SMPTE TUTORIAL

Serial Video Basics

By David Strachan and Randy Conrod
Illustrated by Michel Proulx

In the March issue of the *Journal* we described $4f_{sc}$ composite and 4:2:2 component digital television signals. Now let's take a look at the relative merits of these two signal types.

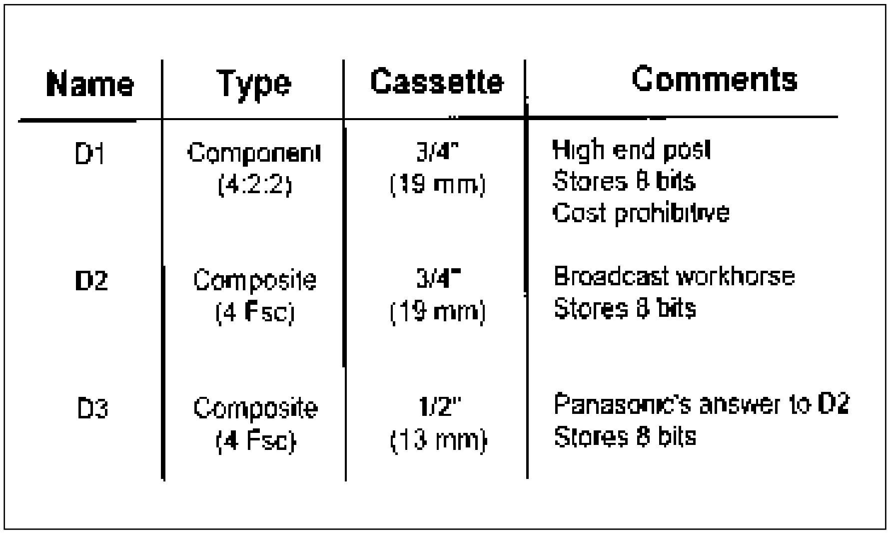
The world has been broadcasting television signals in composite analog form since television was invented. When color TV was first introduced, the authorities decided that the new signal should retain compatibility with monochrome TV sets, so the color information was encoded into a composite signal. Although it was recognized that the encoding and decoding process introduces some abberations to the picture quality, these were minor compared to transmission errors and other picture impairments encountered in the early days of television. More recently, in a quest to improve picture quality, some TV stations, particularly in Europe, have upgraded their facilities to use component analog production and distribution equipment. In France, it is even possible to receive component analog broadcasts via satellite. Betacam and MII component analog tape machines have long been preferred over composite recording devices. All of this evidence might suggest that the 4:2:2 signal would enjoy an overwhelming advantage over $4f_{sc}$, and technically it does. But while we have to deal with a largely composite world, there are still a few remaining benefits to using $4f_{sc}$. Most new digital facilities world-wide have adopted the 4:2:2 format, but many have also chosen $4f_{sc}$.

Tape Formats

If you want to make television pictures, you have to be able to record them — and if you want digital TV, you need a digital VTR. The first commercially available VTRs recorded 8-bit component 4:2:2 signals, and the format became known as D-1. As explained earlier, the great thing about the D-1 format is that it uses the same 13.5-MHz sampling frequency for both 525 and 625-line television signals. This does not mean that a signal that has been recorded on one standard may be played back on the other, but that the hardware is identical for both. The first D-1 machines were large and expensive, but many post-production facilities were prepared to pay the price because these machines offered the advantage of noise-free multigeneration recording for the first time.

Developers were aware that this particular advantage could be realized much more inexpensively by directly digitizing the composite analog signal into a composite digital format, and so D-2 was born and launched at NAB, a few years after D-1. The format was supported by Ampex and Sony; subsequently, Panasonic introduced its composite video D-3 machines. Although the tapes are not compatible between D-2 and D-3, they all record the same $4f_{sc}$ signal type. These machines also have composite analog inputs and outputs, and many of them have simply been used to replace earlier C-format VTRs. As $4f_{sc}$ (4 times the subcarrier) in NTSC is not the same as $4f_{sc}$ in PAL, we have D-2 and D-3 NTSC, and D-2 and D-3 PAL (Fig. 1). The devices have proved to be very practical for broadcasting, and large numbers have been supplied for both 525 and 625-line operation. Their acceptance led to the development of $4f_{sc}$ vision mixers and other devices based on the transmission of digital video clocked at 4 times subcarrier and, as we have seen, some TV stations have adopted the composite digital format as their in-house standard. But although these machines are in wide circulation, they are still composite and still have encoded abberations. So, after securing a lot of D-2 and D-3 sales, back come the manufacturers with less expensive component digital VTRs — and, of course, another battle begins.

