

H.261: A Standard for Video Conferencing Applications

Nimrod Peleg
Update: Dec. 2005

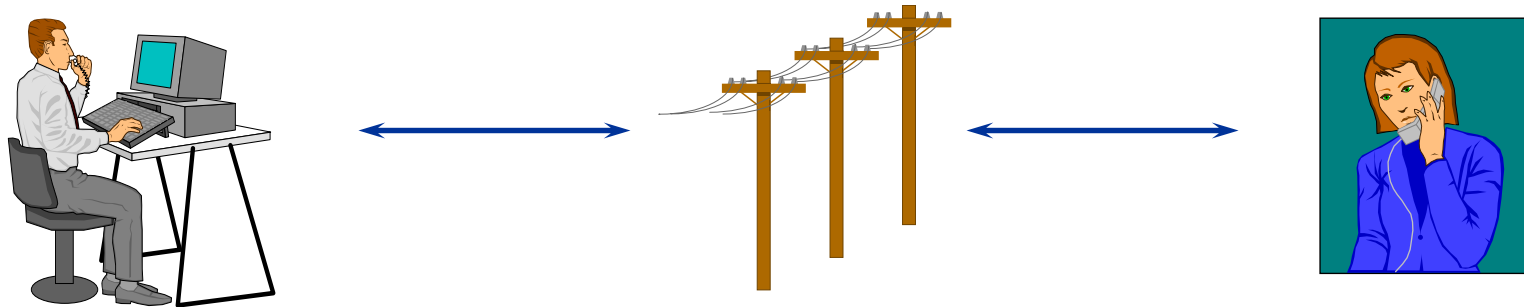


ITU - Rec. H.261 Target (1990)

- “... A Video compression standard developed to facilitate **videoconferencing** (and videophone) services over the integrated services digital network (**ISDN**) at $p \times 64Kbps$ ($p=1..30$) ...”
- Acceptable quality usually above $p=6$ (384Kbps)
- Maximum bitrate over ISDN is 1.92Mbps ($p=30$), better than VHS-quality !

Important Features

- Maximum coding *delay* of 150mSec., due to the need for bi-directional communication.



- *Low-cost* VLSI implementation is possible.



Input Image Format

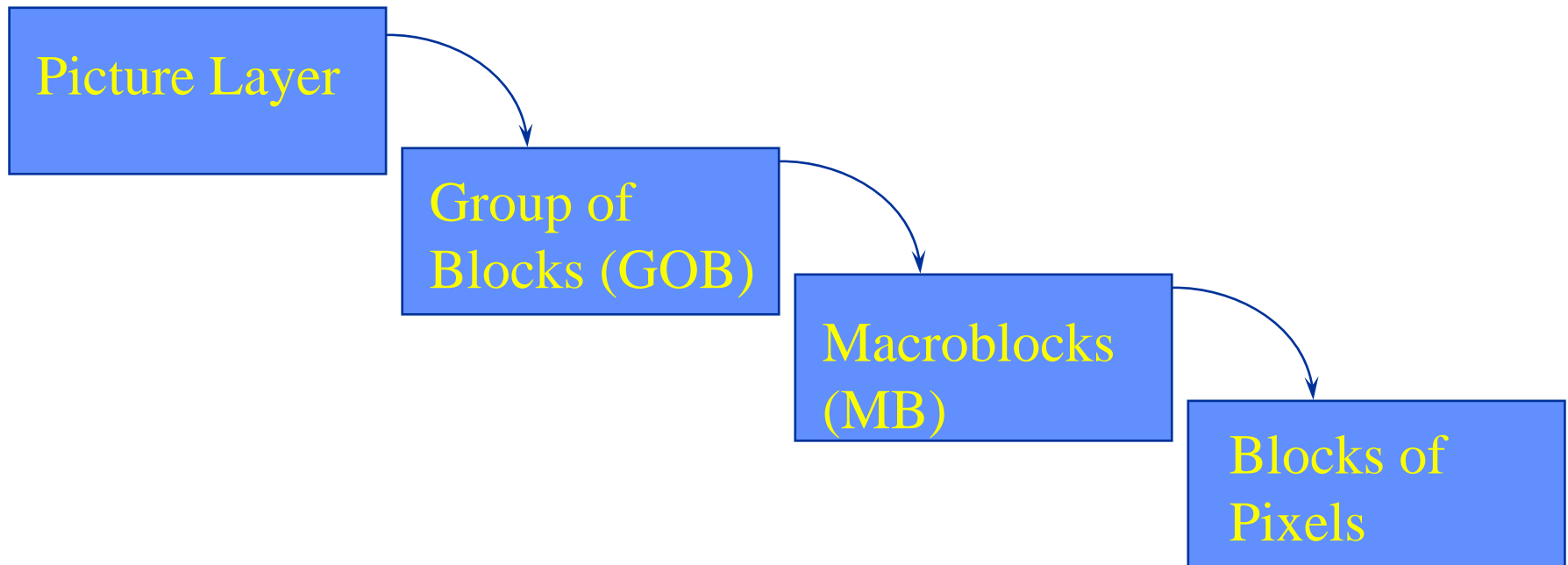
- To enable use of both 525-lines and 625-lines TV standards, a new input format was defined: *Common Intermediate Format (CIF)*
- Maximum rate: CIF, 30fps \longrightarrow 37.3Mbps
for 384Kbps channel rate, 54:1 compression ratio needed
- Minimum rate:, QCIF, 7.5fps \longrightarrow 2.3Mbps
for 64Kbps channel rate, 36:1 compression ratio needed

