Digital Imaging

by Hany Farid

In 1969 the Apollo 11 astronauts landed on the moon, the Woodstock Festival attracted nearly 500,000 concert-goers, and George Smith and Willard Boyle invented the first Charge Coupled Device (CCD). Although greeted with less fanfare than the other events of 1969, the CCD gave rise to digital imaging. By replacing the traditional film with an electronic CCD sensor, digital imaging revolutionized photography ranging from the Hubble telescope in outer-space to the cell phone in your pocket. This entry describes the capture, storage and manipulation of digital images.

Formation

The CCD sensor is the "film" of a digital camera. It consists of a two-dimensional array of photoelectric elements that become electrically charged when exposed to light. Because the amount of charge is proportional to the light's intensity, the CCD's electrical pattern is a faithful representation of the light pattern striking the sensor. Although exquisitely sensitive to light intensity, a CCD element does not differentiate light wavelength (color). To record a color image, a Color Filter Array (CFA) is overlaid on the CCD, Figure 1. With a CFA, each CCD element records a limited range of wavelengths, corresponding to either red, green, or blue.

The pattern of electrical charge on the CCD is transferred to the camera memory, where it is represented as an array of Picture Elements or pixels. A six mega-pixel digital camera, for example, corresponds to a CCD sensor with 6 million elements and a digital image with 6 million pixels. Associated with each pixel is a number between 0 and 255, where 0 corresponds to the minimum charge and 255 to the maximum charge, yielding 256 distinct intensity values. In a full resolution color image, each pixel is assigned three numbers, one for the intensity of red, one for green, and one for blue. But as noted above, the color filter array limits each sensor element to one color and so initially each pixel represents either red, green or blue. For a full resolution RGB color image, the camera must interpolate across neighboring pixels to fill in the missing color values, Figure 1. These three values are sufficient to give rise to tens of millions of colors. For example, red is given by the RGB triple [255,0,0], green by [0,255,0], blue by [0,0,255], yellow by [255,255,0], a dark red by [128,0,0], black by [0,0,0], and white by [255,255,255].

Three values are sufficient to represent millions of colors because a digital camera encodes wavelength in a way that is similar to the human eye. The three filters in the CFA are analogous to the three cones in the eye which selectively absorb light in either the red, green, or blue range of the color spectrum. A digital camera and the human eye invite other comparisons. For example, in the center of the visual field (the fovea), the human eye has an estimated spatial resolution equivalent to tens of millions of pixels, and a dynamic range equivalent to several thousand distinct intensity values. A dynamic range equivalent to millions of intensity values is possible for the human eye due to the fact that the eye can dynamically adjust its lightness sensitivity range. Similar effects can be achieved digitally through High Dynamic Range (HDR) imaging in which multiple images photographed at different exposures are combined.

Format

A digital image in the camera and computer memory can be stored in many different formats. The most basic format, RAW, stores the pixel values directly recorded by the CCD prior to CFA interpolation. This format affords efficient storage since only one number is stored for each pixel, but it requires any subsequent photo editing software to perform the CFA interpolation. The remaining image formats fall into one of two categories: non-lossy and lossy. Non-lossy formats such as TIFF, PNG, and BMP, store the full RGB digital image without any loss of information. The benefit of these formats is that they permit high quality reproductions, but with the drawback that they require significant amounts of memory for storage. The lossy GIF image format achieves compression by limiting the total number of colors in the image from tens of millions to, typically, a few hundred. The image format JPEG, perhaps the most popular lossy format, achieves compression by removing some color and high spatial frequency information (i.e., image details). These lossy image formats reduce the disk space required to store an image, but at the cost of degrading the quality of the image.

Quality

Camera manufacturers typically tout the number of pixels as the single measure of image quality. Many factors, however, contribute to the final quality of a digital image, including the camera lens, the CCD sensor, the CFA filter, and the image compression level. Low quality lenses can lead to chromatic aberrations where the color channels are mis-aligned relative to one another. Lenses may also produce geometric distortions which causes straight lines to appear curved. Low quality CCD sensors can lead to high levels of noise or graininess in the image. Low quality CFA filters and interpolation can lead to poor color reproduction. And high levels of lossy compression can lead to significant loss of color and details. These degradations can range from the slightly to highly visible, and can be particularly problematic when enlarging a digital image for printing.

In high-end digital cameras, these artifacts have been largely eliminated with results that rival those of the best traditional film cameras. In low- to mid-range digital cameras, today's six to eight megapixel cameras (millions of pixels) are of sufficient resolution for most consumer needs (for example, at a high resolution of 300 dots per inch (dpi), a six megapixel digital image can be printed to a size of 8x10 in).

Digital Image Processing

Prior to the digital revolution, the enhancement and manipulation of images required talented artists and technicians to spend long hours in the dark room. With the advent of digital imaging, such alterations are now only a few computer mouse clicks away.

