Multimedia communications
ECP 610

Omar A. Nasr

omaranasr@ieee.org

March, 2015

JPEG

- A Lossy image compression technique
- "JPEG" stands for Joint Photographic Experts Group,
- The original JPEG group was organized in 1986,[7] issuing the first JPEG standard in 1992, which was approved in September 1992 as **ITU-T Recommendation T.81**[8] and in 1994 as **ISO/IEC 10918-1**.
- JPEG standard specifies the codec, which defines how an image is compressed into a stream of bytes and decompressed back into an image, but not the file format used to contain that stream
- JPEG 2000 …. Rarely used!

JPEG

- From last lecture:
  - Huffman coding
  - Discrete Cosine Transform (DCT)
- Run Length (RL) Coding
  - Lossless coding scheme

$$000000001110000001100000111100011000$$

$$8(0)3(1)6(0)2(1)5(0)4(1)3(0)2(1)3(0)$$

Can be extended for *one symbol*

$$0, 0, -3, 5, 1, 0, -2, 0, 0, 0, 0, 2, -4, 3, -2, 0, 0, 0, 1, 0, 0, -2$$
$$\Rightarrow \#2, -3, 5, 1, \#1, -2, \#4, 2, -4, 3, -2, \#3, 1, \#2, -2$$
$$\Rightarrow (2, -3)(0, 5)(0, 1)(1, -2)(4, 2)(0, -4)(0, 3)(0, -2)(3, 1)(2, -2)$$

Peak Signal to Noise Ratio

- Measure of quality

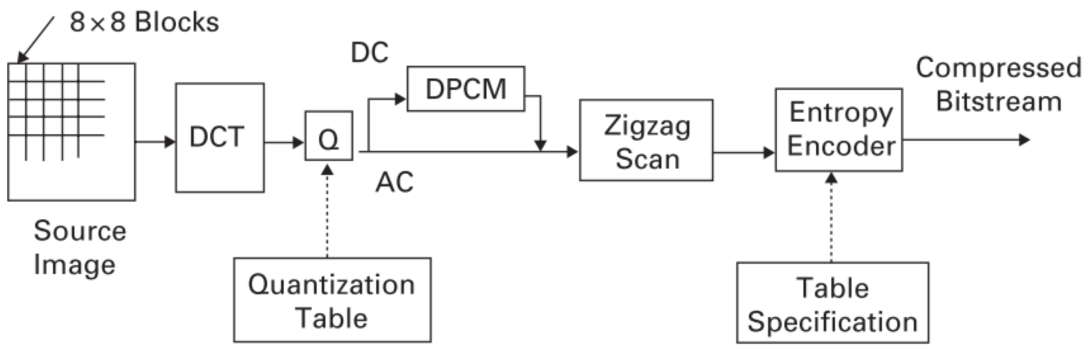
$$PSNR = 10\ \log_{10}\left(\frac{Max_I^2}{MSE}\right)$$

$$MSE = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left\| f(x,y) - \hat{f}(x,y)\right\|^2$$

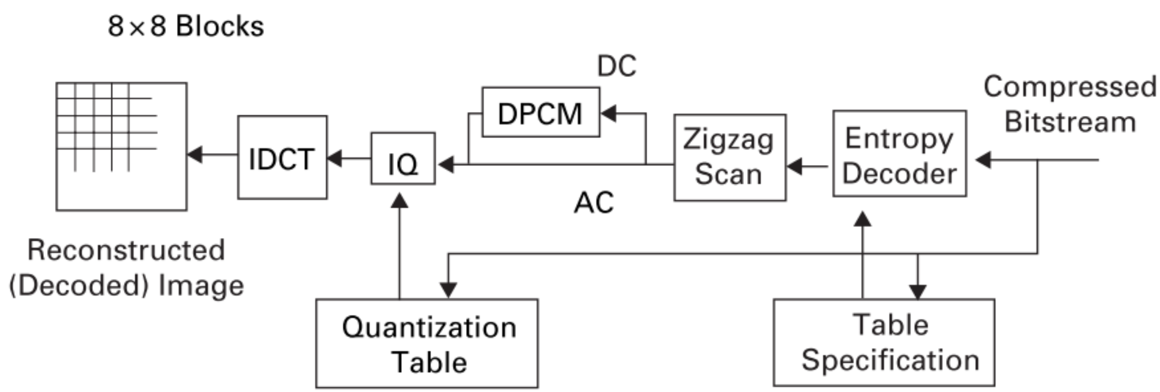
- For RGB images, same metric, with average over three components for the *MSE*

Block diagram

- Encoder



- Decoder

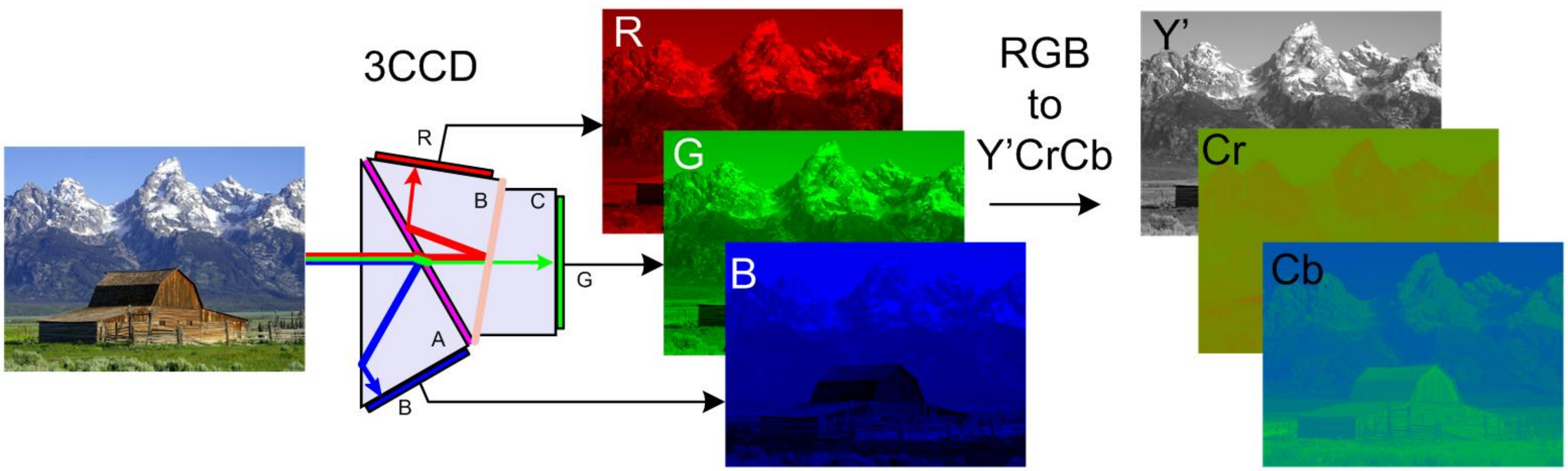


Chroma subsampling

- Human vision system:
  - More sensitive to Luma component than Chroma component

