Chapter 3

Compression Methods

Yu-Chen Kuo

❑3.1 Compression Methods

★reducing the size of blocks of data by removing unused and redundant material

e.g., silence period in telephone call

- Type of redundancy
  - spatial redundancy

    in almost natural image, neighboring pixels are strongly correlated
  - redundancy in scale

    straight edges and constant regions are invariant under rescaling
  - redundancy in frequency

    spectral values for the same pixel location are correlated

    ⇨ DCT coding
  - temporal redundancy

    adjacent frames often very little change

    ⇨ DPCM, motion compensation coding
  - stereo redundancy

    correlation between stereo channels

    ⇨ audio

- Compression characteristics

  See Table 3-1 pp. 37

  Figure 3-1

- Coding techniques

  – Entropy coding independent of lossless characteristics of data stream

  – Source coding take the semantics lossless / into account lossy

  – Hybrid coding H.261, H.263, JPEG MPEG

    Fig. 3-2 pp.39

    See Table 3-2 pp. 38

❑3.2 Basic Coding Methods

■ Entropy Coding

▪ Run-Length Coding

repetitive sequences

uncompressed data :UNNNNNNIMA***NN***HEIM

run-length coded: U!6NIMA***NN***HEIM

★ longer sequences of different characters can also be applied

e.g., WER**!6ABC**IMECAD

▪ Huffman coding

optimal code (minimal number of bits) to represent data

⇨ shortest code for the character occurring most frequently

See Fig. 3-3 pp. 41▪

- Source coding
  - Differential Pulse Code Modulation (DPCM)

    reduce range of numerical input characters

    See Fig. 28.5 pp. 515 and Fig. 28.6 pp. 516

    e,g.,

    uncompressed data :10 12 14 16 18 20

    run-length coded: 10 12 14 16 18 20

    DPCM coded: 12 2 2 2 2 2

    ⇨ run-length coded: 12 !52
  - Discrete Consine Transform
    - Raster → Block

      — dividing an image into 8×8 pixel blocks (components)

      Fig. 30.5 pp.532

- RGB → YUV
  - each pixel is represented by RGB triplet
  - transfer RGB triplet into Y (luminance) UV (chrominance) triplet
  - Y =77/256R + 150/256G +29/256B
    U =-44/256R − 87/256G+131/256B+128
    V =131/256R − 110/256G − 21/256B+128
    Fig. 30.2 pp. 530
- DCT
  - each 8×8 block is mapped from space domain into frequency domain
    (a composition of DCT basis function, Fig. 3-4 pp. 43 with 64 coefficients)
  - human is sensitive at low-intensity levels
  - ⇨ reduce to number of high-frequency DCT coefficients won‘t affect quality

See Fig. 3-5 pp. 43, DCT easy to separate low/high frequencies

- Quantization (compress step)

  See Fig. 30-6 pp. 533

  ⇨ make DCT coefficients with high frequency to be zero

  if careful choice of quantization table

- DPCM encoding the DC coefficients among blocks

  Fig. 30-7 DC coefficients

- Run-length encoding the AC coefficients in a block

  — ZigZag coding sequence: code lower frequency first

  Fig. 30.8

- Huffman encoding

- Motion-compensated prediction
  - — reduce temporal redundancies between two frames
  - — looking blocks in previous and current frames is very closely
  - — record difference signal (DPCM-code)
  - — motion vector
  - ◆ unidirectional prediction

    only previous/subsequent frame is used

    Fig. 3-8 pp. 47
  - ◆ bidirection prediction

    both previous/subsequent frame is used
  - ◆ interpolative prediction

    average two predictions based on pervious frame and future frame (two motion vectors)

- vector quantization
  - dividing data stream into blocks (vectors)
  - predefined a code-book (a set of patterns)
  - finding the best matching pattern in code-book for each vector
  - transmitting the index of the best matching patterns only not the actual data
  - choose an optimal code-book is important

See Fig. 28-7, 28-8 pp.519

❑3.3 Video Compression

▪ JPEG (Joint Photographic Experts Group)

— Sequential DCT-based mode (most JPEG cords)

◆DCT→quantization→entropy coding

◆encode each block by a single scan, left-to-right and top-to-bottom

see Fig. 3-11/Fig. 30.1 pp. 529

— Progressive DCT-based mode

◆ multiple scans for each block rather a single scan

◆ provide an quick transmission and presentation

— Lossless mode

e.g., X-ray

— Hierarchical mode

◆ encode image at multiple resolutions

◆ lower resolution can be accessed without decompress the full resolution

◆ having multiple resolutions within an image file

★ compress ratio 20~25:1 in lossy mode

2:1 in lossless mode

- MPEG-1 video (Motion Picture Experts Group)
  - a bit-stream for synchronized digital audio and video to fit into a bandwidth of not more than 1.5Mbps
  - 1.1 Mbps for **Video**, 128Kbps for **Audio** and others for MPEG **System** data
  - MPEG system stream is responsible for synchronizing one video stream and one/multiple audio streams and allowing random access, fast-forward, rewind
  - motion compensation

    See Fig. 30.11 pp. 538

    - reference frame:

      a frame from which other frames are constructed

    - intracoded frame (I-frame):

      a frame not built from any other frame

      ★ for random access propose

      See Fig. 30.11 pp. 539

- macroblock
  - a 16×16 pixel area
    - four 8×8 blocks for luminance plane
    - two 8×8 blocks for two color difference planes
  - See Fig. 30.13 pp.540
- motion vector
  - indicates the spatial translation of a macroblock between two frames
  - See. Fig. 30.12 pp. 539
- predicted frame (P-frame)
  - a frame only reconstructed from a preceding frame I/P
  - e.g., frame F3 in Fig. 30.10
- bidirectional frame (B-frame)
  - a frame reconstructed from forward /backward frames
  - e.g., frame F2 in Fig. 30.10

- DC coded frame (D-frame)
  - ★ for forward/rewind mode
- frame relation
  - See Fig. 3-17 pp. 58
- data hierarchies of MPEG
  - ◆ GOP (Group of Picture)
    - a fix number of consecutive frames starting at an I-frame
    - IBBBPBBBI
    - IBBPBBPBBI for PAL (352×288)
    - IBBPBBPBBPBBI for NTSC (352×240)
    - increasing #B $\rightarrow$ increasing compress ratio $\rightarrow$ decreasing the correlation
    - See Fig. 3-18 pp. 58

- coding of I-frame
  - like JPEG for a still image
  - See Fig. 30.17 pp. 543
- coding of P- and B-frames
  - ◆ search best matching macroblock for each macroblock
  - ◆ estimate the motion vector→DPCM
  - ◆ take prediction error into account→DCT
  - ◆ if can‘t find any satisfactory matching block
  - → macroblock is simply **intracoded** as I-frame
  - See pp. 30.18 pp. 544

- MPEG-1 Constrained Parameter Set (CPS)
  - See Table 3-5 pp. 59

- MPEG-2 video
  - MPEG-1 for the storage of VCR-quality audio-video sequence on CD-ROMs
  - MPEG-II for recording and transmission of studio-quality motion video at bit rate up to 100Mbps (4~6Mbps, general) over satellite, cable, broadcast channel
  - intends to support HDTV requirements (MPEG-III)
  - MPEG-II video profiles (compression algorithm) and levels (parameter i.e., sample rate)

    See Table 3-6 pp. 61
- MPEG-4 video
  - targets videoconferencing at very low bit rate (4.8~64kbps)
  - frame rate 10fps
  - H.263 using 20kbps for video and 6.5kbps for audio

- MPEG-4編解碼技術意義
  - MPEG-4里是採用了物件的觀念，壓縮之前先將影片中各個主要物件區分開來。
  - 將各個物件分隔開的好處，就是可以針對不同的物件特性採取不同的壓縮編碼技巧，以得到最好的壓縮效果。

❑3.4 Audio Compression

- human auditory is more sensitive than human visual system
- compression ratios is far below than for digital video

- MPEG-1 audio
  - lossy, extending to achieve lossless quality
  - exploiting hearing threshold and auditory masking to remove irrelevant signal
  - hearing threshold: to filter signal into 32 subbands

    See Fig. 3-24 pp. 69
  - auditory masking: strong signal may make weaker signal imperceptible → leave strong signal by quantization

    See Fig. 3-25 pp. 70
  - MPEG-1 audio encoder/decoder

– MPEG-1 audio layer

- three layers compression: Layer I, II, III
- high layer is able to decode audio signal by lower layers
- Layer I: simplest, 192/256 Kbps/channel
- Layer II: 96/128 Kbps/channel, CD-quality
- Layer III: 32~64Kbps/channel, variable bit rate, ISDN
- support one/two audio channels, two independent channels or one stereo signal
- MPEG-II audio has five channels (right, left, center, two surround) and low frequency channel

Characteristics	Description
lossless	Original data can be recovered precisely.
lossy	Not lossless.
intraframe	Frames are coded independently.
interframe	Frames are coded with references to previous and/or future frames, i.e., temporal redundancies between frames are taken into account.
symmetrical	Encoding and decoding time are almost equal.
asymmetrical	Coding time considerably exceeds decoding time.
real-time	Encoding-decoding delay should not exceed 50 ms.
scalable	Frames are coded in different resolutions or quality levels.

Table 3–1 Characteristics of Compression Methods

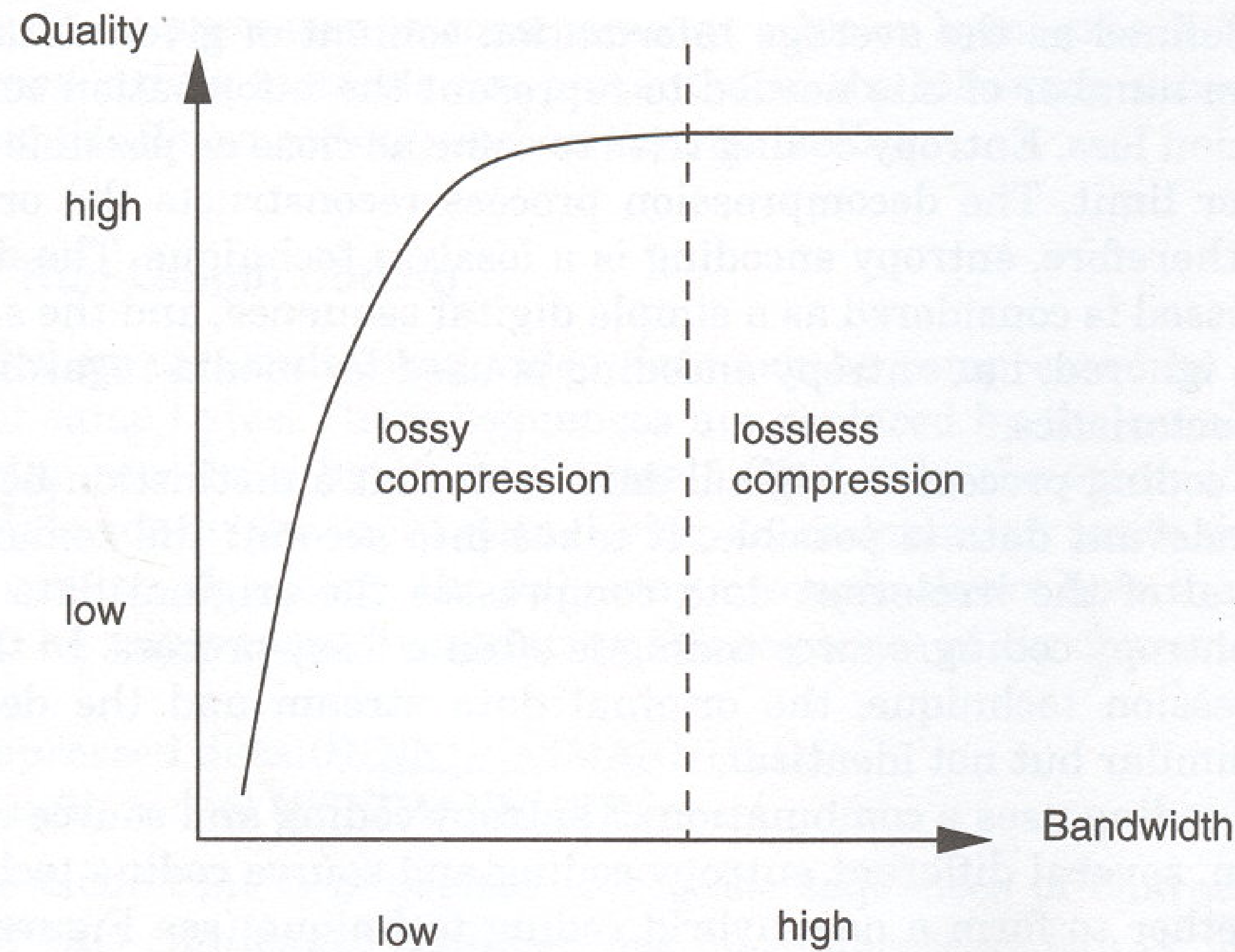


Figure 3–1 Relation Between Perceptible Quality and Required Bandwidth

Coding Technique	Examples
Entropy coding	Arithmetic coding
	Huffman coding
	Run-length coding
Source coding	Differential pulse code modulation
	Discrete cosine transform
	Discrete wavelet transform
	Fourier transform
	Iterated function system
	Motion-compensated prediction
Hybrid coding	Fractal image compression
	H.261
	H.263
	JPEG
	MPEG video
	MPEG audio
	Perceptual Audio Coder
	Wavelet image compression