Input Image Format (Cont'd)

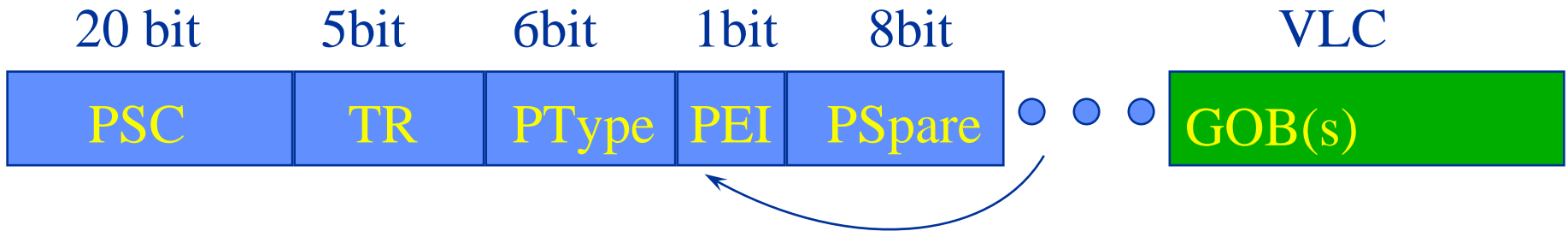
	CIF	QCIF
<u>Active pels/line</u>		
Lum (Y)	360(352)	180(176)
Chroma (U,V)	180(176)	90(88)
<u>Active Lines/picture</u>		
Lum (Y)	288	144
Chroma (U,V)	144	72
Interlacing/Aspect Ratio	1:1 / 4:3	1:1 / 4:3
Temporal Rate	30,15,10,7.5	30,15,10,7.5

Video Multiplex

- Decoder should interpret the received bit stream without any ambiguity
- Hierarchical structure:



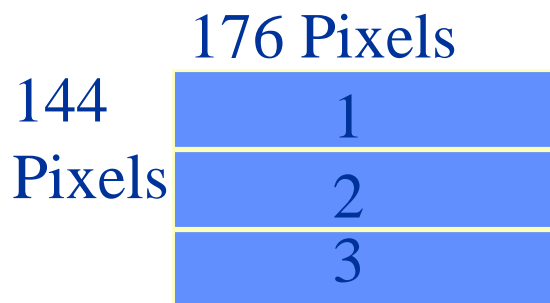
Video Multiplex: Picture Layer



- Picture Start Code: fix word (00010H) .
- Temporal Reference: Position of the picture in the sequence (zero's every 32 pictures !).
- PType: Picture format (CIF, QCIF, NTSC) and type.
- Picture Extra Information: Signaling if PSpare exists.
- Picture Spare: Spare information, repeated by PEI till PEI=0.

