

---

# Video CODEC Basics

---

Tony Gladvin George

[www.drtonygeorge.com/video\\_codec.htm](http://www.drtonygeorge.com/video_codec.htm)

---

# Agenda

- Basics Video Compression
  - New algorithms in Standards
- Implementation strategies
- Future trends

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# Why Video Compression

- A single uncompressed color image or video frame with a medium resolution of 500 x 500 pixels would require 100 s for transmission over an ISDN (Integrated Services Digital Network) link having a capacity of 64 Kbps.

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# How Video compression is possible

- Spatial Redundancy
  - Correlation between adjacent data points
- Temporal Redundancy
  - Correlation between different frames in an image
- Spectral Redundancy
  - Correlation between different color planes or sensors
  - Limitation of Low-level Human Vision System
- Physco-visual redundancy
  - Insensitive to high frequency components

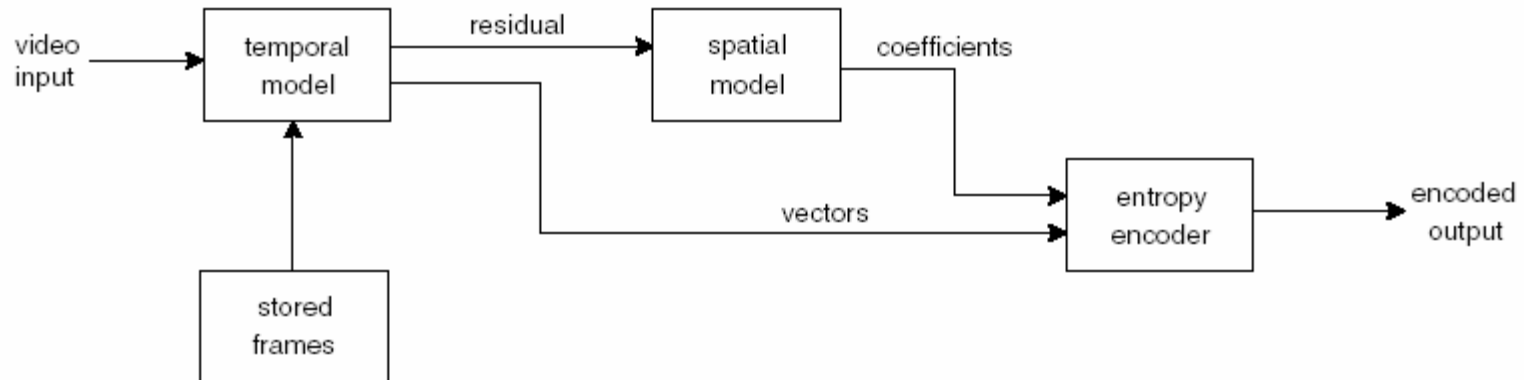
Using strictly spatial redundancy (JPEG) gives video compression ratios from 7:1 to 27:1

Taking advantage of temporal redundancy in video gives 20:1 to 300:1 compression for H.261, or 30:1 to 100:1 for high quality MPEG-2



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

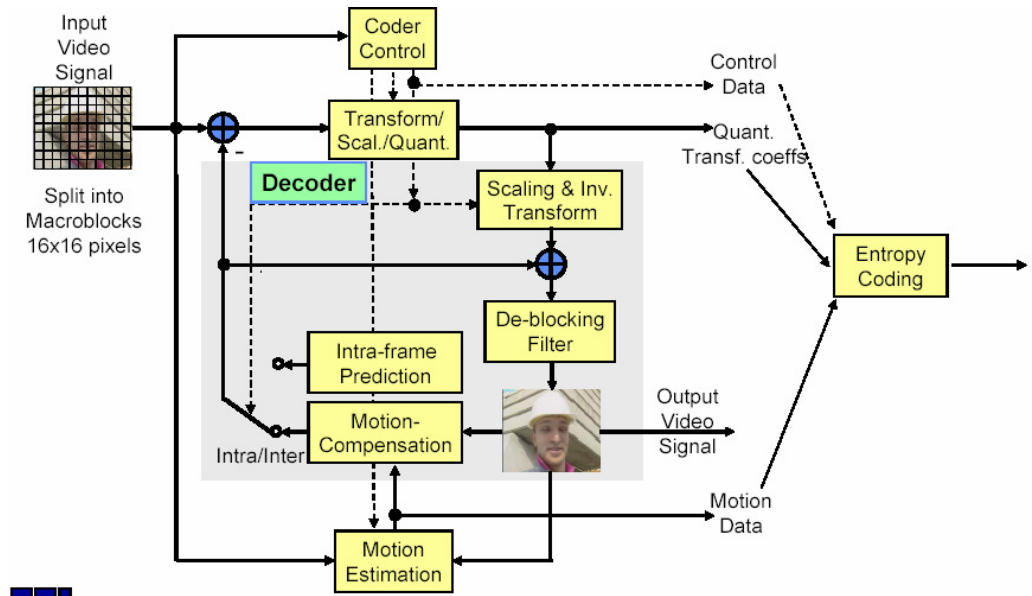
# Video encoder



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

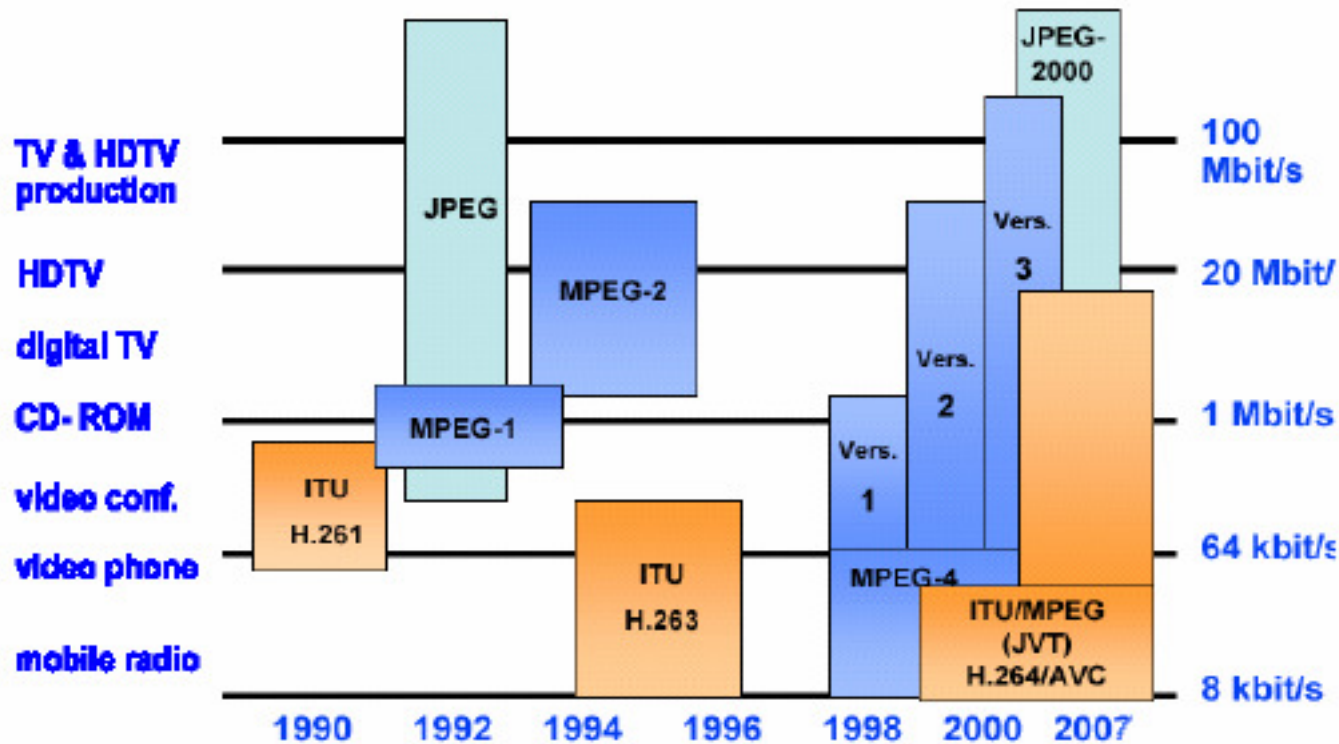
# Video CODEC is a mixture of algorithmic tools

- Tool 1 – Intra frames
- Tool 2 – Motion estimation
- Tool 3 – Search strategies
- Tool 4 – Block Matching
- Tool 5 – B frames
- Tool 6 - Transformation
- Tool 7 - Quantization
- Tool 8 – Entropy coding
- Tool 9 – Deblocking filter



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

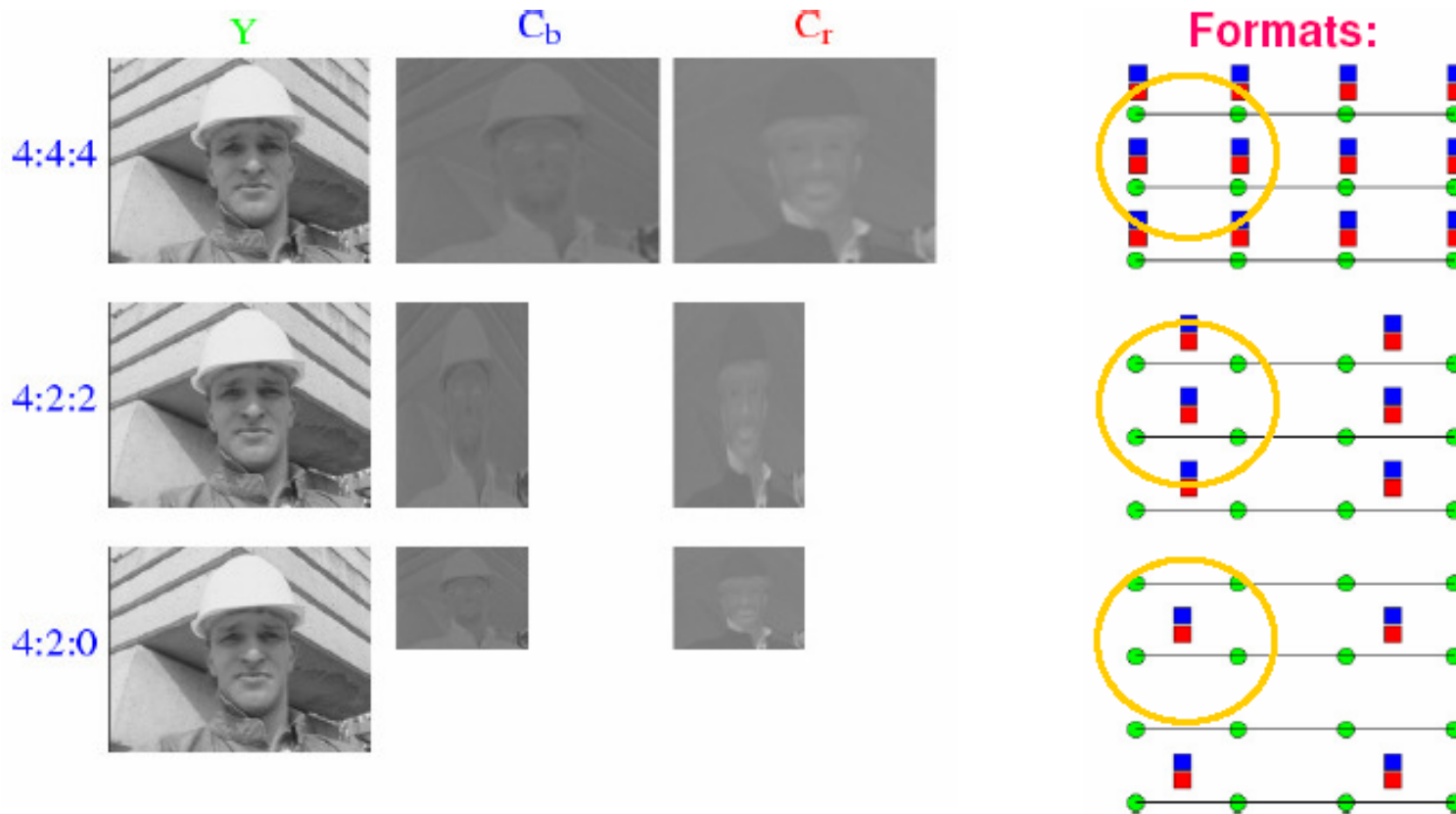
# Compression Standards Overview (ISO/ITU)



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Chrominance sub-sampling and formats

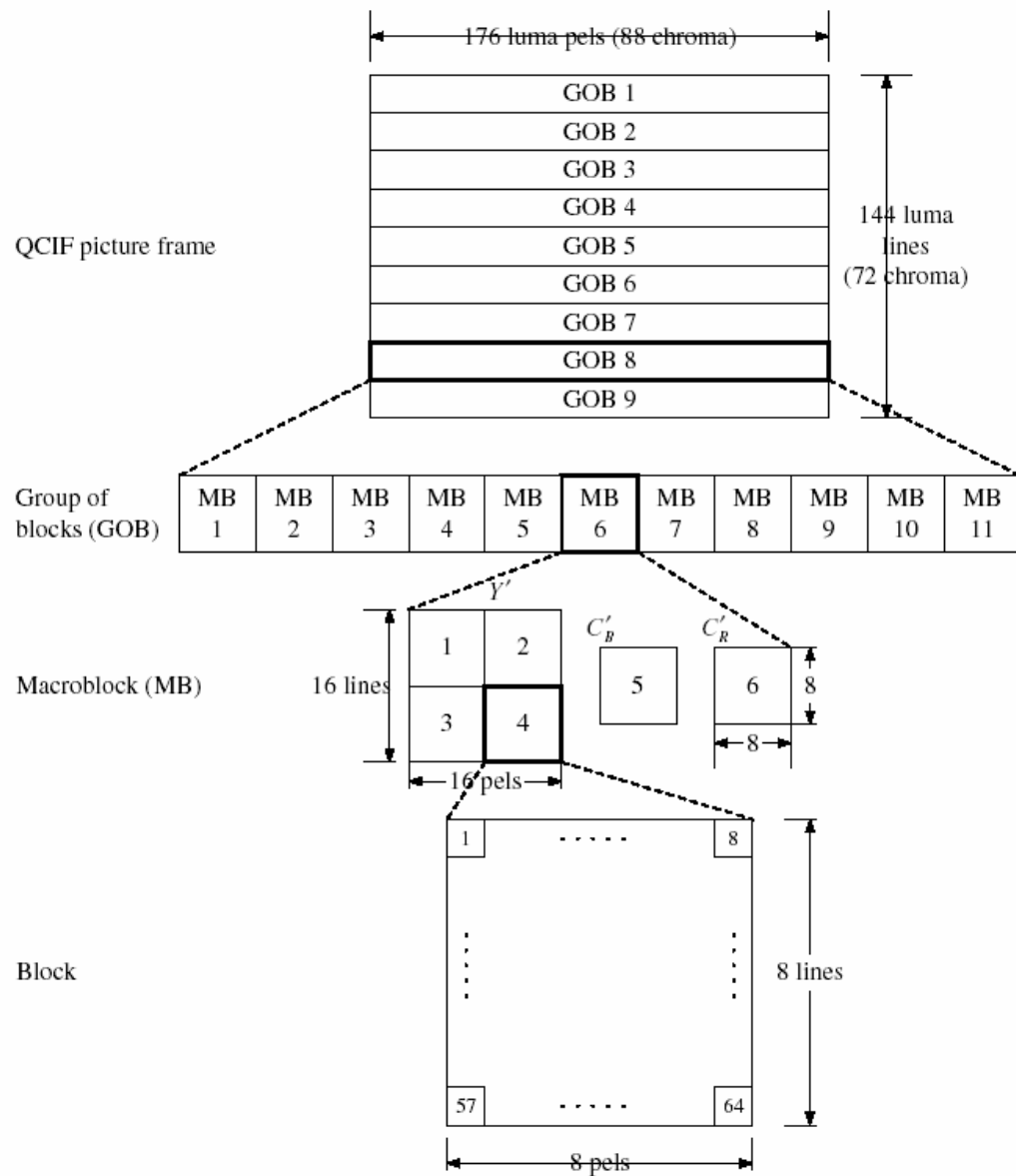
- Luminance/Chrominance commonly used, with Chrominance subsampled due to human vision insensitivity