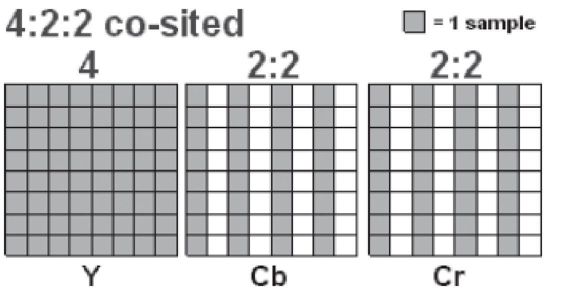
$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

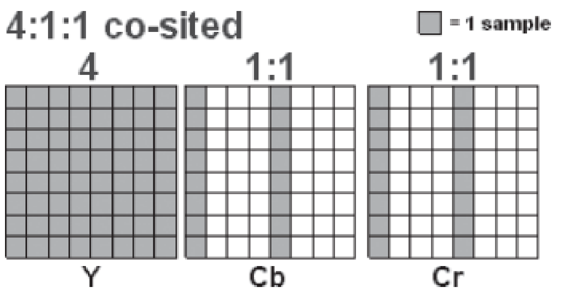
$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0 & 0.0 & 1.4021 \\ 1.0 & -0.3441 & -0.7142 \\ 1.0 & 1.7718 & 0.0 \end{bmatrix} \begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix}$$



Chroma

- Subsampling



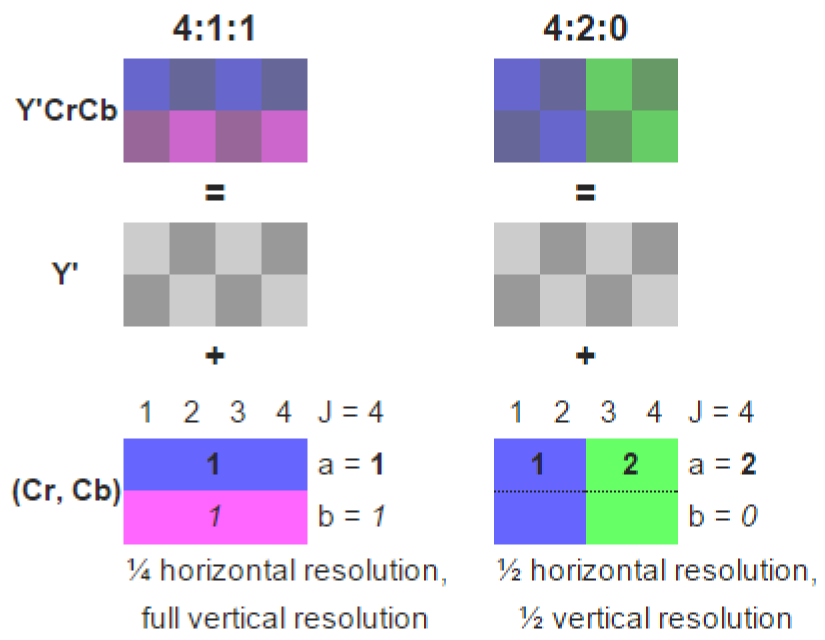


- *J*: horizontal sampling reference

(width of the conceptual region).

- *a*: number of chrominance samples (Cr, Cb) in the first row of *J* pixels.
- *b*: number of changes of chrominance samples (Cr, Cb) between first and second row of *J* pixels.

Chroma subsampling



Block based DCT

- For 8x8 blocks

$$F(u,v) = \frac{1}{4} C(u)C(v) \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right],$$

$$f(x,y) = \frac{1}{4} \sum_{u=0}^{7} \sum_{v=0}^{7} C(u)C(v)F(u,v) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right],$$

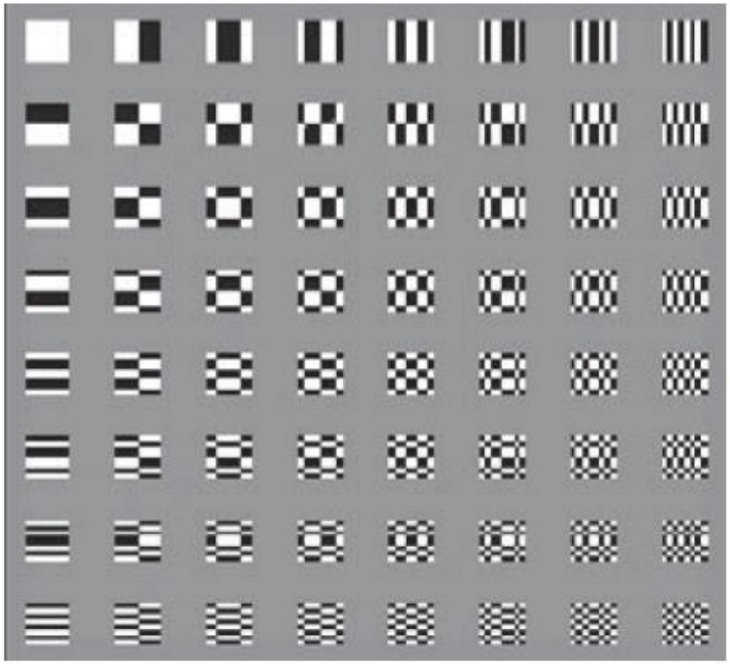
where

$$C(w) = \begin{cases} \frac{1}{\sqrt{2}}, & w = 0, \\ 1 & \text{otherwise.} \end{cases}$$

- Is this lossless or lossy?

