Technology Seminar

MPEG Video Technology

Kit Barritt
Principal Lecturer
Sony Training Centre

SONY

Content

- Standards
- Frame Structure
- Frame Types
- Frame Sequences
- Difference Frames
- Motion Prediction

Content

- MPEG Encoding
  - Inter-frame Process
    - I/P/B Frames
    - Motion Prediction
  - Intra-frame Process
    - DCT
    - Zig Zag Scan & Run length encoding
    - Entropy coding
    - Quantisation

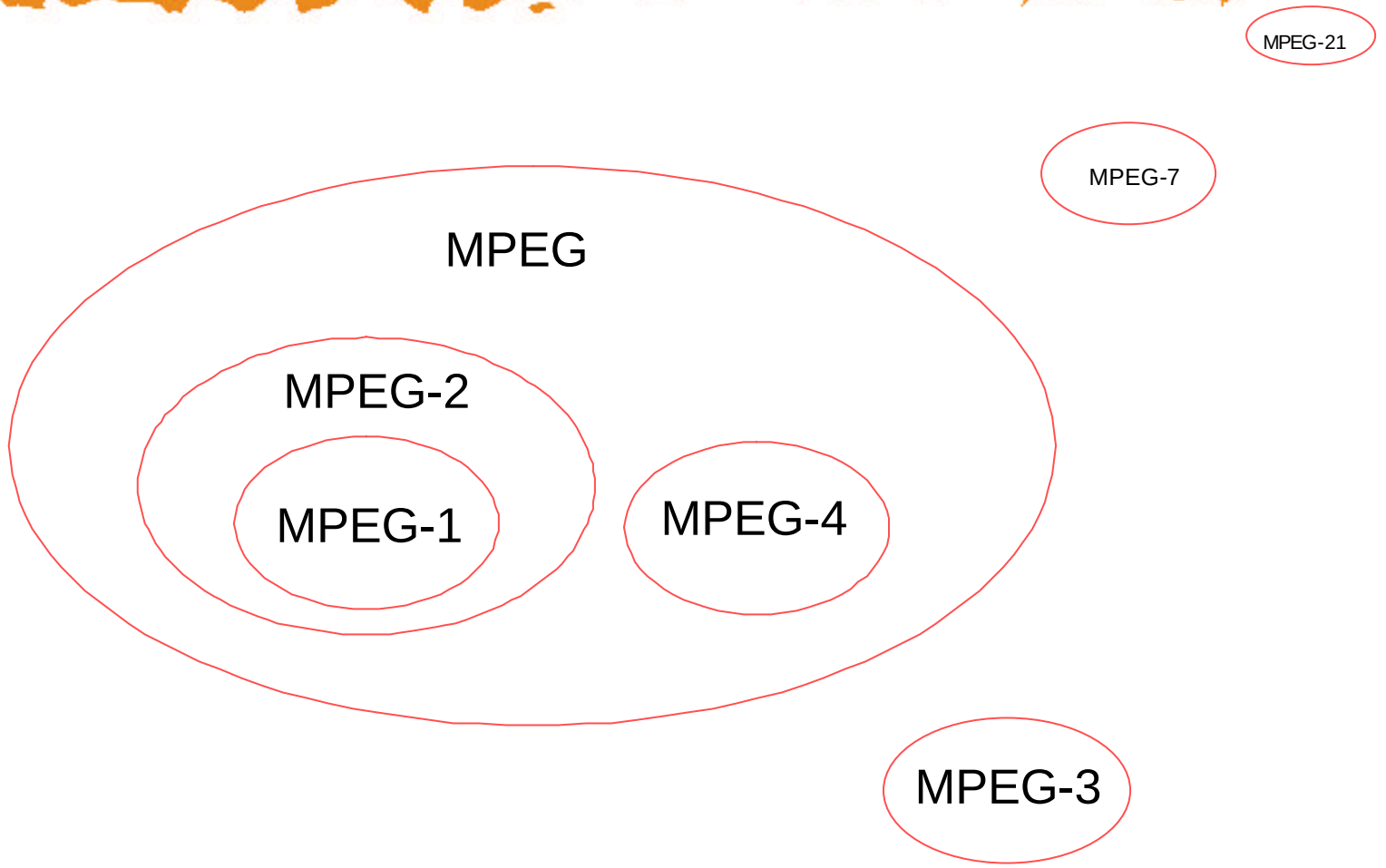
MPEG

- Edit techniques
- Multi-generation considerations
- Compression System Comparisons
  - MPEG-2 422P@ML
  - DVC
- MPEG System Layer
  - MPEG Transmission
    - Programme Stream
    - Transport Stream
    - PID, PAT & PMT
  - Time Clocks

Content

- MPEG-4
  - history
  - Audio coding
  - Video coding

The MPEG Standards Set



SONY

Levels and Profiles

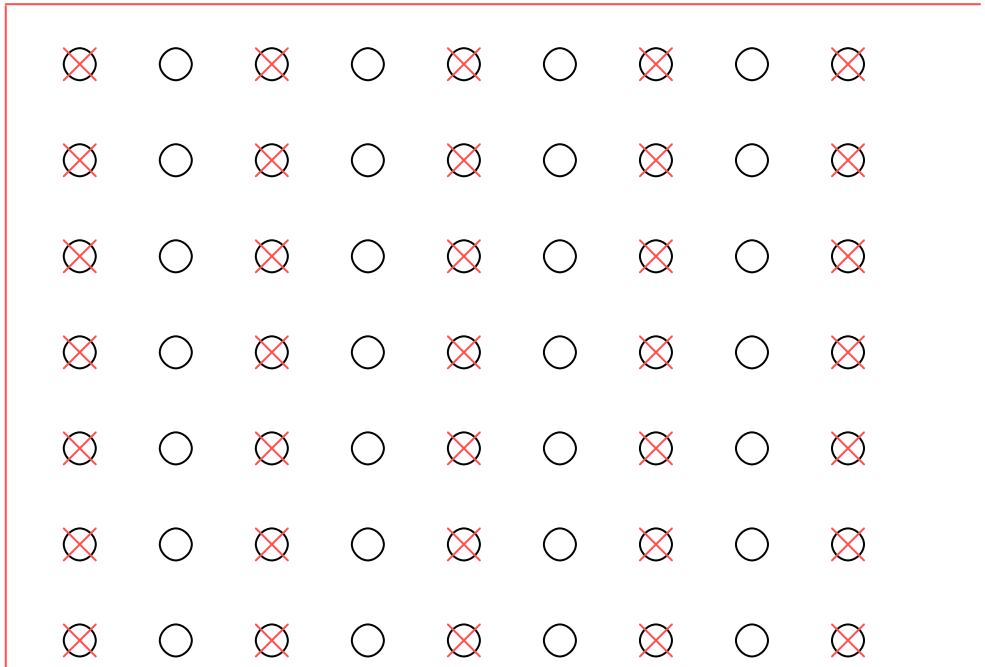
- MPEG-2 Profile @ Level
- Profiles
  - Simple, Main, SNR, Spatial, High, 422 Profile
- Levels
  - High, High-1440, Main, Low

Levels and Profiles

	Profile	Simple	Main	SNR	Spatial	High	422
	Frame Types	I & P	I, P & B	I, P & B	I, P & B	I, P & B	I, P & B
Level	Chroma Sampling	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0 or 4:2:2	4:2:0 or 4:2:2
High	Samples/line Lines/frame Frames/sec Max Bit-rate (MBps)		1920 1152 60 80			1920 1152 60 100	1920 1088 60 300
High 1440	Samples/line Lines/frame Frames/sec Max Bit-rate (MBps)		1440 1152 60 60		1440 1152 60 60	1440 1152 60 80	
Main	Samples/line Lines/frame Frames/sec Max Bit-rate (MBps)	720 576 30 15	720 576 30 15	720 576 30 15		720 576 30 20	720 608 30 50
Low	Samples/line Lines/frame Frames/sec Max Bit-rate (MBps)		352 288 30 4	352 288 30 4			