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



# Picture structure

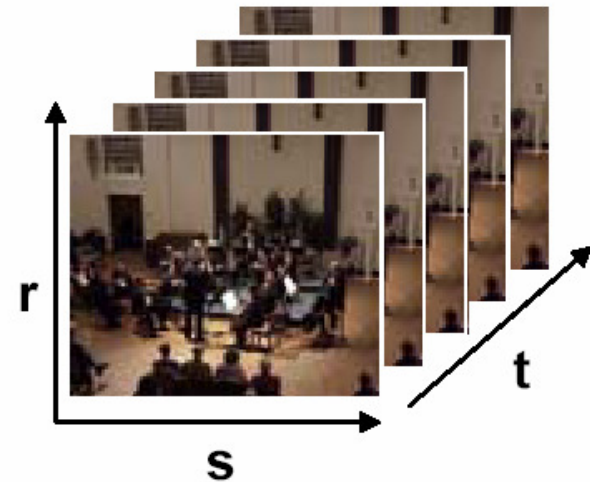


- Processing is done in block level

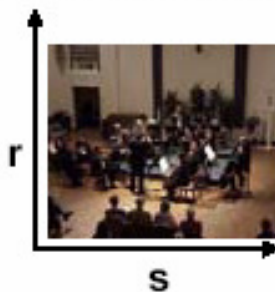
More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Compression by Intra and Inter frame

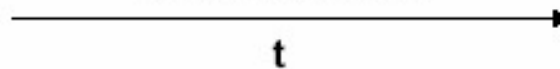
- Intra – within the frame (Spatial)
- Inter – across the frames (Temporal)



**Spatial Redundancy**  
( **Intra-Frame** Correlation )  
self-similarities within one image



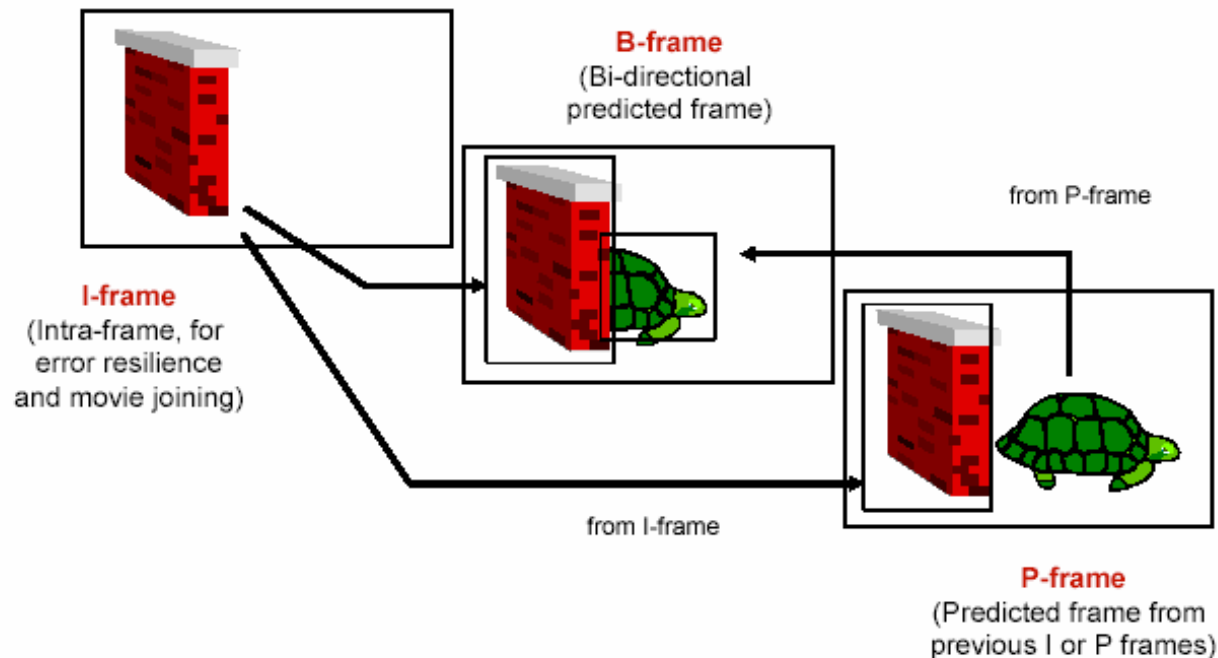
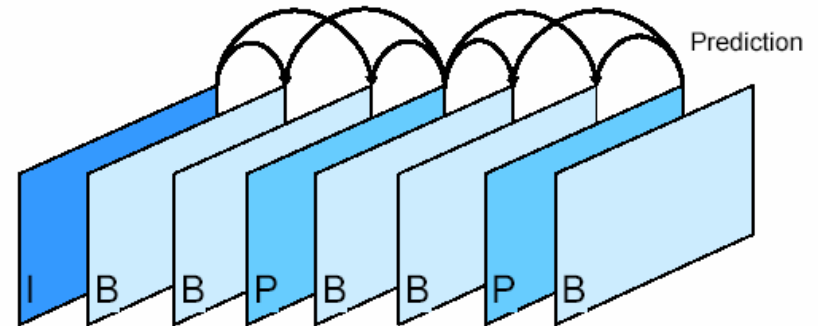
**Temporal Redundancy**  
( **Inter-Frame** Correlation )  
similar content  
in subsequent images



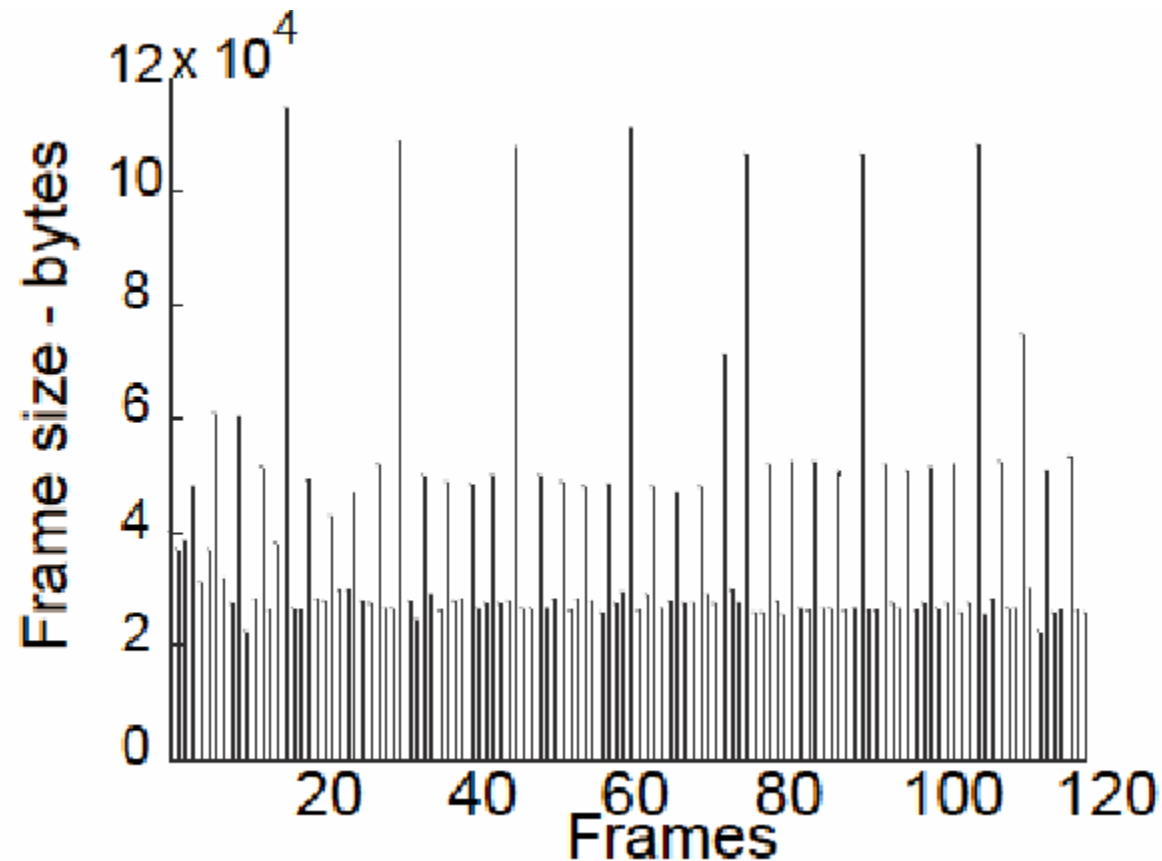
More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# I P B frames

- Three types of frames
- I frame – key frame
- P frame – Predicted frame
- B frame – image order and encoding order are different



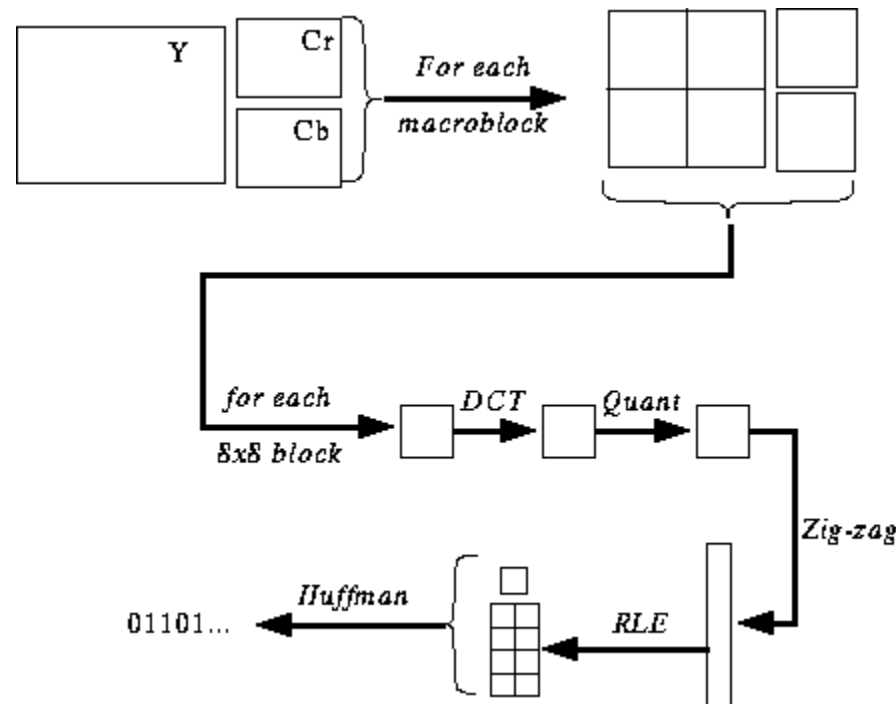
## Size of I, P, B frames



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

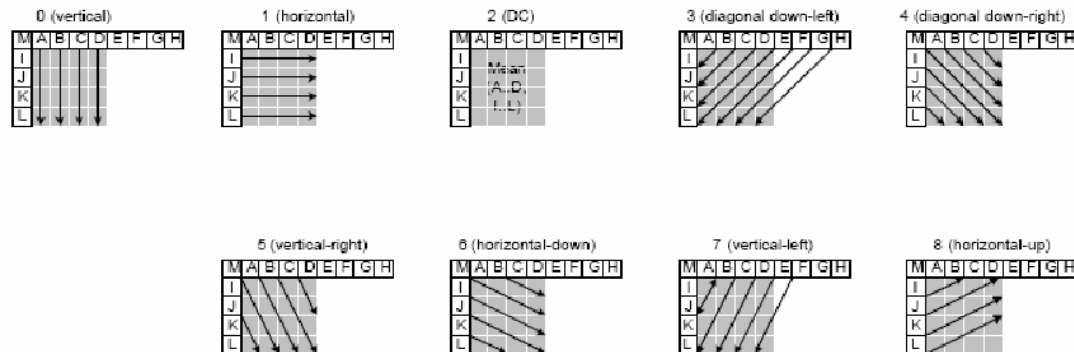
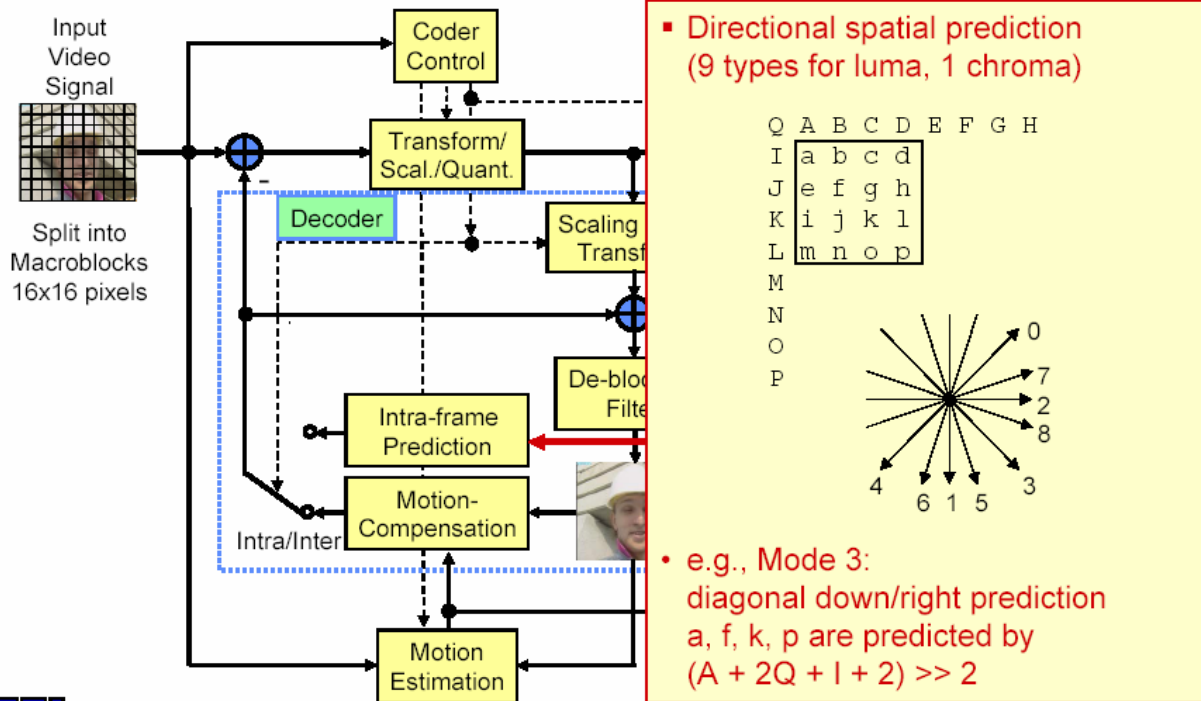
# Intra Frame coding (Tool 1)

- Macroblocks are 16 x 16 pixel areas on Y plane of original image. A macroblock usually consists of 4 Y blocks, 1 Cr block, and 1 Cb block.
- Quantization is by constant value for all DCT coefficients



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

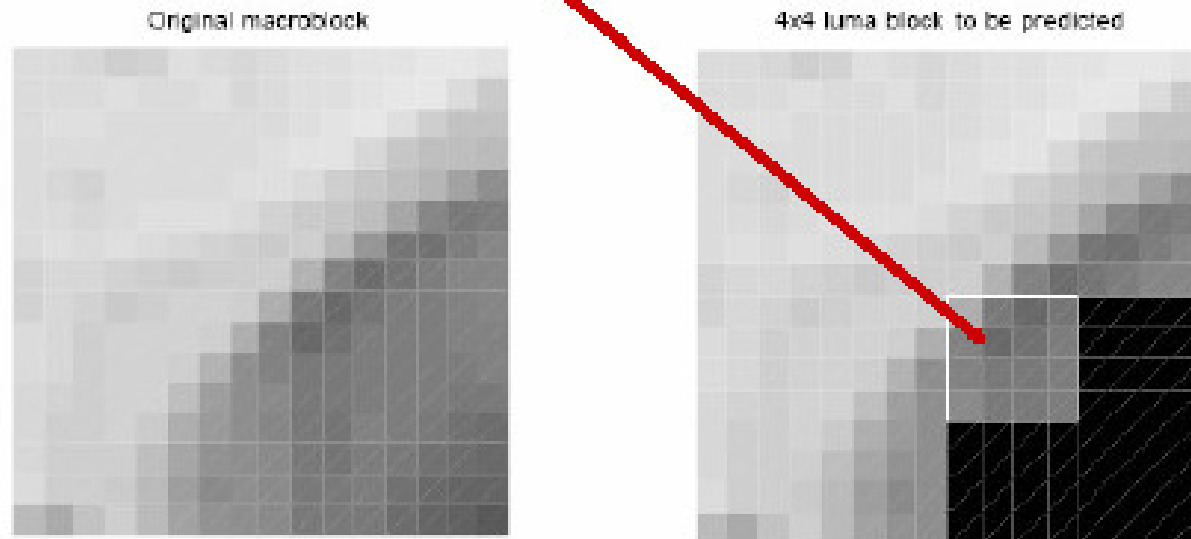
# Intra frame in H.264 (I frame)



