks or paper. Instead it must be used by software in the printer to generate the equivalent of an ICC profile for the image. But it is potentially useful for the direct transfer of images to PIM-enabled printers, in cases that do not require color correction or other processing in the computer.

Out of Gamut

Of course, there is a problem if the image in its working space has a range of colors that exceed the gamut of the output device. Photoshop has a valuable tool to reveal that problem. Select View -> Gamut Warning to show the pixels in an image whose color values lie outside the gamut. You can select a preview color to mark the out-of-gamut pixels; often black or white is a good choice when looking at color images, in order to distinguish them from the rest of the picture.

If only a few pixels lie beyond the gamut of the output device, the best choice is usually to clip them to the maximum saturation that the device can generate but to leave everything else in the image unchanged. **Figure 3a** shows a situation in which this applies. On the other hand, if a significant fraction of the image, or large local areas of the picture, are out of gamut (**Figure 3b**), then it is usually better to squeeze the color range uniformly to fit the extremes into the gamut. This reduces the saturation of all of the colors present and may introduce slight color casts, but maintains the relationship between colors better.

Figure 3. Example of out-of-gamut pixels (shown in white): a) a few scattered pixels; b) large areas.

In the Photoshop Color Settings dialog, you can select the conversion option to use. Choosing “Relative Colorimetric” intent maps the white point of the source and destination color spaces so that the absolute white point corresponds to the white of the paper, and then accurately reproduces all of the colors that lie within the printer gamut. Points that lie outside the gamut are clipped to the nearest color that can be printed. If, instead, you select “Perceptual” intent for the color rendering, the color space is compressed

to fit within the gamut of the output device. This shifts the colors proportionately and thus maintains the relationships between colors, but generally results in a reduced saturation and brightness.

In that same dialog, checking “Use Black Point Compensation” will minimize plugging of shadows by shifting the darkest black value to the darkest value the printer can produce. It results in a scaling of the brightness values, just as the compression of the color range brings all of the colors within the printable gamut, and should always be used.

One thing that the different gamut of displays and printers should emphasize is that converting from one space to another always involves some compromises and loss of information. If clipping is accepted, then color discrimination is lost. If compression is accepted, the dynamic range is reduced and the ability to discriminate nearby colors is lost. The consequence is that conversion between color spaces should be done only once, when the final image is sent to the output device. Switching back and forth between color spaces to perform color balancing and processing operations should be avoided, or at least minimized, because with each conversion the color information in the image is impacted.

In **Chapter 2** on color adjustments, most of the work was performed in RGB space, but mention was made of working in other spaces, such as Lab, HSI or CMYK. Several well-regarded authors recommend that absolutely all such efforts should take place in CMYK space, so that question deserves some review.

CMYK and RGB

One of the continuing debates that rages in books and on the internet news groups that discuss color correction, Photoshop, and related issues is whether color correction, as discussed in **Chapter 2**, should be done in RGB color space or directly in the CMYK pre-press space. For the pre-press professional who is intimately familiar with the SWOP (Specifications for Web Offset Publications) inks, the amounts of those inks that correspond to familiar colors (including neutral greys), and the behavior of those inks as they are deposited on paper, operating directly on the CMYK values makes sense and has the benefit of tapping into a deep knowledge base. But for most photographers, especially if they will be printing out hardcopy on home printers (typically ink jet, possible dye sublimation) rather than the web offset presses used for books, catalogs and magazines, RGB is almost always a better choice. Here's why.

First, the dependence on direct CMYK editing arose when the color on the computer screen bore no particular relationship to that of the final printed copy, and it was necessary to know a lot about CMYK colors and to manipulate the values directly in order to get reasonable color matching. But now the widespread use of ICC profiles means that all devices in the system, such as the monitor and the printer, have their own correction models so that they can properly represent the color in the stored digital image. In other words, color has become device-independent. Once an image has been color corrected on the computer screen, the printout will (assuming the devices are calibrated properly) show the same colors.

The consequence of this change is that the RGB file has become the master repository for the color information. If it is printed out today on one type of printer, the driver software with its ICC profiles will perform the conversion necessary for that device. If it is printed out tomorrow on a very different printer, the conversion process will be carried out again from the original RGB data. This avoids multiple conversions from one color space to another, and back again, which causes a lot of problems, including loss of color gamut.

Even more important, the original reason for working with CMYK values was to deal with the limitations of offset printing using SWOP inks. None of the popular ink jet, dye sublimation or other printers used for personal digital image printing use those inks. True, the basic ink colors are still cyan, magenta and yellow, but they are different versions of those colors than the SWOP inks. Furthermore, most of the printers now use more than just four inks. It is typical to find a seven-color printer designated CcMmYKk that has two different cyans, two different magentas, and two different blacks (one a grey). The printer driver knows how to convert the RGB values to amounts of those inks. If presented with a CMYK image it would first have to convert it back to RGB, losing gamut in the process.

All imaging devices, including cameras and scanners, deliver to the computer a set of RGB values (even though some of them use other color filters in order to actually acquire the original data). The RGB values are the native format for image storage and display. The phosphor dots on your monitor are RGB and so are the little colored squares on an LCD display.

There are several compelling reasons to edit colors in RGB. The most important is that neutral colors are easily defined as having equal amounts of red, green and blue. Adjustments to the RGB channels that keep the proportions of RGB will not introduce color casts or distort the grey balance of the image. If that sounds trivial, it is not. With CMYK values, the amounts of cyan, magenta and yellow that produce neutral grey vary with brightness, and lightening or darkening a portion of an image generally will introduce color casts.

CMYK advocates like to emphasize the ability to adjust the amount of black (K) in the shadow areas of the image to control directly the darkness and visibility of detail in those areas. This is partly true, but the control comes at a price. It is necessary when working with CMYK values to keep the total amount of ink under a limit set by the press and paper type, typically about 300%. Remember that the halftone process uses arrays of dots in each of the four colors to produce the printed copy of the image. If every color was set to maximum (4 X 100% = 400%) the total amount of ink on the page would be too great, creating problems in the press, problems in drying the ink, etc. So depending on the type of paper, the total amount of ink permitted may vary from less than 250% (newsprint) to about 320% (coated magazine stock). Fiddling the amount of black in shadow areas can't be done arbitrarily, and also may shift the color balance for the other inks.

A closely related topic that comes up in these discussions is the meaning of "black." In a halftone printing operation using CMYK inks, you can't generate a visual black with just the black dots. 100%K would leave too much of the white paper uncovered, producing a light grey result. Some other inks have to be added, typically up to that 300% total mentioned above. Since that is a necessity, CMYK advocates try to claim it as an advantage. The so-called "rich black" produced by adding substantial amounts of magenta and cyan, and sometimes a little bit of yellow, is different in appearance than a simple black color produced by just black ink. The difference has to do with the reflection of light from the ink, something that isn't considered in the usual subtractive color theory that ink just absorbs light. Maybe a "rich black" is visually more interesting than just plain black, but it doesn't matter because the conversion of the RGB values (in which R=0, G=0, B=0 is truly black) will still get converted to whatever combination of inks the printer driver with its ICC profiles calculates and the result will still be a "rich black" regardless of what color space the image started out in.

So there really are very few reasons to work in CMYK space. Even if your image is going to be printed with SWOP inks on a web offset press, working on it in RGB space and leaving the final conversion

process to the printer drivers and ICC profiles is still the best policy. Any color that you can create using CMYK combinations can be achieved using RGB as well. To prove that, set up your color monitor to show both RGB and CMYK and observe the equivalent values for various points in an image. It is occasionally convenient to exercise subtle control over shadows by tickling the black value in a CMYK representation, but since the image must be converted back to RGB before it can go to most printers, the loss in gamut and other possible distortions more than overcomes this small advantage.

And if you intend to print the image hardcopy on one of the modern high-quality ink jet printers, there is no reason to even consider CMYK editing because those colors are quite unrelated to the inks that will eventually be used.

RGB is the most natural space for most editing, but there are some reasons to use HSI or Lab space in particular situations. Both separate the brightness information (intensity or luminance) from the color values, so processing of the image (such as the edge-enhancement operations shown in **Chapter 4**) or adjustments to contrast can be made without altering the colors. Also, because the meaning of hue and saturation is easy to understand, it can be very useful to monitor these values while performing editing operations, even though the editing may be performed in the RGB channels.

