

IP 방송제작 기술 21

SMPTE ST 2110-20 Part - 2

글. 조인준 KBS 미디어기술연구소 차장

지난 연재에서 SMPTE ST 2110의 핵심이라 할 수 있는 무압축 비디오 데이터 IP 전송을 위한 SMPTE ST 2110-20 'Professional Media Over Managed IP Networks : Uncompressed Active Video'의 소개를 시작하며 [그림 1]의 Pixel Group Data의 데이터 구조에 관해 소개해 드렸습니다. 이번 연재에서는 Pixel Group Data에 관한 정보를 담고 있는 [그림 2]의 초록색 'RTP Header'와 하늘색 'RTP Payload Header'의 구조에 관해 설명해 드리겠습니다.

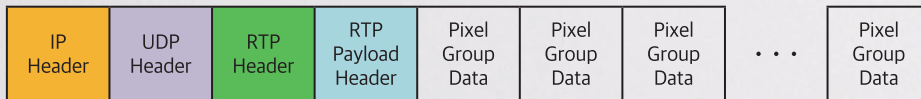


그림 1. SMPTE ST 2110-20 IP 패킷 구조

• SMPTE ST 2110-20 RTP Header

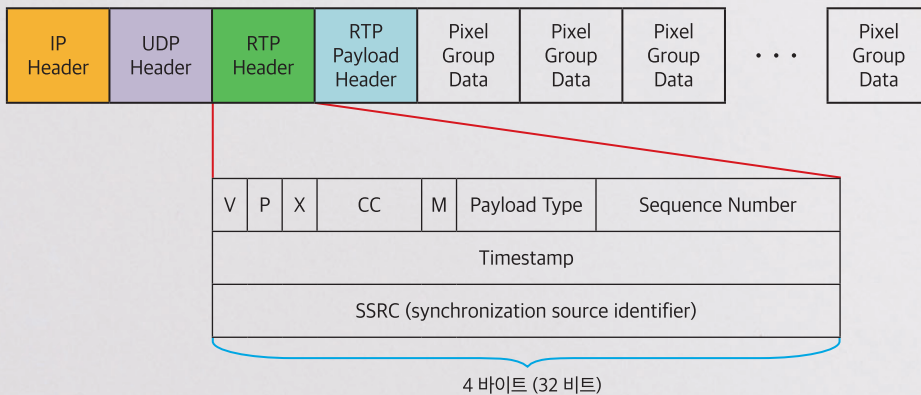


그림 2. SMPTE ST 2110-20 IP 패킷 RTP Header 구조

V (Version)	- RTP 버전을 나타내는 2비트 필드이며 현재 SMPTE 2110-20에서 사용하는 RTP 버전 값은 2임
V (Version)	- RTP로 전송되는 데이터 뒤에 추가로 패딩 되는 데이터 유무를 나타내는 1비트 필드 - 0이면 패딩 데이터 없음 - 1이면 패딩 데이터 있음
X (Extension)	- 헤더 확장 여부를 나타내는 1비트 필드이며, 이 값이 1일 경우 확장 헤더가 SSRC 바로 뒤에 이어짐
CC (Contributor Source Identifier Counter)	- 4비트 데이터 필드 - 영상을 만드는데 사용된 소스 영상을 제공한 장치의 개수를 나타냄 - 영상 소스의 SMPTE ST 2110-20 패킷의 경우 0을 가짐 - 비디오 믹서의 SMPTE ST 2110-20 패킷의 경우 출력 영상을 만드는데 사용된 영상 소스의 개수를 나타내며 최대 15까지 표시할 수 있음 ※ SMPTE ST 2110-20에서는 CC가 0이 아닐 경우 SSRC 뒤에 붙는 CSRC(Contributor Source Identifier)에 대한 별다른 언급이 없어 실제 0 이외의 CC 값을 사용하는지에 관해서는 확실치 않음
M (Marker)	- 비디오 프레임의 마지막 데이터를 나타내는 1비트 필드 - 해당 패킷이 비디오 프레임의 마지막 데이터일 경우 1 - 그 이외의 경우는 0
Payload Type	- 패킷으로 전송되는 데이터의 식별을 위한 7비트 필드 - SMPTE ST 2110-20의 경우 96~127 사이의 값을 동적으로 할당하여 사용 - SMPTE ST 2110-10의 SDP를 사용하여 동적으로 생성된 Payload Type을 수신 장치에 전달하고, 수신 장치는 이 값을 이용하여 수신된 패킷을 해석
Sequence Number	- RTP로 전송되는 패킷의 순서를 구분하기 위한 16비트 필드 - SMPTE ST 2110-20에서는 32비트의 확장형 RTP 패킷 Sequence Counter를 사용하며 RTP Header의 Sequence Number를 통해 32비트 Sequence Counter의 하위 16비트를 전달하고 다음에 설명될 RTP Payload Header의 Extended Sequence Number를 통해 상위 16비트를 전달 - 패킷을 보낼 때마다 1씩 증가
Timestamp	- RTP 패킷의 첫 바이트가 송신 장치에서 처리된 시점을 나타내는 32비트 필드 - 카메라 등의 영상획득 장치의 경우 영상 프레임의 촬영 시간을 나타냄 - 영상 신호 변환기 등의 경우 입력 영상의 Alignment Point(SMPTE ST 2059-10에 정의) 시간을 나타냄 - 동일한 영상 프레임에 속하는 모든 패킷은 같은 Timestamp 값을 가짐
SSRC	- 데이터 송신 소스 ID를 위한 32비트 필드 - 각각의 발송 소스는 무작위 SSRC ID를 가지며, 중복 ID는 없음

