

MPEG-1 Data Structures

The **ISO/IEC 11172** specification defines the audio, video and multiplexing standards collectively and colloquially referred to as the MPEG-1 (**Motion Picture Experts Group**) compression standard. The data structures for the various components in an encoded bitstream are given in a pseudo-C syntax, and are extensively discussed. However, it is difficult to get the big picture from reading the spec.

More practically, in order to parse an MPEG-1 bitstream, it is necessary to know byte offsets within each structure. To make this information more readily accessible, we have condensed it into graphic form. Of course, this is no substitute for the original spec. Where more information is required than can be squeezed into the diagram, references are provided to the spec.

The Big Picture

A multiplexed MPEG-1 stream is composed of distinct **Packs**.

- Each Pack consists of a Pack header and any number of Packets.
- Within those Packets is either video or audio data.

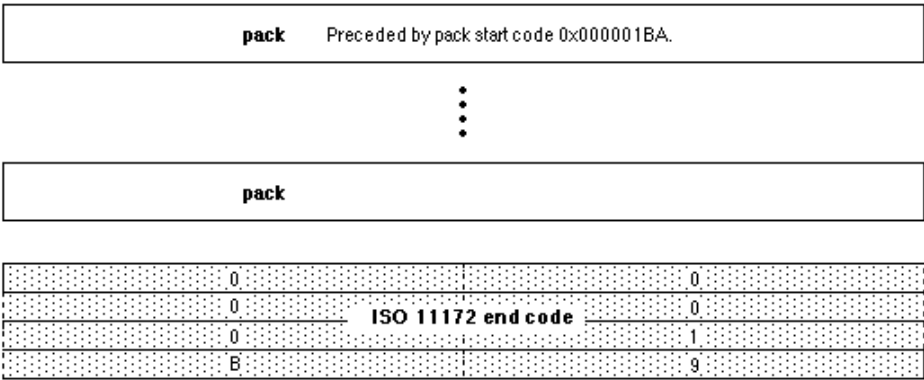
These structures above the video or audio level are called the system layer.

Video or audio data is divided into **Packets** without regard to lower-level structures -- Groups, Pictures, etc. may break across Packet boundaries.

- Video information is composed of individual Pictures. We will not discuss the substructures of Pictures.
- Pictures themselves are of three types:
 - I (intra) Pictures are self-contained, compressing the image using **Discrete Cosine Transform** (DCT) processing.
 - P (predictive) Pictures use less data and are predicted from the preceding I Picture.
 - B (bidirectional) Pictures use the least data and are interpolated using information from surrounding P and I Pictures
- Pictures are organized into Groups of (typically) 15 or so Pictures. If a Group is preceded by a Sequence header, its first Picture is called an entrypoint.

Audio information is composed of Frames. We will not discuss the substructure of Frames. There are no higher-level audio structures.

MPEG-1 Multiplexed Stream



©1994 Philips Interactive Media. All rights reserved.

MPEG-1 Pack

byte	bit 7	6	5	4	3	2	1	0
0	0				0			
1	0				pack start code		0	
2	0				1			
3	B				A			
4	always set to 0010				system clock reference bits 32..30			marker always set
5	SCR bits 29..15 Intended time, in 90 kHz clock cycles, of arrival of byte 8 of this header.							
6	(cont.)							marker always set
7	SCR bits 14..0							
8	(cont.)							marker always set
9	marker always set	multiplex rate (This #) * 400 bits/sec is the rate at which this stream ...						
10	...is to be delivered to the decoder.							
11	(cont.)							marker always set

(Optional)

system header	Preceded by system header start code 0x000001BB.
----------------------	--

packets	Arbitrary number, each preceded by packet start code prefix 0x000001.
----------------	---

© 1994 Philips Interactive Media. All rights reserved.