Block based DCT

- DCT Basis functions



Quantization

- The DCT coefficients for Luma and Chroma are quantized
- Again, human visual system
  - More sensitive to low frequencies than high frequencies
    - Therefore: tolerate more quantization noise for …..
  - What about Luma and Chroma components?
    - ……

Quantization

- The DCT coefficients for Luma and Chroma are quantized
- Again, human visual system
  - More sensitive to low frequencies than high frequencies
    - Therefore: tolerate more quantization noise for higher frequencies
    - Use "larger" quantization step
  - What about Luma and Chroma components?
    - Eye is more sensitive to Luma
    - Use "larger" quantization step for Chroma component at the same frequency

Quantization tables

Table 4.1 (a) The default quantization table for luminance (Y component); (b) the default quantization table for chrominance (C_bC_r components)

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(a)

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

(b)

Quantization of DCT coefficients

$$F^{Q}(u,v) = IntegerRound\left(\frac{F(u,v)}{Q(u,v)}\right).$$

235.6	−1.0	−12.1	−5.2	2.1	−1.7	−2.7	1.3
−22.6	−17.5	−6.2	−3.2	−2.9	−0.1	0.4	−1.2
−10.9	−9.3	−1.6	1.5	0.2	−0.9	−0.6	−0.1
−7.1	−1.9	0.2	1.5	0.9	−0.1	0.0	0.3
−0.6	−0.8	1.5	1.6	−0.1	−0.7	0.6	1.3
1.8	−0.2	1.6	−0.3	−0.8	1.5	1.0	−1.0
−1.3	−0.4	−0.3	−1.5	−0.5	1.7	1.1	−0.8
−2.6	1.6	−3.8	−1.8	1.9	1.2	−0.6	−0.4

(a)

15	0	−1	0	0	0	0	0
−2	−1	0	0	0	0	0	0
−1	−1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

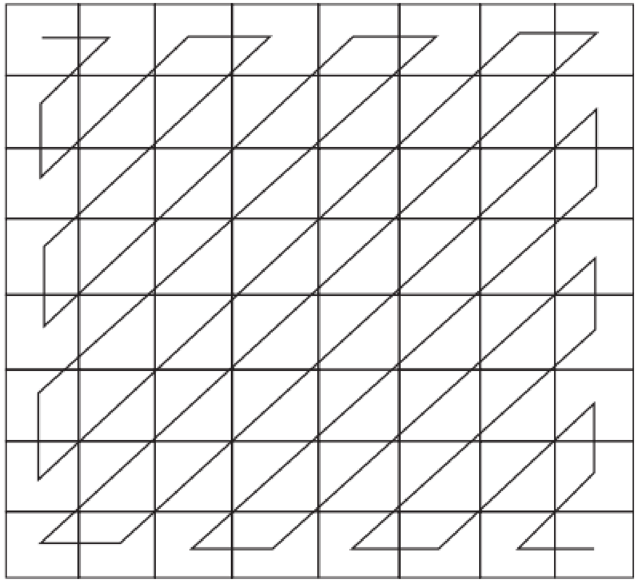
(b)

Quality factor

- Quality factor "Q"
  - Scale the quantization table
  - Medium quality Q = 50% ~ no scaling
  - High quality Q = 100% ~ unit quantization step size
  - Poor quality ~ small Q, larger quantization step
    - visible artifacts like ringing and blockiness

DC coefficient

- DC is encoded as the difference between the dc at each block and the block before.
- AC: Zig Zags



ZigZag

15	0	−1	0	0	0	0	0
−2	−1	0	0	0	0	0	0
−1	−1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$$\{3, 0, -2, -1, -1, -1, 0, 0, -1, EOB\} \Rightarrow 3, (1, -2), (0, -1), (0, -1), (0, -1), (2, -1), EOB$$

Entropy coding – DC difference

Range	DC Difference Category
0	0
−1, 1	1
−3, −2, 2, 3	2
−7, ..., −4, 4, ..., 7	3
−15, ..., −8, 8, ..., 15	4
−31, ..., −16, 16, ..., 31	5
−63, ..., −32, 32, ..., 63	6
−127, ..., −64, 64, ..., 127	7
−255, ..., −128, 128, ..., 255	8
−511, ..., −256, 256, ..., 511	9
−1023, ..., −512, 512, ..., 1023	A
−2047, ..., −1024, 1024, ..., 2047	B
−4095, ..., −2048, 2048, ..., 4095	C
−8191, ..., −4096, 4096, ..., 8191	D
−16383, ..., −8192, 8192, ..., 16383	E
−32767, ..., −16384, 16384, ..., 32767	F

Entropy coding – DC difference

JPEG Standard, Table K3 - Luminance

Category	Code length	Code word
0	2	00
1	3	010
2	3	011
3	3	100
4	3	101
5	3	110
6	4	1110
7	5	11110
8	6	111110
9	7	1111110
10	8	11111110
11	9	111111110

JPEG run-level coding

- *RRRR* – four bits value specifying ac coefficient zero-run of length 0…15
- *SSSS* – four bits specifying a range of magnitudes of the following coefficient ("category")
- Joint Huffman encoding for 8-bit value *RRRRSSSS*
- Append bits for sign and exact magnitude

JPEG coefficient coding categories

Range	DC Difference Category	AC Category
0	0	N/A
−1, 1	1	1
−3, −2, 2, 3	2	2
−7, . . . , −4, 4, . . . , 7	3	3
−15, . . . , −8, 8, . . . , 15	4	4
−31, . . . , −16, 16, . . . , 31	5	5
−63, . . . , −32, 32, . . . , 63	6	6
−127, . . . , −64, 64, . . . , 127	7	7
−255, . . . , −128, 128, . . . , 255	8	8
−511, . . . , −256, 256, . . . , 511	9	9
−1023, . . . , −512, 512, . . . , 1023	A	A
−2047, . . . , −1024, 1024, . . . , 2047	B	B
−4095, . . . , −2048, 2048, . . . , 4095	C	C
−8191, . . . , −4096, 4096, . . . , 8191	D	D
−16383, . . . , −8192, 8192, . . . , 16383	E	E
−32767, . . . , −16384, 16384, . . . , 32767	F	N/A