Table 3–2 Classification of Coding Techniques for Multimedia Systems

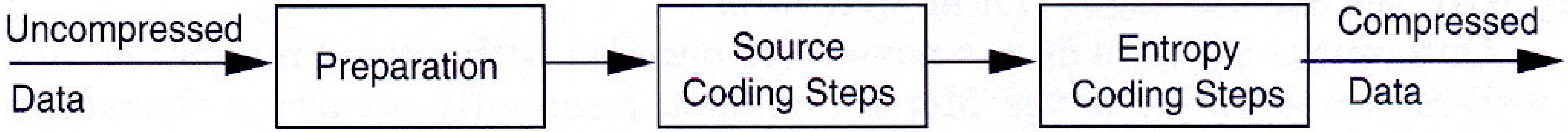


Figure 3–2 Major Encoding Steps of Hybrid Coding Techniques

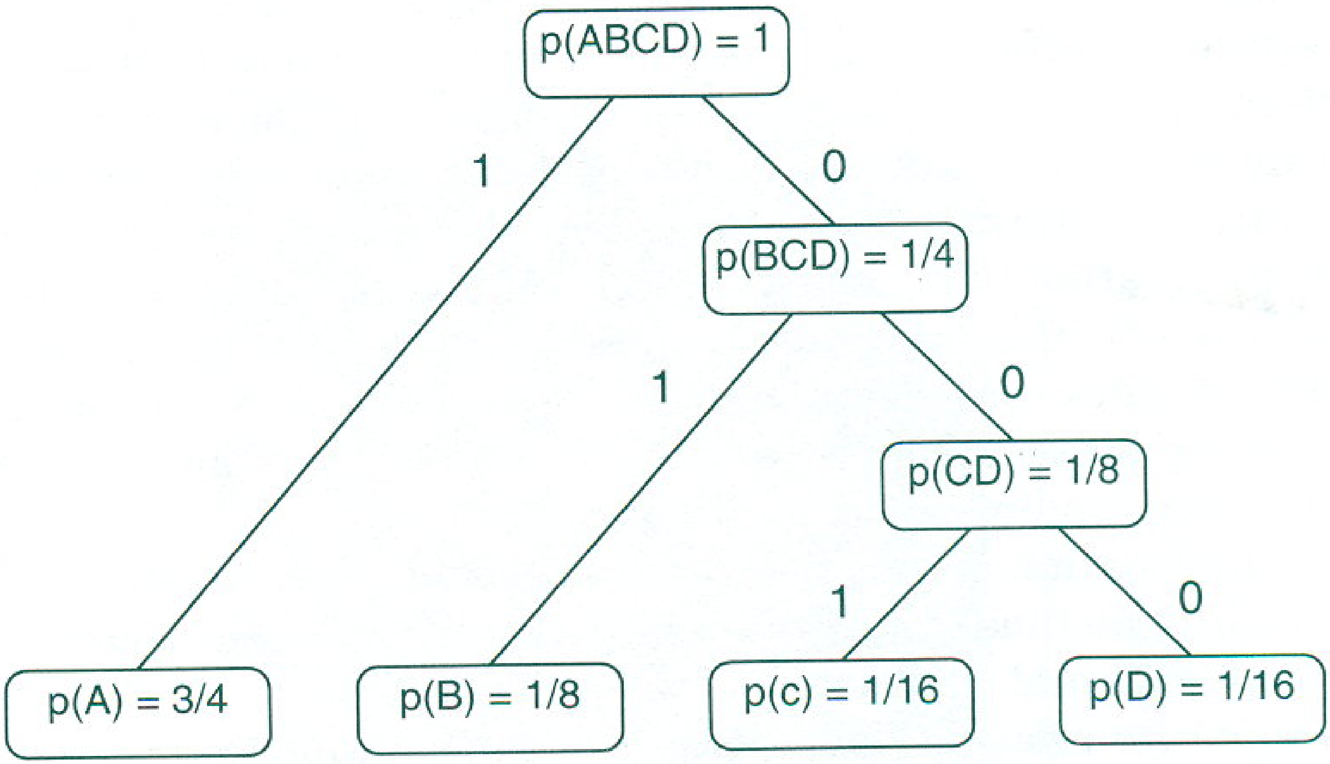


Figure 3–3 Example of Huffman Coding

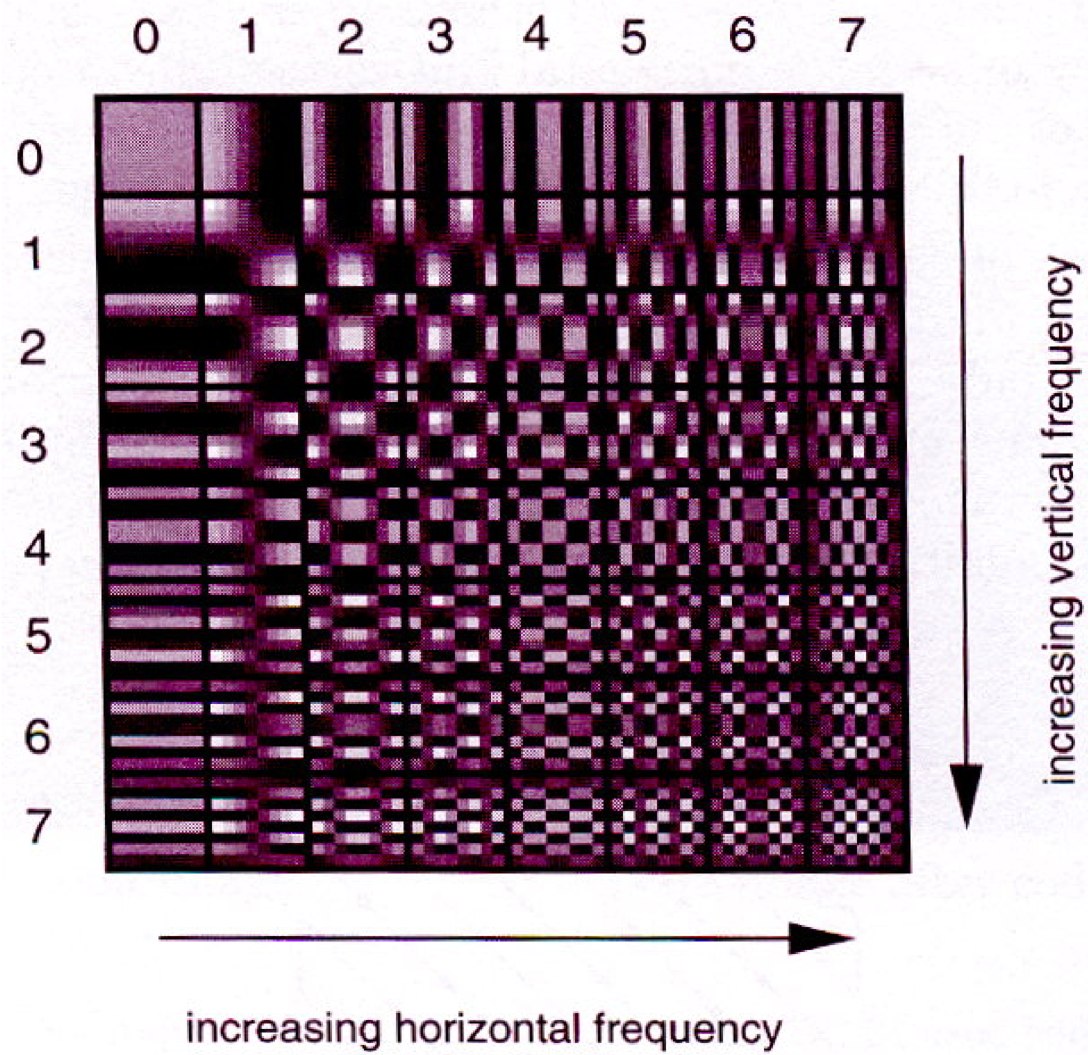


Figure 3–4 DCT Basis Functions

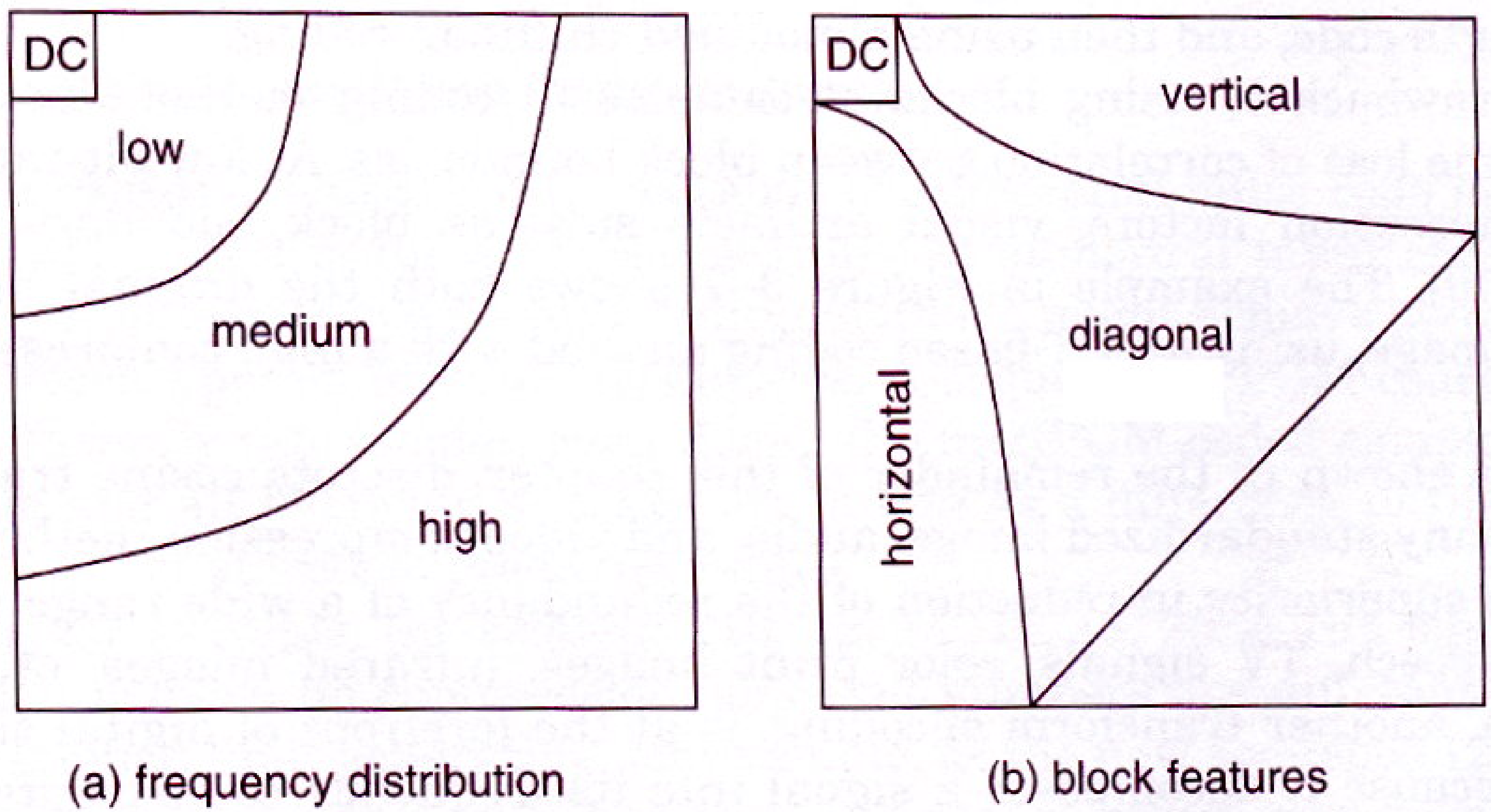


Figure 3–5 Frequency Distribution
(a) of two-dimensional DCT coefficients and block features (b) they represent