MPEG-2 Sample Structure

4:2:2



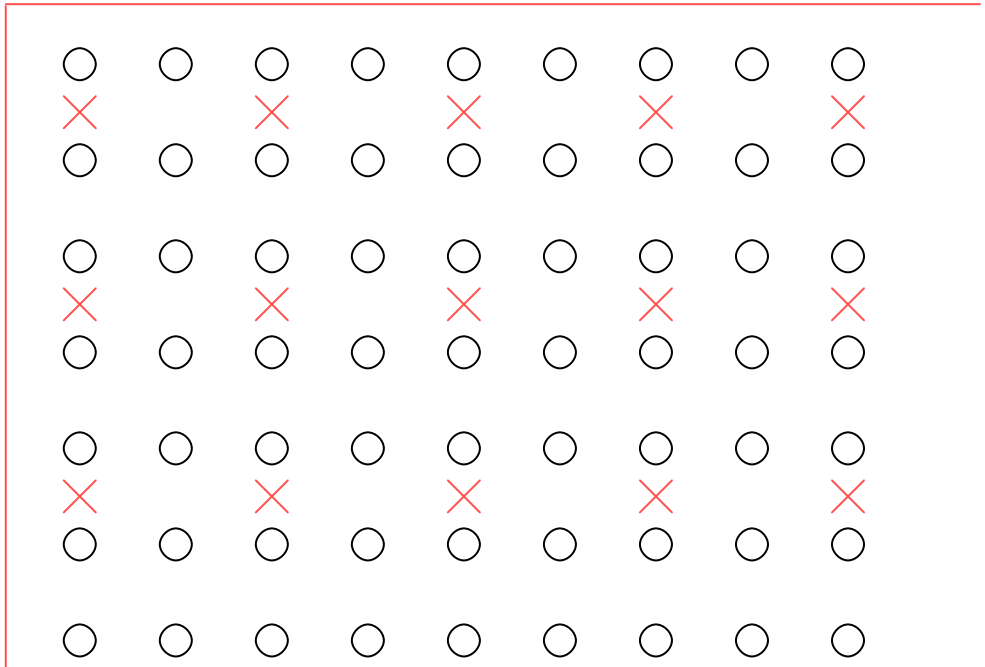
○ Represent Luminance Samples

⊗ Represent Chrominance Samples

SONY

MPEG-2 Sample Structure

4:2:0

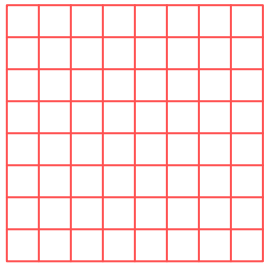


○ Represent Luminance Samples

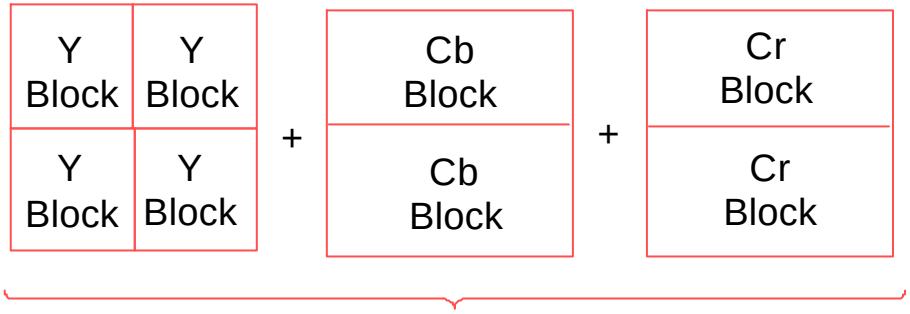
× Represent Chrominance Samples

Skip

MPEG Block Structure



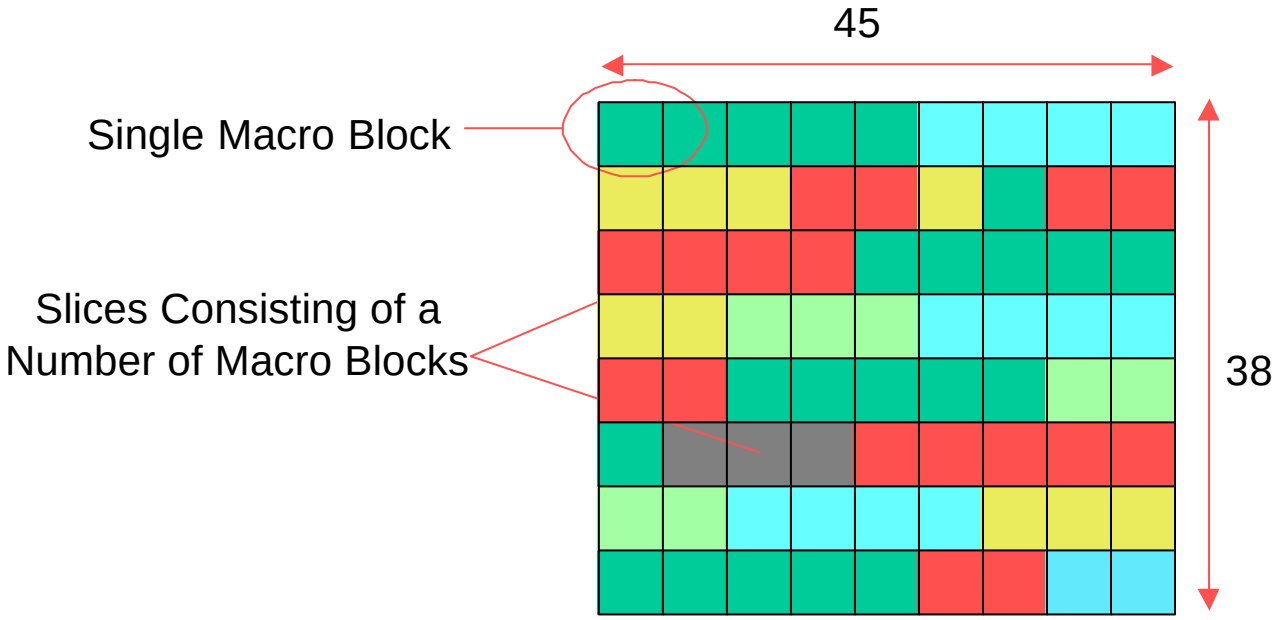
8 x 8 Pixel Block
Y, Cr or Cb



422P @ ML Macro Block consisting of
Y, Cr & Cb Blocks

SONY

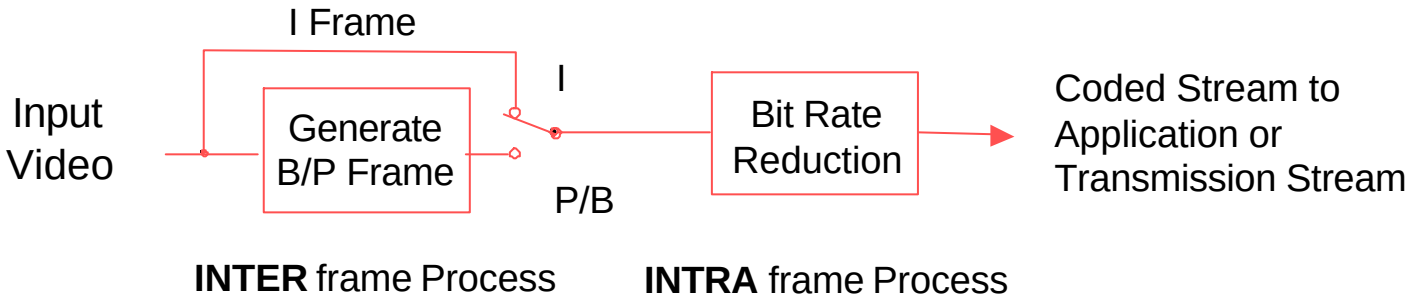
MPEG Frame Structure



An MPEG 422P @ ML Frame of 45 x 38 Macro Blocks

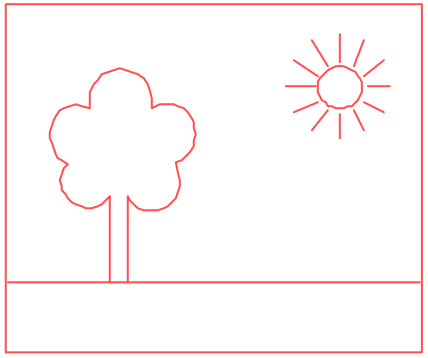
SONY

Inter-field and BRR Process

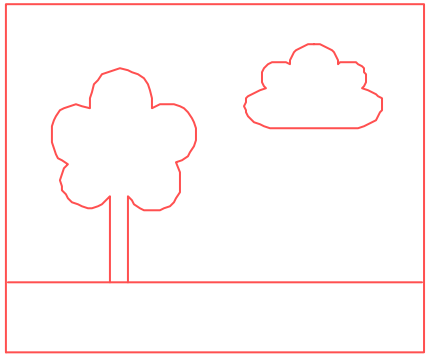


SONY

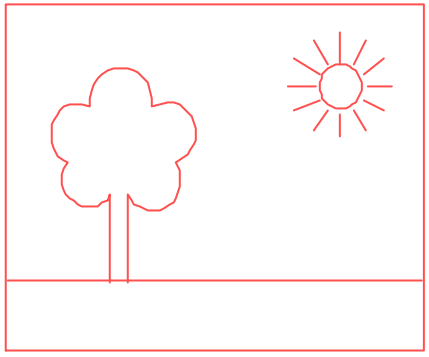
Prediction Frame



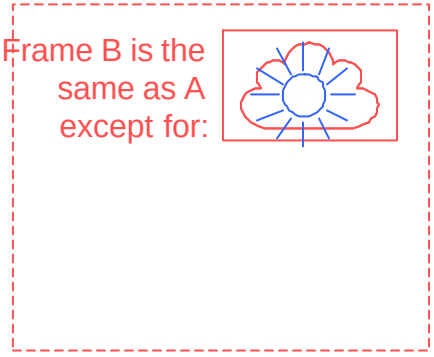
Source Frame A



Source Frame B



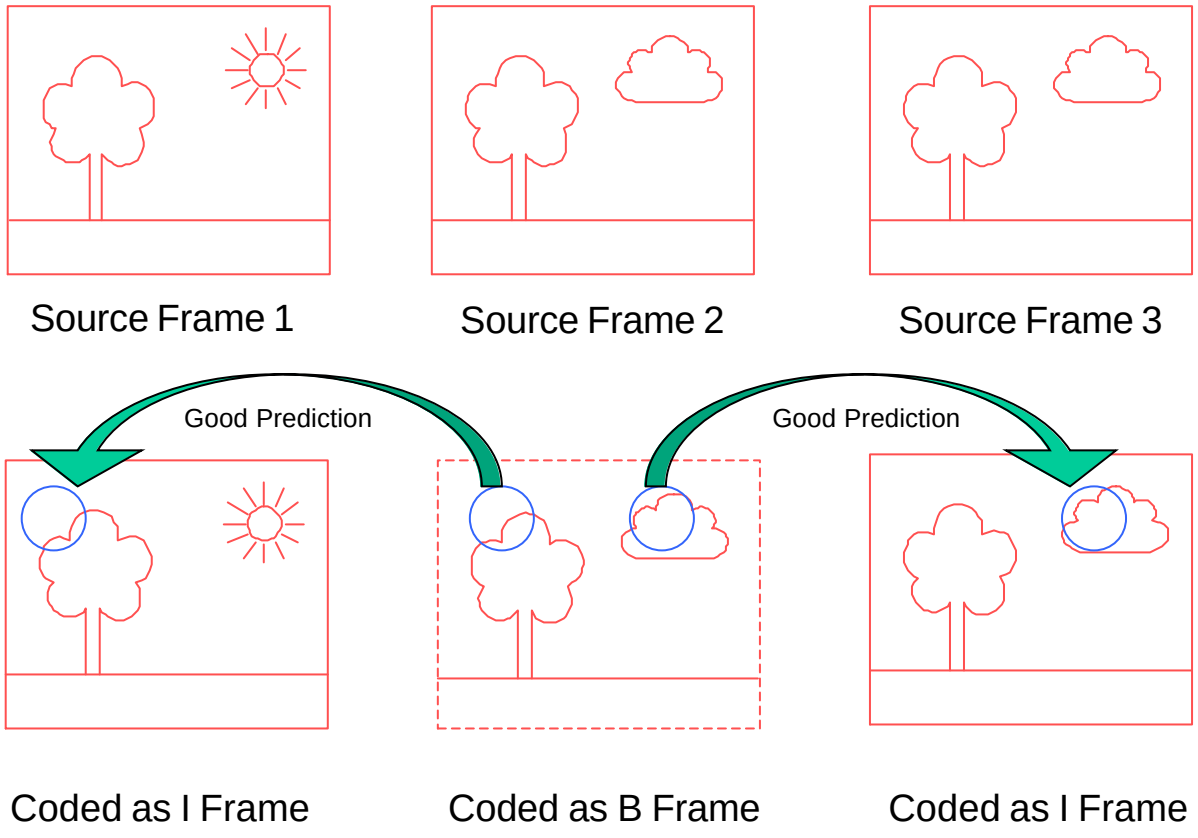
Frame A is encoded as a I Frame



Frame B is encoded as a P Frame

SONY

Prediction Frame

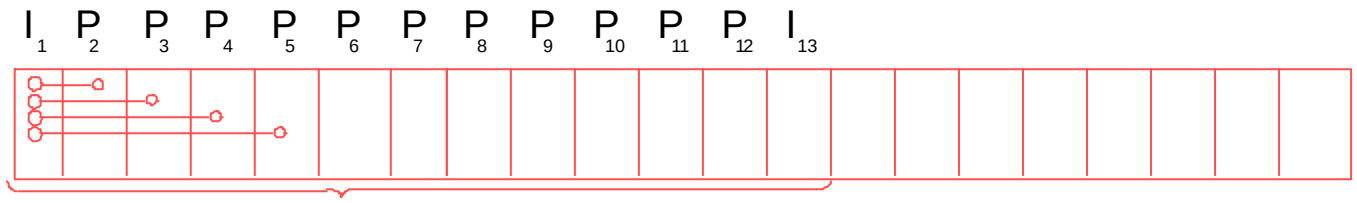


SONY

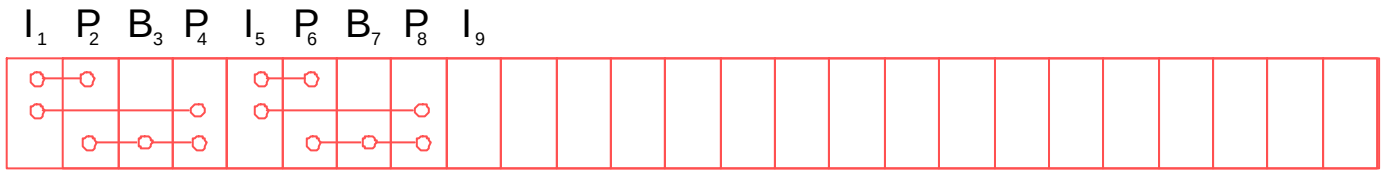
Frame Types

- I Frame
  - Intra frame compression
- P Frame
  - Forward Prediction frame
    - Typically 30% Size of I frame
- B Frame
  - Bi-directional Prediction
    - May refer to I or P frames before or after the coding frame
    - Typically 50% the size of P frame (I.e 15% size of I frame)

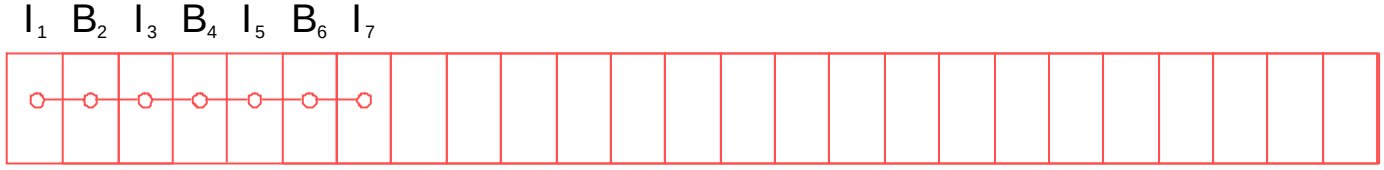
MPEG Frame Sequence



12 Frame GOP



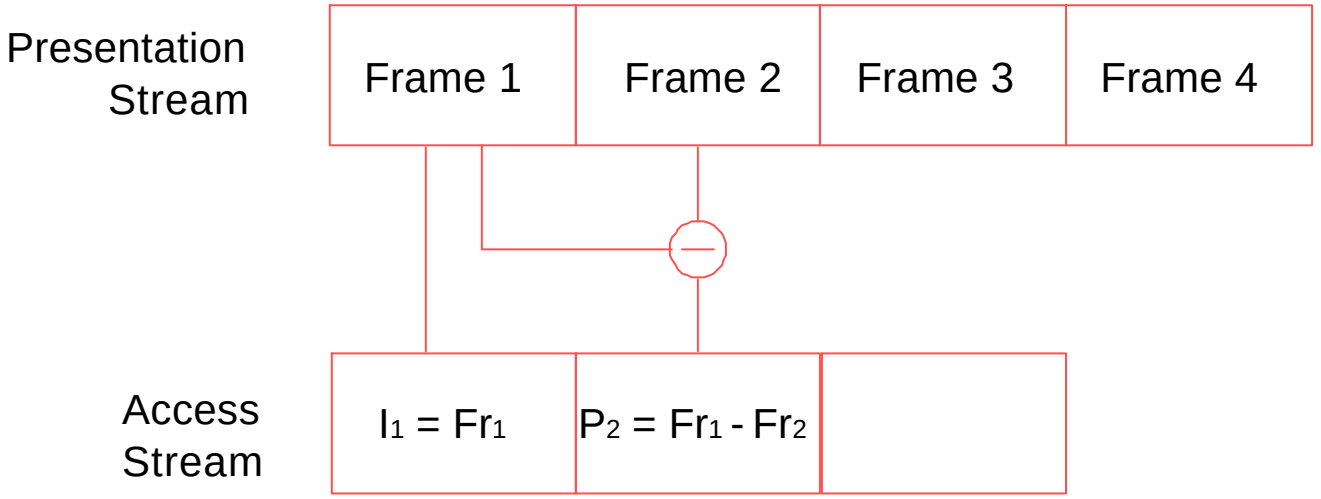
4 Frame GOP



2 Frame GOP

SONY

MPEG P Frame Generation



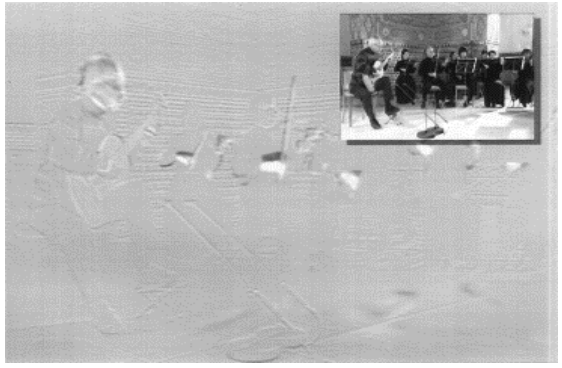
SONY

Difference Frame

Frame N

Frame N+1

(Frame N) - (Frame N+1)



Difference Frame

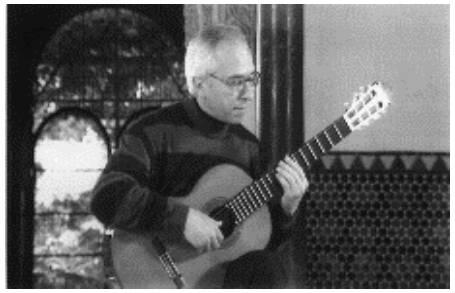
Frame N

Frame N+1

(Frame N) - (Frame N+1)

SONY

Difference Frame

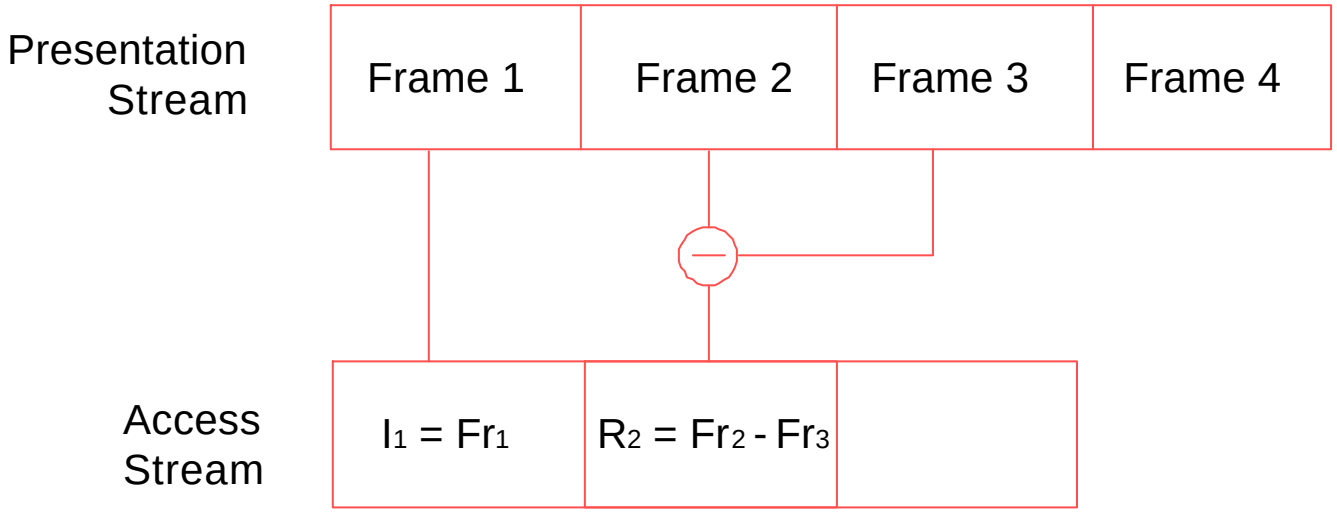


Frame N

Frame N+1

(Frame N) - (Frame N+1)

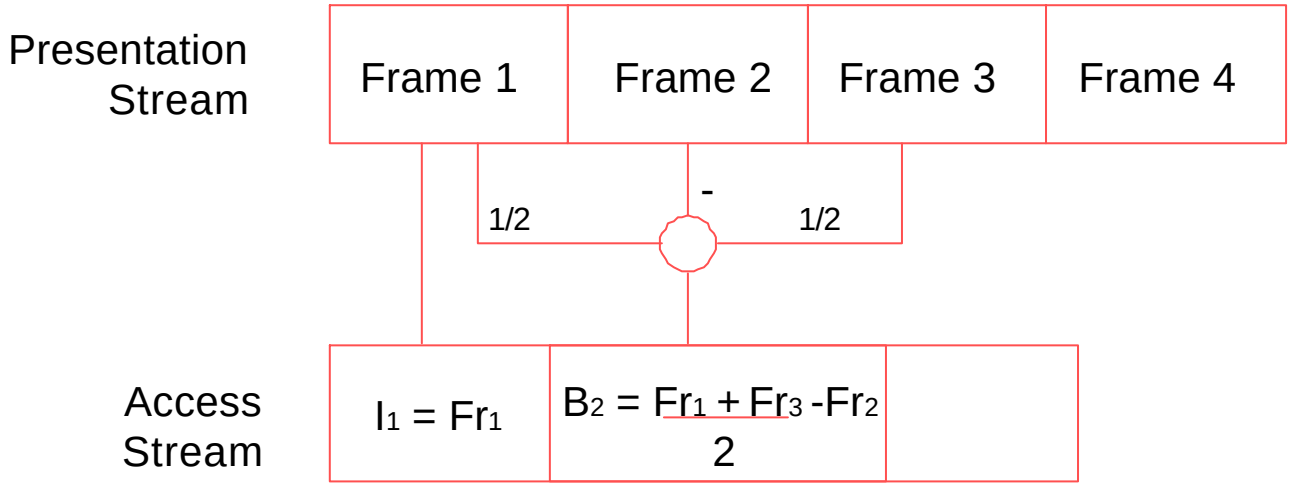
R Frame Generation



Note: MPEG does not support R Frames, they are a type of B Frame

SONY

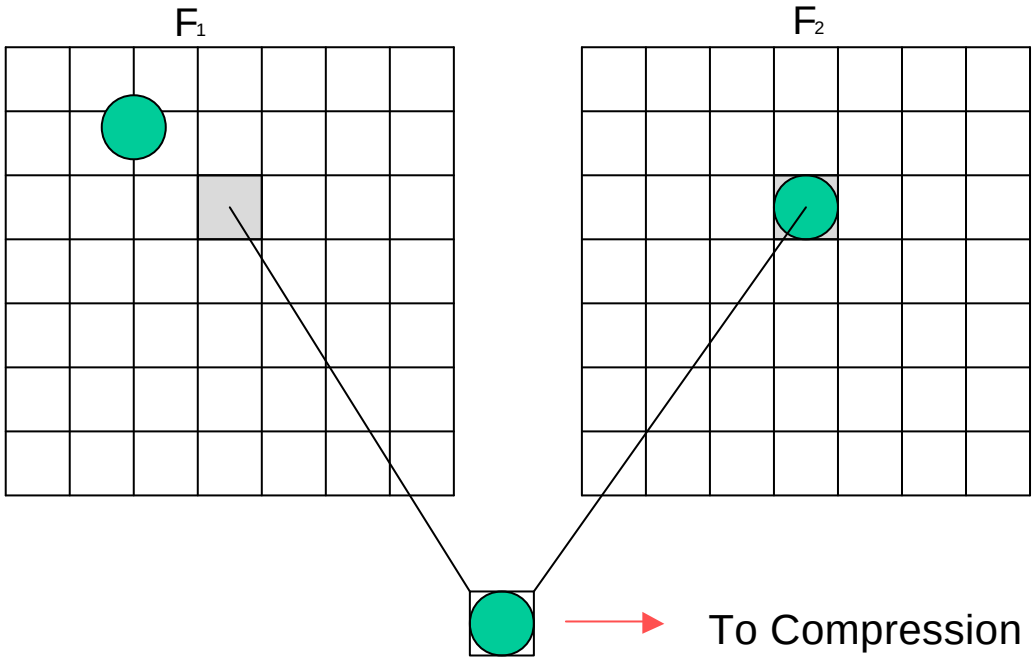
B Frame Generation



Note: B frames can refer to I frames or P frames
Motion Prediction can refer to earlier or later frames

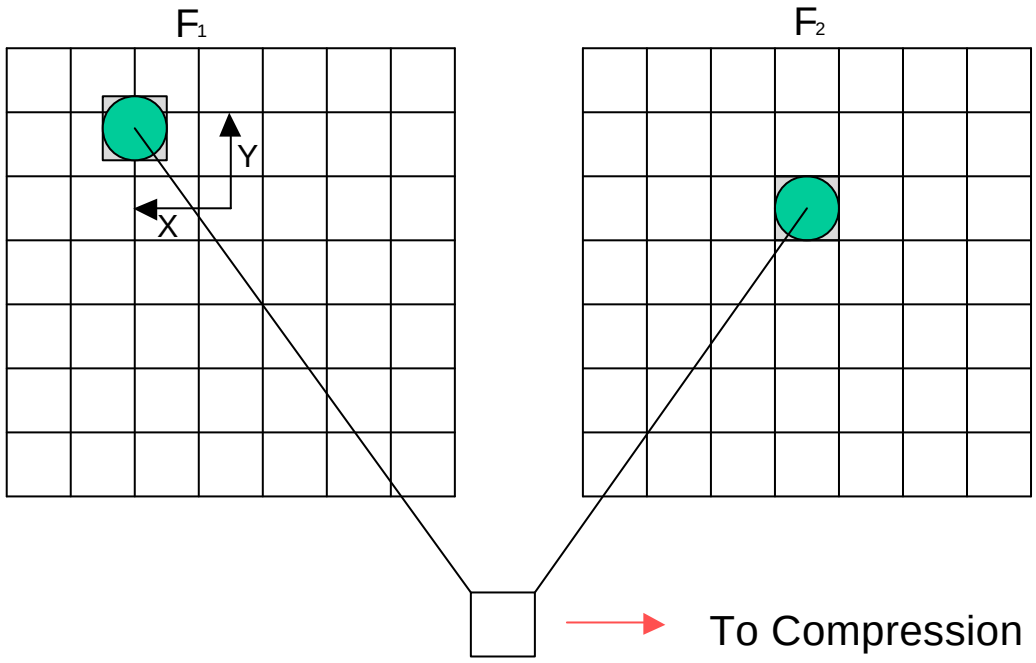
SONY

Motion Prediction



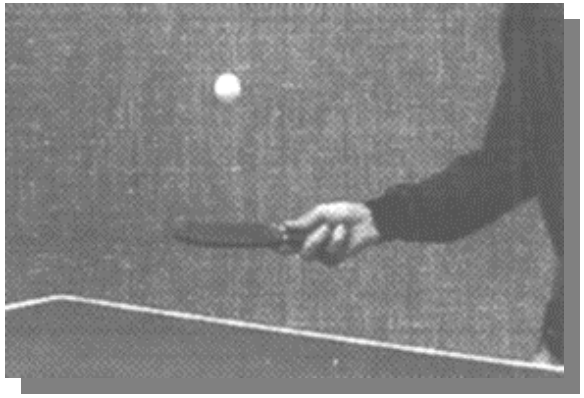
SONY

Motion Prediction

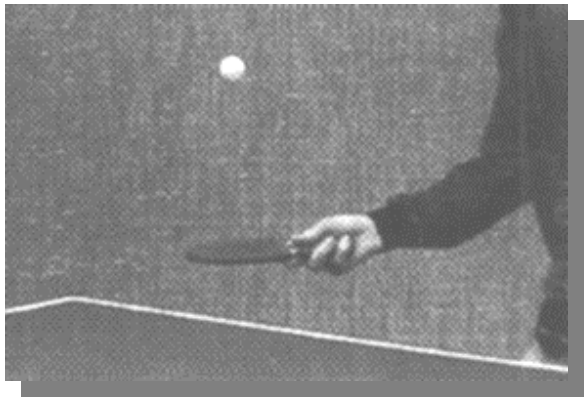


SONY

Motion Prediction

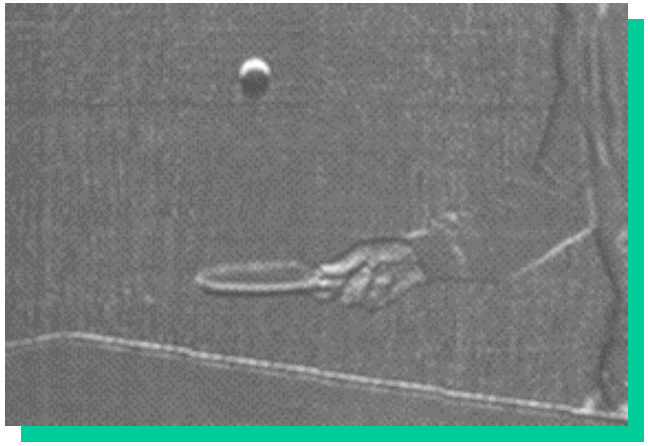


Frame N



Frame N+1

Motion Prediction

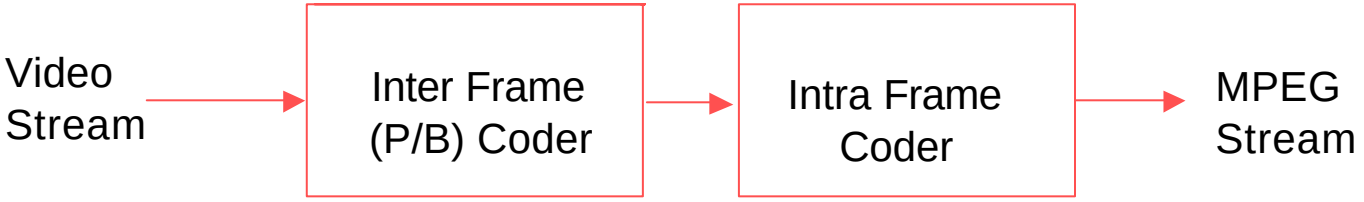


Difference Frame
Without Motion Prediction

Difference Frame
With Motion Prediction

Graphic taken from: “MPEG Video Compression Standard”
Mitchell, Pennebaker, Fogg & LeGall