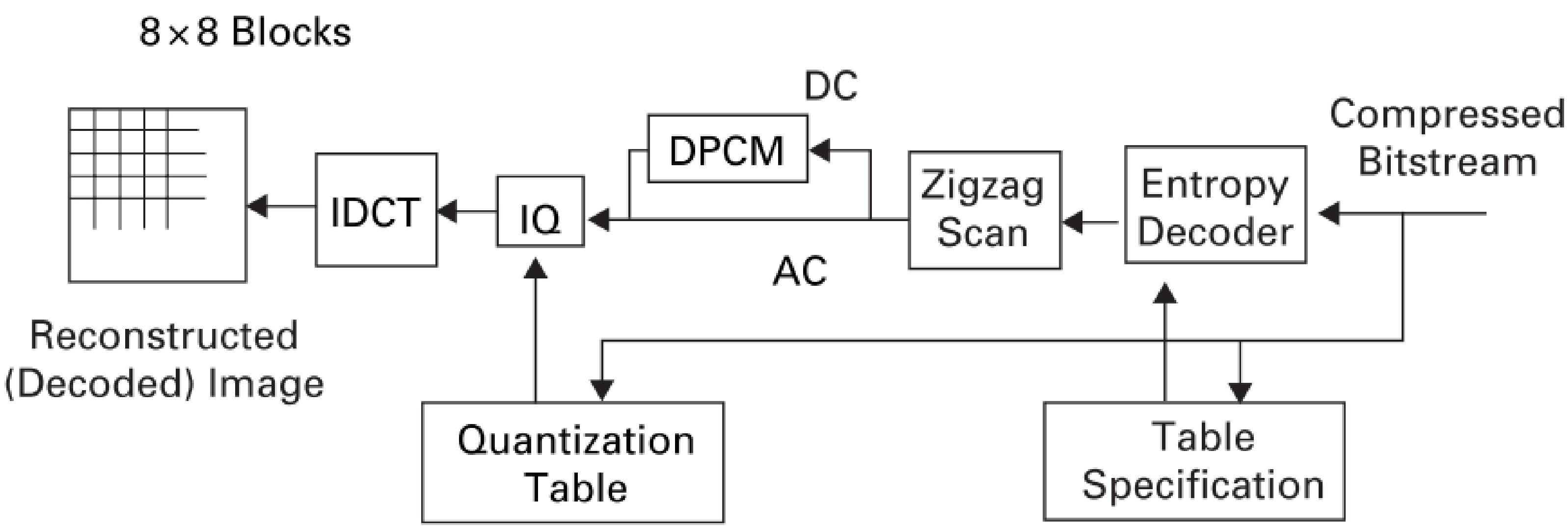
JPEG suggested AC code for luminance

Run/ Category	Base Code	Length	Run/ Category	Base Code	Length
0/0	**1010 (=EOB)**	**4**			
0/1	00	3	8/1	11111010	9
0/2	01	4	8/2	111111111000000	17
0/3	100	6	8/3	1111111110110111	19
0/4	1011	8	8/4	1111111110111000	20
0/5	11010	10	8/5	1111111110111001	21
0/6	111000	12	8/6	1111111110111010	22
0/7	1111000	14	8/7	1111111110111011	23
0/8	1111110110	18	8/8	1111111110111100	24
0/9	1111111110000010	25	8/9	1111111110111101	25
0/A	1111111110000011	26	8/A	1111111110111110	26
1/1	1100	5	9/1	111111000	10
1/2	111001	8	9/2	1111111110111111	18
1/3	1111001	10	9/3	1111111111000000	19
1/4	111110110	13	9/4	1111111111000001	20
1/5	11111110110	16	9/5	1111111111000010	21
1/6	1111111110000100	22	9/6	1111111111000011	22
1/7	1111111110000101	23	9/7	1111111111000100	23
1/8	1111111110000110	24	9/8	1111111111000101	24
1/9	1111111110000111	25	9/9	1111111111000110	25
1/A	1111111110001000	26	9/A	1111111111000111	26
2/1	11011	6	A/1	111111001	10
2/2	11111000	10	A/2	1111111111001000	18
2/3	1111110111	13	A/3	1111111111001001	19
2/4	1111111110001001	20	A/4	1111111111001010	20
2/5	1111111110001010	21	A/5	1111111111001011	21
2/6	1111111110001011	22	A/6	1111111111001100	22
2/7	1111111110001100	23	A/7	1111111111001101	23

JPEG suggested AC code for luminance (cont.)

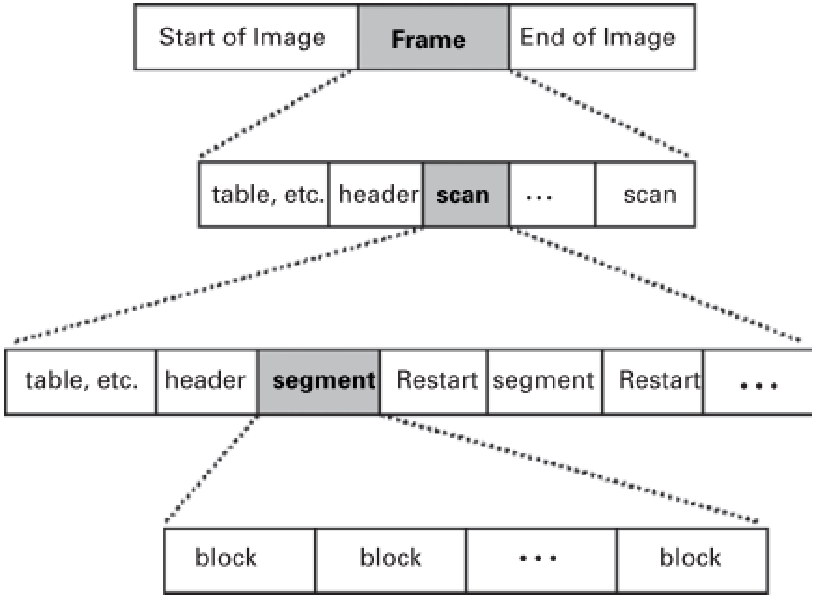
2/8	1111111110001101	24	A/8	1111111111001110	24
2/9	1111111110001110	25	A/9	1111111111001111	25
2/A	1111111110001111	26	A/A	1111111111010000	26
3/1	111010	7	B/1	111111010	10
3/2	111110111	11	B/2	1111111111010001	18
3/3	11111110111	14	B/3	1111111111010010	19
3/4	1111111110010000	20	B/4	1111111111010011	20
3/5	1111111110010001	21	B/5	1111111111010100	21
3/6	1111111110010010	22	B/6	1111111111010101	22
3/7	1111111110010011	23	B/7	1111111111010110	23
3/8	1111111110010100	24	B/8	1111111111010111	24
3/9	1111111110010101	25	B/9	1111111111011000	25
3/A	1111111110010110	26	B/A	1111111111011001	26
4/1	111011	7	C/1	1111111010	11
4/2	1111111000	12	C/2	1111111111011010	18
4/3	1111111110010111	19	C/3	1111111111011011	19
4/4	1111111110011000	20	C/4	1111111111011100	20
4/5	1111111110011001	21	C/5	1111111111011101	21
4/6	1111111110011010	22	C/6	1111111111011110	22
4/7	1111111110011011	23	C/7	1111111111011111	23
4/8	1111111110011100	24	C/8	1111111111100000	24
4/9	1111111110011101	25	C/9	1111111111100001	25
4/A	1111111110011110	26	C/A	1111111111100010	26

