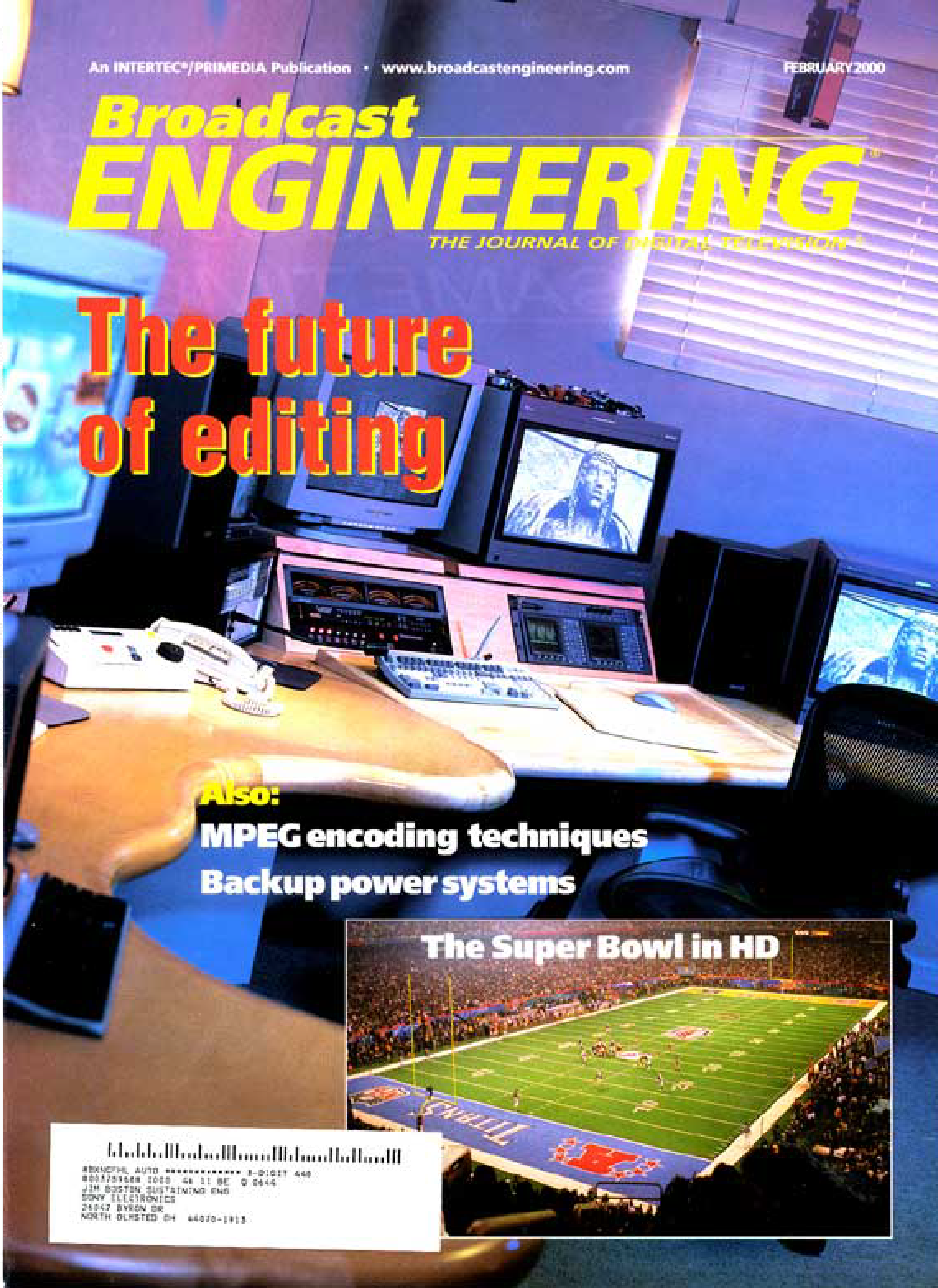
An INTERTEC®/PRIMEDIA Publication • www.broadcastengineering.com FEBRUARY 2000

Broadcast ENGINEERING®

THE JOURNAL OF DIGITAL TELEVISION®



The future of editing

Also:

MPEG encoding techniques

Backup power systems

The Super Bowl in HD

#BXNCFHL AUTO ************** 5-DIGIT 440
#0035759688 2000 46 11 BE Q 0646
JIM BOSTON SUSTAINING ENG
SONY ELECTRONICS
26047 BYRON DR
NORTH OLMSTED OH 44070-1913

