

SMPTE ST 2110-20: Delivering Pixels (Repost)

This article is one in a series introducing the newly published SMPTE standards for fundamental media streaming over IP networks, focusing on video transmission—specifically, the packet format that will replace uncompressed SDI signals.

Uncompressed Active Video

The full name of SMPTE ST 2110-20 is "Professional Media Over Managed IP Networks: Uncompressed Active Video," which clearly indicates its scope. Explicitly, it transmits uncompressed video images (like SDI), but only the data constituting the "active" portion of each video frame. In other words, the image samples (also known as pixels) that form the picture delivered to the audience are transmitted within this new IP video format.

This approach significantly reduces the amount of data required to represent each video frame. For example, a 59.97 Hz 720p video signal using a 1.483 Gb/s SDI signal can be transmitted as an IP packet stream at a data rate below 1.18 Gb/s—a reduction of 20%.

However, this approach has side effects, including eliminating the HANC (Horizontal Ancillary Data) intervals often used for embedding audio signals and the VANC (Vertical Ancillary Data) intervals commonly used to transmit information such as timecode, format descriptions, ad triggers, captions, and many other useful pieces of data. Audio and other non-video data types are transmitted as separate IP streams defined in other parts of the ST 2110 series.

To avoid wasting bandwidth, data samples from adjacent pixels are packed directly together without interleaved header information. This packing is organized using pixel groups whose size and data layout are highly dependent on the specific video format being transmitted. Tables within ST 2110-20 define numerous different video sampling formats that can be transmitted, including most, if not all, formats used in professional video production, such as RGB, Y'C' BC' R, ICTCP, XYZ, and video keying formats. Since pixel group size varies with bits per sample and chroma subsampling format (4:2:2, 4:2:0, etc.), each permitted format is specified within this document.

To ensure interoperability, the order of individual color samples within each pixel group is also specified. For example, a 4:2:2 10-bit video (standard HD-SDI) pixel group is 5 bytes (40 bits) long, containing 4 samples representing two pixels in the order C' B' Y0' C' R Y1'. Up to 286 such pixel groups can fit into a standard UDP datagram, carrying 572 pixels. Other sampling formats will have different numbers of pixel groups and pixels per datagram.

Defining the Image in Under a Thousand Words

The majority of the remainder of the ST 2110-20 document is dedicated to defining the precise format of the transmitted video signal. The Session Description Protocol (SDP), specified in RFC 4566, provides a machine-readable layout for this information.

参数	描述	例子
宽度	图像宽度，像素内	宽度=1920
高度	图像高度，像素形式	高度=1080
精确帧率	视频帧率，两个整数之比	精确帧率=30000/1001
取样	色彩取样格式；允许选择列于ST 2110-20	取样=YCbCr-4:2:2
深度	每个图像取样位数	深度=10
色度	用于呈现颜色的标准 其它选择包括BT601、BT2020	色度=BT709
PM	打包模式，通用或块（每块180字节）	PM=2110GPM
SSN	用于参数的SMPTE标准数字，包含修订年	SSN=ST110-20;2017
TCS	传输特性系统 其它选择包括PQ、HLG	TCS=SDR
隔行	指示视频信号为隔行扫描	隔行

Many different parameters need to be specified to fully define a video format. Figure 1 lists parameters required by ST 2110-20 and several that are only used in certain contexts. The last column in Figure 1 shows the values taken by each listed parameter for an SDP-defined interlaced 1920x1080 video signal running at 29.97 frames per second (adopted by many broadcasters in North America

and Japan). Note that other non-mandatory parameters are also defined within ST 2110-20, such as PAR (Pixel Aspect Ratio) for non-square pixels (used in standard definition video) and MAXUDP for UDP datagrams differing from the standard 1460-byte size.

Each RTP datagram may contain up to 3 Sample Row Data (SRD) headers, each describing a segment of sample row data. Each header indicates the image row (i.e., video line) number and the horizontal offset of the first pixel within the SRD segment. Using this information, the receiving device can reassemble multiple datagrams into a complete video frame. Note that the first row of each image (video frame) is row number 0, which differs from the numbering scheme used for lines within SDI video signals. Care must be taken when converting between SDI and ST 2110-20 line numbering.

Getting the Picture Right

A benefit of the complete video format specification defined in ST 2110-20 is the elimination of ambiguity prevalent in many industry documents. For instance, it doesn't take much searching to find specification sheets listing some equipment as "1080i59.94" and others as "1080i/29.97"—sometimes even from the same manufacturer! With a complete set of SDP required for every video signal, getting the picture right will become much easier.