JPEG decoder



JPEG code stream

- a "Frame" is a picture, a "scan" is a pass through the pixels (e.g., to obtain the luminance component)
- a "segment" is a group of blocks
- a "block" is an 8x8 group of pixel

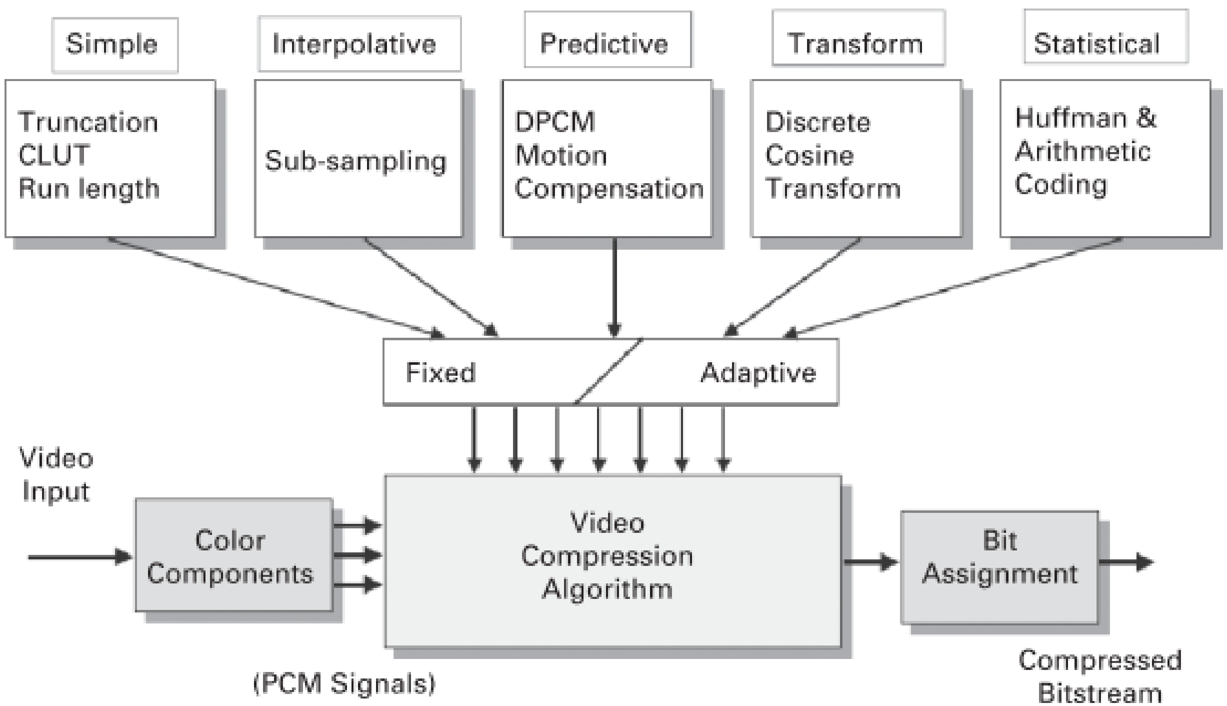


Code stream

- the Frame header includes
  - the sample precision
  - the width and height of the image
  - the number of components
  - the unique ID (for each component)
  - horizontal and vertical sampling factors (for each component)
  - the quantization table to use (for each component)
- the scan header includes
  - the number of components in the scan
  - the component ID (for each component)
  - the Huffman table (for each component)