MPEG Coder



General Structure

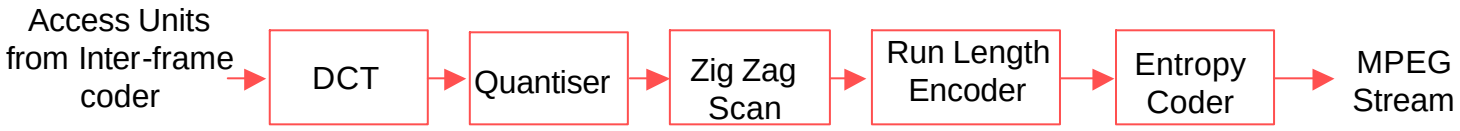
SONY

Typical Picture

Most 8x8 Pixel areas are evenly shaded

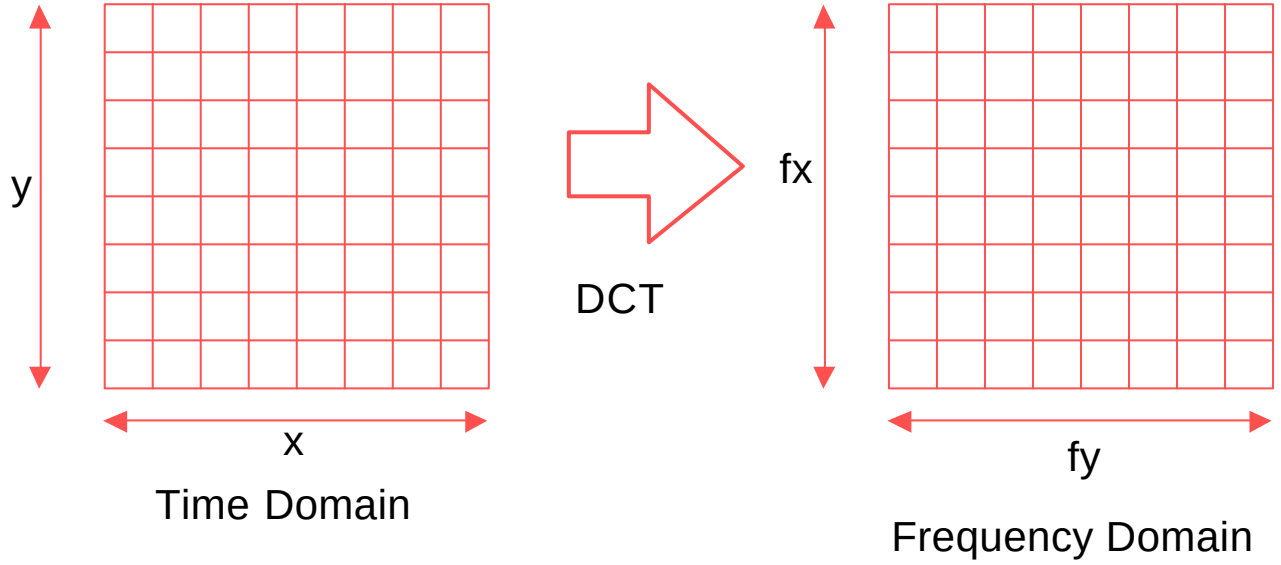
SONY

MPEG Intra -frame Coder

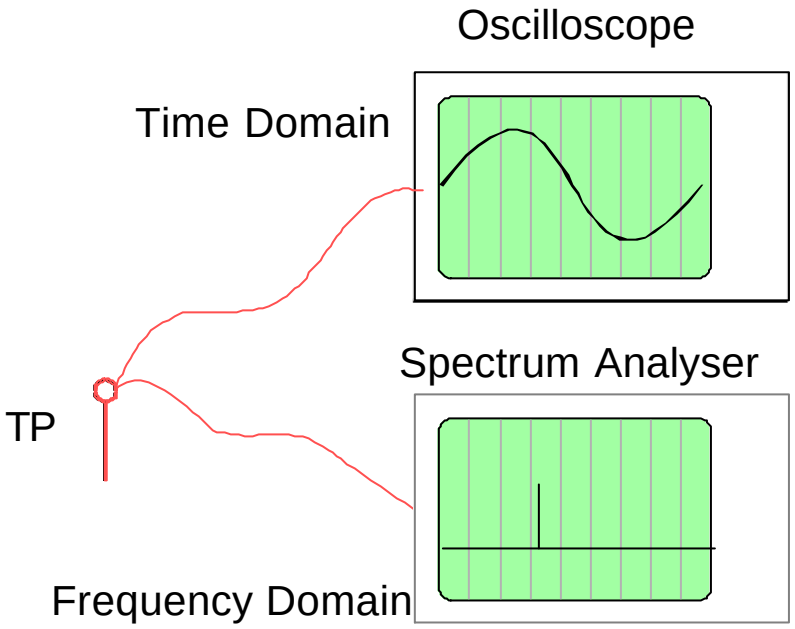


SONY

DCT Process

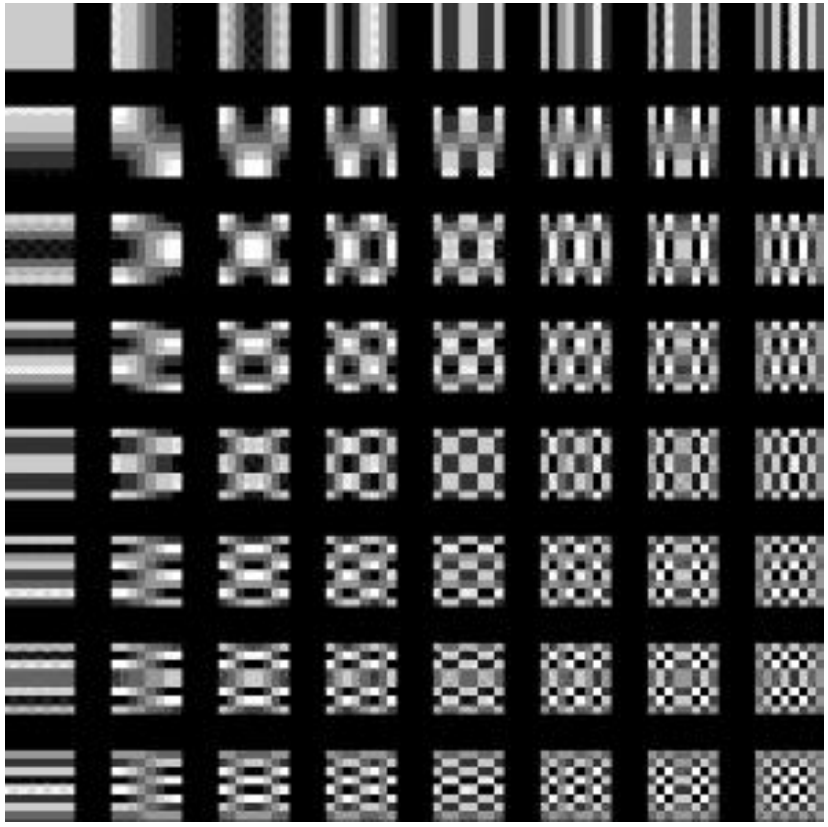


Time and Frequency Domain



SONY

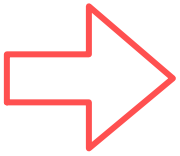
DCT Spacial Frequency Patterns



Typical DCT Transform

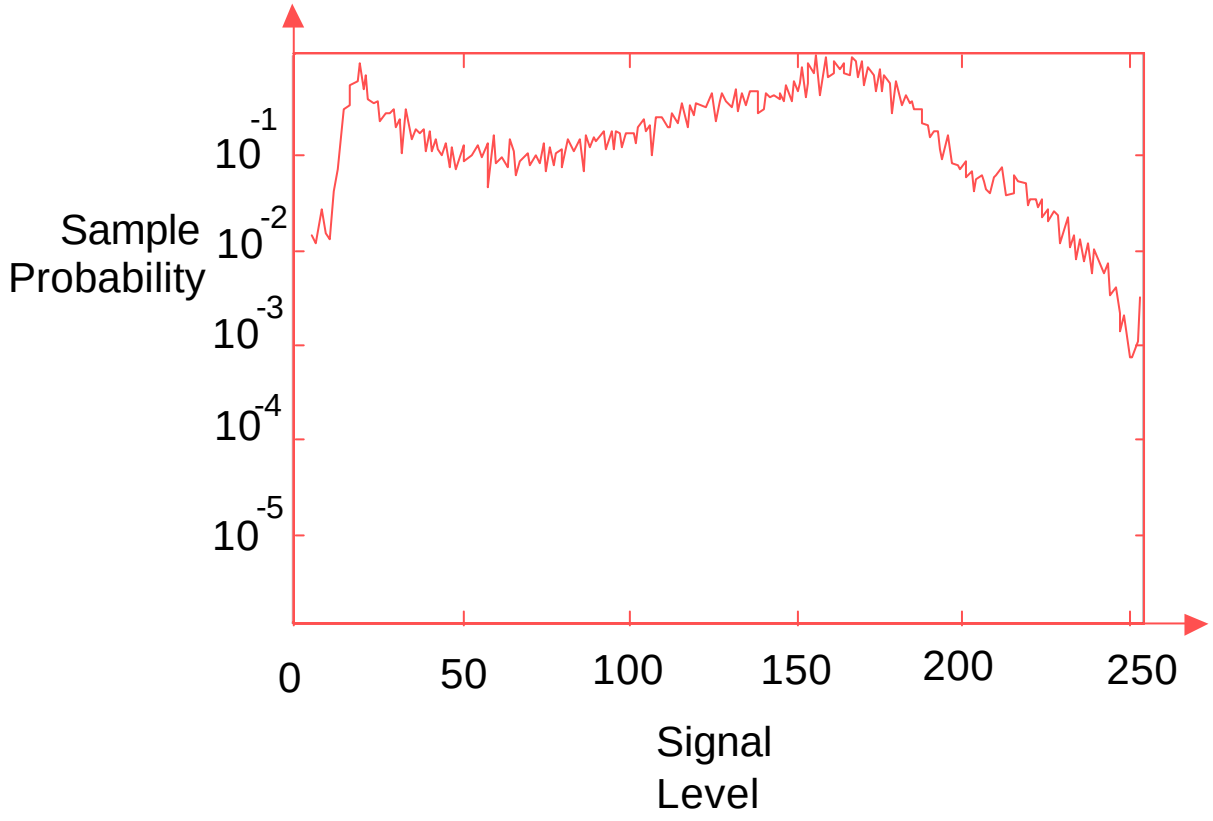
83	87	92	90	89	91	47	95
98	81	77	96	71	44	58	49
27	43	65	40	64	99	61	55
83	21	45	51	59	80	48	63
87	94	56	62	41	74	75	57
98	82	68	79	54	46	52	60
72	42	70	84	69	50	97	67
76	53	85	88	73	66	78	86

DCT



284	18	-12	2	3	-7	0	0
10	9	8	1	-4	1	0	0
-6	2	1	0	2	4	2	-1
2	1	-1	1	-3	0	2	0
0	0	0	1	2	0	0	-1
1	-2	1	0	1	-2	3	0
0	0	0	-1	1	2	0	0
0	0	0	0	0	0	0	0

