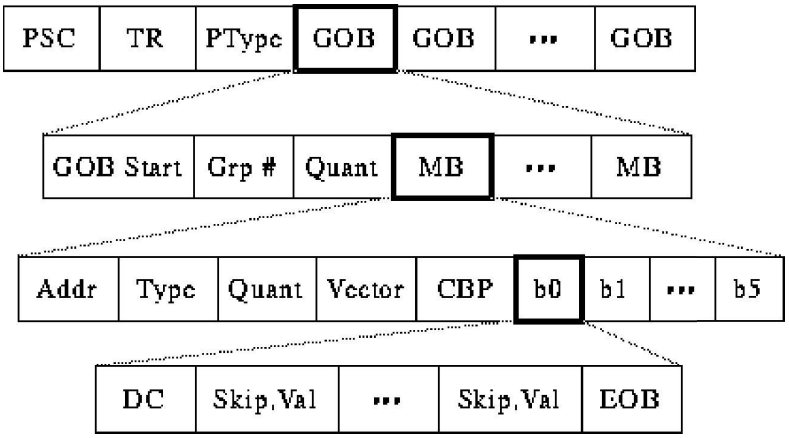
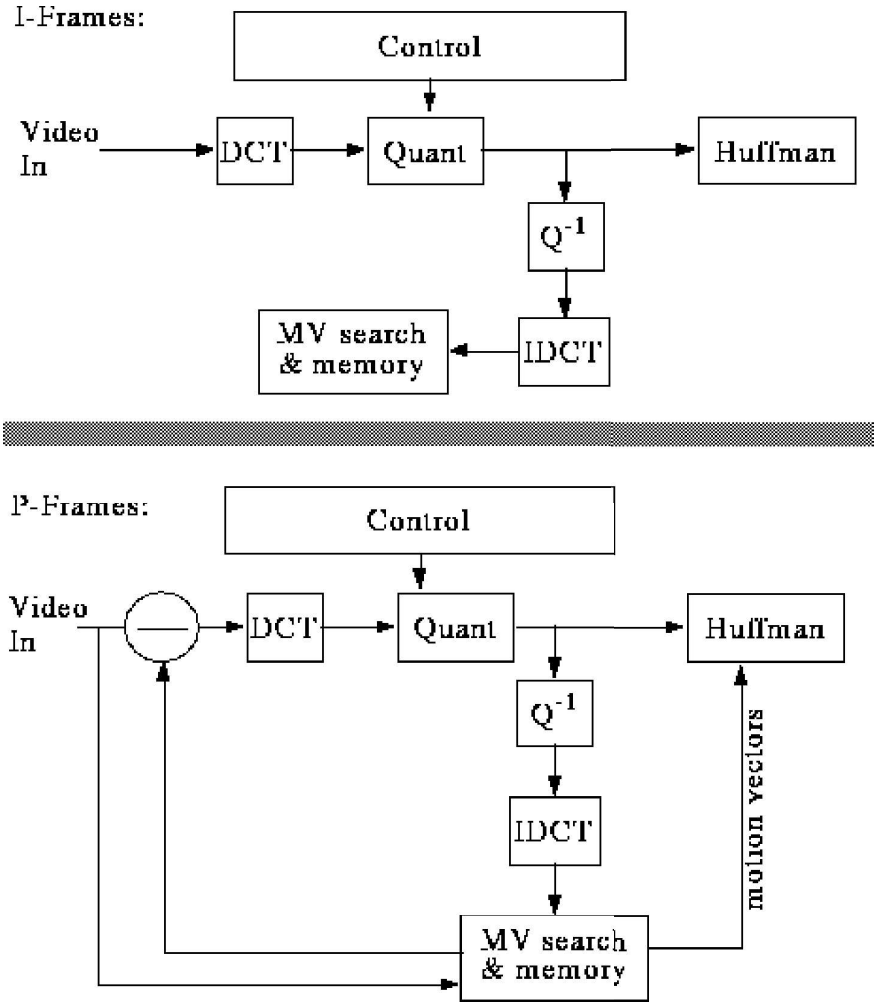
The H.261 Bitstream Structure

The H.261 Bitstream structure may be summarised as follows:



- Need to delineate boundaries between pictures, so send Picture Start Code -> *PSC*
- Need timestamp for picture (used later for audio synchronization), so send Temporal Reference -> *TR*
- Is this a P-frame or an I-frame? Send Picture Type -> *PType*
- Picture is divided into regions of 11x3 macroblocks called Groups of Blocks -> *GOB*
- Might want to skip whole groups, so send Group Number (*Grp #*)
- Might want to use one quantization value for whole group, so send Group Quantization Value -> *GQuant*
- Overall, bitstream is designed so we can skip data whenever possible while still unambiguous.