Digital video coding

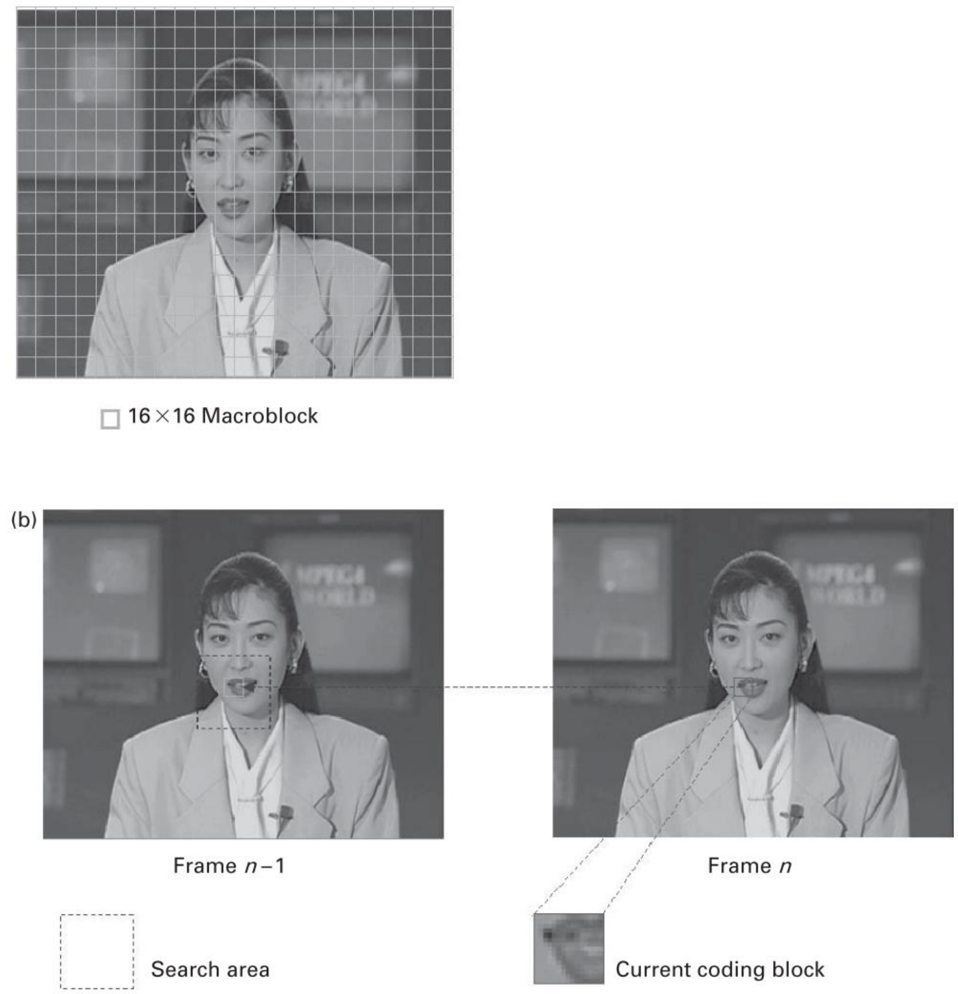
Techniques used in video coding



Predictive coding and motion compensation

- Predictive coding
  - DPCM (Differential Pulse Code Modulation)
- Predictive coding:
  - Use Temporal redundancy
  - Use Spatial redundancy
  - Estimate the "motion vector"
  - Send the motion vector

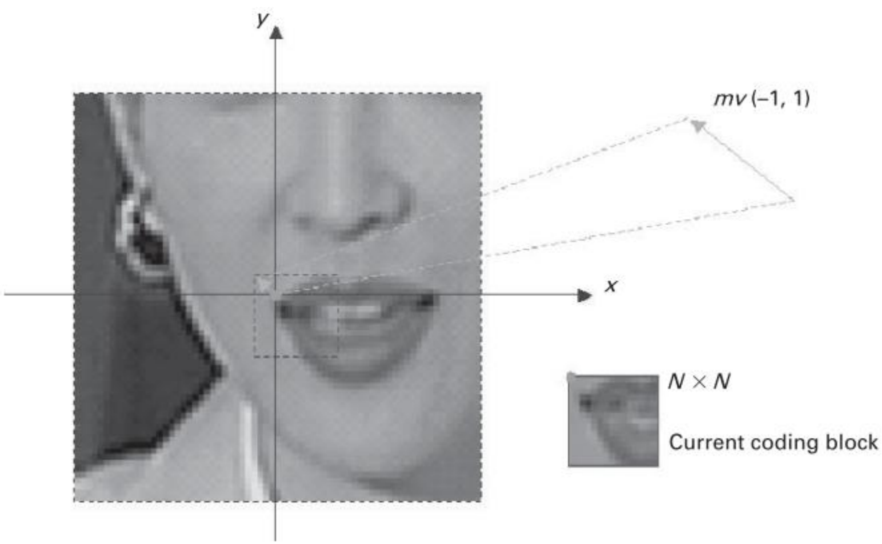
Motion vector estimation



Motion vector estimation

- Sum of absolute difference

$$(u^*, v^*) = \underset{(u,v)}{\arg\min}\, SAD(u,v) = \underset{(u,v)}{\arg\min} \sum_{i=1}^{N} \sum_{j=1}^{N} |a(i,j) - b(i+u, j+v)|$$



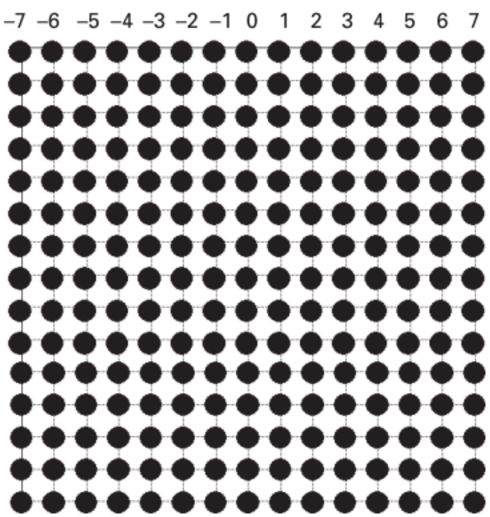
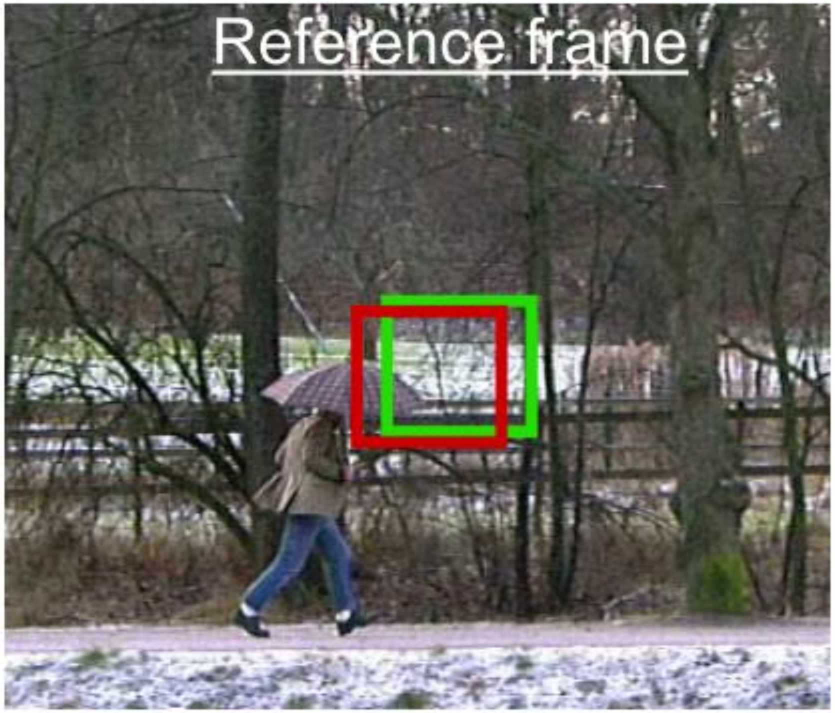


Figure 5.6 An exhaustive search within the motion estimation search area.

