Media Format Analysis and Production

Specification

1. Revision Log

Revision	Date	Changes	Author
0.1	05/04/2010	Initial setup ---ASF, MPEG-1 SS, MPEG-2 PS, MPEG-2 TS, BDAV MPEG-2 TS, MP4, AVI, VC-1 Profile and Levels, H.264 Profile etc	Weifeng Lu
0.2	19/04/2010	Add MPEG-2 Profile and Levels, Elementary Stream(Video), Flash Video (FLV and F4V) , ASF and AVI structure, Modify Source duration etc	Weifeng Lu
0.3	01/05/2010	Change organization structure, Add PES Packet, TS Packet, PAT and PMT, Add I, P, B Frame, Add Chroma Subsampling, Add MKV, Add Priority etc	Weifeng Lu
0.4	14/05/2010	Add MPEG-2 TS (High Level), Relationship between ISO, QuickTime, MP4 and 3GPP&3GPP2, Add MPEG-2 PS Multiplexed, PS Pack etc	Weifeng Lu
0.5	13/07/2010	Add WebM, Ogg, RealMedia, ADTS header, WAV, HTML5 Video, MPEG-2 Video Structure, Elementary Stream(Audio) etc	Weifeng Lu

2. Overview

This document introduces various media containers systematic, relationship between containers and allowed video/audio standard (Codec) for corresponding container, which also provides some useful video production tools and defines media asset directory structure.

This document aims to be familiar with the various media containers and principal video standard---MPEG-2, MPEG-4(Part 2 and Part 10/AVC) and VC-1, get facility in the production of the media asset and reorganize and set up a complete media asset lib to make testing media more conveniently and efficiently.

If you spot any errors, Please contact me via email at lwfwind@gmail.com, I'll get them corrected as soon as possible.

I hope you find this useful and perhaps a tiny bit enjoyable.

3. Source Duration

- About 3 Minutes

Gathering some beautiful, funny or fresh short assets (such as MV) to let testing more active and enjoyable.

- About 1 Hour

The long duration assets are primarily used to test the synchronization of video and audio. SD content fetched from DVD and HD content fetched from BD.

4. Container Format

A container is, like the name says, a construct to contain data - video and audio date and possibly subtitles and navigational information. For instance, you would like to put a soundless video stream and the audio track together in one file. To do that you need a container format

Advanced container formats can support multiple audio and video streams, subtitles, chapter-information, and meta-data (tags) — along with the synchronization information needed to play back the various streams together.

The following table lists the most popular multi-media containers.

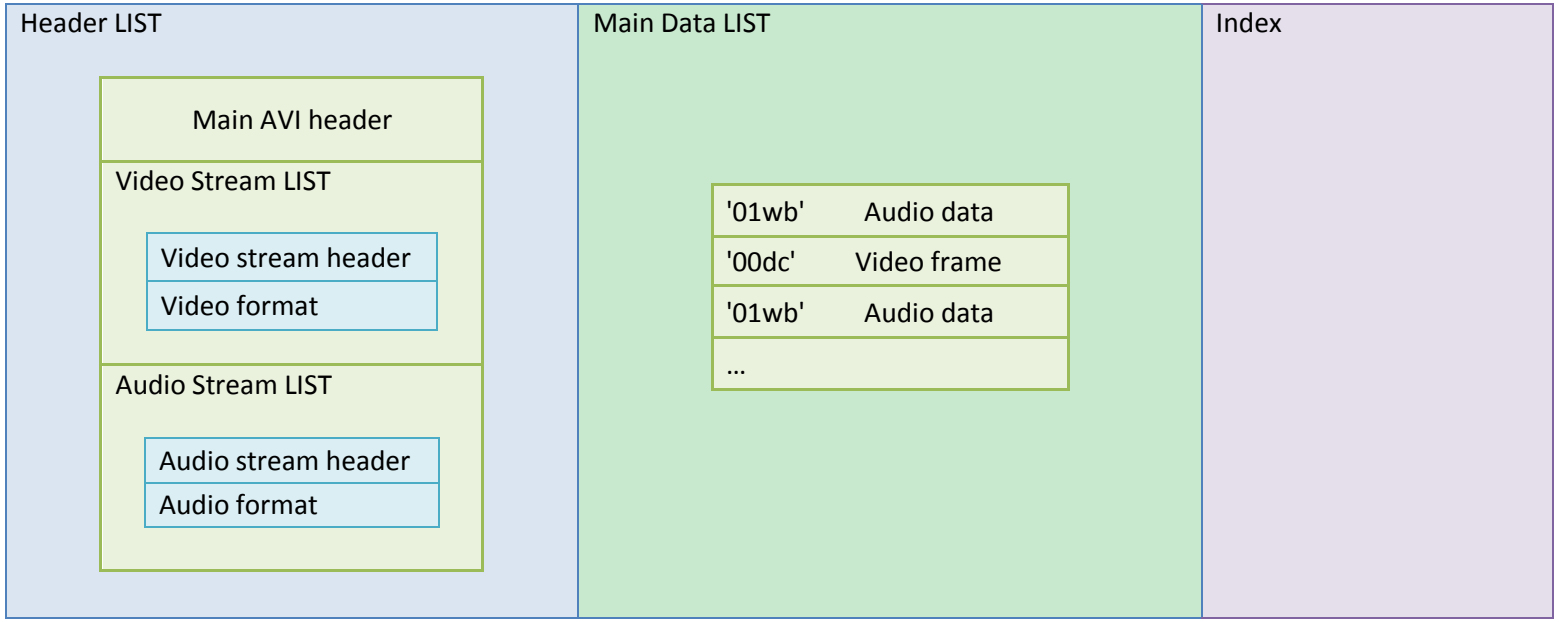
Container	Comments
AVI	The standard Microsoft Windows container, which is based on RIFF structure.
DivX Media Format	The standard for MPEG-4 Part 2 DivX codec from DivX Inc, which is based on AVI.
ASF	The standard container for Microsoft WMA and WMV.
DVR-MS	"Microsoft Digital Video Recording", proprietary video container format developed by Microsoft based on ASF.
WTV	The new container format used to record television shows in Microsoft Windows Vista Media Center starting with Windows Media Center TV Pack 2008, Which is successor of DVR-MS.
MPEG Program Stream	The standard container for MPEG-1 and MPEG-2 elementary streams on reasonably reliable media such as disks; which is also used on VCD and DVD.
MPEG-2 Transport Stream	The standard container for digital broadcasting and for transportation over unreliable media.
BDAV MPEG-2 Transport Stream	The standard container for the Blu-ray Disc Prerecorded application format (BD-ROM) and the Blu-ray Disc Recordable application format (BD-RE), which is based on MPEG-2 Transport Stream.
MPEG-4 Part 14 (MP4)	The standard audio and video container for the MPEG-4 multimedia portfolio, which is based on the ISO base media file format defined in MPEG-4 Part 12.
QuickTime	The standard QuickTime video container from Apple Inc.
3GP/3GP2	The container is used by many mobile phones, which is based on the ISO base media file format.
Flash Video	There are two different container formats (FLV and F4V) defined by Adobe Systems to deliver video over the Internet using Adobe Flash Player
Matroska (MKV)	Not standard for any codec or system, it can hold virtually anything and It is an open source container format.
WebM	The standard container from Google for Xiph.org audio codec Vorbis and video codec VP8.which is based on a subset of Matroska container format
Ogg	The standard container for Xiph.org audio codec Vorbis and video codec Theora
RealMedia	The standard container for RealVideo and RealAudio

4.1 Audio Video Interleave (AVI)

4.1.1. Container & supported video/audio standard (Codec)

Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments:
Audio Video Interleave (AVI)	Microsoft	.avi	Almost anything (Main Codec: Divx, Xvid, 3ivx, DV, Cinepak, Indeo, MJPEG, Uncompressed)	Almost anything	The AVI container was originally introduced by Microsoft in the early 90s, and was designed as a very flexible A/V container format for the video and audio compression formats. AVI cannot support many of the more advanced features that modern audio and video compression formats will offer, such as Variable Frame rate video encoding (VFR). Although AVI was extended with an additional standard called 'Open DML AVI' in the mid/late 90s, overcoming most of its very annoying limitations like the 2 GB file size limit, there is still no proper and spec compliant way to support modern compression formats like the "Advanced Audio Coding" (AAC), "Ogg Vorbis" or "Realvideo 9" (RV9).
Divx Media Format (DMF)	Divx, Inc.	.divx	MPEG-4 Part 2 video encoded to meet Divx profiles	MPEG-1 Audio Layer 3, LPCM, AC-3	Divx container is basically an enhanced AVI format (based on the same RIFF structure, for backward compatibility with existing players and devices).

4.1.2. Sample AVI structure



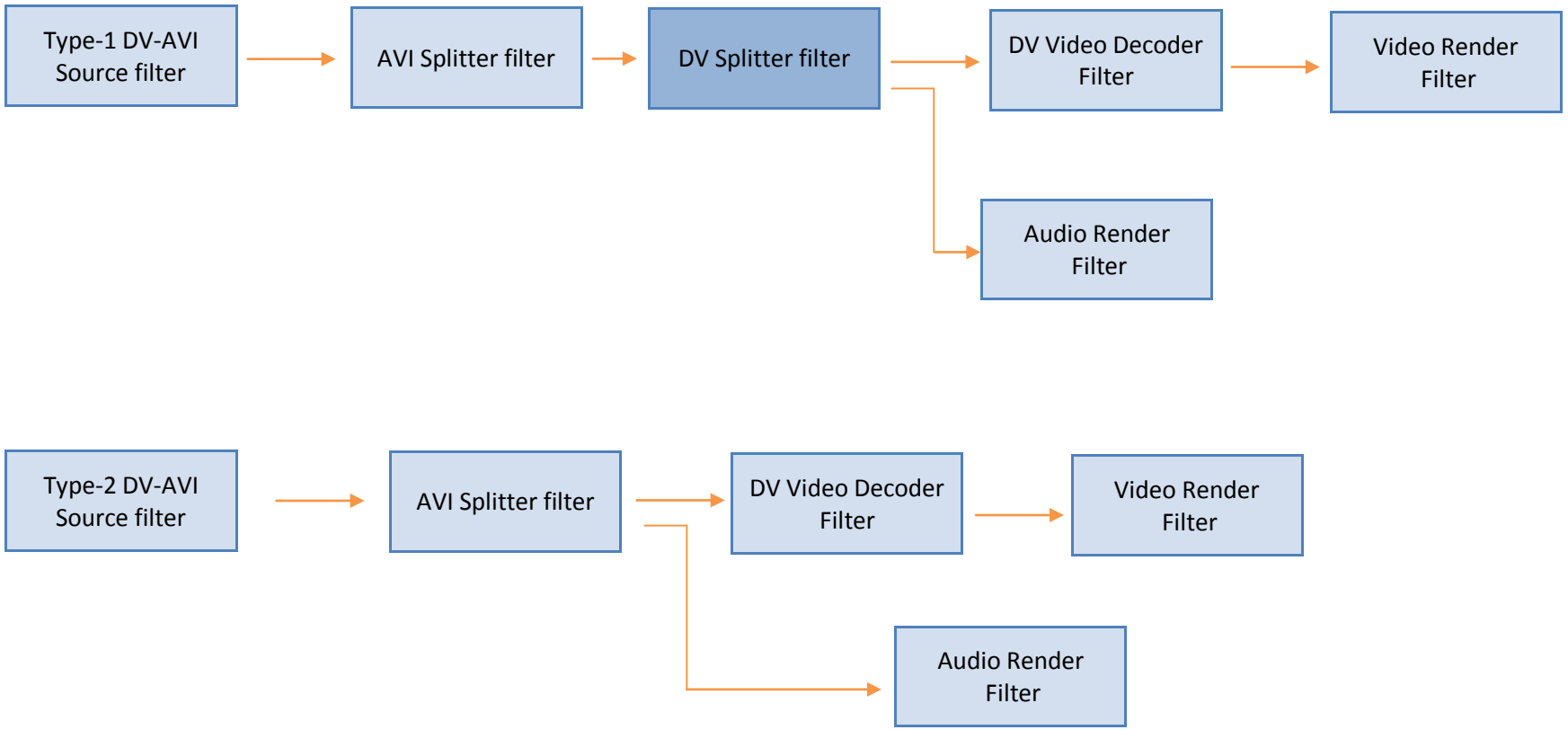
- All AVI files include two mandatory LIST chunks, which define the format of the streams and the stream data, respectively. An AVI file might also include an index chunk, which gives the location of the data chunks within the file.
- The Header LIST begins with the main AVI header, which contains global information for the entire AVI file, such as the number of streams within the file and the width and height of the AVI sequence. One or more Stream LISTs follow the main header, which contains information about one stream in the file.
- Following the header information is a Main Data LIST that contains the actual data in the streams — that is, the video frames and audio samples.
- An optional index chunk can follow the Main Data LIST. The index contains a list of the actual data chunks and their location in the file. Lose the index at the end, and your AVI file will become unseekable or even refuse to play in some players.

4.1.3. DV-AVI Type-1&Type-2 structure

```
RIFF 'AVI'
        LIST 'hdrl'
                avih <Main AVI Header>
                        dwStreams: 1
                        dwWidth: 720
                        dwHeight: 480
        LIST 'strl'
                strh <Stream header>
                        fccType: 'vids'
                        fccHandler: 'dvsd'
                        dwRate: 2997
                strf <Stream format>
        LIST 'strl'
                strh <Stream header>
                        fccType: 'auds'
                strf <Stream format>
                        wFormatTag: PCM
                        nChannels: 2
                        nSamplesPerSec: 32000
                        wBitsPerSample: 16
        LIST 'movi‘
                ##__
        idxl <AVI Index>
```

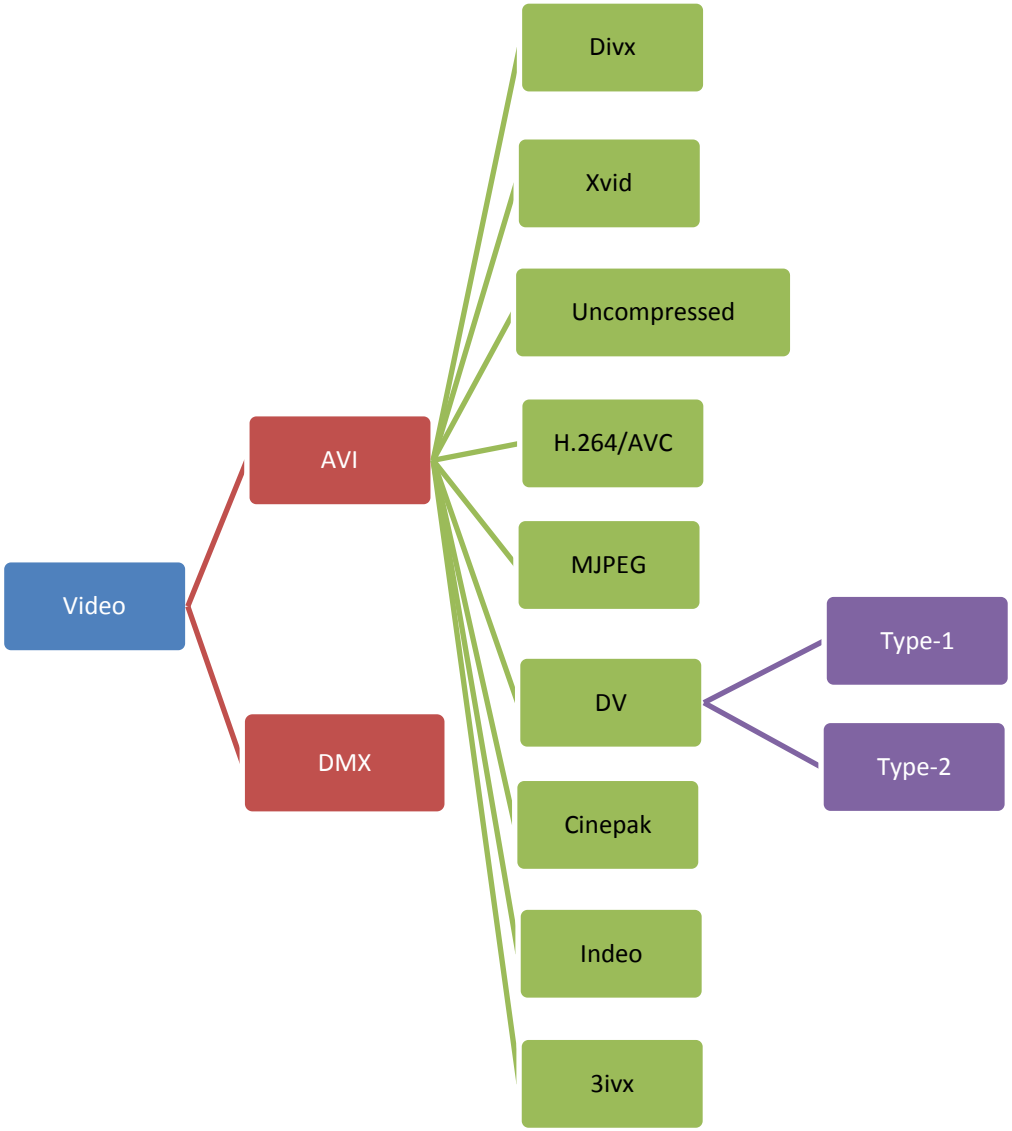
```
RIFF 'AVI'
        LIST 'hdrl'
                avih <Main AVI Header>
                        dwStreams: 1
                        dwWidth: 720
                        dwHeight: 480
                        ...

        LIST 'strl'
                strh <Stream header>
                        fccType: ‘iavs’ <interleaved audio and video
                        stream>
                        FCCHandler: 'dvsd'
                        dwRate: 2997
                        ...

                strf <Stream format>
                        …
        LIST 'movi‘
                ##dc <Compressed video frame>
                ##wb <Audio data>
                …
        idxl <AVI Index>
```

4.1.4. Playback DV-AVI using a DirectShow Filter graph.



4.1.5. Production Tools: VirtualDubMod, Graphedit, DivX Converter, DivXAuthor

4.1.6. Directory Structure

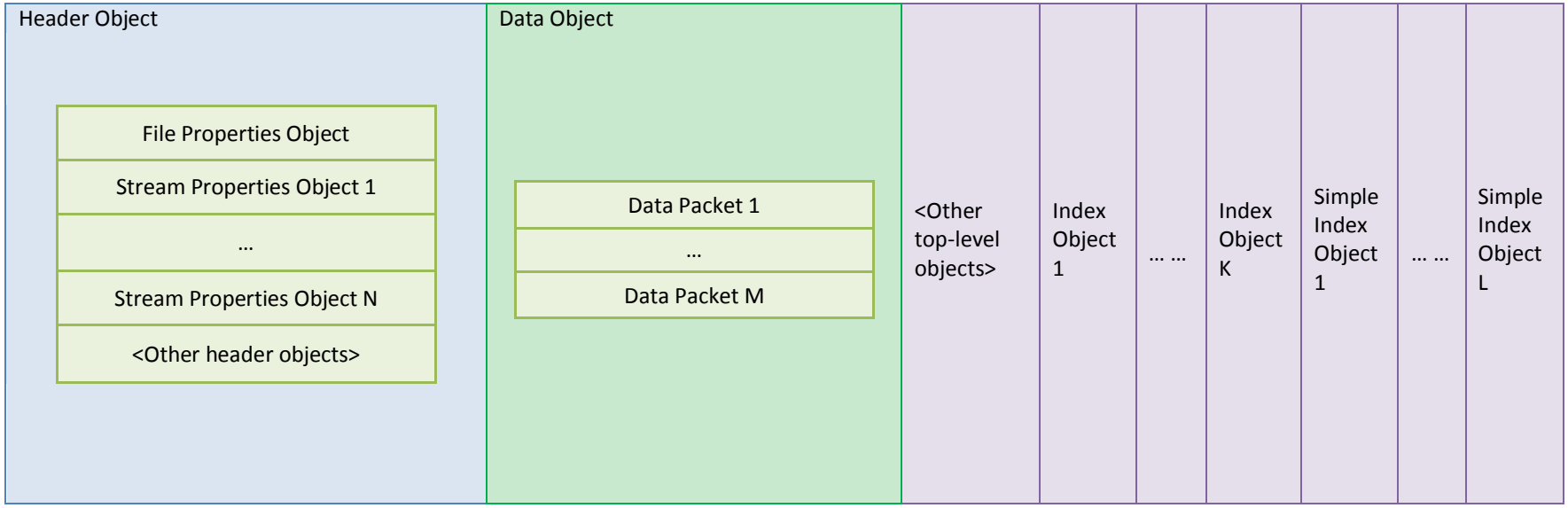

Video
AVI
DMX
Divx
Xvid
Uncompressed
H.264/AVC
MJPEG
DV
Cinepak
Indeo
3ivx
Type-1
Type-2

4.2 Advanced Systems Format (ASF)

4.2.1. Container & supported video/audio standard (Codec)

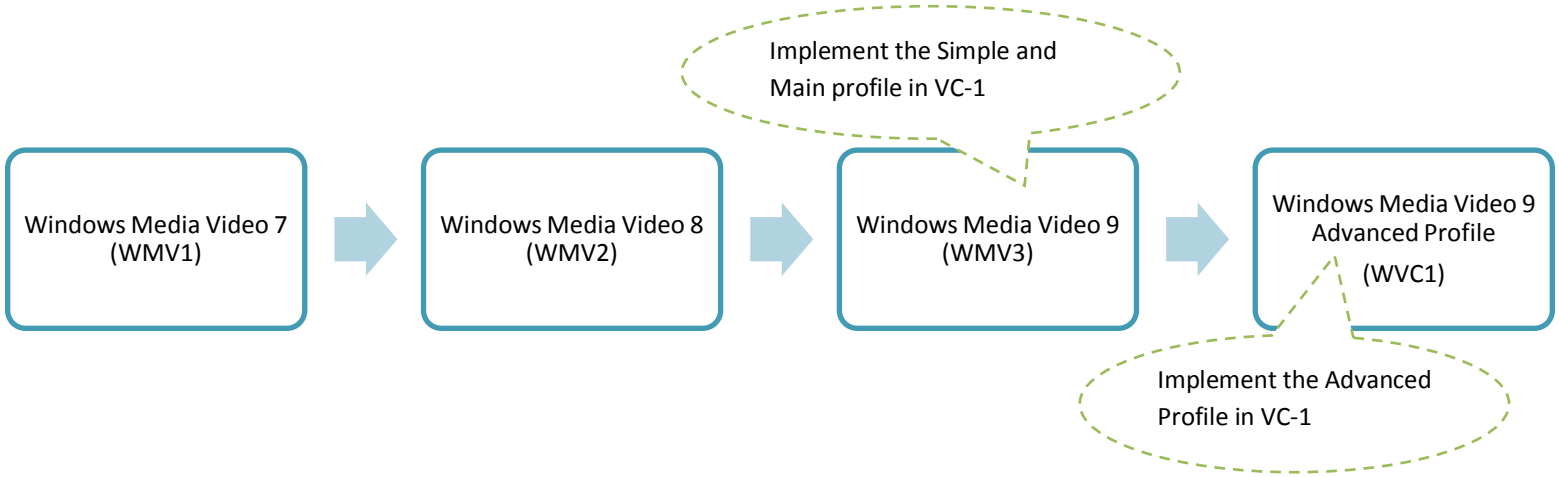
Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments:
Advanced Systems Format (ASF)	Microsoft	.wmv , .asf	Almost anything (Main Codec: WMV, MS MPEG-4 Series)	Almost anything (Main Codec: WMA Series)	The file extension .WMV typically describes ASF container files that use Windows Media Video codec. Microsoft recommends that ASF container files containing non-Windows Media codec use the generic .ASF file extension.
		.wma (A)		WMA	
DVR-MS	Microsoft	.dvr-ms	MPEG-2	MPEG-1 Layer II , AC3	DVR-MS is the format used to record television shows in Microsoft Windows XP Media Center Edition. The format is based on the ASF specification.
WTV	Microsoft	.wtv	MPEG-2	MPEG-1 Layer II , AC3	WTV file format replaces the previously used DVR-MS file format and used to record television shows in Microsoft Win7 Media Center. Unlike DVR-MS files, produced by previous versions of Media Center and based on ASF file format, WTV files are based on a new file format.