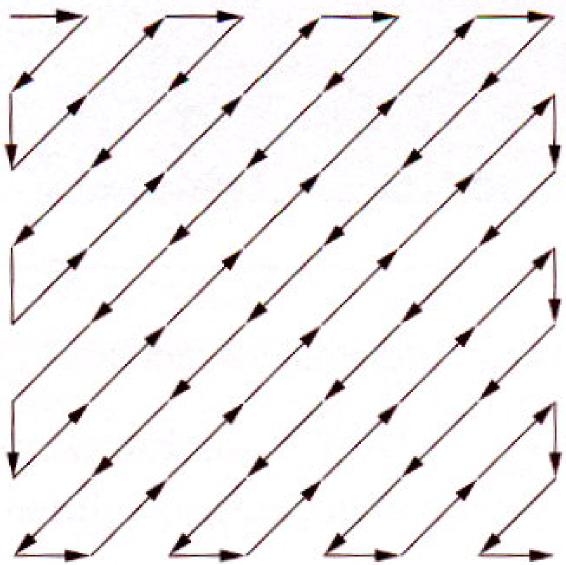


Figure 3–6 Zigzag Sequence: Order of AC Coefficients with Increasing Frequency

Original Lena image | Lena image after coding/decoding

Figure 3–7 Example of blocking artifacts using DCT-based coding methods

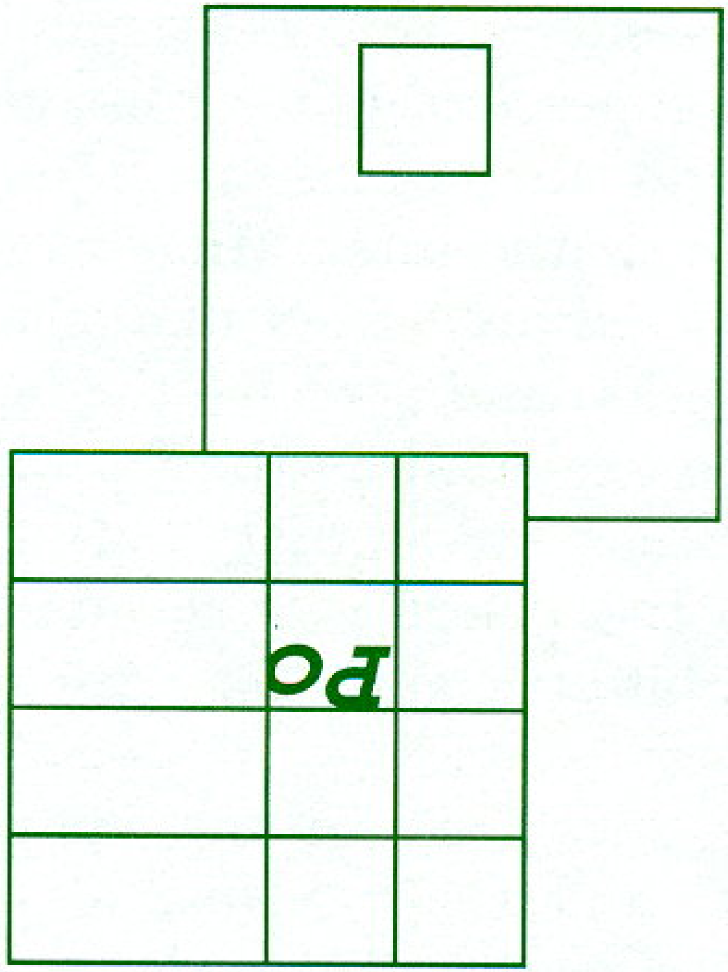


Figure 3–8 Forward Motion-Compensated Prediction

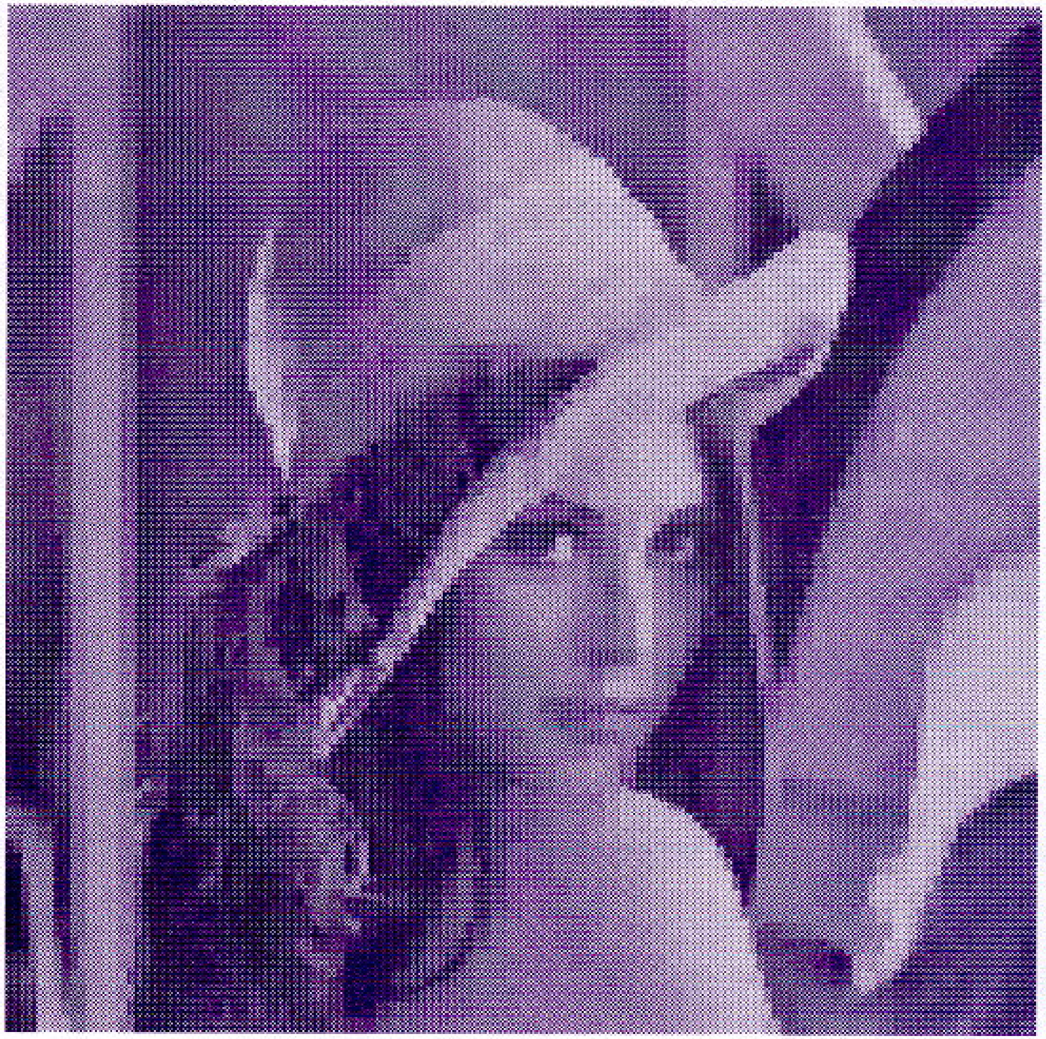


Figure 3–9 Typical Artifacts Caused by a CCC-based Coding Method

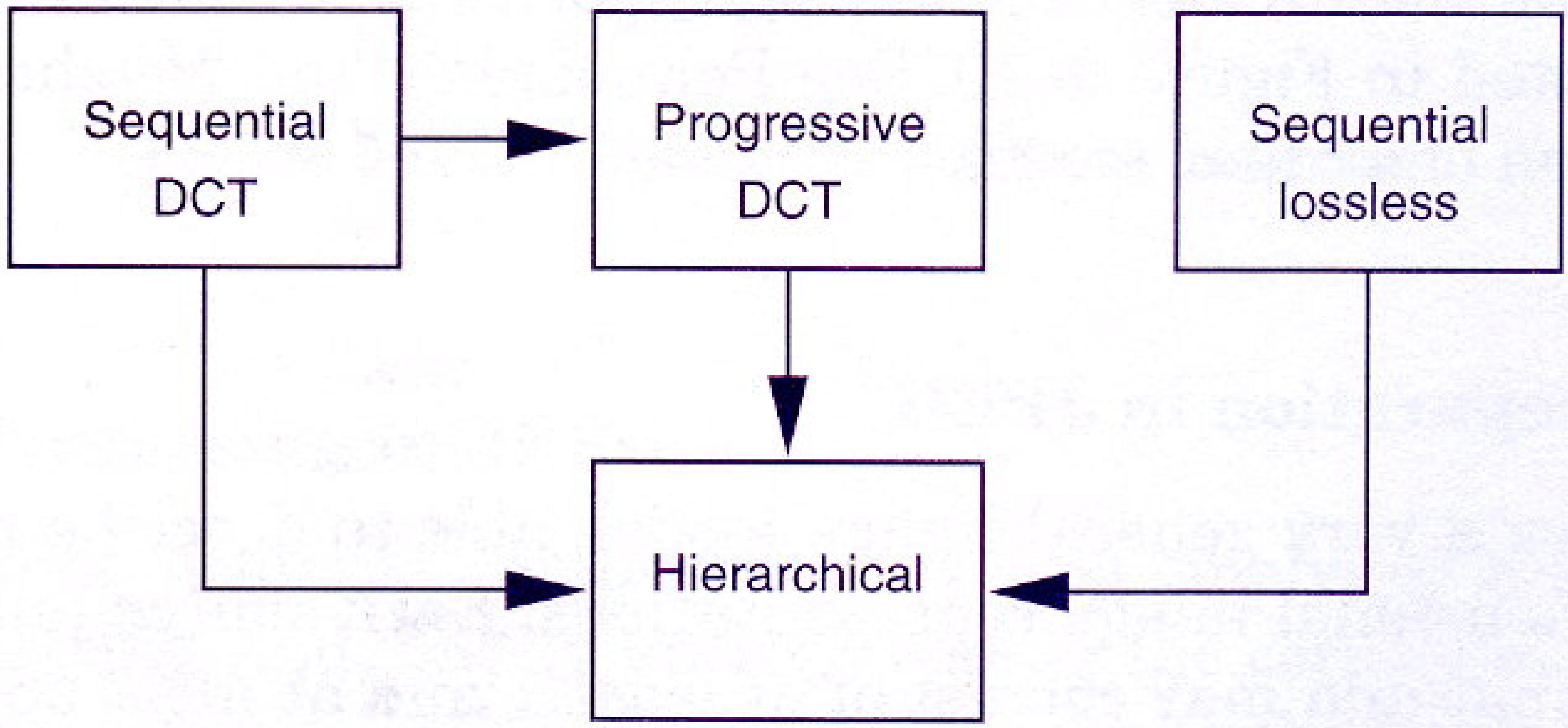


Figure 3–10 Four Coding Modes of JPEG

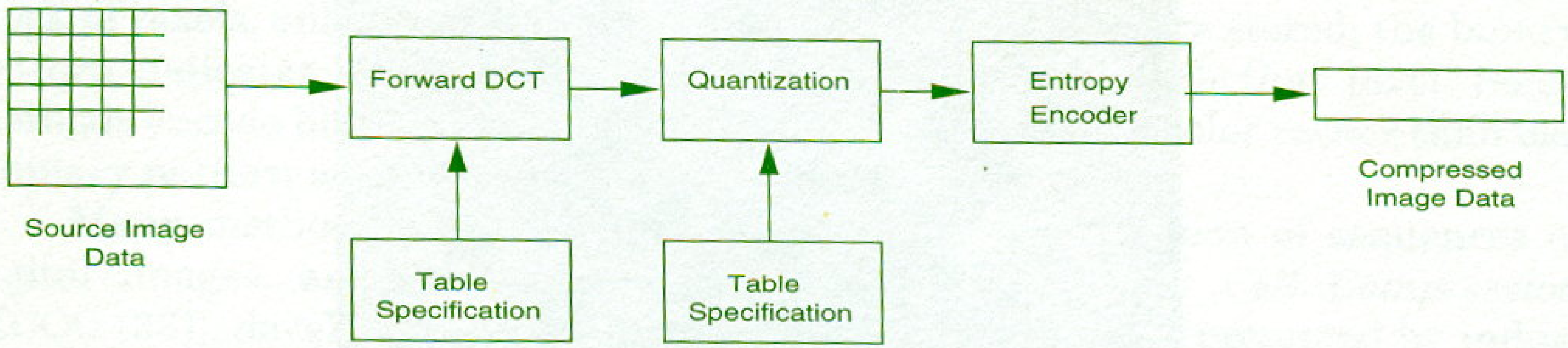


Figure 3–11 Processing Steps of Baseline Mode Encoder

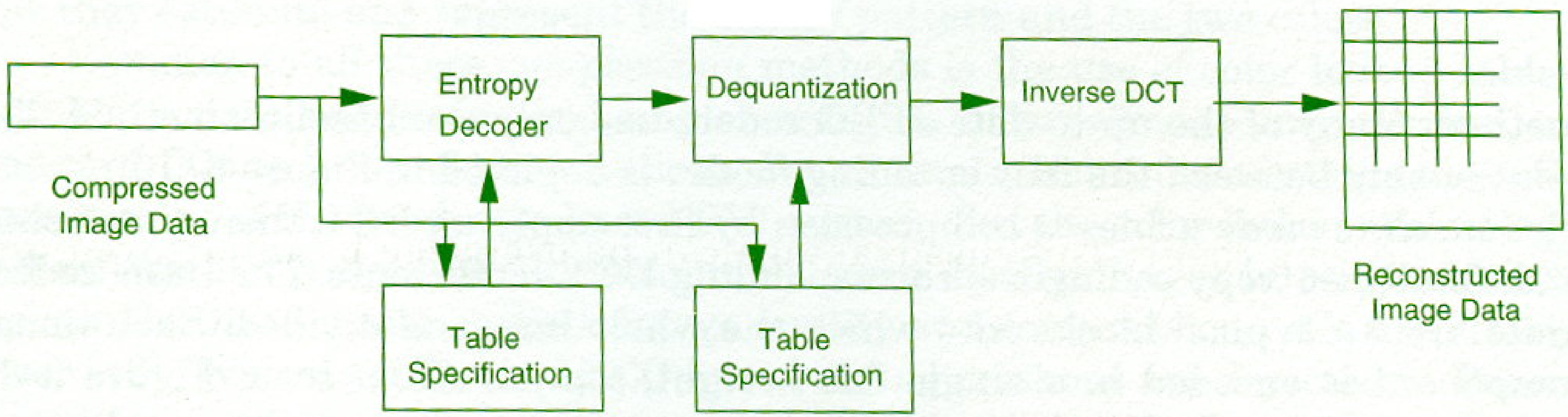


Figure 3–12 Processing Steps of Baseline Mode Decoder

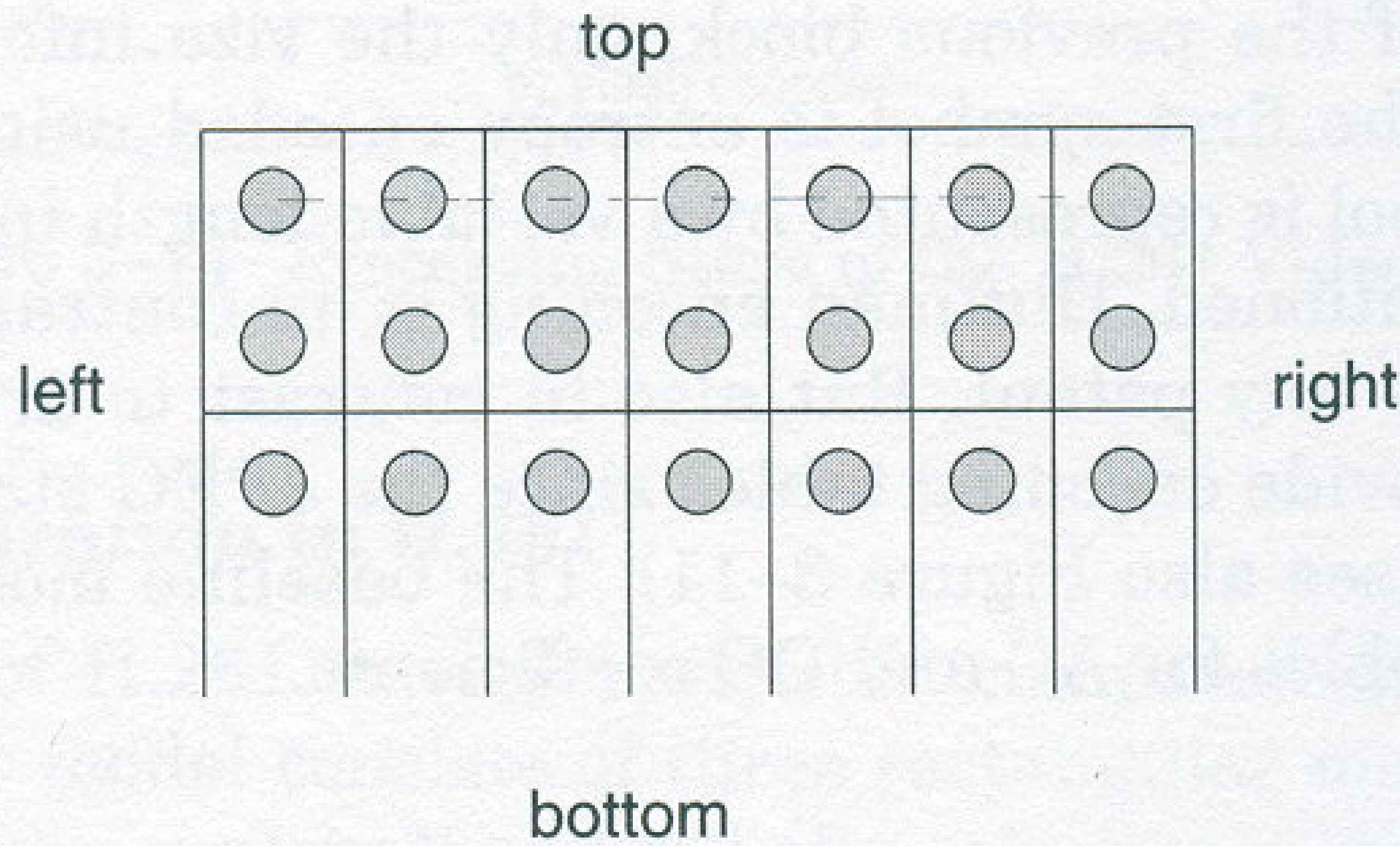


Figure 3–13 Noninterleaved Data Ordering

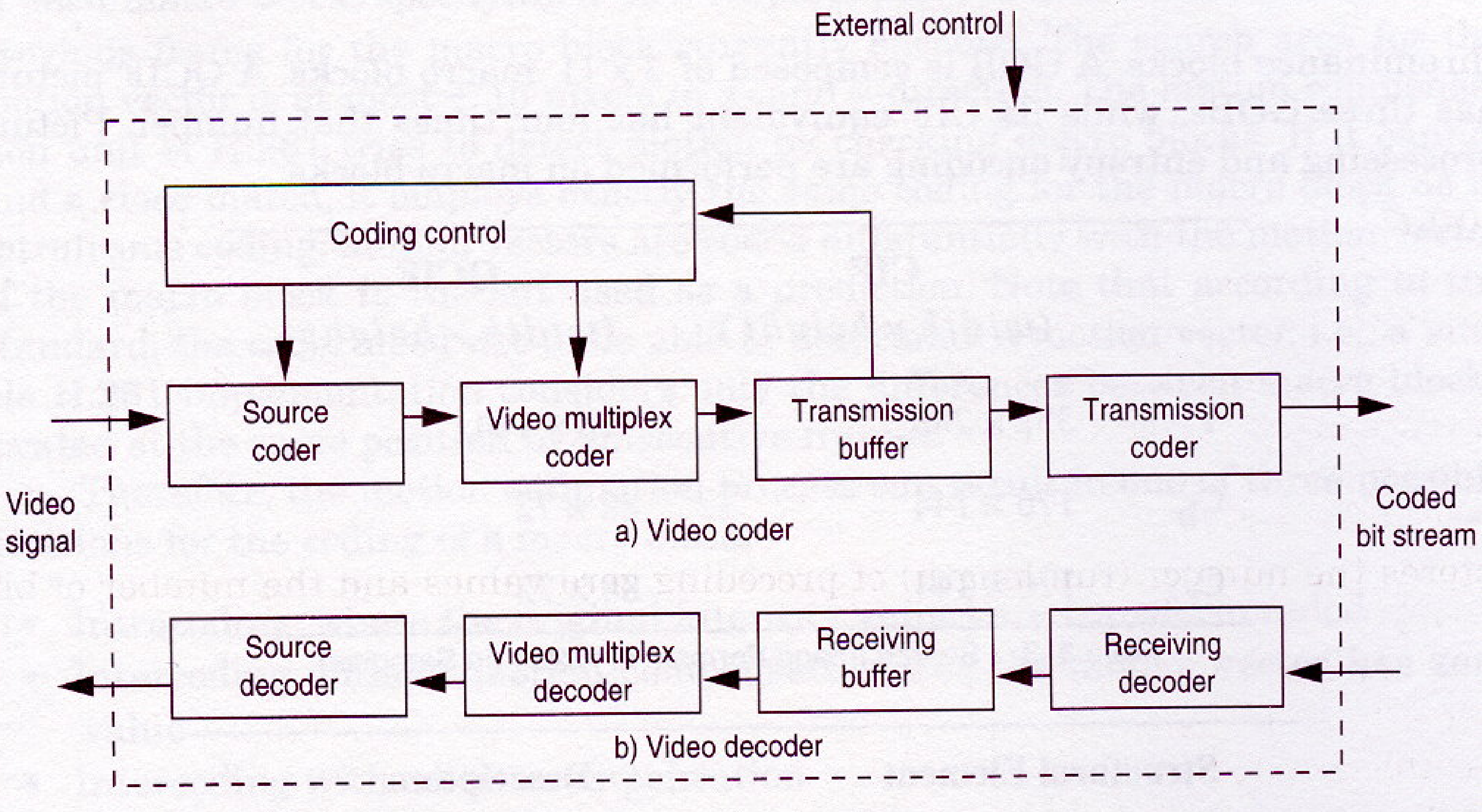


Figure 3–14 Processing Steps of the H.261 Video Codec

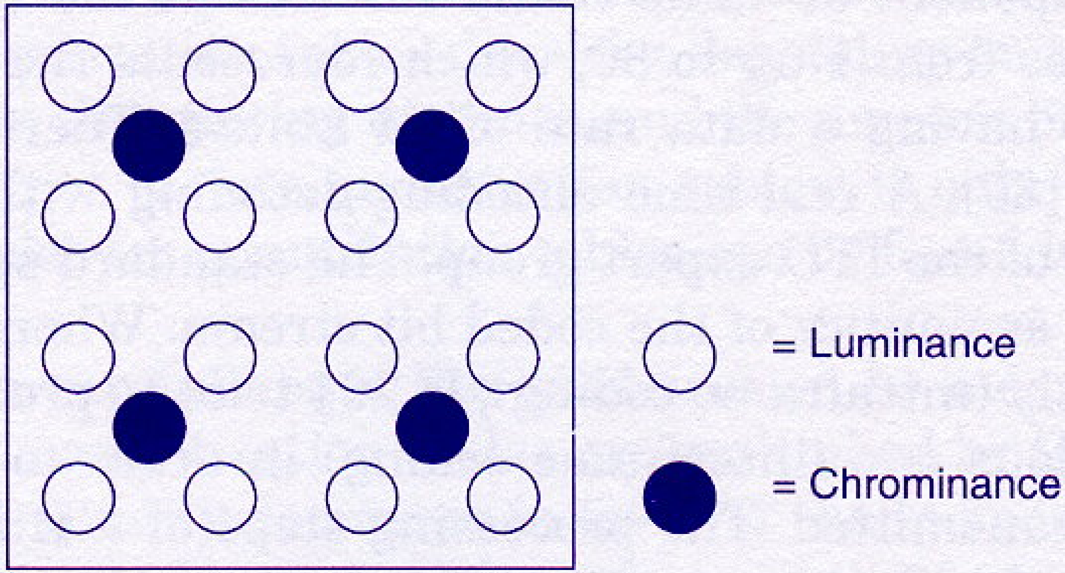


Figure 3–15 Location of Luminance and Chrominance Samples in H.261

	CIF $(width \times height)$	QCIF $(width \times height)$
Y	352×288	176×144
C_b	176×144	88×72
C_r	176×144	88×72