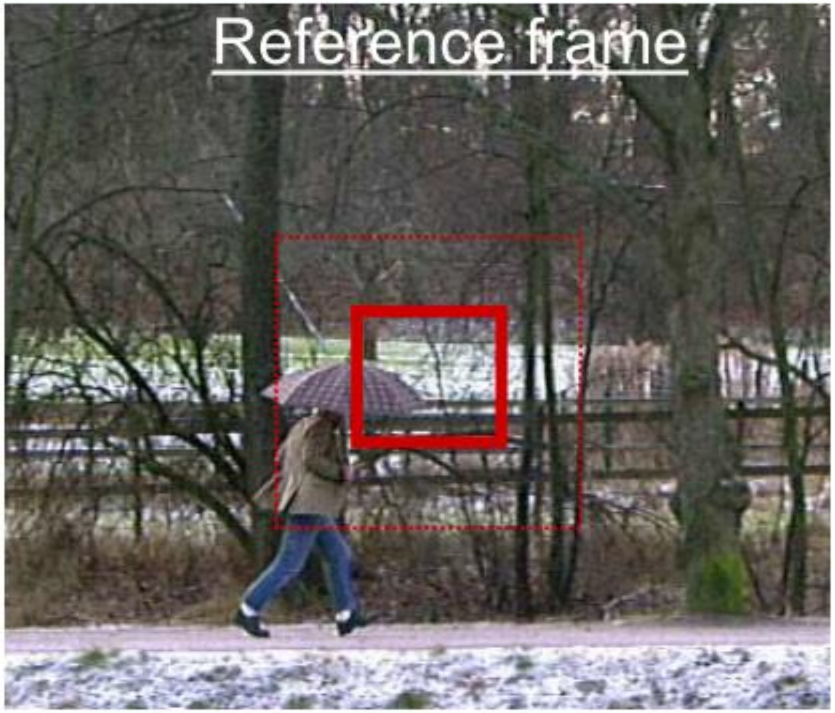
Block-matching algorithm



Block is compared with a shifted array of pixels in the reference frame to determine the best match

Block of pixels is considered

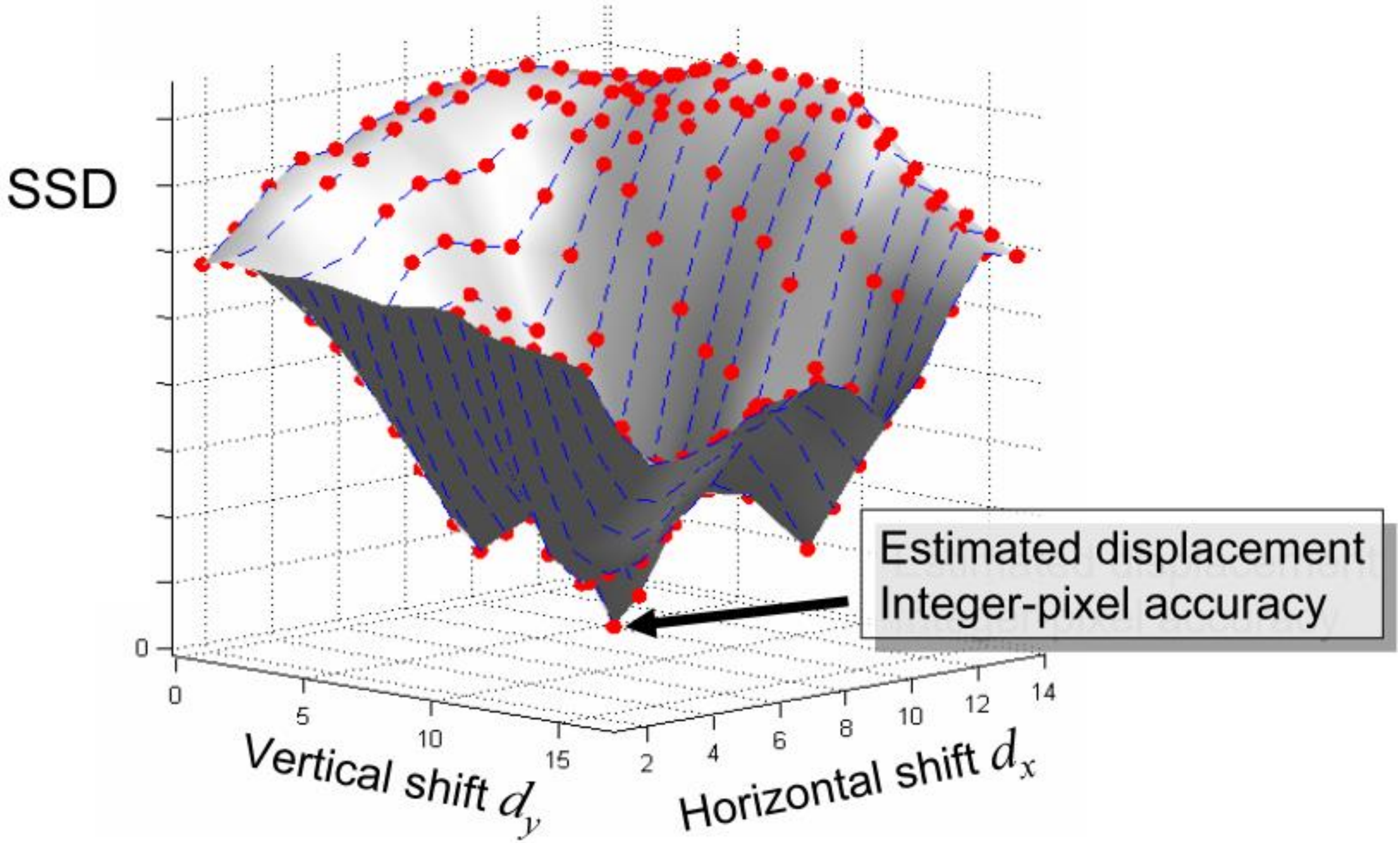
Integer Pixel Shifts



Block is compared with a shifted array of pixels in the reference frame to determine the best match