4.2.2. ASF Structure



ASF files are logically composed of three types of top-level objects: the Header Object, the Data Object, and the Index Object(s). The Header Object is mandatory and must be placed at the beginning of every ASF file. The Data Object is also mandatory and must follow the Header Object. The Index Object(s) are optional, but they are useful in providing time-based random access into ASF files.

4.2.3. WMV Codec Series

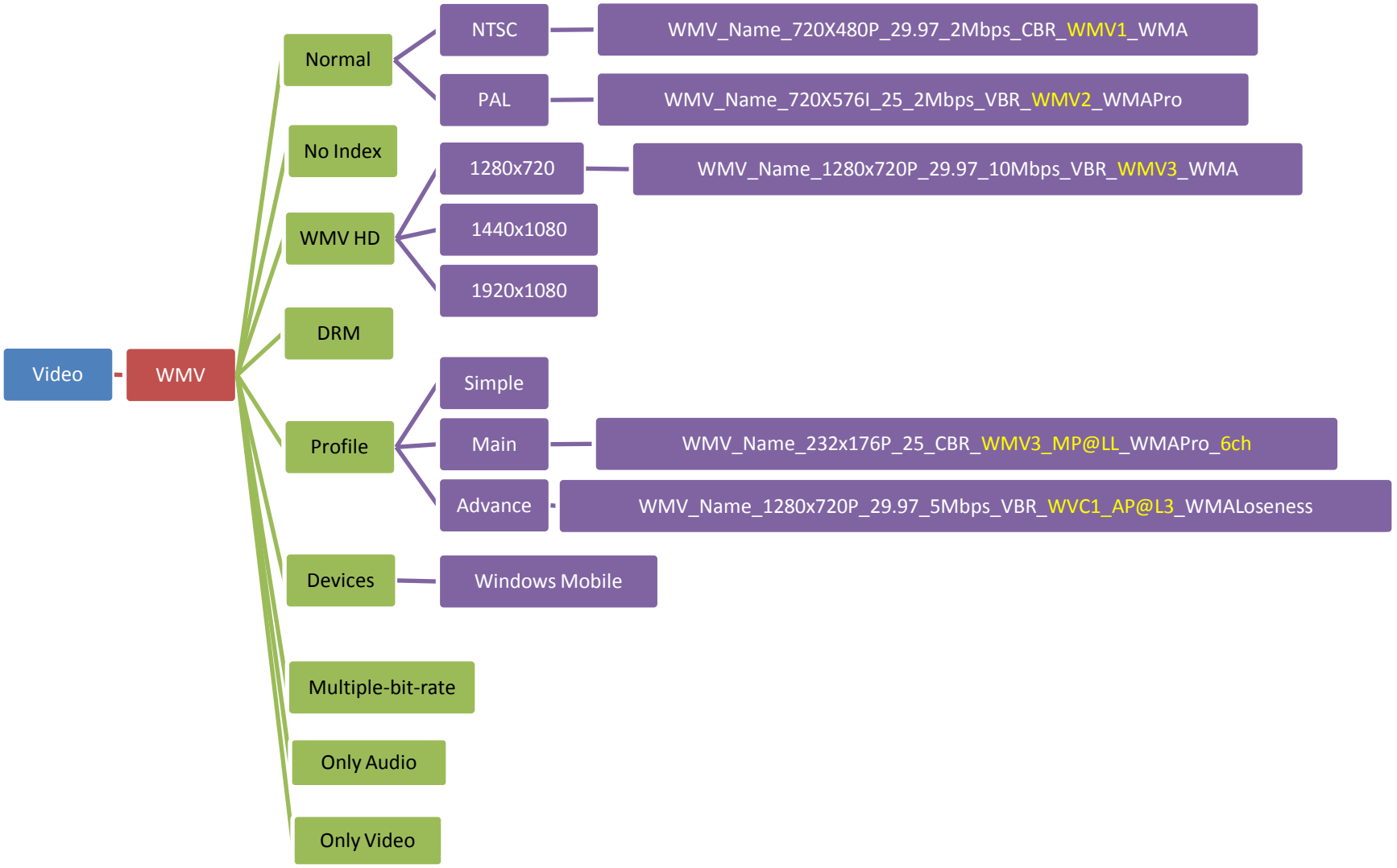


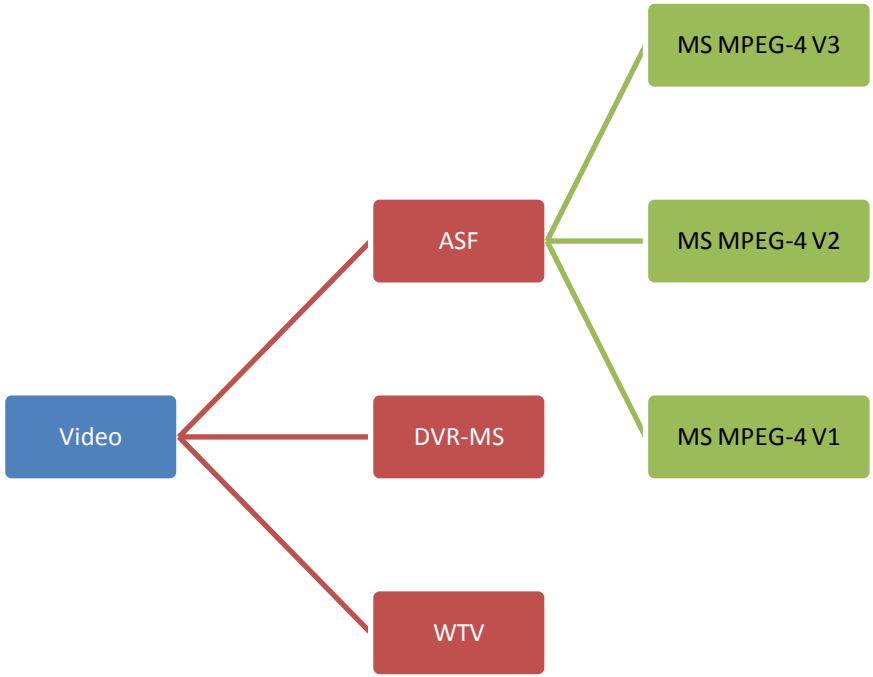
4.2.4. VC-1 Profile & Level

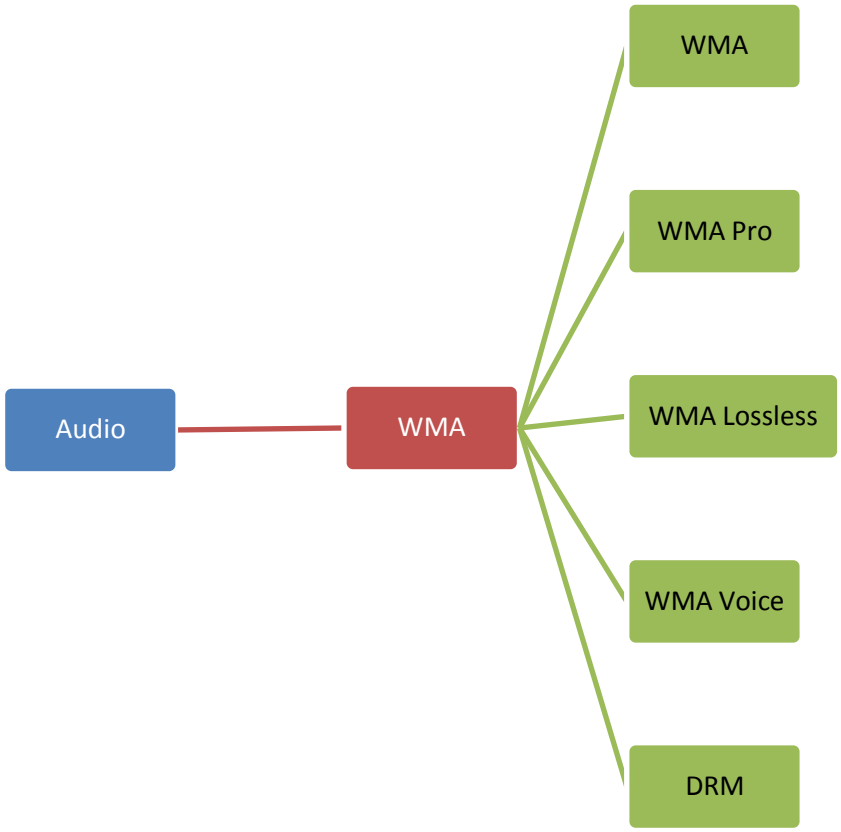
Profile	Level	Max Bit Rate	Representative Resolutions by Frame Rate
Simple	Low	96 Kbps	176 x 144 @ 15 Hz (QCIF)
	Medium	384 Kbps	240 x 176 @ 30 Hz 352 x 288 @ 15 Hz (CIF)
Main	Low	2 Mbps	320 x 240 @ 24 Hz (QVGA)
	Medium	10 Mbps	720 x 480 @ 30 Hz (480p) 720 x 576 @ 25 Hz (576p)
	High	20 Mbps	1920 x 1080 @ 30 Hz (1080p)
Advanced	L0	2 Mbps	352 x 288 @ 30 Hz (CIF)
	L1	10 Mbps	720 x 480 @ 30 Hz (NTSC-SD) 720 x 576 @ 25 Hz (PAL-SD)
	L2	20 Mbps	720 x 480 @ 60 Hz (480p) 1280 x 720 @ 30 Hz (720p)
	L3	45 Mbps	1920 x 1080 @ 24 Hz (1080p) 1920 x 1080 @ 30 Hz (1080i) 1280 x 720 @ 60 Hz (720p)
	L4	135 Mbps	1920 x 1080 @ 60 Hz (1080p) 2048 x 1536 @ 24 Hz

4.2.5. Production Tools: Microsoft Expression Encoder, Windows Media® Encoder 9 Series, Windows Media Tools.

4.2.6. Directory Structure


Video
WMV
Normal
NTSC
WMV_Name_720X480P_29.97_2Mbps_CBR_WMV1_WMA
PAL
WMV_Name_720X576I_25_2Mbps_VBR_WMV2_WMAPro
No Index
WMV HD
1280x720
WMV_Name_1280x720P_29.97_10Mbps_VBR_WMV3_WMA
1440x1080
1920x1080
DRM
Profile
Simple
Main
WMV_Name_232x176P_25_CBR_WMV3_MP@LL_WMAPro_6ch
Advance
WMV_Name_1280x720P_29.97_5Mbps_VBR_WVC1_AP@L3_WMALoseness
Devices
Windows Mobile
Multiple-bit-rate
Only Audio
Only Video





4.3 MPEG-1 System Stream (MPEG-1 SS)

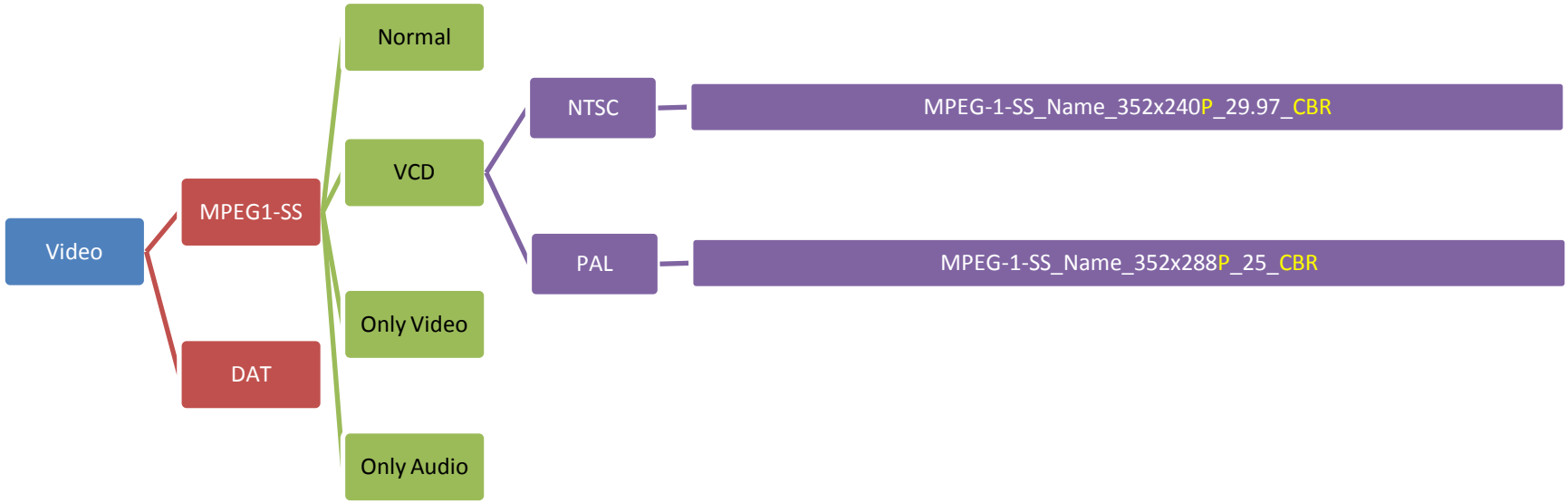
- MPEG-1 System Stream is defined in MPEG-1 Part 1 (ISO/IEC 11172-1).

4.3.1. Container & supported video/audio standard

Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments
MPEG-1 System Stream	MPEG	.mpeg, .mpg, .mpe	MPEG-1	MPEG-1 Layer I II	MPEG-1 System structure was later named a program stream: "The MPEG-1 Systems design is essentially identical to the MPEG-2 Program Stream structure." This terminology is more popular.
DAT	Philips	.dat (VCD)	MPEG-1	MPEG-1 Layer II	DAT files are basically MPEG-1 files with an additional information and certain specific file structure -- they are NOT "real" MPEG-1 files (their header data is slightly different when stored on VCD)

4.3.2. Production Tools: MainConcept Reference

4.3.3. Directory Structure


Video
MPEG1-SS
DAT
Normal
VCD
Only Video
Only Audio
NTSC
PAL
MPEG-1-SS_Name_352x240P_29.97_CBR
MPEG-1-SS_Name_352x288P_25_CBR

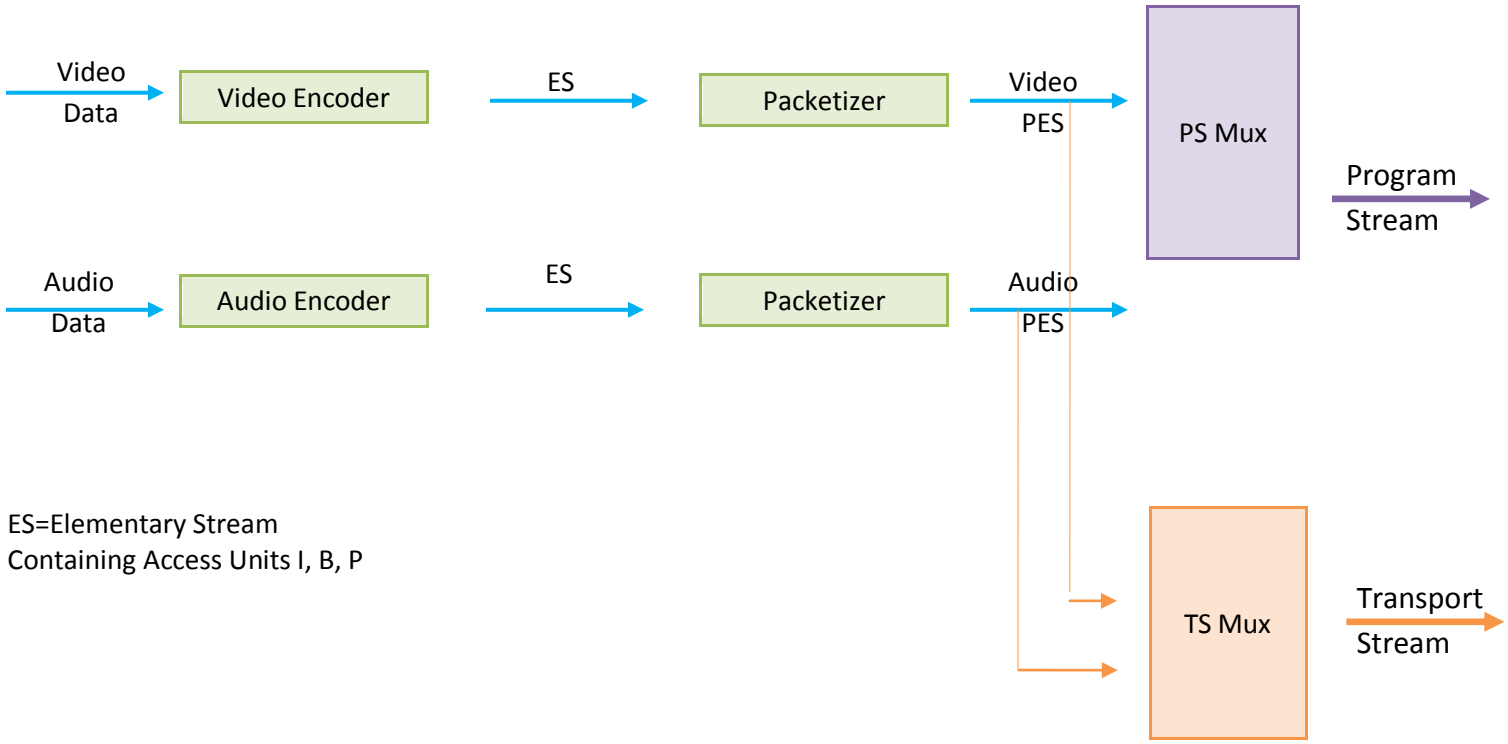
4.4 MPEG-2 Program Stream (MPEG-2 PS)

- MPEG-2 Program Stream is defined in MPEG-2 Part 1, Systems (ISO/IEC standard 13818-1/ITU-T H.222.0).

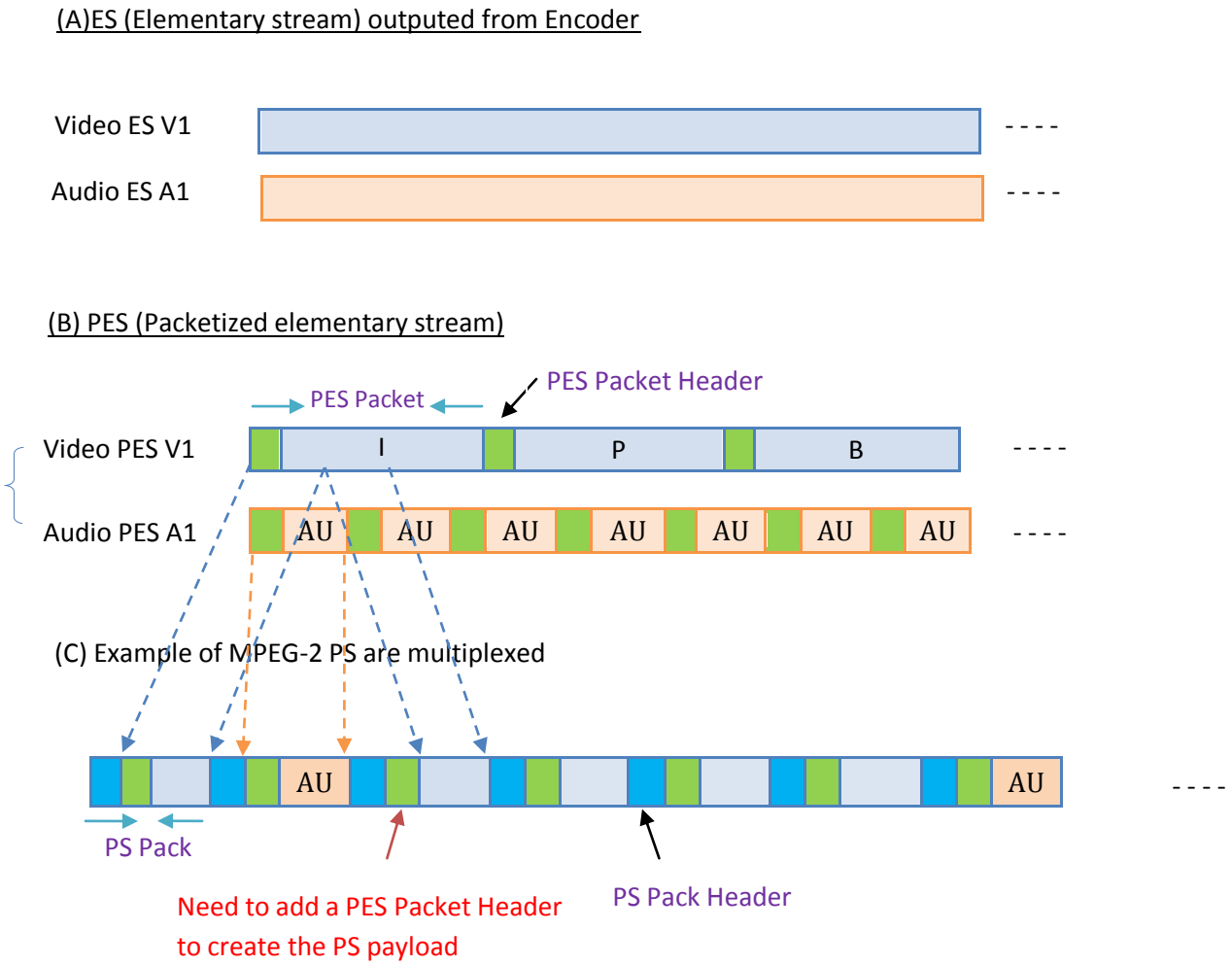
4.4.1. Container & supported video/audio standard

Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments
MPEG-2 Program Stream	MPEG	.mpeg, .mpg, .mpe, m2p	MPEG-2	LPCM, MPEG-1 Layer I II, AC3	1. MPEG-2 Program Stream is designed for reasonably reliable media such as disks, in contrast to MPEG Transport Stream which is for data transmission in which loss of data is likely. 2. An MPEG-2 Program Stream contains one, and only one, content channel. A Program Stream is like a single-passenger car without shock absorbers: drive it on a smooth road.
VOB	DVD Forum	.vob (DVD)	MPEG-1, MPEG-2	AC3, LPCM, MPEG-1 Layer II, DTS	1. A VOB file is an MPEG-2 program stream. This means that it complies 100% with the MPEG-2 system level standard. However, VOB files are a very strict subset of the standard. One good example is the restriction on the coded size of a picture: MPEG-2 Main Profile @ Main Level allows max coded frame size is 720x576 pixels, another example is the restriction on the bit rate: The maximum bit rate of 9.8 Mbps is more restrictive than MP@ML's 15 Mbps limit. So while all VOB files are MPEG-2 program streams, not all MPEG-2 system streams comply with the definition for a VOB file. 2. The DVD-Video format uses MPEG-2 video primarily, but MPEG-1 support is explicitly defined in the standard.