Table 3–3 Source Image Formats of H.261 (in Samples)

Structural Element	Description
Picture	1 video picture (frame)
Group of blocks	33 macro blocks
Macro block	16×16 Y, 8×8 C_b, C_r
Block	8×8 pixels (coding unit for DCT)

Table 3–4 Hierarchical Block Structure of H.261

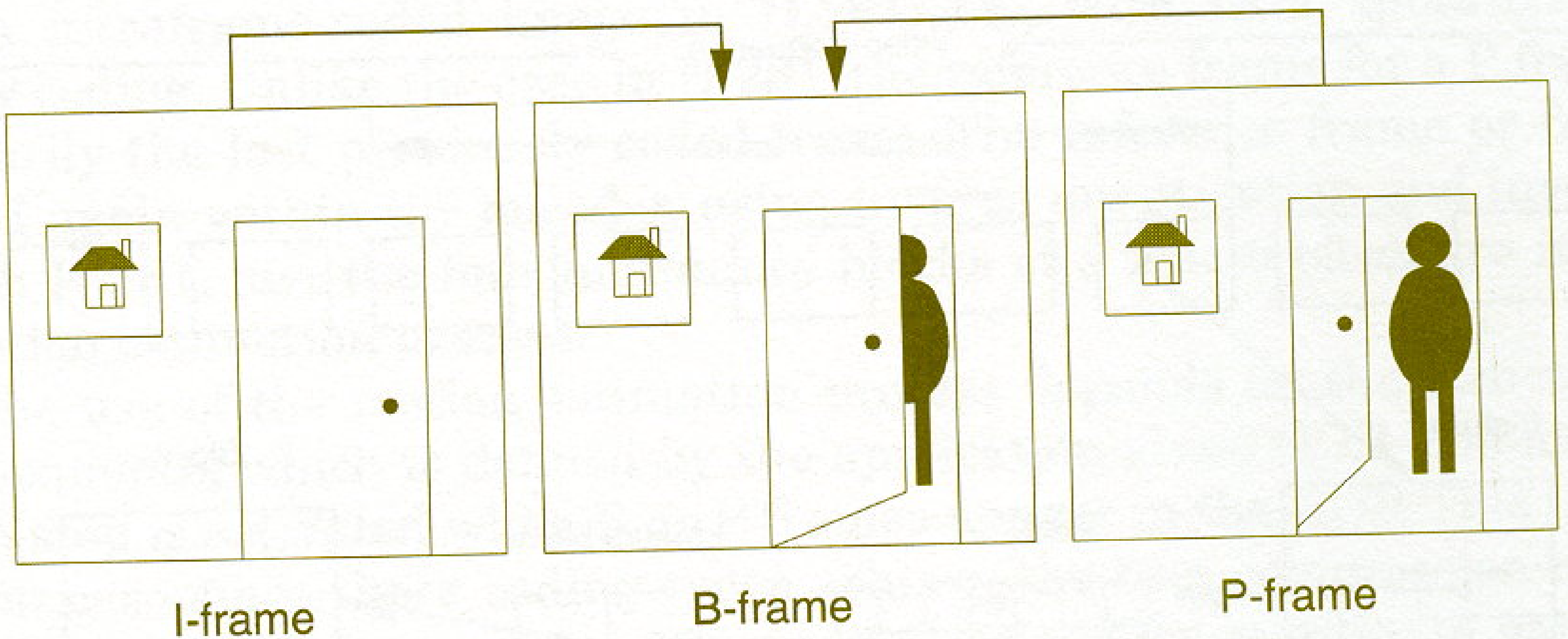


Figure 3–16 Use of B Frames in MPEG Video

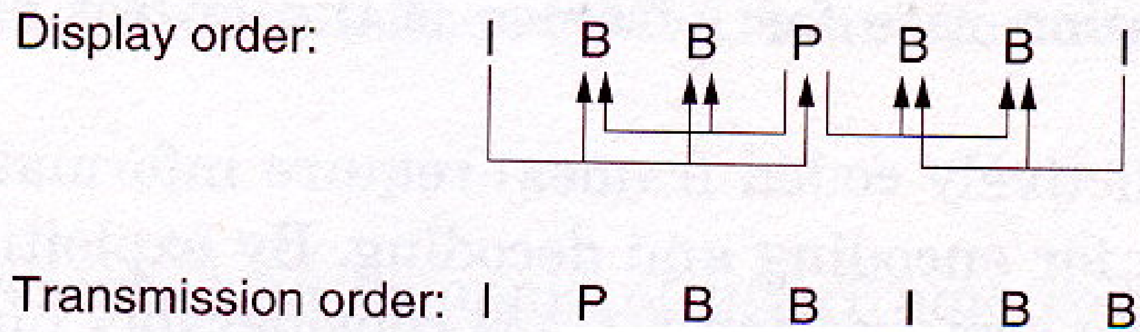


Figure 3–17 Display and Transmission Order in MPEG-1 Video

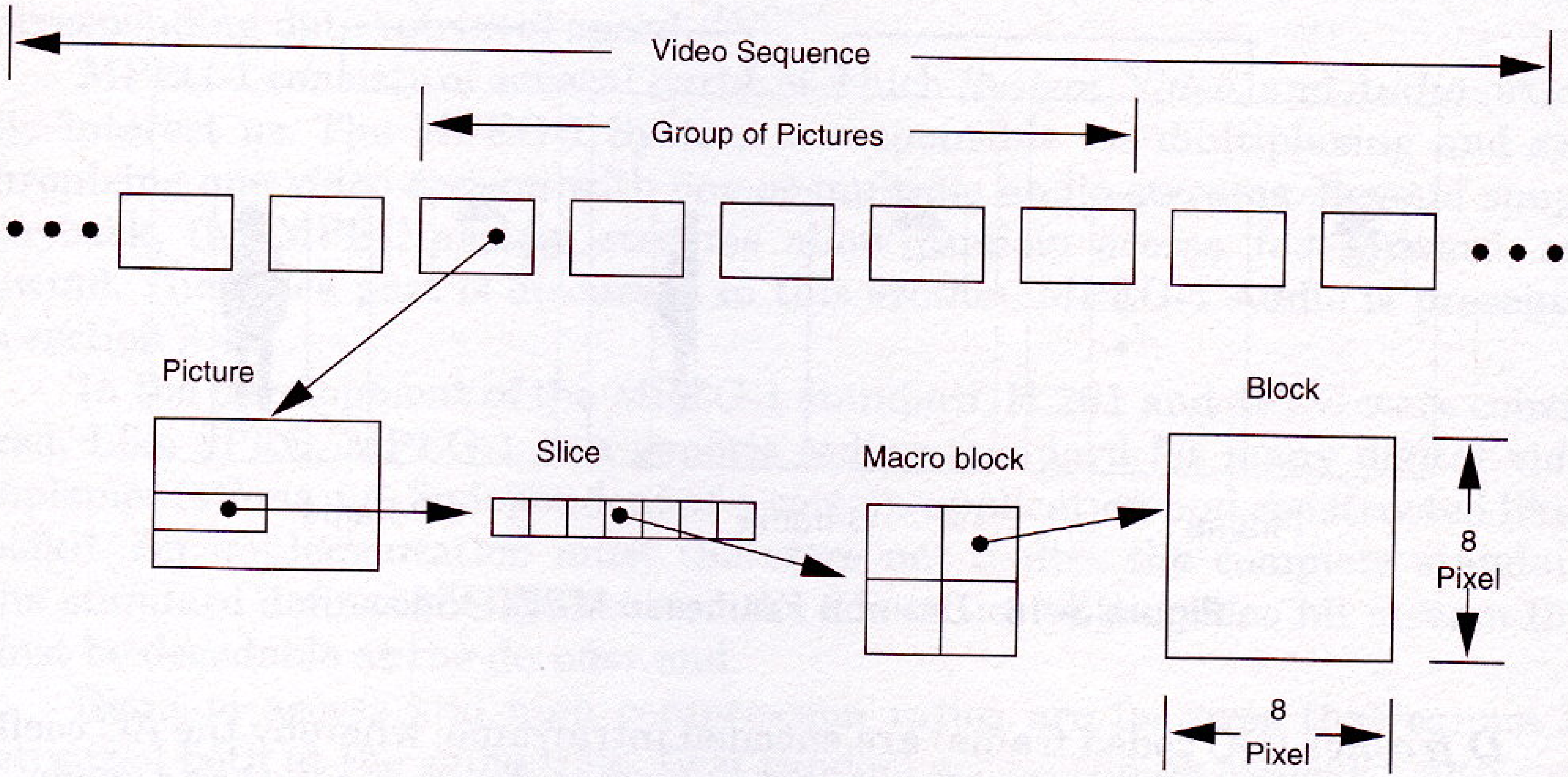


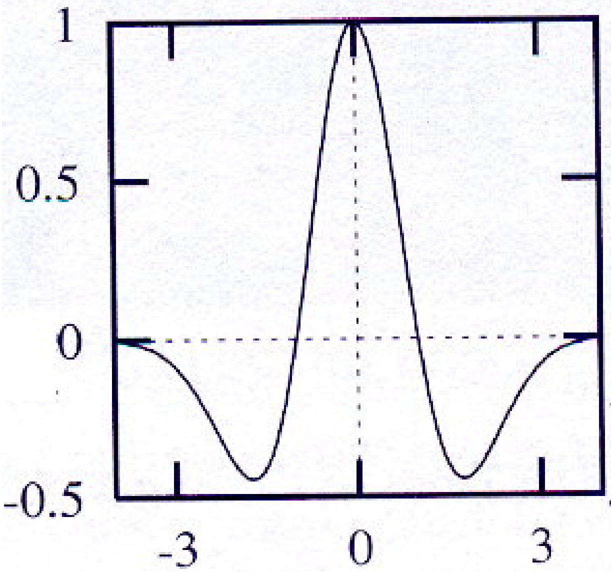
Figure 3–18 Data Hierarchies of MPEG Video Streams

Parameter	Restrictions
Horizontal resolution	≤ 768 pixel
Vertical resolution	≤ 576 lines
Macro blocks/s	≤ 25 macro blocks/s
Frames/s	≤ 30 Hz
Motion vector range	≤ (-64 / + 63,5) pixel
Input puffer size	≤ 327.680 bit
Bit rate	≤ 1*856 Mbps

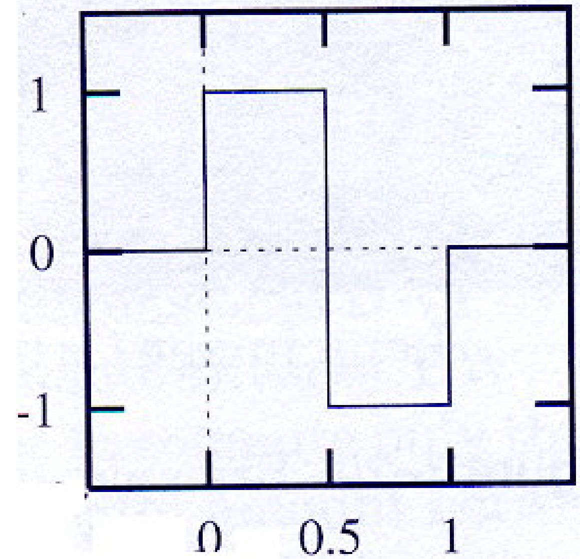
Table 3–5 The MPEG-1 Constrained Parameter Set

	Simple Profile (no B frames, not scalable)	Main Profile (B frames, not scalable)	SNR Scalable Profile (B frames, SNR scalable)	Spatially Scalable Profile (B frames, spatial or SNR scalable)	High Profile (B frames, spatial or SNR scalable)
High level (1920×1152 ×60)		≤80 Mbps			≤100 Mbps
High-1440 level (1440×1152 ×60)		≤60 Mbps		≤60 Mbps	≤80 Mbps
Main level (720×576×30)	≤15 Mbps	≤15 Mbps	≤15 Mbps		≤20 Mbps
Low level (352×288×30)		≤4 Mbps	≤4 Mbps		

Table 3–6 MPEG-2 Video Profiles and Levels with the Most-Important Characteristics (Cells in the table without entries are not defined)



Mexican hat wavelet



Haar wavelet

Figure 3–19 Two Mother Wavelets

Lena image (1:16)

Lena image (1:256)

Figure 3–20 Examples of Artifacts When Using DWT-based Coding Methods

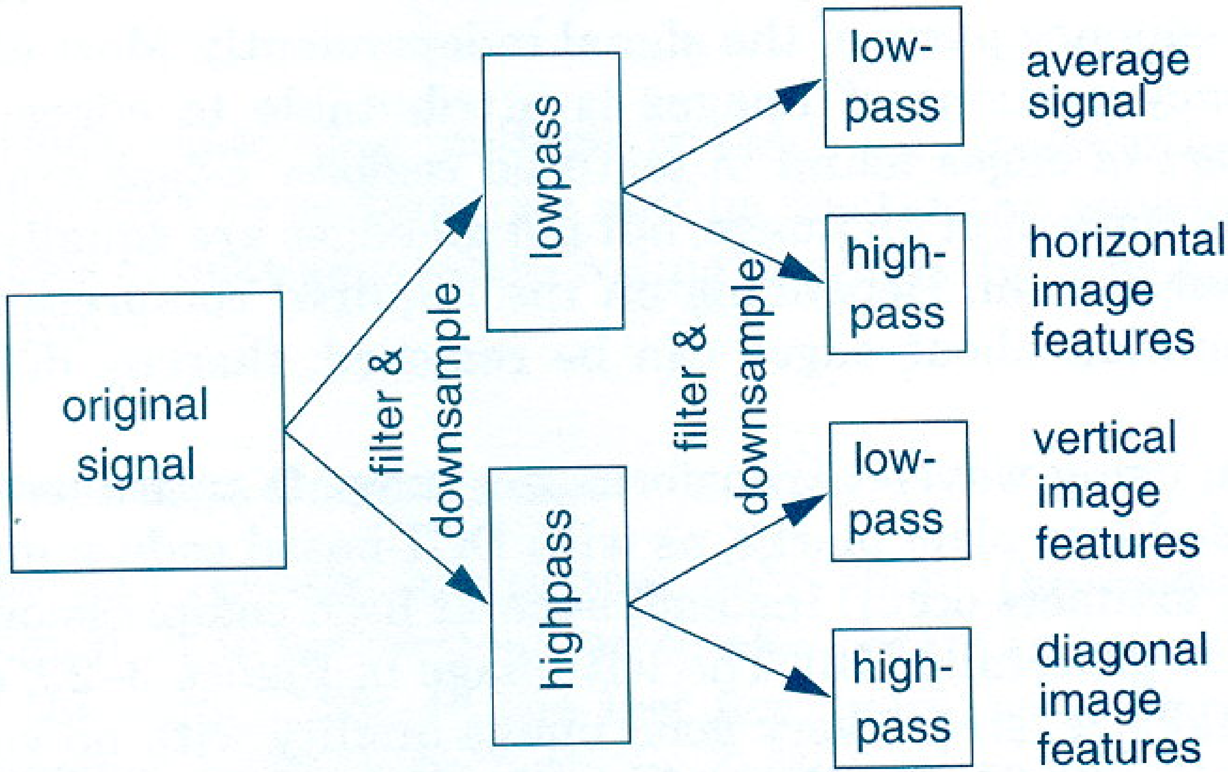


Figure 3–21 Block Diagram of Two-Dimensional Forward Wavelet Transform.

