

## RTP Payload Format for BT.656 Video Encoding

### Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its Areas, and its Working Groups. Note that other groups may also distribute working documents as Internet Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced or obsoleted by other documents at any time. It is not appropriate to use Internet-Drafts as reference material or to cite them other than as ``work in progress.''

To learn the current status of any Internet-Draft, please check the ``[1id-abstracts.txt](#)'' listing contained in the Internet-Drafts Shadow Directories on [ftp.is.co.za](#) (Africa), [nic.nordu.net](#) (Europe), [munnari.oz.au](#) (Pacific Rim), [ds.internic.net](#) (US East Coast), or [ftp.isi.edu](#) (US West Coast).

Distribution of this document is unlimited.

### Abstract

This document specifies the RTP payload format for encapsulating ITU Recommendation BT.656-3 video streams in the Real-Time Transport Protocol (RTP). Each RTP packet contains all or a portion of one scan line as defined by ITU Recommendation BT.601-5, and includes fragmentation, decoding and positioning information.

### 1. Introduction

This document describes a scheme to packetize uncompressed, studio-quality video streams as defined by BT.656 for transport using RTP [[1](#)]. A BT.656 video stream is defined by ITU-R Recommendation BT.656-3 [[2](#)], as a means of interconnecting digital television equipment operating on the 525-line or 625-line standards, and complying with the 4:2:2 encoding parameters as defined in ITU-R

Internet-Draft

[draft-tynan-rtp-bt656-02.txt](#)

March 10th, 1998

Recommendation BT.601-5 (formerly CCIR-601) [3], Part A.

RTP is defined by the Internet Engineering Task Force (IETF) to provide end-to-end network transport functions suitable for applications transmitting real-time data over multicast or unicast network services. The complete specification of RTP for a particular application requires the RTP protocol document [1], a profile specification document [4], and a payload format specification. This document is intended to serve as the payload format specification for studio-quality video streams.

## [2.](#) Definitions

For the purposes of this document, the following definitions apply:

Y: An 8-bit or 10-bit coded luminance sample (as per BT.601-5 [3]). Each luminance value has 220 quantization levels with the black level corresponding to level 16 and the peak white level corresponding to 235.

Cb, Cr: An 8-bit or 10-bit coded color-difference sample (as per BT.601-5). Each color-difference value has 225 quantization levels in the centre part of the quantization scale with a color-difference of zero having an encoded value of 128.

True Black: BT.601-5 defines a true black level as the quad-sample sequence 0x80, 0x10, 0x80, 0x10, representing color-difference values of 128 (0x80) and a luminance value of 16 (0x10).

SAV, EAV: Video timing reference codes which appear at the start and end of a BT.656 scan line.

## [3.](#) Payload Design

ITU Recommendation BT.656-3 defines a schema for the digital interconnection of television video signals in conjunction with BT.601-5 which defines the digital representation of the original analog signal. While BT.601-5 refers to images with or without color subsampling, the interconnection standard (BT.656-3) specifically requires 4:2:2 subsampling. This specification also requires 4:2:2 subsampling such that the luminance stream occupies twice the bandwidth of each of the two color-difference streams. For normal 4:3 aspect ratio images, this results in 720 luminance samples per

scan line, and 360 samples of each of the two chrominance channels. The total number of samples per scan line in this case is 1440. While this payload format specification can accomodate various image sizes and frame rates, only those in accordance with BT.601-5 are currently supported.

Due to the lack of any form of video compression within the payload and sampling-rate compliance with BT.601-5, the resultant video stream can be considered ``studio quality''. However, such a stream can require approximately 20 megabytes per second of network bandwidth. In order to maximize packet size within a given MTU, and to optimize scan line decoding, each video scan line is encoded within one or more RTP packets.

To allow for scan line synchronization, each packet includes certain flag bits (as defined in BT.656-3) and a unique scan line number. The SAV and EAV timing reference codes are removed. Furthermore, no line blanking samples are included, so no ancillary data can be included in the line blanking period. It is the responsibility of the receiver to generate the timing reference codes, and to insert the correct number of line blanking samples.