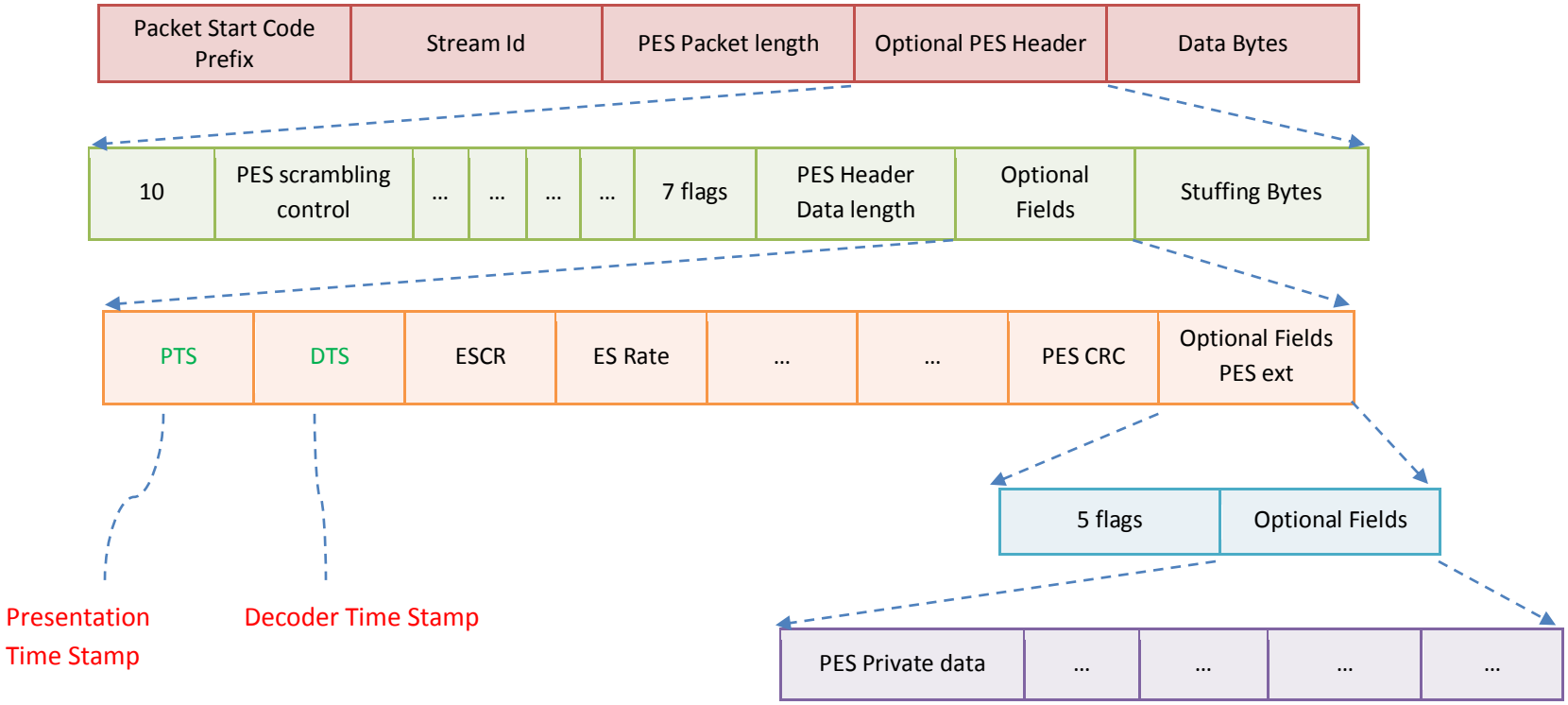
4.4.2. MPEG-2 System



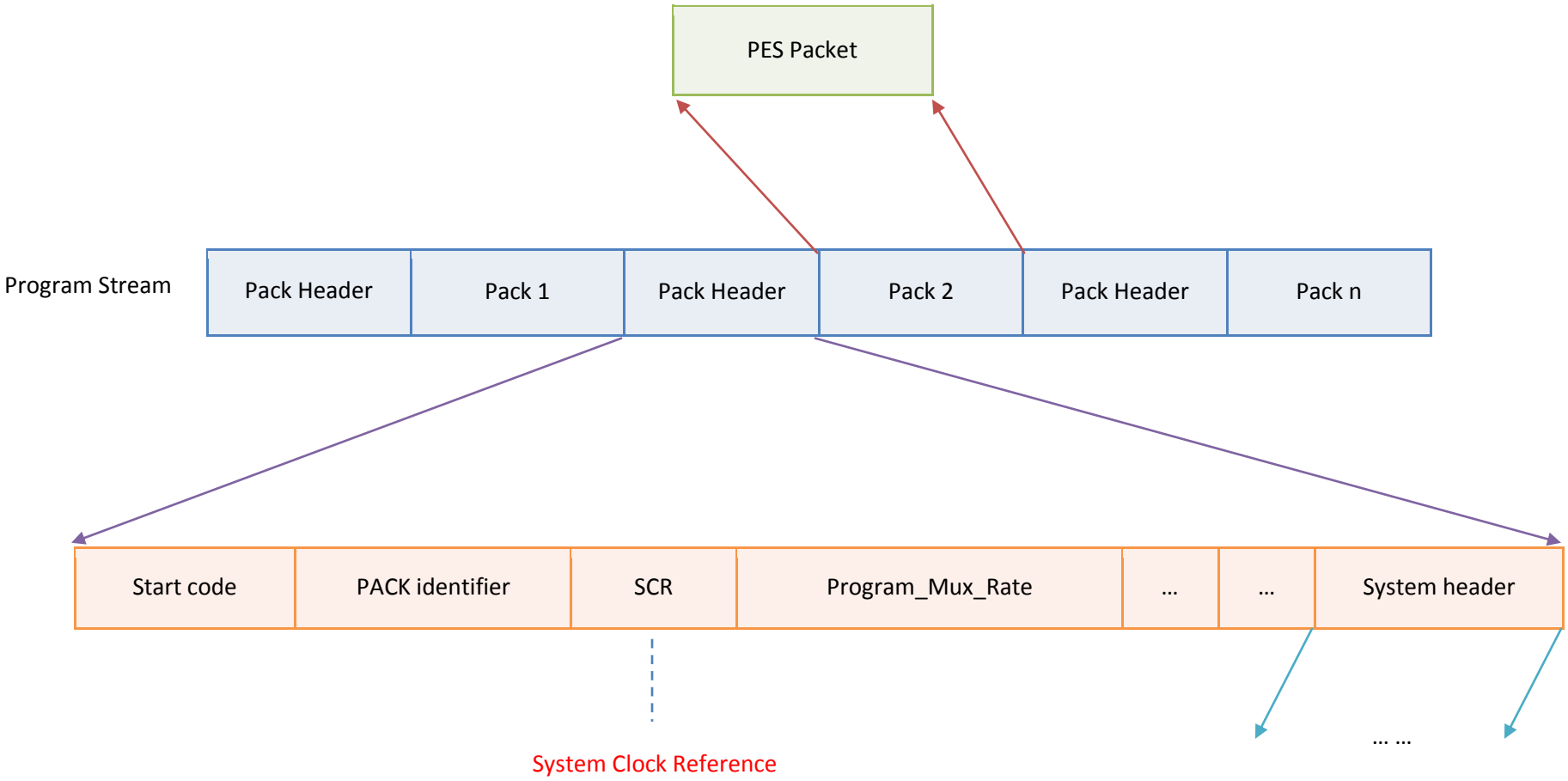
4.4.3. MPEG-2 Program Stream Multiplexed



4.4.4 PES Packet



4.4.5 PS Pack



4.4.6. MPEG-2 Profiles

Abbr.	Name	Picture Coding Types	Chroma Format	Aspect Ratios	Scalable modes
SP	Simple Profile	I, P	4:2:0	square pixels, 4:3, or 16:9	None
MP	Main Profile	I, P, B	4:2:0	square pixels, 4:3, or 16:9	None
422P	4:2:2 Profile	I, P, B	4:2:2 or 4:2:0	square pixels, 4:3, or 16:9	None
SNR	SNR Scalable Profile	I, P, B	4:2:0	square pixels, 4:3, or 16:9	SNR (signal-to-noise ratio) scalable
Spatial	Spatially Scalable Profile	I, P, B	4:2:0	square pixels, 4:3, or 16:9	SNR- or spatial-scalable
HP	High Profile	I, P, B	4:2:2 or 4:2:0	square pixels, 4:3, or 16:9	SNR- or spatial-scalable

4.4.7. MPEG-2 Levels

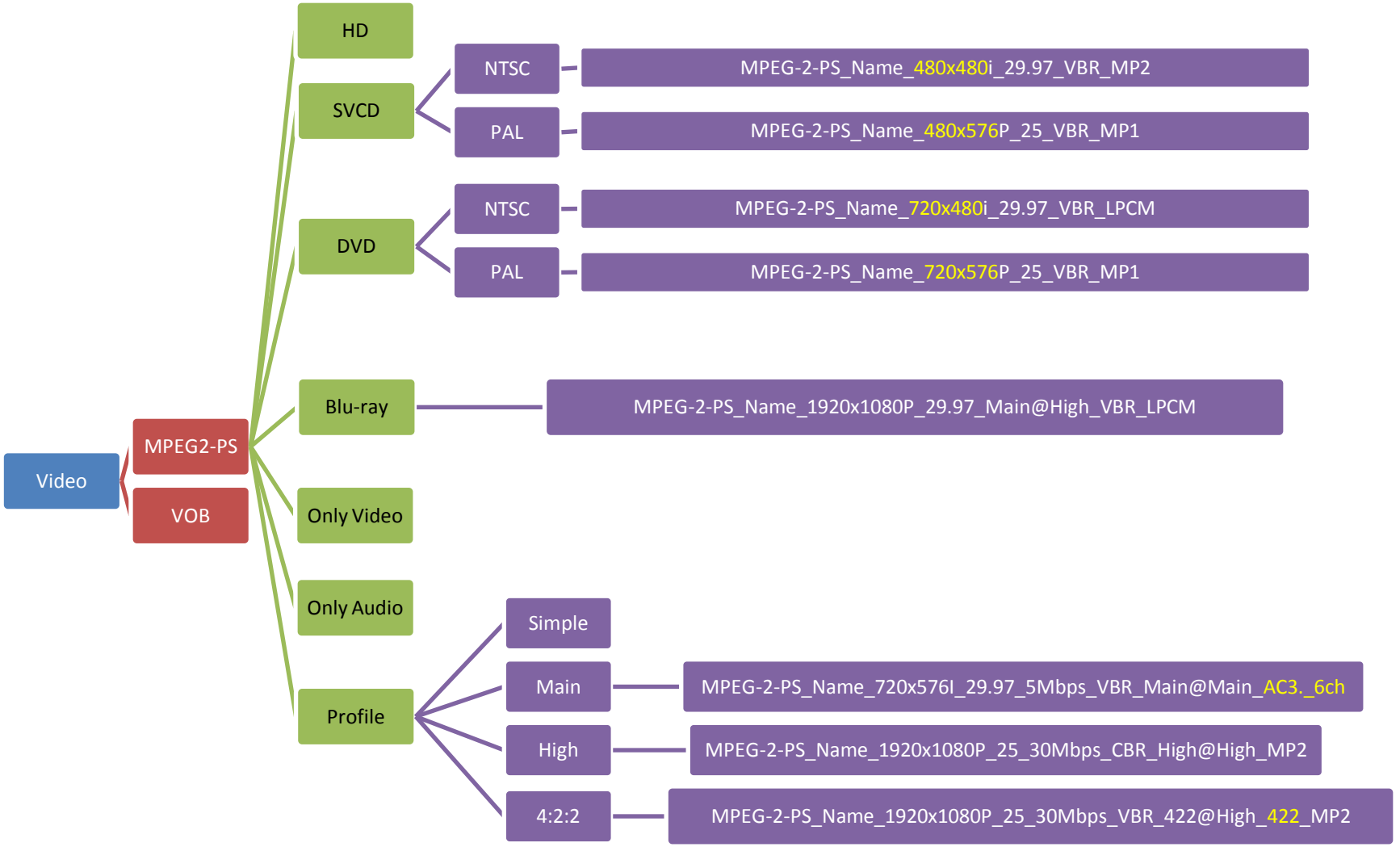
Abbr.	Name	Frame rates (Hz)	Max horizontal resolution	Max vertical resolution	Max bit rate (Mbit/s)
LL	Low Level	23.976, 24, 25, 29.97, 30	352	288	4
ML	Main Level	23.976, 24, 25, 29.97, 30	720	576	15
H-14	High 1440	23.976, 24, 25, 29.97, 30, 50, 59.94, 60	1440	1152	60
HL	High Level	23.976, 24, 25, 29.97, 30, 50, 59.94, 60	1920	1152	80

4.4.8. MPEG-2 typical application

Profile @ Level	Maximum Resolution	Maximum Frame Rate	Sampling	Rate	Comments
Main Profile @ Low Level	352 x 288	30	4:2:0	4Mbps	Set-top boxes
Main Profile @ Main Level	720 x 576/480	30	4:2:0	15Mbps; limited to 9Mbps for DVDs	DVD, SD-DVB
Main Profile @ High 1440	1440 x 1080	60	4:2:0	60Mbps; limited to 25Mbps for DV tape	HDV
Main Profile @ High Level	1920 X 1080	60	4:2:0	80Mbps; limited to 19.4Mbps for terrestrial transmission	ATSC (DTV/HDTV Standard in US), HD-DVB
4:2:2 Profile @ Main Level	720 x 576/480	30	4:2:2	50Mbps	Sony MPEG IMX using I-frame only

4.4.9. Production Tools: MainConcept Reference

4.4.10. Directory Structure


Video
MPEG2-PS
VOB
HD
SVCD
NTSC
PAL
DVD
NTSC
PAL
Blu-ray
Only Video
Only Audio
Profile
Simple
Main
High
4:2:2
MPEG-2-PS_Name_480x480i_29.97_VBR_MP2
MPEG-2-PS_Name_480x576P_25_VBR_MP1
MPEG-2-PS_Name_720x480i_29.97_VBR_LPCM
MPEG-2-PS_Name_720x576P_25_VBR_MP1
MPEG-2-PS_Name_1920x1080P_29.97_Main@High_VBR_LPCM
MPEG-2-PS_Name_720x576I_29.97_5Mbps_VBR_Main@Main_AC3._6ch
MPEG-2-PS_Name_1920x1080P_25_30Mbps_CBR_High@High_MP2
MPEG-2-PS_Name_1920x1080P_25_30Mbps_VBR_422@High_422_MP2

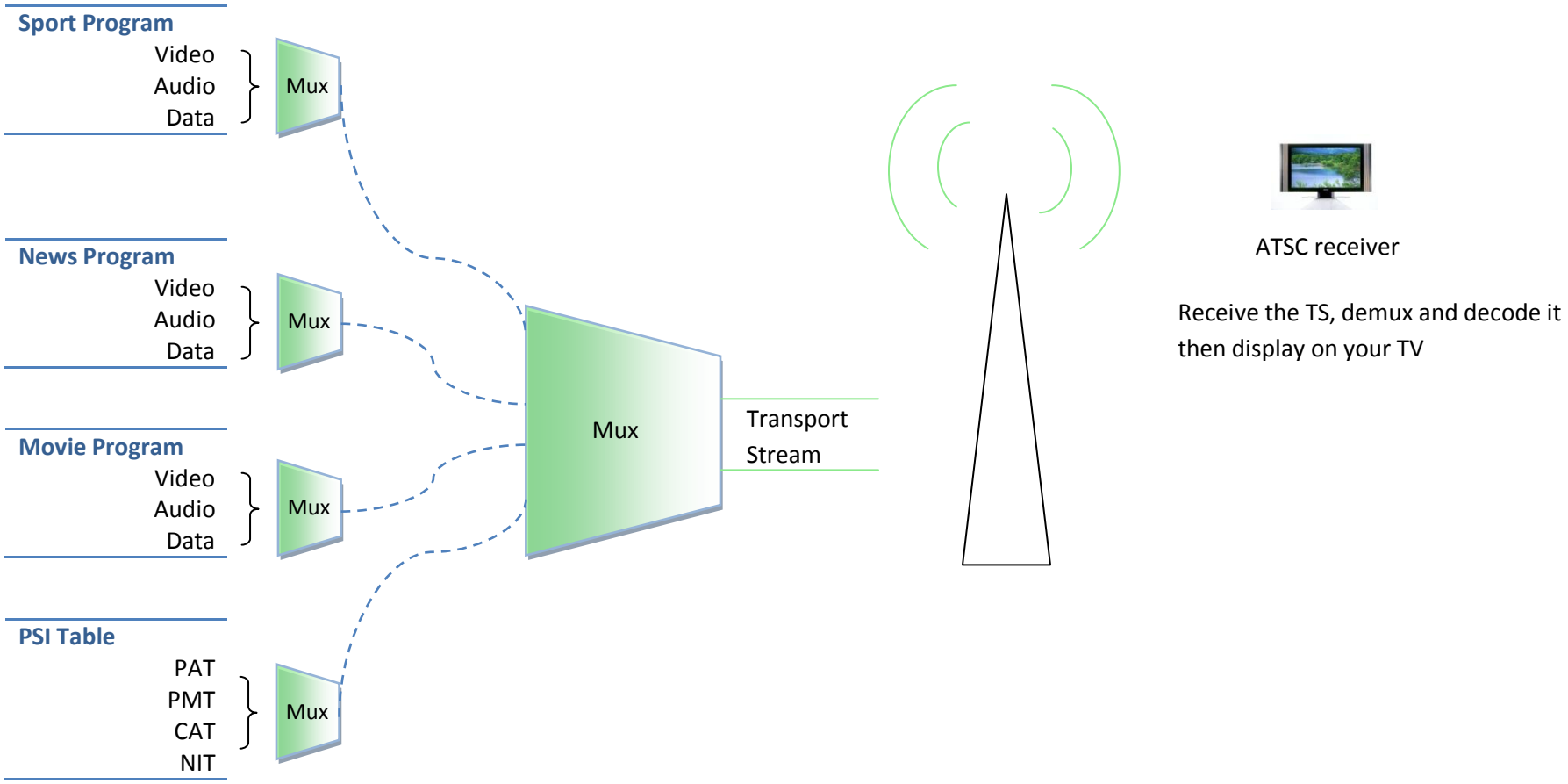
4.5 MPEG-2 Transport Stream (MPEG-2 TS)

- MPEG-2 Transport Stream is defined in MPEG-2 Part 1, Systems (ISO/IEC standard 13818-1 or ITU-T Rec. H.222.0).
- BDAV MPEG-2 Transport Stream is defined in Blu-ray Disc Association BD ROM - Audio Visual Application Format Specifications.