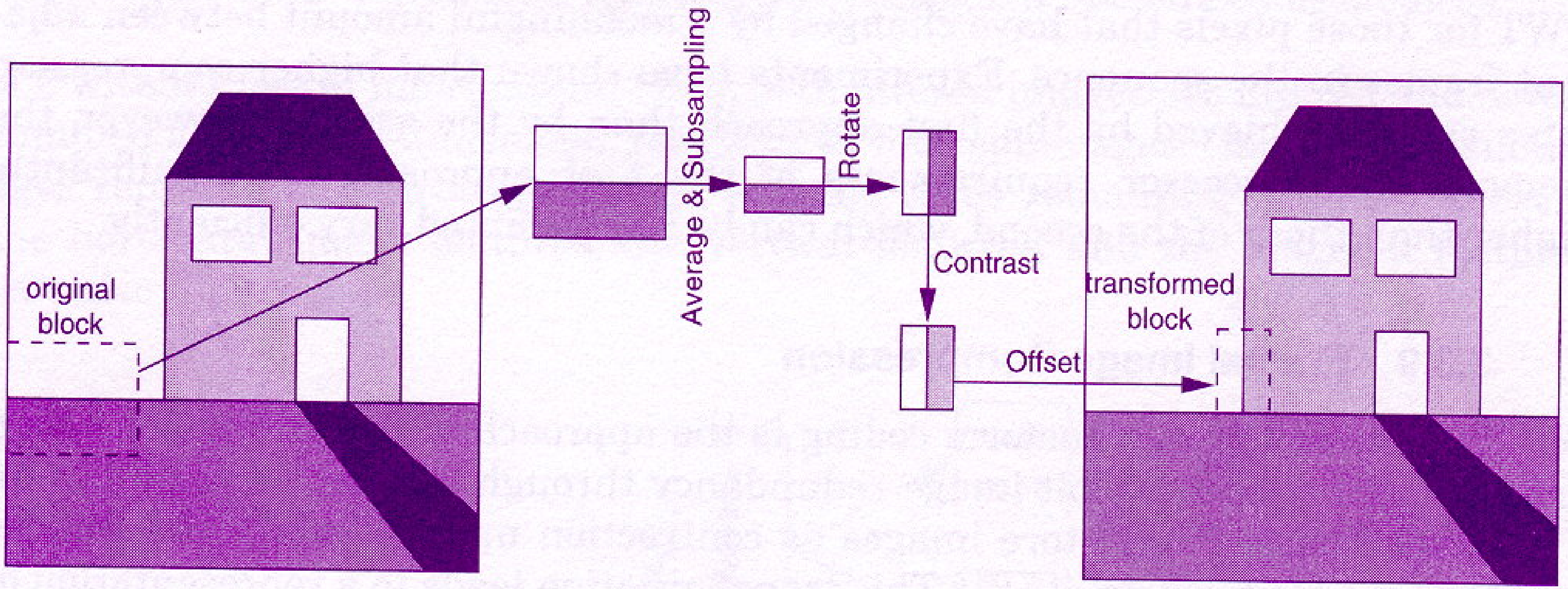


Figure 3–22 Transformation of an Image Block

Lena image (1:16)

Lena image (1:256)