Two years ago, Sony introduced Digital Betacam, Ampex introduced DCT, and Panasonic introduced D-5



Name	Type	Cassette	Comments
D1	Component (4:2:2)	3/4" (19 mm)	High end post Stores 8 bits Cost prohibitive
D2	Composite (4 Fsc)	3/4" (19 mm)	Broadcast workhorse Stores 8 bits
D3	Composite (4 Fsc)	1/2" (13 mm)	Panasonic's answer to D2 Stores 8 bits

Figure 1. Established DVTR formats.

A contribution from David Strachan, Randy Conrod, and Michel Proulx, Leitch Video International Inc., North York, Ont., Canada M3B 1V7. This is the second in a series of tutorial papers by these authors. Copyright © 1995 by the Society of Motion Picture and Television Engineers, Inc.

Name	Type	Cassette	Comments
DCT	Component (4:2:2)	3/4" (19 mm)	Ampex's new DVTR Stores 8 bits Uses compression
Digital Betacam	Component (4:2:2)	1/2" (13 mm)	Stores 10 bits Some models play Betacam Uses compression
D5	Component (4:2:2)	1/2" (13 mm)	Stores 10 bits Compatible with D3 No Compression

Figure 2. New DVTR formats.

into the marketplace. Ampex machines record and play back an 8-bit compressed signal (more on compression later), Sony machines record and play back a 10-bit compressed signal, and Panasonic machines record and play back a 10-bit uncompressed signal. None of these tape formats are compatible with each other (Fig. 2), but competition has forced the price down and 4:2:2 is now a very viable recording format.

Other Considerations

The composite $4f_{sc}$ digital television signal benefits from another advantage. When converting from NTSC (and especially from PAL) into component digital video, expensive decoders are required. So if you have many analog TV signals coming into your TV station and they are to remain largely unchanged before being retransmitted, you can spare some expense by leaving the signals in composite form. It is when you want to add effects, generate graphics, save stills, and do all the other post-production things that tend to take place in a modern TV facility, that you see the real benefits of 4:2:2. Let's now examine the two signal types in more detail.

Signal Formats

$4f_{sc}$

When a signal is sampled at $4f_{sc}$, a 10-bit word (1024 levels) is written at a rate of 14.3 MHz in NTSC and 17.73 MHz in PAL. The entire signal is sampled, including the horizontal and vertical blanking intervals, containing sync pulses and the color burst. As is the case in analog video, the vertical and horizontal blanking intervals in the $4f_{sc}$ signal also provide room to carry additional information along with the digital video signal. In the $4f_{sc}$ sampled signals, ancillary data can be placed in the tips of the synchronizing signals (horizontal sync, vertical sync, and pre- and post-equalizing pulses). Any additional information that can be added is called ancillary data.

4:2:2

When a signal is sampled at 4:2:2, a 10-bit word (1024 levels) is written by taking a sample across all three channels (Y, R-Y, B-Y). This is called a co-sited sample and is referred to as Y, Cr, Cb. A sample of the Y channel alone is then taken and then another co-sited sample, and so on, alternating between Y and Y, Cr, Cb. These samples are taken at a frequency of 13.5 MHz for the Y channel and 6.75 MHz for the color-difference channels R-Y and B-Y. When multiplexed, the resultant sample frequency is 27 MHz. In the 4:2:2 sampling standard, the blanking intervals are not sampled. Instead, end of active video (EAV) and start of active video (SAV) data words are inserted in the data stream as markers. Ancillary data is placed between the EAV and SAV data words. In fact, in the component digital domain, there is spare space for over 55 Mbits of ancillary data, enough room to carry digital audio, time code, error, detection, and handling information with plenty of room to spare for future use.