# Prediction of Intra 4x4 block (I frame)

## □ Example

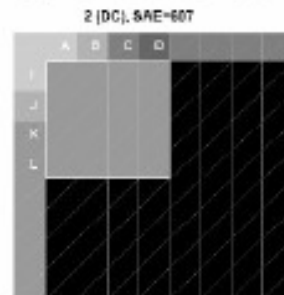
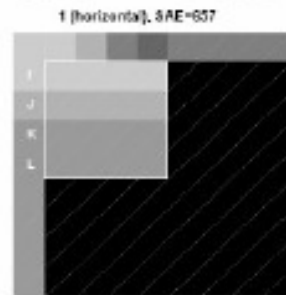
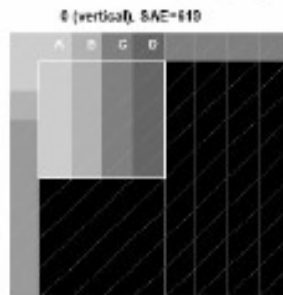
- Predict the following 4x4 block



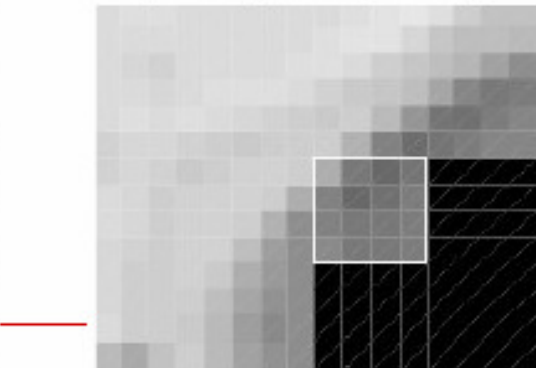
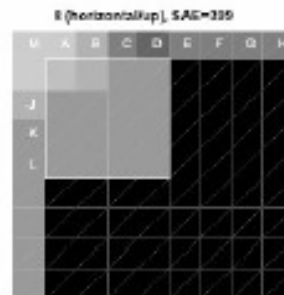
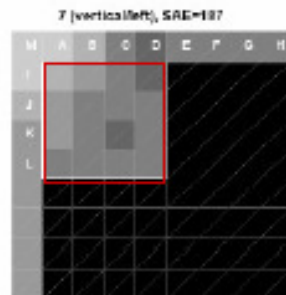
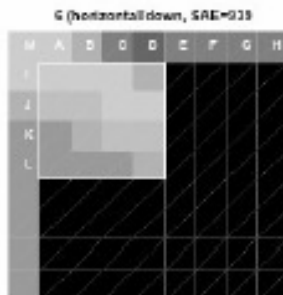
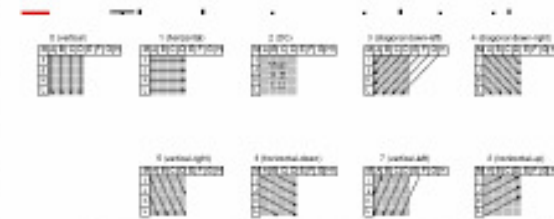
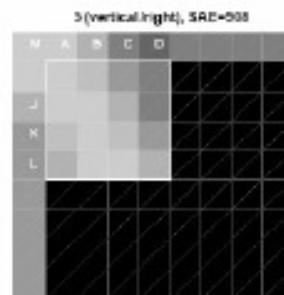
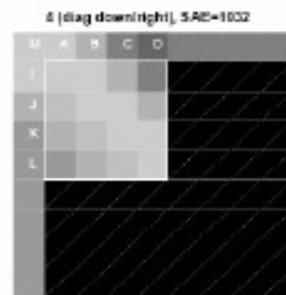
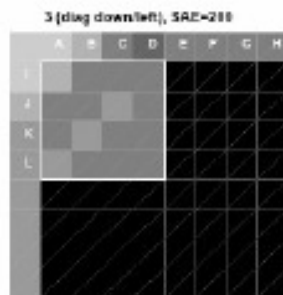
More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Prediction of Intra 4x4 block (I frame)

- The 9 prediction modes (0-8) are calculated for the 4x4 block



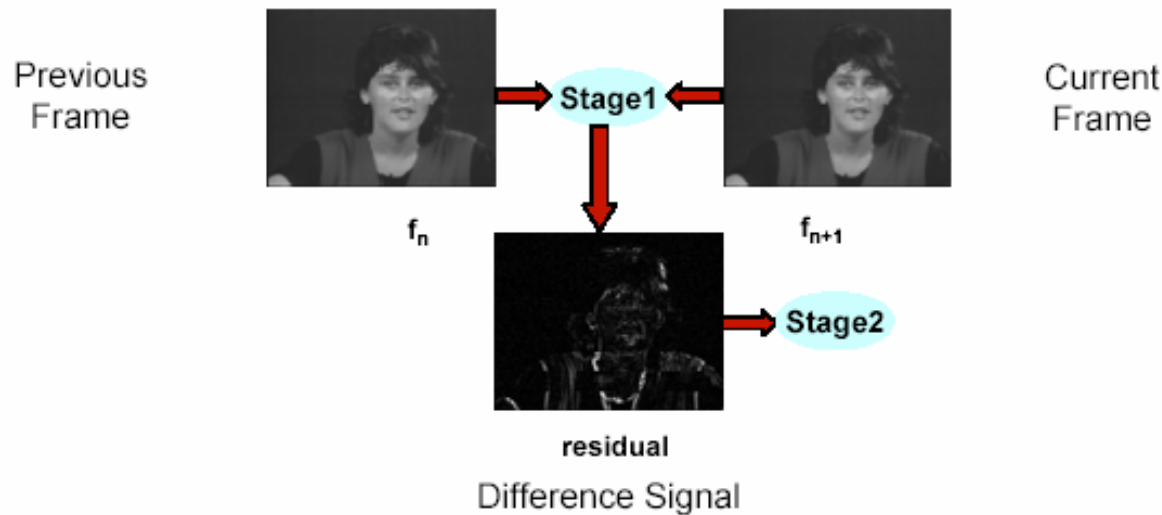
- The Sum of Absolute Errors (SAE) for each prediction indicates the magnitude of the prediction error



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



# Inter Frame coding



- Two types of inter-frames
  - P-frame predicted from previous frames
  - B-frame predicted from previous and future frames
- Intra frames are generated by Motion Estimation

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

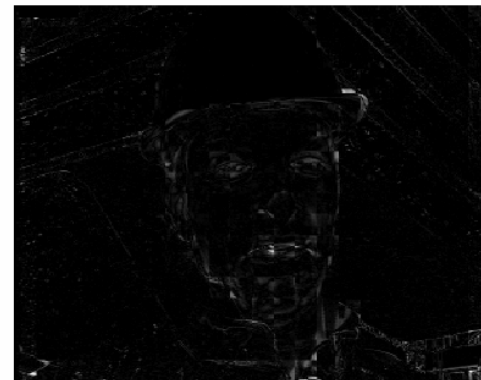
# Significance of Motion estimation (Tool - 2)



Absolute Difference w/o Motion Compensation



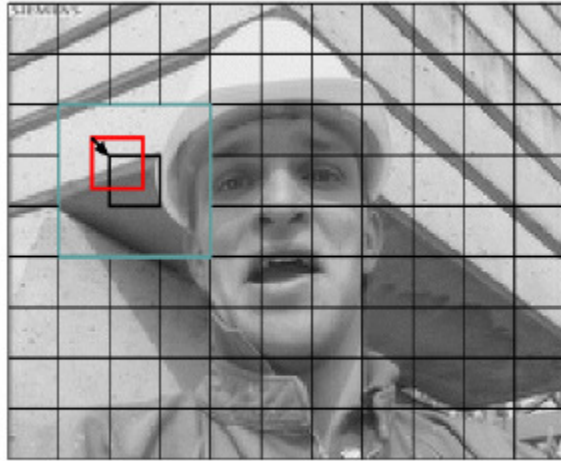
Absolute Difference with Motion Compensation



---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Motion Estimation

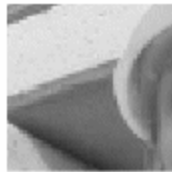


$g_{n-3}$



$g_n$

- Search starts from predefined search area
- From the search window motion vector will be predicted



Search area



Reference block  
from current image



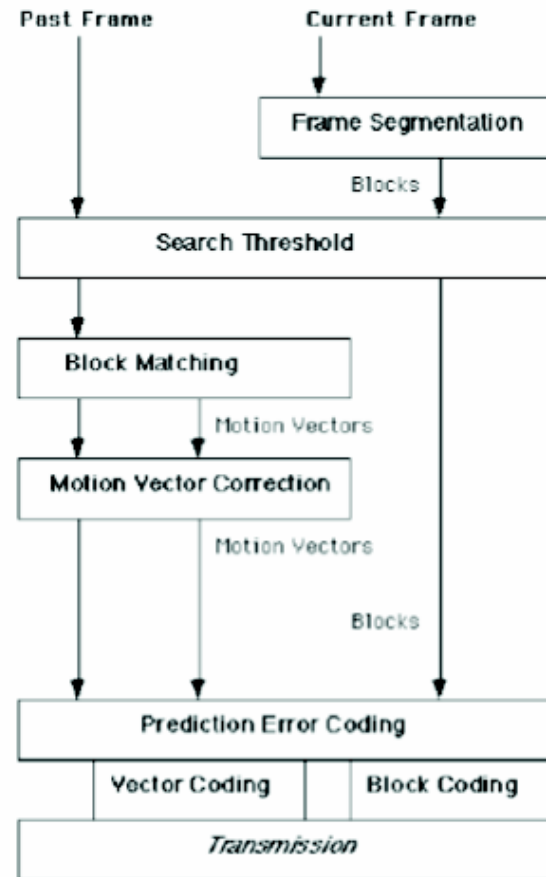
Most similar block  
from previous image

More Details Go to

[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Steps in Motion compensation

- Frame Segmentation
- Search Threshold
- Block Matching
- Motion Vector Correction
- Vector Coding
- Prediction Error Coding



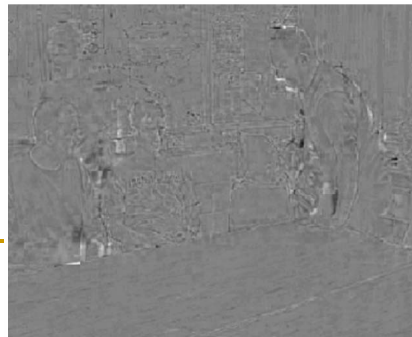
More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



■ Input frame



■ Reconstructed frame



■ Residuals



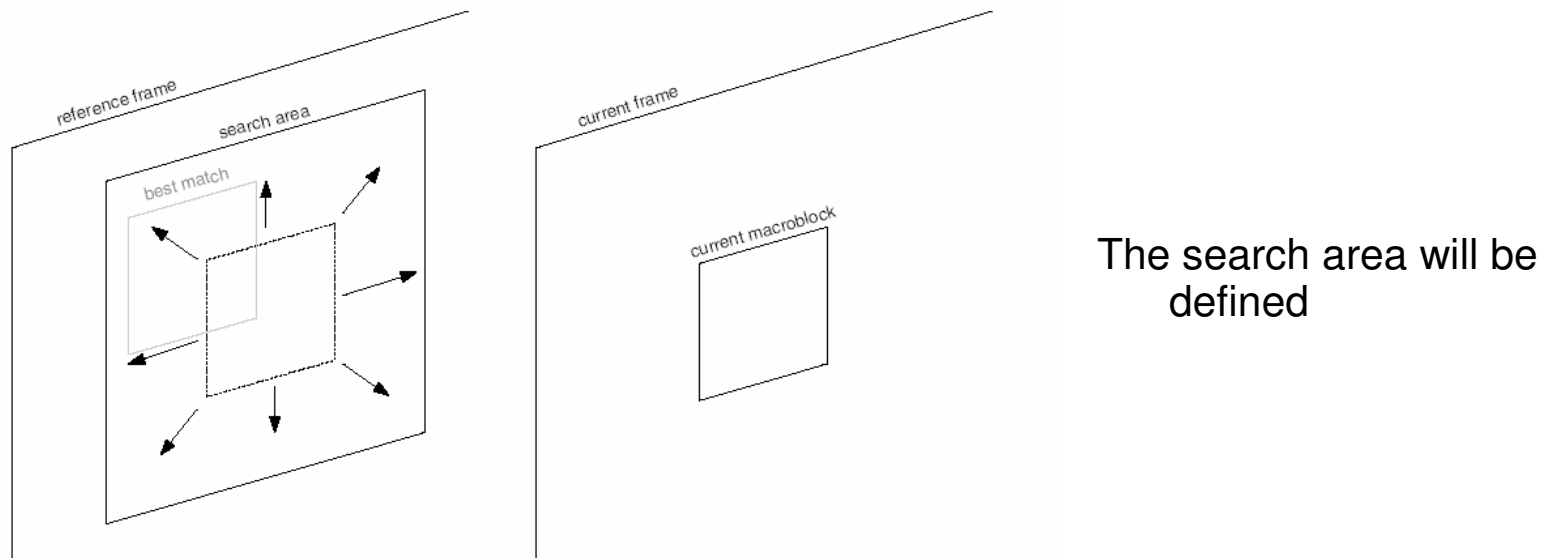
■ Motion vectors

More Details Go to

[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

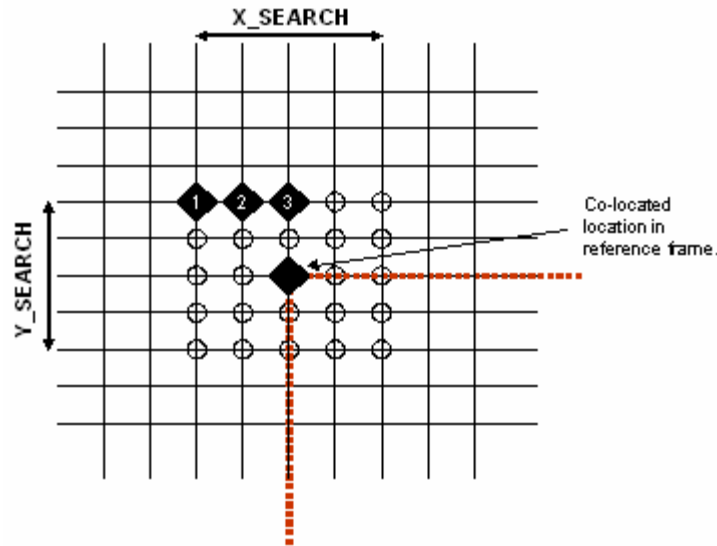
# Search Strategies (Tool - 3)

- full search
- 3 step search
- logarithmic search (assumption: convex error progression)
- Diamond search



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Full-Search Motion Estimation



- guarantees a minimal error, but computational expensive.

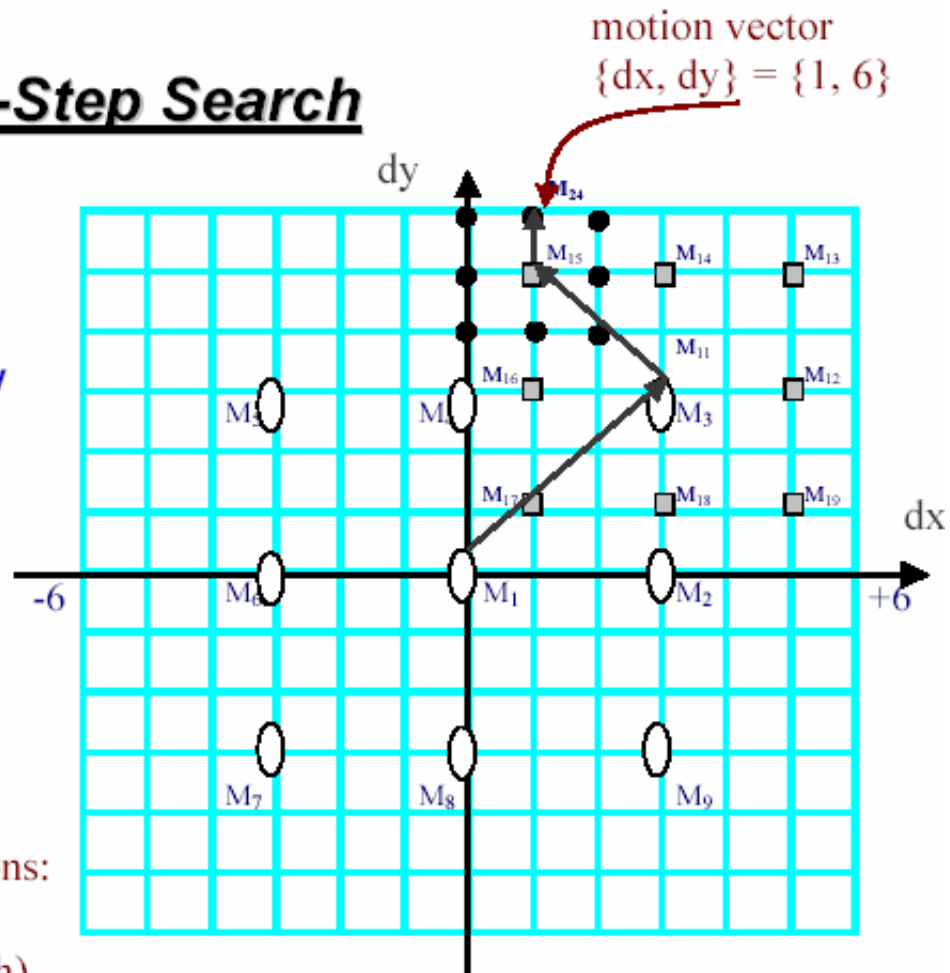
More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# 3-step search