MPEG-1 System Header

byte	bit	7	6	5	4	3	2	1	0	
0			0					0		
1			0	system header start code				0		
2			0					1		
3			B					B		
4		header length				Bytes in header, from byte 6.				
5		(cont.)								
6	marker always set	rate bound				Max value of the multiplex rates of all packs in the stream.				
7		(cont.)								
8		(cont.)							marker always set	
9	audio bound						Max number of audio streams in this ISO stream.		fixed flag Set for fixed bitrate.	CSPS flag Set for constrained.
10	system audio lock flag †	system video lock flag †	marker always set	video bound						Max number of video streams in this ISO stream.
11	reserved									Currently should be set to FF.

† Set if there is, for all (audio/video) elementary streams in the ISO multiplexed stream, a constant rational relationship between the (audio sampling rate/video picture rate) and the system clock frequency in the decoder. See §2.4.4.2.

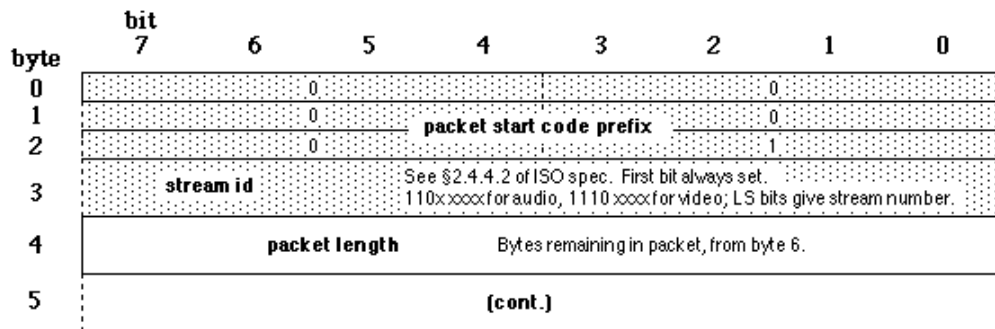
The following stream specs are present only if the first bit is a 1. Any number of stream specs may follow.

12	stream id		See §2.4.4.2 of ISO spec. First bit always set. 110xxxxx for audio, 1110xxxx for video; LS bits give stream number.	
13	Always set to 11.	STD buffer bound scale *	STD buffer size bound	Largest required buffer over all ...
14	... packets in this stream. If bound scale bit is cleared, units are 128 bytes; if set, 1024 bytes.			

* Cleared for audio stream, set for video; varies for others.

© 1994 Philips Interactive Media. All rights reserved.

MPEG-1 Packet



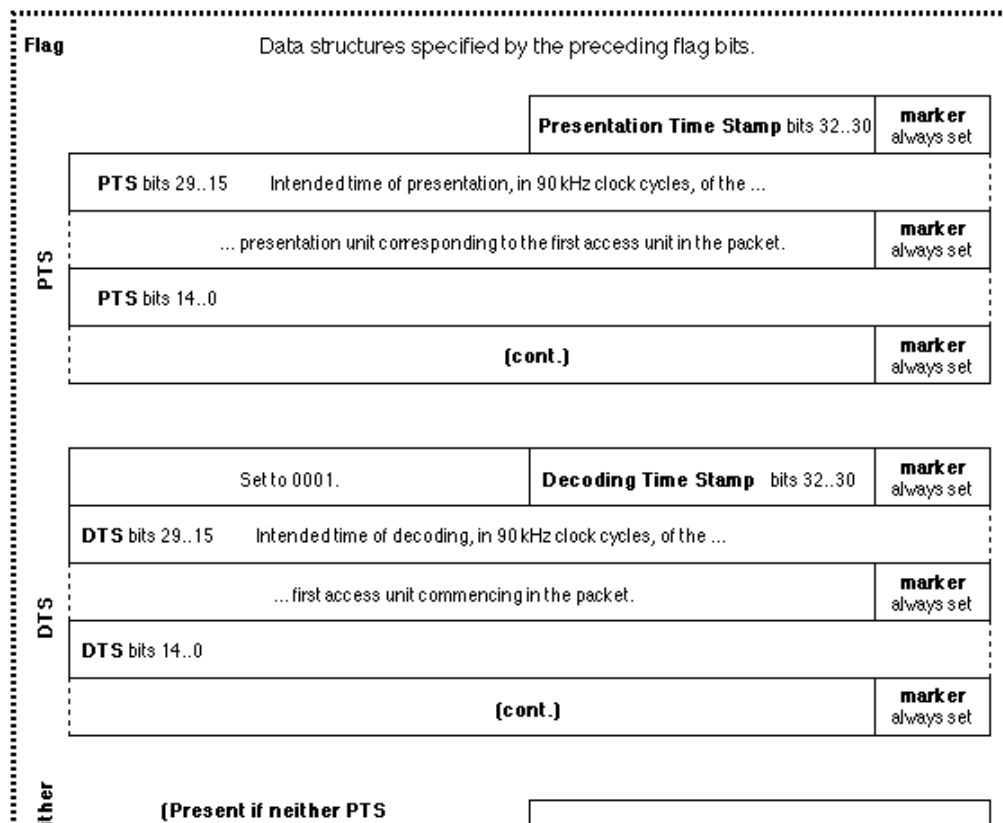
(Optional)