How printers work

The discussion on printers and inks above has neglected the important topic of just how modern printers work. All printers create color images by depositing a minimum of three inks or dyes onto paper. With the exception of dye-sublimation (also called dye-diffusion) printers, which use specially coated paper into which the dyes diffuse and mix, the colors are usually deposited on the surface of the paper and do not mix. In fact, most printers try to place the dots so that they are next to each other rather than one on top of the other, although this is imperfect and it consequently matters which color goes down first and which one last.

It should be clear from the preceding discussion that unlike emissive displays like RGB monitors which build up color by adding together the light from the different phosphors, printers work by subtracting light that would otherwise reflect from the white paper. The basic subtractive color space, introduced back in **Chapter 2**, is CMY: C=Cyan, the opposite color to red on the color wheel, M=Magenta, opposite to green, and Y=Yellow, opposite to blue. These colors add together to produce the other hues, for instance yellow and magenta make red.

Without the complexity of the black ink, it would seem that converting from RGB to CMY would be simple, based on this complementarity. Alas, it isn't so. Since C+M+Y cannot produce a very dark black (more of a muddy brown), most printers include as a fourth color a black ink or dye (hence CMYK). The rules for how much black to put down, and how to alter the amount of the colored inks as this is done, vary considerably from one printer driver to another.

The code words in this field are GCR (grey component replacement) and UCR (undercolor removal). These both describe sets of rules that control the need to add in black ink to build up shadows in a smooth fashion, and the removal of some of the colored inks as black is added, to prevent building up too great a total amount of ink on the page. If you ever need to know more about this, to control the conversion to a finer degree than the standard ICC profiles do it for you, read Dan Margulis "Professional Photoshop" book (2001, John Wiley, New York). Then go talk to your printer and let him handle it for you - this is a highly specialized field.

The method by which printers deposit small flecks and dots on ink, dye or toner and trick people into seeing the result as an image is called halftoning. It is easiest to see in a newspaper image like that shown in **Figure 4**. The "dots" that make up the image are halftone cells. Large ones cover up a lot of white paper and so look dark, and vice versa. So varying the amount of coverage in each cell is the way to generate an intensity scale. The printers build up each cell by depositing much smaller dots onto the paper. As shown in **Figure 5**, if the dots are one-sixth of the size of the cell, as in the illustration, then it is possible to put generate 37 different shades of grey in the cell when it is viewed from far enough away that only the integrated brightness and not the internal detail can be seen.

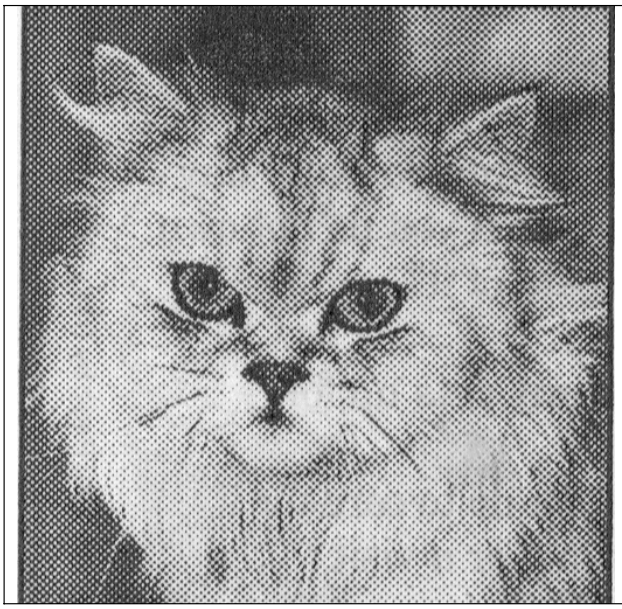


Figure 4. Typical newspaper halftone image, showing the use of a coarse grid of cells whose black area is varied to produce the visual effect of grey scale.

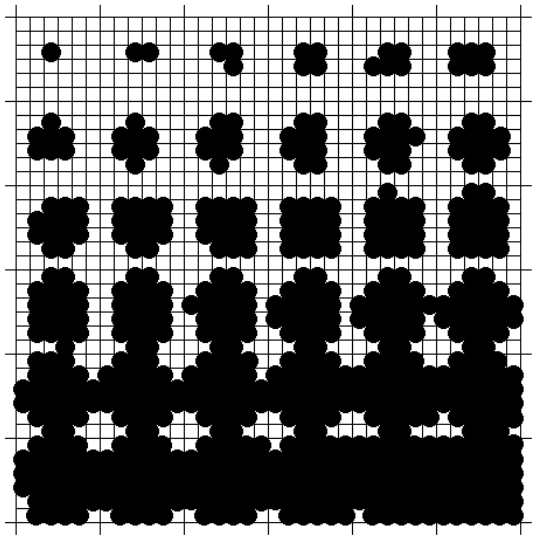


Figure 5. Constructing a grey scale in halftone cells by placing individual printer dots in each cell. The order in which the dots are used to fill the cell varies widely, and of course they are not perfectly regular in shape or spacing.

That is where a lot of the confusion in printing arises. Printers are often specified in terms of the number of dots per inch (dpi) that they can place on paper. These are the dots that make up the coverage of the cell. The resolution of the image as it is viewed is controlled by the number of cells, which is usually

called the lines per inch (lpi) meaning the number of lines of halftone cells on the paper. Ideally, we might want each cell to correspond to one pixel in the image. If the goal is to print the image so that individual pixels are as small as the eye can see at close viewing distance, about 0.1 mm, then we would want to have about 250 cells per inch. If we printed the pixels larger than that the image would look increasingly blurred. Photographer call this "empty magnification." Enlarging a 35 mm. negative to poster size results in an image that is fine when viewed from far away but is dominated by the film grain and looks blurry when you get close.

So to print a sharp image for close viewing, we want about 250 cells per inch. This is generally called the lines per inch (lpi) setting. The name arises from ancient history, even before modern halftone printing, when printing plates were etched manually with fine lines to create textures interpreted as grey scale. Within each cell we also want a tonal resolution, or ability to distinguish between different brightness levels, that captures the subtlety of our picture. We've been using 256 brightness levels on the CRT, which is somewhat more than the eye can distinguish. Considering the nonlinearity of human vision, which detects equal ratios of brightness rather than equal increments, it is sufficient to have about 100 brightness levels to prevent any ability to see contouring or banding in the image. From the example shown above, this would require that the halftone cell have a 10x10 array of dots within it, to produce 100 levels of brightness.

Additional problems arise because the dots aren't really just little squares and some interesting artefacts arise when they touch. In order to completely cover the paper with black dots, they must overlap somewhat. This gives rise to "dot gain" that makes the printed grey values nonlinear. Depending on the dot pattern in the cells, the sudden transition as the number of dots per cell is increase from a continuous white background with separated black dots to a continuous black tone with separated white dots can produce a visual discontinuity in the image. Typically, these nonlinearities are built into the ICC profiles for the printer.

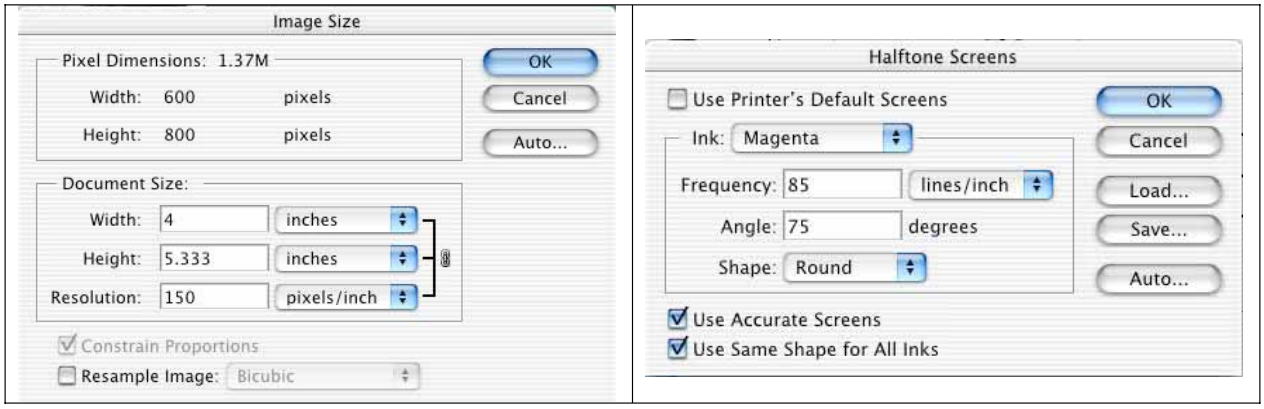
So far we've only considered grey scale images and not color. But even so, the assumptions above lead to the conclusion that the printer should have a dpi (dots per inch) specification on the order of 2500 in order to generate an lpi (lines per inch) resolution of 250, with 100 brightness steps. Not coincidentally, that is a typical resolution for imagesetters that use film to produce CMYK separations for web printing. Such resolution can produce visually sharp grey scale images. Newspaper images aren't that good. They generally have only about 85 lines per inch, and it shows in the coarse appearance of the example above.