## **Fast Algorithm: 3-Step Search**

- Search candidates at 8 neighbour positions
- Step-size cut down by 2 after each iteration
  - Start with step size approx. half of max. search range

Total number of computations:  
 $9 + 8 \times 2 = 25$  (3-step)  
 $(2R+1)^2 = 169$  (full search)



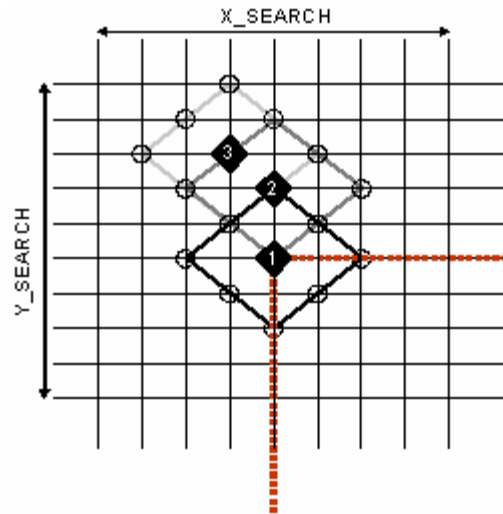
(From Ken Lam – HK Poly Univ.  
short course in summer 2001)

More Details Go to

[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



# Diamond Search

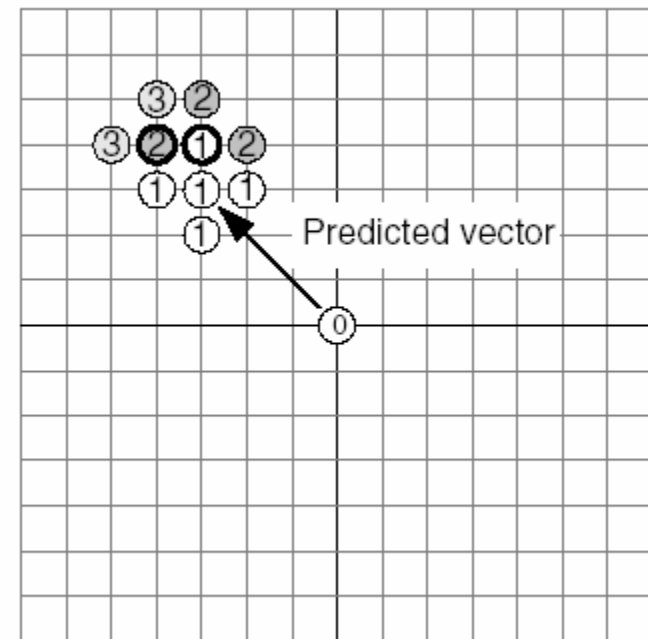
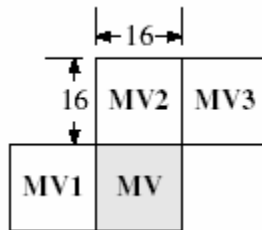


- The search stops once the position of the difference is located in the center of the diamond.

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Predictive Algorithms

- use neighboring blocks with already computed motion vectors to estimate the initial pattern of the search
- Used in H.264



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

## Block matching (Tool – 4)

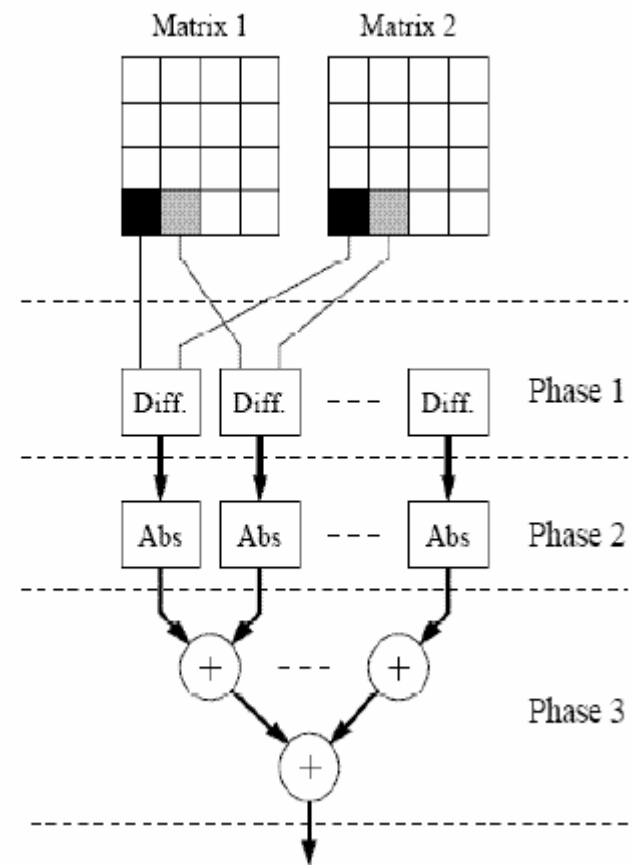
- SAD – Sum of Absolute differences
- SSE – Sum of Squared Error
- SATD – Sum of Absolute Transform Differences

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

## Sum of Absolute Differences (SAD)

$$Dist(u, v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \|a_{i,j} - b_{i,j}\|_u$$



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

## Sum of Absolute Transformed Differences (SATD)

$$SATD = \sum_{i=1}^4 \sum_{j=1}^4 |c_{i,j}|$$

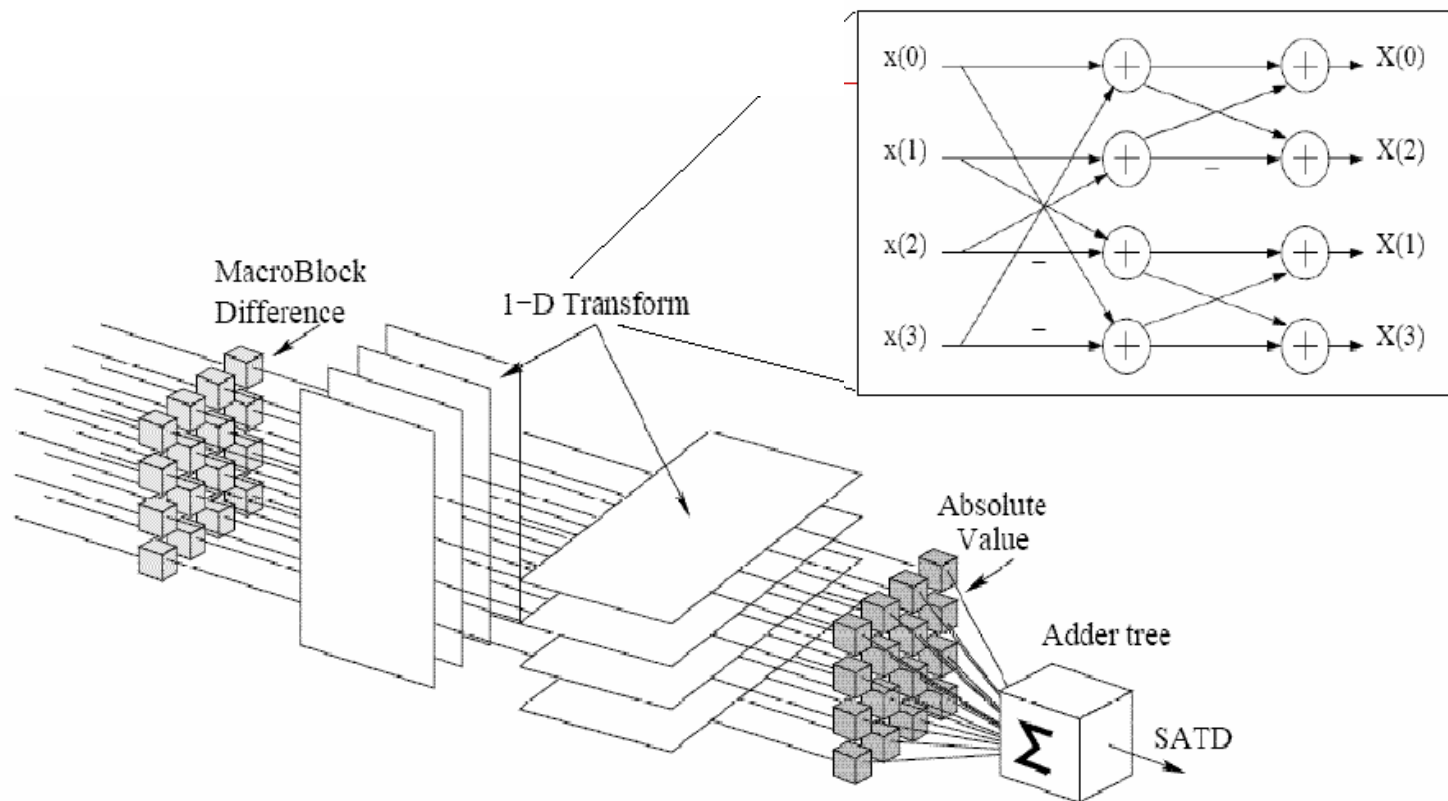
$$C = H_4 D H_4^T$$

Hadamard-Walsh transform

$$W = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

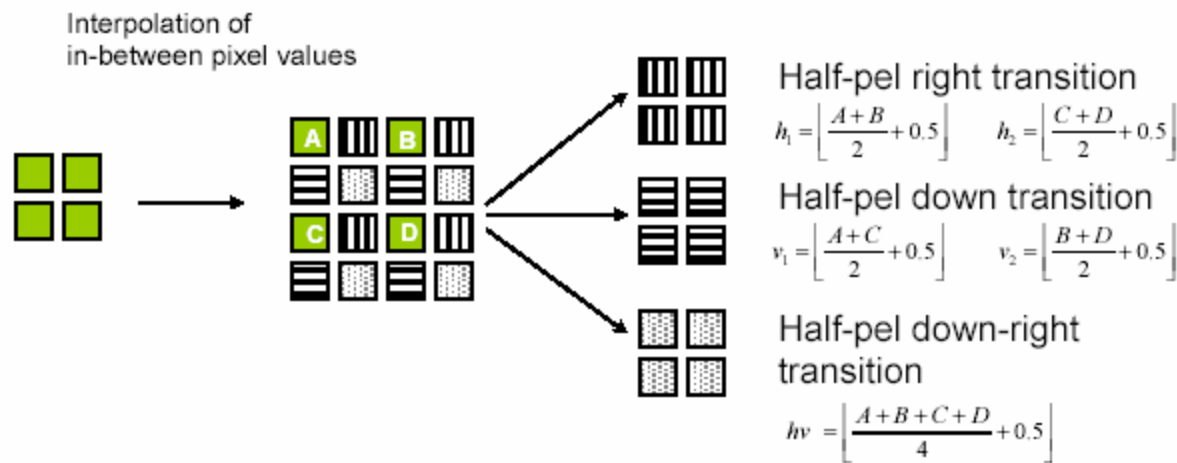
# SATD hardware



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Half Pixel Resolution

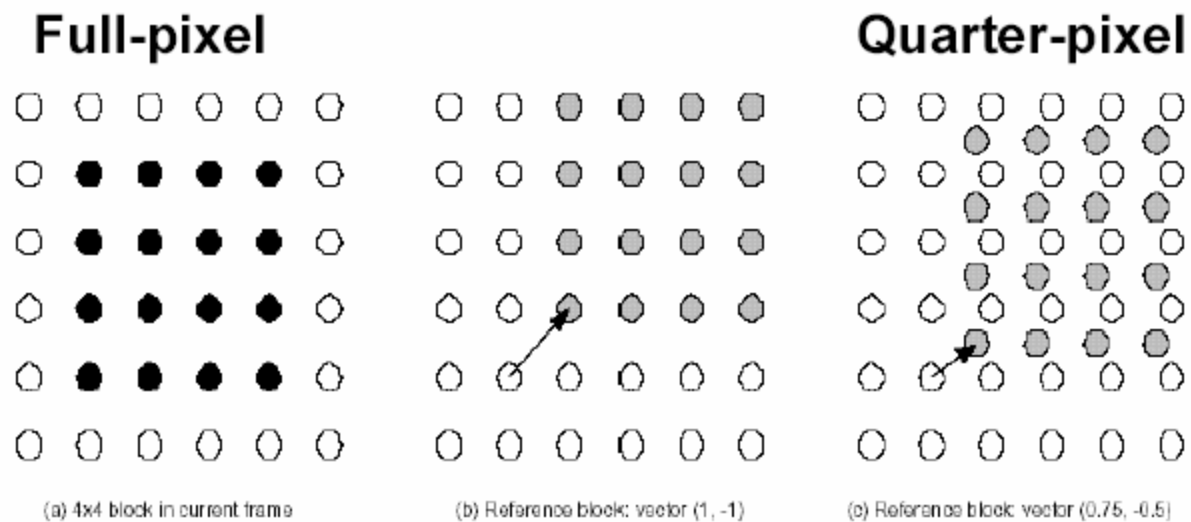
- Using motion vectors with half-pixel resolution allows for increased efficiency in motion compensation



- Minute motions are captured here

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

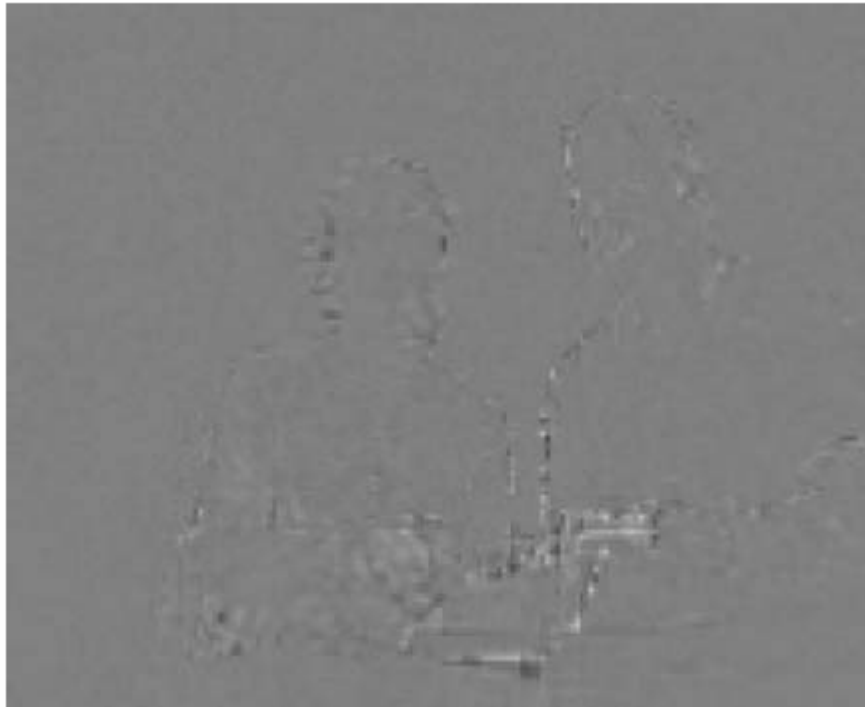
# Quarter - Pixel



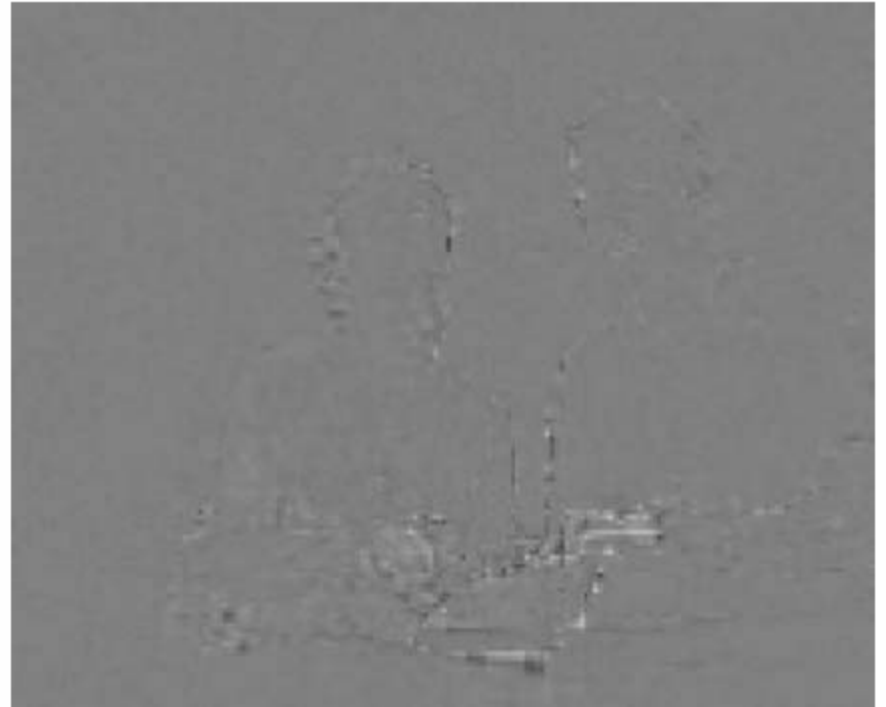
- Used in H.264

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)





