What's Your Bit Depth?

Why having more "bits" of information garners better latitude. ▪ By C.R. Caillouet

caption to go here and here and here and here too! caption to go here and here and here and here too! caption to go here and here and here and here too! caption to go here and here and here and here too!

What does it mean to have 8-, 16-, 24- or 32-bit video? How does the number of bits affect the latitude of a camera? What's the difference between bits in the "front end" of a camera, bits at the camera output and bits on tape? And while we're at it, exactly how much is a bit anyway?

Anytime an analog signal is digitized, it's sampled, usually at a regular time interval, and quantized; in other words, its "amplitude" is measured. The value at each interval is represented digitally as a binary number, as a sequence of ones and zeros.

Each one or zero is a "bit" of information, and the total number of bits in the signal amplitude number determines the number of distinct levels available to describe the signal. Eight bits yield 256 levels; 10 bits, 1024 levels; 14 bits, 16,384 levels and so on.

The number of bits doesn't limit the amplitude of the signal because we can interpret the digital number any way we like, but it does limit the "granularity" or "precision" of the measurement.

If the maximum value is to be contained, then the smallest value is limited by the size of the digital number—the number of bits—and the darkest areas might lose detail, sometimes called "blocking up." If the detail in dark areas is maintained, then highlight detail may be lost.

In the real world, scene contrast range (the ratio of the brightest to darkest visible brightness values) may be as high as 1,000,000:1—or 20 bits of brightness information. In a single scene, it may range from 100:1 to 10,000:1 or 14 stops. To represent this range of actual light values would require at least 14 bit numbers. If quantized brightness values were to be maintained in the original form through a video system, luma (or brightness) values of 14 bits would be required. For a tri-color (red, green, blue) representation of the scene, a total of 42 bits (3 x 14) would be required per sample (pixel).

Most video systems can't handle the bit rates necessary to sustain continuous 42-bit sampling through recording, processing, transmission and display, so some sort of bit-rate reduction is needed.

Luckily, the human visual system doesn't respond linearly to scene brightness. The eye/brain system also needs to be as efficient as possible so that it can maximize scene resolution and contrast range without wasting resources. So if we could figure out how human vision deals with the problem, we might be able to build more efficient systems.

In fact, the eye responds logarithmically, not linearly, to brightness. The signal output from the sensitive components in the eye seems to be a string of pulses, the frequency of which varies with brightness but increases in even steps with each doubling of light input.

Humans sense a visible change in contrast with approximately a one-percent change in brightness. Because one percent of a dark stimulus is much smaller than one percent of a bright one, the brightness steps required in dark areas are much smaller than in bright areas. With this contrast "compression," the visual system perceives smooth gradations at both ends of the contrast range with significantly fewer bits than if linear light were to be represented directly. The minimum brightness change detectable in a dark video scene is roughly 0.1 percent of reference white.

A video camera contains a similar function called gamma. Gamma is a nonlinear brightness compression that follows a power function with an exponent of 0.45, close to a log function in shape. So dark values are expanded, and bright ones are compressed.

Video gamma, as defined in the current international standard (ITU Rec 709), provides for a maximum gain in dark areas of 4.5x and a corresponding reduction in bright areas. The result is a 9:1 difference in the value of one step between dark and light areas, or a gain of more than three bits (23 = 8) of precision over a linear representation.

So the 14 bits required to represent that 10,000:1 contrast range is reduced to 11 bits. With a little creative control of highlights from a knee circuit, 14 stops worth of scene brightness can be represented in about 10 bits, the number of luma (brightness) bits available in a serial digital-video stream often used in modern video systems.

If we were planning to view the output of the video system directly, without further processing or editing, then we could make use of the fact that the eye can't really respond to that 10,000:1 contrast range at one time. For a small screen, where the viewer takes in the entire scene at once, the available range is more like 1000:1 or less, and a corresponding bit depth of 8 bits in a gamma corrected signal will do.

That's why we can capture and process in a camera with 14 bits, distribute and edit with 10 bits, and record and display with eight. If we're careful and don't significantly change the contrast range in postproduction, we can actually get by with those 8 bits through the editing process. Many high-definition programs today are captured and recorded using HDCAM or DVCPRO HD 8-bit systems, then edited and delivered at the same bit depth.

Editing an 8-bit signal at a higher bit depth can help reduce processing losses and help to deliver usable video programs from sources with limited bit depth, but aggressive color or contrast correction can introduce contouring in smooth gradations. Recording with more bits per pixel is always desirable, if not always practical.

Notice that in this discussion (even though the term compression was used), bit-rate reduction isn't from reduced resolution or from eliminating redundant video content. Those are the techniques usually referred to as digital video compression. HDVP