4:2:2:4 and Others

The 4:2:2 sampled signal is the data signal used between such devices as tape transports, switchers, etc. But there are other sampling standards used, as well. When a key signal is needed with the video signal, the

Format	Parallel Clock Rate	Serial Data Rate
4 Fsc - NTSC	14.3 MHz	143 Mbps
4 Fsc - PAL	17.7 MHz	177 Mbps
4:2:2	27.0 MHz (13.5, 6.75, 6.75)	270 Mbps
16:9	13.5 MHz	270 Mbps
	18 MHz	360 Mbps

Figure 3. Sampling and serial rates.

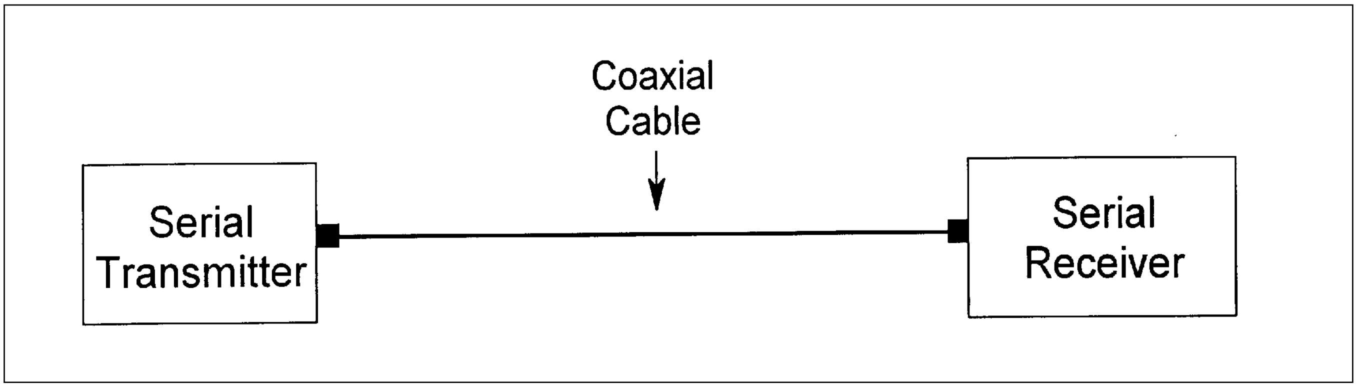


Figure 4. The serial digital interface.

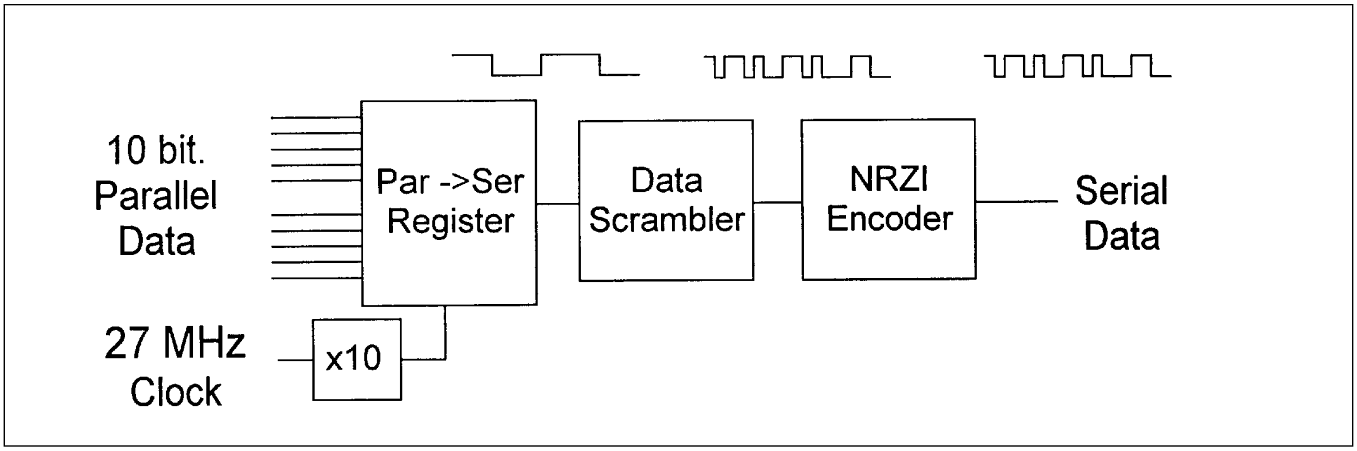


Figure 5. Simplified block: serial transmitter.