Similarly, there is no requirement that the frame blanking samples be provided. However, it is possible to include frame blanking samples if such samples contain relevant information, such as a vertical-interlace time code (VITC), or teletext data. In the absence of frame blanking samples, the receiver must generate true black levels as defined above, to complete the correct number of scan lines per field. If frame blanking samples are provided, they must be copied without modification into the resultant BT.656-3 stream.

Scan lines must be sent in sequential order. Error concealment for missing scan lines or fragments of scan lines is at the discretion of the receiver.

Both 8-bit and 10-bit quantization types as defined by BT.601-5 are supported. 10-bit samples are considered to have two extra bits of fixed-point precision such that a binary value of 10111110.11 represents a sample value of 190.75. Using 8-bit quantization, this would give a sample value of 190. An application receiving 8-bit samples for a 10-bit device must consider the sample as reflecting



Each RTP packet starts with a fixed RTP header. The following fields of the RTP fixed header are used for BT.656-3 encapsulation:

Marker bit (M): The Marker bit of the RTP header is set to 1 when the last packet of a frame (or the last fragment of the last scan line if fragmented), and set to 0 on all other packets.

**Payload Type (PT):** The Payload Type indicates the use of the payload format defined in this document. A profile may assign a payload type value for this format either statically or dynamically as described in [4].

Timestamp: The RTP Timestamp encodes the sampling instant of the video frame currently being rendered. All scan line packets within the same frame will have the same timestamp. The timestamp should refer to the '0v' field synchronization point of the first field. For the payload format defined by this document, the RTP timestamp is based on a 90kHz clock.

## 5. Payload Header

The payload header is a fixed four-octet header broken down as follows:

[illegible]

F: 1 bit

When 0, indicates the first field of a frame (line 4 through 265 inclusive for Type=0 or 2, and line 1 through 312 inclusive for Type=1 or 3). Any other scan line is considered a component of the second field, and the F bit will be set to 1. This bit is copied directly from the BT.656-compliant stream by the transmitter, and directly inserted into the stream by the receiver.

V: 1 bit

When 1, indicates that the scan line is part of the vertical

interval. Should always be 0 unless frame blanking data is sent. In which case, the V bit should be set to 1 for scan lines which do not form an integral part of the image. This bit is copied directly from the BT.656-compliant stream by the transmitter, and directly inserted into the stream by the receiver. For receivers which do not receive scan lines during the vertical interval, BT.656 vertical interval data must be generated by repeating the quad-sample sequence 0x80, 0x10, 0x80, 0x10, representing a true black level.

Type: 4 bits

This field indicates the type of frame encoding within the payload. It must remain unchanged for all scan lines within the same frame. Currently only four types of encoding are defined. These are as follows;

- 0 - NTSC (13.5MHz sample rate; 720 samples per line; 60 fields per second; 525 lines per frame)
- 1 - PAL (13.5MHz sample rate; 720 samples per line; 50 fields per second; 625 lines per frame)
- 2 - High Definition NTSC (18MHz sample rate; 1144 samples per line; 60 fields per second; 525 lines per frame)
- 3 - High Definition PAL (18MHz sample rate; 1152 samples per line; 50 fields per second; 625 lines per frame)

Tynan

[Page 5]

---

Internet-Draft

[draft-tynan-rtp-bt656-02.txt](#)

March 10th, 1998

P: 1 bit

Indicates the required sample quantization size. When 0, the payload is comprised of 8-bit samples. Otherwise, it carries 10-bit samples. This bit must remain unchanged for all scan lines within the same frame.

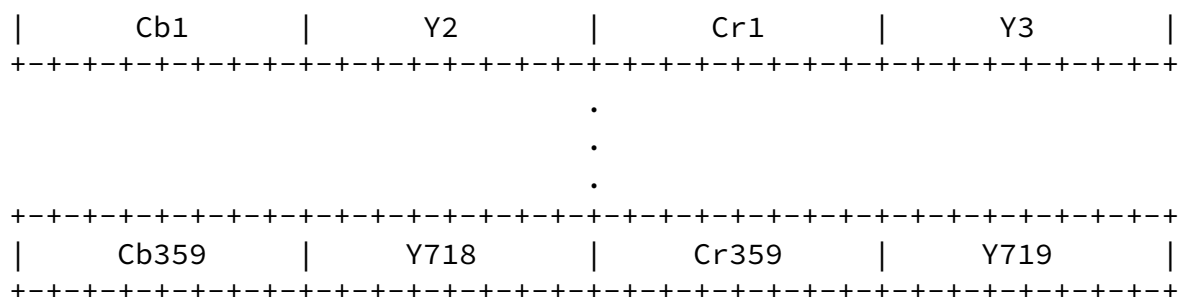
Z: 2 bits

Reserved for future use. Must be set to zero by the transmitter and ignored by the receiver.