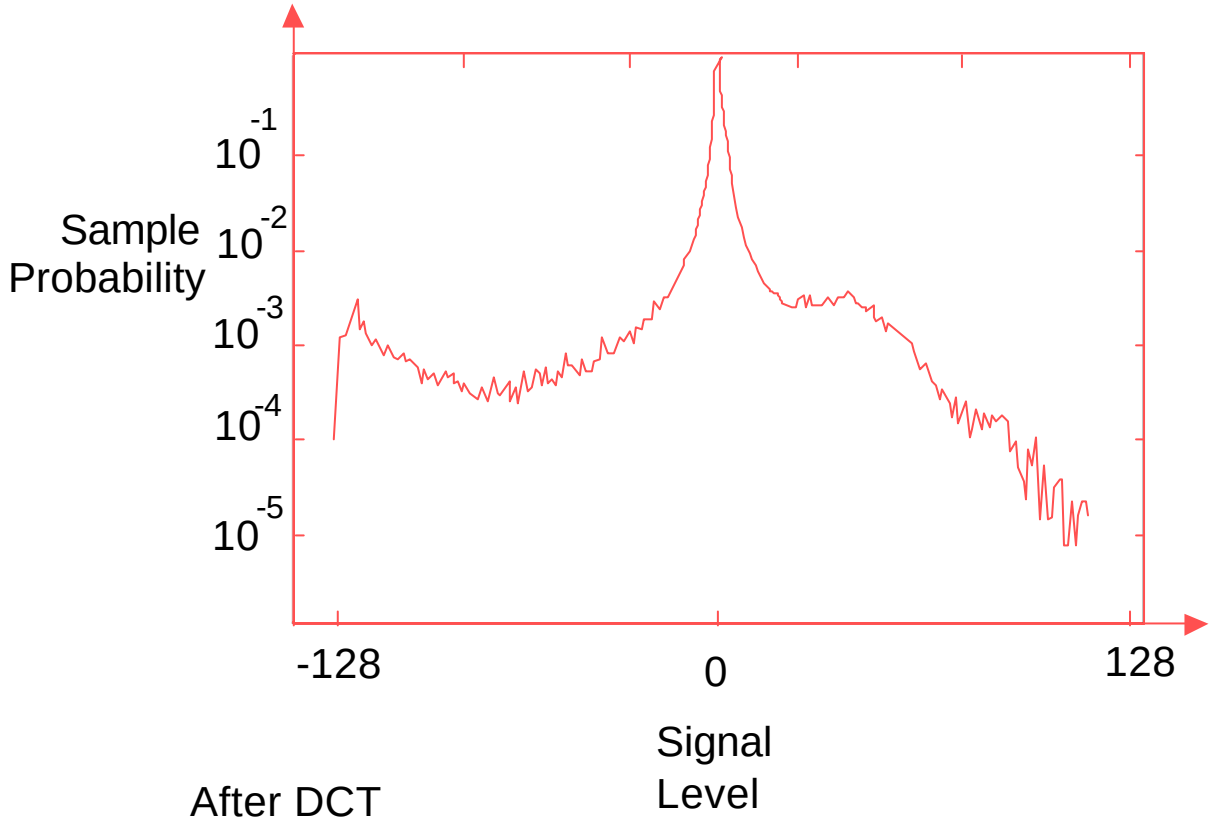
Signal Entropy



Before DCT

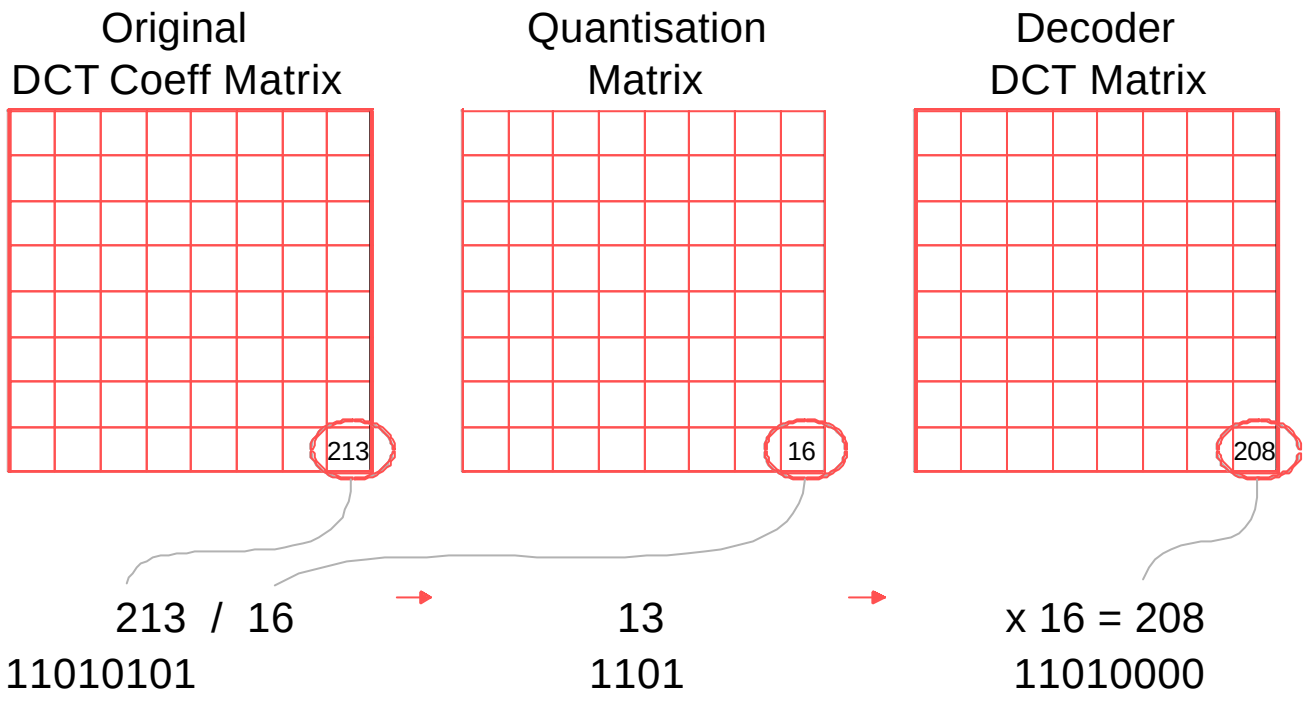
SONY

Signal Entropy



SONY

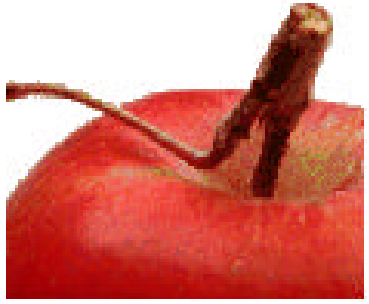
Quantisation



Quantisation Error: 213-208 = 5

SONY

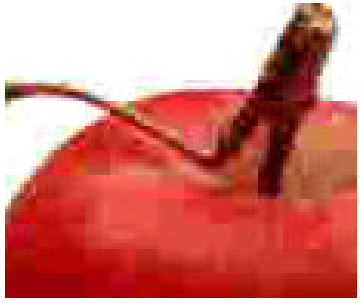
Mosquitos



Source File

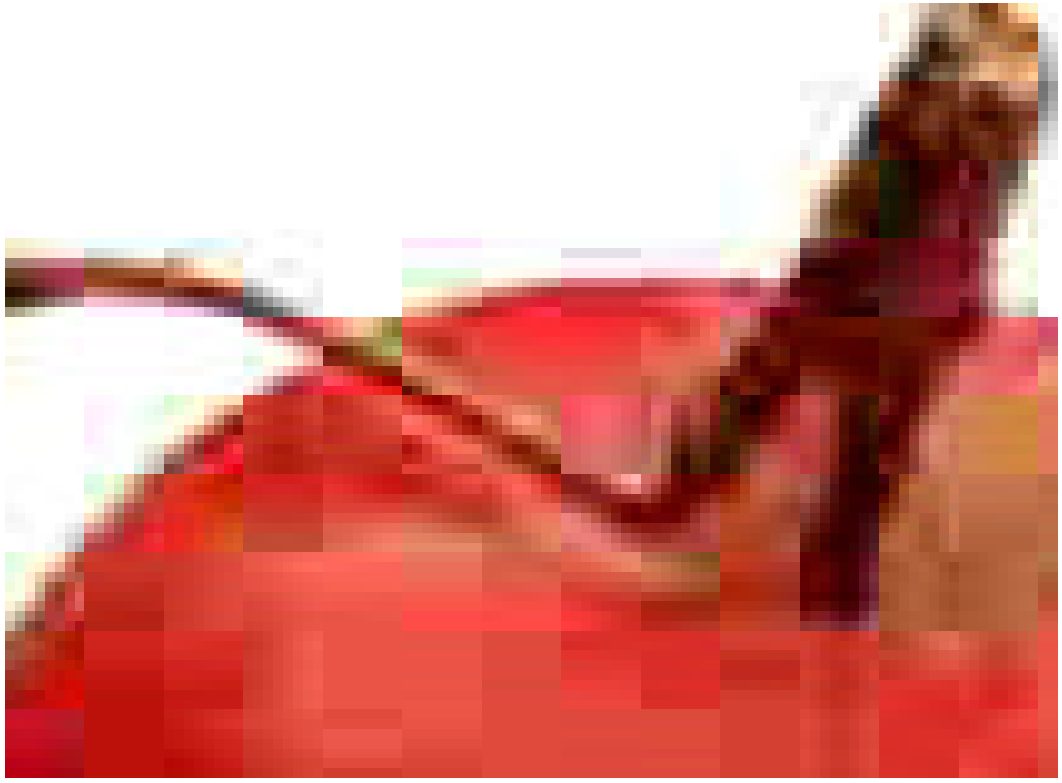
2:1 Compressed JPEG

4:1 Compressed JPEG



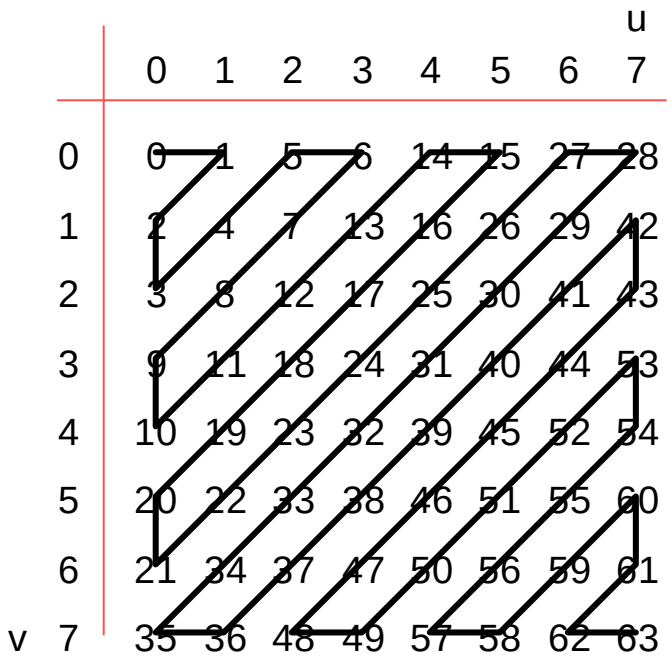
12:1 Compressed JPEG

Mosquitoes

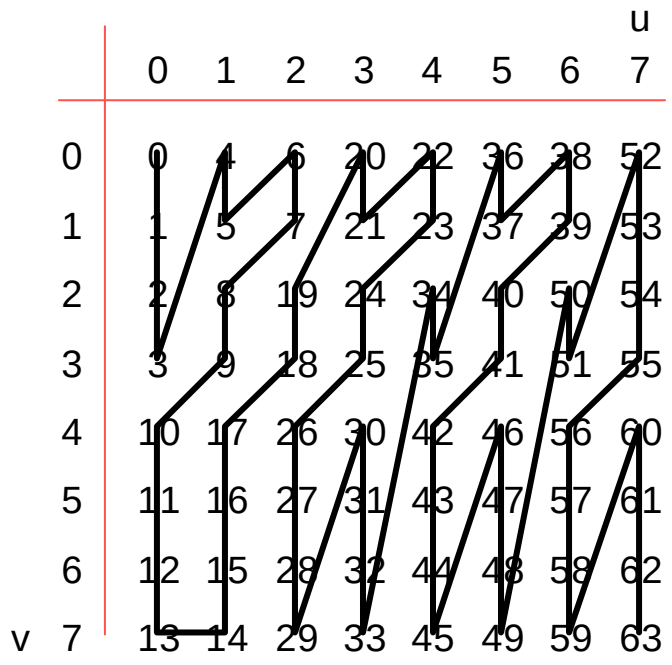


12:1 Compressed JPEG

Zig Zag Scan Pattern



Zig Zag Scan
Definition (0)



Alternate Zig Zag Scan
Definition (1)

SONY

Huffman Code Table

Variable Length Code	Run Length	Level
0000 110 s	0	4
0010 0110 s	0	5
0010 0001 s	0	6
0010 0101 s	1	3
0000 100 s	2	2
0010 0100 s	3	2
0001 01 s	6	1
0001 00 s	7	1
0000 111 s	8	1
0000 101 s	9	1
0010 0111 s	10	1
0010 0011 s	11	1
0010 0010 s	12	1
000001	Escape	
10	End_of_block	

Escape Codes

Run Length 0 - 63	Size -2047 to +2047
6 Bits	12 Bits

To System Comparison

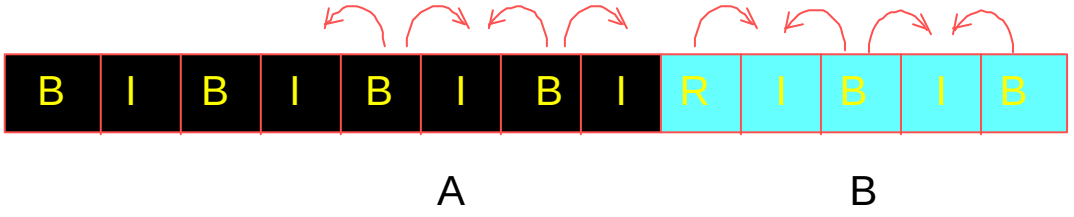
Editing Betacam SX

- There are three forms of editing:
  - Camcorder based backspace Edit
  - Tape to tape editing
  - Hard disk based editing

Editing Betacam SX

- The unit recorded on tape is a GOP (B-frame, I-Frame)
- Recordings must start and finish at GOP boundaries
- B frame must be re-coded if the reference I-frame is removed

Editing Betacam SX



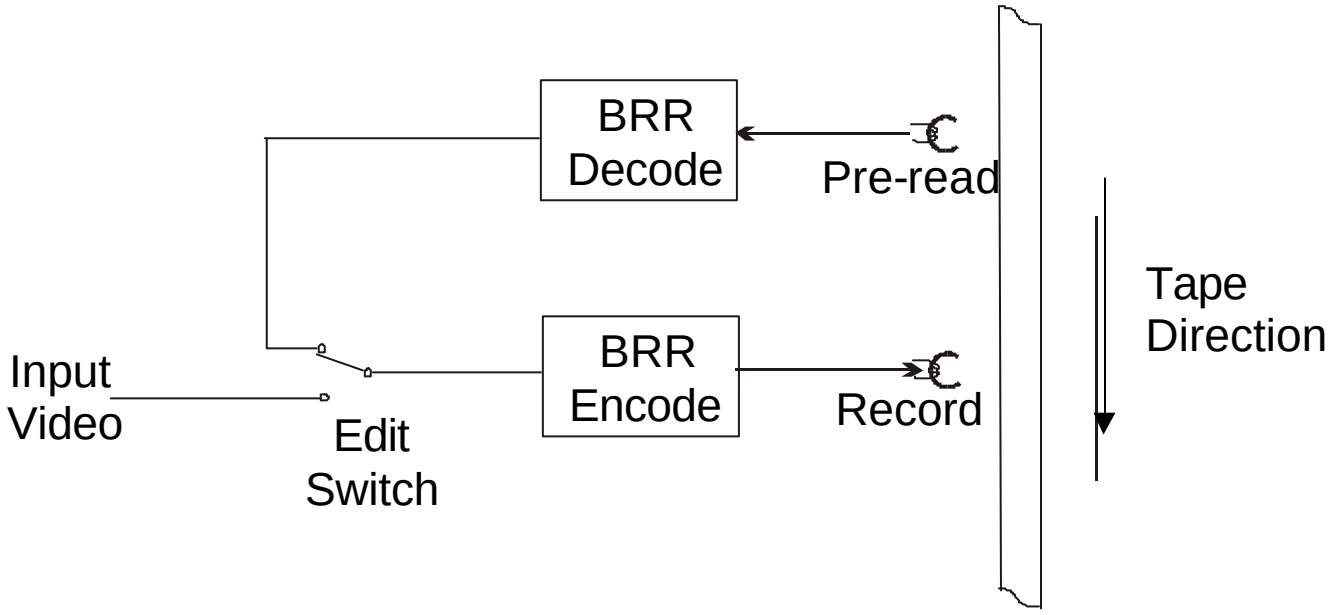
Backspace/Assembly Edit
performed with an R frame

SONY

Insert Editing in a SX VTR

- The two machines are connected via SDI
- Edit is decode - re-code
- Frame accurate editing is possible
- Existing frames around the edit point must be re-coded
- Pre-read heads are used

Pre-read Process

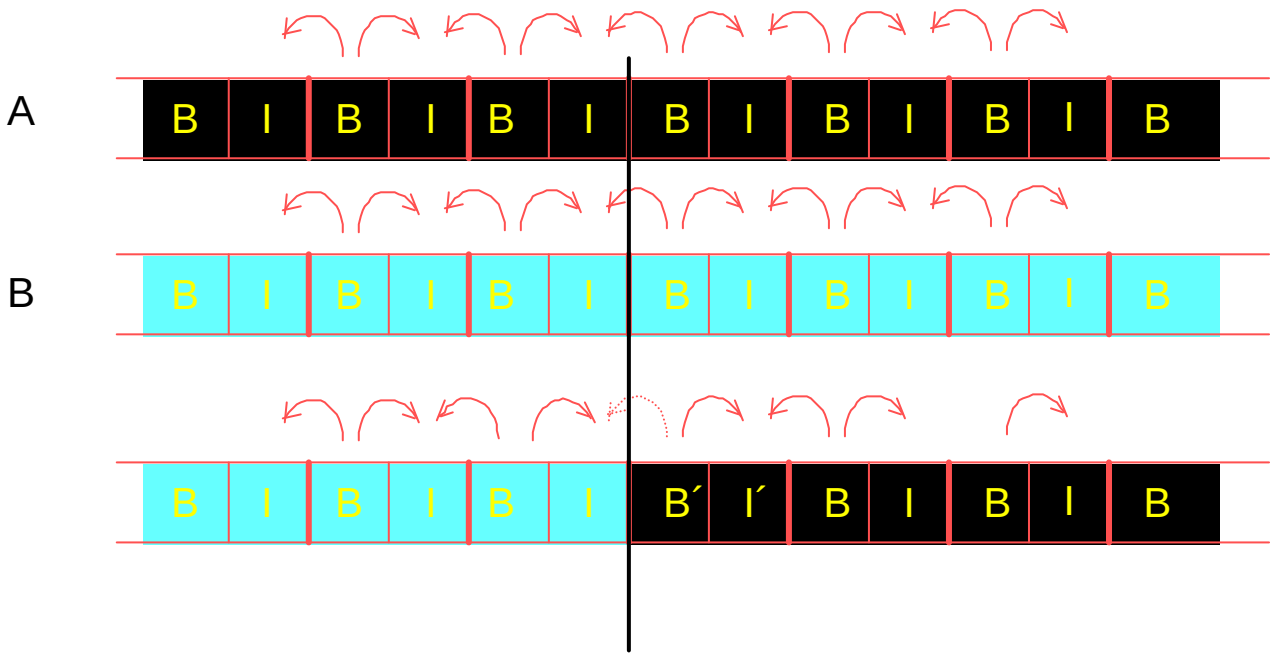


SONY

Tape to tape editing

- Four possible scenarios exist
  - In-point is
    - B Frame
    - I frame
  - Out point is
    - B frame
    - I frame