• SMPTE ST 2110-20 RTP Payload Header

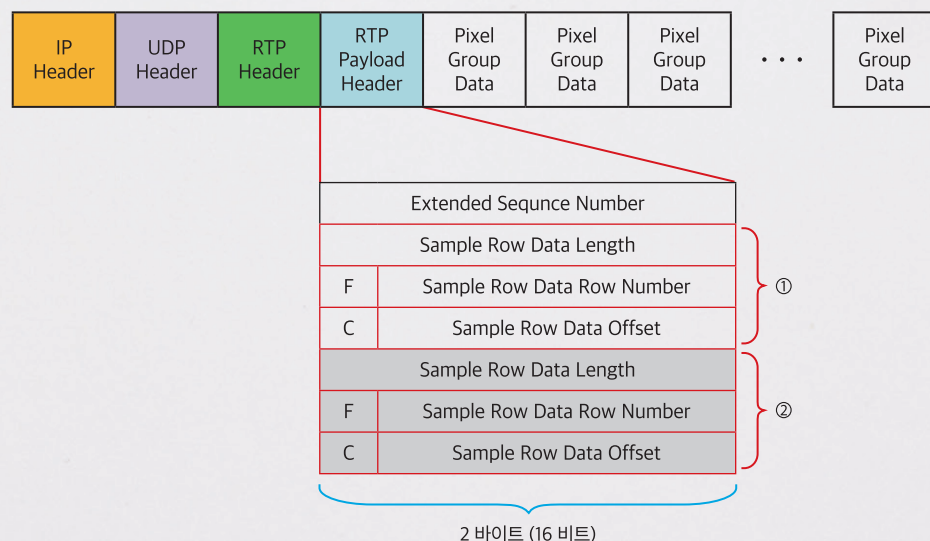


그림 3. SMPTE ST 2110-20 IP 패킷 RTP Payload Header 구조

Extended Sequence Number	- 2비트의 확장형 RTP 패킷 Sequence Counter에 사용되는 상위 16비트 - 하위 16비트는 앞서 설명된 RTP Header의 Sequence Number를 사용
Sample Row Data Length	- 패킷을 통해 전달되는 영상 데이터의 바이트 수를 나타내는 16비트 필드 - Pixel Data Group 사이즈의 배수가 되어야 함
F(Field Identification)	- 패킷을 통해 전달되는 영상의 필드 구분을 위한 1비트 필드 - 인터레이스 영상의 경우 패킷으로 전송되는 데이터가 첫 번째 필드이면 0, 두 번째 필드이면 1 - 프로그레시브 영상을 전송하는 경우 0, 단, Progressive segmented Frame (PsF)의 경우 F값은 세그먼트를 나타내기 위해 사용
Sample Row Data Number	- 패킷을 통해 전달되는 비디오 데이터의 행을 나타내는 15비트 필드 - 비디오의 첫 번째 행은 0부터 시작
C(Continuation)	- [그림 3]에 붉은색 외곽선으로 표시된 Sample Row Data Length부터 Sample Row Data Offset까지의 데이터 반복 여부를 표시하는 1비트 필드 - 이 값이 1이면 Sample Row Data Length ~ Sample Row Data Offset까지의 데이터가 뒤에 붙음 - 이 값이 0이면 현재의 Sample Row Data Length ~ Sample Row Data Offset이 마지막
Sample Row Data Offset	- Sample Row Data Length ~ Sample Row Data Offset에 대응되는 Pixel Group Data가 시작되는 비디오 행의 화소 위치를 나타내는 15비트 필드