Scan Line (SL): 12 bits

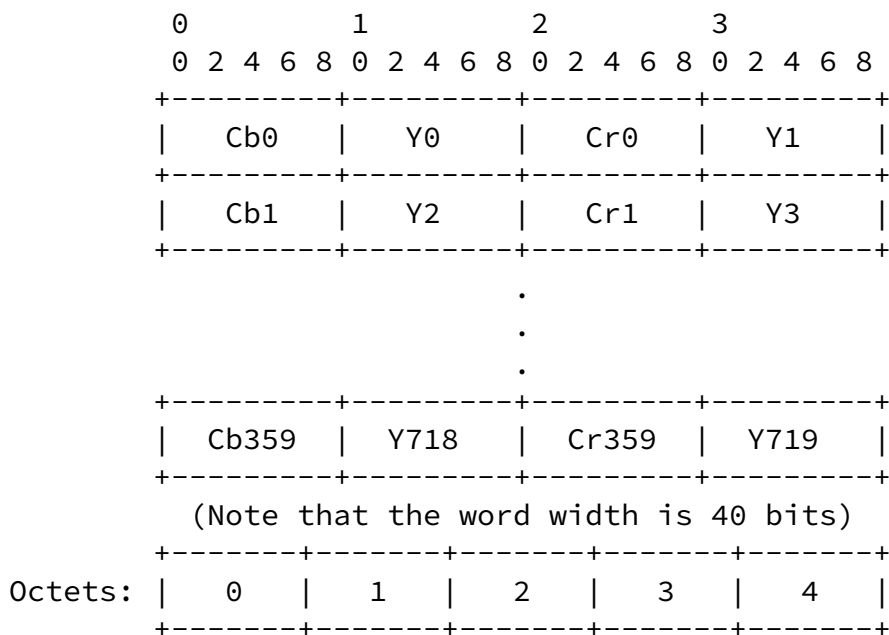
Indicates the scan line encapsulated in the payload. Valid values range from 1 through 625 inclusive. If no frame blanking data is being transmitted, only scan lines 23 through 310 inclusive, and





1144 and 1152 sample packets should increase the packet size accordingly while maintaining the sample order.

For 10-bit quantization, each group of four samples must be encoded into a 40-bit word (five octets) prior to transmission. The sample order is identical to that for 8-bit quantization. The following is a representation of a 720 sample packet with 10-bit quantization:



The octets shown in these diagrams are transmitted in network byte order, that is, left-to-right as shown.



## [7.](#) Security Considerations

RTP packets using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification [[1](#)]. This implies that confidentiality of the media streams is achieved by encryption. Because the payload format is arranged end-to-end, encryption may be performed after encapsulation so there is no conflict between the two operations.

This payload type does not exhibit any significant non-uniformity in the receiver side computational complexity for packet processing to cause a potential denial-of-service threat.

## [8.](#) References

- [1] RTP: A Transport Protocol for Real-Time Applications, H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, [RFC 1889](#).
- [2] Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems operating at the 4:2:2 Level of Recommendation ITU-R BT.601 (Part A), ITU-R Recommendation BT.656-3, 1995.
- [3] Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios, ITU-R Recommendation BT.601-5, 1995.
- [4] RTP Profile for Audio and Video Conference with Minimal Control, H. Schulzrinne, [RFC 1890](#).
- [5] Path MTU Discovery, J. Mogul, S. Deering, [RFC 1191](#).

## [9.](#) Authors Address

Dermot Tynan  
Claddagh Films Limited  
3 White Oaks  
Clybaun Road  
Galway  
Ireland

Email: [dtynan@claddagh.ie](mailto:dtynan@claddagh.ie)  
Tel: +353 91 529944