4.5.1. Container & supported video/audio standard (Codec)

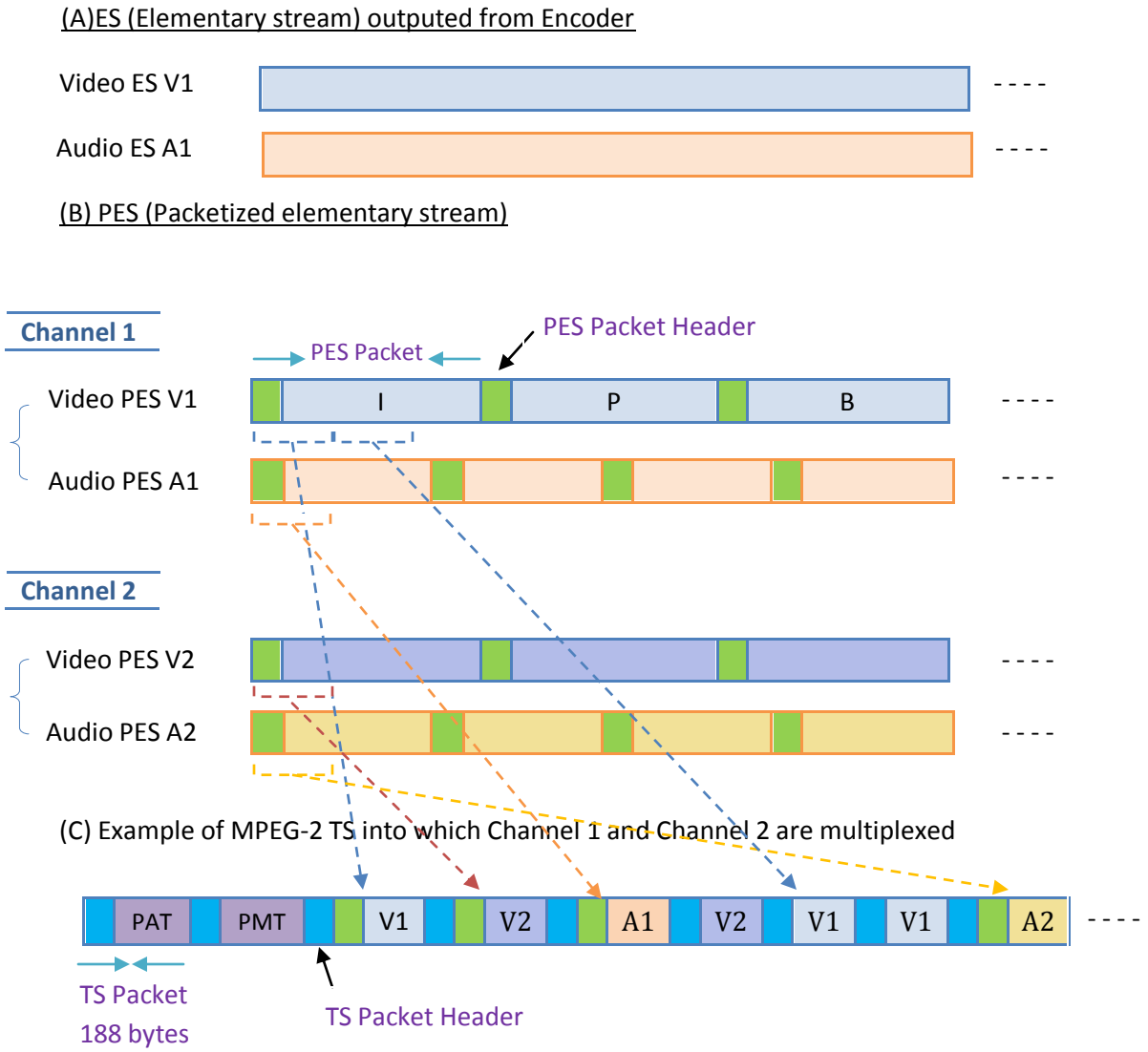
Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments
MPEG-2 Transport Stream	MPEG	.m2t (HDV)	MPEG-2	MPEG-1 Layer II, LPCM	1. Transport Stream offers features for error correction for transportation over unreliable media, and is used in broadcast applications such as DVB and ATSC. It is contrasted with MPEG Program Stream, designed for more reliable media such as DVDs. 2. An MPEG-2 Transport Stream can contain one or more content channels. A Transport Stream is like a high end car or bus: it can carry one or more passengers and has a good suspension system that can handle the bumps in the road.
		.tp, .trp, .ts (DTV/HDTV)	MPEG-2 H.264/MPEG4-AVC	MPEG-1 Layer III , AC-3, AAC,DTS	
BDAV MPEG-2 Transport Stream	Blu-ray Disc Association	.m2ts (BD)	H.264/MPEG4-AVC VC-1 MPEG-2	AC-3(M), DTS(M), LPCM(M), DTS-HD(O), Dolby Digital Plus(O)， Dolby TrueHD (lossless / O)	1. Audio, video and other streams are multiplexed and stored on Blu-ray Discs in a container format based on the MPEG Transport Stream. It is also known as BDAV MPEG-2 Transport Stream. This container is commonly used for high definition video on Blu-ray Disc and AVCHD and HD Camcorders. 2. Blu-ray Disc Video use MPEG Transport Streams, compared to DVD's MPEG Program Streams. This allows multiple video programs to be stored in the same file so they can be played back simultaneously (e.g., with "Picture in picture" effect). 3. Audio formats listed as mandatory(M) must be supported by every player.
		.m2ts, .mts (AVCHD)	H.264/MPEG-4 AVC	AC-3, LPCM	

4.5.2. MPEG-2 Transport Stream (High Level)

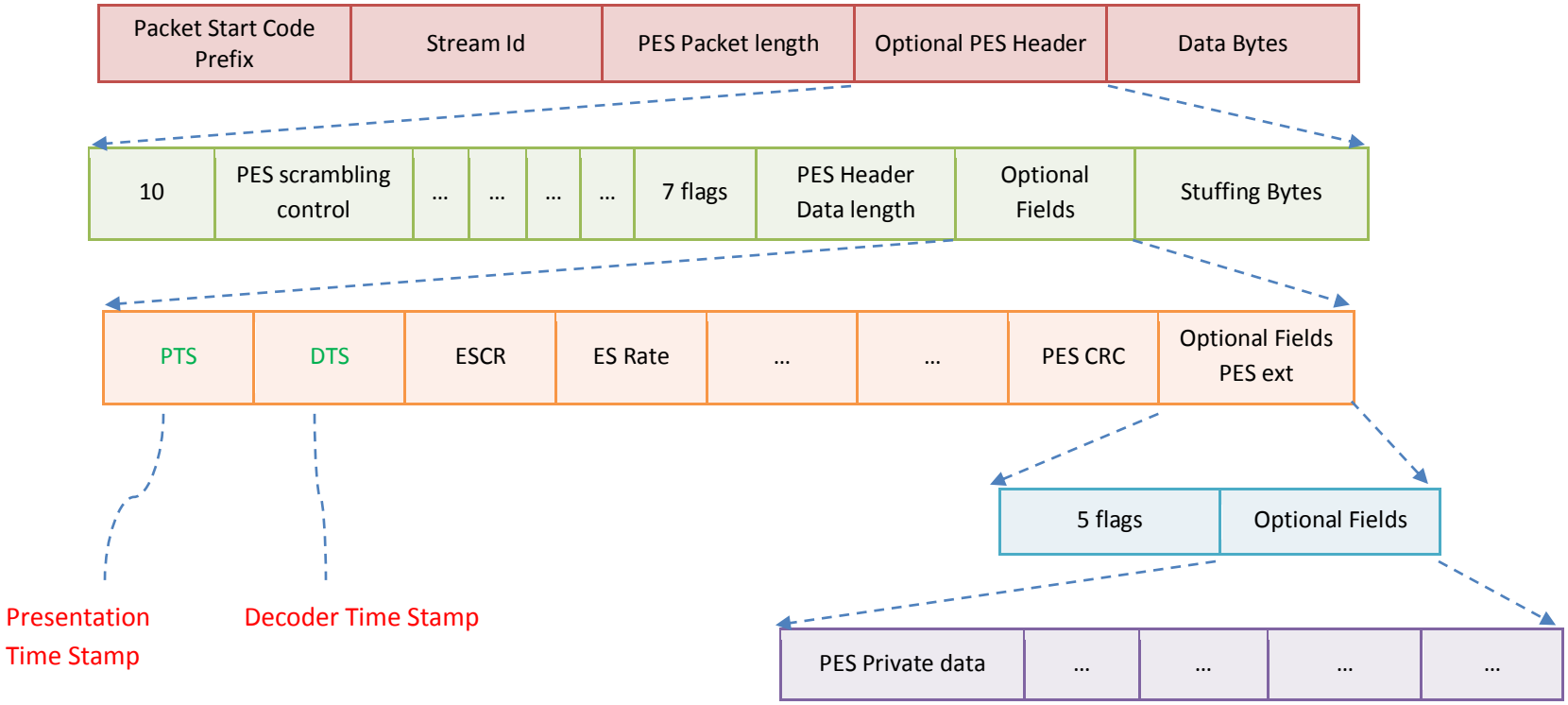


Multiple MPEG programs are combined into MPEG-2 Transport Stream then sent to a transmitting antenna. In the US broadcast digital TV system, an ATSC receiver actually receives all of the Programs (channels) at once (it has no choice!), but it only demultiplexes and then decodes the selected content, one at a time, from the delivered Transport Stream.

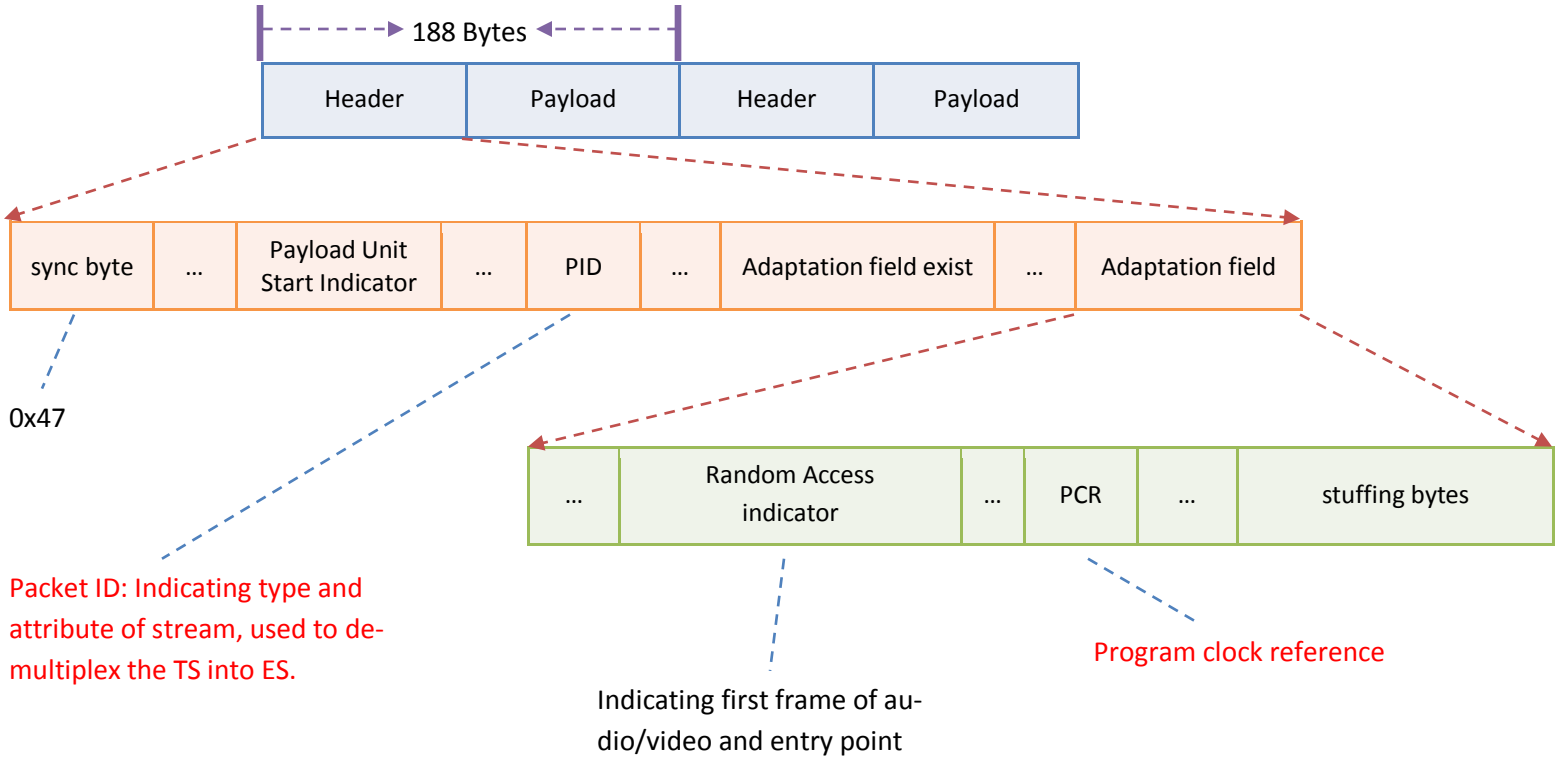
4.5.3. MPEG-2 Transport Stream Multiplexed



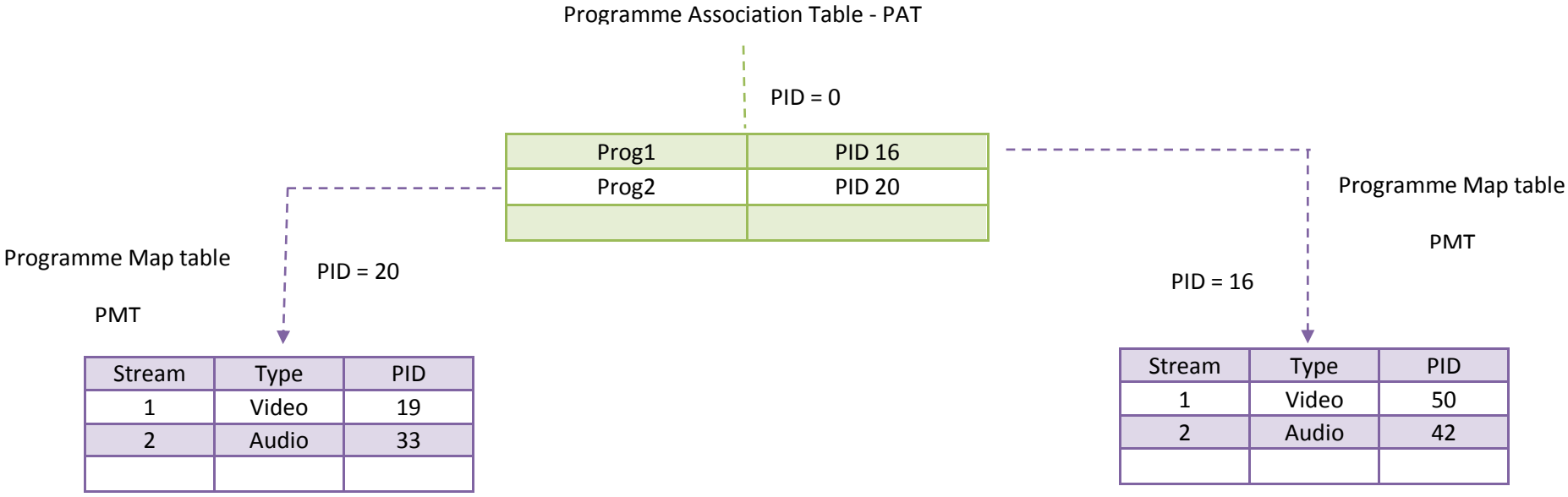
4.5.4 PES Packet

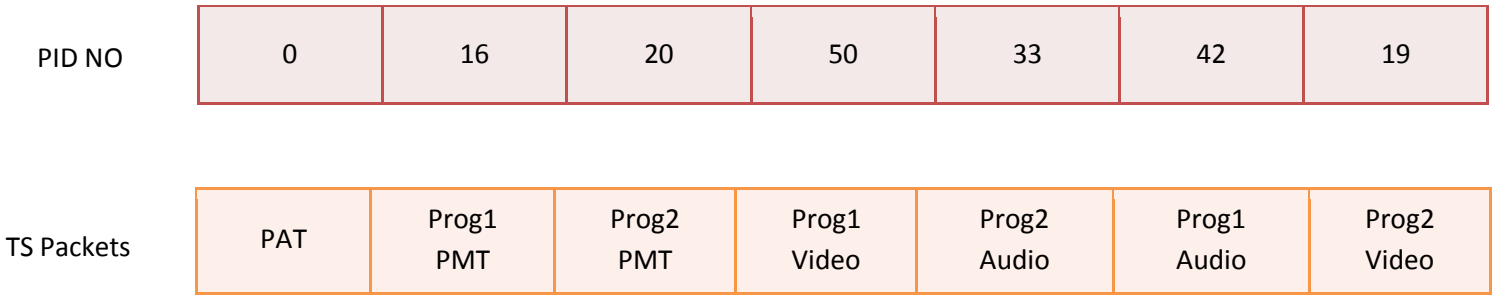


4.5.5 TS Packet



4.5.6 PAT & PMT





4.5.7. BDAV MPEG-2 TS Data Flow

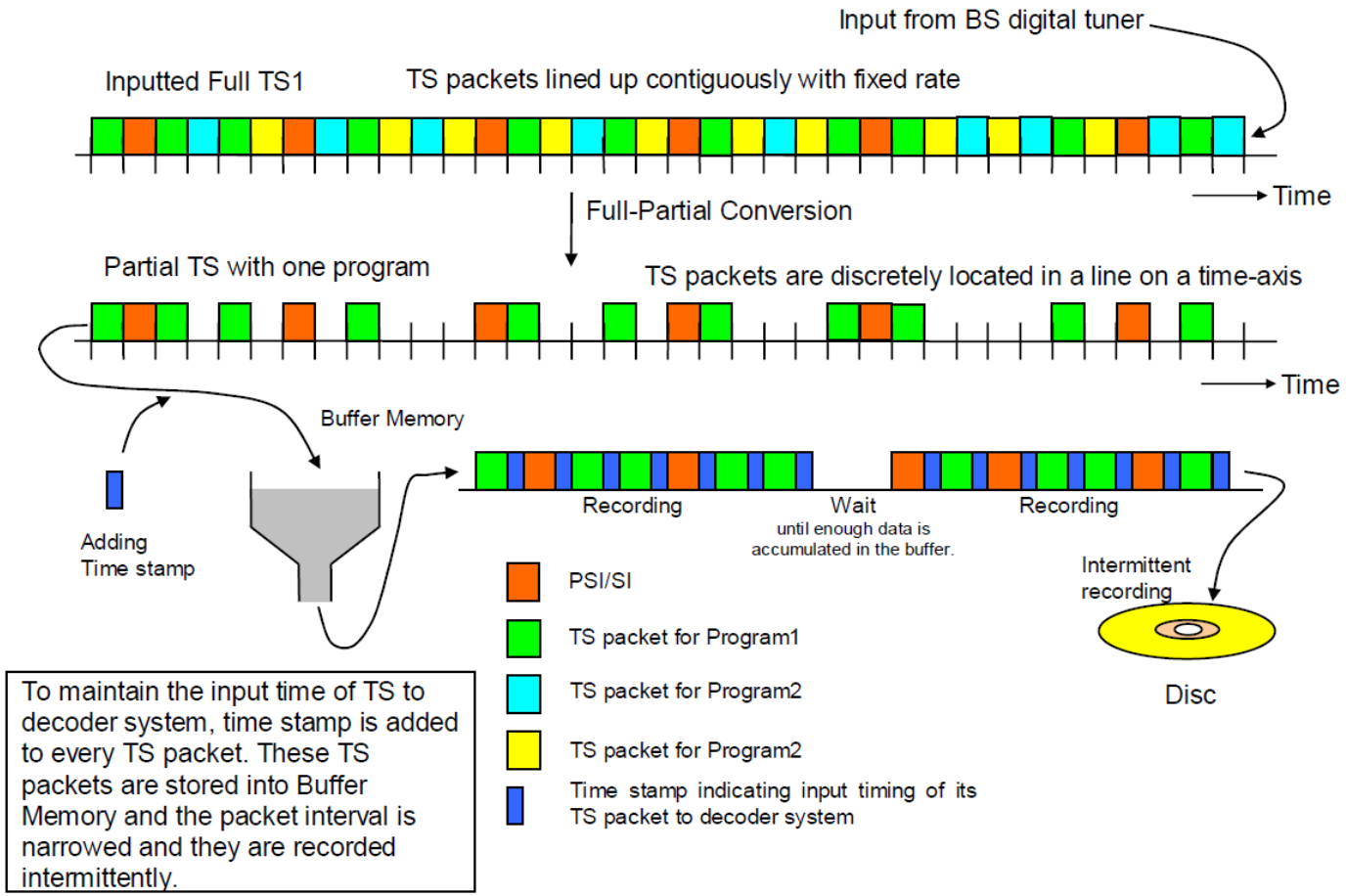
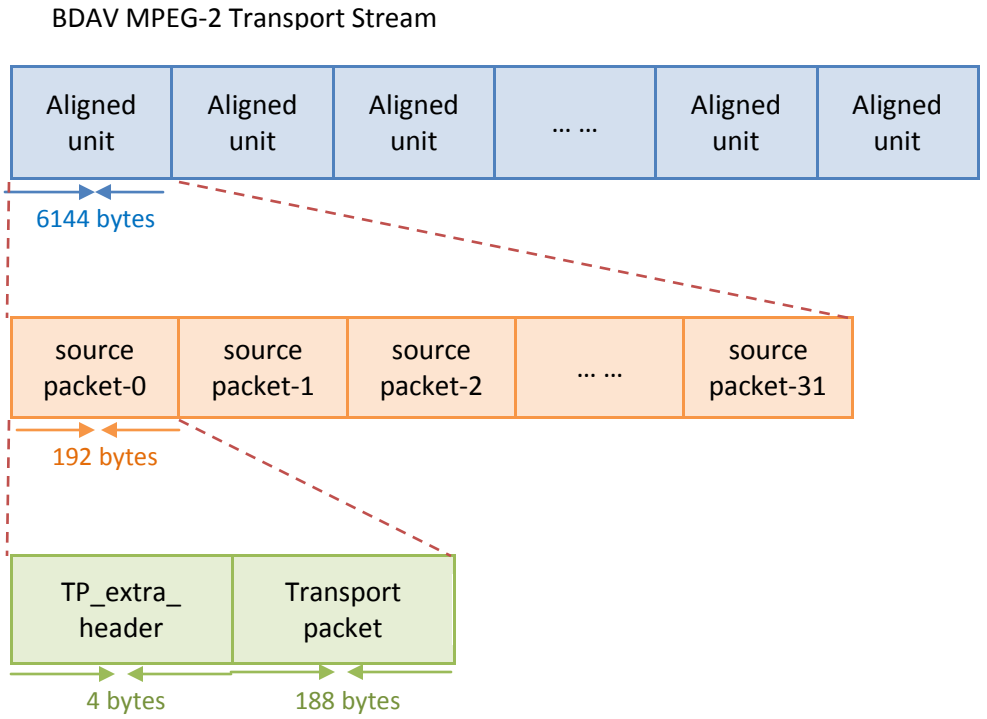


Figure 3.1.2.3: Data flow until the inputted Transport Stream is recorded to the disc

4.5.8. BDAV MPEG-2 TS Structure



Aligned units are recorded in three consecutive logical sectors on the BD-ROM disc. The size of one logical sector is 2048 bytes.