Some basic image manipulations involve systematically remapping pixel values using a lookup table (LUT). This remapping is represented graphically in Figure 2(a), which shows how a LUT can be used to alter the brightness of an image. The horizontal axis of the LUT corresponds to the original intensity value (in the range 0 to 255), and the vertical axis corresponds to the re-mapped intensity value (also in the range 0 to 255). By sliding the LUT down, for example, all pixel values between 0 and 80 are mapped to 0, while the remaining pixel values between 81 and 255 are linearly mapped into the range 1 to 175 (resulting in a darkening of the image). Functionally, consider a single pixel whose value is in the range 0 to 255. Find this value on the horizontal axis and draw a line straight up to intersect the LUT. Now draw a straight line leftward to the vertical axis to determine the new value for the pixel. This process is effectively repeated for each pixel to uniformly change the appearance of an image. Also shown in Figure 2(b) is the intensity histogram after the brightness adjustment. The horizontal axis of the histogram corresponds to the intensity value, in the range 0 to 255, and the vertical axis corresponds to the number of pixels that have the specified gray value. Darkening corresponds to a leftward shift of the histogram. Similar to darkening, brightening is achieved by shifting the LUT upward which corresponds to a rightward shift of the histogram.

As shown in Figure 2(c), image contrast can be increased by increasing the slope of the LUT. Similarly, the contrast can be decreased by decreasing the slope of the LUT. Note that both brightness and contrast enhancement can saturate pixels, where a range of values is collapsed to a single value. As a result, these operations are not reversible since once the values are mapped en masse to 0 or 255, their original values are lost. This saturation can be seen in the histogram where there is an increase in the number of pixels in the bin at 0 (black) and at 255 (white) corresponding to all of the pixels that were mapped to these values. In order to increase contrast while avoiding saturation, an image can be auto-scaled by adjusting the LUT such that the smallest pixel value in the image is mapped to 0 and the largest pixel value is mapped to 255.

Image contrast can also be increased or decreased through a non-linear LUT. Shown in Figure 2(d) is the result of one such contrast enhancement, gamma correction, where each pixel value, divided by 255, is raised to the power 2, and then re-multiplied by 255 (the scaling by 255 is so that the pixel values of the gamma corrected image remain in the range 0 to 255). Varying amounts of gamma correction are often automatically applied by digital cameras and in computer monitors used to view digital images. Unlike brightness and contrast enhancement, gamma correction does not saturate any pixels. Lastly, shown in 2(e) is the result of quantizing a digital image to only two gray values (also known as thresholding). The corresponding LUT is a step function that uniformly collapses a range of intensity values to a single value.

In addition to basic intensity manipulations, digital filtering can remove or enhance information in an image. Shown in Figure 2(f) is an image that has been blurred to remove image details. Shown in Figure 2(g) is an image that has been sharpened to enhance image details. And shown in Figure 2(h) is the result of filtering that isolates edges. Each of these filtering operations involves adjusting pixel values based on the values of neighboring pixels. An image is blurred by replacing each pixel value by the average of its neighboring pixel values -- the larger the neighborhood, the greater the blurring. An image is sharpened by replacing each pixel with the difference between itself and its neighboring pixels. And, an image's edges are enhanced by replacing each pixel with a combination of horizontal and vertical differences in pixel values.

Many photo-editing software packages perform these LUT and filter operations, along with many more sophisticated manipulations: red-eye can be removed from flash photographs, objects or people can be added or removed from an image, several images can be stitched together to create a wide-angle mosaic, and much much more.

Future

Advances in digital and computer technology have revolutionized photography in ways that could not have been imagined when the CCD was first invented. These advances have had an exciting impact on science, medicine, and art. These advances have also had the problematic consequence that they make it easy to doctor images in ways that are difficult to detect. Doctored images are appearing with a growing frequency and sophistication in tabloid magazines, fashion magazines, main-stream media outlets, scientific journals, political campaigns, courts of law, and our email in-boxes. As a result, our trust in photography has been diminished. This in turn has given rise to the field of digital forensics and the development of techniques to authenticate digital images. In the coming years, digital and computer technology will (of course) continue to develop while the field of digital forensics will try to help us regain our trust in photography. Although it is impossible to imagine precisely how these fields will develop, it is clear that digital imaging has forever changed our relationship with visual imagery.

Further Readings and References

1. H. Farid (2008). Digital Doctoring: can we trust photographs? In B. Harrington (Ed.) *Deception: Methods, Motives, Contexts and Consequences*, Stanford University Press.