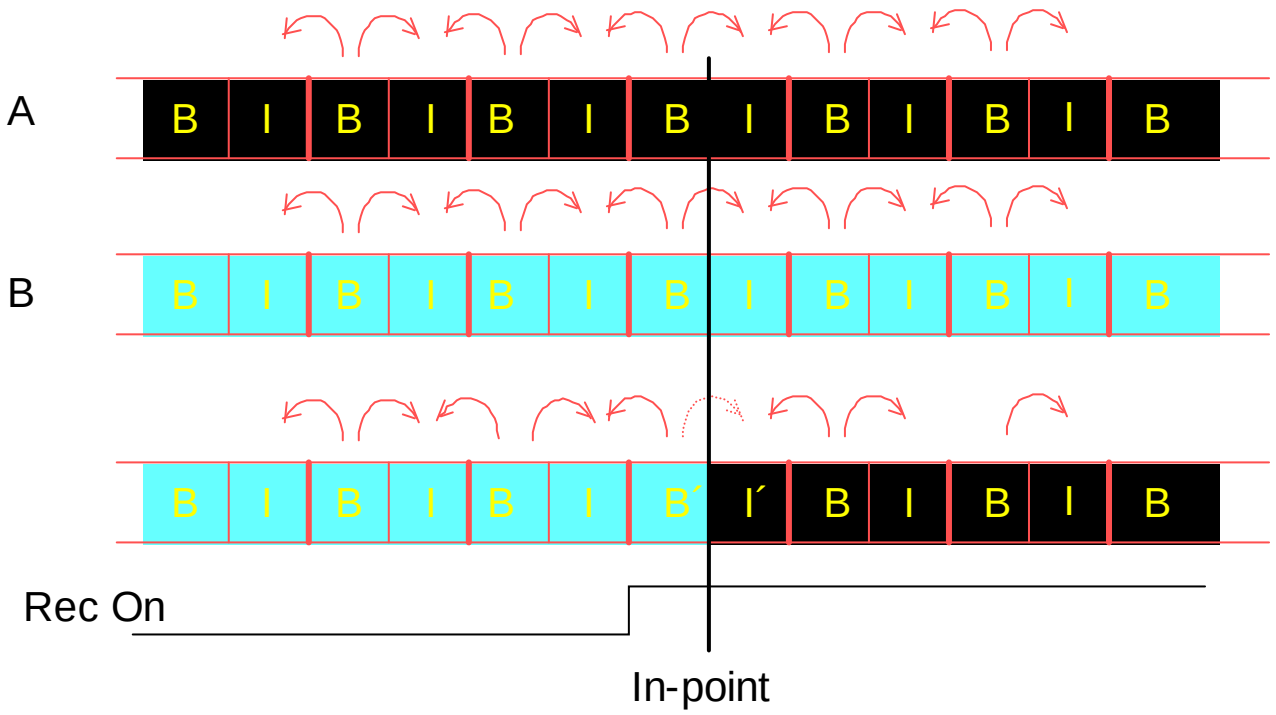
In-point B-frame



Insert Edit
performed in an A220

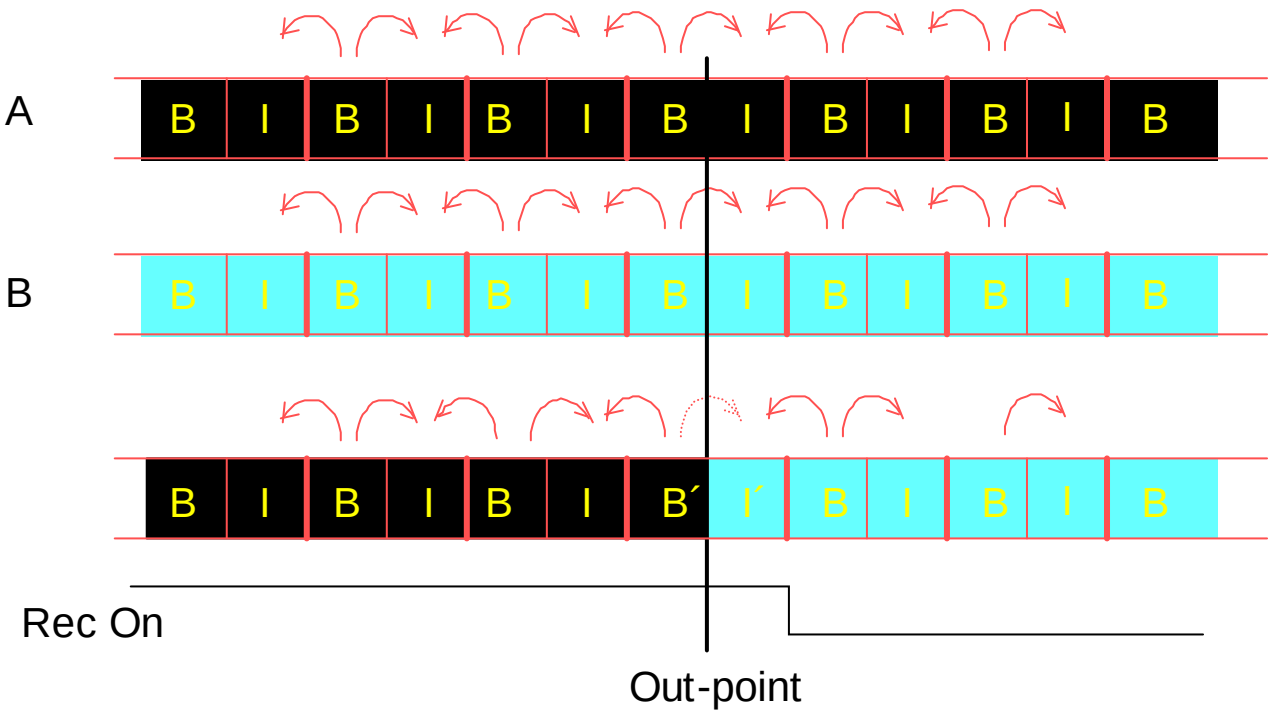
SONY

In-point I-frame



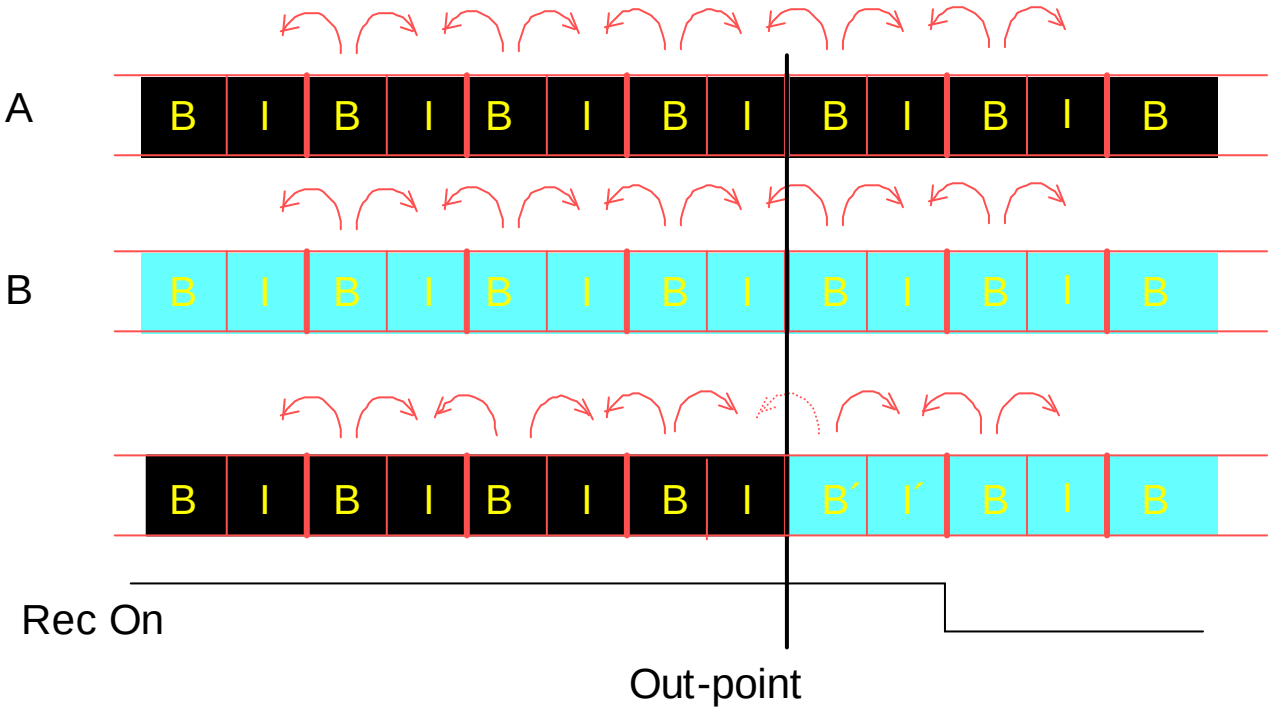
SONY

Out-point B-frame



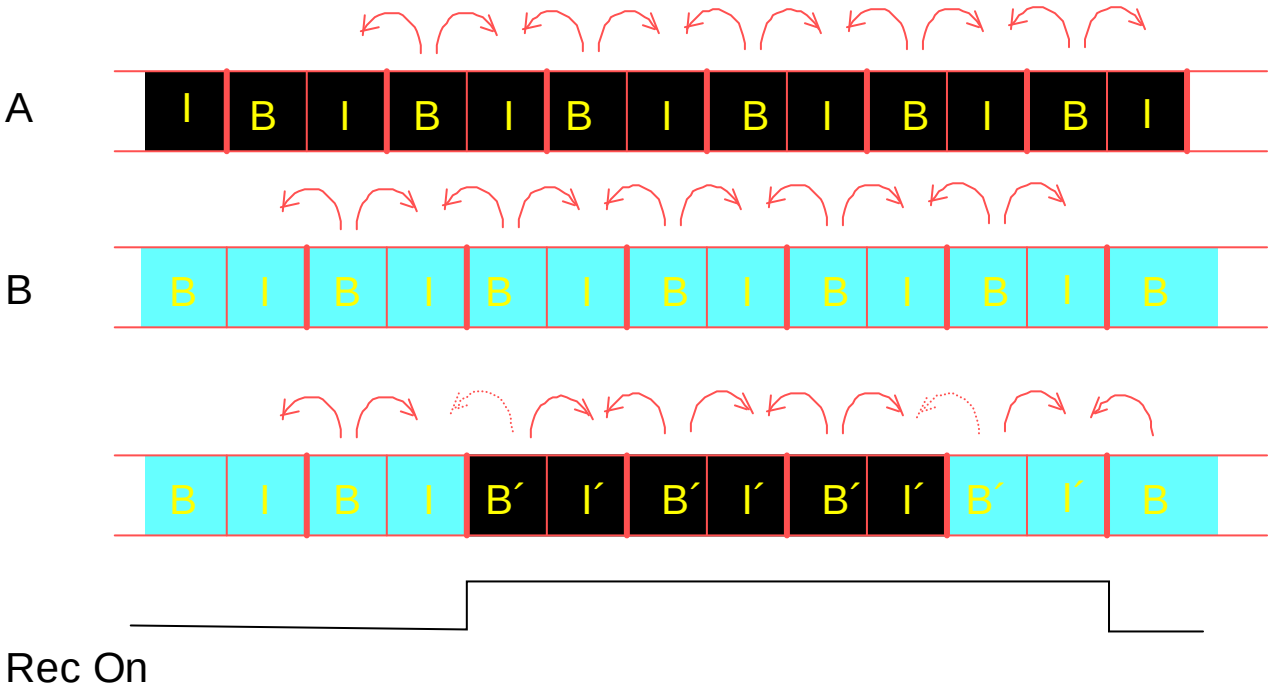
SONY

Out-point I-frame



SONY

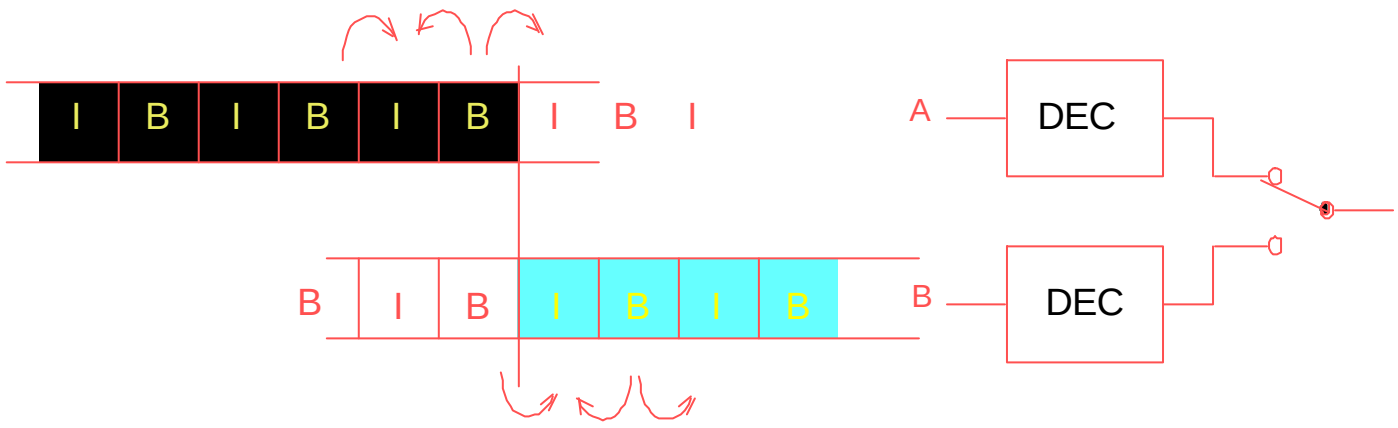
Editing Betacam SX



Frame accurate Editing

SONY

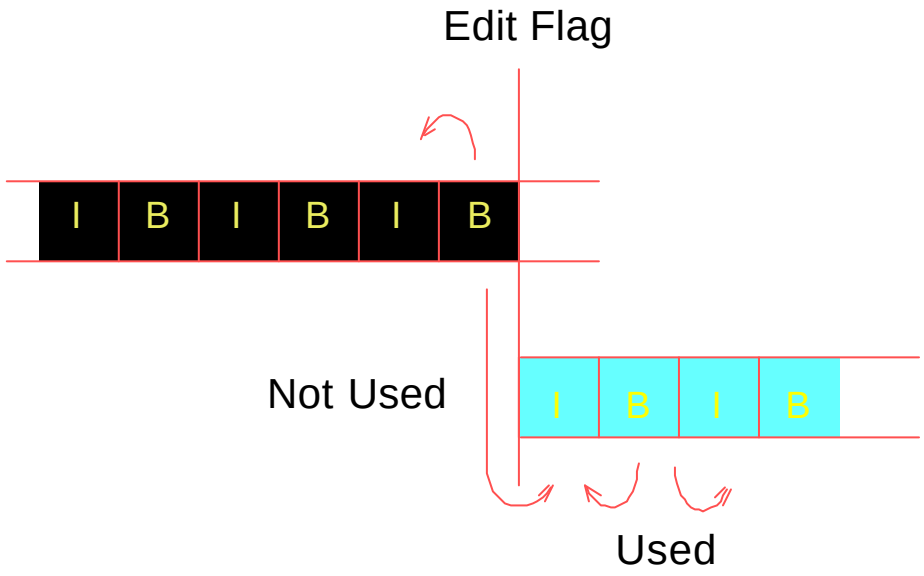
Editing MPEG



Overlap Edit
Hard disc edit where scene A+B are played out concurrently

SONY

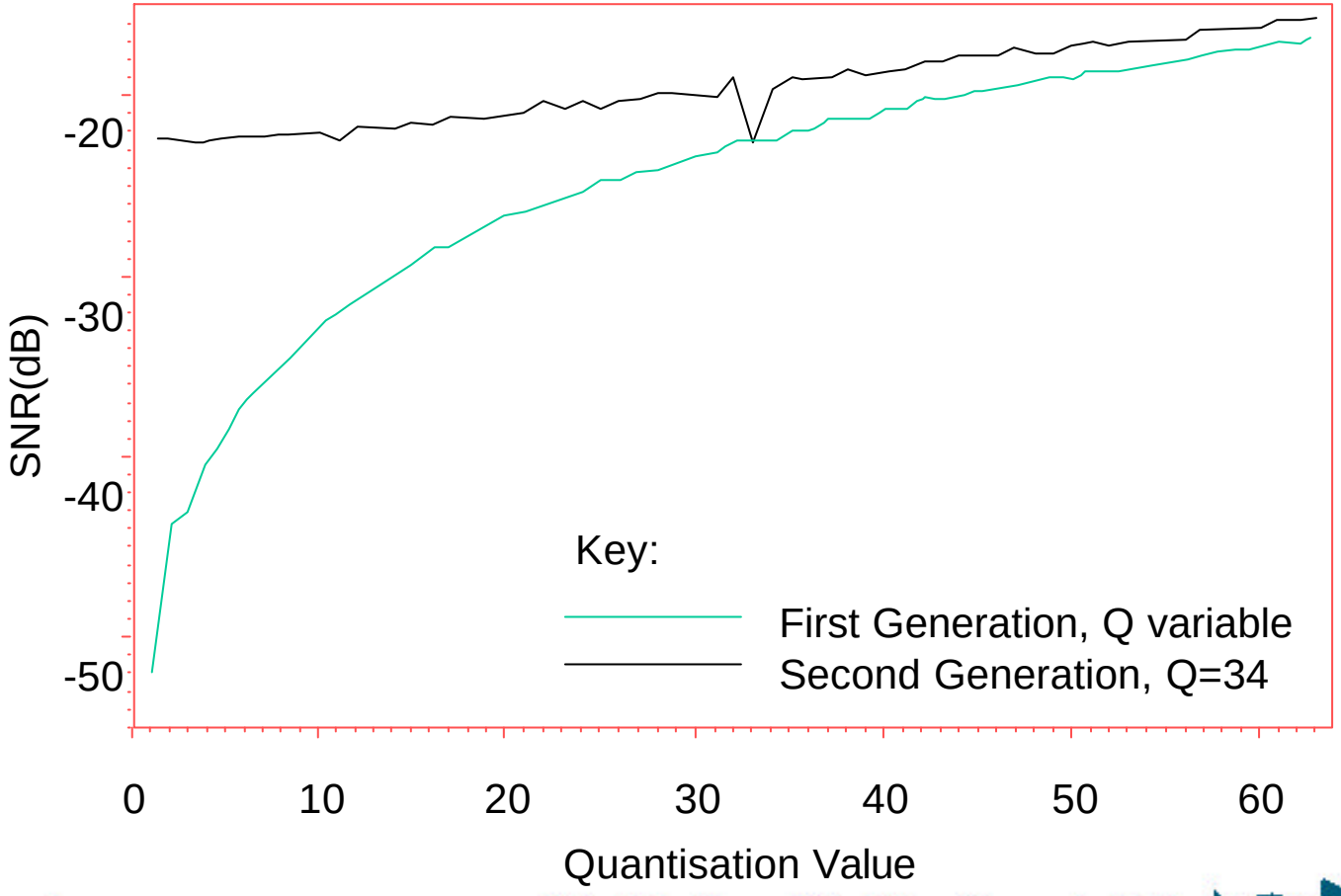
Editing MPEG



MPEG transmitted as SDDI/4 x SDDI with an edit flag,
in decoder first B frame after an edit is processed as an R frame

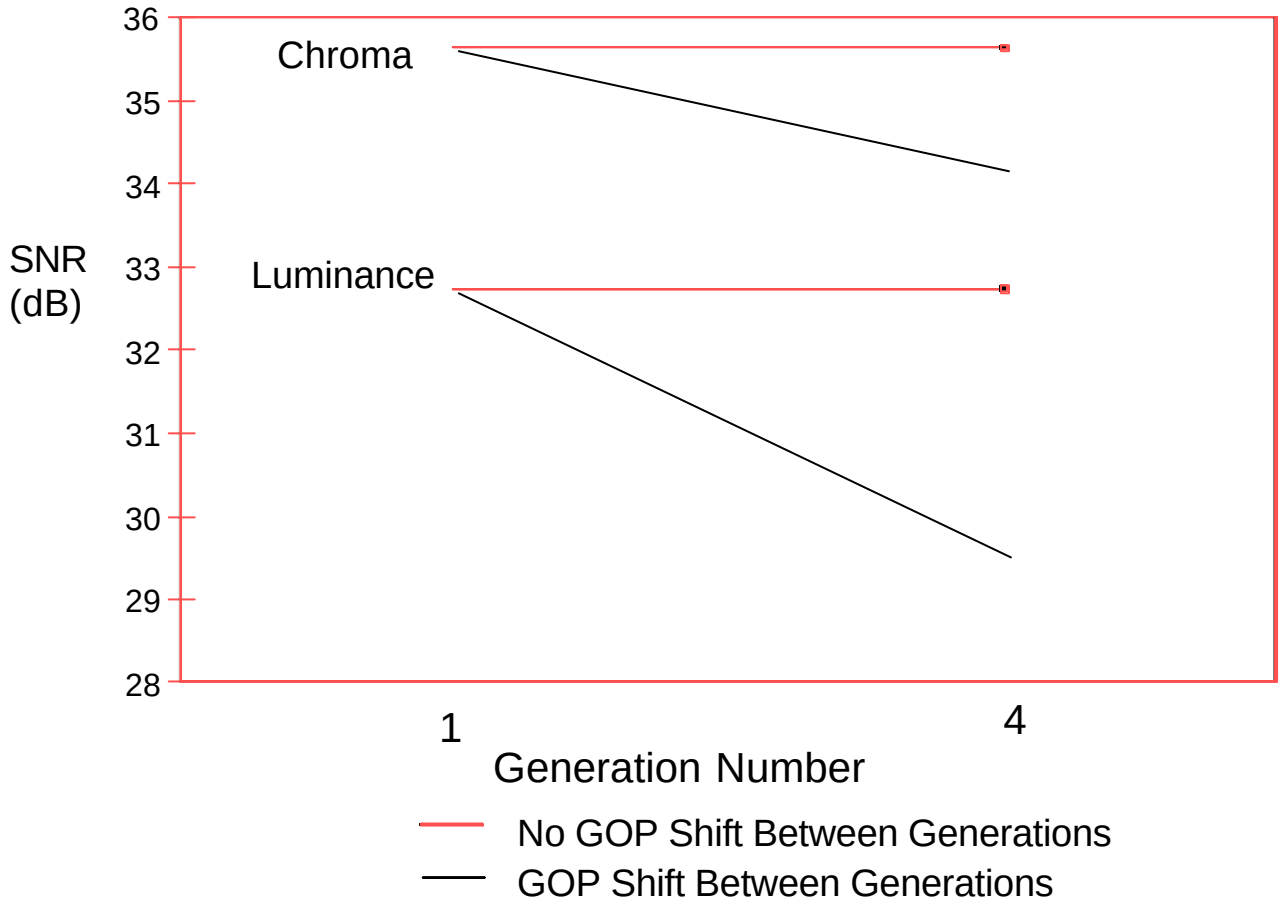
SONY

Multigeneration Performance



SONY

GOP Shift



SONY

System Comparison

- Betacam SX
  - 4:2:2 sampling
  - 608 lines
  - MPEG-2 422P@ML
  - 10:1 Compression
  - ‘Full picture’ compression ratio 9.1:1
  - Video Data Rate = 18 MBits/s
  - Audio - 4 Channel, 48KHz, Uncompressed
  - Recorded Data Rate = 43.8Mb/s

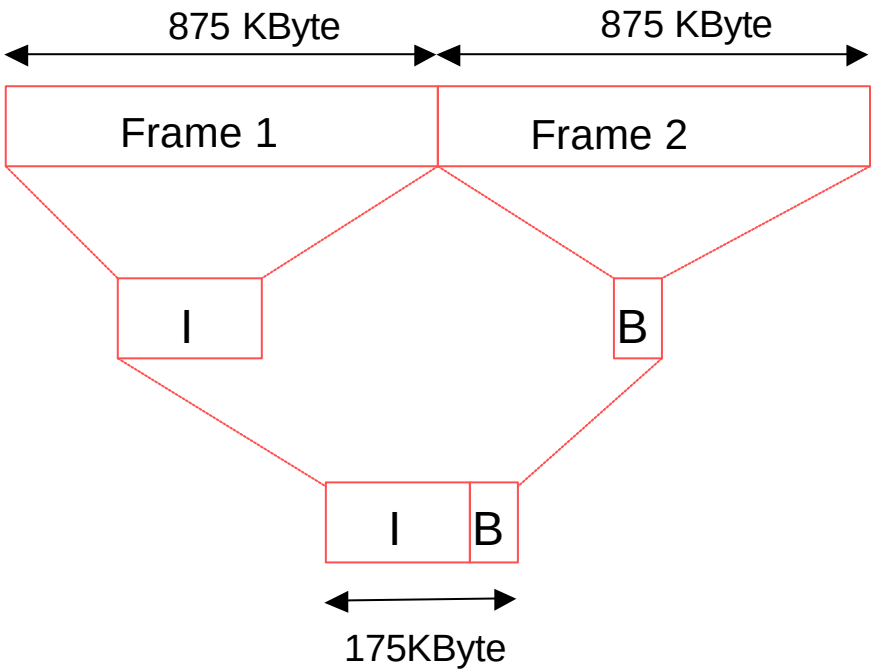
System Comparison

- DVC
  - 4:2:0 sampling
  - 576 lines
  - DVC Compression (intra frame)
  - ‘5:1’ Compression
  - Video data rate = 25 MBits/s
  - ‘Full picture’ compression ratio 6.6:1
  - Audio - 2 channel, 48KHz, uncompressed
  - Recorded data rate = 40.4Mb/s

Compression Ratios

- Is compression ratio a good measure of quality?
- What else must be considered?