Printing "resolution"

For any dpi specification on the printer, there is a tradeoff that can be made between spatial resolution and tonal resolution. Using a smaller halftone cell improves the former and degrades the latter, as shown in **Figure 6**. One of the major sources of confusion that seems to plague many users of digital image processing programs like Photoshop with home printers is the relationship between the "resolution" of the image and that of the printer. The variables that matter are:

1. The number of pixels in the image, Height X Width
2. The image "resolution," which may be as low as 72 pixels per inch for screen display, or as high as 300 pixels per inch for fine printing.
3. The intended size of the image hardcopy, usually given as the Width and Height in inches or cm (or, for printers, sometimes in points where 1 inch = 72 points).

Figure 6. Printing the same image with different “lines per inch” setting on a laser printer: a) original grey scale image; b) 100 lines per inch; c) 75 lines per inch; d) 50 lines per inch.



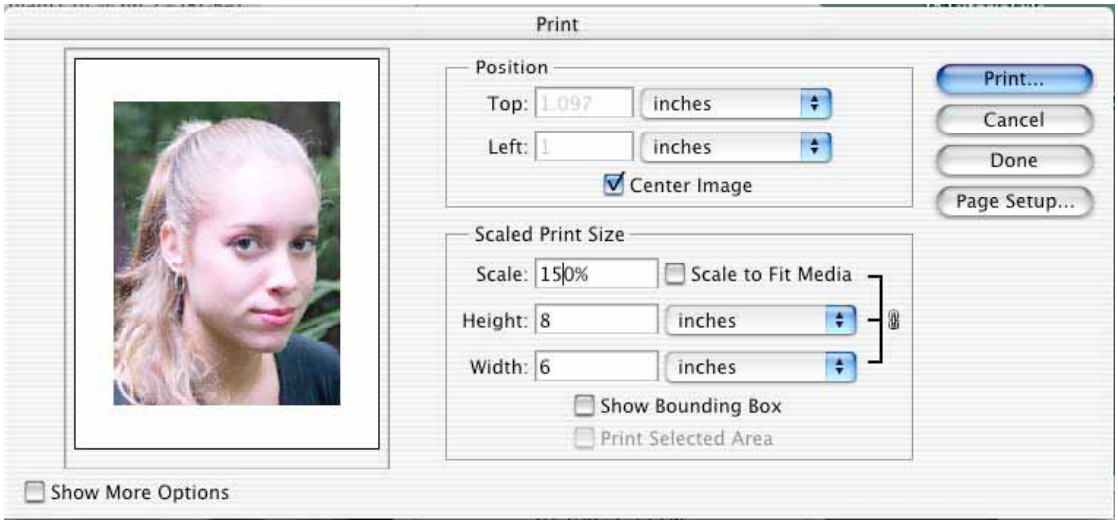


Figure 7. Dialogs controlling the relationship between the stored pixels and the printed image size: a) Image Size dialog; b) Halftone screen dialog (selected from c); c) Print with Preview dialog.

The relationship between these is most clearly seen in the Image Size dialog, shown in **Figure 7**. If you uncheck the Resample Image box so that the program does not alter the number of pixels in the image by reinterpolating an image of different size, then changing the document size (width or height) changes the resolution (pixels per inch) with which it will be printed. In the Print with Preview dialog there is another chance to change the printed size of the image, also shown in the figure. This dialog also allows you to enter the screen angle and spacing, now called lines per inch. This information is sent to the printer in Postscript, but is generally ignored except by high end imagesetters used for creating offset masters. Home laser printers ignore Postscript commands that they cannot implement, such as the screen orientation, and most other printers aren't controlled with Postscript anyway, and many don't print with traditional screens.

But the basic principle relating the intended image size and the resolution still applies. You can print a 160x200 pixel image to cover a full 8 x 10 inch area on a page, but the result won't be very sharp because the pixels in the original image will be one twentieth of an inch across, much greater than the ability of the human eye to resolve detail. The image will have empty magnification (i.e., visually without detail or information) but might be acceptable as a poster viewed from far away.

Conversely, you can tell the printer to print a 2560x1920 pixel image from a 5 megapixel camera as a 3 inch wide picture, but that corresponds to more than 850 pixels per inch. The printer can't put down that many small dots to show the image with grey scale or color information, even by reducing the number of brightness values in each cell, so the software skips most of the pixels and just samples the image. And even if all of the pixels could be printed, you wouldn't be able to see them without a microscope.

Generally, a good strategy for printing images is to set the resolution to a value of about 150-300 pixels per inch for hardcopy printing, depending on the resolution of your printer. In Photoshop, the image will still be displayed on the screen with all of the pixels shown at 100% viewing scale. Depending on your monitor and display hardware, this will typically be in the range of 75-100 pixels per inch. If you print your images with this setting, the image quality will be good, but the size of the print may not be as large as you would like. If you enlarge the image using the Print with Preview dialog, you can fill the

page but at some cost in the sharpness of the result. This may still be acceptable depending on your purpose.

Contours

Sacrificing tonal resolution for spatial resolution, by using too small a halftone cell, produces a visual artefact called contouring. This results when each step in the brightness scale is large enough to be visible. In a region of an image with gradually varying brightness, as shown in **Figure 8**, this produces the effect of a visible contour line that degrades the image appearance. In general, this is more likely to happen in the dark or shadow regions of an image, where small variations in brightness are more easily distinguished. **Figure 9** shows that the same problems can arise in color images. The steps in the R, G, B channels produce slight shifts in perceived hue as well as brightness.

Figure 8. Contouring produced by expanding the contrast in the dark range of values.

Figure 9. Contouring in a color image (on cheek and neck).

Contouring is also evident on the computer monitor if the interface supports only a limited bit depth (thousands of colors rather than millions). Most modern systems provide 8 bits each for red, green and blue which is enough to provide a visually continuous color and brightness range. However, the image itself may produce visual contours if it has been captured with too small a range of brightness values, or if processing operations have stretched the brightness scale. Typically this happens when an image is initially captured with an incorrect exposure, so that the histogram does not cover the full brightness range. Stretching the histogram with the Adjustments -> Levels or Adjustments -> Curves functions, as discussed in **Chapter 2**, shows the contrast and detail the very light or very dark regions but spreads out the brightness values so that the differences are apparent (**Figure 10**).