2. R.C. Gonzalez and R.E. Woods (2008). *Digital Image Processing, 3rd edition*, Prentice Hall.

3. J.C. Russ (2006). *The Image Processing Handbook, 5th edition*, CRC Press.

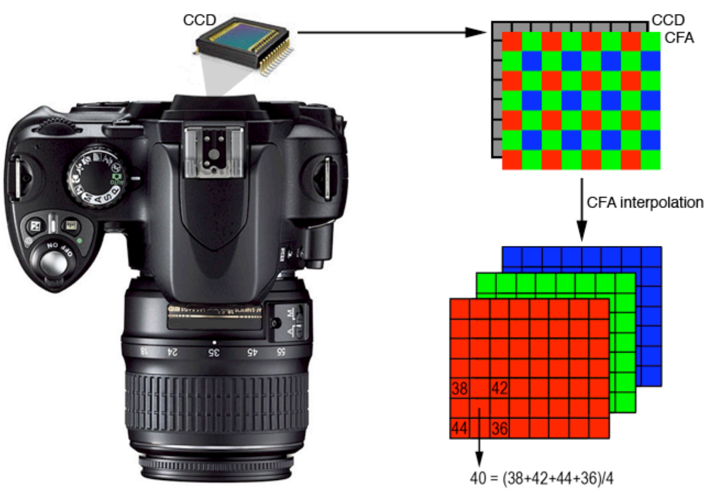

Figure 1: A color filter array (CFA) overlaid onto a CCD sensor. Each CCD element records a limited range of wavelengths, corresponding to either red (gray), green (light gray), or blue (dark gray). A full three channel RGB color image is created by interpolating the missing color pixels by, for example, averaging the recorded values.

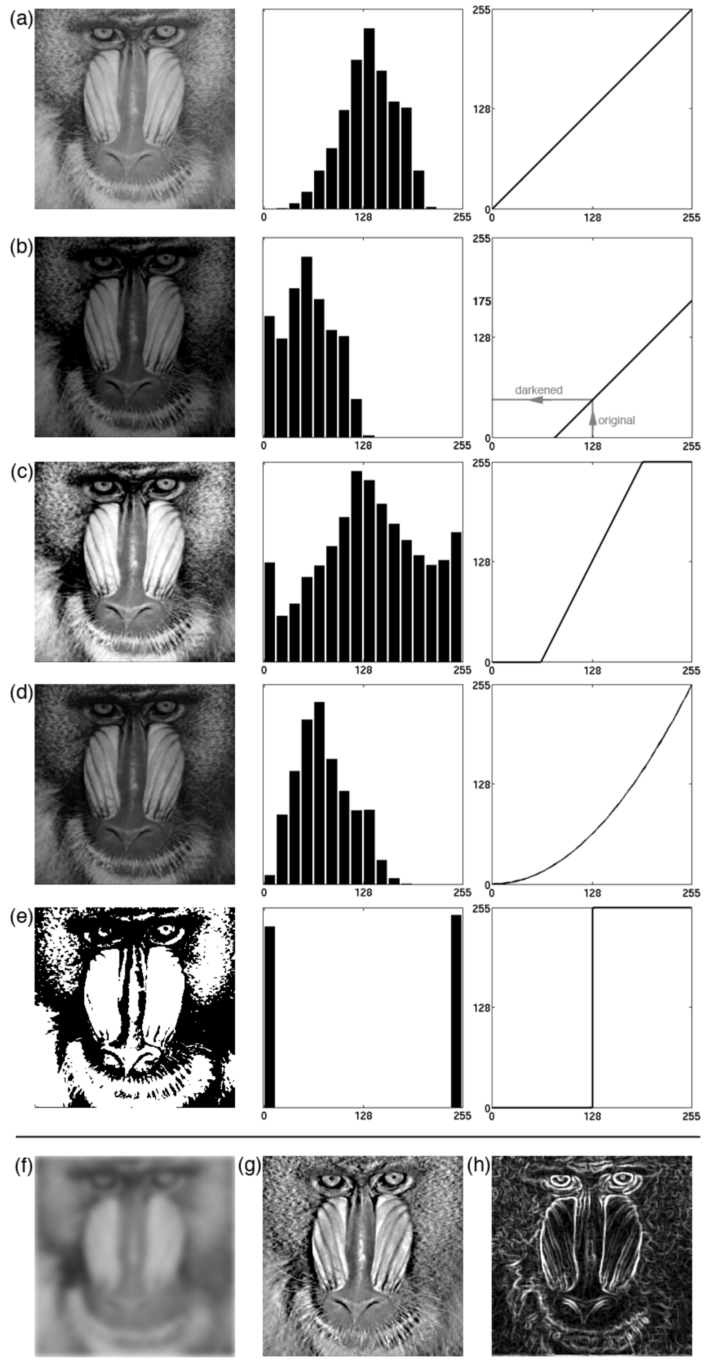

Figure 2: Shown is (a) an original grayscale image that has been: (b) darkened, (c) contrast enhanced, (d) gamma corrected, (e) quantized (thresholded), (f) blurred, (g) sharpened, and (h) edge extracted. Shown in the second column of panels (a)-(e) is the corresponding intensity histogram, and in the third column is the lookup table (LUT) that embodies the transformation.