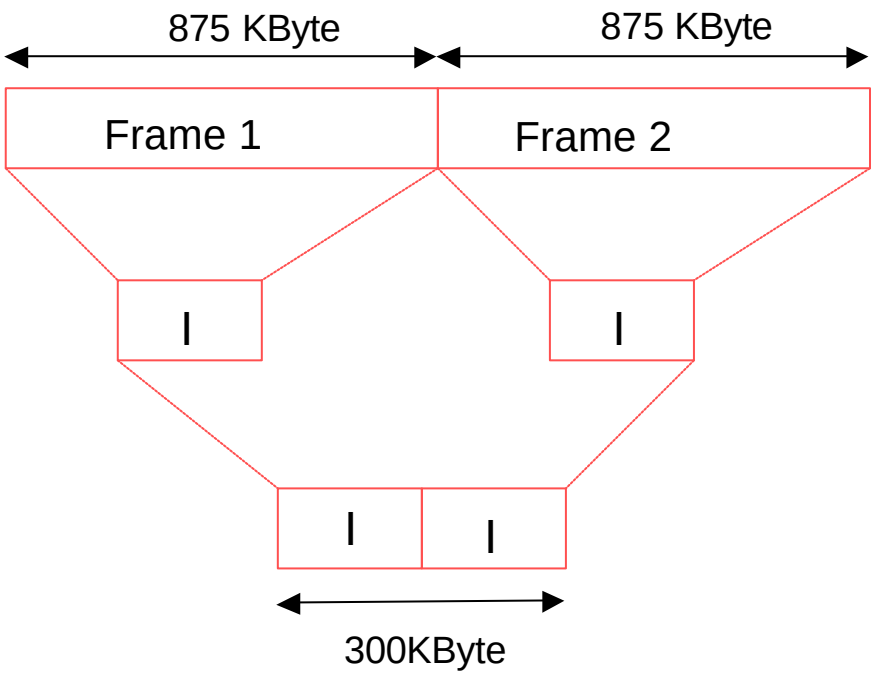
Compression Ratio



Betacam SX
Ratio: 10:1

SONY

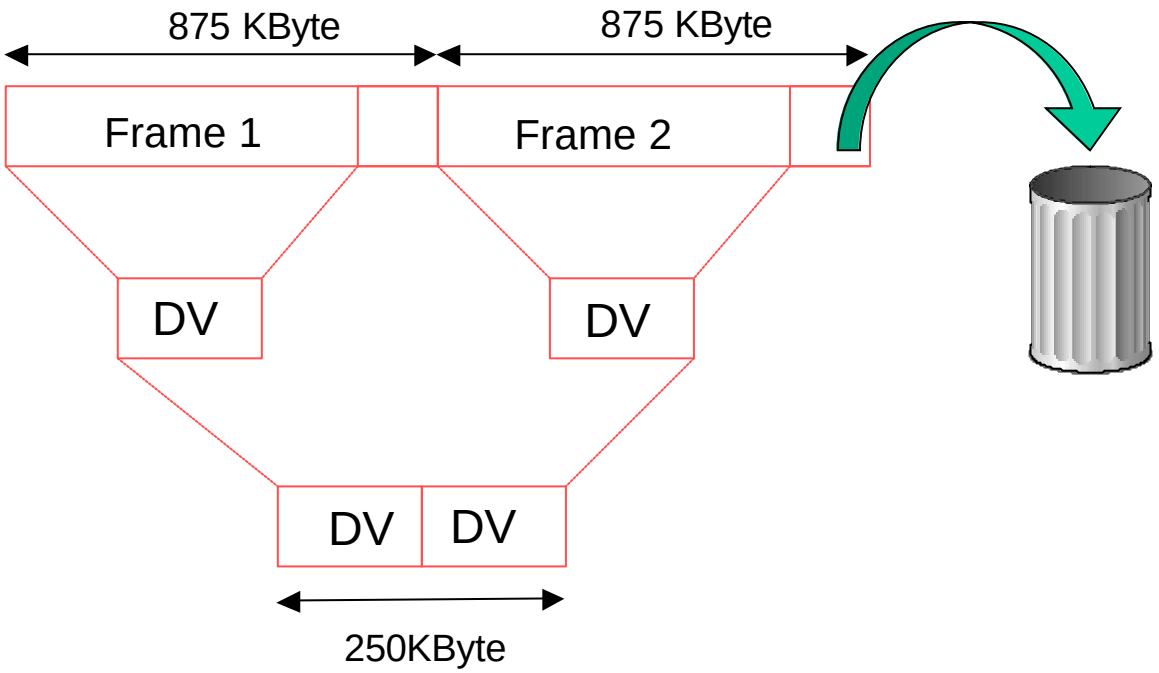
Compression Ratio



I frame only MPEG
Ratio: ~6:1
(equiv. 30 Mbps)

SONY

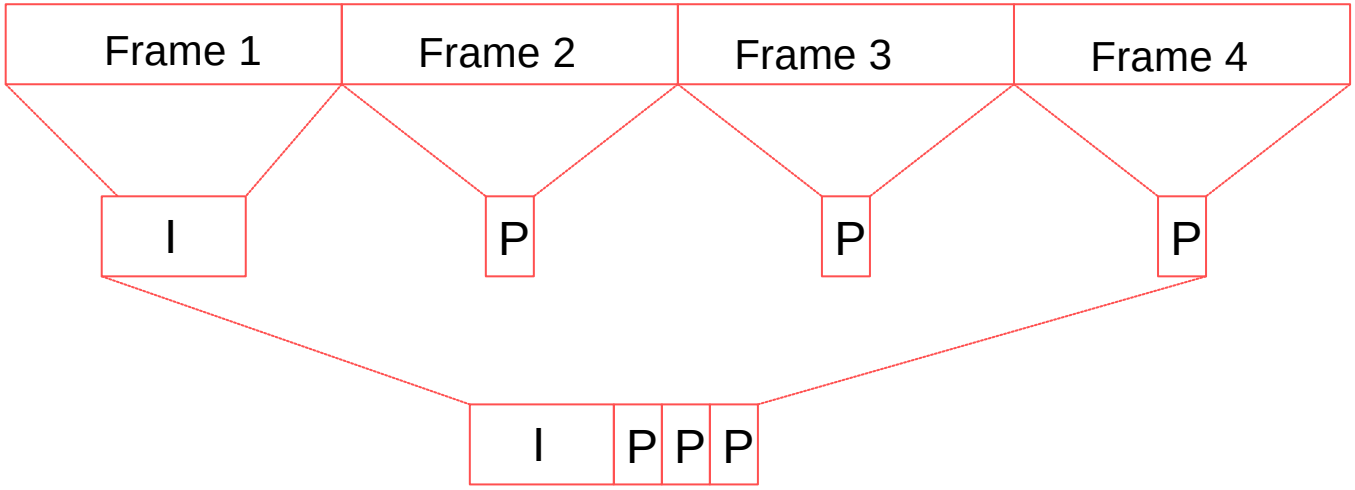
Compression Ratio



DV
Ratio: ‘5:1’ (quoted)
‘full picture’ ratio 6.6:1

SONY

Compression Ratio



Data rate 4-8Mbps typically

MPEG Transmission
Ratio: ~20:1

SONY

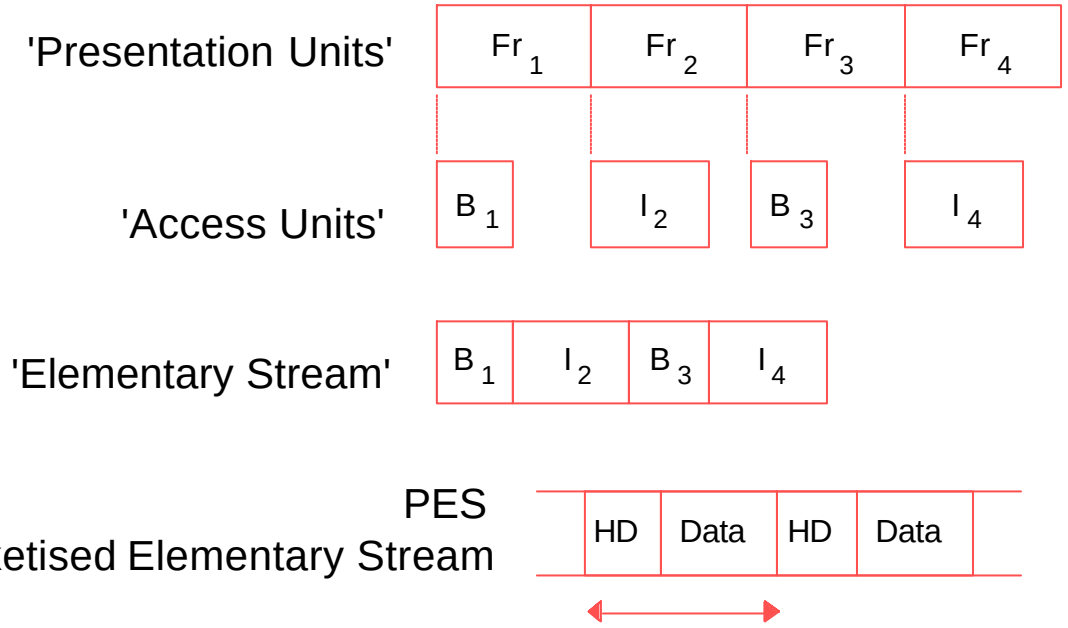
Quality Comparison

- Cannot use analogue measurement techniques
  - all systems produce ‘perfect’ results
- Data rate or Compression ratio is not a good measure of quality
- Different systems have different efficiencies
- Must measure system based on visual impairment

Textronix PQA

- Picture Quality Assessor
  - Compares reference sequence with recorder output
  - weights data on the basis of how we see
  - produces a PQR (Picture Quality Rating)
    - 0: Perfect - No measurable error
    - 0-5, Good - No noticeable error
    - 5-10 Fair - Discernable errors, but ‘Broadcast Quality’
    - 10+ Poor - Not ‘Broadcast Quality’

MPEG System Layer



Variable Length <64k (Except Video)

SONY

Stream Conversion

- SX native is output on SDTI
- Stream convertor turns this into TS
- This is data re-ordering and there is no quality loss
- TS can be interfaced to ATM
  - again this is loss-less re-odering

SX is MPEG

&

Stream Conversion is a loss-less Process

Transcoding

- Transcoding is used by different people to mean different processes
- **We** use it to mean:

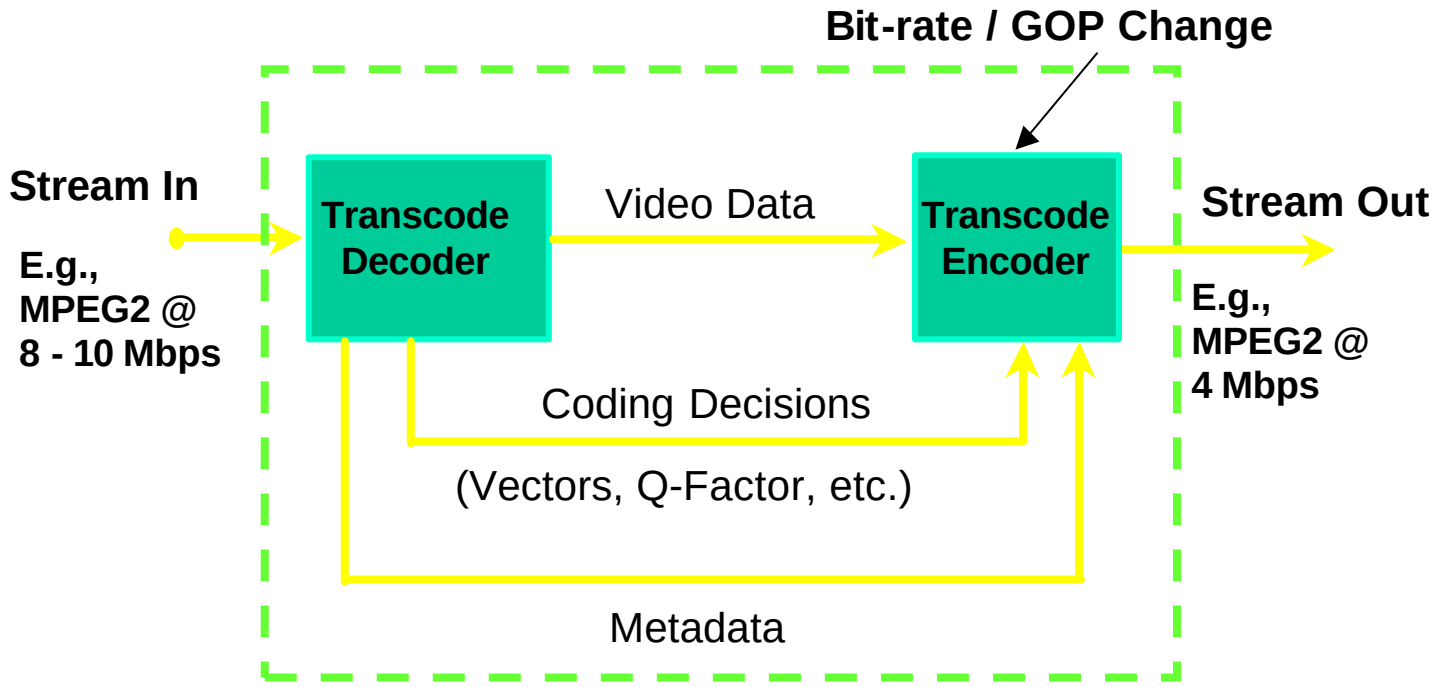
The conversion of one MPEG bit-stream into another

without decoding back to baseband (SDI)

Transcoding

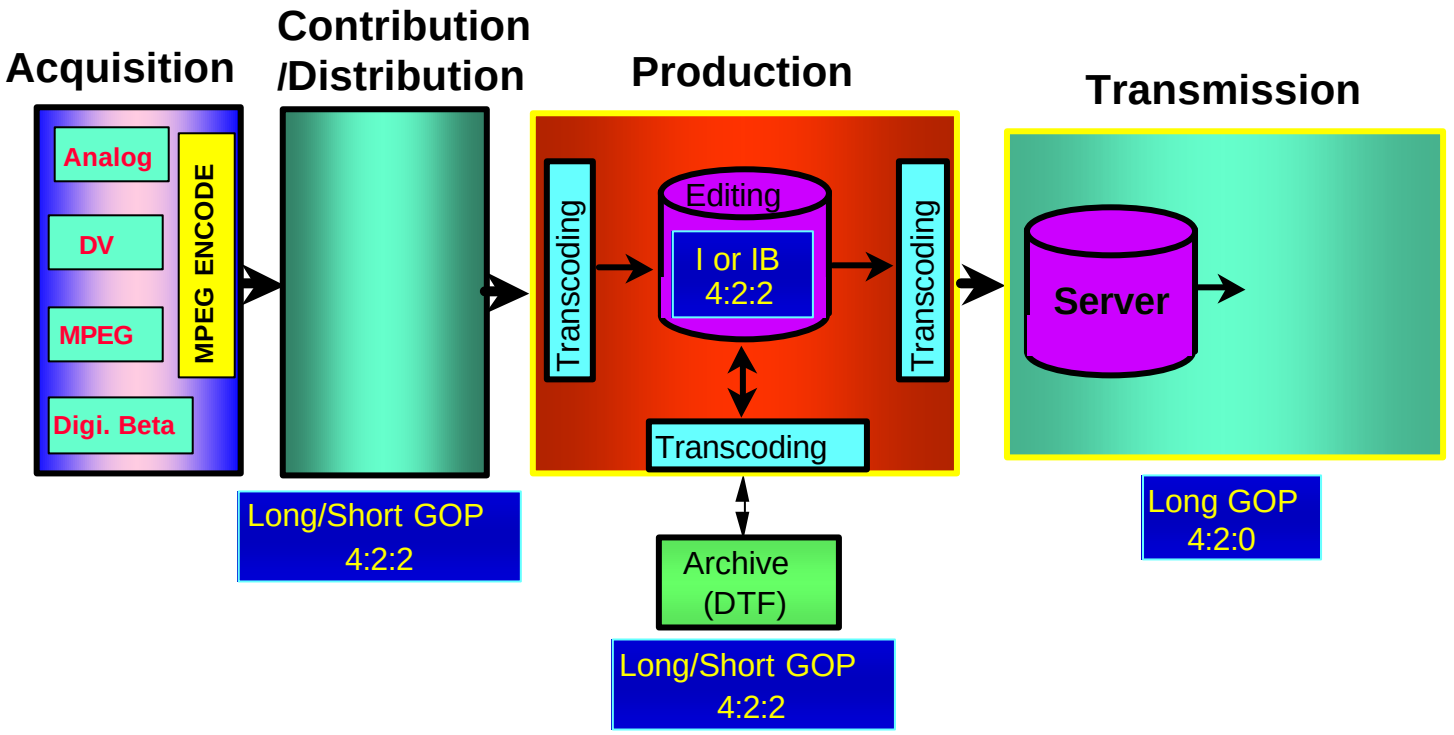
- Transcoding is a key technology
- Many companies are working on transcoding
  - Atlantic project
- Sony have a chip set and can demonstrate its function
- Transcoding has the advantage of:
  - minimising the loss of quality during the process
  - allowing coder decisions to be reused in a subsequent process
  - improving signal quality
- Principally transcoding keeps the signal in the MPEG domain