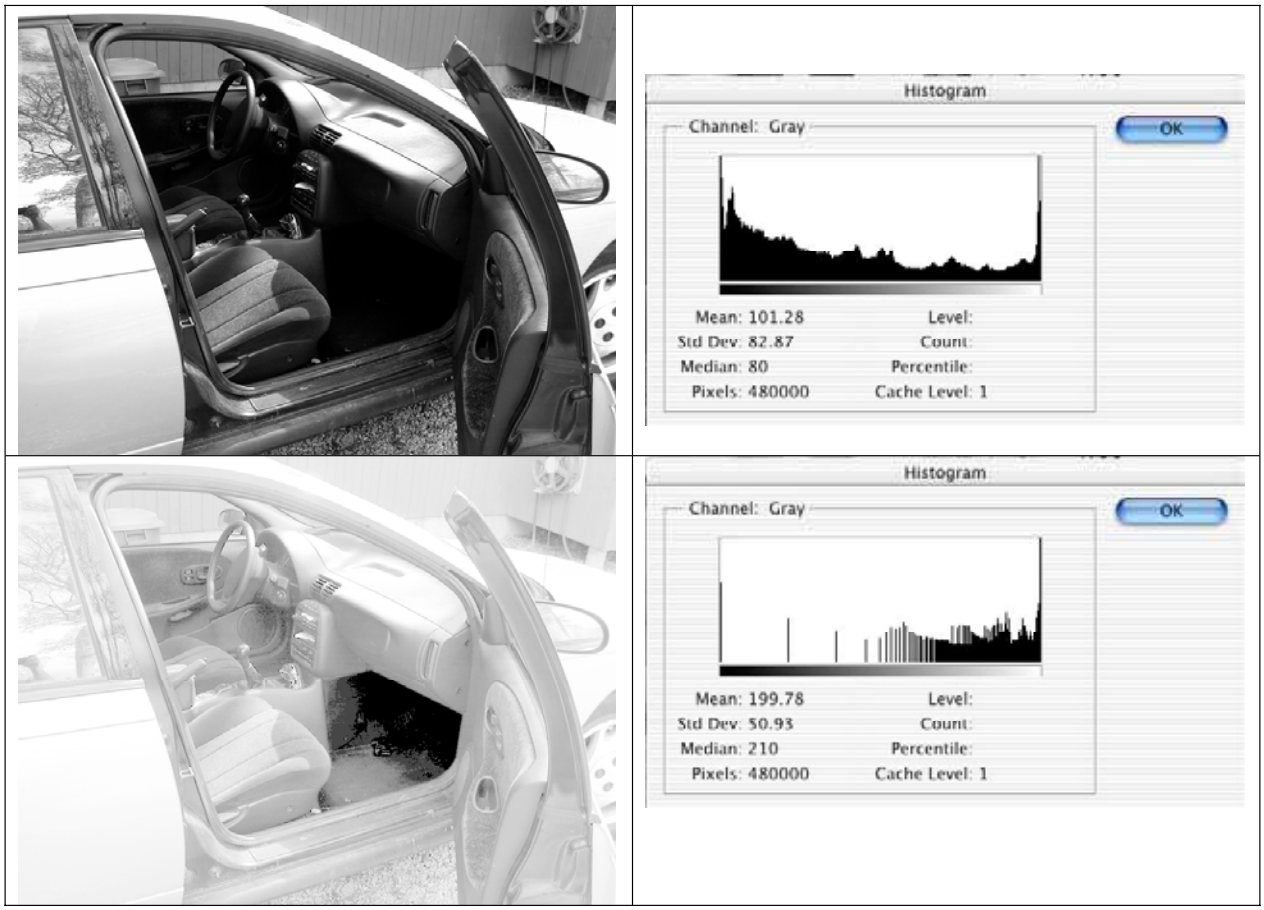


Figure 10. Photograph recorded to capture forensic evidence at a crime scene: a) original, with deep shadow hiding information on the car floor; b) histogram of a; c) adjustment of gamma to reveal gun in shadow area; d) histogram of c.

If the images are captured with more than 8 bits, this problem can be alleviated. Some digital cameras save the "raw" image with 10 or 12 bits per channel, while most film scanners provide 12 bits per channel or more to capture the full dynamic range of the film. If these images are opened in Photoshop they are treated as 16 bit per channel images. As shown in **Figure 11**, stretching the contrast in 16 bit space does not open up gaps between adjacent brightness values, and so does not produce contouring.

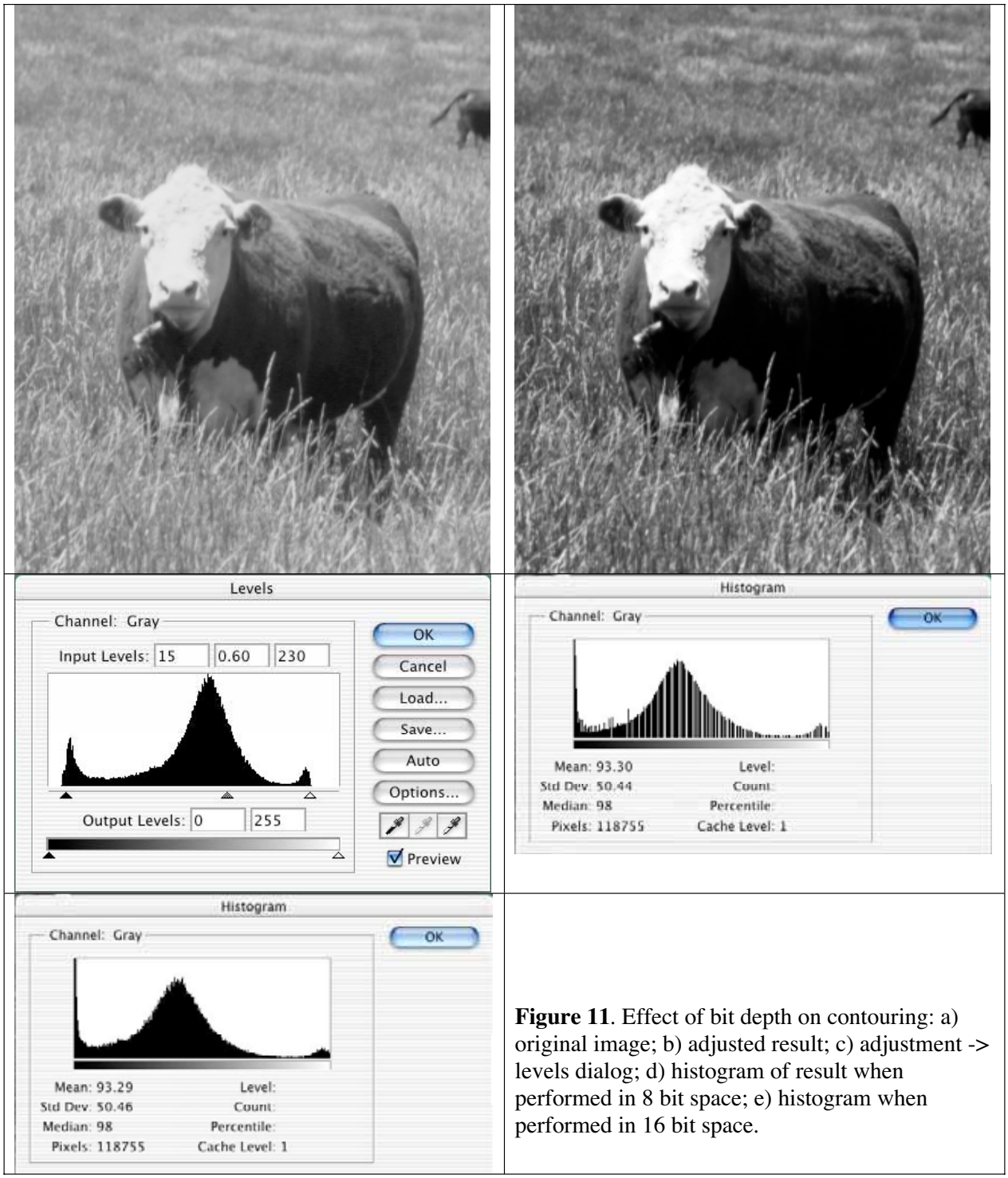


Figure 11. Effect of bit depth on contouring: a) original image; b) adjusted result; c) adjustment -> levels dialog; d) histogram of result when performed in 8 bit space; e) histogram when performed in 16 bit space.

Capturing and processing images in 16 bits is always advantageous because in the contrast enhancement or local processing operations described in preceding chapters, the loss of information if two values are compressed to one, or the gaps in the histogram that give rise to contouring, are eliminated. On the other hand, it is often more cumbersome to work in 16 bits. Many of the tricks shown in preceding chapters used layers and selections. Layers are not supported for 16 bit per channel images, and the wand

selection tool does not function. There are work-arounds for these limitations. It is possible, for instance, to duplicate a 16 bit image, change one copy to 8 bits, make the selections on that image, save the selection, and open the selection up as an alpha channel on the 16 bit image, and use it there. But, admittedly, this is more time consuming and requires more understanding on the part of the user.

Also, 16 bit per channel images require twice the storage space as 8 bit images, and take longer to process. And finally, many of the built-in routines and third-party plug-ins for Photoshop do not function on 16 bit per channel images. But the trend is clearly toward working with images of greater bit depth, the memory, storage and speed of computers continue to increase, and software is being adapted to function with these images.

Whenever it is possible to work through all of the image processing and adjustment steps using 16 bit images, it is best to do so. If images are being captured with an 8 bit camera, minimal adjustments to color are being applied, and then they are being printed out, then staying in an 8 bit mode is fine. But if more precision is available in the original image, or if aggressive adjustment or processing operations is applied such as discussed in earlier chapters, then staying in 16 bit mode until the final step is wise. Reducing the final image to 8 bits per channel in order to print it is fine, because the printers do not have even 8 bits of tonal resolution, nor does the eye.