■ Half pixel

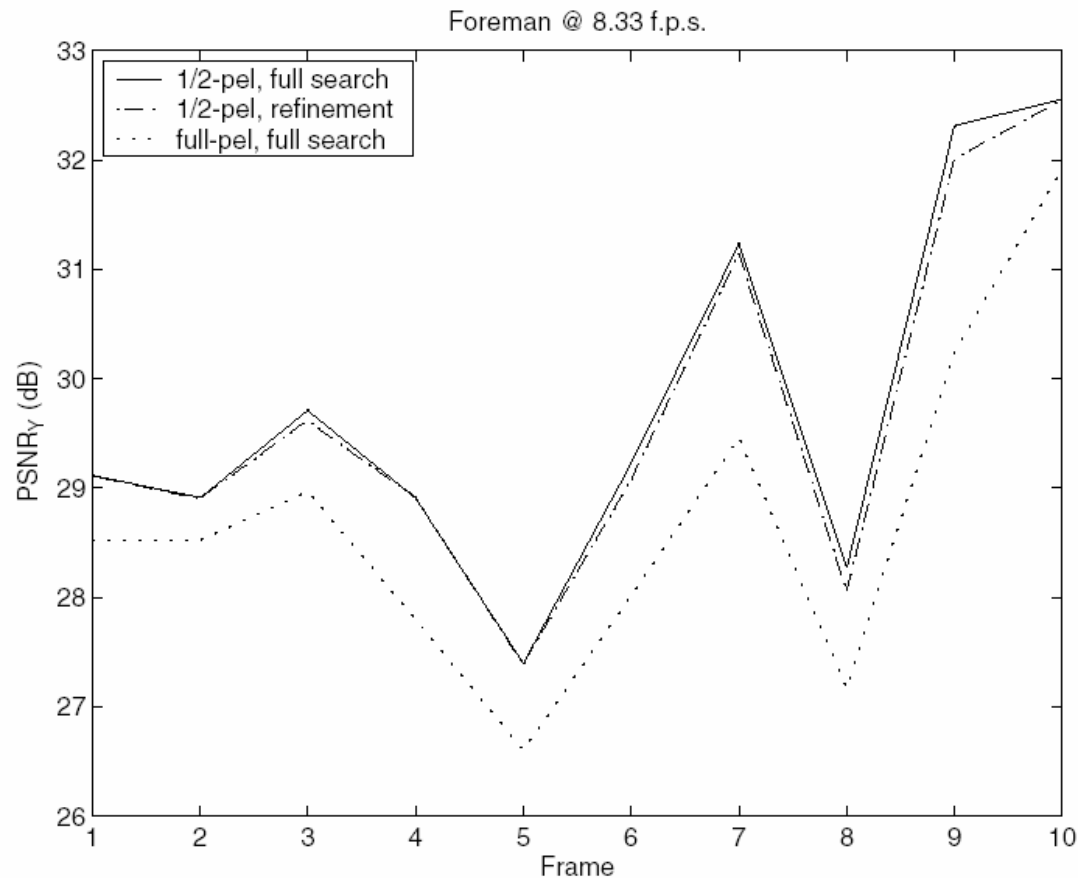


■ Quater pixel

---

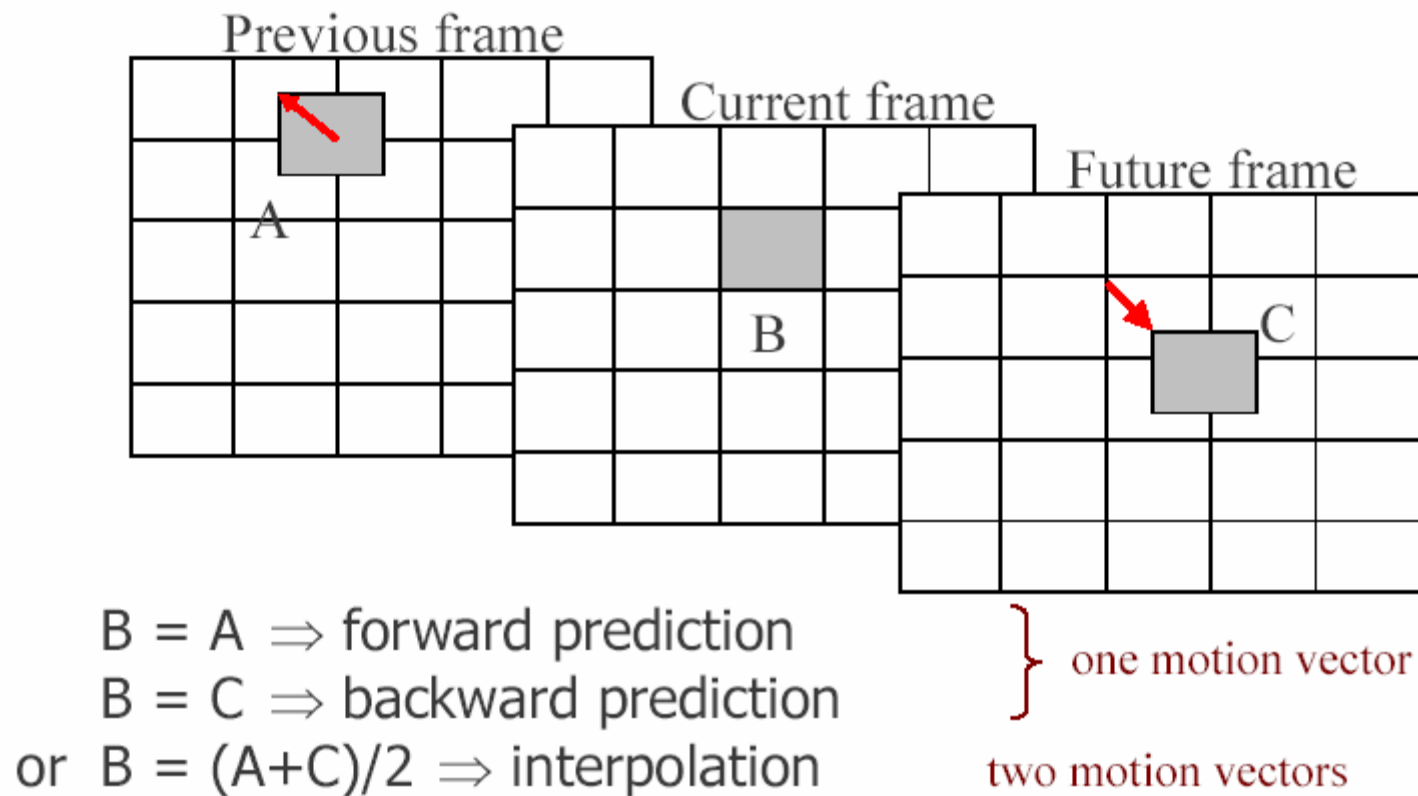
More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Full-pixel vs Half-pixel PSNR



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

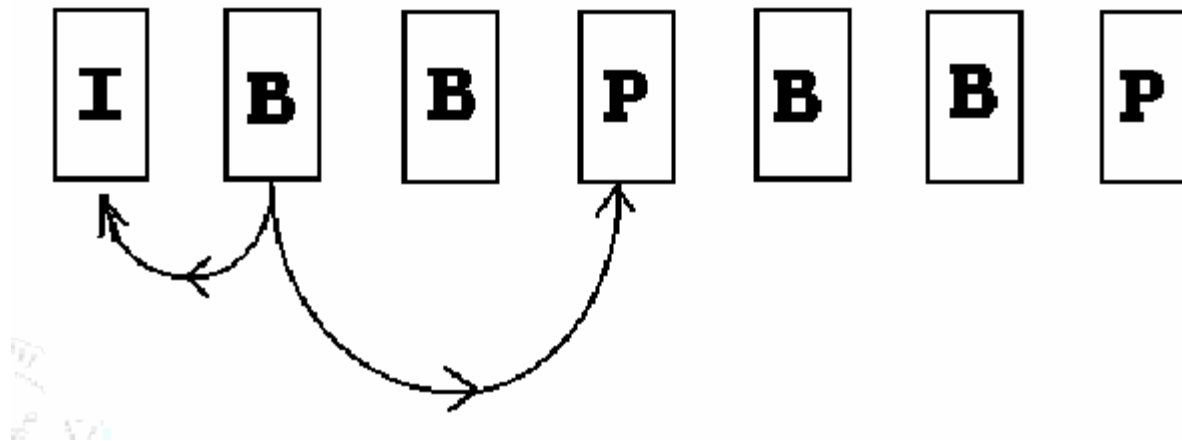
# Bi-Prediction Frame (Tool – 5)



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Frame coding example

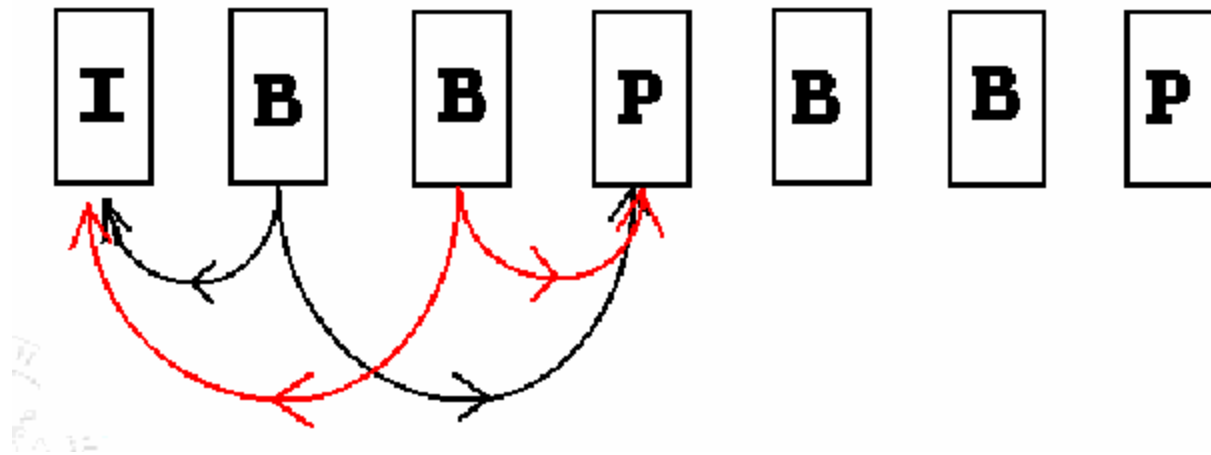
- The first B-frame needs **two references**: one backward (I-frame) and one forward (P-frame)



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Frame coding example

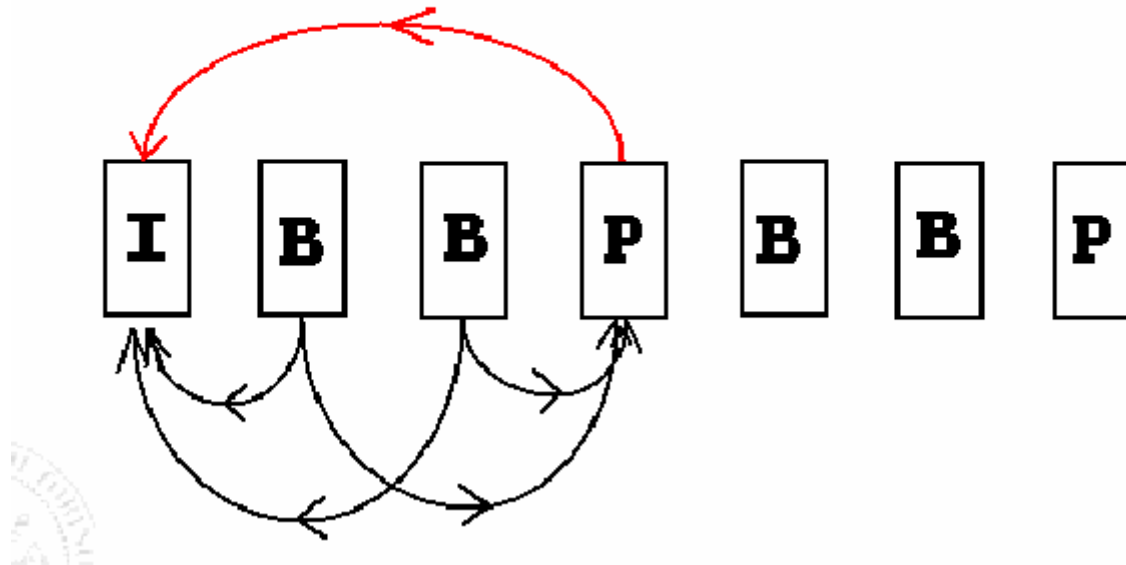
- The second B-frame cannot use the previous B-frame as a reference



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Frame coding example

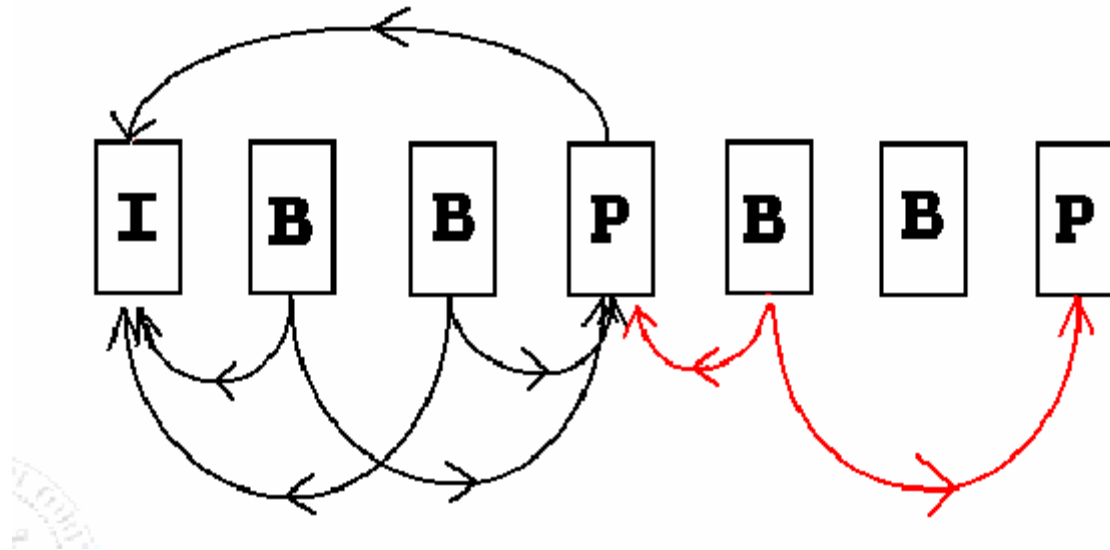
- P-frames are predicted from the previous I-frame



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Frame coding example

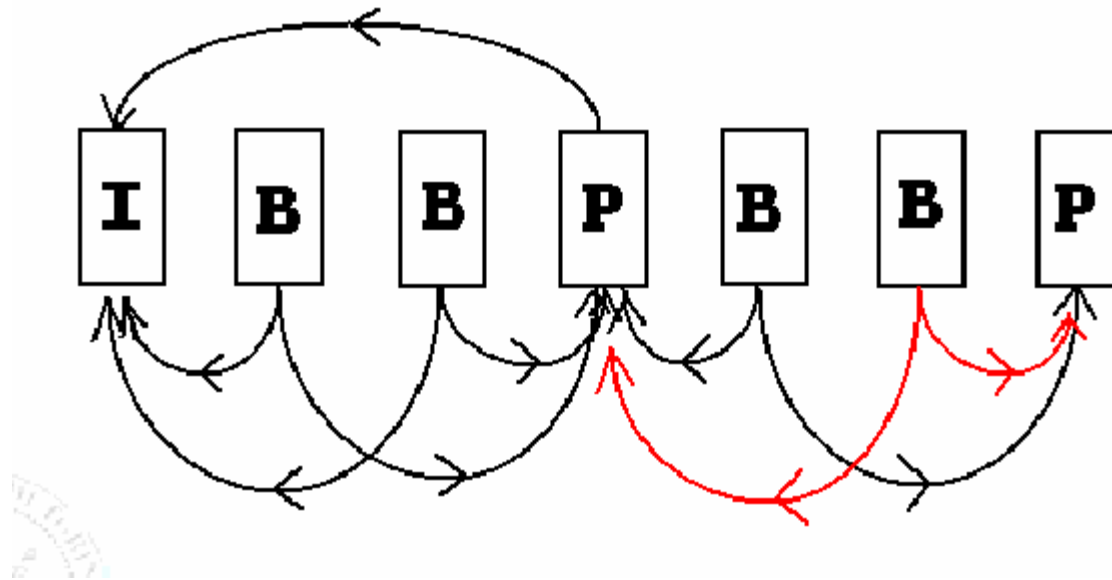
- The third B-frame refers to two P-frames