Transcoding



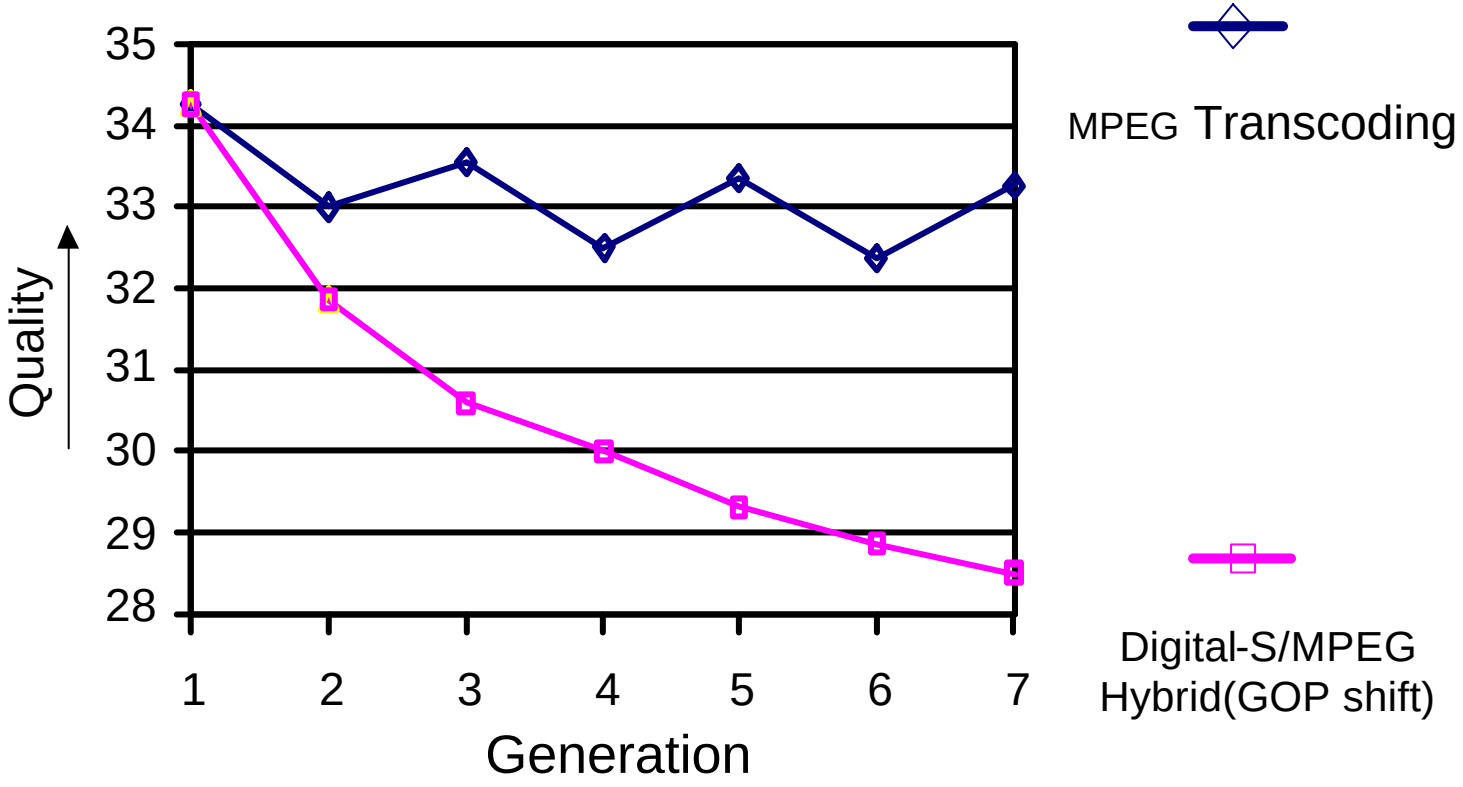
SONY

Transcoding



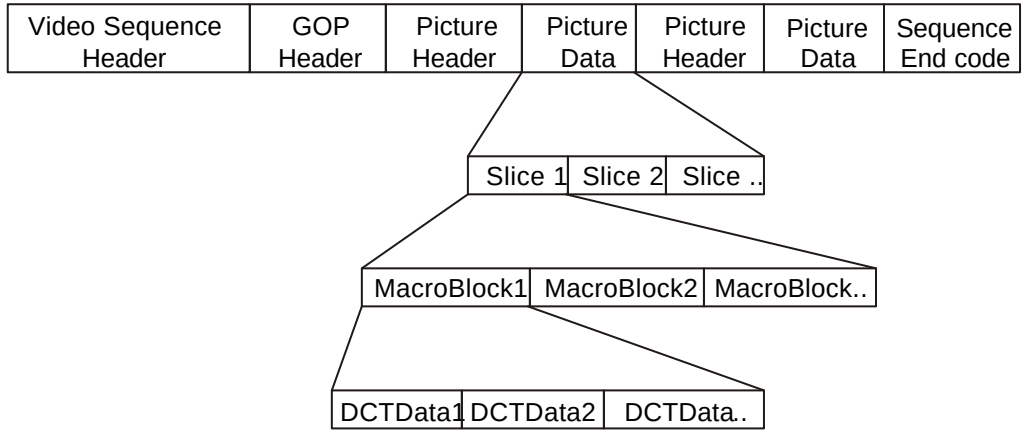
SONY

Transcoding



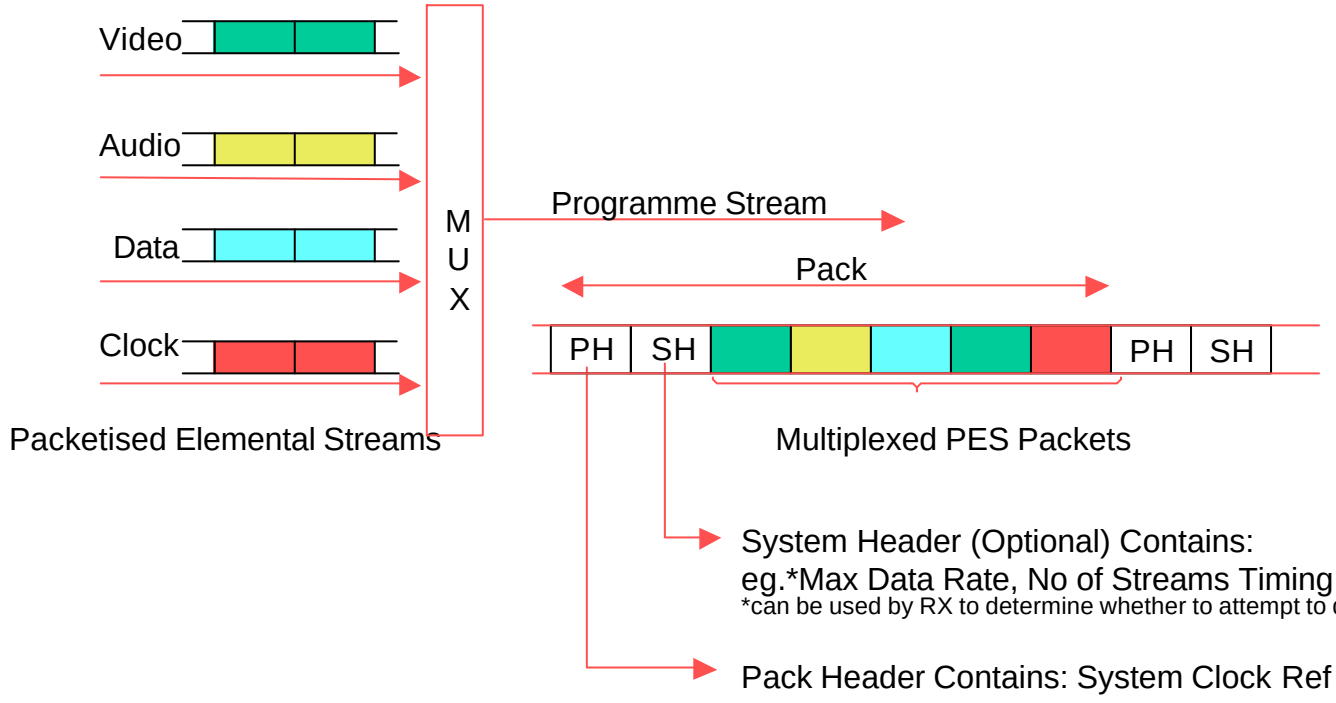
SONY

MPEG ES


Video Sequence Header
GOP Header
Picture Header
Picture Data
Picture Header
Picture Data
Sequence End code
Slice 1
Slice 2
Slice ..
MacroBlock1
MacroBlock2
MacroBlock..
DCTData1
DCTData2
DCTData..

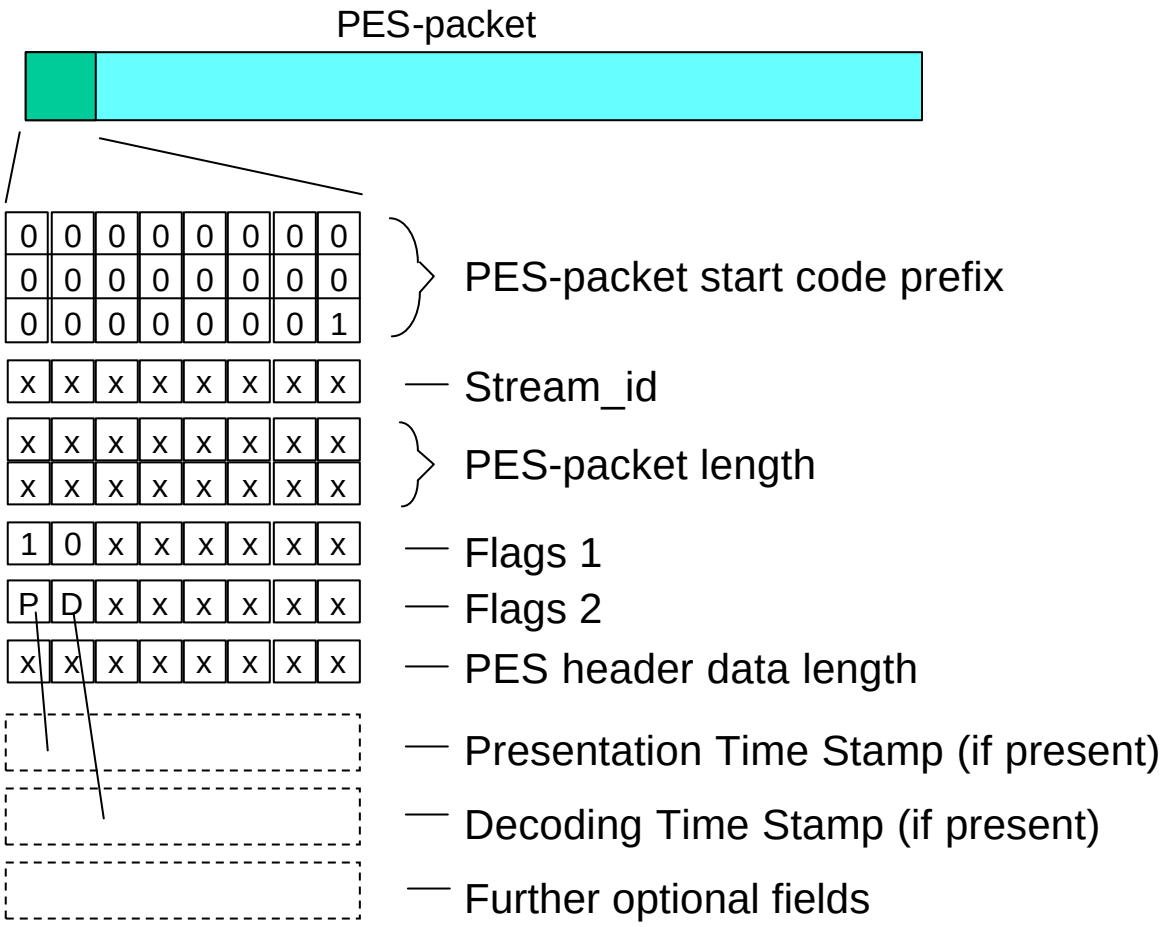
SONY

MPEG Programme Stream



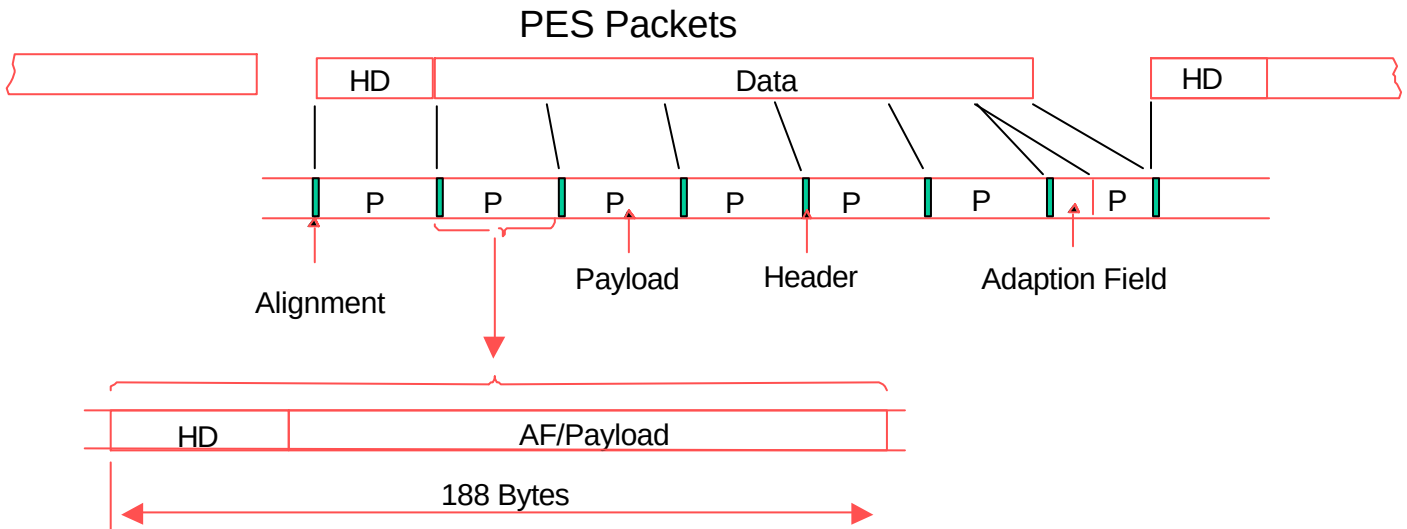
SONY

PES Header



SONY

MPEG Transport Stream

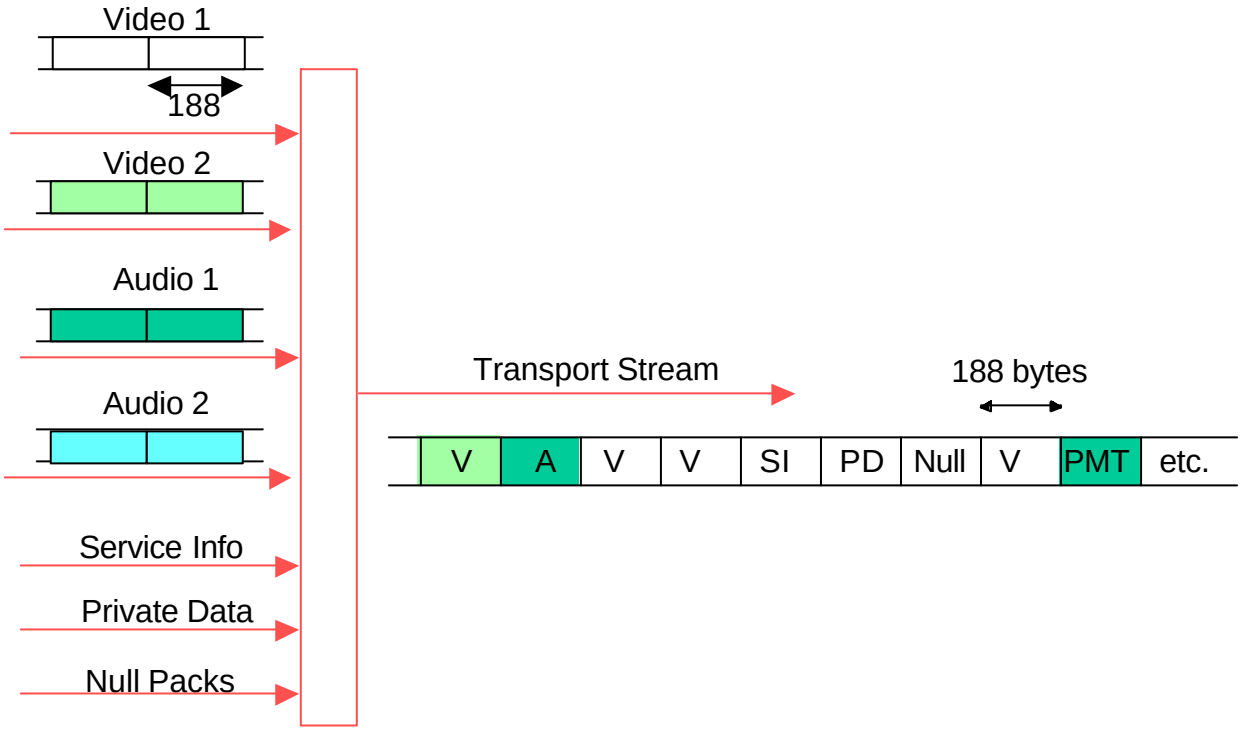


Sync	8	47H/01000111
TEI	1	Transport Error Indicator
PUSI	1	Payload Unit Start Indicator
TP	1	Transport Priority
PID	13	Packet ID (which prgramme in stream) 17 have special meaning, therefore 8175 left
TSC	2	Transport Scramble Control
AFC	2	Adaption Field Control
CC	4	Continuity Counter (increments every 188 bytes)