Photoshop has one more trick that it uses to alleviate contour problems as well. Dithering is the process of adding small amounts of random noise to an image to hide the contours. When an image is reduced from 16 bit per channel mode to 8 bit per channel mode, mathematically it would be simple to just divide by 128 (not 256 - in Photoshop 16 bit mode the values actually range from 0 to 32768, or 15 bits). But Photoshop then randomly perturbs the lowest order bit so that smooth gradual changes in value do not have abrupt contour lines. One step in 256 is smaller than the eye can see (or the printers can record). But of course, this should be done only as the final step before printing since even tiny differences can be amplified by processing until the random “noise” becomes visible.

Compressing dynamic range

Photographic prints have much less dynamic range than film, and likewise prints from digital cameras cannot reproduce the full range of tones captured by the camera. This is particularly true for 16 bit per channel images. Film photographers have dealt with this problem by dodging and burning areas of the image in the printing process so that the darker regions are lightened and lighter regions darkened, while retaining their local contrast. The justification for the procedure is that human vision does not measure absolute brightness values but compares local differences in the image. That means the size of the dodging and burning tools needs to be just large enough to cover the important local detail, and small enough to reduce the differences between large regions in the image.

It is possible to do the same thing automatically with digital images. The difference of Gaussians operator discussed in Chapter 4 compares the brightness of a pixels in a small region to those in a surrounding neighborhood, mimicing the behavior of the human visual system. Combining the result from this operator with the original image produces a reduction in overall contrast while increasing local contrast, as illustrated in **Figure 12**. This procedure is carried out using just the luminance or grey-scale brightness of the pixels, without affecting the color (by converting the image to an H-S-I or L-a-b space), and reduces the overall contrast range so that the picture can be printed without losing the details in bright and dark regions.

Figure 12. Range compression using a difference-of-Gaussians: a) original image produced by averaging together the exposures in Chapter 3, Figure 1; b) averaging the difference of Gaussians result with the original.

The problem with using just a single difference of Gaussians operator is that the scale of detail in different parts of the image may not be the same. It is relatively straightforward to extent the same logic to use several different sizes for the center and surround comparison, as shown in **Figure 13**. The fine scale detail in the trees, medium scale detail in the rocks, and large scale detail in the sky, can all be seen, whereas in the original image the rocks were dark, and the trees very dark, so that the detail was hidden, even on the computer screen. In a printout, with the lower range of available contrast, the problem would be much worse.

Figure 13. Range compression using multiple difference-of-Gaussians (performed using the dynamic range compression plugin in Optipix from Reindeer Graphics): a) applied to Figure 12a; b) applied to Figure 31a from Chapter 4, to reveal detail in the extreme shadow regions.

When slide film is scanned to produce digital images, or when capturing high dynamic range images with a digital camera (either one that has more than 8 bits per channel, or by combining multiple images as discussed in Chapter 3), the ability to perform dynamic range compression becomes essential to producing viewable and printable images in which the detail in light and dark regions is retained.

Color

Grey scale printing can be difficult, but life gets a lot more complicated when color is involved. For CMYK printing, we need to put down arrays of halftone cells for all four ink colors. Preferably, they should not overlap. That makes each cell smaller, which is why the amount of each color (and hence the maximum saturation) of printed images is reduced for printed images. The non-overlapping requirement is handled on a web offset press by rotating the halftone screens at different angles for each color, as shown in **Figure 14**. **Figure 15** shows an enlarged example of a printed color image in a book.

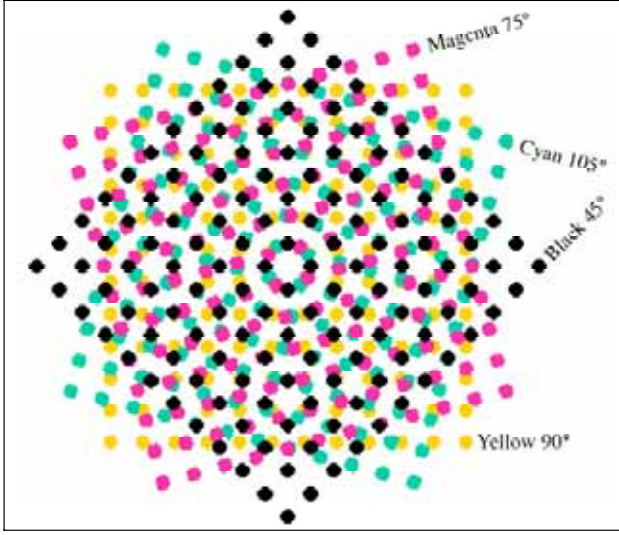


Figure 14. Diagram of CMYK printing showing the rotation of the halftone screens.

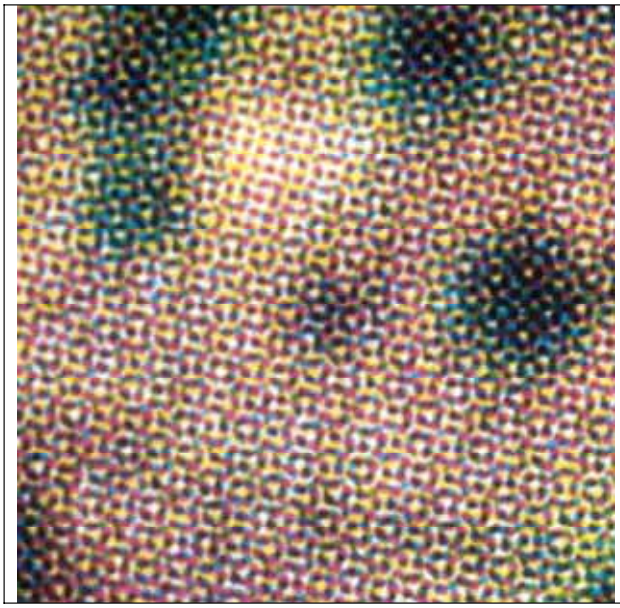


Figure 15. Enlarged view of a color image printed by a web offset press.

Home printers accomplish this in different ways. Color laser printers deposit the four CMYK colored toners. The stated resolution of 600 to 1200 dpi is the size of the toner dots, and typically for printing images the halftone cell size may be 8x8 dots (65 theoretical brightness levels, somewhat less in reality because of nonlinear visual response and dot overlaps). That would produce a resolution of 1200/8 =

150 cells per inch, which is fair but not as good as a photograph. More important is the limited saturation and brightness range. **Figure 16** shows an enlarged example of output from a color laser printer.

Figure 16. Enlarged view of a color image printed by a laser printer.

The individual toner dots are printed with the same grid alignment, and partially overlap each other. In order to generate maximum color, they must entirely cover the white paper, which produces a darker image. The maximum color saturation and brightness are much poorer than a halftone press (and also less than other types of home printers, discussed below). The advantage of the laser printer is the relatively lower cost of supplies and higher speed of printing, which makes them suitable for office use and perhaps for layout proofs but not for high quality image.

Figure 15 compares the output quality of several types of personal printers that are available for personal use. Dye sublimation printers are generally considered to produce the smoothest visual appearance. This is due to the fact that the dyes diffuse into the gel coating on the special paper used, and the three or four colors (CMY or CMYK) mix together instead of being separate dots. The color saturation and brightness rival photographic prints, and even when the spatial resolution is poorer the smoothness of the image presents no visual artefacts that alert the viewer to the lack of resolution.

The major drawbacks to the dye sublimation or dye diffusion printer, which has limited their use, is the high cost of consumables, low speed, and relatively poor stability for the prints. Dye sub prints are sensitive to heat, light and other environmental effects which can cause fading or discoloring. The cost is due to the need for special coated paper to receive the dyes, and the fact that most printer designs use an entire sheet of dye-coated material for each color and for each page. This is placed in contact with the paper and a moving “print” head heats the dye sheet at each point to control the amount of that color dye transferred to the gel coating. Because of this control, the tonal variations in dye sub prints can be very subtle. But generally the maximum color saturation is less than achievable with ink jet printers, because the dyes are inside the gel coating while ink jet printing places the inks on top of the coated paper surface.