4.5.9. HDV Spec

Format		HDV 720p	HDV 1080i	HDV 1080p
Video	Scanning type	progressive	interlaced	progressive
	Frame aspect ratio	16x9		
	Frame size in pixels	1280 x 720	1440 x 1080	
	Pixel aspect ratio	1.0	1.33	
	Video signal	720/60p, 720/30p, 720/50p, 720/25p, 720/24p	1080/60i, 1080/50i	1080/30p, 1080/25p, 1080/24p
	Video Compression	MPEG2 Video (profile & level: MP@H-14/HL)	MPEG2 Video (profile & level: MP@H-14)	
	Sampling frequency for luminance	74.25 MHz	55.6875 MHz	
	Chroma sampling format	4:2:0		
	Quantization	8 bits (both luminance and chrominance)		
	Compressed video bitstream rate	~19.7 Mbps	~25 Mbps	
Audio	Compression	MPEG-1 Audio Layer II, PCM	MPEG-1 Audio Layer II	
	Sampling frequency	48 kHz		
	Quantization	16 bits		
	Audio modes and data rate	Stereo (2-channel) at 384 kbit/s (192 kbit/s per channel); optional 4-channel at 96 kbit/s per channel.		
System	Stream type	MPEG-2 Transport Stream		
	Stream interface	IEEE 1394 in alpha mode (also known as FireWire 400 or i.LINK)		
	File extension	m2t (generally)		
	Media	Same as DV format (DV and/or Mini DV cassette tape)		

4.5.10. AVCHD Spec

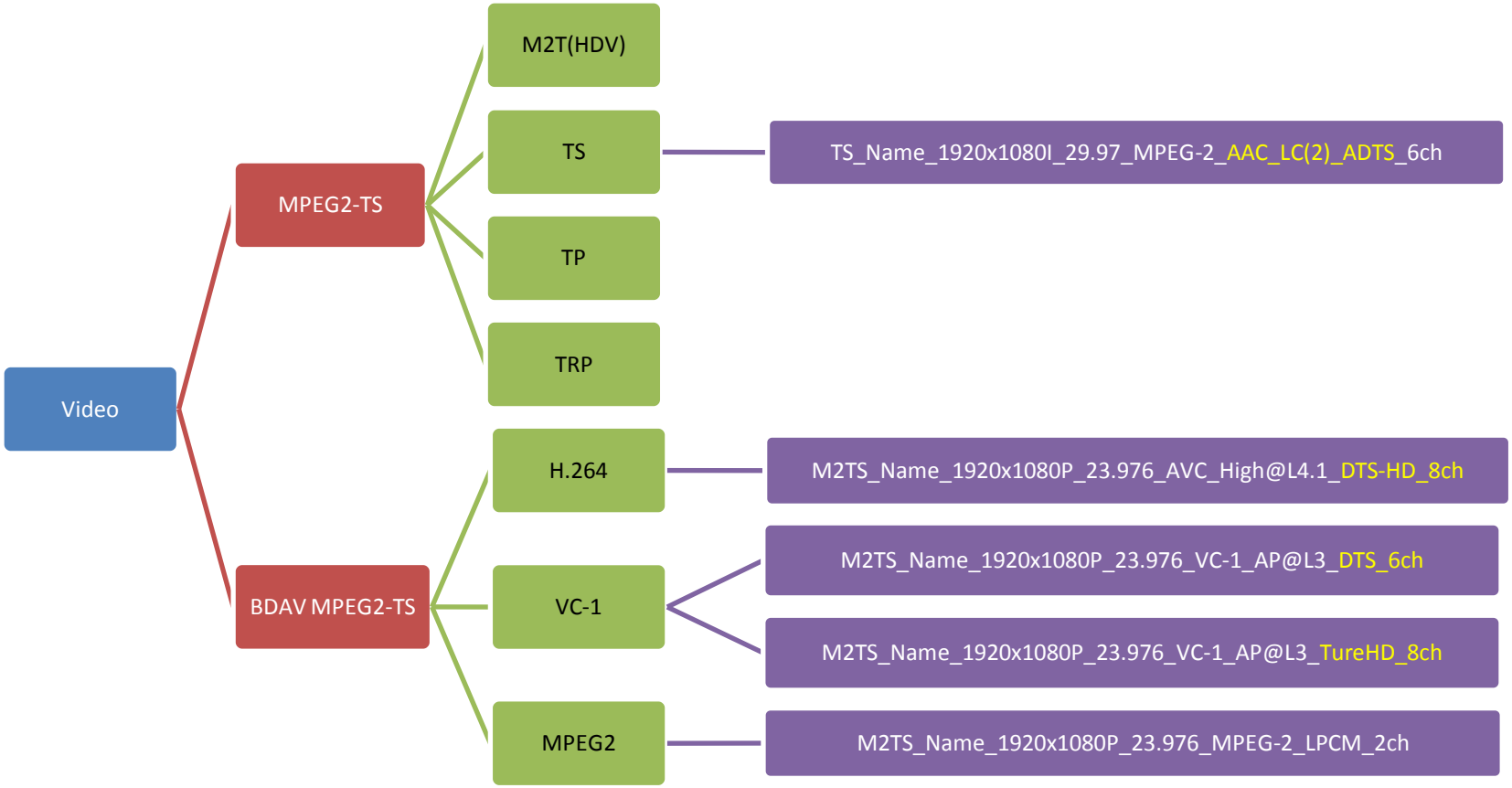
Format		1080/60i 1080/50i 1080/24p	720/60p 720/50p 720/24p	480/60i	576/50i
Video	Frame size in pixels	1920×1080 1440×1080	1280 x 720	720×480	720×576
	Frame aspect ratio	16:9		4:3, 16:9	
	Video Compression	MPEG-4 AVC/H.264 (Main Profile Level-4.0 or High Profile Level-4.1, depending on vendor)			
	Luminance sampling frequency	74.25 MHz 55.7 MHz	74.25 MHz	13.5 MHz	13.5 MHz
	Chroma sampling format	4:2:0			
	Quantization	8 bits (both luminance and chrominance)			
Audio	Compression	Dolby Digital (AC-3)		Linear PCM	
	Compressed audio bitstream rate	64 to 640 kbps		1.5 Mbps (2 channels)	
	Audio mode	1-5.1 channels		1-7.1 channels	
System	Stream type	MPEG-2 Transport Stream			
	System data rate	up to 24 Mbps (AVCHD conforming to H264 High-Profile, Level 4.1) up to 17 Mbps (AVCHD conforming to H264 Main-Profile, Level 4.0) up to 18 Mbps for DVD media			
	File extension (generally)	mts (on camcorder), m2ts (after import to computer)			
	Media	8 cm optical media (DVD) SD/SDHC Memory Card "Memory Stick" Built-in hard-disk or flash Media			

4.5.11. BD-ROM Video streams Spec

Video	Standard	MPEG-2: MP@HL and MP@ML
		MPEG-4 AVC: MPEG-4 AVC: HP@4.1/4.0 and MP@4.1/4.0/3.2/3.1/3.0
		SMPTE VC-1: AP@L3 and AP@L2
	Max. bitrate	40Mbps
	HD	1920x1080x59.94-i, 50-i (16:9) 1920x1080x24-p, 23.976-p (16:9) 1440x1080x59.94-i, 50-i (16:9) MPEG-4 AVC / SMPTE VC-1 only 1440x1080x24-p, 23.976-p (16:9) MPEG-4 AVC / SMPTE VC-1 only 1280x720x59.94-p, 50-p (16:9) 1280x720x24-p, 23.976-p (16:9)
	SD	720x480x59.94-i (4:3/16:9) 720x576x50-i (4:3/16:9)

4.5.12. Production Tools: HDV, AVCHD, tsMuxeR

4.5.13. Directory Structure


Video
MPEG2-TS
M2T(HDV)
TS
TS_Name_1920x1080I_29.97_MPEG-2_AAC_LC(2)_ADTS_6ch
TP
TRP
BDAV MPEG2-TS
H.264
M2TS_Name_1920x1080P_23.976_AVC_High@L4.1_DTS-HD_8ch
VC-1
M2TS_Name_1920x1080P_23.976_VC-1_AP@L3_DTS_6ch
M2TS_Name_1920x1080P_23.976_VC-1_AP@L3_TureHD_8ch
MPEG2
M2TS_Name_1920x1080P_23.976_MPEG-2_LPCM_2ch

4.6 MPEG-4 Part 14 (MP4)

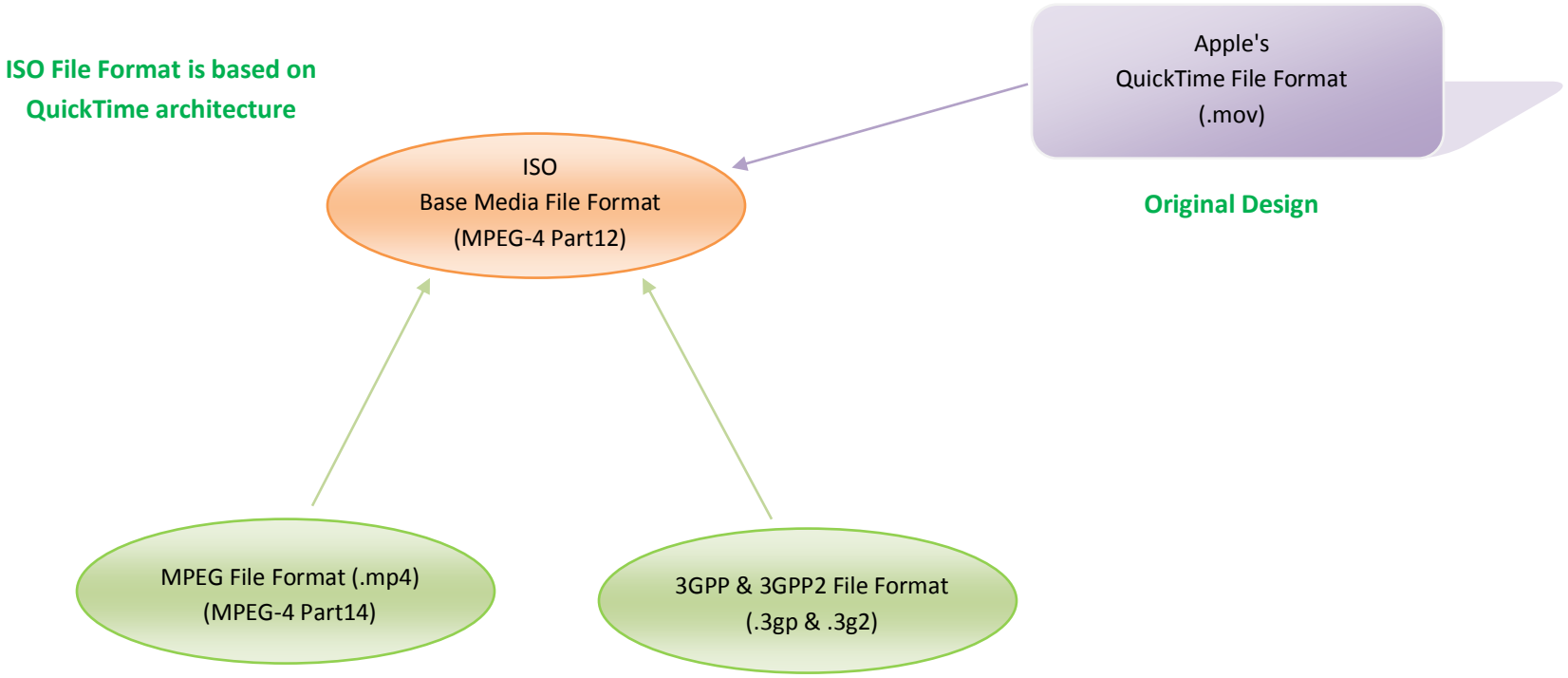
- MP4 file format is defined in MPEG-4 Part 14(ISO/IEC 14496-14).

4.6.1. Container & supported video/audio standard (Codec)

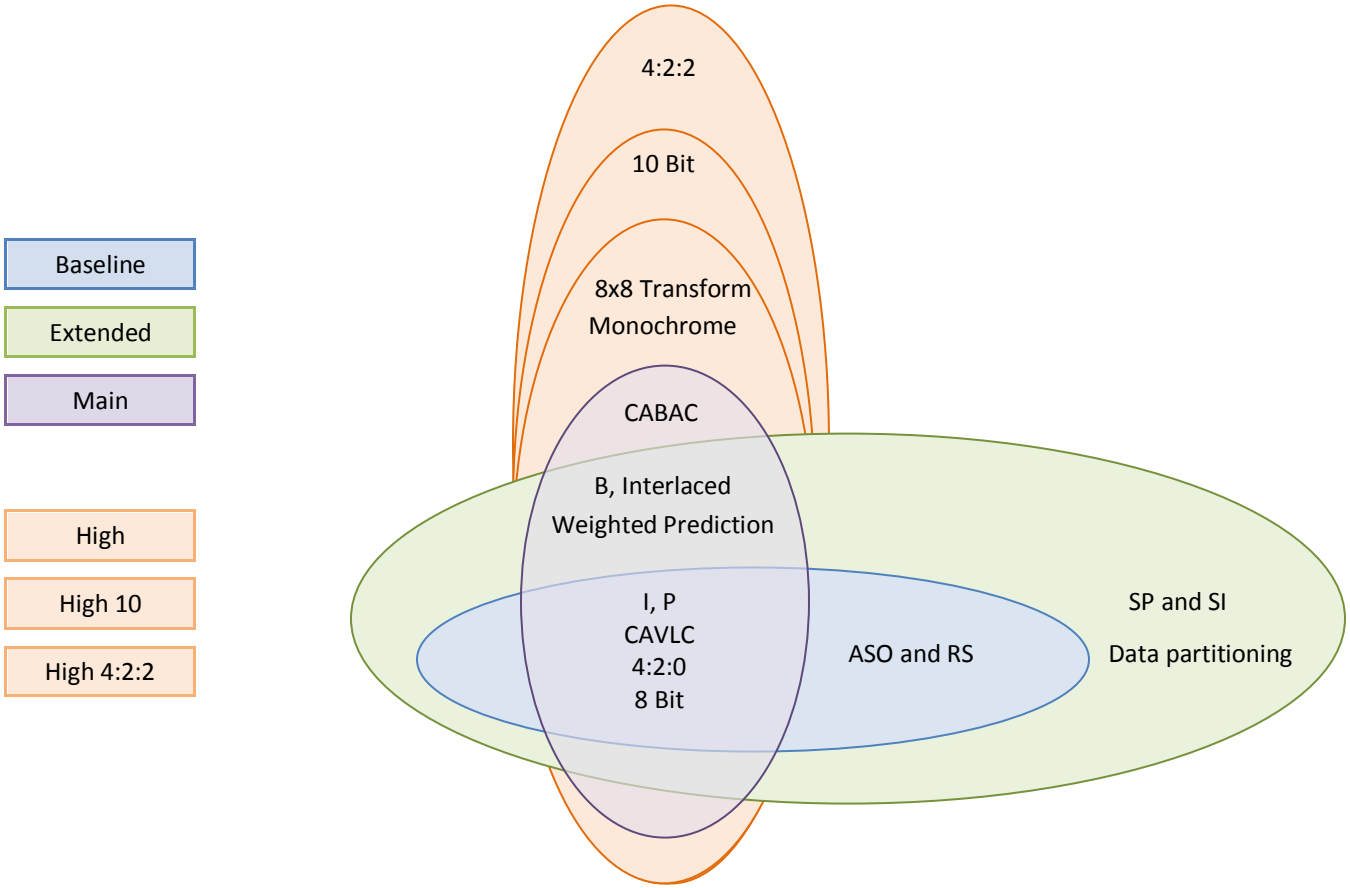
Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments:
QuickTime	Apple	.mov, .qt	Limited to what is available to the QuickTime codec manager	Limited to what is available to CoreAudio	1. MPEG-4 Part 14(MP4) is not directly based upon Apple's QuickTime (MOV) container format. However, Apple's QuickTime (MOV) container format was indeed taken as a starting point for MP4. MP4, being an international standard, has more support than QuickTime container. This is especially true on hardware devices, such as the Sony PSP, PS3, Microsoft Xbox 360 and various DVD players. 2. MP4 on the other hand itself was the starting point for the development of the 3GP format (mainly for mobile phones), which has more in common with MP4 as MP4 with QuickTime for example, but still is a different format. 3. The common but non-standard use of the extensions .m4a and .m4v is due to the popularity of Apple's iPod, iPhone, and iTunes Store. Audio-only MP4 files generally have an .m4a extension. This is especially true of non-protected content. 4. MP4 files with audio streams encrypted by FairPlay Digital Rights Management as sold through the iTunes Store use the .m4p extension. 5. A .qt file is (for historical reasons) an alternate extension for QuickTime movie files (.qt's and .mov's are identical).
MP4	MPEG	.mp4(A,V, A+V)	MPEG-1, MPEG-2, H.263, MPEG-4 part2, H.264/MPEG4-AVC, VC-1	MPEG-1 Layers I, II, III, AAC, AC-3	
		.m4v (A+V)	H.264/MPEG4-AVC	AAC, AC3	
		.m4a (A)		AAC	

3GPP	3GPP	.3gp (A,V,A+V)	MPEG-4 part2, H.264/MPEG4-AVC, H.263	AMR-NB, AMR-WB, AMR-WB+, AAC-LC, HE-AAC and HE-AAC v2	3GPP, defined by a group of telecommunications standards bodies called the 3rd Generation Partnership Project (3GPP), was created for use on Global System for Mobile Communication (GSM) networks, the most popular type of 3G networks across the globe. 3GPP2 was defined by a different group of telecommunications bodies called 3rd Generation Partnership Project 2 (3GPP2) for use on the second most predominate type of 3G network, Code Division Multiple Access (CDMA) 2000. The 3GPP and 3GPP2 formats are very similar, as both are based on the QuickTime file format and contain MPEG-4 and H.263 video, AAC and AMR audio, and 3G Text. 3GPP2 adds the option to use QCELP audio and Movie Fragments, a technology that allows multimedia content to be delivered incrementally over standard TCP wireless networks, providing a more immediate viewing experience for the end user.
3GPP2	3GPP2	.3g2 (A,V,A+V)		AMR-NB, AMR-WB, AAC-LC, HE-AAC, EVRC, EVRC-B, EVRC-WB, QCELP, SMV or VMR-WB	

4.6.2. Relationship between ISO, QuickTime, MP4 and 3GPP&3GPP2

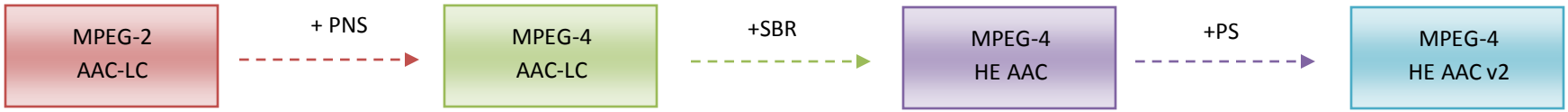


4.6.3. H.264/MPEG-4 AVC Profiles



4.6.4. AAC Profiles

Profile	Features	Typical Applications	Typical Bitrate
AAC-LC (Low Complexity AAC)	High performance audio codec for excellent audio quality at low bit rates	– Apple iPod – iTunes – ISDB television broadcasting (Japan)	128 kbit/s (Stereo)
HE-AAC (High Efficiency AAC)	High performance audio codec for good quality at bit rates of 28 kbit/s per channel and below	– XM Radio – Mobile music download – Digital Radio Mondiale	56 kbit/s (Stereo)
HE-AAC v2	Highest performance audio codec for good quality at bit rates below of 24 kbit/s per channel	– 3GPP music download – Digital radio DAB+	48 kbit/s (Stereo)



PNS: Perceptual Noise Substitution SBR: Spectral Band Replication PS: Parametric Stereo

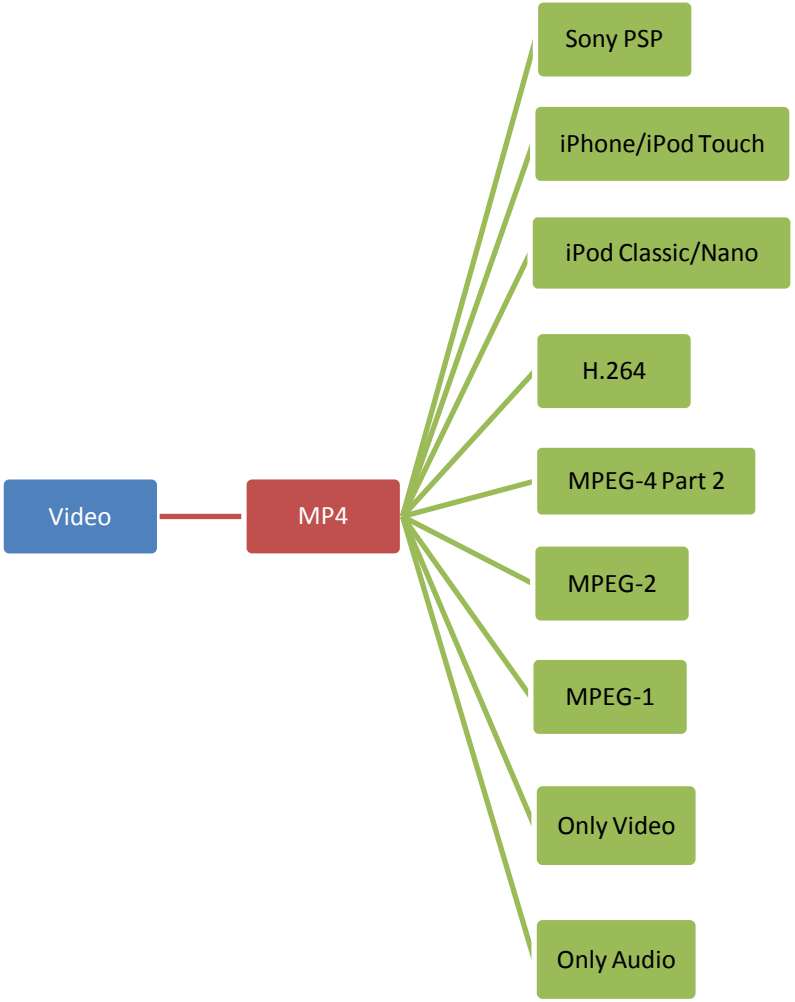
AAC IS defined in MPEG-2 Part 7(ISO/IEC 13818-7) and MPEG-4 Part 3(SO/IEC 14496-3)

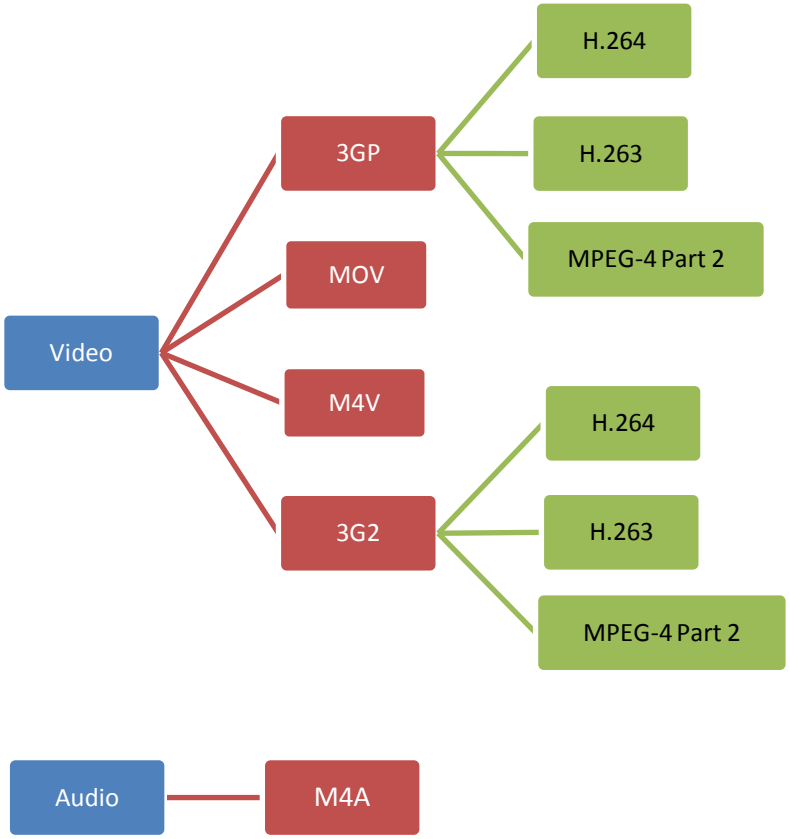
4.6.5. Apple iPhone & iPod spec

File Formats	Video	Audio
.m4v, .mp4, and .mov	H.264 video, up to 1.5 Mbps, 640 by 480 pixels, 30 frames per second, Low-Complexity version of the H.264 Baseline Profile	AAC-LC audio up to 160 Kbps, 48kHz, stereo audio
	H.264 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Baseline Profile up to Level 3.0	
	MPEG-4 Part2 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Simple Profile	