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Frame coding example

- The same applies to the fourth B-frame

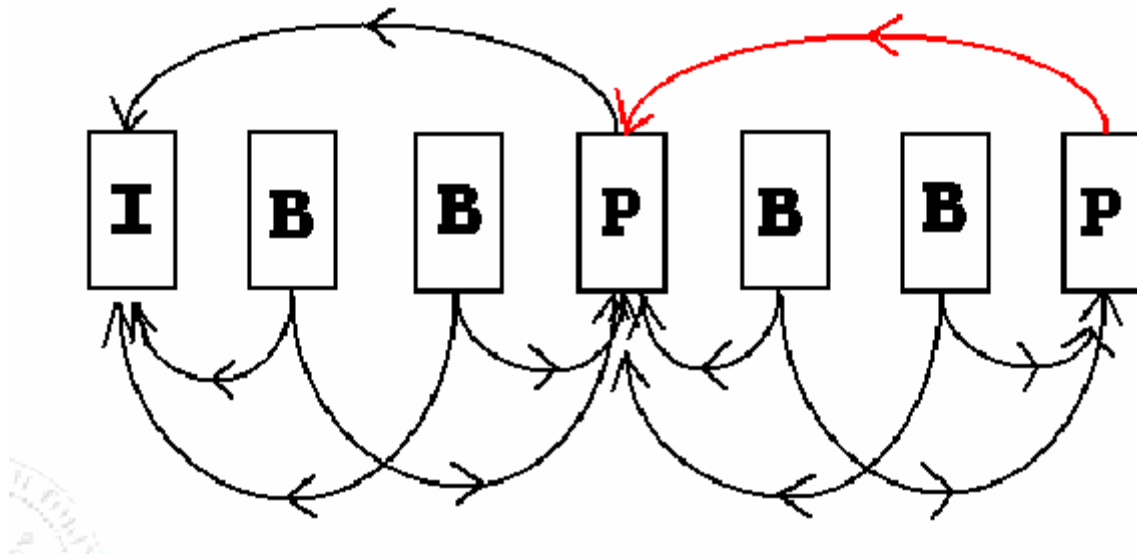


More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



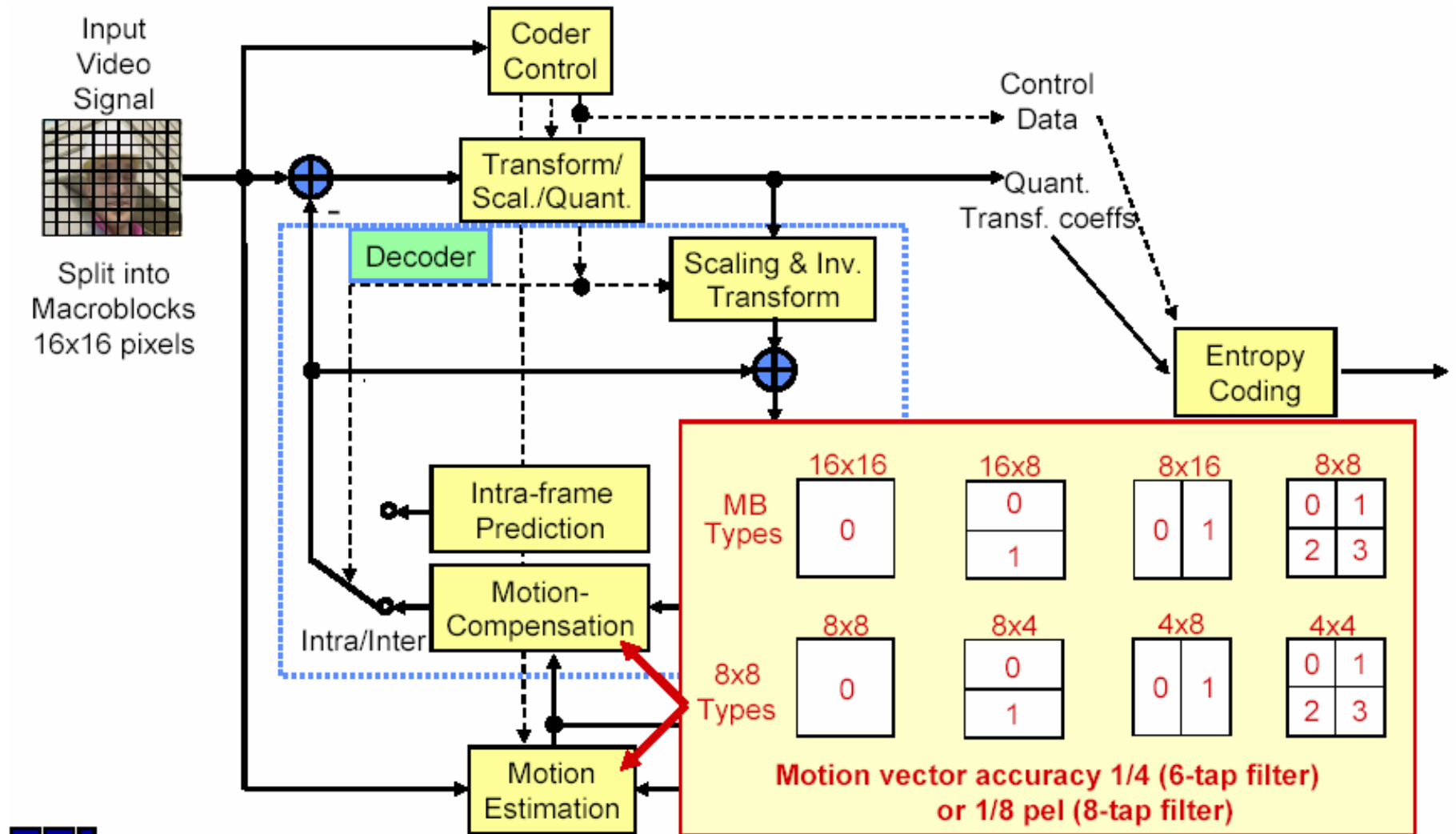
# Frame coding example

- Finally the second P-frame refers to the previous P-frame



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Motion compensation accuracy in H.264



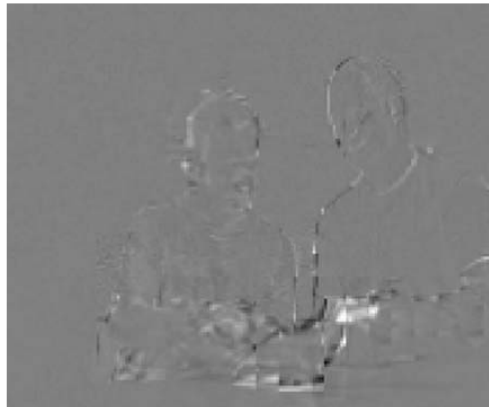
More Details Go to

[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

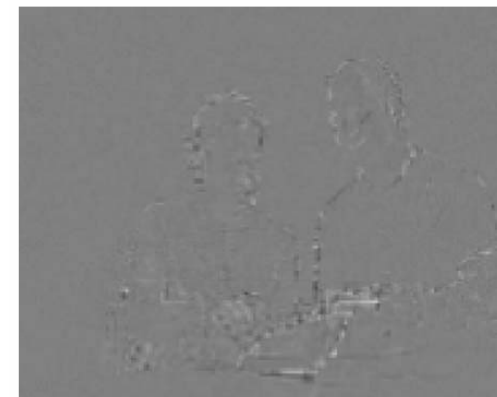
# Fixed block vs Variable block



16x16



8x8

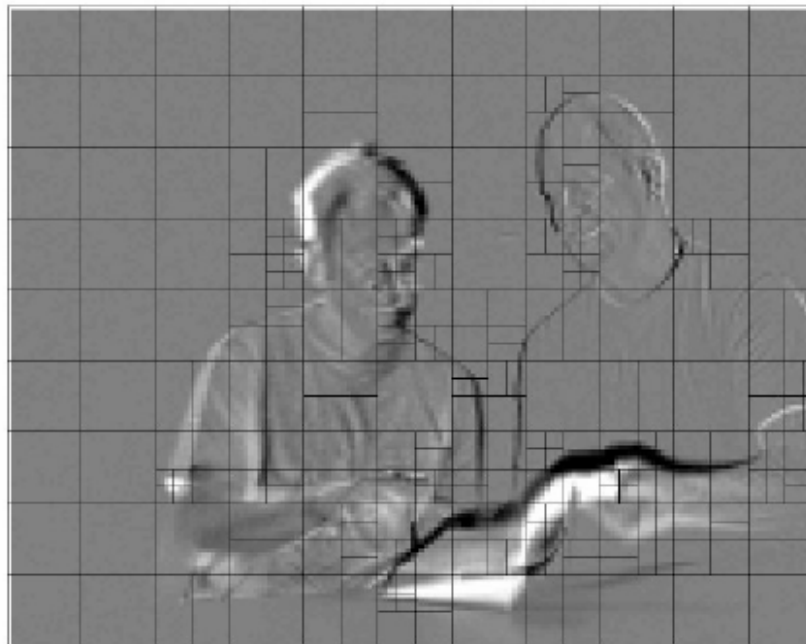


4x4

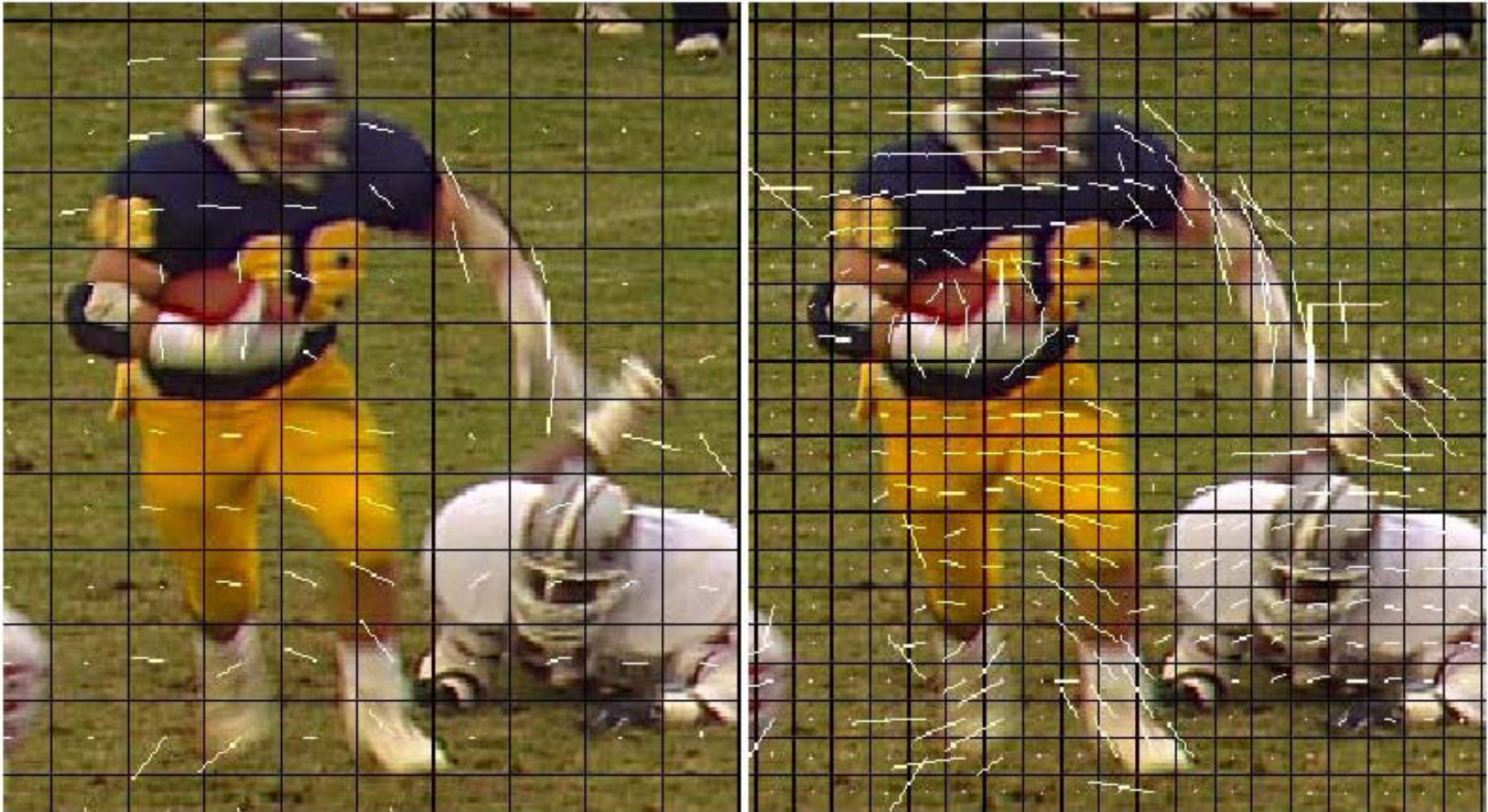
Residuals

**Note:**

Edges require smaller  
block sizes for more  
accurate prediction



# Motion vectors at different resolutions



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

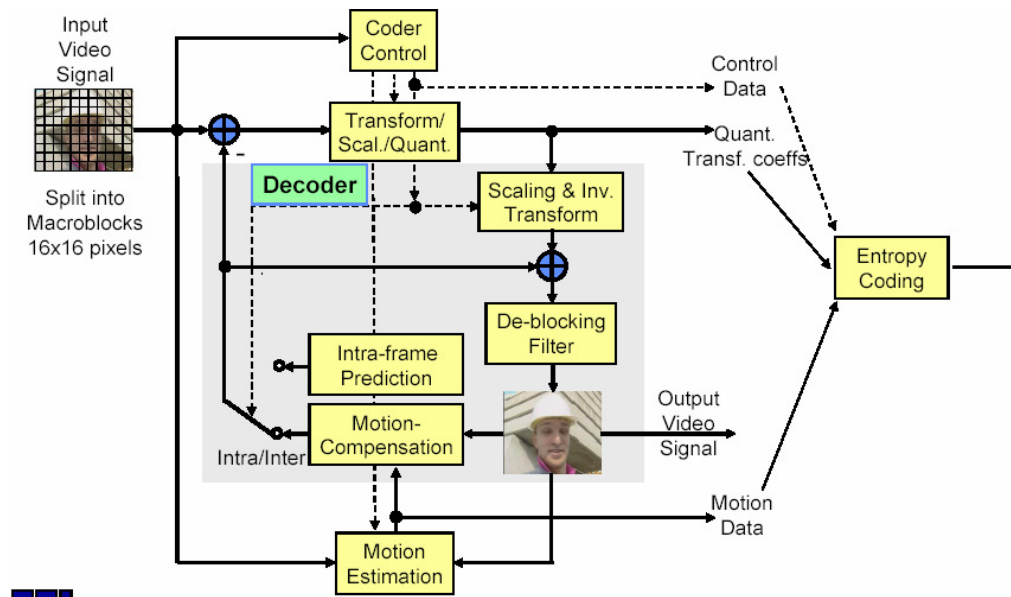
---

# Advanced: Motion Estimation and Compensation

- *Pixel-based* motion compensation
  - Large computation overhead
- *Block-based* motion compensation
  - Simple, hardware based implementation possible (most often used today)
- *Object-based* motion compensation
  - Object-detection complex and error-prone. Still in research

