

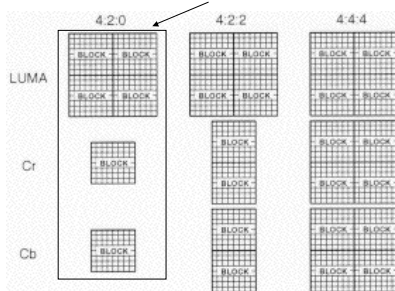
General Reading

- Halsall “Multimedia Communications”
 - Audio: pages 88-96, 173-189
 - Video: pages 96-110, 193-215
 - H.261/263: pages 203-215

H.261

- Standardised by ITU (in 1990)
- Data-rates: $n \times 64\text{kbps}$ ($1 \leq n \leq 30$)
- Designed for video-conferencing over ISDN (128kbps)
- 2 resolutions: QCIF (176x144) and optionally CIF (352x288)
- Colourspace: YUV4:2:0
- Uses transform coding, inter-frame prediction, motion compensation, variable-length coding

Basic unit: Macroblock



Group of Blocks (GOBs)

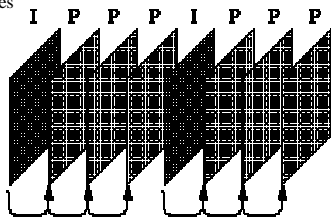
- GOB = 11x3 MacroBlocks
- CIF (2x6 GOBs) QCIF (1x3 GOBs)

1	2	1
3	4	2
5	6	3
7	8	
9	10	
11	12	

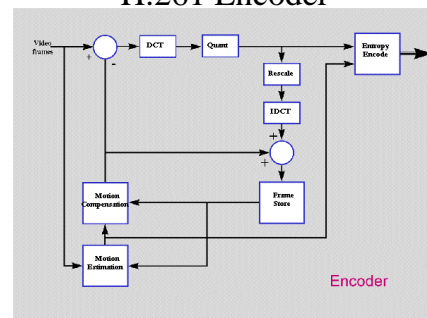
- Purpose: resync after transmission error

Frame types

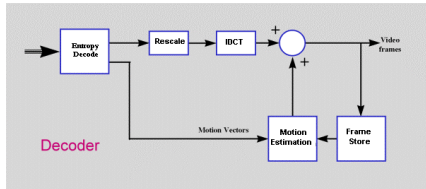
- I-frames (Intra): same as JPEG, except quantisation is by constant value for all DCT coefficients (ie no table).
- P-frames (Predicted): pseudo-differences from previous frames



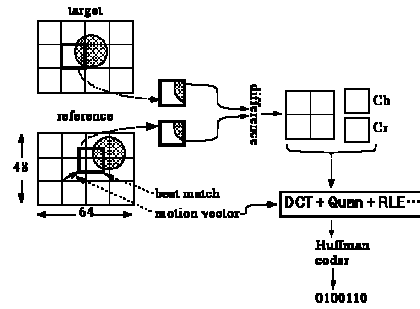
H.261 Encoder



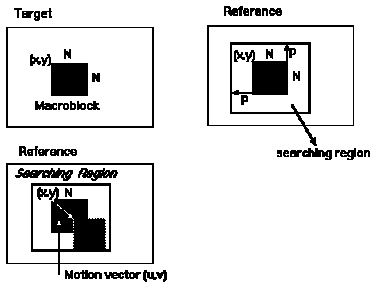
H.261 Decoder



Motion Estimation



Motion Vector Search



Mean Absolute Error (MAE)

- Goal is to find a vector (i, j) such that $MAE(i, j)$ is minimum, where

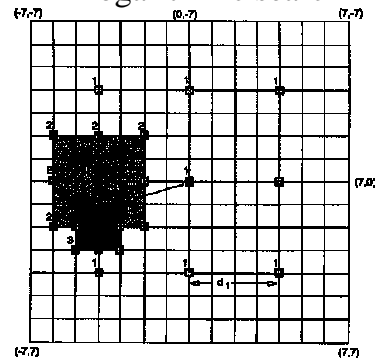
$$MAE(i, j) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} |C(x+k, y+l) - R(x+i+k, y+j+l)|$$

- $C(x+k, y+l)$ - pixels in the macroblock with upper left corner (x, y) in the Target.
- $R(x+i+k, y+j+l)$ - pixels in the macroblock with upper left corner $(x+i, y+j)$ in the Reference.

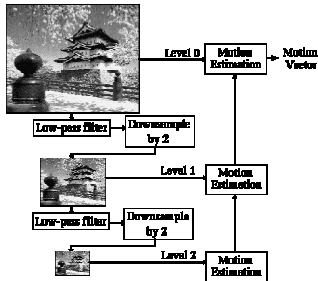
Full Search

- Search the whole $[-p, p]$ space
- Calculate MAE at each location
- Computationally expensive!

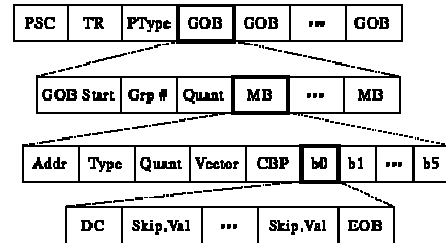
2D logarithmic search



Hierarchical Search



Bitstream structure



Macroblock format

Addr	Type	Quant	Vector	CBP	b0	b1	...	b5
------	------	-------	--------	-----	----	----	-----	----

- Many macroblocks will be exact matches (or close enough). So send address of each block in image -> *Addr*
- Sometimes no good match can be found, so send INTRA block -> *Type* (Intra/Inter?)
- Will want to vary the quantisation to fine tune compression, so send quantisation value -> *Quant*
- Motion vector -> *Vector*
- Some blocks in macroblock will match well, others match poorly. So send bitmask indicating which blocks are present (Coded Block Pattern, or *CBP*).
- b0...b5: Send the blocks (4 Y, 1 Cr, 1 Cb) as in JPEG

Higher level structures

- delineate boundaries between pictures: Picture Start Code (*PSC*)
 - timestamp (used for audio synchronization): Temporal Reference (*TR*)
 - CIF or QCIF? Picture Type (*Ptype*)
 - Might want to skip whole groups: Group Number (*Grp #*)
 - Might want to use one quantisation value for whole group, so send Group Quantisation Value (*Gquant*)
- [Bitstream is designed so we can skip data whenever possible while still unambiguous].

H.263

- Half pixel precision is used for motion compensation.
- Some parts of the hierarchical structure of the data-stream are now optional, so the codec can be configured for a lower data-rate or better error recovery.
- There are now four optional negotiable options included to improve performance: Unrestricted Motion Vectors, Syntax-based arithmetic coding, Advance prediction, and forward and backward frame prediction.
- Support for SQCIF, 4CIF, and 16CIF. SQCIF is approximately half the resolution of QCIF. 4CIF and 16CIF are 4 and 16 times the resolution of CIF respectively.