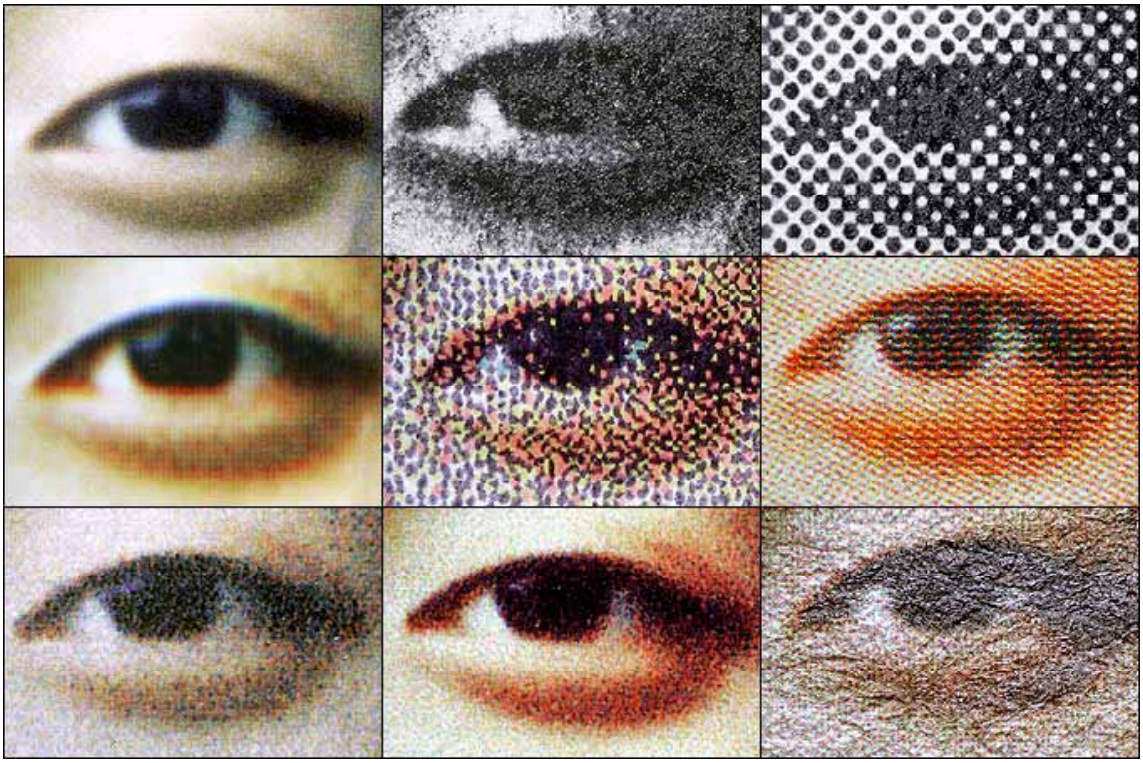


Figure 15. Magnified (30x) views of the same image printed using different devices: a) photograph; b) Xerox copy; c) black and white laser printer; d) dye-sublimation printer; e) phase change solid ink printer; f) color laser printer; g) ink jet printer in low resolution mode; h) ink jet printer in high resolution mode; i) ink jet printer using uncoated bond paper (images from M. Jurgens, Queens Univ.)

Ink jet printers

The dominant type of printer for color images has become the ink jet. The printers themselves are quite inexpensive; the consumables (ink and paper) are moderately costly but the total cost is still much less (and the convenience much greater) than traditional photographic materials. The current generation of ink jet printers place tiny (a few picoliter) droplets of inks onto paper with a resolution of more than 2000 dots per inch. The relationship between dpi, lpi and tonal resolution described above indicates that these printers should produce results that match very well with the maximum resolution (both spatial and tonal) of human vision. Indeed, with some of the newest multiple color inks, specialty papers, and advanced printing software, the results can surpass that of traditional photographic prints.

The first issue to decide is what kind of inks to use. Dye-based inks give the widest color gamut and greatest saturation. They also are capable of better resolution (smaller minimum droplet size) and show less metamerism (change in visually perceived color with changes in lighting) than the pigment based inks. However, they do not provide archival permanence (for that matter, neither do the dyes in traditional photographic paper, as anyone opening a box of color prints stored under less than perfect conditions of temperature and humidity will realize). Pigment based inks applied to the appropriate type of paper can produce results with extremely good permanence, important in some applications.

On the other hand, if the images are stored digitally in a form that might be readable on computers at some future time (i.e. assuming that *.tif files on a CD-R will be more meaningful in 2020 than a stack of IBM punch cards or reel of Dectape is now) then perhaps archival quality in the print is not so important. In that case, the dye inks have the clear advantage.

Some printers now use 6 or 7 different inks, not just a simple CMYK but perhaps CcMmYKk with extra cyan, magenta and grey (light black) colors that improve the ability to cover the entire color space smoothly. Of course, it requires specialized software to control these printers and to transform the color information from the values stored in the computer to those required by the printhead. All of the driver software used in these printers starts with an RGB image, so the theoretical advantages of using other color spaces for editing become moot. Transforming the image back to RGB so the printer can convert it again to its own space would compress the color gamut or clip values if editing was done in CMYK or Lab space.

Ink jet printers do not correspond exactly to the halftone cell illustrations used above for a web press. The colored dots are deposited on the paper in a more-or-less regular pattern that covers the paper with varying amounts of ink to produce a range of brightness and saturation as shown in **Figure 16**. As the total amount of ink (color) increases, the brightness decreases.

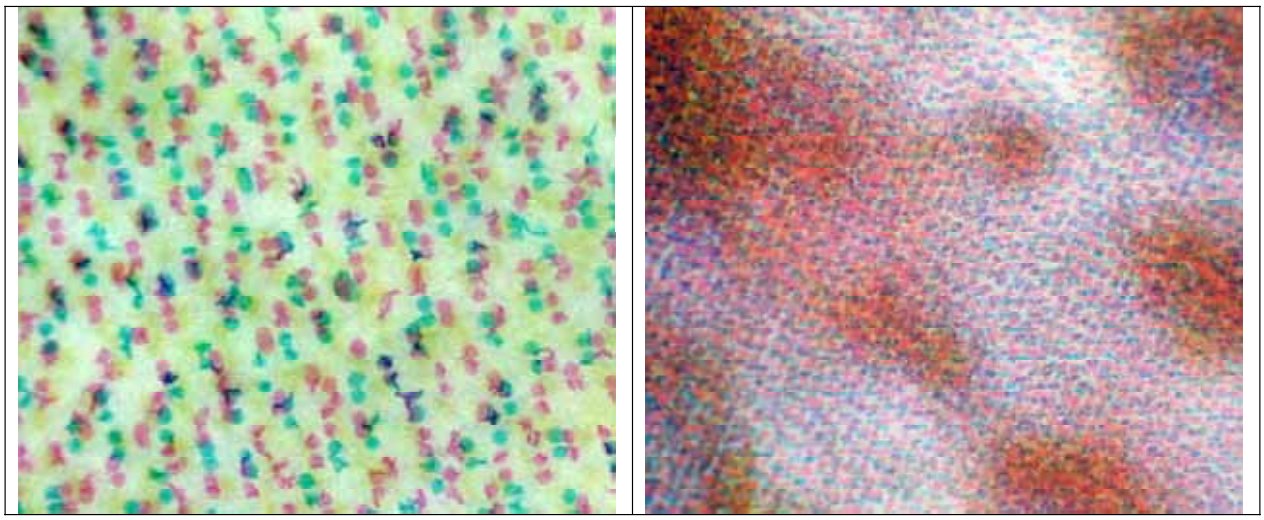


Figure 16. Ink jet printer output (six color CcMmYK inks): a) light area with well defined separate dots; b) superposition of more colors produces greater saturation but covers more white paper for a darker result.

One very interesting development for ink jet printers is the use of multiple grey inks to produce “black and white” monochrome prints. Ink jet printers are generally known for their relatively wide color gamut, and were initially used for computer graphics of the “comic-book” variety in terms of colors. But by formulating special inks and writing a custom printer driver to dispense them, Jon Cone (www.Piezography.com) has been able to generate black and white prints that are superior in dynamic range to traditional photographic darkroom results, and far beyond the possibilities of traditional printing methods.

The piezography method produces an extended high-precision grey scale range along with high spatial resolution. The differences are particularly striking at the ends of the brightness scale, where very light

grey values are produced that are discernibly different from white, and dark grey values that are not quite black. In traditional printing, and even in photographic printing, these values are often lost. Admittedly this is a niche market, but the combination of superior quality, cost and convenience is compelling if your interest is in black and white prints. **Figure 17** shows an example of the results achievable with this approach.