Figure 3–23 Examples of Artifacts When Using a Fractal Image Coder

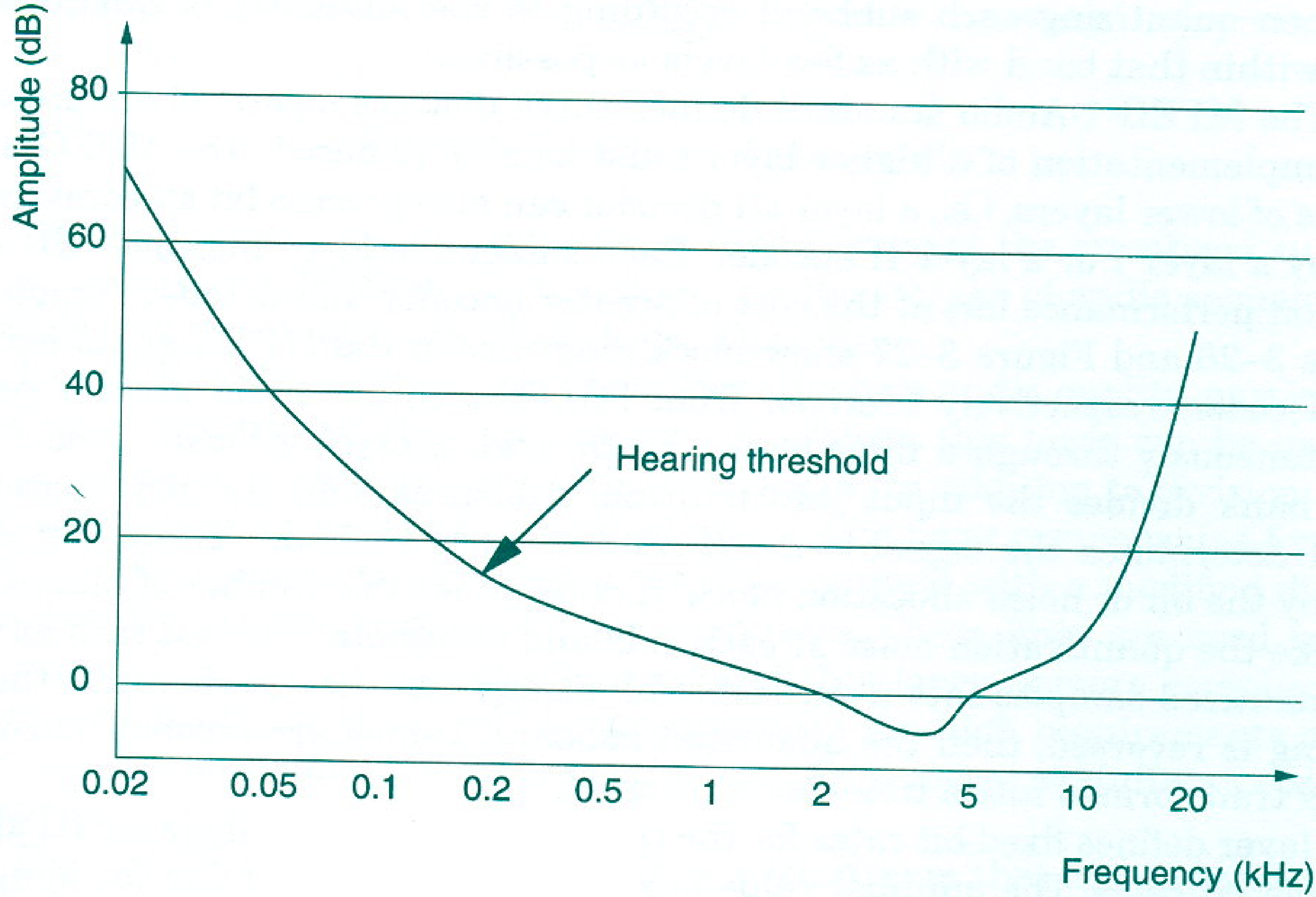


Figure 3–24 Hearing Threshold of the Human Ear

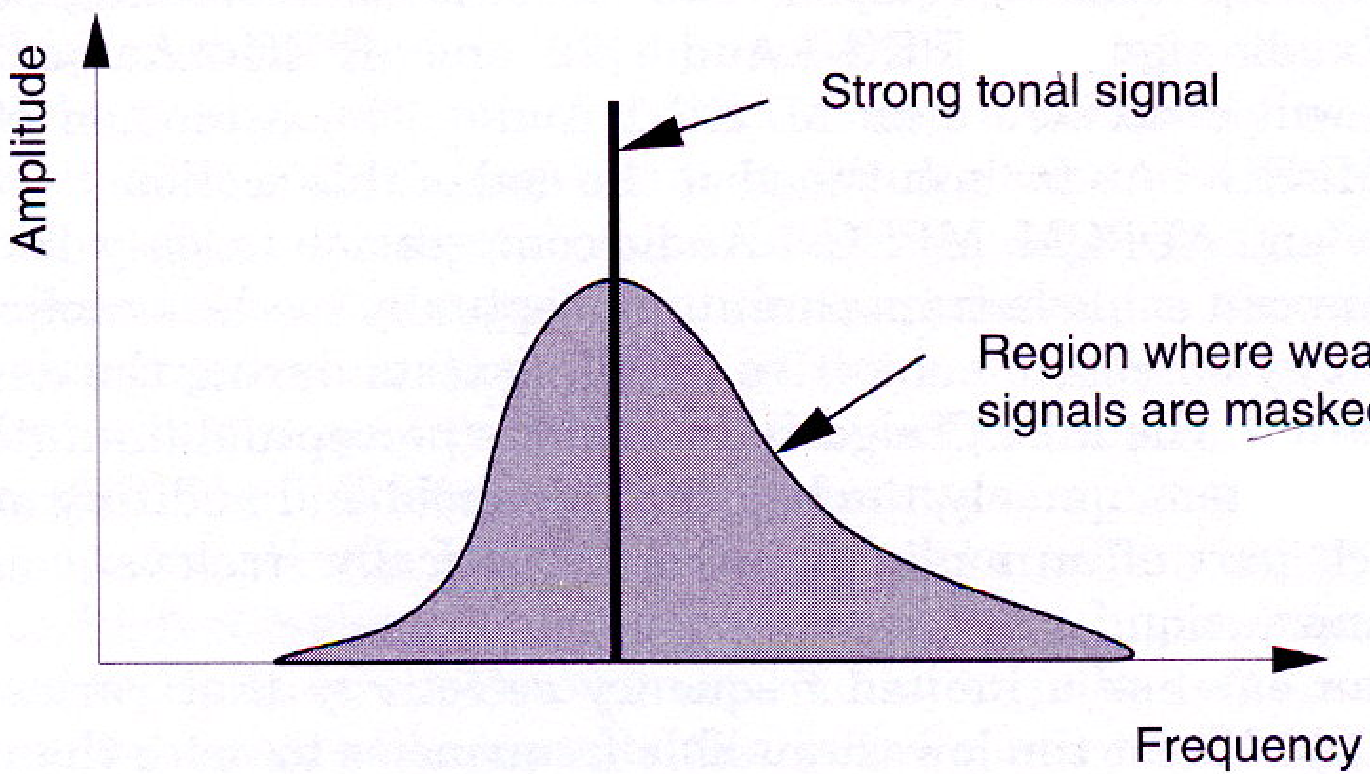


Figure 3–25 Audio Noise-Masking Property of the Human Auditory System

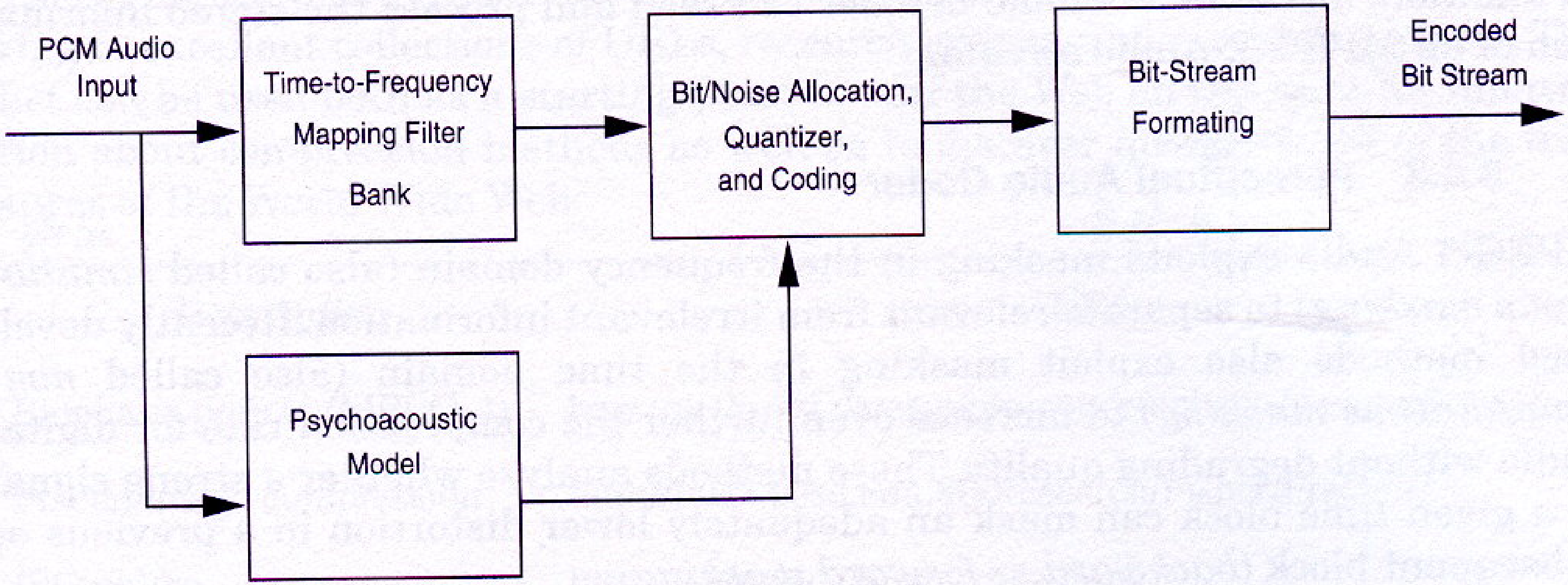


Figure 3–26 MPEG-1 Audio Encoder

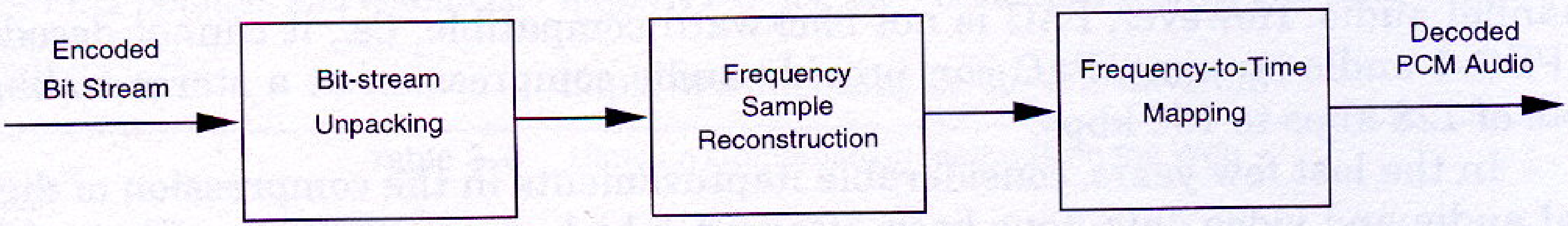


Figure 3–27 MPEG-1 Audio Decoder

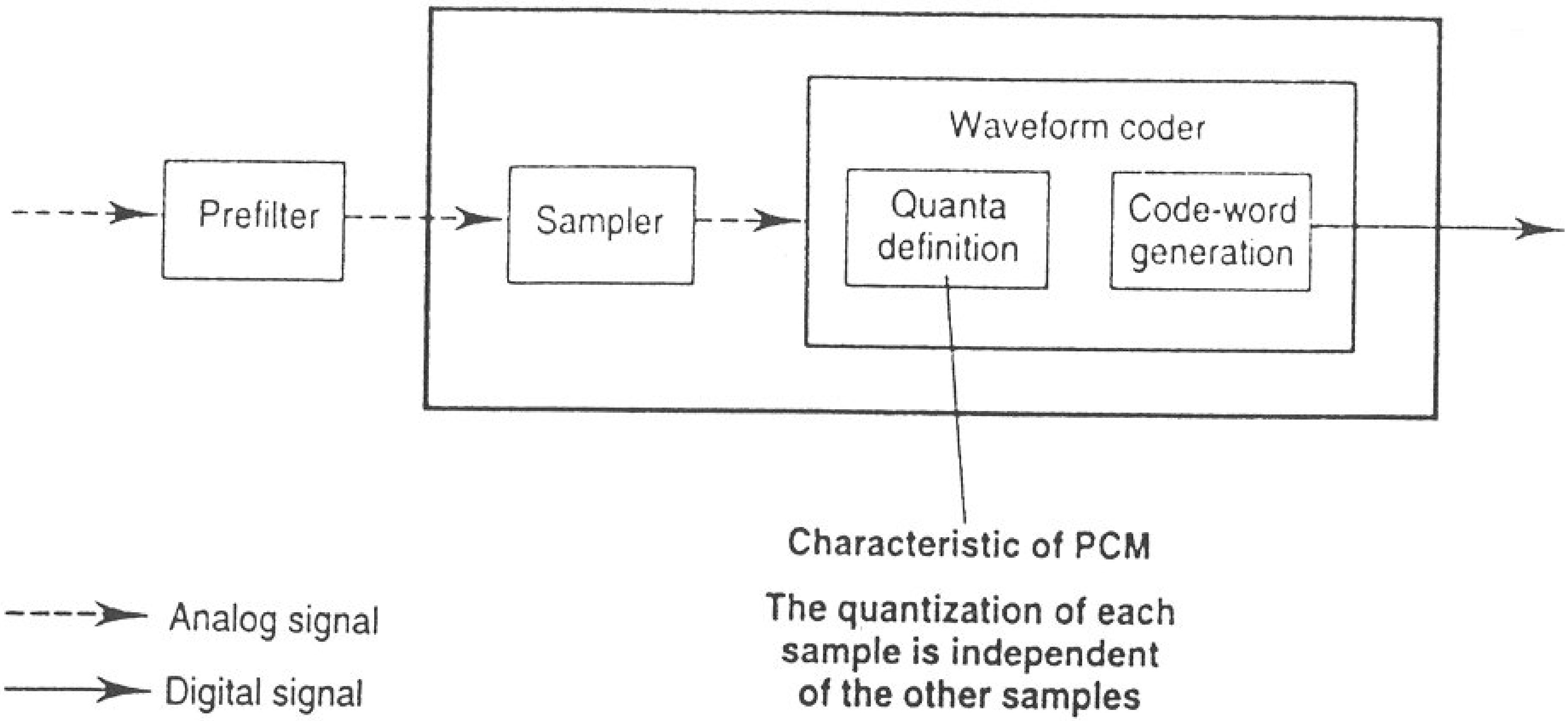


Figure 27.1 PCM analog-to-digital converter

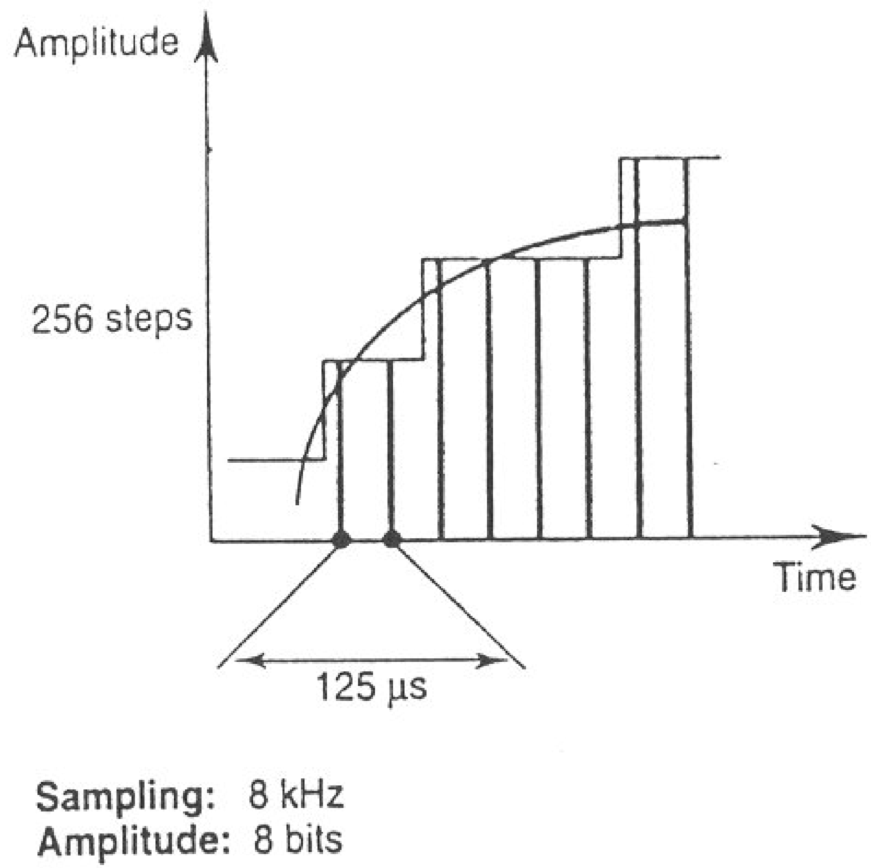


Figure 27.2 PCM sampling rate and amplitude depth according to ITU-TS G.711

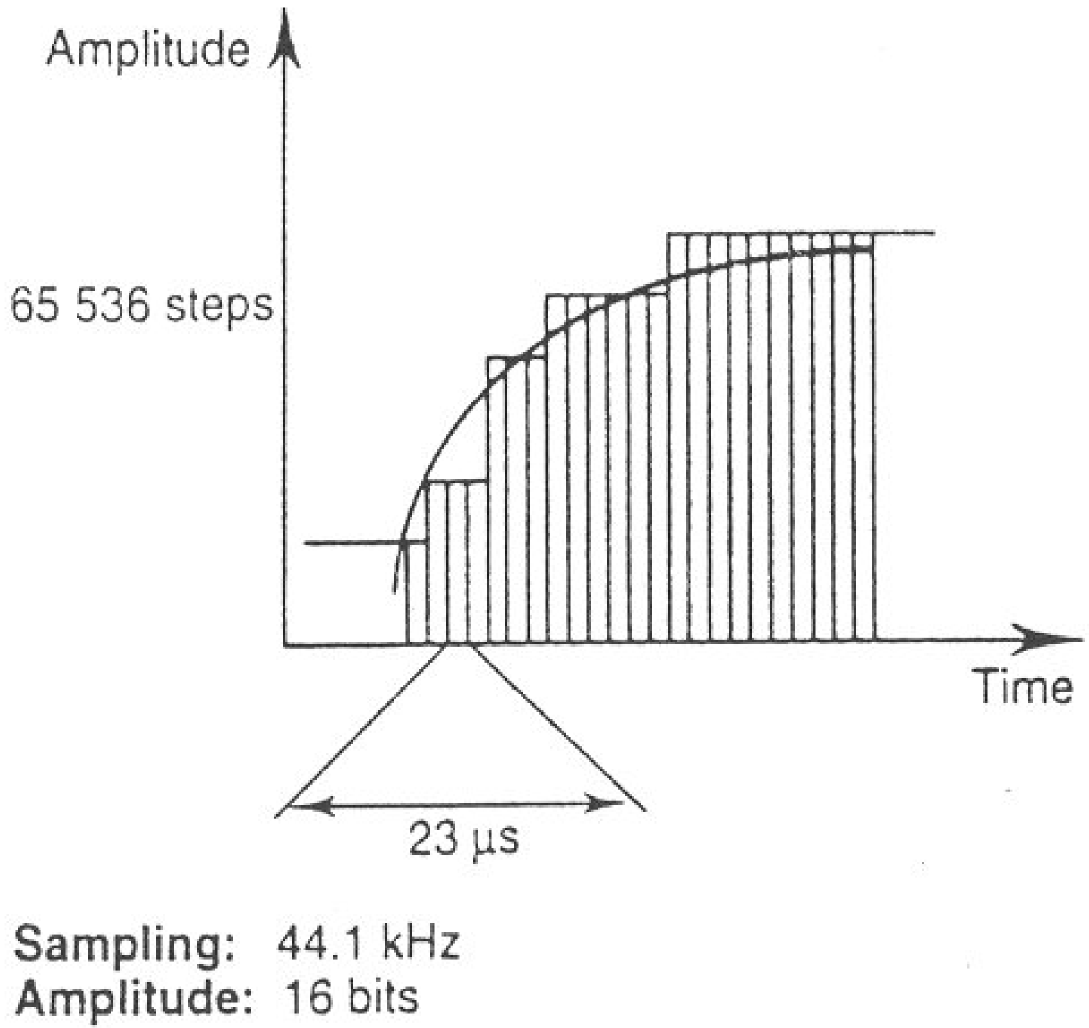
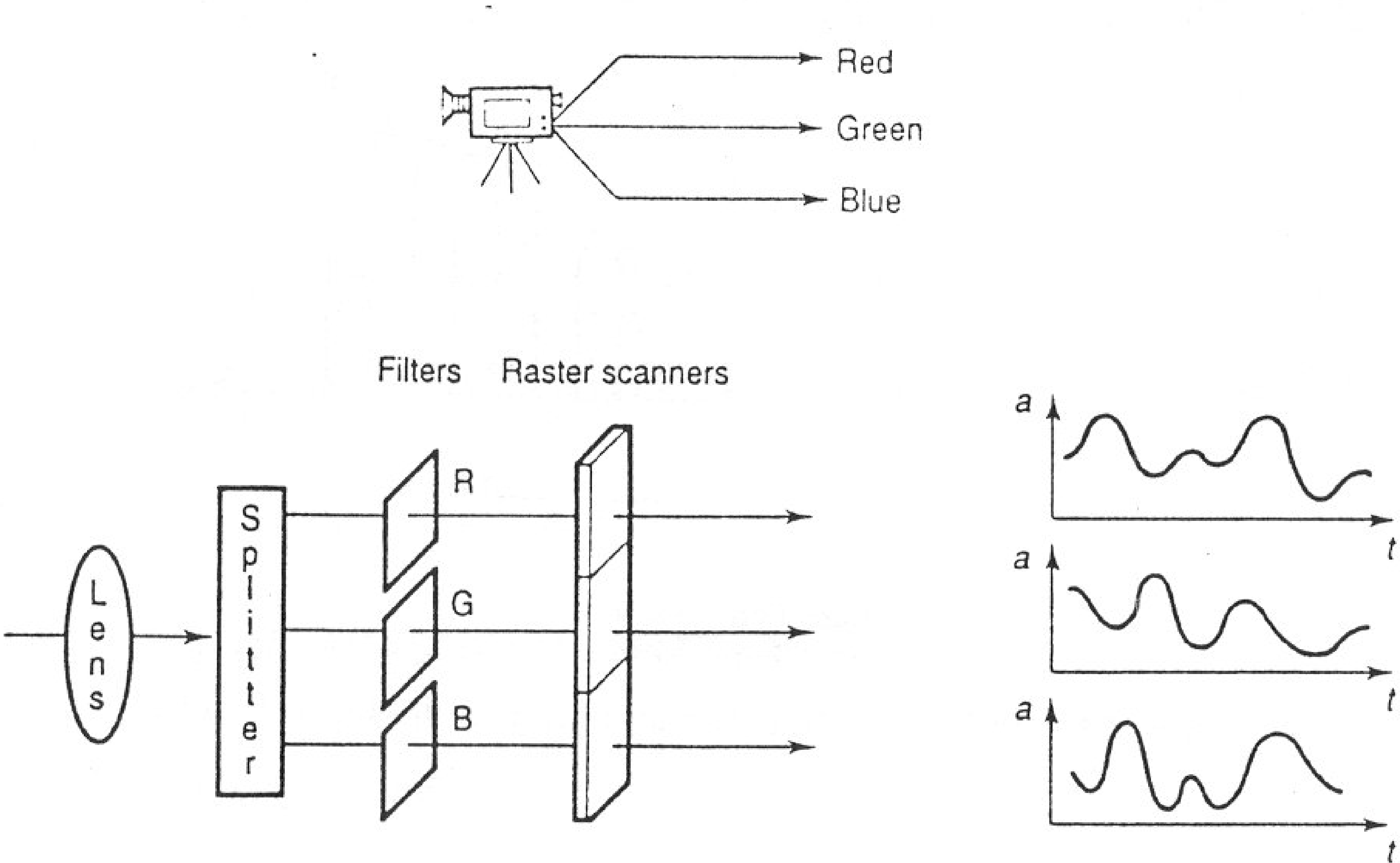
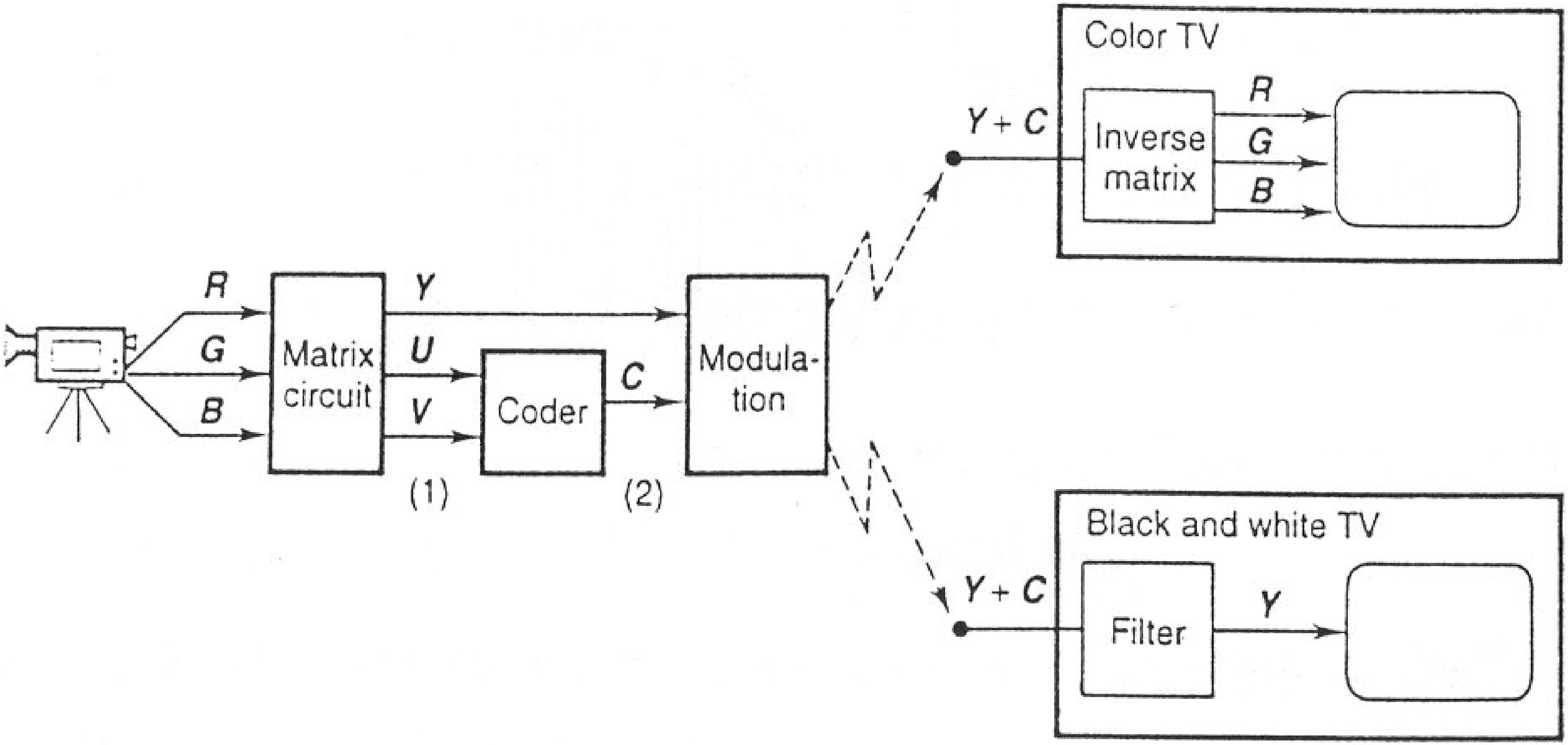


Figure 27.3 PCM sampling rate and amplitude depth for CD-audio



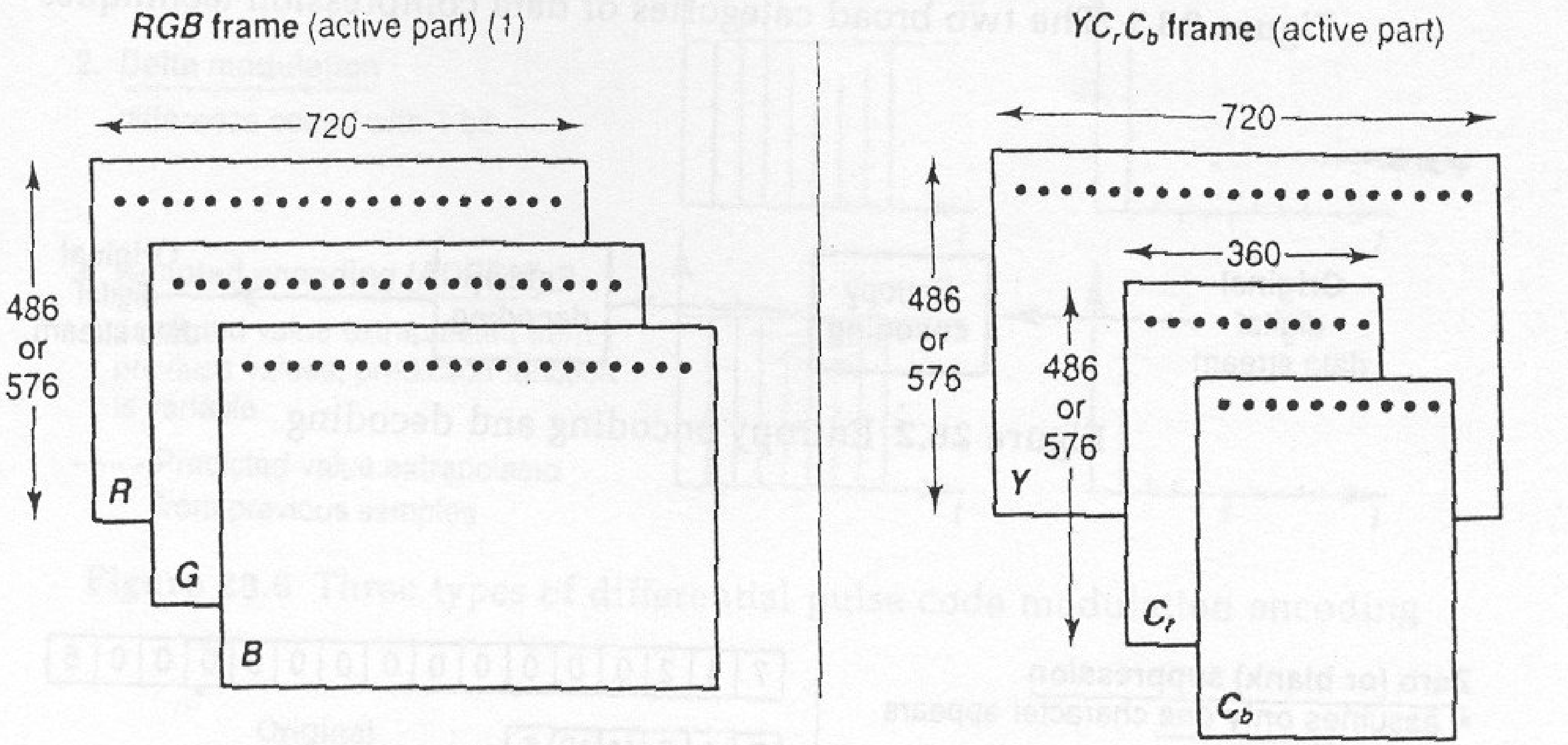
Simplified TV camera

Figure 27.5 TV cameras produce a tristimulus signal



(1) *YUV* are the luminance and chrominance signals in PAL.
YIQ are the corresponding signals in NTSC
(2) The *U* and *V* signals are combined into a chroma signal *C*