The overall H.261 Codec is summarised in Fig below.

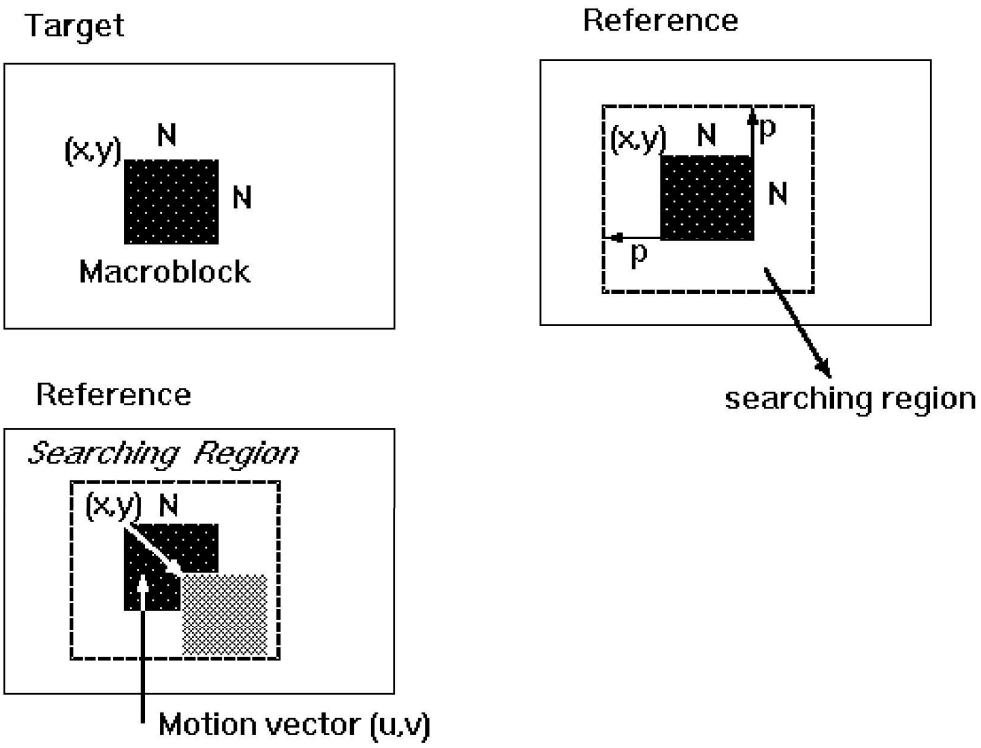


Hard Problems in H.261

There are however a few difficult problems in H.261:

- **Motion vector search**
- **Propagation of Errors**
- **Bit-rate Control**

Motion Vector Search



- $C(x+k,y+i)$ - pixels in the macro block with upper left corner (x,y) in the Target.

  $R(X+i+k,y+j+l)$ - pixels in the macro block with upper left corner $(x+i,y+j)$ in the Reference.

  Cost function is:

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

  Where MAE stands for *Mean Absolute Error*.

- Goal is to find a vector (u, v) such that MAE (u, v) is minimum

Full Search Method:

1. Search the whole $[-p, p]$ searching region.

2. Cost is:

$$\frac{IJF}{N^2}(2p+1)^2 \times N^2 \times 3$$

operations, assuming that each pixel comparison needs 3 operations (Subtraction, Absolute value, Addition).