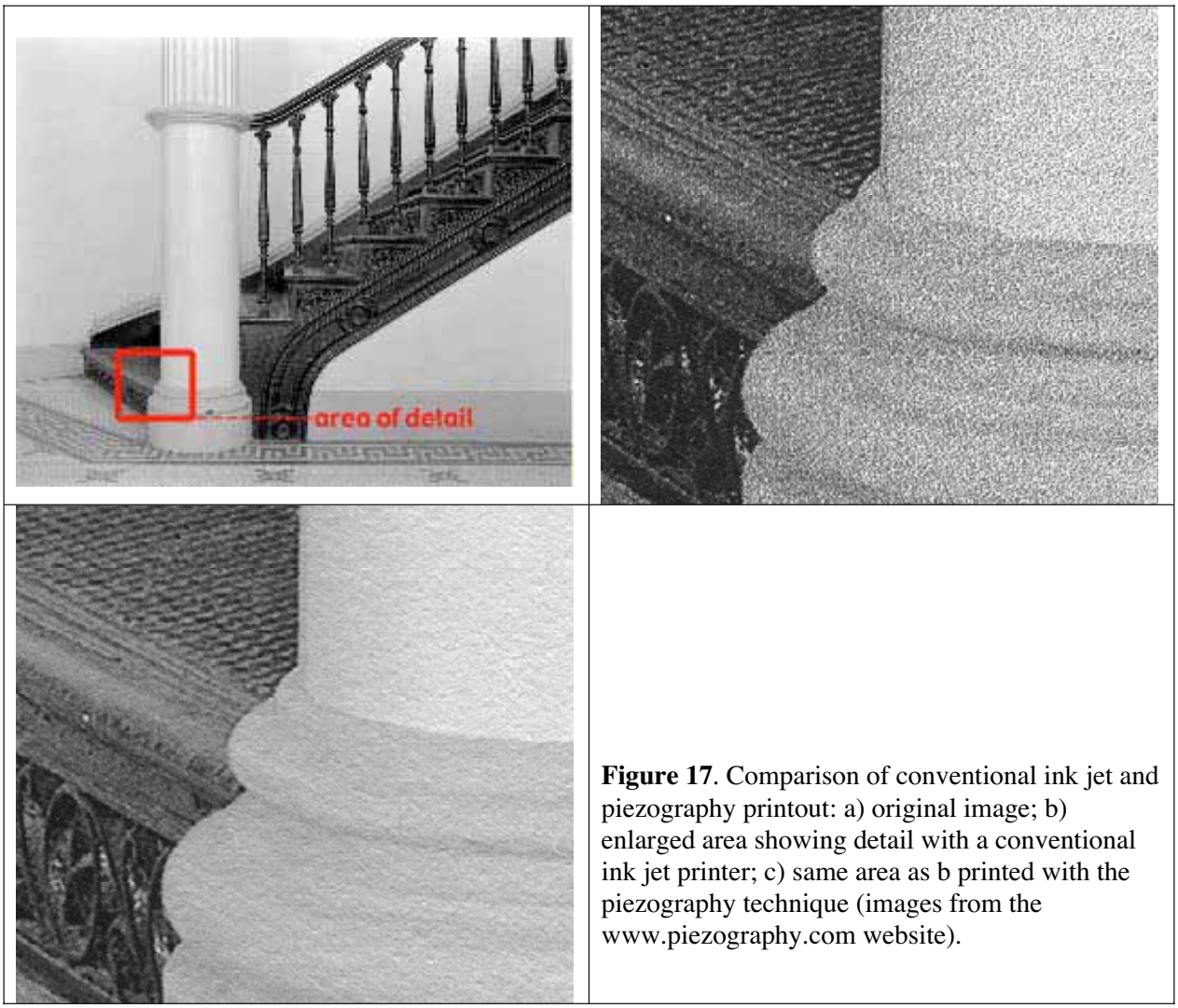


Figure 17. Comparison of conventional ink jet and piezography printout: a) original image; b) enlarged area showing detail with a conventional ink jet printer; c) same area as b printed with the piezography technique (images from the www.piezography.com website).

For either color or black and white printing, papers are an important and all too often overlooked parameter affecting results. You can't just use some handy duplicator or laser printer paper to get decent results from an ink jet printer. Most of the printer and some photographic companies market several grades of paper intended for ink jet printers. Simple coated stock is fine for printing text, because the coating prevents the ink from diffusing into the paper fibers and spreading out to produce a ragged result (**Figure 18**). But for printing images the higher grade "photo" quality paper should be used. The coating prevents the ink from spreading and maintains sharply defined, small color dots that give the best resolution.

Figure 18. Wicking of ink into the fibers on an uncoated paper surface produces a ragged effect for type, as well as images.

Coated papers are generally available with either matte or glossy finish. It is the matte finish that produces the best results, truest colors, and best overall results. This is particularly true with pigmented inks, which exhibit significant color shifts with lighting (metamerism) on glossy paper, and sometimes give problems of adherence as well. Textured papers also create difficulties for ink adherence and color shadows, although some of the more recent papers show marked improvements. Special acid-free papers, often using cotton instead of wood fibers, are available for archival storage. These should be used with pigmented inks rather than dyes, to produce results that are chemically stable. Prints should still be protected from strong sunlight (especially UV), unstable environments, and extremes of temperature or humidity.

Note that for every combination of inks and papers, there must be an appropriate ICC profile. Most printer drivers have at least a few generic ones built in, for instance for glossy or matte paper. Generally these are finely tuned to work well with the printer manufacturer's own inks and papers. Substituting a third party ink supplier, particularly one who recycles used cartridges, can result in significant shifts in color and quality, and should be done with caution. Saving a few cents per page on a picture that you've invested hours of time preparing is not usually a wise choice.

To get the truest color results, or to handle inks or papers that differ from the generic categories provided for in the printer driver software, it may be desirable to perform your own calibration. This has become much easier with the introduction of software and hardware from companies like GretagMacbeth (www.gretagmacbeth.com), Monaco (www.monacosys.com), Praxis (www.praxisoft.com), and others, which can handle every part of the calibration and profiling process, including cameras, scanners and monitors. For the printer, the procedure is typically to print an image of a test pattern supplied as part of the kit. The printed color swatches are either measured photometrically or compared visually to a standard, and the results used to generate the profile. Most of the suppliers include detailed step-by-step guides that make the process quite straightforward.

Preparing images for printing

Once the monitor and printer have been calibrated and the correct ICC profiles are in place, it may be expected that the appearance of color on the monitor will be matched by that on the hardcopy. Depending on the intent selection for rendering an image from one color space to another, various tradeoffs can be made to handle color values that do not fit within the gamut of the printer. As noted above, the most usual setting is Relative Colorimetric, which first scales the white of the source space to the white of the target space, adjusting all other colors relative to that white. Then it matches the adjusted colors in the source space that are inside the gamut of the target space exactly, and clips out-of-gamut colors to the nearest reproducible hue, sacrificing lightness and saturation.

For images that would have large areas of pixels clipped by the relative colorimetric method, the best choice is perceptual rendering. This compresses the gamut of the source space into the gamut of the destination space in such a way that the overall relationships between the colors is preserved, even though all the colors may change in the process. This preserves the overall appearance of the image when examined by itself but may introduce slight color casts, and the output result will not match side-by-side with the original. The other two choices, Saturation rendering (which sacrifices color accuracy in order to produce maximally saturated printouts, suitable for business graphics) and absolute colorimetric (which is like relative colorimetric except that it does not correct for the different white points of the monitor and paper) are not generally useful for photographic applications.

Adjusting the colors in the image using the monitor presentation can either be done for visual impression or to maintain color accuracy. The latter requirement, as discussed in previous chapters, typically requires photographing a standard color chart under the same lighting conditions as the images to be processed. While important for applications such as clothing catalogs, it is not the goal of most photographers. The usual purpose of color adjustments is to produce an image with visual interest and impact, and with no colors that would be seen as questionable by the viewer. That leaves a large range for enhancement and personal intent.

But once the image is “looking good” on the screen, it isn’t quite time to print yet. The next issue is whether the image size in pixels, the printer resolution (spatial and tonal), and the desired print size are compatible. If the required size of the printout is larger than can be generated with the full resolution of the printer, the printer driver software will try to scale it up to the larger size. Typically this is done just by repeating each pixel value as many times as needed, and it produces a blocky effect in the hardcopy. Much better enlargement of images can be performed using either the Photoshop Image Size routine with bicubic interpolation to fill in the gaps, or by one of several third-party plug-ins that apply other enlargement methods.

Of course, it is impossible to increase the amount of information in the image. Enlarging it requires filling in the spaces between the original pixels with new values. As shown in **Chapter 4**, the bicubic interpolation method usually works pretty well. But other techniques such as spline fitting, sinc interpolation or fractal interpolation are also used. Spline fitting is best for preserving the sharp boundaries of features, rather than smoothing the boundary by spreading it over a greater distance. Sinc interpolation has nice mathematical properties that should also preserver sharpness, but in practice often results in ringing around edges that is visually distracting. Fractal interpolation manufactures new detail that is not actually present. This makes the enlarged image appear to have greater resolution than is actually there, but in many cases that detail, being false, disrupts the interpretation of the image. The introduction of detail borrowed from other locations in the image also produces some color shifts.