stuffing byte	Up to 16 of these, if needed for storage medium reasons. Set to FF.
----------------------	---

Present only if first two bits are 01. * Cleared for audio stream, set for video; varies for others.

Set to 01.	STD buffer scale #	STD buffer size	Specifies buffer size in target decoder.
If buffer scale bit is cleared, units are 128 bytes; if set, units are 1024 bytes.			

Always set to 00.	PTS 01 not allowed.	DTS
PTS/DTS flags		



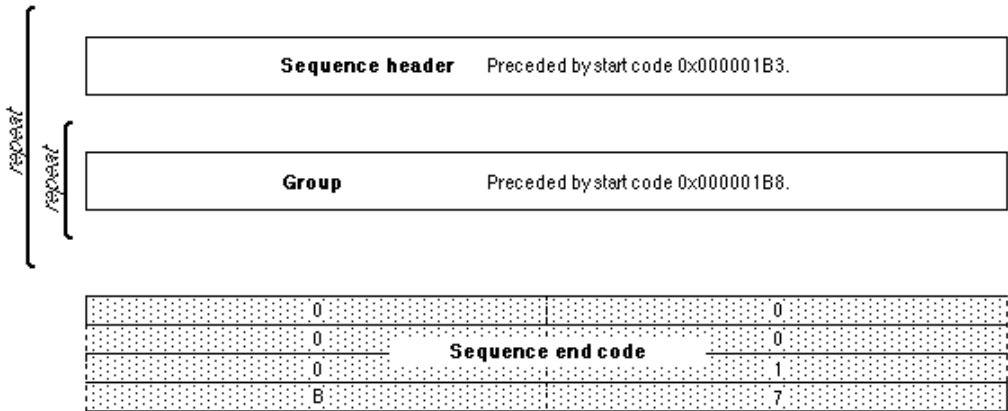
ne

or DTS flags were set.)

Set to 0xF.

MPEG-1 Video Sequence

(Skip to next start code)



©1994 Philips Interactive Media. All rights reserved.

MPEG-1 Sequence Header

byte	bit 7	6	5	4	3	2	1	0
0	0				0			
1	0				Sequence header code		0	
2	0						1	
3	B						3	
4	horizontal size Width in pixels of the displayable luminance picture. The encoded width...							
5	... in macroblocks is (displayable + 15) div 16.				vertical size			
6	(cont.)							
7	pixel aspect ratio Respectively: forbidden, 1.0, 0.6735, 0.7031, 0.7615, 0.8055, 0.8437, 0.8935, 0.9157, 0.9815, 1.0255, 1.0695, 1.095, 1.1575, 1.2051, rsvd.				picture rate Respectively: forbidden, 23.976, 24, 25, 29.97, 30, 50, 59.94, 60, rsvd ... rsvd.			
8	bit rate Stream bitrate in units of 400 bits/s. Zero is forbidden; 3FFFF means variable bitrate.							
9	(cont.)							
10	(cont.)		marker Always set.	VBV buffer size		16 * 1024 * (this #) is minimum VBV...		
11	... size in bits required to decode the sequence.					constrained Set if true.	load intra Q matrix	