Figure 27.6 *RGB* to *YUV* conversion in broadcast TV



(1) In ITU-R 601 digital TV, the frame is sent in two successive fields each containing half the lines

Figure 27.7 Subsampling of the color difference signals in studio-quality TV (ITU-R 601 recommendation)

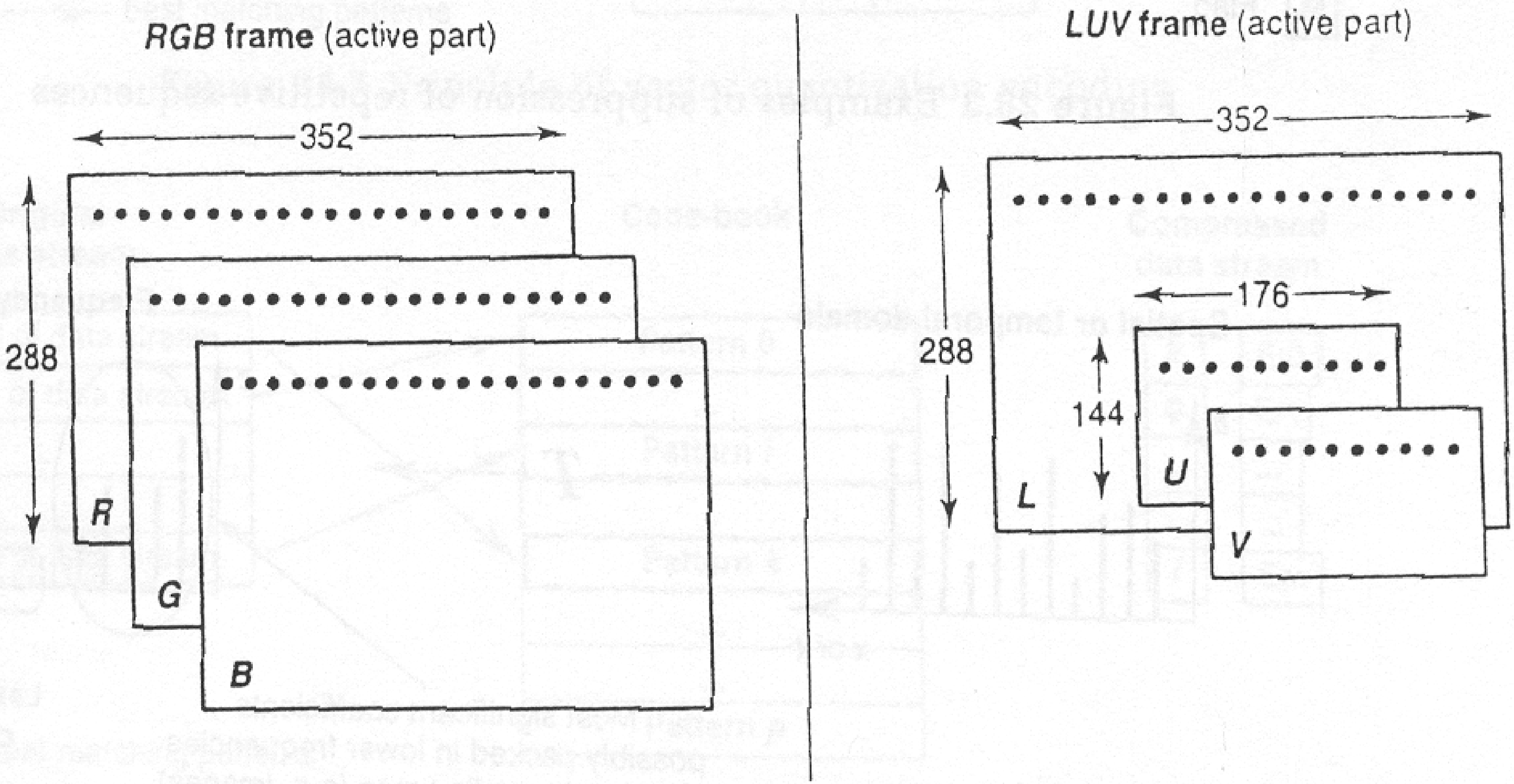


Figure 27.8 Subsampling of the color difference signals in videoconferencing (ITU-TS H.261 recommendation, when using CIF)

Entropy encoding	Independent of characteristics of of the data stream	Lossless compression
Source encoding	Takes into account the **semantics**	**Lossless or lossy compression**

Figure 28.1 The two broad categories of data compression techniques

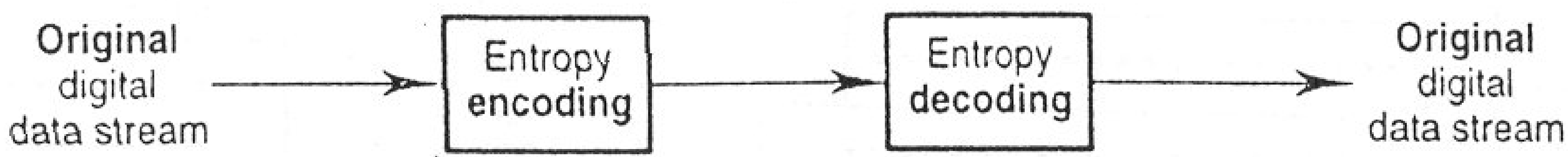


Figure 28.2 Entropy encoding and decoding

Zero (or blank) suppression • assumes **only one** character appears frequently	7 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 7 | 4 | 2 | **M** | 12 | 5
Run-length encoding • four or more successive occurrences of **any** character replaced by • the character + flag (M) + number of occurrences	N | a | m | e | : | . | . | . | . | . | . | . | . | . | . | CR N | a | m | e | : | . | **M** | 10 | CR