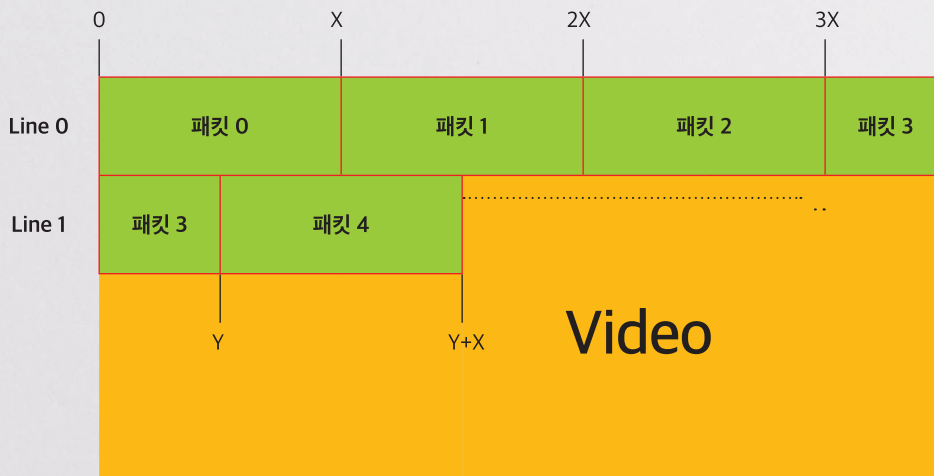


그림 4. 비디오와 Sample Row Data

RTP Payload Header의 경우 위 각 필드의 설명만으로 헤더 정보가 어떻게 비디오 데이터로부터 구성되는지 알기 힘들므로 [그림 4]를 통한 예를 통해 구체적인 설명을 이어가 보겠습니다. [그림 4]의 녹색 사각형은 영상 데이터와 해당 SMPTE ST 2110-20 패킷의 대응 관계를 보여줍니다. [그림 4]의 경우 RTP Payload Header의 Sample Row Data Length는 녹색으로 표시된 패킷에 대응하는 영상 화소의 색상 값들로부터 만들어진 Pixel Group Data의 바이트 수이며, Sample Row Data Length로부터 패킷에 포함된 Pixel Group Data의 데이터 사이즈를 알 수 있습니다. Sample Row Data Number는 패킷과 대응되는 영상 데이터가 속한 라인의 번호입니다. [그림 4]에 표시된 것과 같이 영상의 첫째 줄 데이터의 Sample Row Data Number는 0이고, 둘째 줄 데이터의 Sample Row Data Number는 1이며, 줄이 바뀔 때마다 1씩 증가합니다. Sample Row Data Offset은 각 패킷의 Pixel Group Data가 시작되는 화소의 가로방향 순서입니다. 대부분의 경우 패킷의 사이즈는 일정하게 유지되므로 [그림 4]와 같이 패킷 0의 Sample Row Data Offset은 0, 패킷 1의 Sample Row Data Offset은 X라고 하면 패킷 2의 Sample Row Data Offset은 2X가 되며, 이러한 패턴이 반복됩니다. 여기서 한 가지 주목할 점은 패킷 3과 같은 경우입니다. 패킷 3은 Line 0과 Line 1에 걸친 화소 값들로 만들어진 Pixel Group Data를 전송합니다. 그러므로 Line 0과 Line 1의 화소로부터 만들어진 Pixel Group Data의 정보를 각각 분리해서 RTP Payload Header에 실어야 하며, 이 경우 다른 패킷 0, 1, 2, 등은 [그림 3]의 ①번 블록만을 쓰지만 패킷 3은 ①번, ②번 블록을 모두 사용합니다. 이를 [표 1]에 정리하였습니다.