Video Multiplex: GOB

GOB Layer: Every picture is divided into 12 GOBs for CIF or 3 GOBs for QCIF:



QCIF



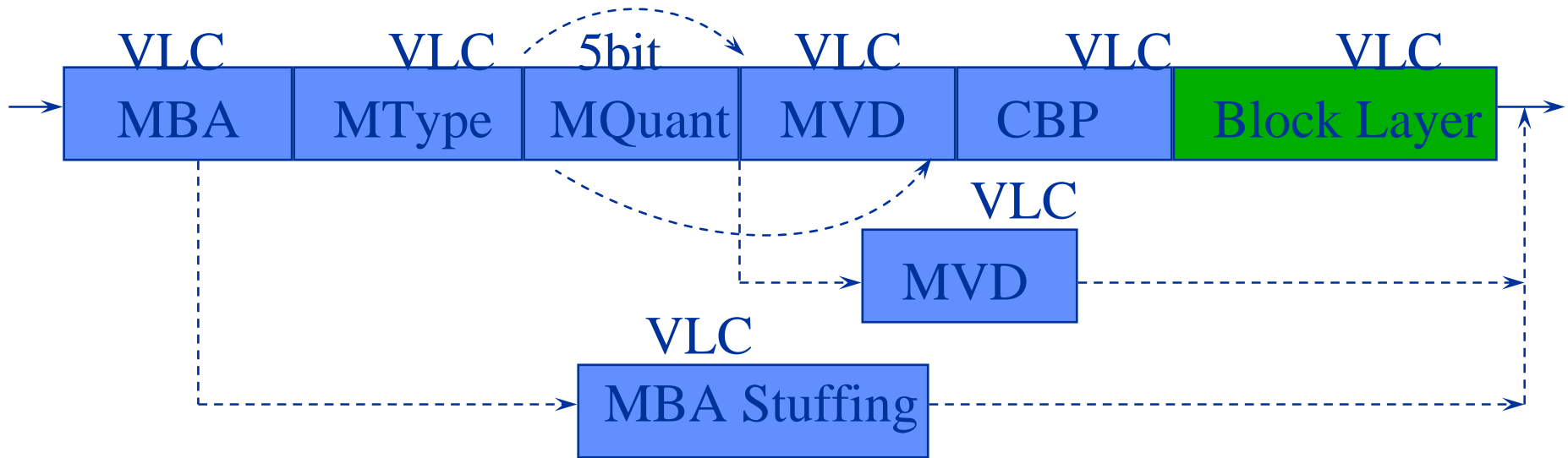
CIF

Video Multiplex: GOB (Cont'd)



- GOB Start Code: fix word (0001H) .
- GOB Number: Position of the group in the picture (zero's every 16 GOBs !).
- GQuant: GOB Quantization step (step size= $2 * GQuant$), fixed till changed by MQuant (see later).
- GOB Extra Information: Signaling if GSpare exists.
- GOB Spare: Spare information, repeated by GEI till $GEI=0$.

Video Multiplex: MB (Cont'd)



- **MacroBlock Address:** Position within the GOB, 1st MB has absolute address, others: differential.

Video Multiplex: MB (Cont'd)

- MType: Information about coming MB (Inter or Intra, MV included or not, MQuant exists, etc.)
- MQuant: Replacing GQuant till the end of the GOB or a new Mquant.
- Motion Vector Data: Motion vector for the MB, relative to the former picture and differential from former MB. Absolute value in several cases:
 - MB is first in the line (1, 12, 22).
 - Former MB is not attached (MBA not 1).
 - Last MB was not of MC type .

Video Multiplex: MB (Cont'd)

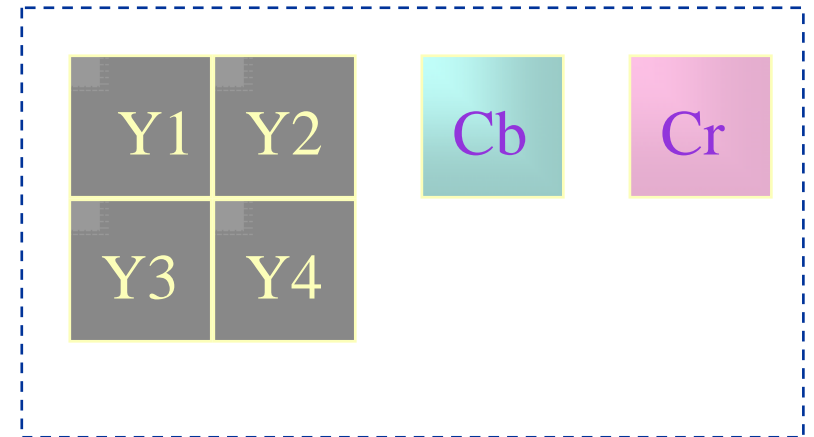
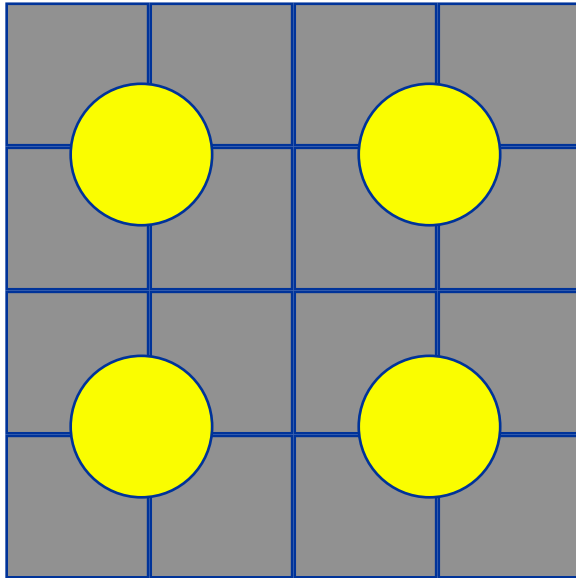
- The MV includes two words: Horizontal change and Vertical change
- Coded Block Pattern: Shows which blocks in the MB were transmitted:

$$CBP = 32P_1 + 16P_2 + 8P_3 + 4P_4 + 2P_5 + P_6$$

$$P_n = \begin{cases} 1 & \text{At least one coeff. was transmitted} \\ 0 & \text{No coeff. transmitted} \end{cases}$$

Video Multiplex: Block Layer

- A MB contains 6 Blocks, 8x8 pixels each:
4 Luminance (Y) and 2 Chrominance (Cb,Cr)



Composition of MacroBlock

Position of Lum. And Chroma Pixels

Video Multiplex: Block (Cont'd)

- Coeff. are **Run-Length** , **Huffman** coded.
- For **Intra** Blocks, all 64 coeff. transmitted.
- All other cases: CBP points which blocks are transmitted.
- Coeff. consists of 2 words: *Run* and *Level* according to *Zig-Zag* scan.
- Every block ends with the code: 1H .

Video Compression Algorithm

- Two main modes:
 - *Intra Mode*: JPEG-like compression.
 - *Inter Mode*: Temporal prediction employed, with or without MC. Then, prediction error is DCT encoded.
- For each mode, several options can be selected (quantization, filters etc.)

Inter frame coding steps

- Estimate (one) MV for each MB, max. value: ± 15 .
 - motion estimation technique is NOT mentioned !
- Select a compression mode for each MB, based on Displaced Block Difference criterion (*dbd*):

$$dbd(x,k) = b(x,k) - b(x-d, k-1)$$

b: block *x*: pixel coordinates *k*: time index

d: displacement vector (*k* frame vs. *k-1*)

if *d=0*, then *dbd* becomes block difference (*bd*)

- Process each MB to generate header + data bitstream, according to **chosen compression mode**.

Compression modes

<u>Prediction</u>	<u>MQuant</u>	<u>MVD</u>	<u>CBP</u>	<u>TCoeff</u>	<u>Code</u>
Intra				+	0001
Intra	+			+	0000 001
Inter			+	+	1
Inter	+		+	+	0000 1
Inter+MC		+			0000 0000 1
Inter+MC		+	+	+	0000 0001
Inter+MC	+	+	+	+	0000 0000 01
Inter+MC+Fil		+			001
Inter+MC+Fil		+	+	+	01
Inter+MC+Fil	+	+	+	+	0000 01

Compression modes (Cont'd)

Table codes:

- MQuant: + indicates a new value.
- MVD: Motion vector data exists.
- CBP: If at least one transform coeff. is transmitted.
- TCoeff: Transform coeff. are encoded.
- Code: indicating the compression mode.

Compression modes (Cont'd)

- *Inter + MC* is selected if $\text{var}(dbd) < bd$
Transmission of the prediction error (*TCoeff*) is optional.
- Otherwise, no MV sent. If original MB has a small variance, *Intra* mode selected (DCT computed). In both *Inter* and *Inter+MC* blocks, prediction error is DCT encoded.
- For MC blocks, prediction error can be modified by 2-D (separable) *spatial Filter* .

Coding Model

- Quantized coefficients are *Zig-Zag* scanned, and *Events* are defined and then entropy coded.
- Events are defined as combination of run-length of zero coeff. preceding a non-zero coefficient.

That is:

Event = (Run, Level)

Rate and Buffer Control

Options for rate control are:

- PreProcessing
- Quantizer step size
- Block significance criterion
- Temporal sub-sampling

All options are NOT subject to the recommendation !

H.263 Demo ...