		(If load intra Q matrix is set)		
		64 8-bit unsigned integers defining the intra quantizer matrix.		
12 -				
75		:		
				load non-intra Q matrix
		(If load non-intra Q matrix is set)		
		64 8-bit unsigned integers defining the non-intra quantizer matrix.		
76 -				
139		:		

(Skip to next start code)

Sequence extension data	Reserved for MPEG-2. Preceded by start code 0x000001B5. End is signaled by the presence of 0x0000001.
--------------------------------	---

User data (optional)	Preceded by start code 0x000001B2. End is signaled by the presence of 0x0000001.
-----------------------------	--

©1994 Philips Interactive Media. All rights reserved.

MPEG-1 Group of Pictures

byte	bit 7	6	5	4	3	2	1	0
0	0					0		
1	0		Group start code				0	
2	0						1	
3	B						8	
4	drop frame	hours (0-23)					minutes (0-59)	
5	(cont.)				marker always set	seconds (0-59)		
6	(cont.)			picture (0-59)				
7	(cont.)	closed gop *	broken link †	† Cleared during encoding. Set if editing has removed info needed to decode B-pictures after first I-picture of Group.				


* Set if the Group is encoded without prediction vectors pointing to the previous Group. A closed Group may more easily be edited after encoding.


(skip to next start code)

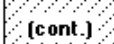
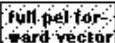
Group extension data	Reserved for MPEG-2. Preceded by start code 0x000001B5. End is signaled by the presence of 0x000001.
User data (optional)	Preceded by start code 0x000001B2. End is signaled by the presence of 0x000001.
Pictures	Arbitrary number, preceded by Picture start code 0x00000100.

MPEG-1 Picture

byte	bit	7	6	5	4	3	2	1	0
0	0					0			
1	0					0			
2	0					1			
3	0					0			
4	temporal reference Unsigned integer. Set to zero for first displayed Picture of Group, then incremented mod 1024.								
5	coding type Resp.: forbidden, I, P, B, D, rsvd...rsvd			YBY delay Time in 90 kHz ...					
6	... clock cycles needed to fill YBY buffer from empty state at target bitrate to correct level at start of play.								
7	For non-constant bitrate, is set to FFFF.								

 : For P- or B- pictures only.

 : For B-pictures only.

	full back-ward vector *	backward f code †		full pel forward vector *	forward f code †
---	--------------------------------	--------------------------	---	----------------------------------	-------------------------

★ Full vector flags are set if motion vector values decoded represent integer pixel offsets rather than half-pixels.

† F codes are unsigned non-zero integers describing decoding motion vectors as per §2.4.4.3, ISO 11172.

(Optional arbitrary number of 9-bit extra information structures.)	extra bit (set)	extra information
(Present only if extra bit is set. Always one byte long.)	...	extra bit (cleared)

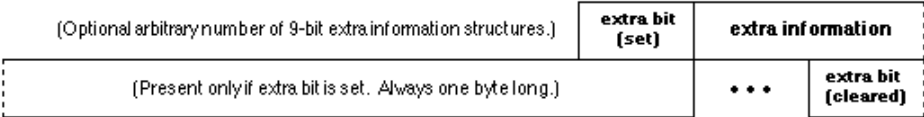
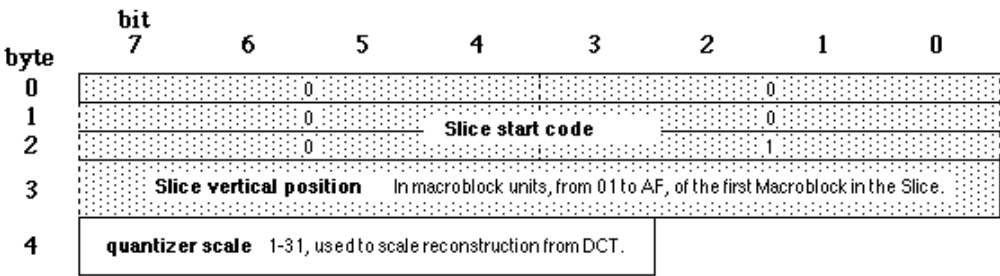
Extra information structures are terminated by a cleared extra bit.

