Gamma, Perception, Posterization and Raw Conversion

To reproduce on paper or on a monitor a realistic representation of a scene we have photographed, we merely have to accurately reproduce the apparent brightness of each area of the scene. Exactly how we represent these brightnesses in the original raw file as recorded by the camera or other intermediate files is irrelevant. All we need is a one-to-one mapping between brightness of points in the original scene and how we store the information in files, and the same for an inverse mapping (which could well be media dependent) when we use the files to reproduce the image on paper or a monitor. Exactly what sort of mapping or representation that we use becomes important however, when efficiency, simplicity, and cost are also considered. For example, we may want to store and reproduce an image using the smallest possible file size (let us ignore file compression here) such that the human perceptual system cannot detect that a reduced file size and that a finite number of levels of greyness are being used. Imposition of these later considerations leads to the need for deeper analysis and, for example, illustrates why gamma in digitial photography is used. It also provides an explanation of why poserization in dark areas of images occurs more often than in bright areas.

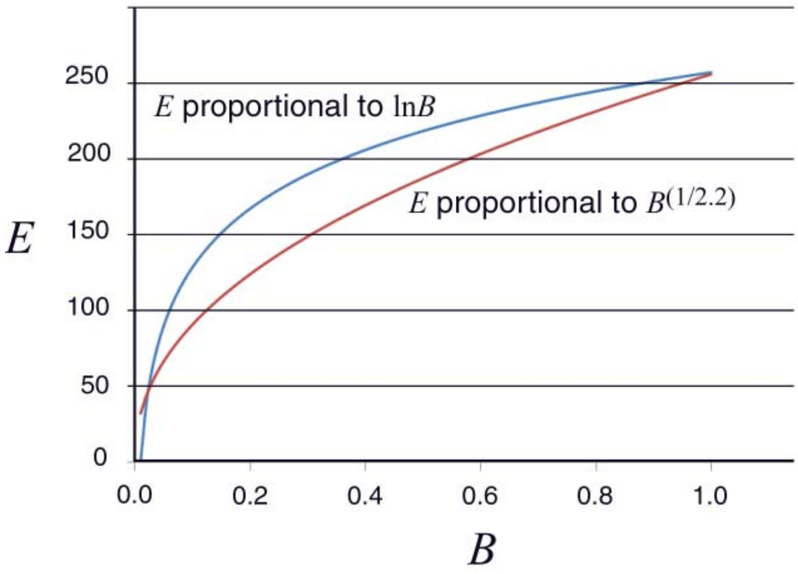
Over a fairly broad range, the human visual system obeys Weber's law in that two levels of gray can just be distinguished as different if they differ from one another by a certain percentage, and this percentage is the same from quite low to quite high levels of brightness. As a result, two levels of gray can be discerned as different if one has a brightness of 1 and the other is $1 \times (1+c)$times as bright. The next just discernable gray level would have a brightness of $1 \times (1+c) \times (1+c)$ and so on. Suppose we wish to present an image using a finite number of gray levels whose levels have been chosen to minimize our visual system's ability to detect that the image is composed of a series of discrete gray levels. That is, across the range of brightnesses in the scene, we should use a set of greys such that each differs from the next by a constant percentage.

How should the gray levels be chosen if, for example, the brightness difference between the darkest and brightest part of the image is to be 100 (which is about the maximum which can be achieved on a print) and we will allow 256 different levels of gray to be used? The choice of 256 comes about if we allow the 256 levels to be described by 8 binary bits, a useful bit depth for computer files. In this case then, if the darkest gray is a brightness of 0.01, the brightest white will be 1. Thus, $0.01(1+c)^{255} = 1.0$. Consequently, c = 0.0179, and from this, we can calculate the level of all the intermediate levels of gray. That is, the 256 levels of gray are now determined and the brightness of the nth gray level is $0.01(1+c)^{n}$ and n ranges from 0 to 255. The relationship between the actual brightness B of an area in the original image (for example, in the raw file produced by the camera) and the encoded or converted value, call it E (where E ranges from 0 to 255) is $B = 0.01(1+c)^{E}$. The fact that c is small permits the approximation

$(1 + c) \approx e^c$. Hence, $B = 0.01e^{cE}$. Equivalently, $E = {}^1\!/_c \ln(100B)$ where $\ln$ is natural log. This then is the ideal mapping between actual brightness and the representation of brightness. It maximizes image quality when only a finite number of gray levels can be used. If a sufficient number of grey levels are used, no posterization will be visible from the darkest to the lightest area of the image. If an insufficient number of grey levels are used, their presence will be equally perceptible across the entire tonal range of the image, that is, posterization will be no more visible in the dark shadows than it will be in the lightest parts of the image.