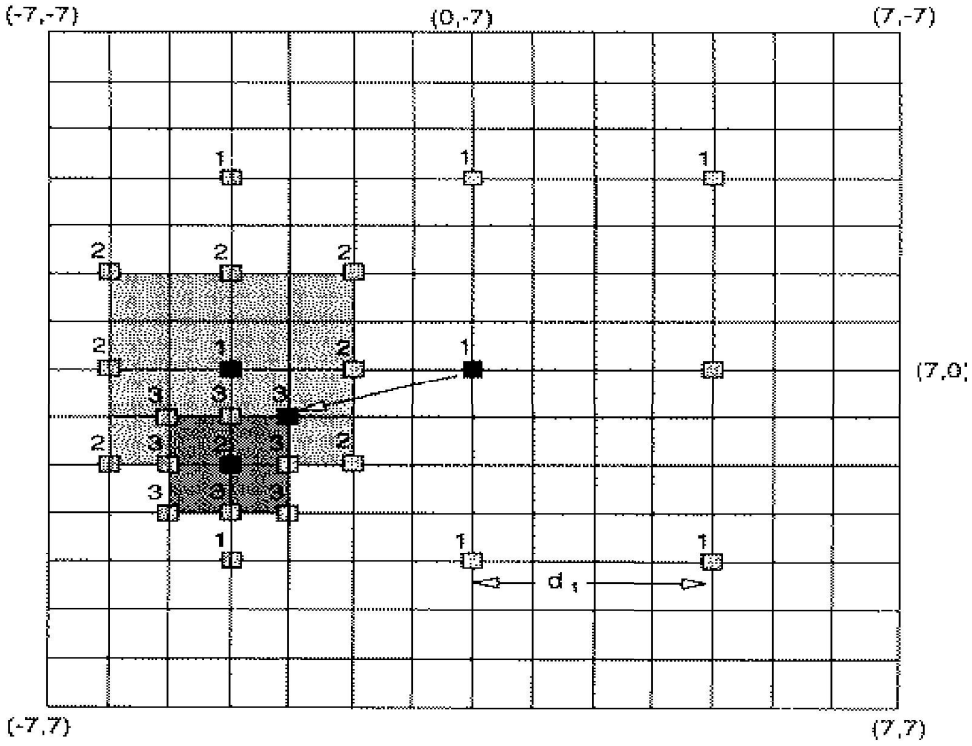
- **Two-Dimensional Logarithmic Search:**

Similar to binary search. MAE function is initially computed within a window of $[-p/2, p/2]$ at nine locations as shown in the figure.

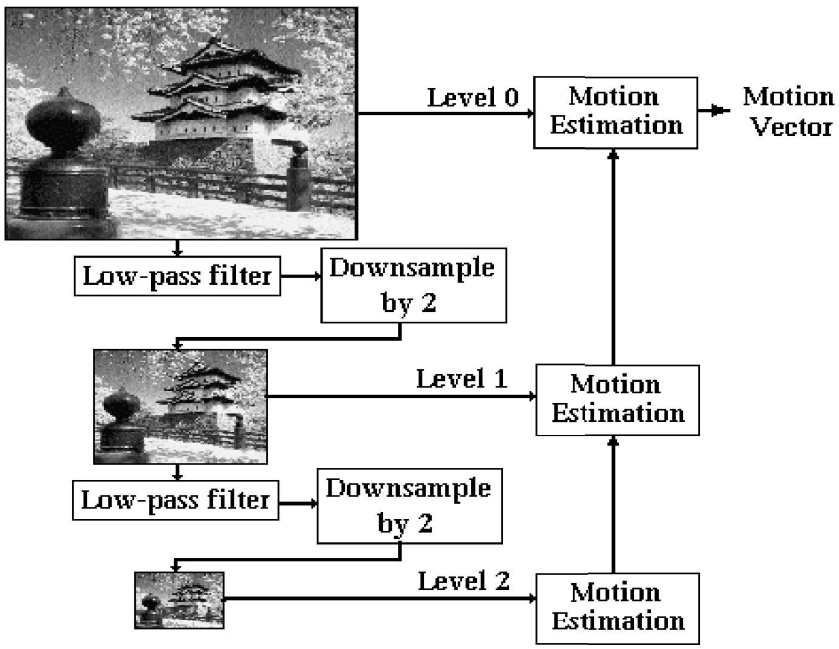
Repeat until the size of the search region is one pixel wide:

1. Find one of the nine locations that yields the minimum MAE.

2. Form a new searching region with half of the previous size and centered at the location found in step 1.



- **Hierarchical Motion Estimation:**



1. Form several low resolution version of the target and reference pictures
2. Find the best match motion vector in the lowest resolution version.
3. Modify the motion vector level by level when going up.

Performance comparison:

Search Method	Operation for 720x480 at 30 fps	
	p = 15	p=7
Full Search	29.89 GOPS	6.99 GOPS
Logarithmic	1.02 GOPS	777.60 MOPS
Hierarchical	507.38 MOPS	398.52 MOPS

Propagation of Errors

- Send an I-frame every once in a while
- Make sure you use decoded frame for comparison

Bit-rate Control

- Simple feedback loop based on "buffer fullness"

  If buffer is too full, increase the quantization scale factor to reduce the data.