(skip to next start code)

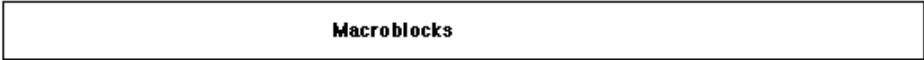
Sequence extension data	Reserved for MPEG-2. Preceded by start code 0x000001B5. End is signaled by the presence of 0x000001.
User data (optional)	Preceded by start code 0x000001B2. End is signaled by the presence of 0x000001.
Slices	Preceded by slice start codes 00000101 - 000001AF.

©1994 Philips Interactive Media. All rights reserved.

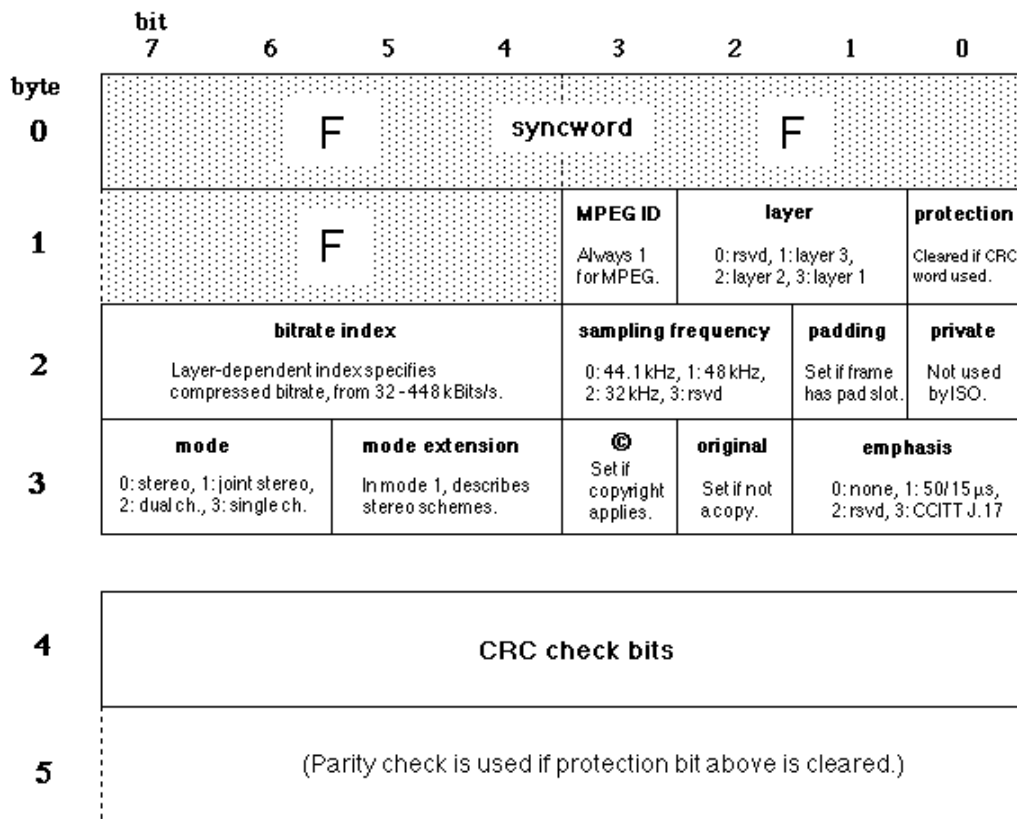
MPEG-1 Slice



Extra information structures are terminated by a cleared extra bit.



MPEG-1 Audio Frame Header



©1994 Philips Interactive Media. All rights reserved.