Total 32 bits

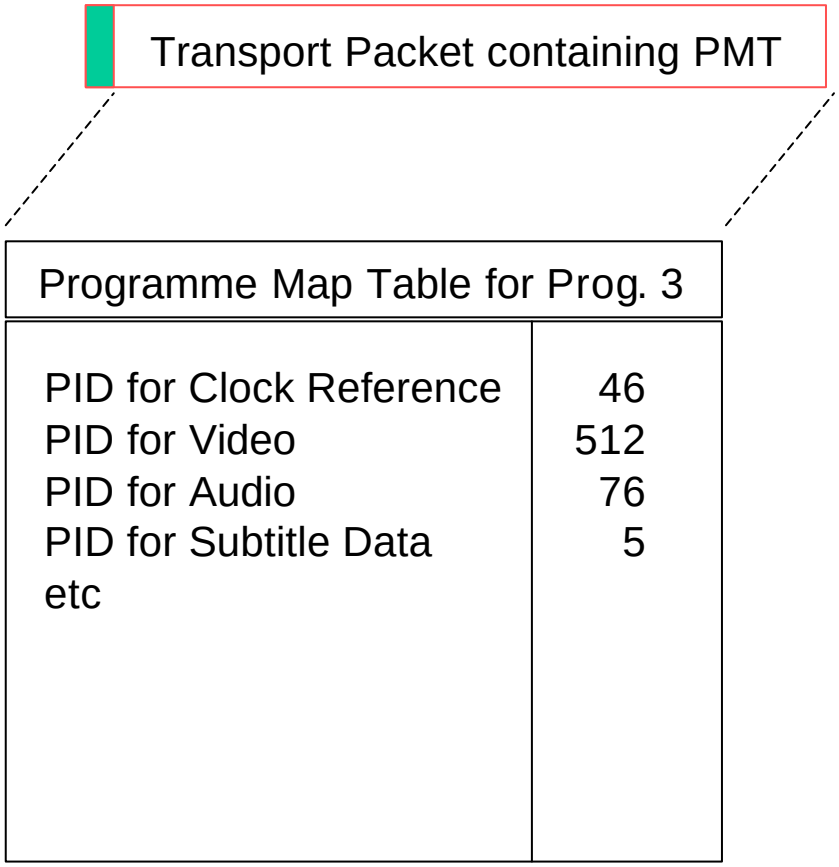
SONY

MPEG Transport Stream Generation



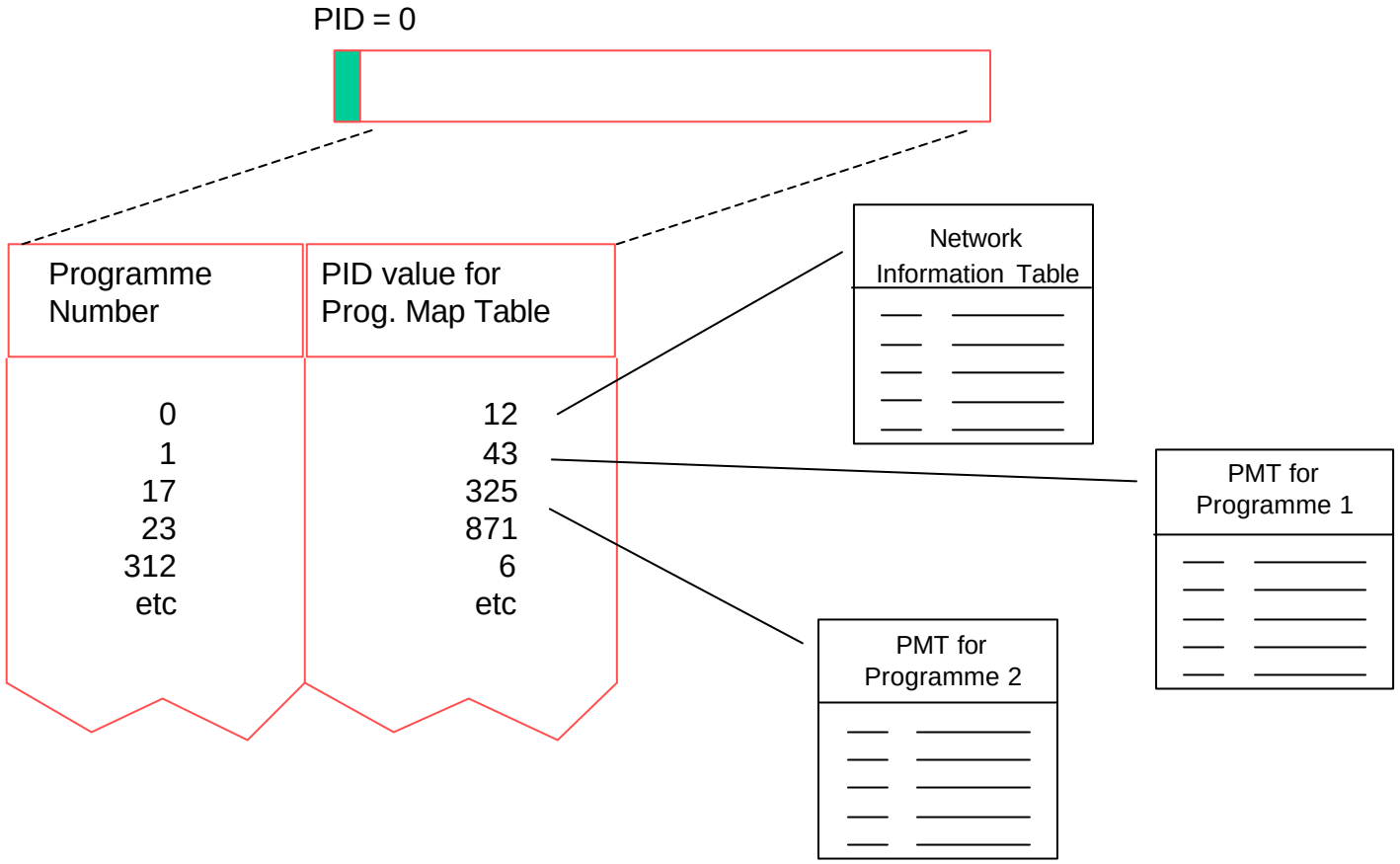
SONY

Programme Map Table



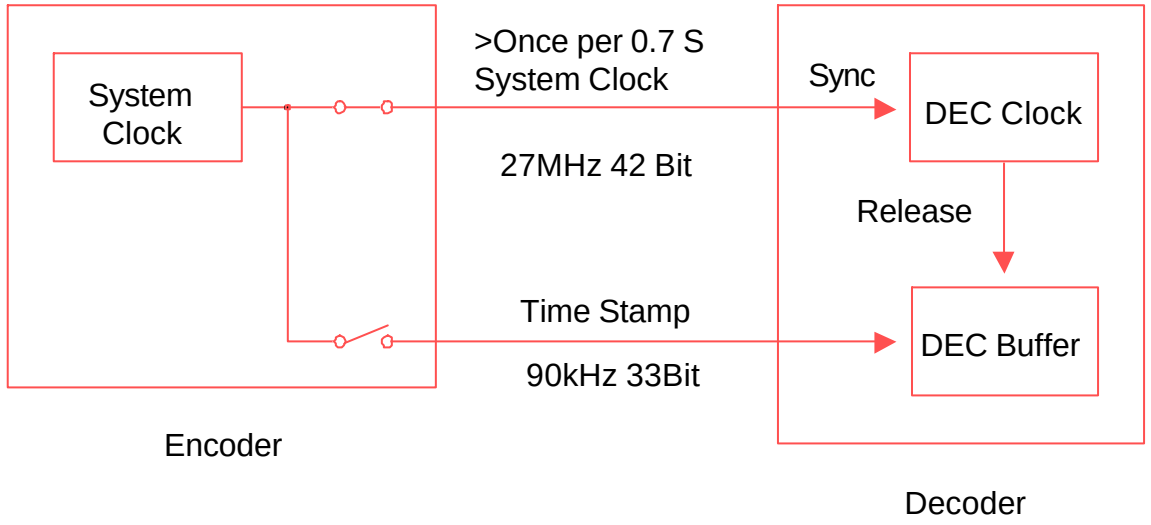
SONY

Programme Control



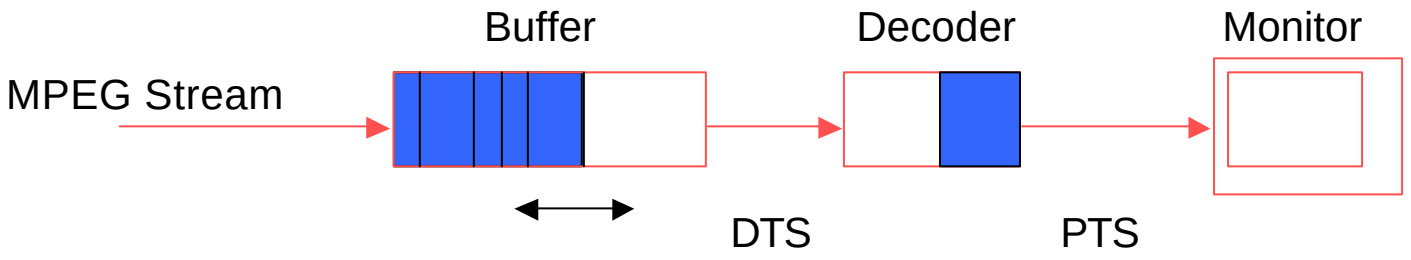
SONY

MPEG Clock System



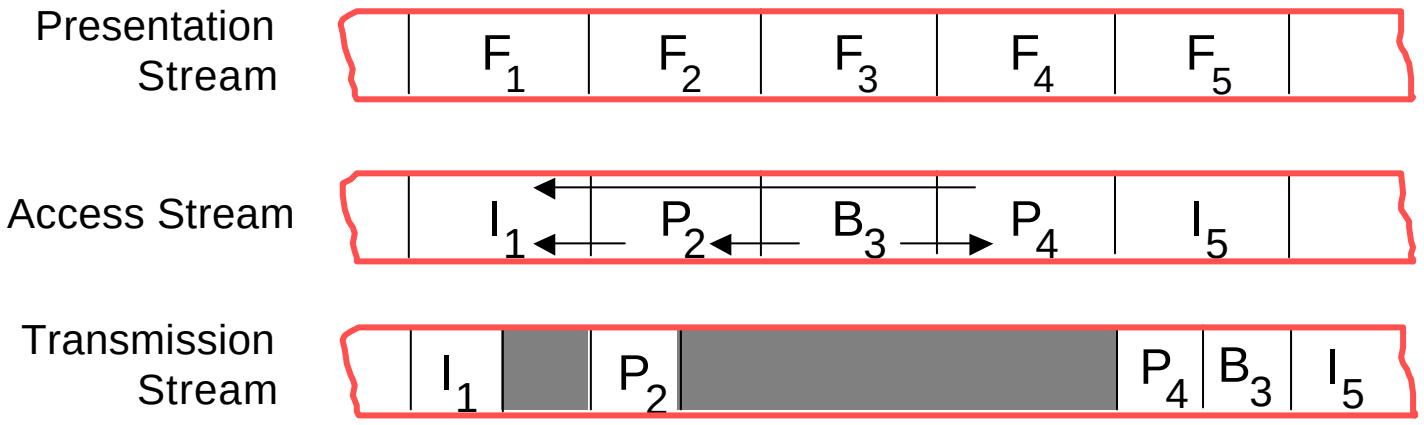
SONY

Time Stamps



SONY

MPEG Transmission Sequence



SONY

MPEG-4

- Originally intended for very low data rate for:
  - Video conferencing
  - Portable video phones
- Low data limit raised (now 5Kbps to 10Mbps)
- Wide range of resolutions supported
- Progressive and interlace supported
- Now intended for complex multimedia type applications

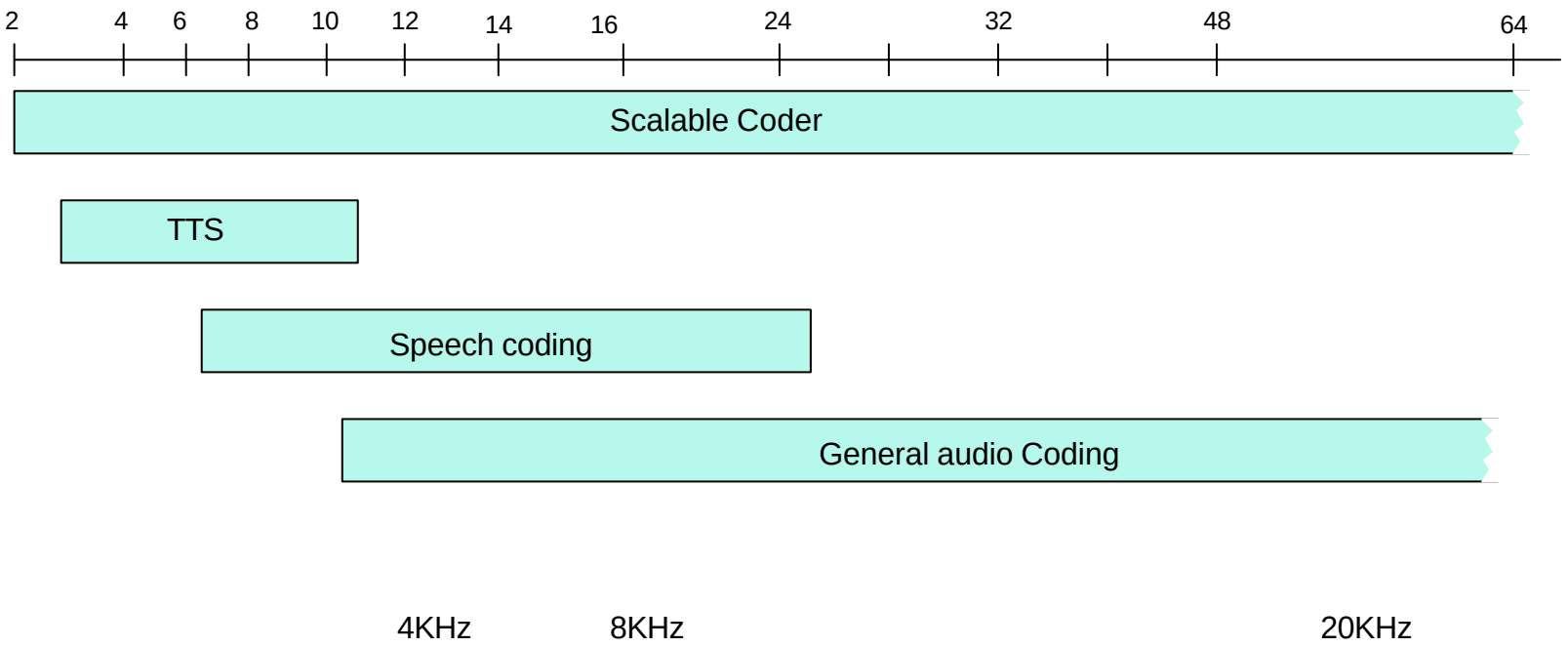
MPEG-4

- Audiovisual scene made up of media objects:
  - Objects can be:
    - Natural or Synthetic
    - Audio or Video
  - Objects are composited into the scene
  - Objects can interact with objects at the receiver's end

Audio Coding

- Profiles:
  - Speech
  - Synthesis
  - Scalable
  - Main
  - High quality Audio
  - Low delay audio
  - Natural audio
  - Mobile audio internetworking

Natural Audio Coding



Typical Audio Bandwidth

SONY

Audio Coding

- Synthetic Methods
  - TTS (Text to speech)
    - 200bps to 1200bps
    - text plus prosodic parameters
    - interface standard not normative sythesizer
  - Score driven
    - SAOL (Structured audio orchestra language)
    - Synthesis can be by many processes
      - wavetable, FM,additive and others
    - control is via “score”
    - MIDI can be used

Video Coding

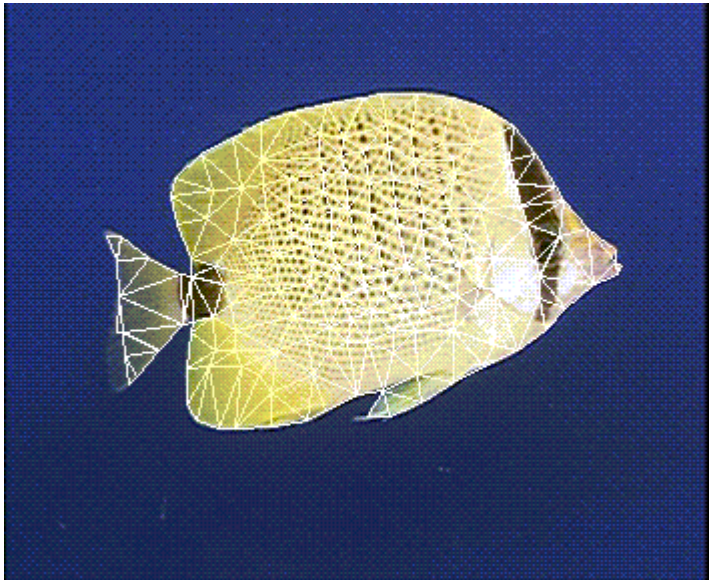
- Profiles
  - for natural objects
    - Simple
    - Simple scalable
    - Core
    - Main
    - N-bit
  - for synthetic objects
    - Simple facial animation
    - Scalable textural
    - Basic Animated 2D Texture
    - Hybrid visual

Video Coding

- Synthetic objects
  - parametric coding of
    - faces
    - bodies
    - static and dynamic meshes
- Face Animation
  - Facial description parameters
  - Facial animation parameters
  - Face animation table
  - Face interpolation technique

Video coding

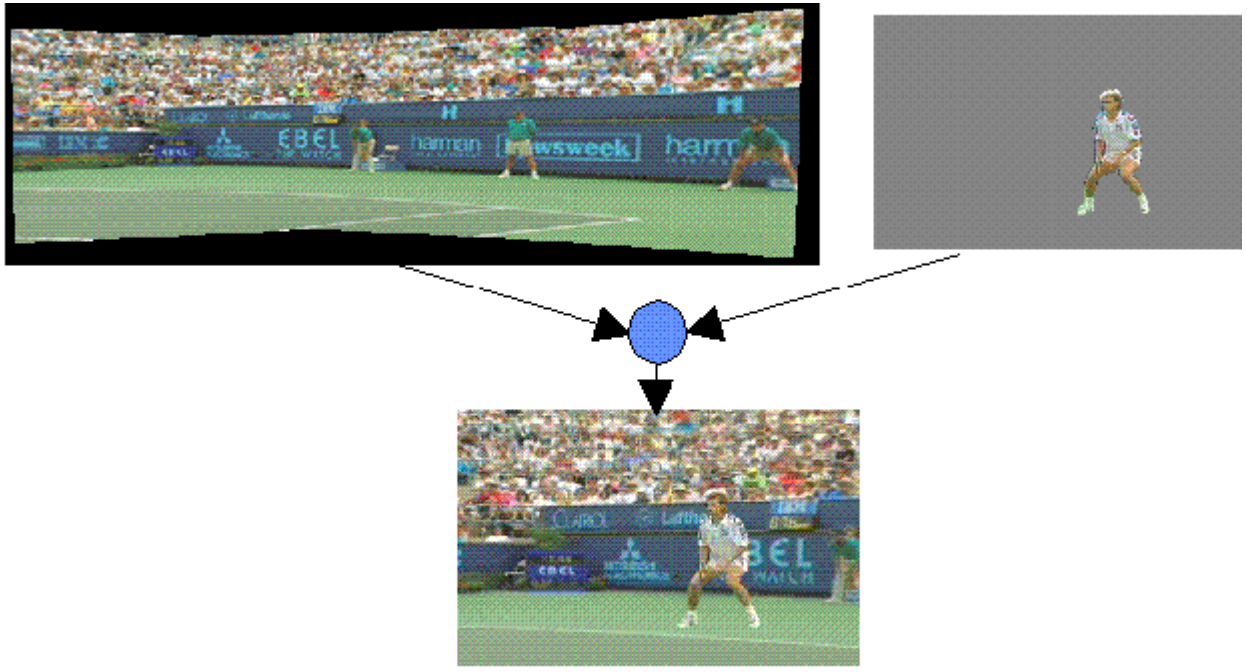
- 2D animated meshes
  - region of polygons
  - could have texture map



Example of a mesh from the MPEG-4 Standard

Sprite Coding

- MPEG-4 supports moving sprite over static background



Example of sprite coding from the MPEG-4 Standard

MPEG Products

- Broadcast
  - VTRs - Betacam SX, IMX
  - Servers - MAV-70, MAV-555
- Contribution
  - EBU Standard
- DVD
  - players and authoring systems
- Transmission
  - Satellite and terrestrial

MPEG Open World

- Pro-MPEG Forum
  - Consortium of many companies working toward greater Interoperabilty

www.pro-mpeg.org

Thank You

Thank You

SONY