Transition to Digital

Video compression

BY JIM BOSTON

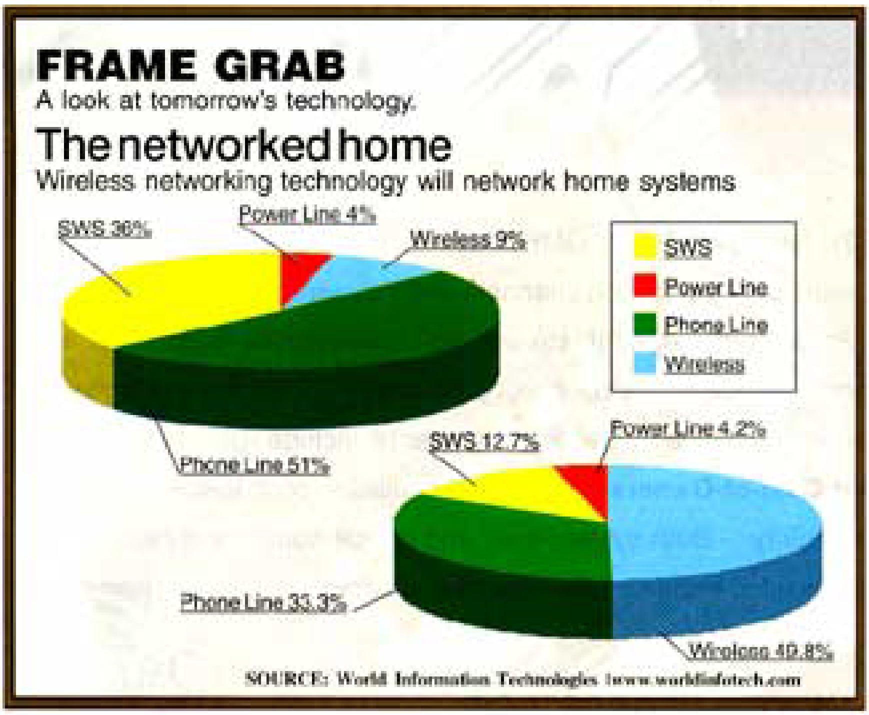
Remember when we used to build television facilities? We would string coax around and terminate it eventually in an AM transmitter. At the same time we would lay an audio web on top of the video. This audio layer would eventually terminate in a FM transmitter built into the same box that housed the AM half. Today, it seems we are building data facilities that happen to be used for television. Sure, the original two layers are still there, but now they are wrapped with numerous other layers. These new layers include SD and HD digital video paths, AES paths, and a web of data infrastructure. Because these new paths contain digital data, the bandwidth requirements have exploded. It is likely that digital would still be found only in a few niche applications today if JPEG (followed by MPEG) had not come along.

Digital baseband video creates too many bits to cope with in some paths, at least from an economical standpoint.

MPEG's place today

Not many have decided (yet) to move MPEG around their facilities. But MPEG can be found at most of the bandwidth choke points that digital creates. These points are generally any paths into or out of the facility, along with paths leading to or from digital video storage. Digital baseband video creates too many bits to cope with in some paths, at least from an economical standpoint. Methods were devised to create bit streams with just enough bits to adequately describe the video or audio content. This is done by throwing away some of the information in the picture, hence lossy compression. Compression is done in the spatial dimension (JPEG, MPEG I frames), and in the temporal or time dimension (MPEG B, P frames). Spatial compression is accomplished by converting video from the time domain to the frequency domain, and then eliminating frequency components that are not noticed (hopefully). (See Figure 1.) In reality, some lossless compression is also performed after the lossy compression. Temporal compression is accomplished by looking at differences between I (anchor) frames. These differences are used to create motion vectors. Video blocks from earlier frames that have moved are not sent again. Instead, motion vectors describing where this earlier video has moved are sent. Also a differential picture is created between the previous anchor and the current B or P frame. Both these data sets are then compressed. JPEG (I frame) spatial compression can deliver good quality video at up to eight times compression, MPEG temporal compression can increase the compression rate by another factor of five.