Why doesn't the ideal mapping that was derived in the previous paragraph contain something including the well known gamma? The reason for the use of gamma and its commonly chosen value of 2.2 comes from the properties of cathode ray tubes (early television) that once were used to display images. CRT tubes have an approximate 2.2 power law relation between their apparent brightness and an input voltage. In the early days, for convenience, image information was converted in the cameras by the transformation $B^{(1/2.2)}$. After transmission and receipt of this transformed signal, it was directly applied to the CRT tube. The 2.2 power dependence of the display produced an output of the input signal to the 2.2 power. In terms of the original brightness B, the brightness of the output was $(B^{(1/2.2)})^{2.2} = B$. That is, the original brightness is produced on the display screen without requiring any explicit transformation or remapping at the receiving end. This scheme has several virtues. First, it eliminates the need for special circuity to perform the inverse mapping when displaying because the CRT intrinsically does this. Second, properly used, it can compress brightness information into a relatively small number of discrete levels. In this coding-decoding scheme, an image will be perfectly restored if the number of gray levels that represent the original signal is sufficiently high that the human perceptual system cannot discern the difference between the adjacent and discrete values. How close does this scheme compare to the ideal mapping function that was derived in the previous paragraph?

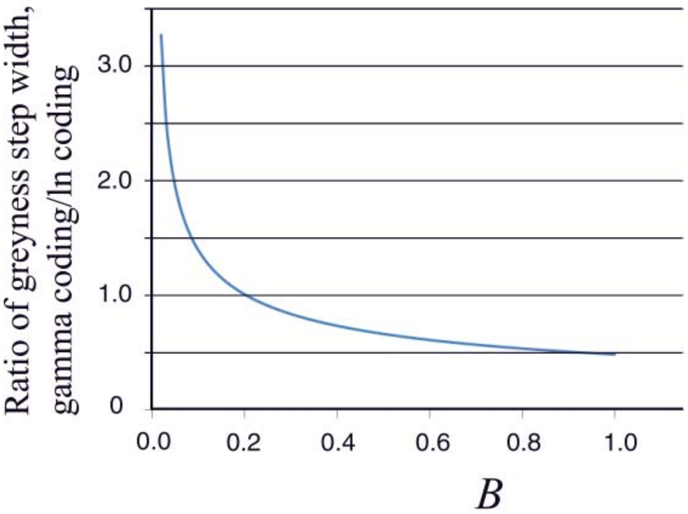
The graph shown below displays the ideal mapping function where E is proportional to lnB and also the mapping using power law mapping function that has come to be used because of the early use of CRT tubes. Both mapping functions encode the range of brightnesses ranging from 0.01 to 1 in 255 levels of gray.



As an aside, note the jump from the point (B,E)= (0,0) to the point (B,E)=(1,32) in the power law mapping. In the definition of the sRGB color space, this discontinuity is eliminated by changing the relationship between B and E over the very bottom part of the curve. A straight line from the origin connects to an early point on the curve. In reality, the situation becomes a little more complex than this, as is described in the site giving the full technical definition of sRGB, http://www.w3.org/Graphics/Color/sRGB.html. In the Adobe RGB color space a linear portion is not included, but the space covers a wider range of B and the linear approximation is not needed.

The figure above explains why posterization of dark tones occurs more easily than posterization of bright tones. For $B < 0.21$, the slope of the power law conversion curve is less than the slope of the logarithmic conversion curve. This means that for the power law conversion, in this region, any change in E produces a greater change in grayness of B than the ideal change as is described by the logarithmic conversion. That is, the grayness levels are further apart than ideal. In the region $B > 0.21$, the opposite is the case and hence grayness levels in bright areas are closer together than ideal.

The perceptual "size" (ratio of brightnesses) of the greyness steps for the gamma power law compared to step size ln-exponental encoding is plotted below. In the darkest areas of an image the steps in greyness are more than three times the ideal step size, and in the brightest areas, they are only one half of ideal, making a factor of more than six separating the perceptual step size between the darkest and lightest areas of an image. This shows why posterization of shadow areas is much more likely to occur than posterization of bright areas. One should also note that this perceptual variation in the step size is a consequence of the less than ideal nature of using a power law relationship to encode brightness.



Problems with posterization in the image files themselves are eliminated by using 16 bits in files to represent brightness rather than 8 bits. Since red, green and blue are each represented, such color files are sometimes referred to as 48 bit files. The larger file size does not fully eliminate the possibility of posterization on one's computer display however, unless the display uses more than 256 levels of brightness. Many displays and operating systems now use 10 bits (called 32 bit color since 10 bits are used for red, green, and blue, and 32 sounds better than 30) to encode brightness. This is sufficient to prevent discernable posterization in most situations despite the use of gamma=2.2 power law encoding.