# The reference frame

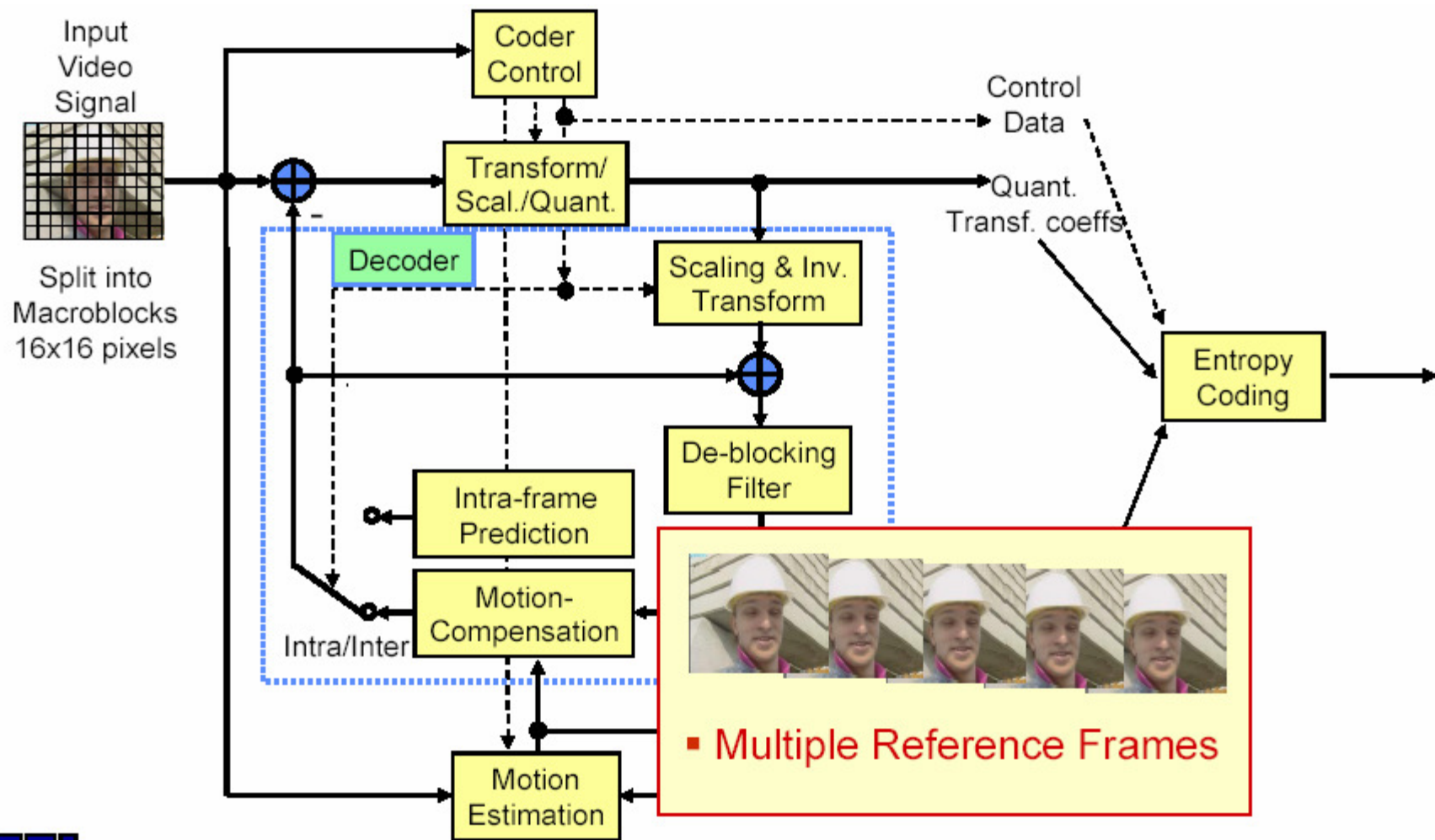
- The encoder have both the encoding loop as well as the decoding loop.
- In the encoder loop, the predictions are made out of the reconstructed images from decoder loop to reduce the error on the decoder side



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



# Reference frames in H.264

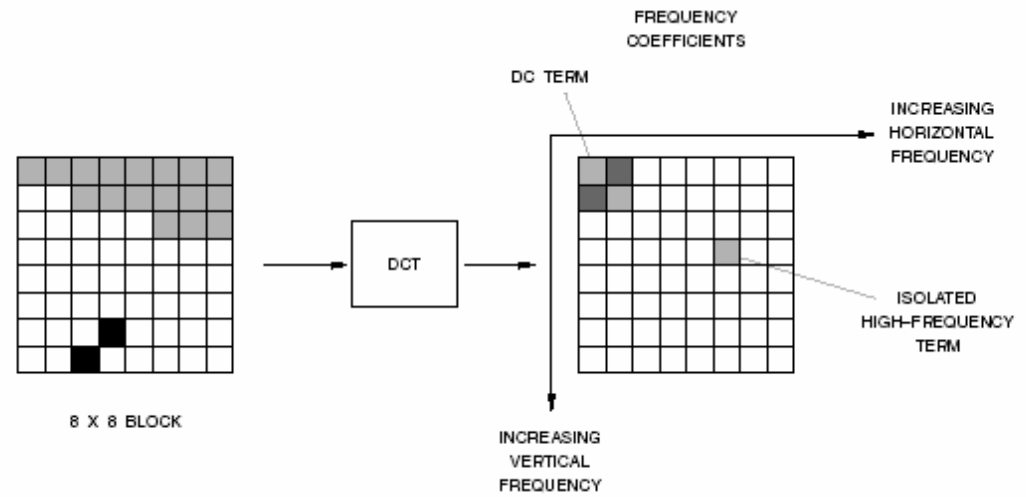


More Details Go to

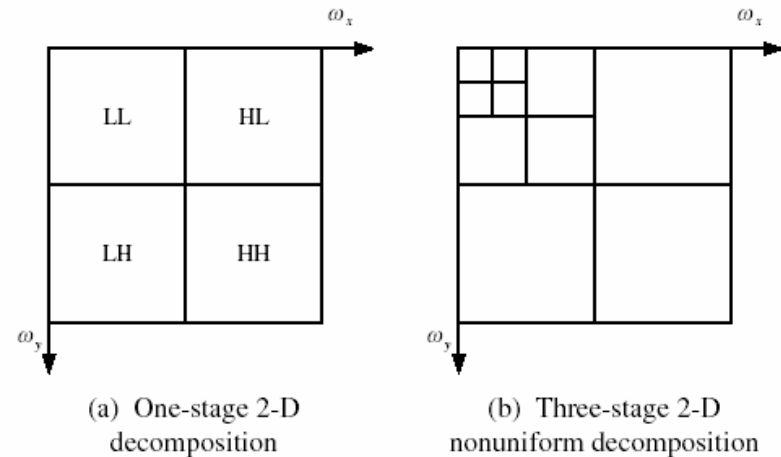
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Transform coding (Tool - 6)

- DCT is used to convert data in time domain to data in frequency domain.



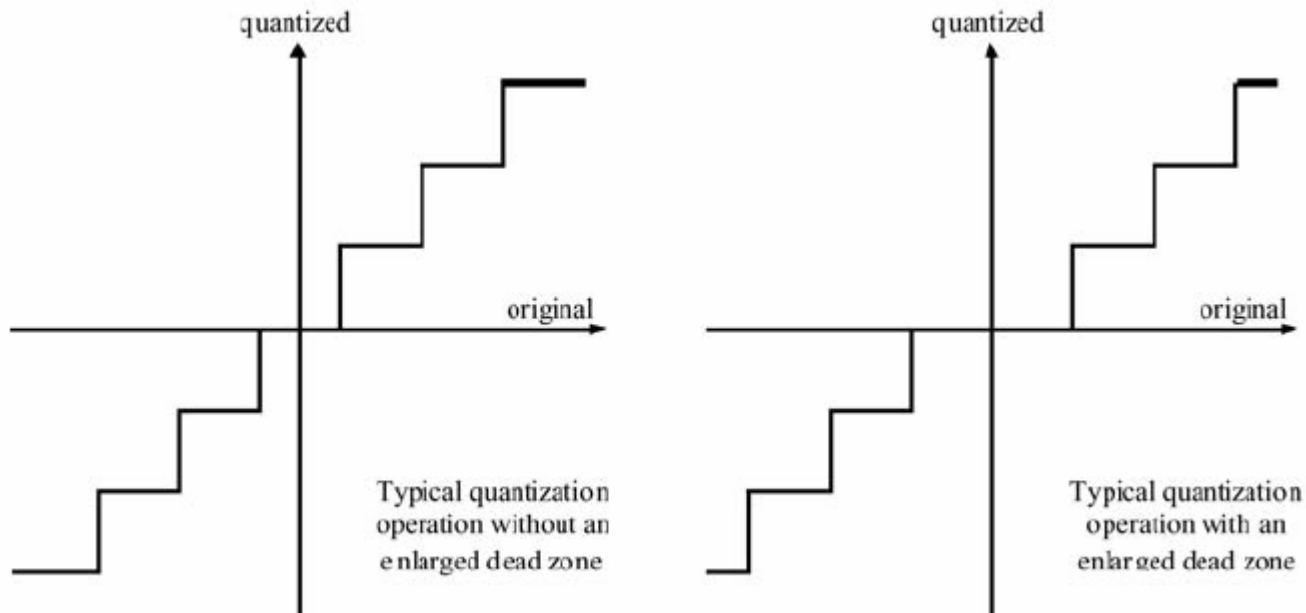
- Wavelets



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



# Quantization and dead zone (Tool - 7)



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Quantization for I-frame (I-block) & M.C. Residues

- Quantizer for I-frame (I-block)
  - Different step size for different freq. band (similar to JPEG)
  - Default quantization table
  - Scale the table for different compression-quality
- Quantizer for residues in predictive block
  - Noise-like residue
  - Similar variance in different freq. band
  - Assign same quantization step size for each freq. band

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

(a) quantization table for intra coding.

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

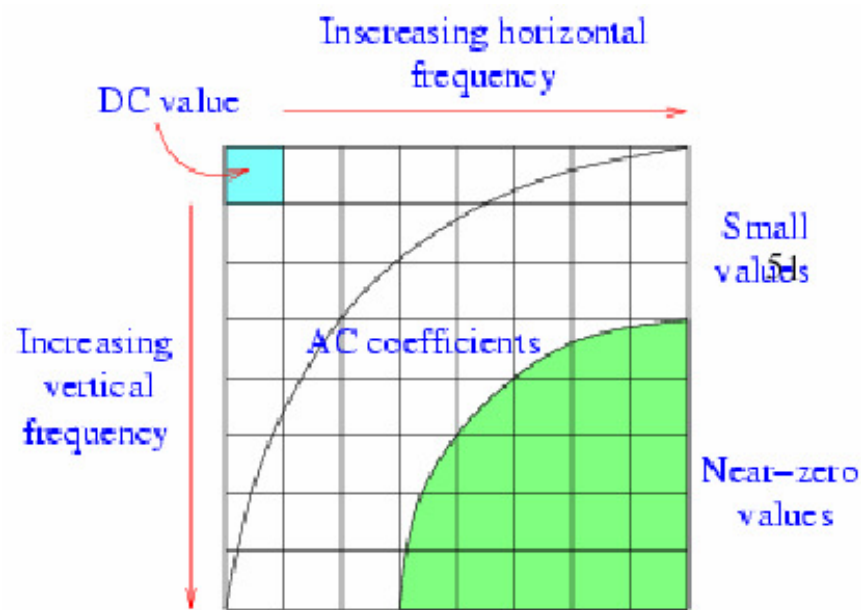
(b) quantization table for non-intra coding.



More Details Go to

[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

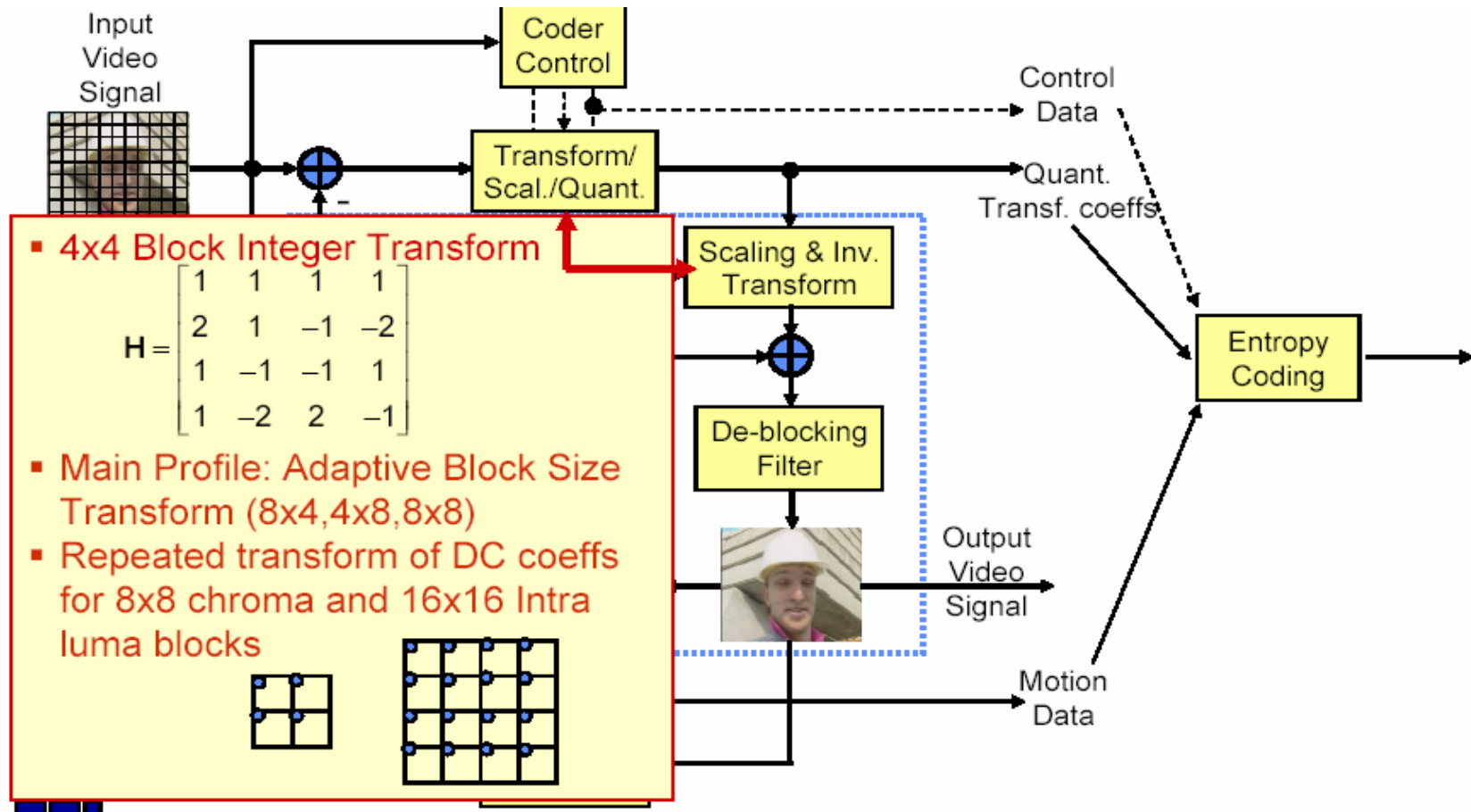
# The distribution of 2-D DCT Coefficients



68	3	5	-2	0	0	-2	0
-10	0	-4	3	0	0	0	0
9	3	0	0	0	-2	0	0
3	2	0	3	0	2	-2	0
0	0	2	-2	0	0	0	0
0	2	-2	2	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

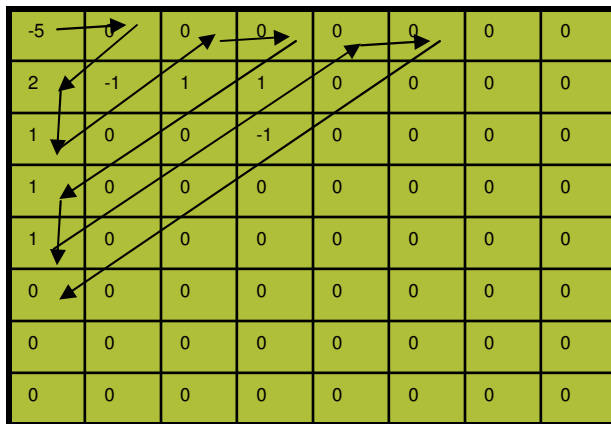
# Transform coding in H.264



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Scanning and Huffman Encoding

- Spatial Frequencies scanned in zig-zag pattern (note high frequencies mostly zero)

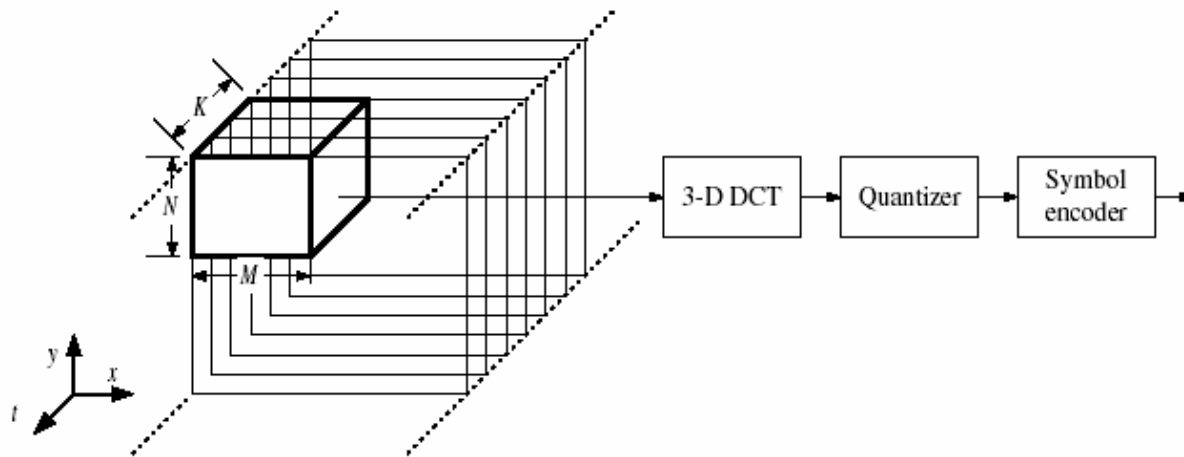


0,2,1,-1,0,0,1,0,1,1,0,0,1,0,0,0,-1,0,0,... 0

Can be stored as:

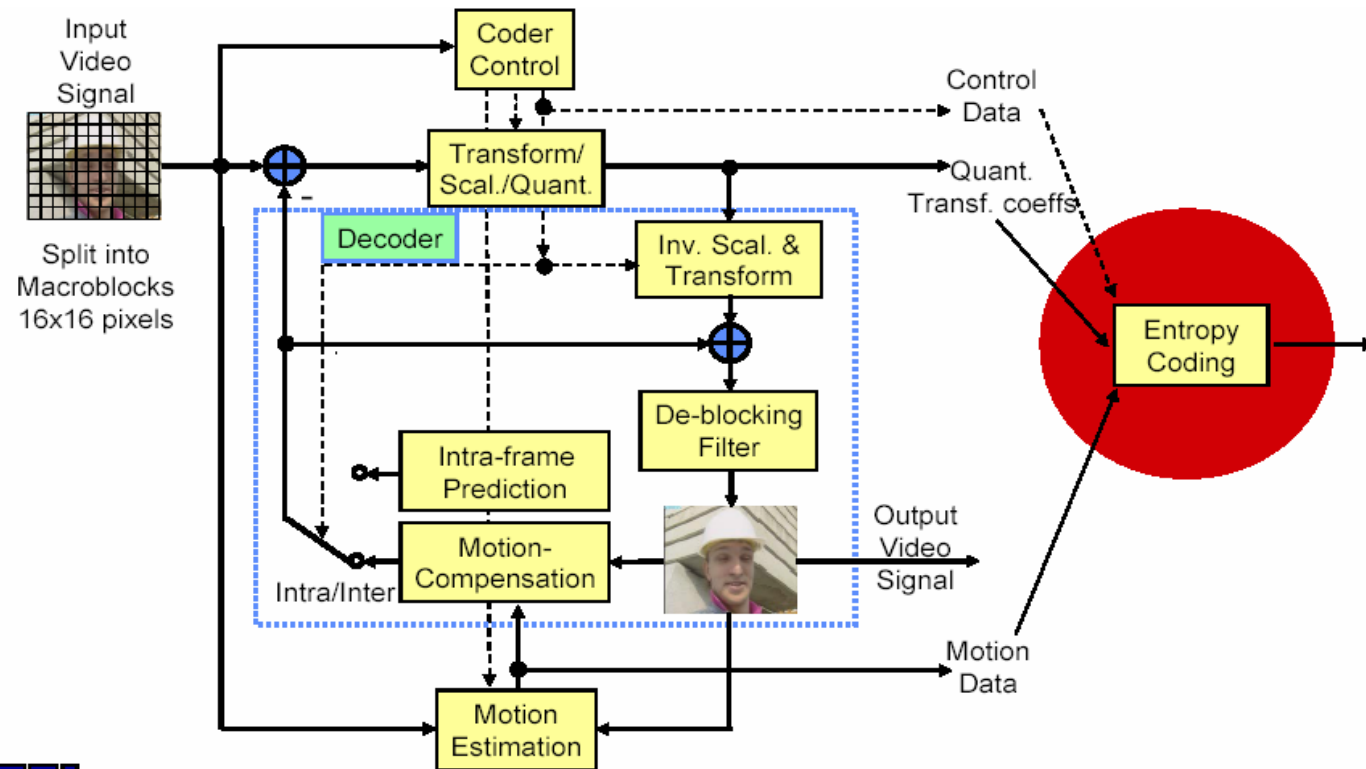
(1,2),(0,1),(0,-1),(2,1),(1,1),(0,1),(0,1),(2,1),(3,1),EOB

# 3D DCT is of research interest



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Entropy Coding (Tool – 8)



## ■ Lossless compression

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# CAVLC

- Exp-Golomb code is used universally for all symbols except for transform coefficients
- Context adaptive VLCs for coding of transform coefficients
  - No end-of-block, but number of coefficients is decoded
  - Coefficients are scanned backwards
  - Contexts are built dependent on transform coefficients

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



---

# CABAC

- Usage of **adaptive** probability models for most symbols
- Exploiting symbol correlations by using **contexts**
- Restriction to **binary arithmetic coding**
  - **Simple and fast adaptation** mechanism
  - Fast binary arithmetic codec based on table look-ups and shifts only

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# CAVLC in Detail

- Context-adaptive feature in the H.264 coding efficiency is much higher
- CAVLC algorithm is adaptive
- 5 syntax
  - The number of coefficients and trailing ones
  - The sign of each TrailingOnes
  - The levels of the remaining non-zero coefficients
  - The total number of zeros before the last coefficient.
  - Each run of zeros

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# CAVLC Example

4 × 4 block:

0	3	-1	0
0	-1	1	0
1	0	0	0
0	0	0	0

Reordered block:

0,3,0,1,-1,-1,0,1,0,...

- TotalCoeffs = 5 (indexed from highest frequency, 4, to lowest frequency, 0)
- total zeros = 3
- TrailingOnes = 3 (in fact there are four trailing ones but only three can be encoded as a 'special case')

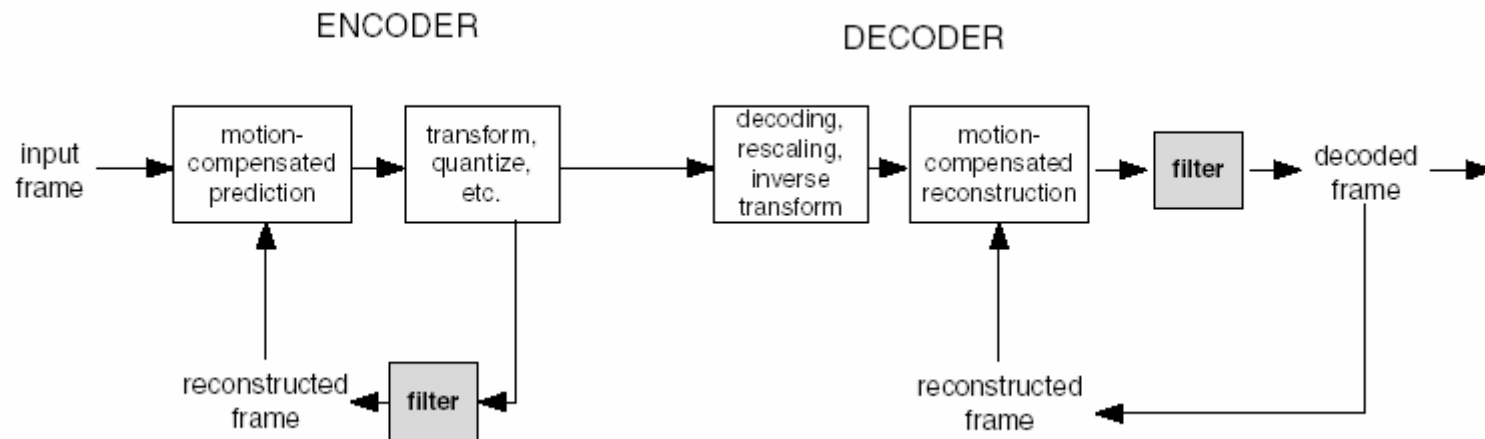
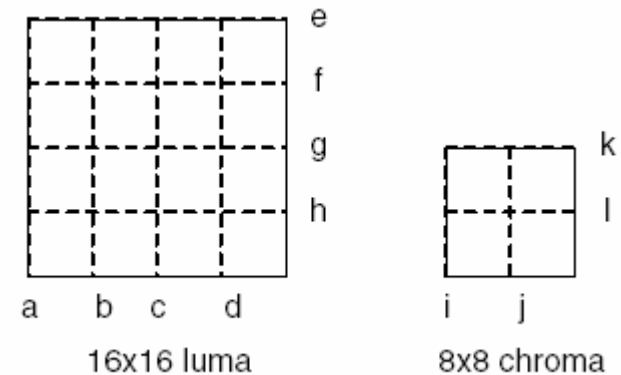
Element	Value	Code
coeff_token	TotalCoeffs = 5, TrailingOnes = 3 (use Table 1)	0000100
TrailingOne sign (4)	+	0
TrailingOne sign (3)	-	1
TrailingOne sign (2)	-	1
Level (1)	+1 (use suffixLength = 0)	1 (prefix)
Level (0)	+3 (use suffixLength = 1)	001 (prefix) 0 (suffix)
total zeros	3	111
run_before(4)	ZerosLeft = 3; run_before = 1	10
run_before(3)	ZerosLeft = 2; run_before = 0	1
run_before(2)	ZerosLeft = 2; run_before = 0	1
run_before(1)	ZerosLeft = 2; run_before = 1	01
run_before(0)	ZerosLeft = 1; run_before = 1	No code required; last coefficient.

The transmitted bitstream for this block is 000010001110010111101101.

# Deblocking filter in H.264 (new)

## (Tool – 9)

- Improves subjective visual *and* objective quality of the decoded picture. Is significantly superior to post filtering.
- Filtering affects the edges of the 4x4 block structure
- Highly content adaptive filtering procedure mainly removes blocking artifacts and does not unnecessarily blur the visual content



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# Deblocking: Subjective result for Intra

Highly compressed first decoded intra picture  
at a data rate of 0.28 bit/sample



1) Without Filter

2) with H264/AVC Deblocking

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# Deblocking: Subjective result for Inter

Highly compressed decoded inter picture



1) Without Filter

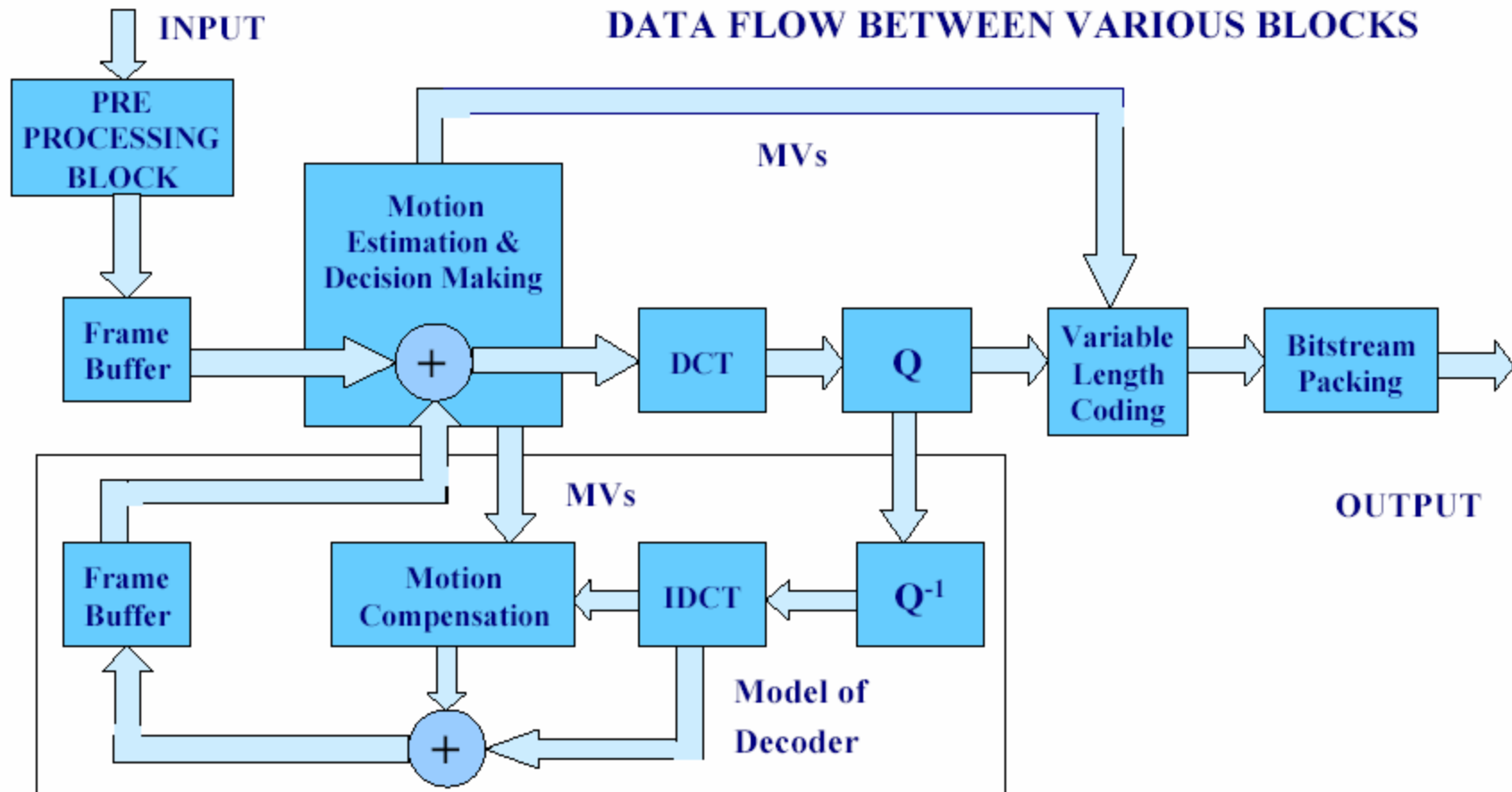


2) with H264/AVC Deblocking

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Putting blocks together



More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# Approaches For Implementing Video

- SW Implementation
  - Internet and PC application with video CODEC SW running on PC's processor
  - Portable application price and power consumption of processor a limitation
  - Currently works at 6fps for CIF format, with a target of 15 fps.
- HW and SW Implementation
  - Execute compute intensive calculations in HW and less demanding in SW
  - Decrease need for calculation power of processor
  - Motion estimation is typically heaviest calculation required, so if implemented in HW causes a high bus traffic from HW to SW
- Full HW Implementation
  - Compute intensive calculations in HW and only control functions to be done in SW
  - Calculations and interim results held internally Processor load very light and is needed only to handle control software
  - Minimal power consumption, minimal use to processor and minimal bus traffic
- Research is required for SW/HW partitioning

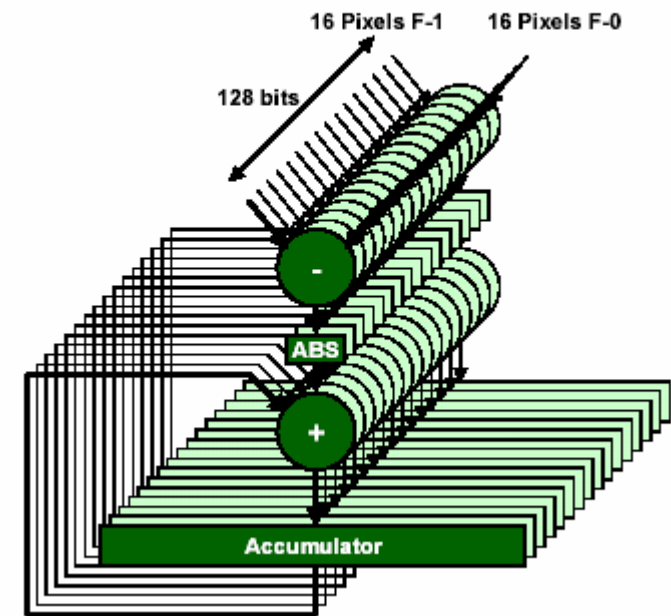
---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)



- Sequential processors
  - ❑ Desktop processors
  - ❑ ARM series
  - ❑ DM series from TI
  - ❑ VLIW processors
- Parallel processors (the new wave)
  - ❑ FPGA
  - ❑ Xtensa
  - ❑ Pact

Figure 8: Adding a SIMD SAD computational engine reduces the computational load by 46x



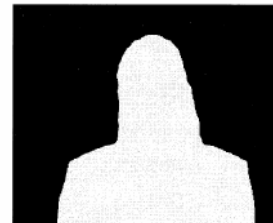
More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

# Future Trends

- The future trends for CODECs will be highly based on the implementation possibility.
- Sprite movement is very effective for compression but too computational intensive. So MPEG4 is futuristic.
- 3D video compression (future)
- Video transcoding in Compressed domain



Original



Binary Segmentation Mask



Background Layer VOP



Foreground Layer VOP

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# Parallel algorithms

- The high end researches are sponsored by processors people. So the algorithms are also mostly sequential.
- In the emergence of parallel computing units, we have to come up with parallel algorithms.
- More researches should come in parallel architectures.

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)

---

# THANK YOU

■ Questions?

---

More Details Go to  
[www.drtonygeorge.com/Video\\_Codec.htm](http://www.drtonygeorge.com/Video_Codec.htm)