However, there are issues before any compression takes place that can play a great part in the final quality of the video in the compressed bitstream. Noise makes compression engines work very hard. Randomness breaks compression systems, and nothing is as random as noise. Some MPEG encoders provide an indication of DCT coefficient quantification. What does that mean? DCT is the process that takes blocks of video from the time domain to the frequency domain. The coeffi-



Video bit rates (active video)

4:4:4 bit rate (720*486) = (720+720+720) * 486 * 30 * 10 = 315 Mb/s
4:2:2 bit rate (720*486) = (720+360+360) * 486 * 30 * 10 = 210 Mb/s
4:2:0 bit rate (720*486) = (720*486) + (360*243) + (360*243) * 30 * 8 = 126 Mb/s
4:1:1 bit rate (720*486) = (720+180+180) * 486 * 30 * 8 = 126 Mb/s

Note: 4:2:0 & 4:1:1 are eight bit only (to further reduce bit rate) whereas the other two are eight or 10 bits.

Table 1. The uncompressed bit rate is based on the number of samples, the number of bits per sample and the sampling structure used.

cients in essence are values that describe the values of frequencies needed to reproduce the video block. This is much like the frequency values produced by a Fourier transform. If the DCT process creates too many coefficients, the encoder increases the value by which all the coefficients are divided or scaled. The higher the dividing number, the more likely that coefficients will be rounded to zero after division. Noise increases these coefficient values, with the result being that everything must be compressed further to fit within a given bandwidth.

Sampling

Another issue that affects compression is chroma subsampling. SMPTE 259 subsamples the chroma at half the luminance rate. We have always sent less chroma information than luminance information. NTSC provides considerably less chroma info than luminance. This was done because our eyes are less sensitive to chrominance than luminance. SMPTE 259 digital component uses the 4:2:2 sampling scheme. This means that there are half the chroma samples of luminance samples per horizontal line. But vertically there is no chroma subsampling. To allow for chroma subsampling in both the horizontal and vertical directions 4:2:0 is offered. One method is a quincunx pattern in that on odd horizontal lines you throw away even chroma samples, and on even lines the odd chroma samples are dropped. This creates an interleaved 4:1:1 sample rate. (See Figure 2.) 4:2:0 sampling creates 25 percent less data than 4:2:2 sampling. So an argument arose as to whether it is better to under-sample the chroma information up front and therefore end up with less information (at least for chroma) to compress. Or is it better to have more chroma info to start with and to compress a little harder. The EBU and the CBC addressed that issue a couple of years ago. They found that picture quality was essentially the same (4:2:2 slightly better) unless the bit rates got extremely low. Then 4:2:0 had a slight advantage.

But what was borne out in the test was that with multiple compression/decompression cycles 4:2:2 became the clear winner. This illustrates that starting out with more is still better than starting with less. Sometimes we forget why we buy broadcast-quality equipment instead of industrial or consumer stuff. The transfer function, comparing the output to the input, is never 1 to 1 in any analog box. Each pass through a box creates a slight reordering of frequencies and phase, and usually some rolloff of the high end of the spectrum. The more information you start with means the more you will end up with. Now with digital, once we're in that domain and we make no changes, we should incur no change as we cascade through digital boxes. But it is at points were we make domain changes that we incur quality degradation. Analog to digital, digital to analog, composite to component (digital or analog), component to composite, baseband digital to compressed digital, and vice versa, all these domain changes impact the video.

When we talk levels in MPEG we mean sample structure and bit rate. Sampling structure and bit rate have a great effect on how many times we can compress and decompress before the quality gives out. MPEG profiles refer to the tool sets available for temporal (time) compression, including I, B and P frames. More B and P frames between I frames (long GOPs) means that frames towards the end of the

Randomness breaks compression systems, and nothing is as random as noise.