combination is known as 4:2:2:4. The 4:2:2 describes the video signal and the last 4 describes the full-bandwidth key signal. In reality, the key signal, when sampled, becomes 4:0:0, because it is monochrome and the color-difference signals are set to zero when sampled. Two cables are required to distribute a 4:2:2:4 signal, whether it be parallel or serial. The 4:4:4 is full-bandwidth sampling of the Y, R-Y, and B-Y channels.

Co-sited samples are taken at a rate of 13.5 MHz, with a resultant sampling rate when multiplexed of 40.5 MHz. This signal is typically used inside a device for high-quality digital processing (e.g., color correction). The 4:4:4 signals can be distributed in a parallel or serial fashion using two cables in a facility. When a 4:0:0 key signal is used with a 4:4:4 signal, it is called 4:4:4:4 and three cables are required if interfacing equipment. If an even higher resolution is required, the video may be sampled at 8:8:8. This is 27-MHz co-sited sampling on all channels and is used for internal processing only.

Distribution of Digital Signals

Now that we have an understanding of how an analog signal is sampled, the question is, how do we move this signal around a facility? There are two forms of digital interfaces: parallel and serial. The parallel interface consists of 25-pin DB-style connectors at each end of a 25-conductor cable. Distances of up to 50 m (150 ft) are allowed. Beyond that point, the data and related clock information begin to “skew” in time because of their differing frequency content. The clock and data information arrive at different times, causing data errors. In addition, the cable and connectors are bulky, difficult to use in a large facility, and expensive to implement.

The serial interface uses BNC connectors at each end of a 75-Ω coaxial cable, the same type as that used for conventional analog video signals. Parallel data must be serialized into a 1 bit wide path and transmitted down the coaxial cable at 10 times the original data rate. With serial transmission, distances of up to 300 m (1000 ft) are possible before the data can no longer be recovered. Existing analog cables may be used but, as serial signals have a very wide bandwidth, care should be taken to avoid cables with corroded connectors and poor grounds, as bit errors may result, ultimately causing the loss of the signal. A note for systems design purposes: passive looping inputs typically are not provided at the input to a serial digital device, as adequate return loss characteristics cannot be maintained. Reflections in the cable caused by poor return loss can lead to bit errors in the digital stream.

Sample Rates and Serial Data Rates

Figure 3 shows sample rates for different standards. Parallel NTSC signals are sampled at 14.3 MHz, so

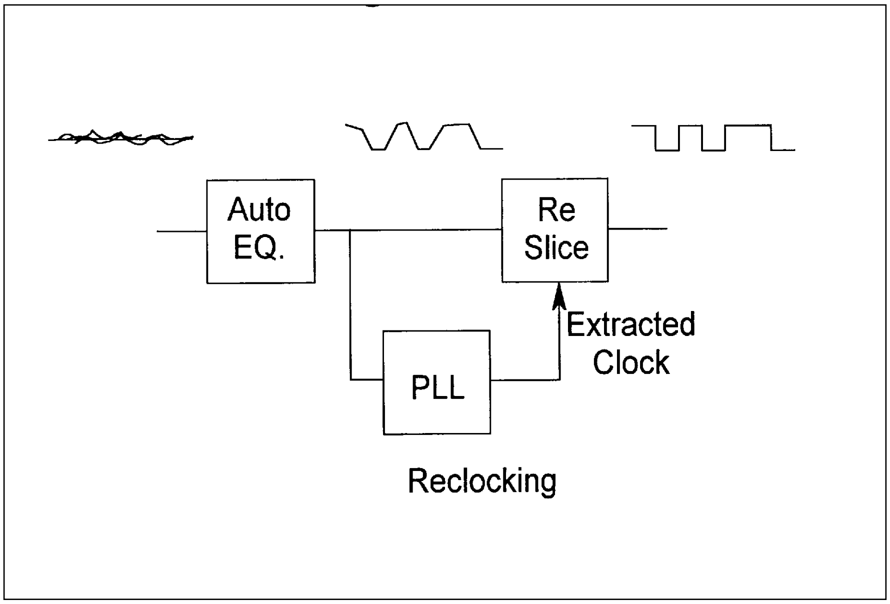


Figure 6. Reclocking serial receiver.