4.6.6. Production Tools: Yamb, QuickTime 7 Pro, MainConcept Reference, Nero AAC Codec

4.6.7. Directory Structure


Video
MP4
Sony PSP
iPhone/iPod Touch
iPod Classic/Nano
H.264
MPEG-4 Part 2
MPEG-2
MPEG-1
Only Video
Only Audio


Video
3GP
H.264
H.263
MPEG-4 Part 2
MOV
M4V
3G2
H.264
H.263
MPEG-4 Part 2
Audio
M4A

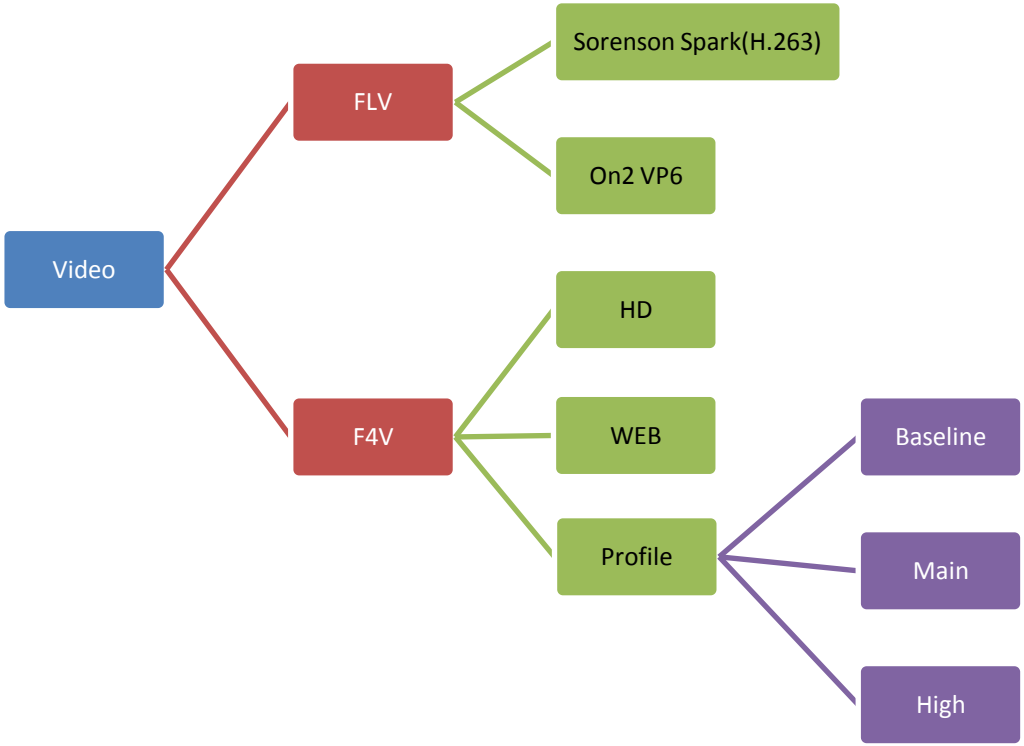
4.7 Flash Video (FLV and F4V)

4.7.1 Container & supported video/audio standard (Codec)

Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments:
FLV (Flash Video)	Adobe Systems	.flv	Sorenson Spark(H.263), On2 VP6, Macromedia Screen Video	MP3, Nellymoser, ADPCM, LPCM, AAC, Speex	It is possible to place H.264 and AAC streams into the traditional FLV file, but Adobe strongly encourages everyone to embrace the new standard file format (F4V). There are functional limits with the FLV structure when streaming H.264 which couldn't be overcome without a redesign of the file format. This is one of the reasons Adobe is moving away from the traditional FLV file structure. Specifically dealing with sequence headers and enders is tricky with FLV streams.
F4V (Flash Video)	Adobe Systems	.f4v	H.264/MPEG-4 AVC	MP3, AAC-LC, HE-AAC	The latter F4V file format is based on the ISO base media file format and is supported starting with Flash Player 9 update 3

4.7.2. Production Tools: Adobe Media Encoder

4.7.3. Directory Structure

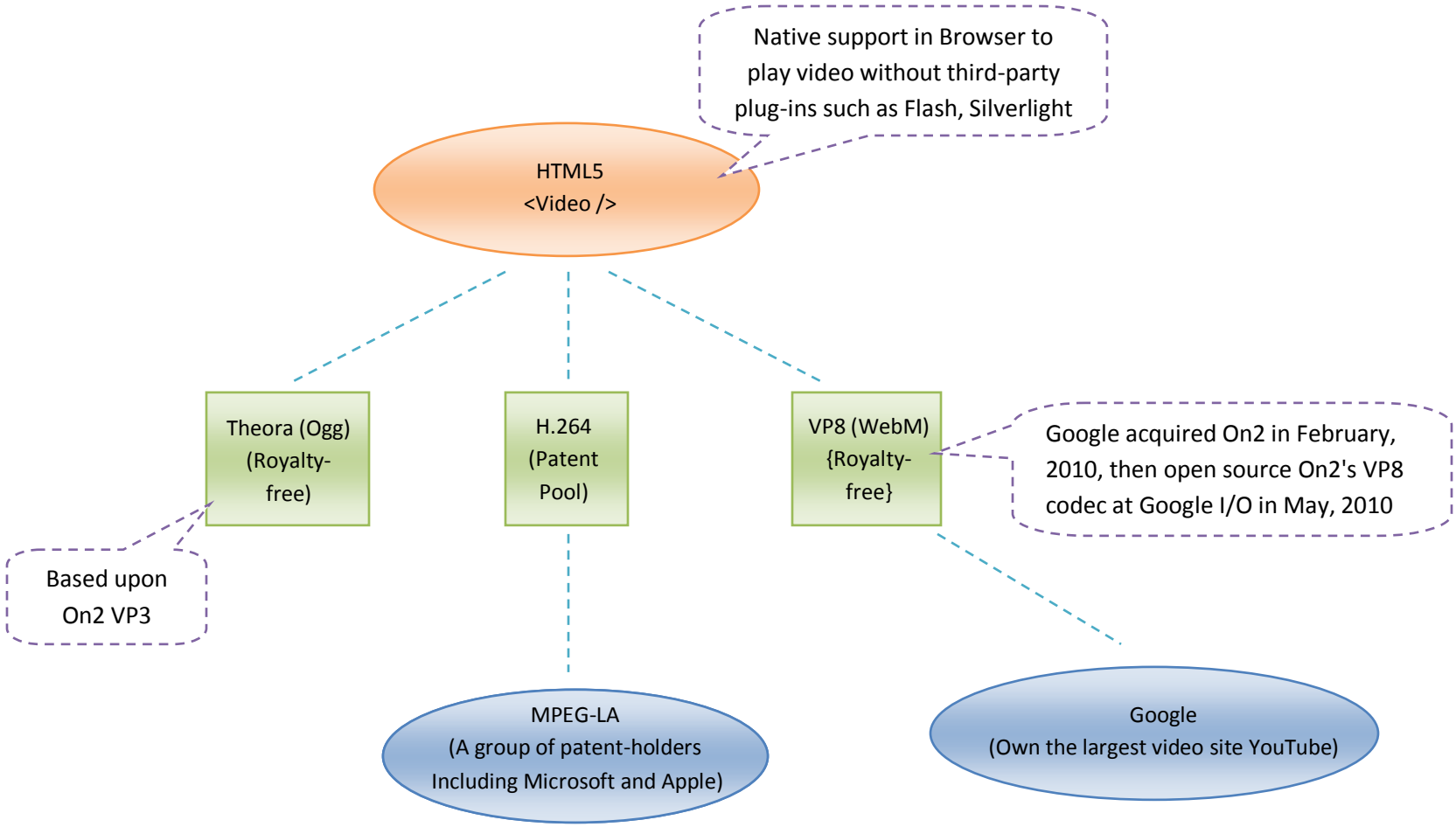

Video
FLV
Sorenson Spark(H.263)
On2 VP6
F4V
HD
WEB
Profile
Baseline
Main
High

4.8 Matroska (MKV)

4.8.1 Container & supported video/audio standard (Codec)

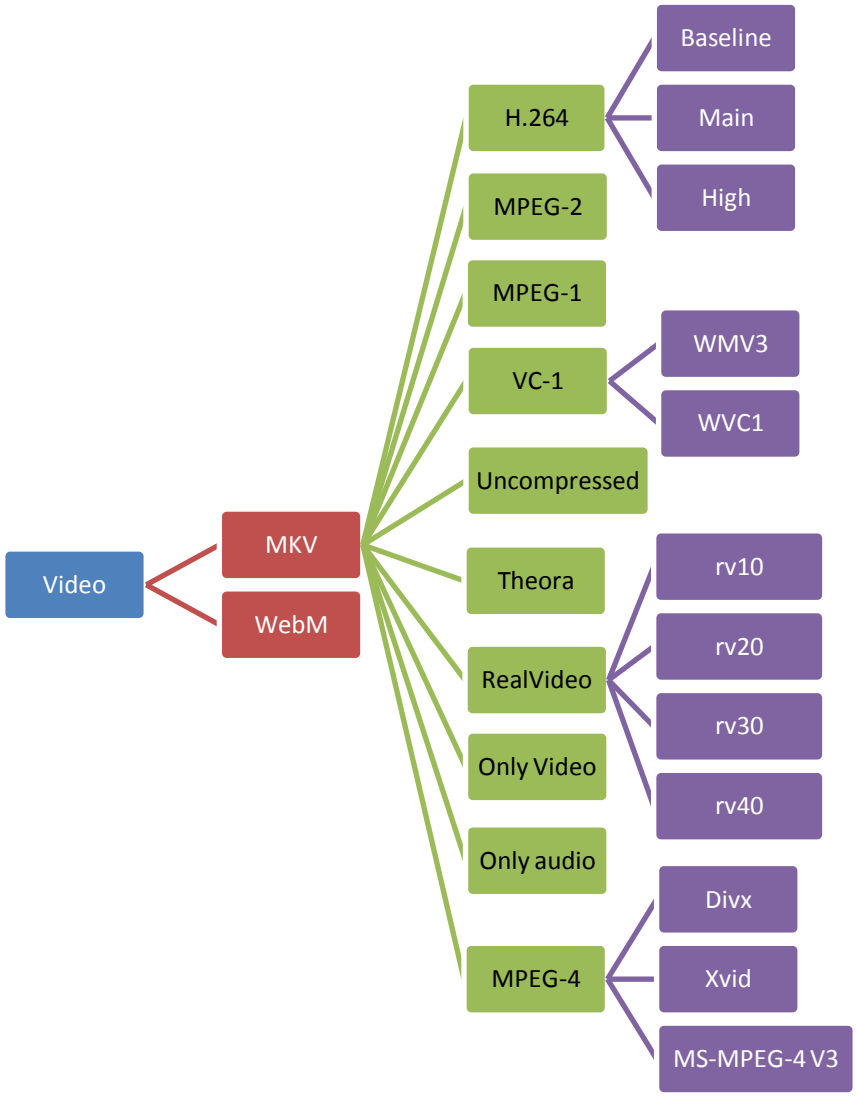
Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments:
Matroska	CoreCodec	.mkv (A,V,A+V)	virtually anything	virtually anything	The matroska is a free, opensource container format, aiming to be able to offer a lot of advanced features, which older formats like AVI can't handle, on an extensible basis. Matroska supports for example the storage of Variable Bitrate audio content (VBR) without any hassles, Variable Framerates (VFR), Chapters, attachment of files, Error Detection (EDC) and modern A/V Codecs like "Advanced Audio Coding" (AAC), "Ogg Vorbis" or "Realvideo 9" (RV9), next to nothing handled by AVI.
		.mka (A)		virtually anything	
WebM	Google	.webm	On2 VP8	Ogg Vorbis	WebM is an open, royalty-free media file format designed for use with HTML5 video. WebM files consist of video streams compressed with the VP8 video codec and audio streams compressed with the Vorbis audio codec. The WebM file structure is based on a subset of the Matroska media container.

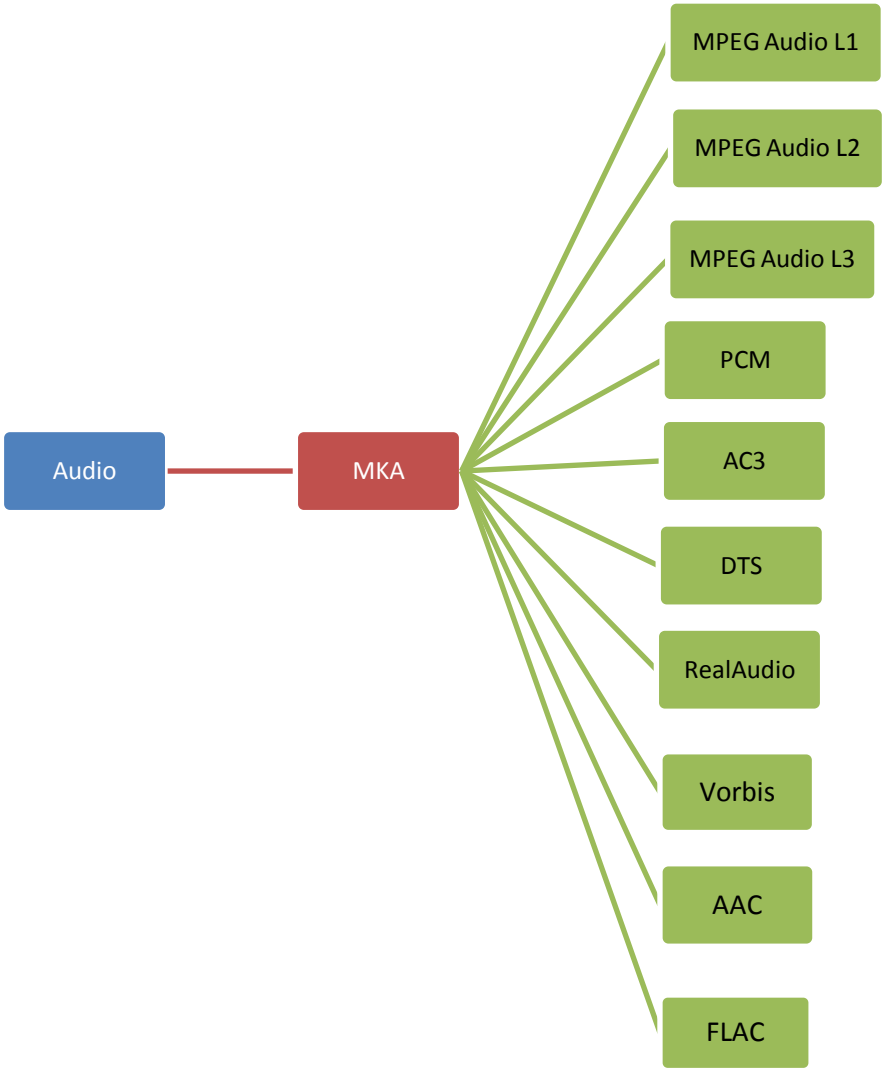
4.8.2. HTML5 Video



4.8.3. Production Tools: MKVtoolnix, VirtualDubMod, AVI-Mux GUI, ffmpeg

4.8.4. Directory Structure


Video
MKV
WebM
H.264
MPEG-2
MPEG-1
VC-1
Uncompressed
Theora
RealVideo
Only Video
Only audio
MPEG-4
Baseline
Main
High
WMV3
WVC1
rv10
rv20
rv30
rv40
Divx
Xvid
MS-MPEG-4 V3


Audio
MKA
MPEG Audio L1
MPEG Audio L2
MPEG Audio L3
PCM
AC3
DTS
RealAudio
Vorbis
AAC
FLAC

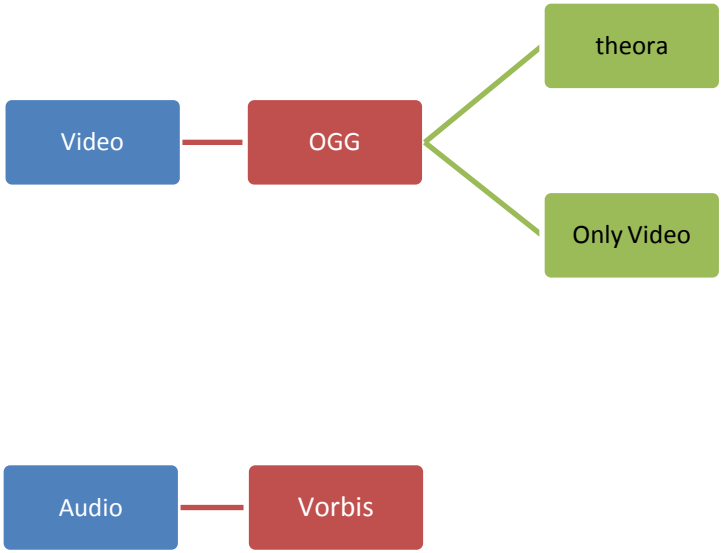
4.9 Ogg

4.9.1 Container & supported video/audio standard (Codec)

Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments:
Ogg	Xiph.org	.ogv, ogg (V,A+V)	Theora	Vorbis, FLAC, Speex, OggPCM	Ogg is a free, open standard container format maintained by the Xiph.Org Foundation. Before 2007, the .ogg filename extension was used for all files whose content used the Ogg container format. Since 2007, the Xiph.Org Foundation recommends that .ogg only be used for Ogg Vorbis audio files. The Xiph.Org Foundation decided to create a new set of file extensions and media types to describe different types of content such as .oga for audio only files, .ogv for video with or without sound (including Theora), and .ogx for applications
		.oga, .ogg (A)		Vorbis, FLAC, Speex, OggPCM	

4.9.2. Production Tools: ffmpeg2theora

4.9.3. Directory Structure



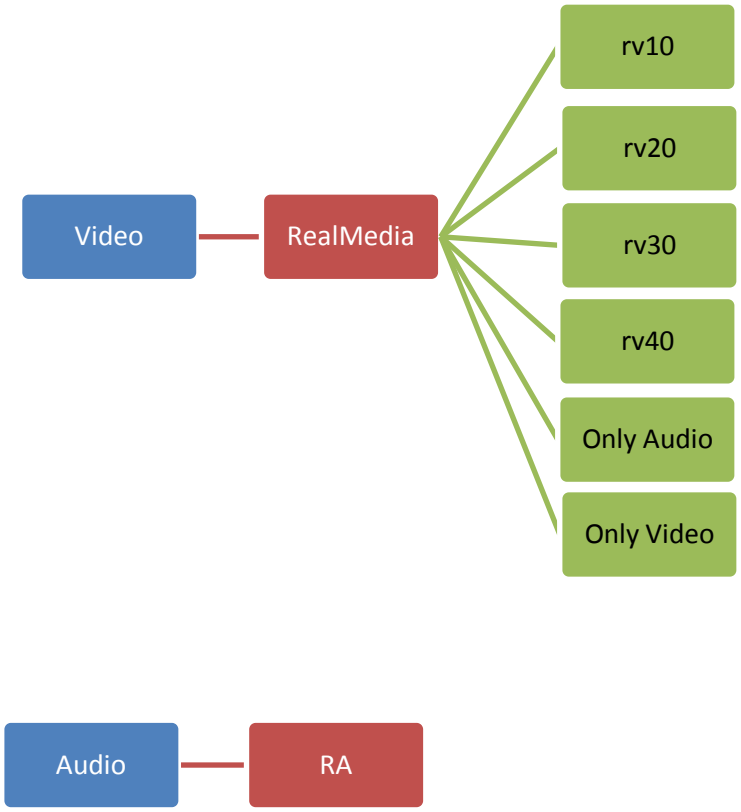
4.10 RealMedia

4.10.1 Container & supported video/audio standard (Codec)

Container	Standard owner	File extension	Video formats supported	Audio formats supported	Comments:
RealMedia (RM)	RealNetworks	.rm, .rmvb (A,V,A+V)	RealVideo 8, RealVideo 9, RealVideo 10	Cook, ACELP.net, AAC-LC, AAC-HE, RealAudio Lossless	1. RealMedia is a proprietary multimedia container format created by RealNetworks. Its extension is ".rm". It is typically used in conjunction with RealVideo and RealAudio and is used for streaming content over the Internet. 2. RealMedia Variable Bitrate (RMVB) is a variable bitrate extension of the RealMedia multimedia container format developed by RealNetworks. As opposed to the more common RealMedia container, which holds streaming media encoded at a constant bit rate, RMVB is typically used for multimedia content stored locally. Files using this format have the file extension ".rmvb".
		.ra (A)		Cook, ACELP.net, AAC-LC, AAC-HE, RealAudio Lossless	

4.10.2. Production Tools: Helix RealProducer

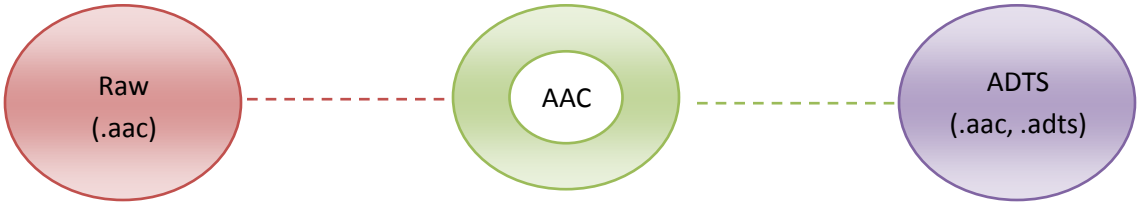
4.10.3. Directory Structure


Video
RealMedia
rv10
rv20
rv30
rv40
Only Audio
Only Video
Audio
RA

4.11 Audio Data Transport Stream (ADTS)

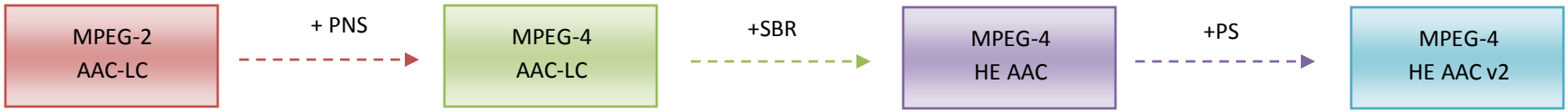
ADTS (Audio Data Transport Stream) is a container for AAC encoded audio. It consists of a series of frames, each frame having a header followed by the AAC audio data. The formats are defined in MPEG-2 Part 7, but are only considered informative by MPEG-4.

There are two common AAC types, raw AAC streams with no headers and AAC in an audio data transport stream (ADTS).