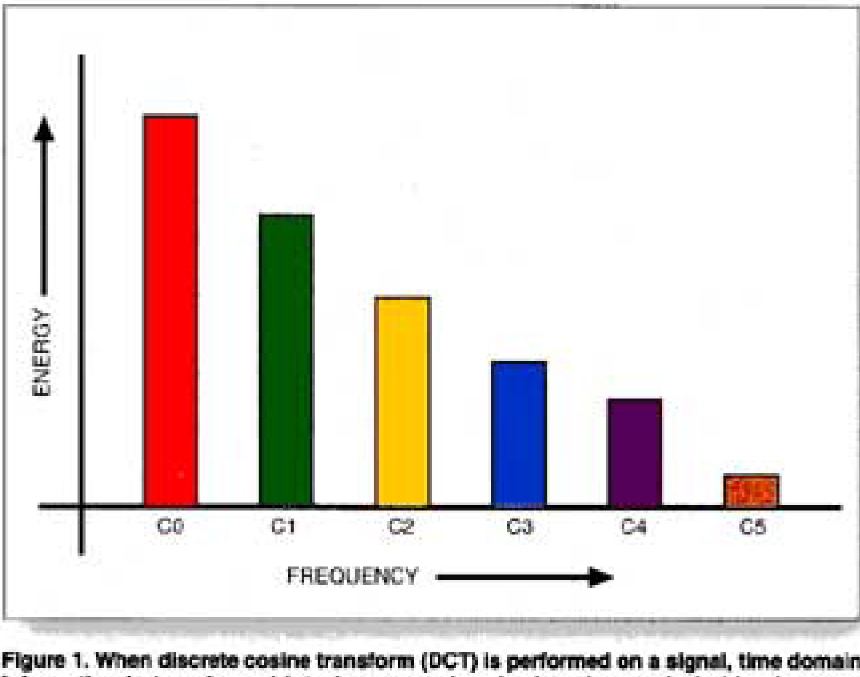


Figure 1. When discrete cosine transform (DCT) is performed on a signal, time domain information is transformed into frequency domain data. In a typical video image, a majority of the frequency components can be discribed by a single coefficient (C0), harmonics are then described using additional coefficients (C1-5). Each additional coefficient requires extra bandwidth.

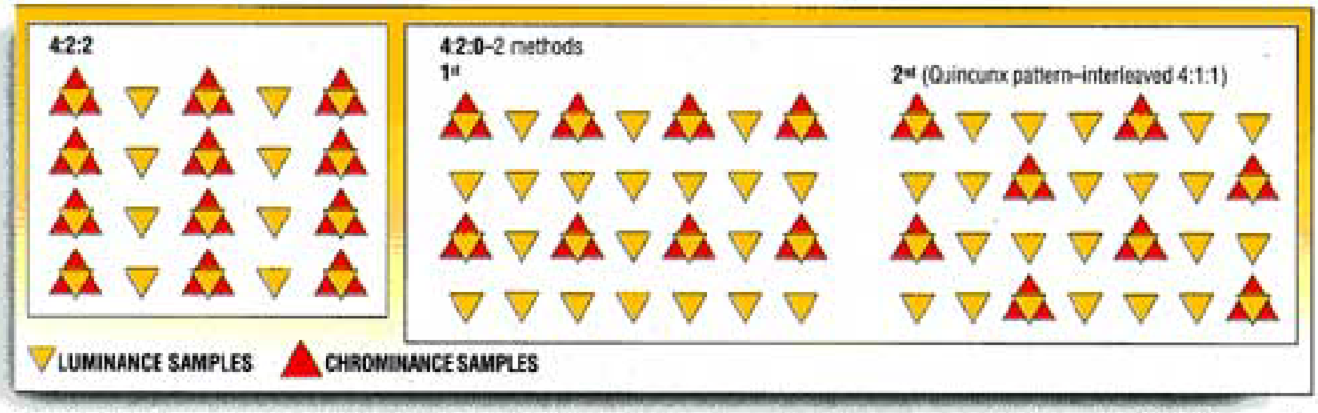


Figure 2. Not all sampling structures use the same techniques for sampling video. The arrangement of the co-sited Cr/Y/Cb samples depends on the way the sampling structure is defined.

GOP will become less exacting versions of the original video. Multiple compression/decompression cycles only exacerbate the situation. Required bitrates to get the job done are a rapidly moving target. Contribution rates are still generally high, 20Mb/s and above. These high bit rates insure that the received video can be acted upon, stored, taken to baseband and back a few times and stay reasonably intact. Bit rates sent to the end user can be quite a bit less because they are expected displayed by the user and discarded. What happens when DTV VCRs finally come on the scene is an unanswered question. Many are sending virtual SD DTV programs at 8Mb/s, while others think technology has reached the point where 6Mb/s is good enough. On the HD side 15Mb/s is often used, but some are starting to look to 12Mb/s as enough.

Many seem to think that working with digital video today means the technical quality judgements required in the past are no longer needed. This simply is not true. What you do in the analog, and digital domains, along with the decisions you make moving from the baseband to compressed domains will mean that your product will stand out from your competitors, either favorably or as an example of not what to do. ■

Jim Boston is a project engineer for Scripps Howard based in Cincinnati.

Send questions and comments to: brad_dick@intertec.com