Block of pixels is considered

SSD Values Resulting from Blockmatching

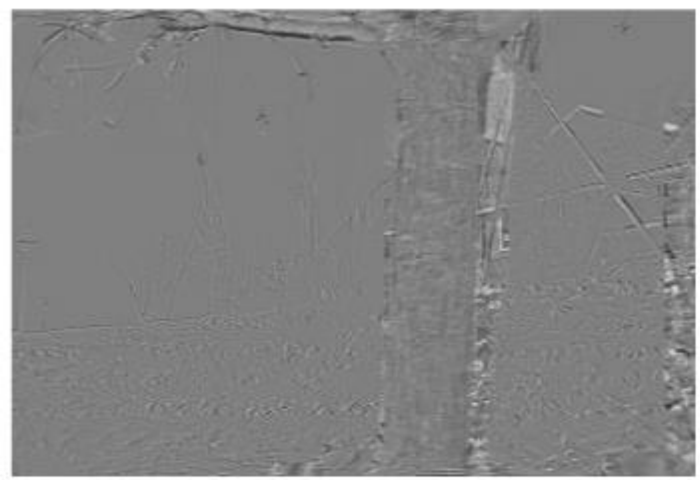


Motion-compensated prediction: example

Previous frame

Current frame

Current frame with displacement vectors



Motion-compensated Prediction error

Fast motion vector estimation

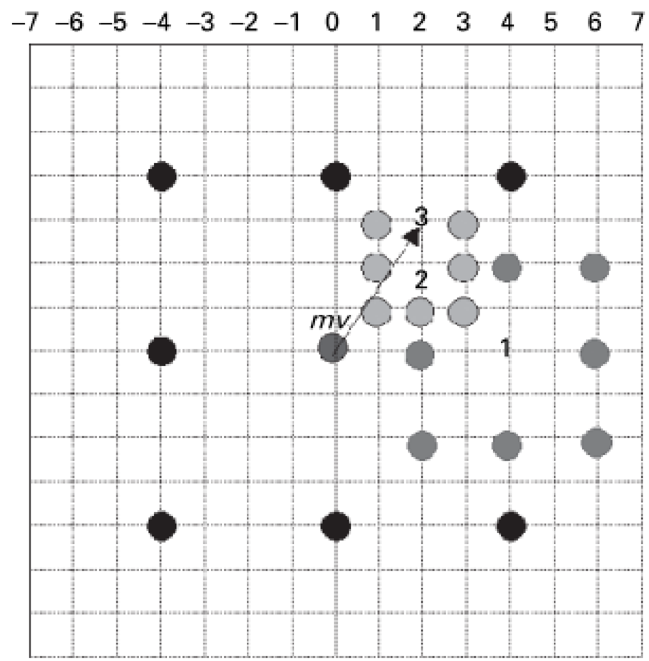


Figure 5.7 The three-step search (3SS) for motion estimation.

Fast motion vector estimation

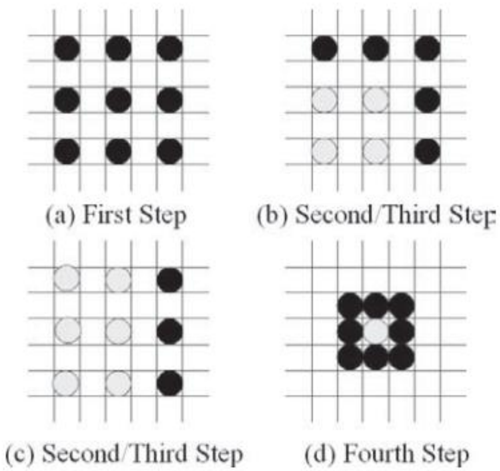
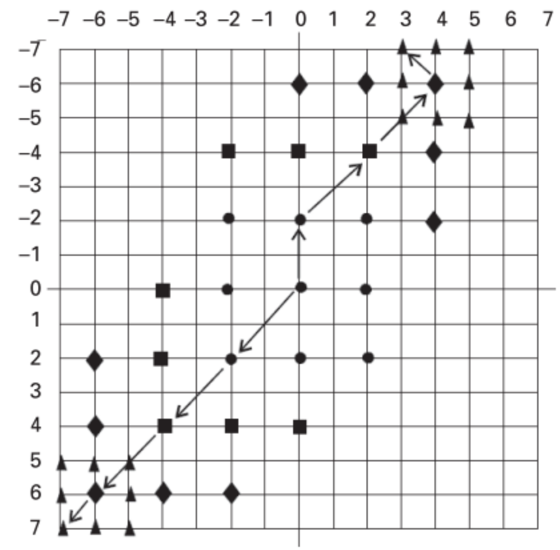


Figure 5.8 The four-step search (4SS) algorithm for motion estimation [18] (© IEEE 1996).



Rate distortion curve

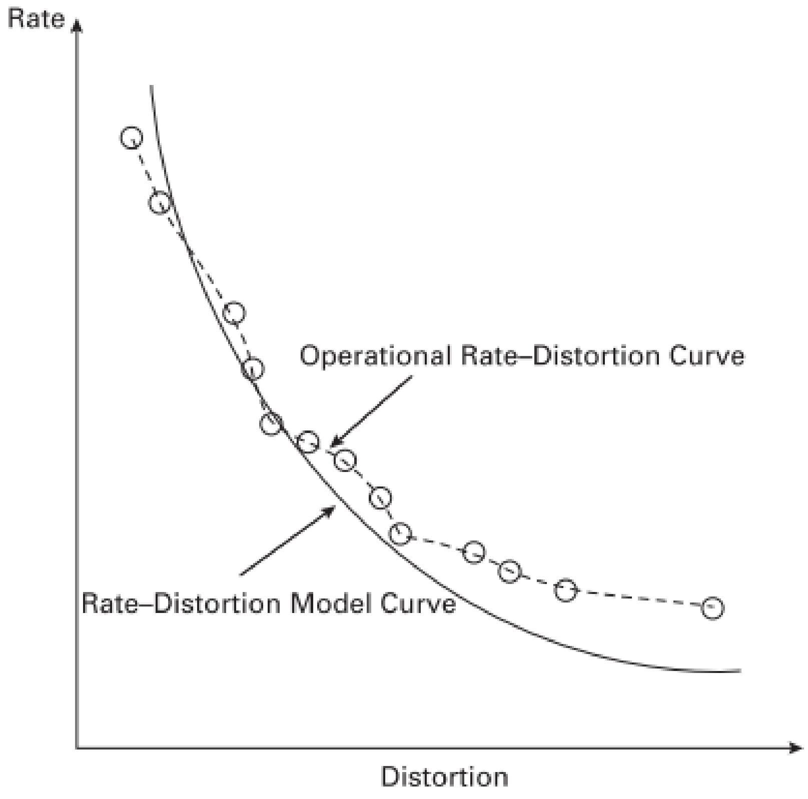


Figure 5.13 According to rate–distortion (R–D) theory, the distortion D is a decreasing function of the bitrate R [21].