4.11.1. AAC Profiles

Profile	Features	Typical Applications	Typical Bitrate
AAC-LC (Low Complexity AAC)	High performance audio codec for excellent audio quality at low bit rates	– Apple iPod – iTunes – ISDB television broadcasting (Japan)	128 kbit/s (Stereo)
HE-AAC (High Efficiency AAC)	High performance audio codec for good quality at bit rates of 28 kbit/s per channel and below	– XM Radio – Mobile music download – Digital Radio Mondiale	56 kbit/s (Stereo)
HE-AAC v2	Highest performance audio codec for good quality at bit rates below of 24 kbit/s per channel	– 3GPP music download – Digital radio DAB+	48 kbit/s (Stereo)

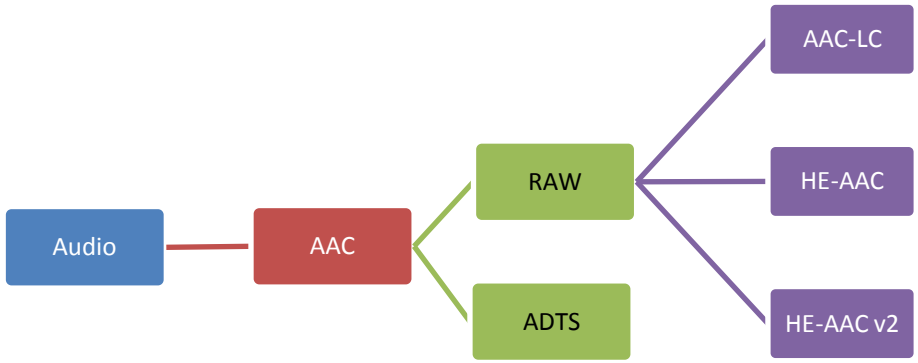


PNS: Perceptual Noise Substitution SBR: Spectral Band Replication PS: Parametric Stereo

AAC IS defined in MPEG-2 Part 7(ISO/IEC 13818-7) and MPEG-4 Part 3(SO/IEC 14496-3)

4.11.2. Production Tools: Nero AAC Codec

4.11.3. Directory Structure



4.12 Waveform Audio File Format (WAV)

WAV or WAVE, short for Waveform Audio File Format, is a Microsoft and IBM audio file format standard for storing an audio bitstream on PCs.

WAV file container format is based on RIFF structure, which commonly used for storing uncompressed (PCM), Wave files can also contain data encoded with a variety of (lossy) codecs (for example the GSM or mp3 codecs).

Directory Structure

5. Elementary Stream (ES)

An elementary stream is a single (video or audio) stream without container. For instance a basic MPEG-2 video stream (.m2v or .mpv) is an MPEG-2 ES, and on the audio side we have AC3, MP2, etc files that are ES.

5.1. Elementary Stream (Video)

Stream	File extension	Video formats	Comments:
Elementary Stream	.mpv, .m1v, m2v, m4v	MPEG-1/MPEG-2/MPEG-4 part2	1. MPV is the umbrella term for raw MPEG video streams that have been demuxed from full transport streams. The video will not have any audio or subtitles. MPEG-1 streams have a .m1v extension and MPEG-2 streams have a .m2v extension. Raw MPEG-4 Visual bit-streams are named .m4v. 2. There are two definitions for the term M4V. The first is that raw MPEG-4 Visual bit-streams are named .m4v. The second and much more likely, is that you have legally down-loaded a video file from the Apple iTunes store and it has the M4V extension. If your file is unprotected, you can change the file extension from ".m4v" to ".mp4“.
	.264, .h264, .avc	H.264/MPEG-4 AVC	
	.vc1	VC-1	

5.2. Elementary Stream (Audio)

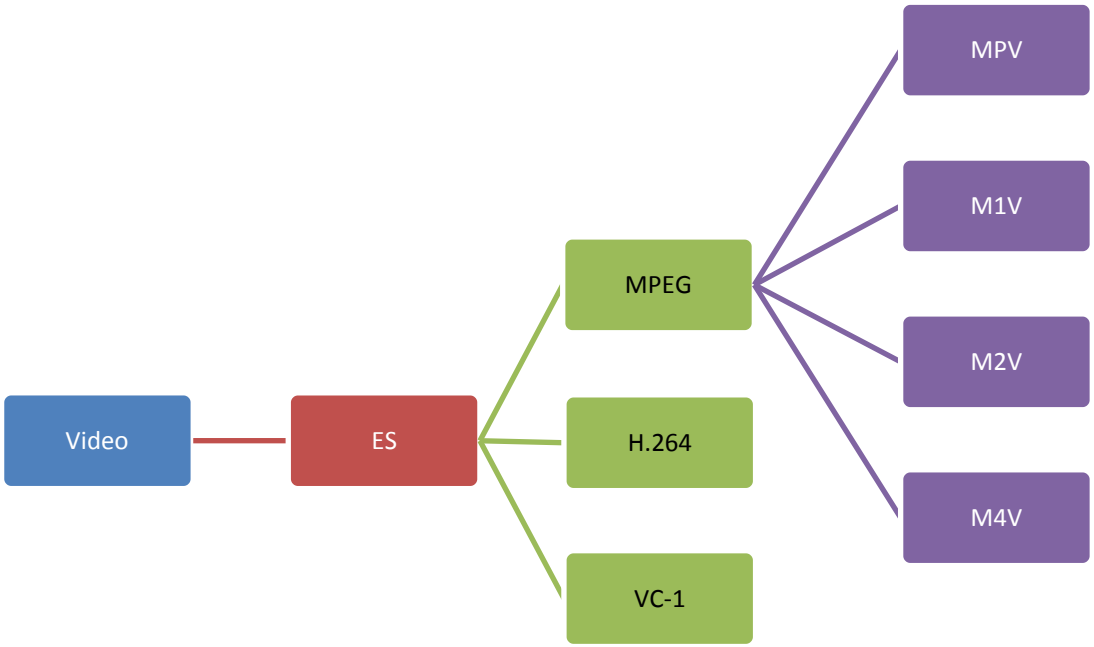
Stream	File extension	Audio formats	Comments:
Elementary Stream	.mpa	MPEG Audio Layer I, II, III	
	.mp2	MPEG 1 Audio Layer II, MPEG 2 Audio Layer II	
	.mp3	MPEG 1 Audio Layer III, MPEG 2 Audio Layer III	
	.ac3	AC3	
		TrueHD	
	.dts	DTS	
		DTS-HD	There are two DTS-HD audio formats: DTS-HD Master Audio and DTS-HD High Resolution Audio. DTS-HD Master Audio is an optional lossless audio format for Blu-ray Disc.
	.flac	FLAC	
	.amr	AMR(NB)	

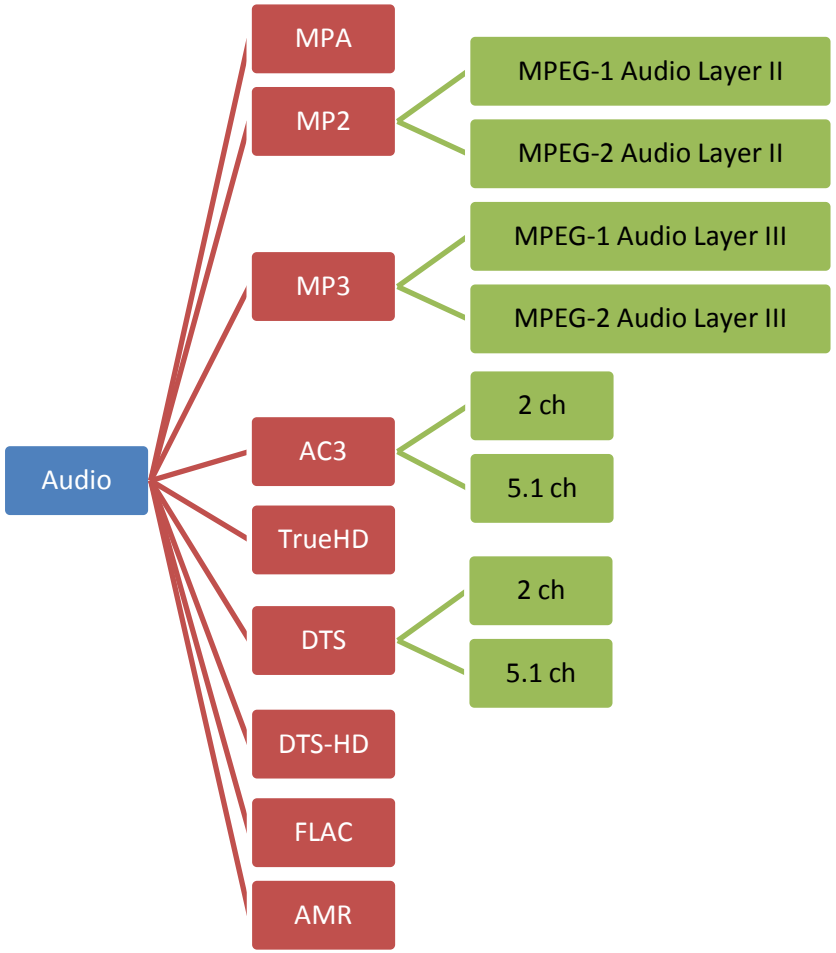
5.3. MPEG Audio

Audio compression format	Sample Rate	Bit rate	Multichannel
MPEG 1 Audio Layer II (MP2)	32, 44.1, 48 kHz	32, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256, 320, 384 kbit/s	No
MPEG 2 Audio Layer II (MP2)	16, 22.05, 24 kHz	8, 16, 24, 32, 40, 48, 56, 64, 80, 96, 112, 128, 144, 160 kbit/s	up to 5 full range audio channels and an LFE-channel (Low Frequency Enhancement channel)
MPEG 1 Audio Layer III (MP3)	32, 44.1 and 48 kHz	32, 40, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256 and 320 kbit/s	No
MPEG 2 Audio Layer III (MP3)	16, 22.05 and 24 kHz	8, 16, 24, 32, 40, 48, 56, 64, 80, 96, 112, 128, 144, 160 kbit/s	No

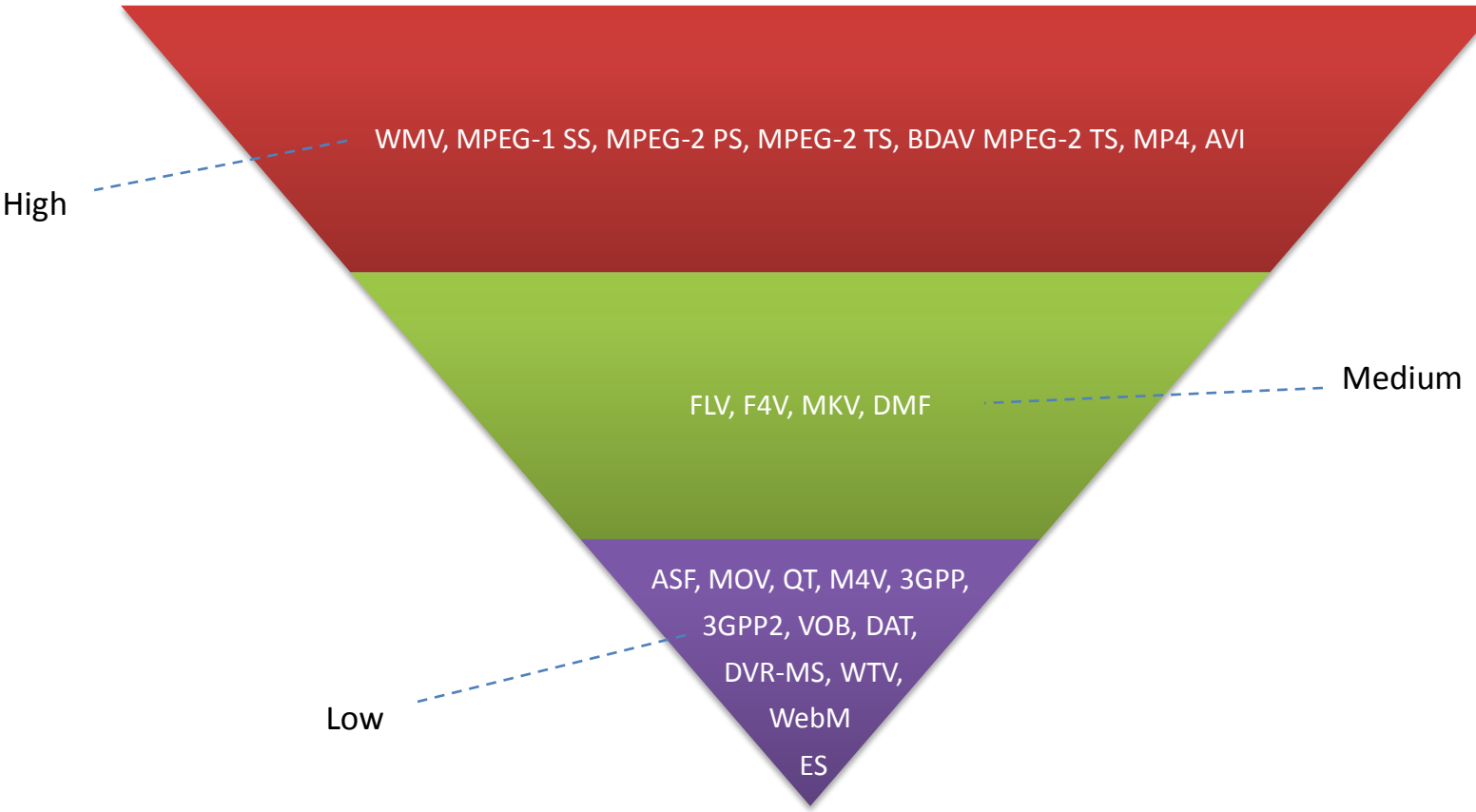
MPEG-1 Audio Layer I, II and III is defined in MPEG-1 Part 3(ISO/IEC 11172-3) and MPEG-2 Audio Layer I, II and III is defined in MPEG-2 Part 3(ISO/IEC 13818-3).

5.4. Directory Structure




Audio
MPA
MP2
MPEG-1 Audio Layer II
MPEG-2 Audio Layer II
MP3
MPEG-1 Audio Layer III
MPEG-2 Audio Layer III
AC3
2 ch
5.1 ch
TrueHD
DTS
2 ch
5.1 ch
DTS-HD
FLAC
AMR

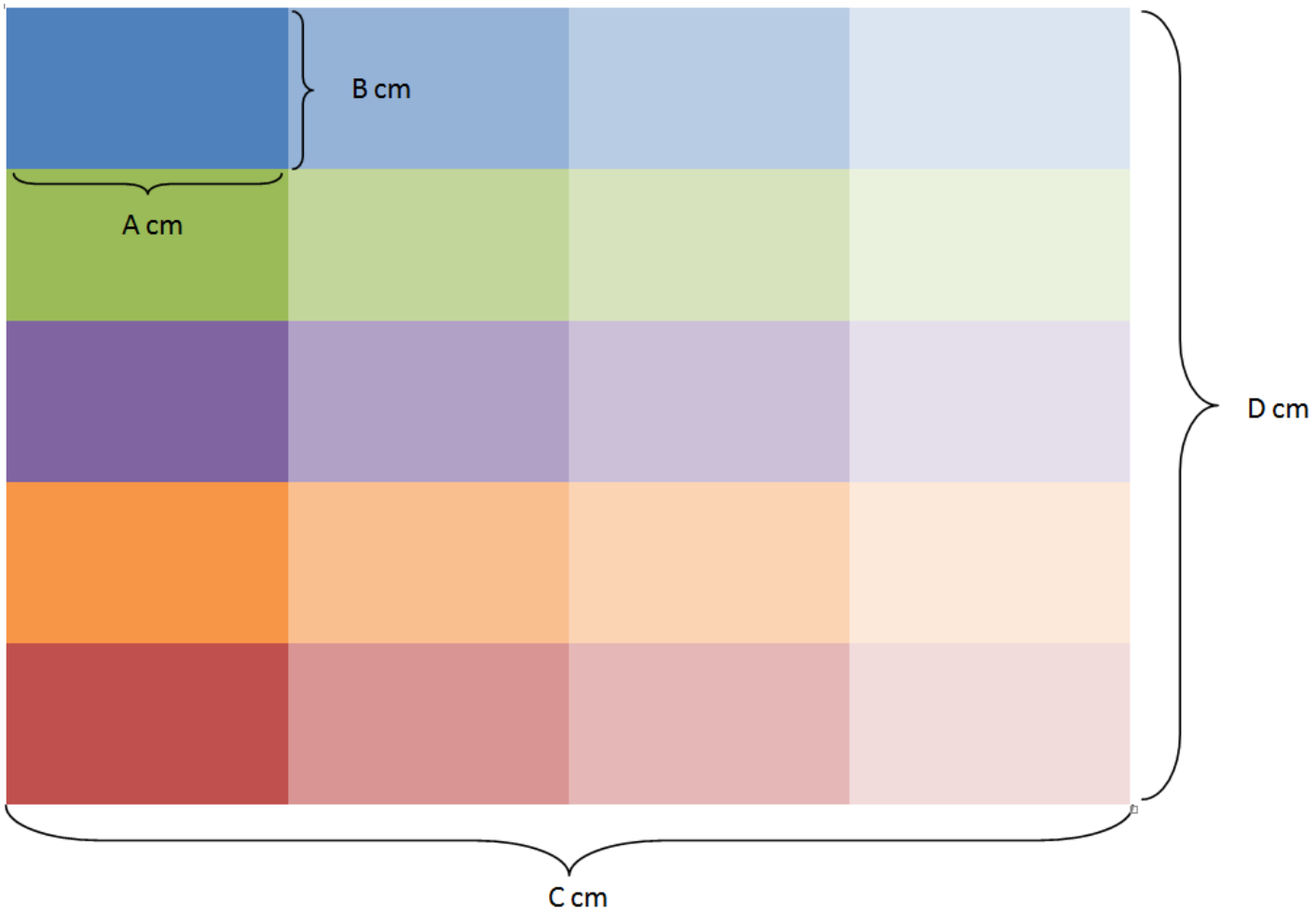
6. Priority



Glossary

Display Aspect Ratio (DAR) and Pixel Aspect Ratio (PAR)

- Display aspect ratio (DAR) describes the ratio of width to height in the dimensions of a frame. For example, DV NTSC has a frame aspect ratio of 4:3. A typical widescreen frame has a frame aspect ratio of 16:9.
- Pixel aspect ratio (PAR) describes the ratio of width to height of a single pixel in a frame.
- Shown below is a diagram of a frame which's resolution is 4x5 Pixels, there is a relationship between DAR and PAR:
- 4 X A = C, 5 X B = D => C/D (DAR) = 4/5 (Width/Height) X A/B (PAR)
- More info please refer to http://msdn.microsoft.com/en-us/library/bb530115(v=VS.85).aspx



A: Pixel Width

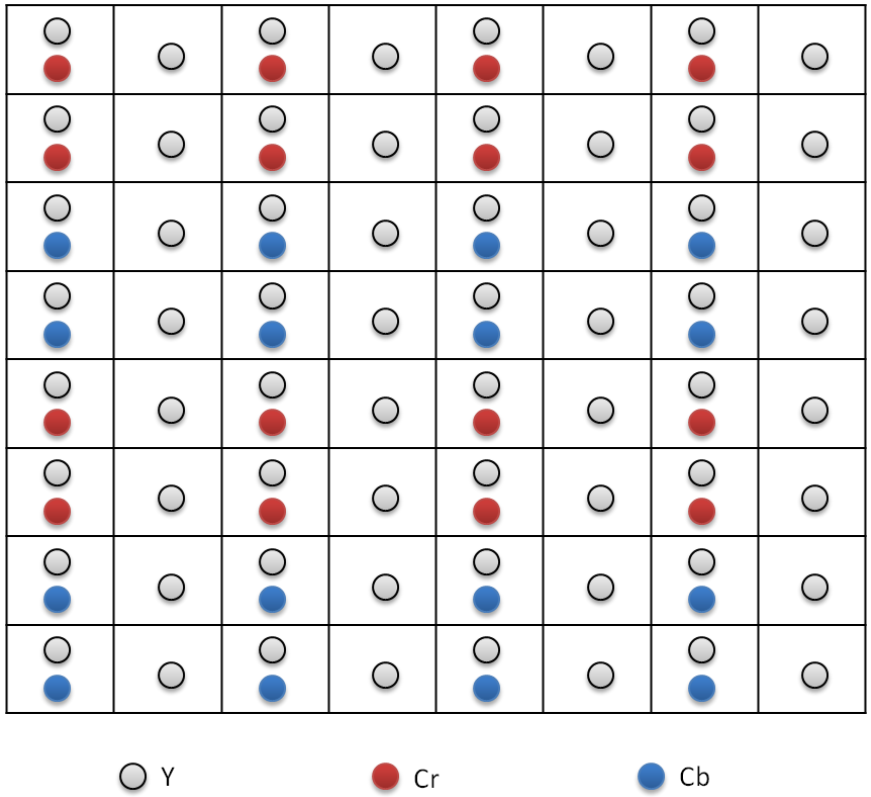
B: Pixel Height

C: Frame Width

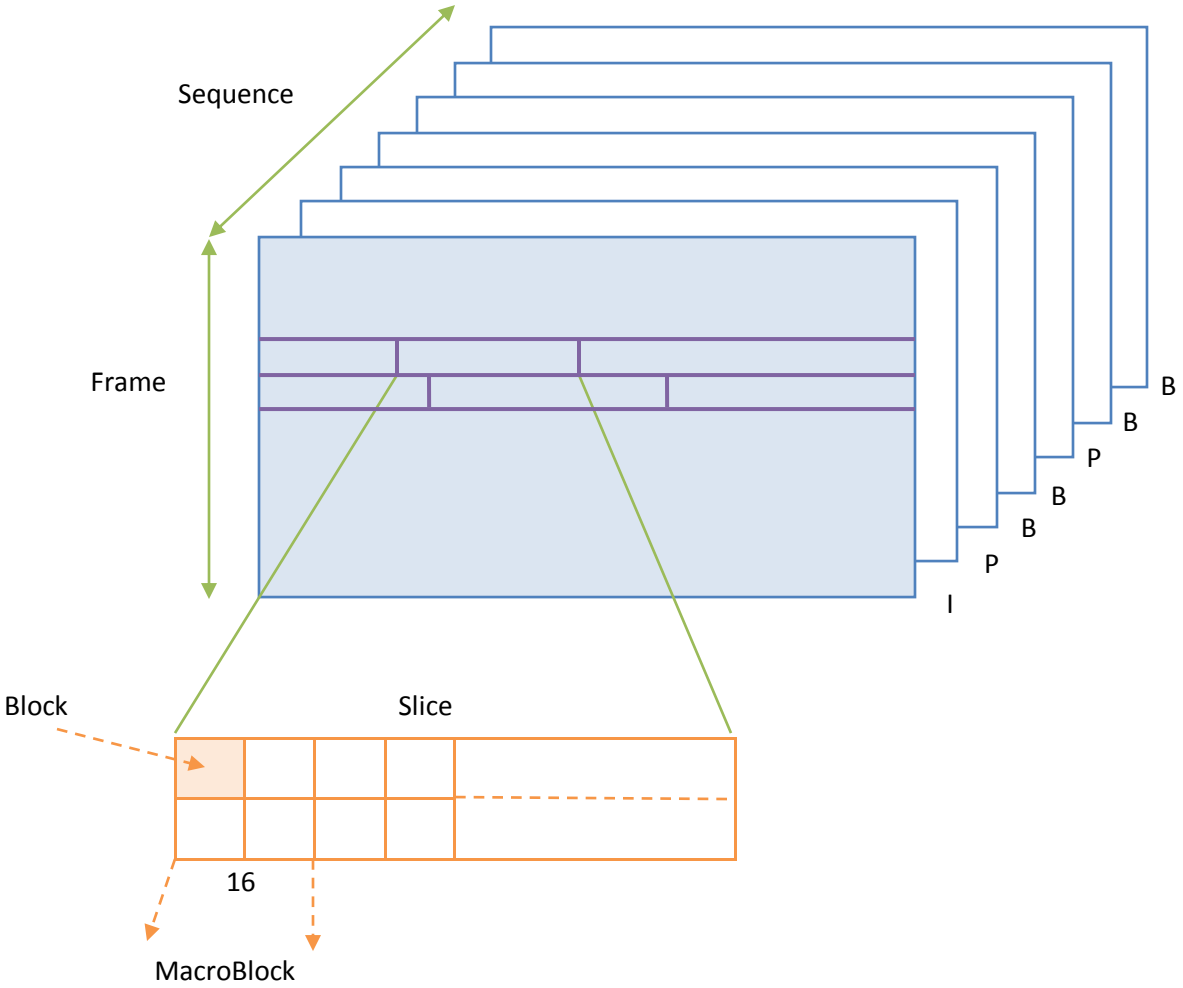
D: Frame Height

Chroma Subsampling

- Research into the Human Visual System has shown that the eye is most sensitive to changes in luminance, and less sensitive to variations in chrominance, so it makes sense that MPEG should operate on a color space that can effectively take advantage of the eye different sensitivity to luminance and chrominance information. As such, MPEG uses the YCbCr color space to represent the data values instead of RGB, where Y is the luminance signal, Cb is the blue color difference signal, and Cr is the red color difference signal.
- Shown below is a diagram of MPEG Chroma Subsampling 4:2:0