표 1. Sample Row Data 예

	[그림 3] ①번 블록	[그림 3] ②번 블록
패킷 0	Sample Row Data Length : P Sample Row Data Number : 0 Sample Row Data Offset : 0 F : 비디오 포맷에 따라 0 또는 1 C : 0	없음
패킷 1	Sample Row Data Length : P Sample Row Data Number : 0 Sample Row Data Offset : X F : 비디오 포맷에 따라 0 또는 1 C : 0	없음
패킷 2	Sample Row Data Length : P Sample Row Data Number : 0 Sample Row Data Offset : 2X F : 비디오 포맷에 따라 0 또는 1 C : 0	없음
패킷 3	Sample Row Data Length : p1 Sample Row Data Number : 0 Sample Row Data Offset : 3X F : 비디오 포맷에 따라 0 또는 1 C : 1	Sample Row Data Length : p2 Sample Row Data Number : 1 Sample Row Data Offset : 0 F : 비디오 포맷에 따라 0 또는 1 C : 0
패킷 4	Sample Row Data Length : P Sample Row Data Number : 1 Sample Row Data Offset : Y F : 비디오 포맷에 따라 0 또는 1 C : 0	없음

※ [표 1]에서 $p1 + p2 = P$, P는 패킷 0, 1, 2 등에 포함된 전체 Pixel Data Group들의 바이트 사이즈

그렇다면 얼마나 많은 패킷이 네트워크를 통해 전송되게 될까요? 이를 알아보기 위해 HD 비디오 제작 신호를 이용한 간단한 산수를 해보겠습니다.

HD 비디오 제작 신호

· 해상도 : 1920×1080 · 픽셀포맷 : 4:2:2 10bit · 프레임율 : 29.97

SMPTE ST 2110의 경우 네트워크에서의 효과적 전송을 위해 UDP 패킷의 최대 사이즈는 헤더를 포함하여 1460바이트로 제한하고 있습니다. 이 때문에 1460바이트를 이용하여 실제로 전송할 수 있는 UDP Payload 사이즈는 아래와 같습니다. RTP Payload Header를 14바이트로 계산한 것은 [그림 3]의 ①번 블록과 ②번 블록이 모두 다 있는 경우를 통해 Payload 용량이 줄어든 최악의 상황을 가정한 것입니다.

UDP Payload 사이즈

$1460 - 8(\text{UDP Header}) - 12(\text{RTP Header}) - 14(\text{RTP Payload Header}) = 1426\text{바이트}$

UDP로 전송 가능한 Payload 바이트 수인 1426을 4:2:2 10bit 포맷 영상의 Pixel Group Data 바이트 수인 5로 나누면 패킷당 전송할 수 있는 Pixel Group Data 개수를 알 수 있습니다.

패킷당 Pixel Group Data 개수

1426바이트 / 5바이트 $\approx$ 285개

※ Pixel Group Data 관련 내용은 이전 연재 참고

패킷당 전송 가능한 Pixel Group Data 개수인 285에 Pixel Group Data 당 화소 수인 2를 곱하면 패킷당 커버되는 화소 수를 구할 수 있습니다.

패킷당 커버되는 화소 수

$285\text{개} \times 2\text{개} = 570\text{개}$

HD 영상 프레임의 총 화소 수인 2,073,600(1920x1080)을 패킷당 커버되는 화소 수인 570으로 나누면 영상 프레임당 필요 패킷 수인 3,638을 구할 수 있습니다.

영상 프레임당 필요 패킷 수

$2,073,600\text{개} / 570\text{개} \approx 3,638\text{개}$

HD 영상은 1초에 29.97장의 영상 프레임이 포함되므로 HD 비디오 제작 신호의 초당 패킷 수는 109,031개가 됩니다.

HD 비디오 제작 신호의 초당 패킷 수

$3,638 \times 29.97 \approx 109,031\text{개}$

보시는 것처럼 HD 비디오 제작 신호 전송에 1초에 10만 개가 넘는 패킷이 네트워크를 통해 전송됩니다. UHD의 경우 해상도 4배, 프레임율 2배이므로 같은 4:2:2 10bit를 사용한다면 초당 $109,031 \times 8 = 872,248$ 개의 패킷이 네트워크를 통해 전송됩니다. 독자 여러분 예상보다 놀라운 숫자일 것 같습니다.

다음 연재에서는 현재 설명한 비디오 패킷의 네트워크상 동작 관련 표준인 SMPTE ST 2110-21에 대해 설명해 드리겠습니다. 