the resulting serial data rate (1-bit word) is 143 Mbits/sec. In PAL, the signal is sampled at 17.7 MHz, resulting in a serial data rate of 177 Mbits/sec. The multiplexed sample rate for sampling analog component is 27 MHz. When we serialize the 27-MHz signal, the resulting serial data rate is 270 Mbits/sec. A 16:9 format is being developed and two sampling frequency options are being considered for the Y channel (13.5 MHz and 18.0 MHz). The resulting data rates are 270 Mbits/sec and 360 Mbits/sec, respectively. These future formats will be discussed in a paper on the 16:9 aspect ratio (HDTV, ATV, EDTV, etc.), to be published in a later issue.

Serial Digital Concepts

The serial digital interface, or SDI (Fig. 4), is made up of the serial transmitter (the circuitry just ahead of the output BNC connector), a length of coaxial cable with BNC connectors at each end, and a serial receiver at the end of the cable (the circuitry just after the input BNC connector). At the transmit end, the parallel data is shift registered using a 10 times parallel rate clock. The resulting 1 bit wide serial data is passed through a data scrambler. A known algorithm is applied to ensure that there are a sufficient number of transitions to recover the clock signal at the receiving end and to limit any DC content on the signal. The data is then passed through a nonreturn to zero inverted (NRZI) encoder, which takes a transition in the signal and changes it to a 0 and takes a nontransition in the signal and changes it to a 1. This further limits any DC content in the signal. The serial data is then sent down the coaxial cable (Fig. 5).

The square wave is affected by the losses in the cable (one per frequency) and is attenuated and distorted. Take note that the higher the data rate, the shorter the cable length that the signal can pass through. The 270-Hz (component digital) signals will not pass through as much cable as a signal at 143 Mbits/sec (composite digital). The signal starts out with a peak-to-peak voltage of 800 mV ± 10%. At the end of a 220-m (600-ft) cable run using a 270-Mbit/sec signal, the peak-to-peak level will be attenuated down to only about 30 mV. (A 143-Mbit/sec signal after 300 m is also 30 mV.) Not only will the amplitude be reduced, but also the square-shaped pulses will have become more like a sine wave. So in reality, it is amazing that the signal can be fully recovered at the receive end and reproduced exactly the same as the original signal.

There are two types of receivers — reclocking and nonreclocking. Reclocking receivers use a phase lock loop (PLL) circuitry to clock the incoming pulses and restore the original square-wave shape (Fig. 6). Nonreclocking receivers simply pass the rounded pulses. Either type may incorporate automatic equalization circuitry to restore the level to 800 mV once again. Clearly, input stages with automatic equalization and reclocking are needed when cable runs approaching 300 m are involved. There is also the possibility of using edge-shaping techniques instead of reclocking, but when edge shaping is used, jitter is not removed from the signal and reliable recovery is not possible after 50 m of cable at 270 Mbits/sec. (Remember that when a cable length is specified, the data rate must be included in the specification.)

The Cliff Effect

With digital video, you might be tempted to look at the picture on a monitor and say, “Looks great! Why worry about EQ and reclocking?” This is where we can be lured into a false sense of security. In digital television, either you have a perfect picture or you have nothing. If you take a 200-m piece of coaxial cable and replace it with a 250-m piece, you may see no difference in the signal. But as you add longer and longer lengths, there comes a point when, by adding only one or two meters, the signal becomes unusable. This is known as the cliff effect. Without warning, your signal has suddenly disappeared.

In the analog domain, we can see the effects of cable as a loss of amplitude and high-frequency response, which happens gradually as the cable is lengthened. In the digital domain, it is possible to be on the “edge of the cliff” and not realize it. There is no indication that you are close to the edge of the “cliff.” What happens, as we approach the edge of the cliff, is that errors appear in the bit stream and increase dramatically as additional obstacles are inserted in the path. As digital video is supposed to be entirely error-free, we can use this assumption to provide an early warning system. This is known as error detection and handling (EDH) and will be discussed in a later article in this series.