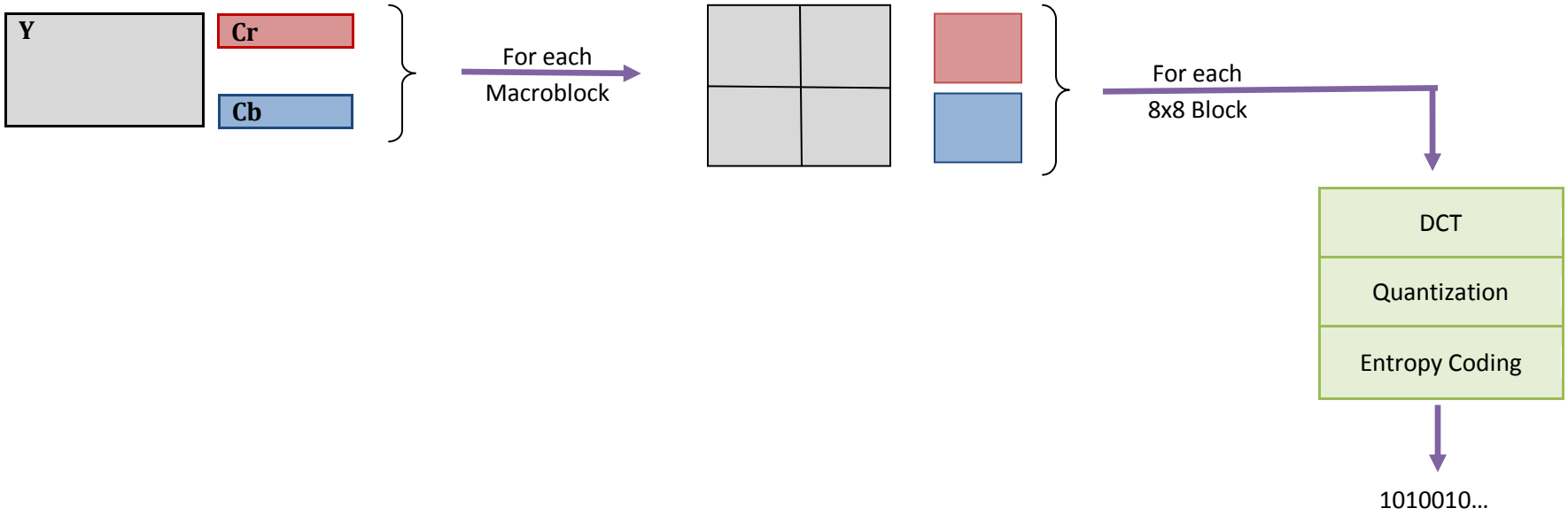


MPEG-2 Video Structure



I Frame

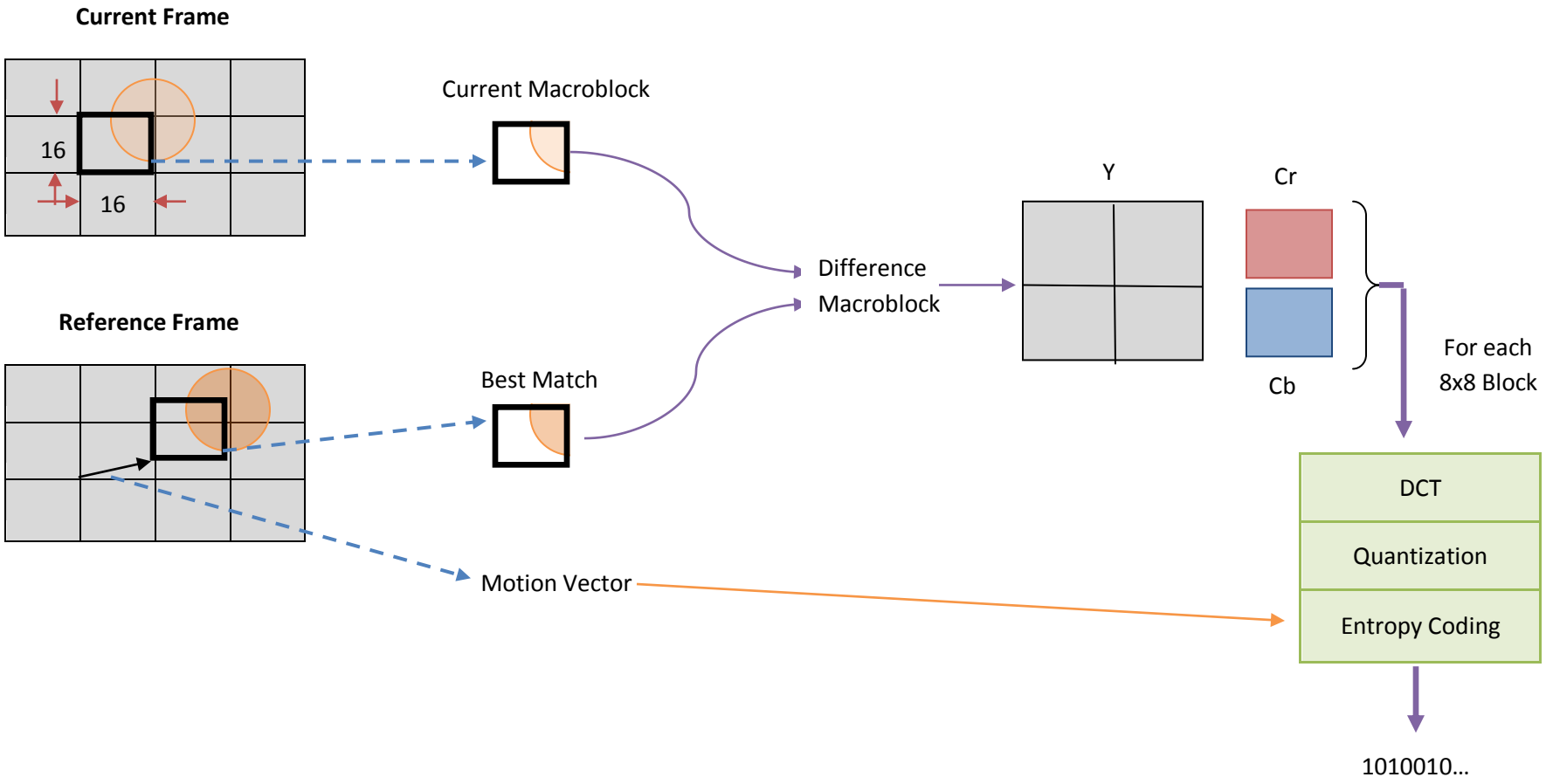
- An Intra frame, or I-Frame, is coded using only information present in the frame itself and not relative to any other frame in the video sequence. No temporal processing is performed outside of the current frame. I frame provides potential random access points into the compressed video data.



- Frame coding is depicted in above diagram. Here Macroblocks are of size 16 × 16 pixels for the Y frame, and 8 × 8 for Cb and Cr frames, since 4:2:0 chroma subsampling is apllied. A macroblock consists of four Y, one Cb, and one Cr 8 × 8 blocks. For each 8 × 8 block a DCT transform is applied, the DCT coefficients then go through quantization zigzag scan and entropy coding.

P Frame

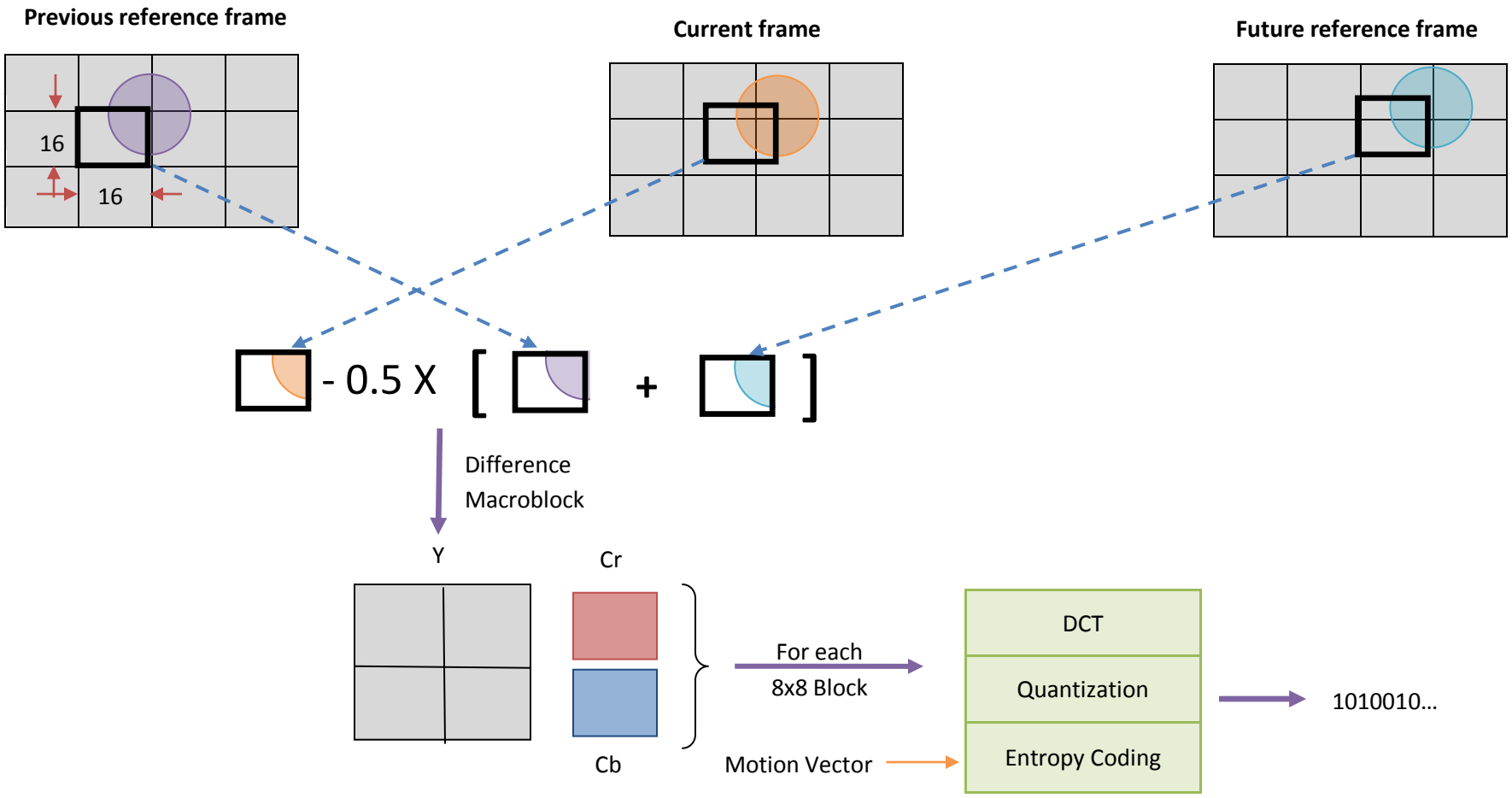
- A Predicted frame, or P-Frame, is coded with respect to the nearest previous I- or P-frame. This kind of prediction tries to take advantage from temporal redundancy between neighboring frames allowing achieving higher compression rates.



- Above diagram gives an overview on P-frame coding. The P-frame coding encodes the difference macroblock (not the Target macroblock itself). Sometimes, a good match cannot be found, i.e., the prediction error exceeds a certain acceptable level. The MB itself is then encoded (treated as an Intra MB) and in this case it is termed a non-motion compensated MB. For motion vector, the difference MVD is sent for entropy coding.

B Frame

- A Bidirectional frame, or B-frame, uses both a past and future frame as a reference.



- Each Macroblock from a B-frame will have up to two motion vectors (MV's) (one from the forward and one from the backward prediction). If its matching in both directions, then two MVs will be sent and the two corresponding matching MBs are averaged before comparing to the Target MB for generating the prediction error. If an acceptable match can be found in only one of the reference frames, then only one MV and its corresponding MB will be used from either the forward or backward prediction.

Profiles & Levels

- Profile is sub-set of the entire bitstream syntax (for example 4:2:0 only, I/P frames only, field DCT, etc). The profile determines the codec features that are available, and thereby determines the required decoder complexity.
- Level is a defined set of constraints imposed on parameters in the bitstream (for example max Picture Size, max BitRate, max Vertical Motion Vector, max Buffer Size, etc).

CBR/VBR

- CBR (Constant Bit Rate) encoding is designed to work optimally in a variety of streaming scenarios. You can constrict the bit rate to guarantee consistent playback across a wide range of systems. The bit rate remains fairly constant and close to the target bit rate over the course of the stream. The disadvantage of CBR encoding is that the quality of the encoded content is not constant. Because some pieces of content are more difficult to compress than others, some parts of a CBR stream are of lower quality than others. In addition, CBR encoding gives you inconsistent quality from one stream to the next. In general, quality variations are more pronounced at lower bit rates.
- VBR (Variable Bit Rate) encoding is designed to work optimally in high bandwidth scenarios and is especially suited for encoding content that is a mixture of simple and complex data. The encoder allocates fewer bits to the simple parts of the content, leaving enough bits available to produce good quality for the more complex portions. Content with consistent data (for example, a "talking head" news story) would not benefit from VBR encoding. However, when it is used on content with varying complexity, VBR encoding produces a much better output than CBR encoding, even if both methods produce files of identical size. In some cases, a VBR-encoded file might have the same quality as a CBR-encoded file that is twice as large because the VBR encode compressed the less complex portions much more than the CBR method did. However, VBR would give more bandwidth to complex portions, unlike CBR encoding.

One Pass/Two Pass

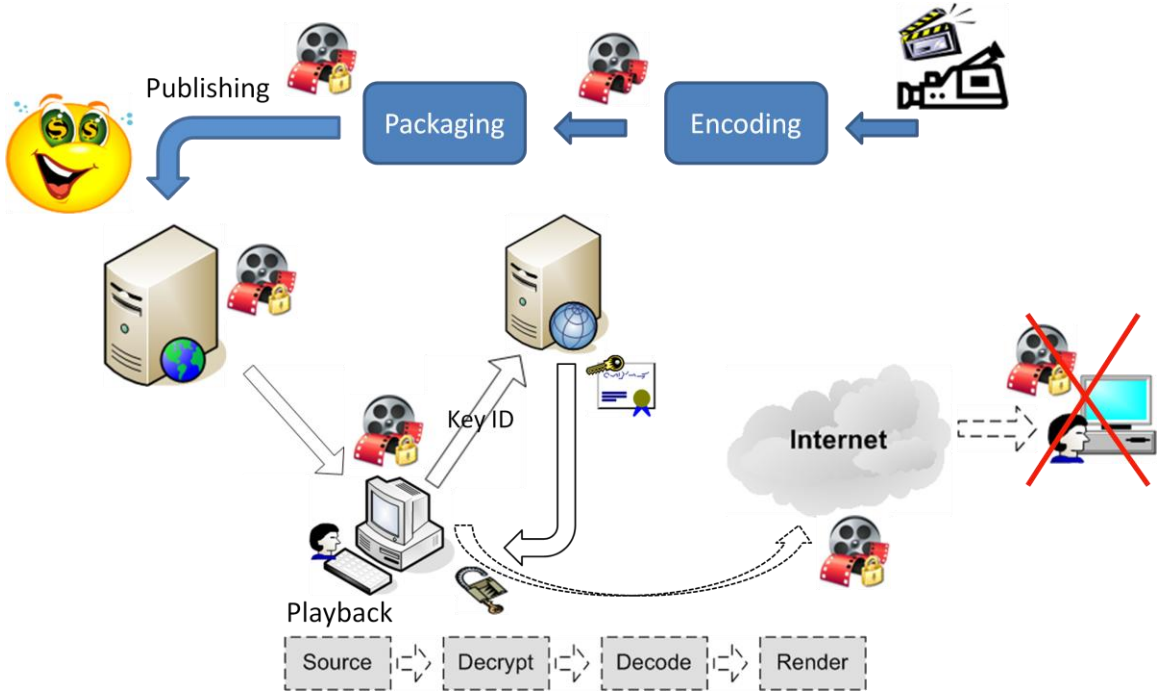
- One pass VBR has a moving window over which it analyzes the video. That is, the encoder looks ahead a second or two decides how complex the video is and allocated bits based on that information. Because it is only looking ahead a short time in the future it has no idea how complex the video is much later in the movie so the allocation is not as optimal as it could be
- Two pass VBR analyzes the entire movie for video complexity in the first pass. It then stores this information. It can take advantage of knowing the variation in complexity of the entire movie and allocate bits more effectively than the 1-pass approach.
- For example, perhaps the beginning of the movie is all very high action but the end is people sitting quietly talking. The high action is much more difficult to encode and should be allocated many more bits than the talking part of the movie. The two pass algorithm can take advantage of knowing the complexity of the entire movie while the one pass cannot.

CAVLC/ CABAC

- AVC/H.264 defines two, more advanced tools for entropy coding of the bitstream syntax than MPEG-4 ASP: Context-Adaptive Variable Length Coding (CAVLC) and Context-Adaptive Binary Arithmetic Coding (CABAC).
  CABAC, compared to CAVLC which is the default method in AVC/H.264, is a more powerful compression method, being said to bring down the bitrate additionally by about 10-15% (especially on high bitrates). CABAC (as CAVLC) is a lossless method and therefore will never hurt the quality, but will slow down encoding and decoding.
- CAVLC is supported in all H.264 profiles, unlike CABAC which is not supported in Baseline and Extended profiles.

DRM

- Digital rights management (DRM) is a generic term for access control technologies that can be used by hardware manufacturers, publishers, copyright holders and individuals to try to impose limitations on the usage of digital content and devices.
- How DRM Works?



Multiple-bit-rate

- If you plan to deliver your content as a unicast stream from a Windows Media server, you can encode a multiple-bit-rate (MBR) stream. This provides users with better quality content during times of network congestion. When MBR content is received by a player, only the bit rate that is the most appropriate for network bandwidth conditions is streamed. The process of selecting the appropriate stream is handled by the Windows Media server and the player and is invisible to the user.
- When streaming single-bit-rate streams or files, a Windows Media server is designed to handle network congestion smoothly. If congestion occurs during the broadcast, the stream is "thinned", which means that the frame rate is reduced. If this is insufficient, the video portion of the stream is frozen and only the audio portion is streamed.

Reference

- Official MPEG http://mpeg.chiariglione.org/
- Blu-ray White Papers http://www.blurayjukebox.com/blu_ray_whitepapers.html
- MPEG Industry Forum http://www.mpegif.org/
- MSDN DirectShow http://msdn.microsoft.com/en-us/library/dd375454(v=VS.85).aspx
- MSDN Media Foundation http://msdn.microsoft.com/en-us/library/ms694197(v=VS.85).aspx
- Doom9 http://www.doom9.org/
- Afterdawn http://www.afterdawn.com/
- Video Help http://www.videohelp.com/
- MultimediaWiki http://wiki.multimedia.cx/index.php?title=Main_Page
- Wikipedia http://en.wikipedia.org/wiki/Main_Page
- Streaming Learning Center http://www.streaminglearningcenter.com/
- FFmpeg for Windows Help http://ffmpeg.arrozcru.org/
- Samples from MPlayer/FFmpeg http://samples.mplayerhq.hu/
- Streams from VideoLAN ftp://streams.videolan.org/streams-videolan/

- GDCL http://www.gdcl.co.uk/
- DVD-Video Information http://dvd.sourceforge.net/dvdinfo/index.html
- VBrick White Papers http://www.vbrick.com/support/white_papers.asp
- Moviola http://edu.moviola.com/edu/rc/kz
- Matroska Media Container http://www.matroska.org/
- Haali Media Splitter http://haali.su/mkv/
- Nero AAC Codec http://www.nero.com/eng/downloads-nerodigital-nero-aac-codec.php
- Wave File Specifications http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/WAVE.html
- Huifang Sun, Xuemin Chen, Tihao Chiang, Digital Video Transcoding for Transmission and Storage, December 28, 2004
- Hervé Benoît, Digital television: MPEG-1, MPEG-2 and principles of the DVB system, September 24, 2002
- 陆其明， DirectShow 实务精选 ，July， 2004
- 陆其明，DirectShow 开发指南 ，December， 2003

Thanks

This work was pushed by Ming Luo.

I gained many opinions, suggestions, and feedback in this document from following colleagues: Jianqi Zhu, Yan Liu, Liu Hu, Weiping He and Qinghua Chen. All practical work (Production) will be done by Zhaoqing Li, Ge Chen and Me.

Here I just want to say from the bottom of my heart, Thank you, Thank you all for your support and help.