In some special situations it may be necessary to deal with color trapping. The appearance of color fringes along the boundaries between different color regions occurs because there is a slight offset in most printing methods between the C, M and Y inks. So, for example, at a boundary between red (magenta and yellow) and green (cyan and yellow), where the magenta ends and before the green starts there would be a yellow line, or if the magenta and cyan overlapped there would be a black line. This is not usually a concern in printing photographic images, because edges between strongly saturated uniform color regions are rare. The problem typically arises with generated computer graphics, and is handled by special software.

Finally, before printing, it may be worthwhile to apply some additional image sharpening. **Chapter 4** has discussed at length techniques such as the unsharp mask (in its various forms) as tools for increasing the visual sharpness and crispness of images by expanding the contrast at edges and in detail. Presumably that has already been done to your image by now. But the printing process itself is a low-pass filter. All of the printing methods described above, even the highest resolution film separations for a web offset press, have less resolution than the very sharpest resolution of the human eye. That means that high-frequency information that is needed visually to produce really crisp detail and edges will be blurred slightly.

The usual solution is to apply some extra sharpening beforehand to offset the blurring effects of printing. In most cases, this is done simply using an unsharp mask. The same considerations and modifications shown in **Chapter 4** still apply.

After all of these issues have been addressed, it's time to make your printouts. If you use a modern ink jet printer (or, better yet, a dye sub printer) with high quality paper, the results will rival if not surpass the tonal range of a photographic print, with adequate spatial resolution and reasonable longevity. And, of course, you will still have the image in digital form in case you want to make other printouts at another time, perhaps after further adjustments and editing.

aberration (see lens aberrations)
adaptive equalization 131-132
adjustment curves 42, 56-58, 107
adjustment layers 43, 51, 169-170
AdobeRGB working space 190-191
aligning images 90-91, 97-101, 167
alpha channel (also see selections, quickmask) 163, 167
Apply Image function 165-166
archival prints 206-211
averaging exposures 81-83, 85-86, 89-90, 138
background 16, 101-102, 114-116
Bayer pattern 11
bit depth (8 vs. 16) 7, 166-167, 201-203
blemish removal (see retouching)
blending multiple images 75-79
blown-out whites 188
blue (see RGB color representation)
blur (see Gaussian smooth)
boundaries - blurring 18, 109-110
boundaries - color 108
boundaries - enhancement (see edges)
boundaries - selecting 23, 166
brightness (see HSI color representation)
camera sensors 8, 9
CCD (charge coupled device) 9, 10
Channel Mixer adjustment 51
CIE color representation 48, 187-188
clipping 7, 51-52
clone tool 175-181
CMOS (complementary metal oxide on silicon) 9, 10, 13
CMYK color representation 44, 78, 152-153, 188, 194-195, 204-206
color adjustments 60-61, 66-71
color artefacts 11
color channels 22, 92, 127, 181
color matching (see color adjustments)
color median (see median filter)
color resolution 110
color sampler (see control points)
color selection (see selection-color)
color shifts 25-27
color temperature 193
compression 31-33, 108, 191
conditional smoothing (see also thresholding) 25
contouring 200-201
contrast adjustment 8, 43, 58
control points 60-69
cooled chip cameras 15

CRT displays 190-191, 212
Custom filter 16-17, 36, 90, 123-124, 134, 138-139
cut and paste 168
cyan (see CMYK color representation)
dark image (see background image)
DCF (Design Rules for Camera Files) 191
dead pixels 21, 51, 144
demosaicking 11
density 7
difference of Gaussians 124, 139
digital camera review 6
distortion (see lens distortion)
dithering 203
dodging and burning 183-184
dot gain 189, 197
dpi (dots per inch) 196-199
drawing layer 169
dust and scratches 144-145
dye sublimation printers 195, 206-207
dynamic range 6, 11, 28
edges 126-130, 135-137
embossing 133-135
empty magnification 9, 12
enhancement of detail (see also edges) 131
enlargement (see also interpolation) 153-156
equalization (see histogram equalization)
exposure 28, 14, 15, 75-78
extended focal depth 87-90
f-stop 76, 77-78
feathering 163
fiducial marks 91
fill factor 10
film resolution 6
focal depth (see extended focal depth)
foreground isolation 101-103
Fourier space 149-151
Foveon chip 12, 13
full well capacity 10
gamma adjustment 54-55
gamut 28, 187-189, 192-194
Gaussian smooth 16-18, 24, 118-119, 138-140, 143, 144, 184-186
GCR (grey component replacement) 195
gradient blending 96-97, 173-174
green (see RGB color representation)
GretagMacbeth chart 73
halftoning 148, 151-153, 196
healing tool 175-181
highlights 83-85
histogram 14-16, 41-42, 51-58, 67, 111, 201-202

histogram equalization 57, 104
HSI color representation 26, 43, 46-47, 64, 135, 193, 195
Hue & Saturation adjustment 51
hue (see HSI color representation)
human vision 15, 38, 65, 105-107, 131, 133, 187
hybrid median (see median filter)
ICC profiles 187, 190-192, 210-211
inhibition 40, 106
ink jet printers 191, 193, 207-212
ink types 207-208
intensity (see also HSI color representation)
International Color Consortium (see ICC profiles)
interpolation 153-156, 199, 210
jpeg artefacts 32-37
jpeg compression (see compression)
Lab color representation 22, 26, 27, 32, 43, 49, 122, 123, 125, 135, 147, 193, 195
Laplacian (see sharpening filter)
laser printers 198, 205-206
layer blending modes 78-81, 101, 117, 165-166, 170-173
layer masks 173-174
LCD displays 190
lens aberrations 28-30
lens distortion 29, 97-99
Levels function 25, 43, 53, 58
locked pixels 21, 51, 144
logarithmic intensity scale 7, 14
lpi (lines per inch) 196-199
luminance (see also Lab color representation) 49-50
magenta (see CMYK color representation)
masking 101
median filter - color 11, 22-23
median filter - hybrid 11
median filter - neighborhood shape 19, 20
median filter 16, 19, 132-133, 140-141, 143-146
merging selections 78, 8788
metamerism 209
midtones 57, 63
moiré 147
montage assembly 953
multiple images 75-103, 167
neutral grey 27, 60-62, 67, 70, 114
noise - additive 14 - 16
noise - blue channel 13, 14, 142, 144-145
noise - fixed pattern 10
noise - line 20
noise - multiplicative 14-16
noise - periodic 148-153
noise - random 10, 14, 121-124, 137-143

noise - shot 21-23, 143-144
noise - visible 14, 15
nonuniform lighting 113-116
opacity 167-168, 171
paper types 208-211
patch tool 175-181
perceptial intent 192, 210
periodic noise (see noise)
perspective distortion 99-100
piezography 208-209
PIM (Print Image Matching) 191-192
pixel 6, 7
plugged blacks 189
quartertones 57, 63
quickmask (see also selections) 163
radius of filter 119-121, 127, 138
random noise (see noise)
red (see RGB color representation)
red eye 181-183
region selection (see selections)
relative colorimetric intent 192, 211
retouching 105, 174-181
RGB color representation 22, 26, 27, 43-44, 48, 107, 121, 123, 125, 135, 142-143, 181, 190, 193-195
rich black 194
saturation (see also HSI color representation) 28, 76, 192, 205-206, 211
scanners 6, 7, 9
selections 115-116, 156-167, 203
sharpening filter 90, 118-130, 184-186, 212
shot noise (see noise)
single chip camera 10, 11, 13
spatial resolution 6-8, 12, 112
sRGB working space 190-191
stereo viewing 91-93
storage 31
SWOP (Specifications for Web Offset Publications) 188, 193
text layer 169
three chip camera 10, 12
threshold 25, 3, 120, 164
tolerance 157, 159
transfer function (see adjustment curves)
trapping 211
tristimulus correction 72-73
true color 28, 72-73
UCR (undercolor removal) 195
unsharp mask 7, 107, 118-130, 132, 165-166, 212
variance equalization 131-132
variance filter 137
vergence 39, 92

vignetting 29, 117
web color 187
white point 27, 191
YCC color representation 27, 191
yellow (see CMYK color representation)