M Flag

Figure 28.3 Examples of suppression of repetitive sequences

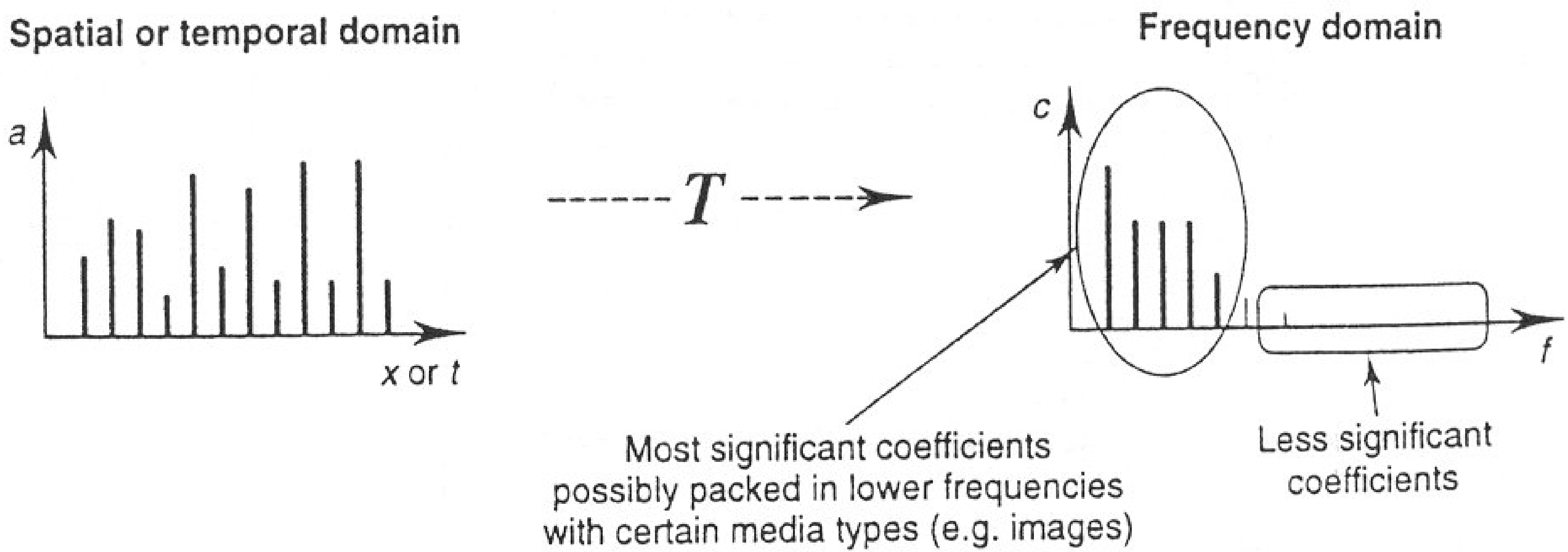


Figure 28.4 Principle of transform encoding

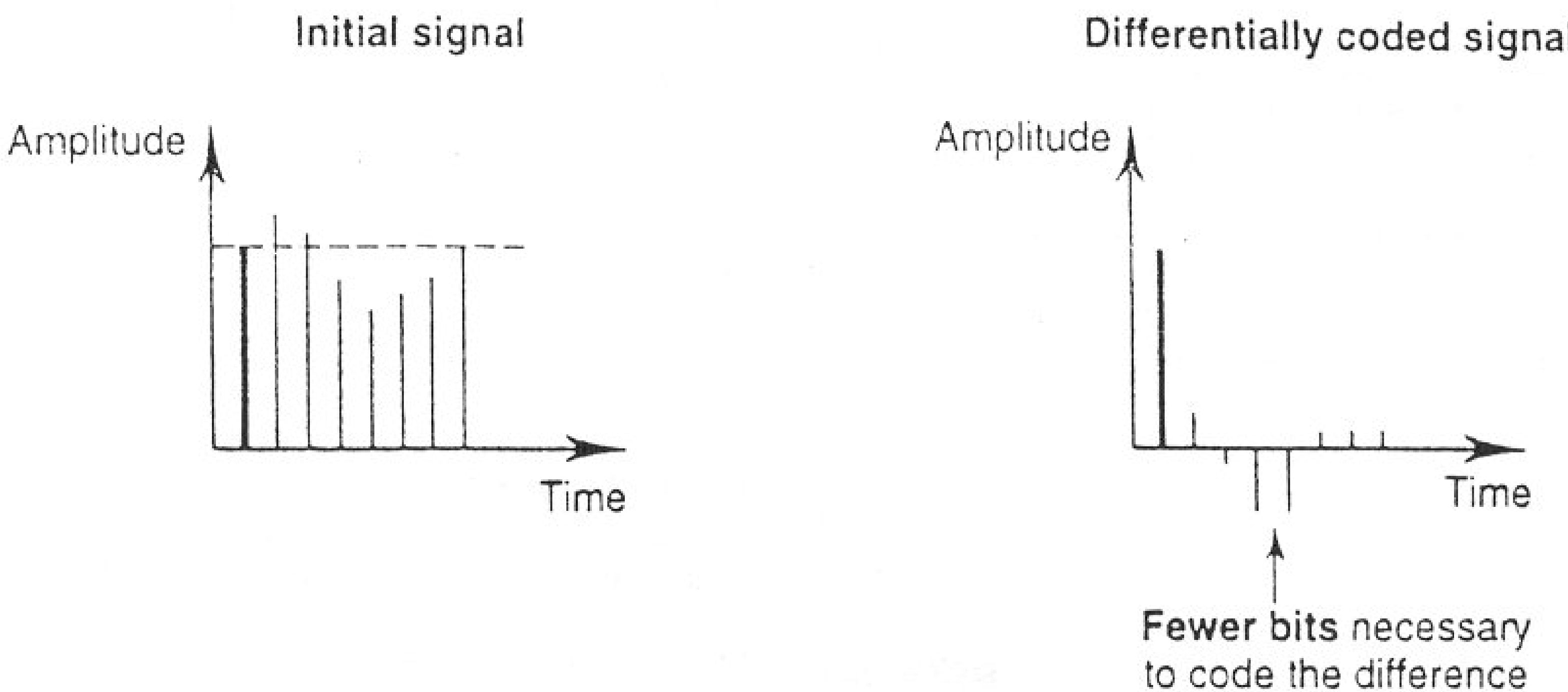


Figure 28.5 Principle of differential pulse code modulation encoding

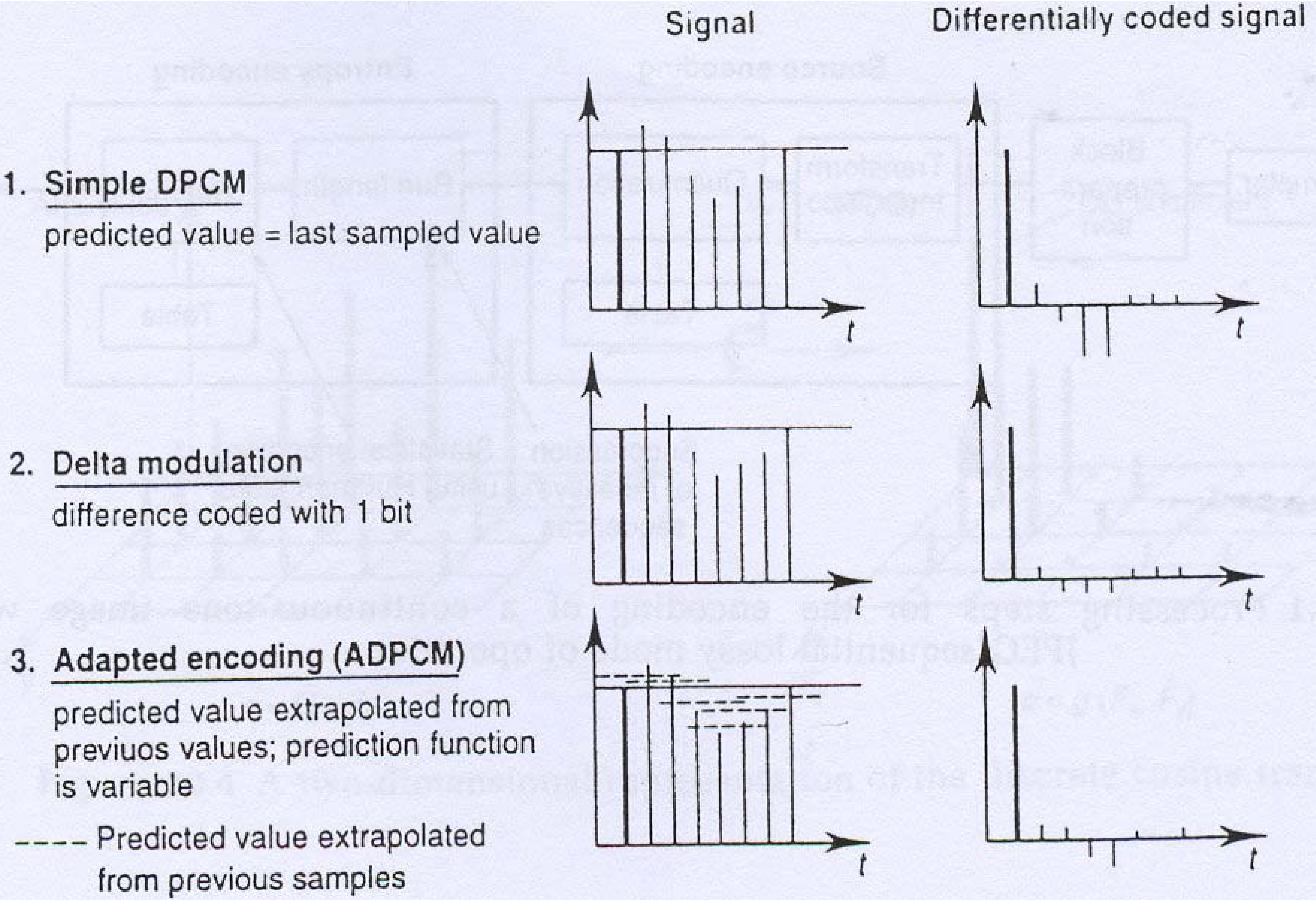


Figure 28.6 Three types of differential pulse code modulation encoding

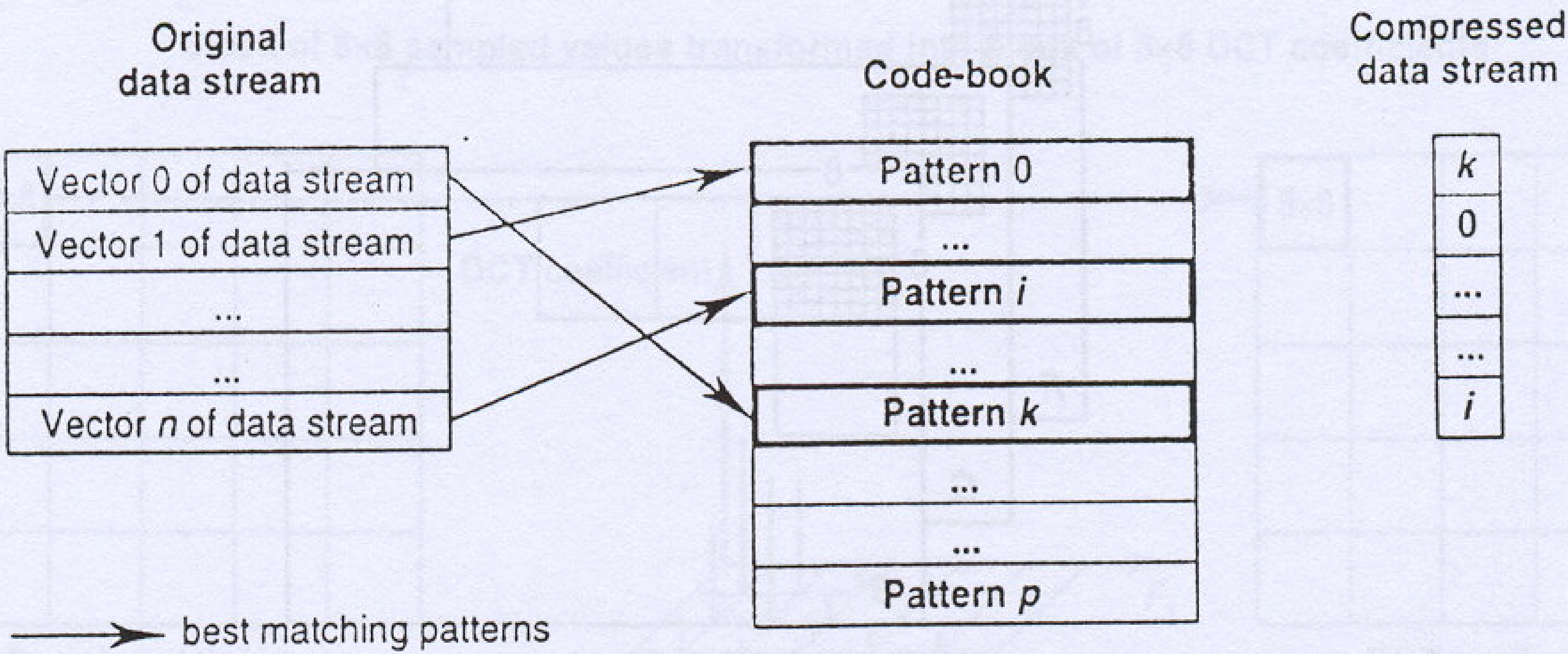


Figure 28.7 Principle of vector quantization encoding

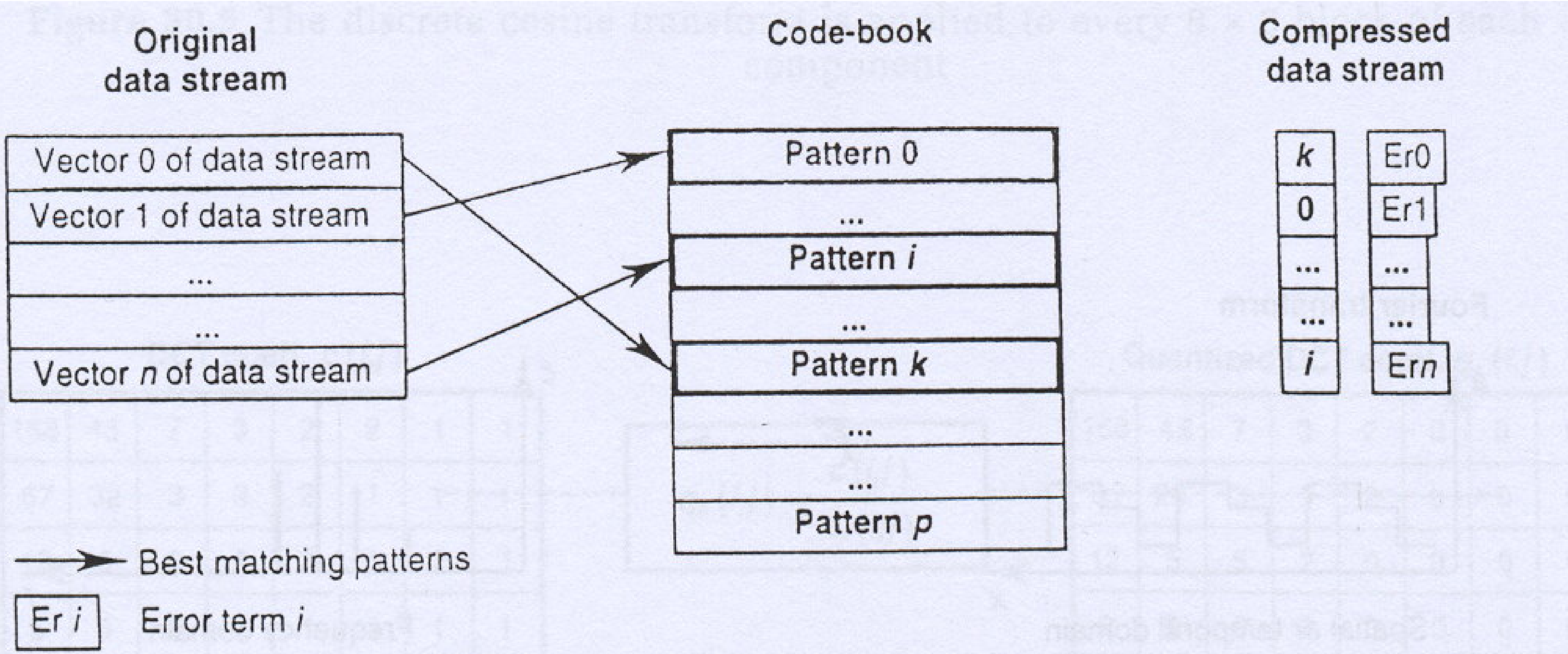


Figure 28.8 Principle of vector quantization encoding with transmission of the error term

Entropy coding	Repetitive sequence suppression	• Zero suppression • Run-length encoding
	Statistical encoding	• Pattern substitution • Huffman-like encoding
Source coding	Transform encoding	• FFT • DCT • Others
	Differential encoding	• DPCM • Delta modulation • ADPCM
	Vector quantization	

Figure 28.9 The major data compression techniques

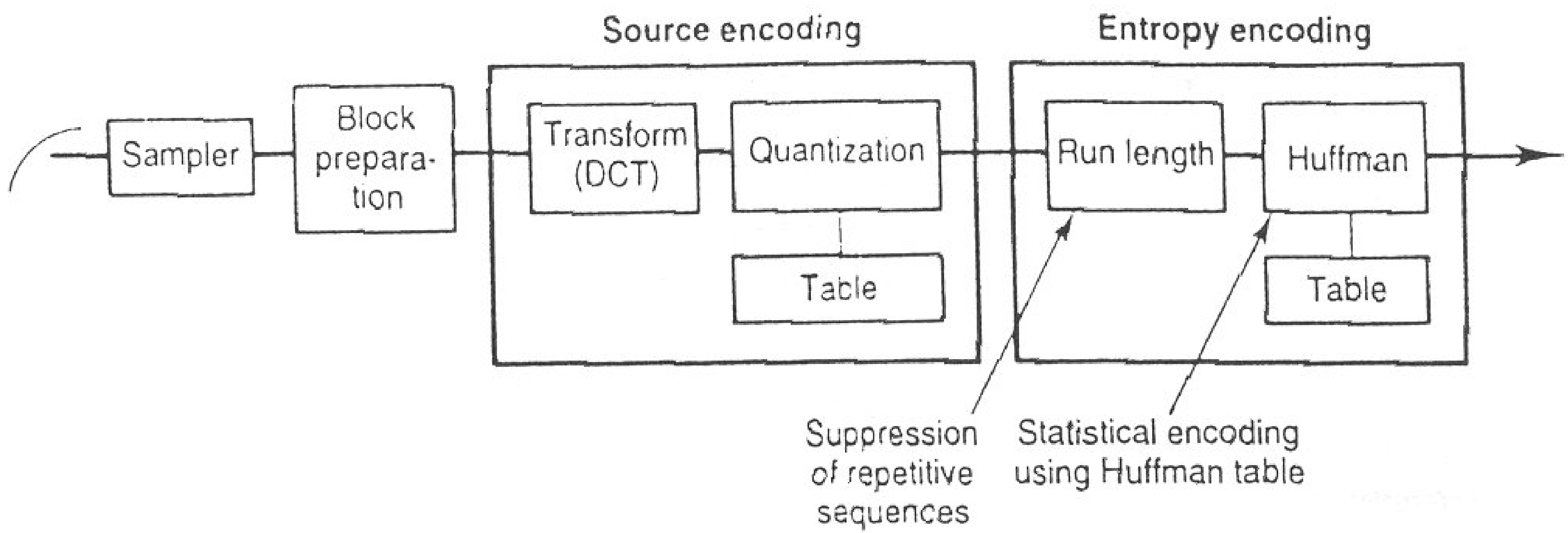


Figure 30.1 Processing steps for the encoding of a continuous-tone image with JPEG/sequential lossy mode of operation

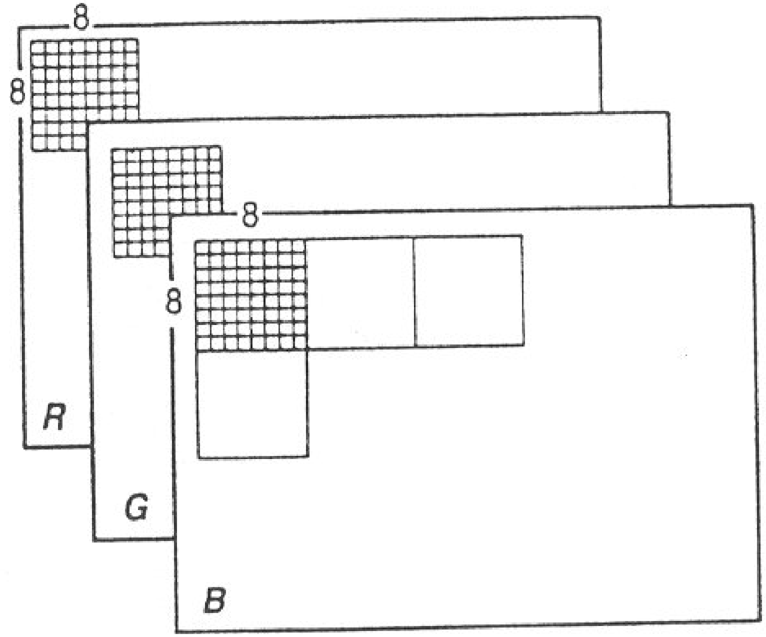


Figure 30.2 Example of block preparation for JPEG. Assumes the image to be compressed is represented by the three *R*, *G*, *B* components. Each component is divided into 8 × 8 pixel blocks

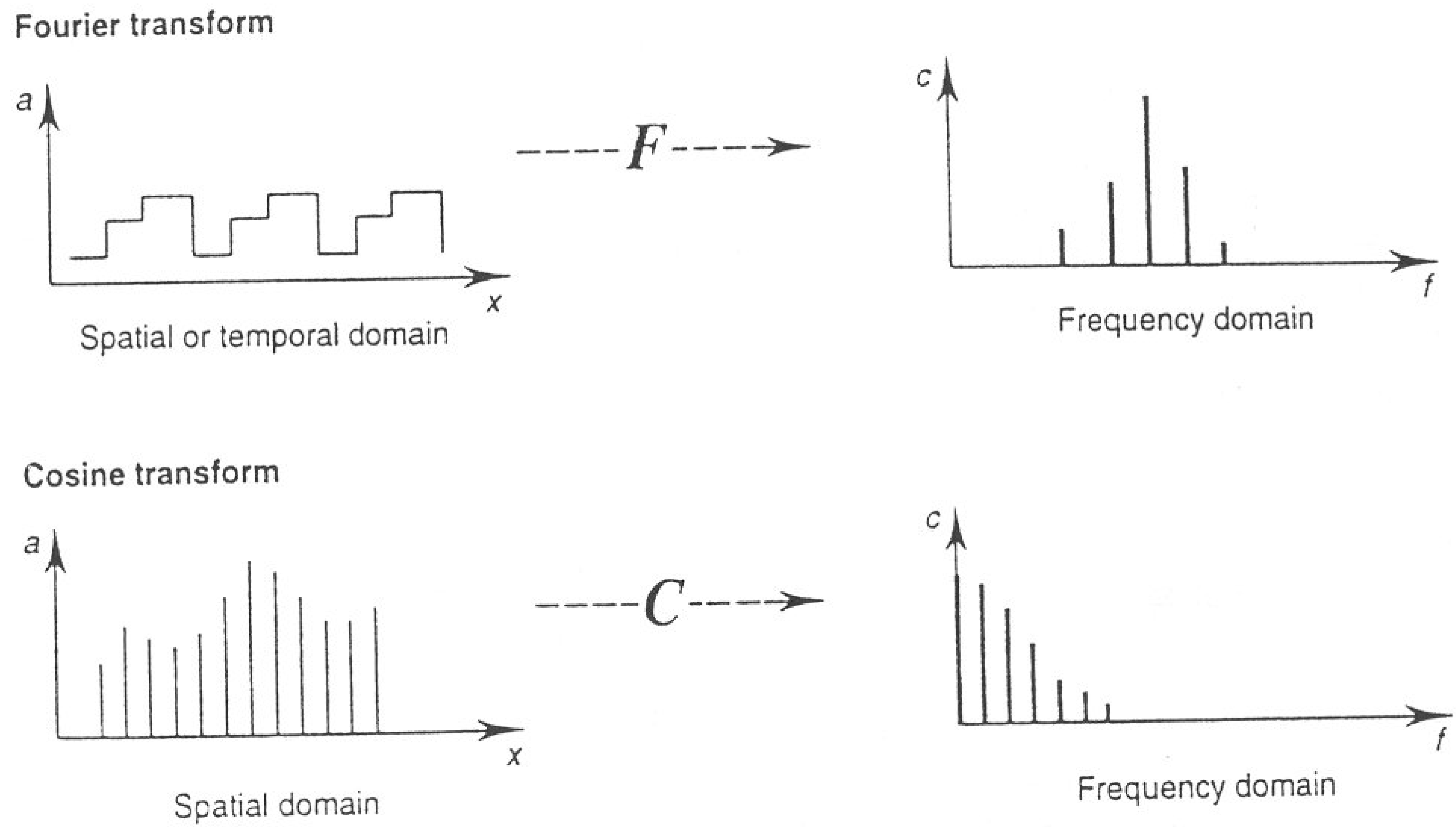


Figure 30.3 The principles of the Fourier transform and discrete cosine transform

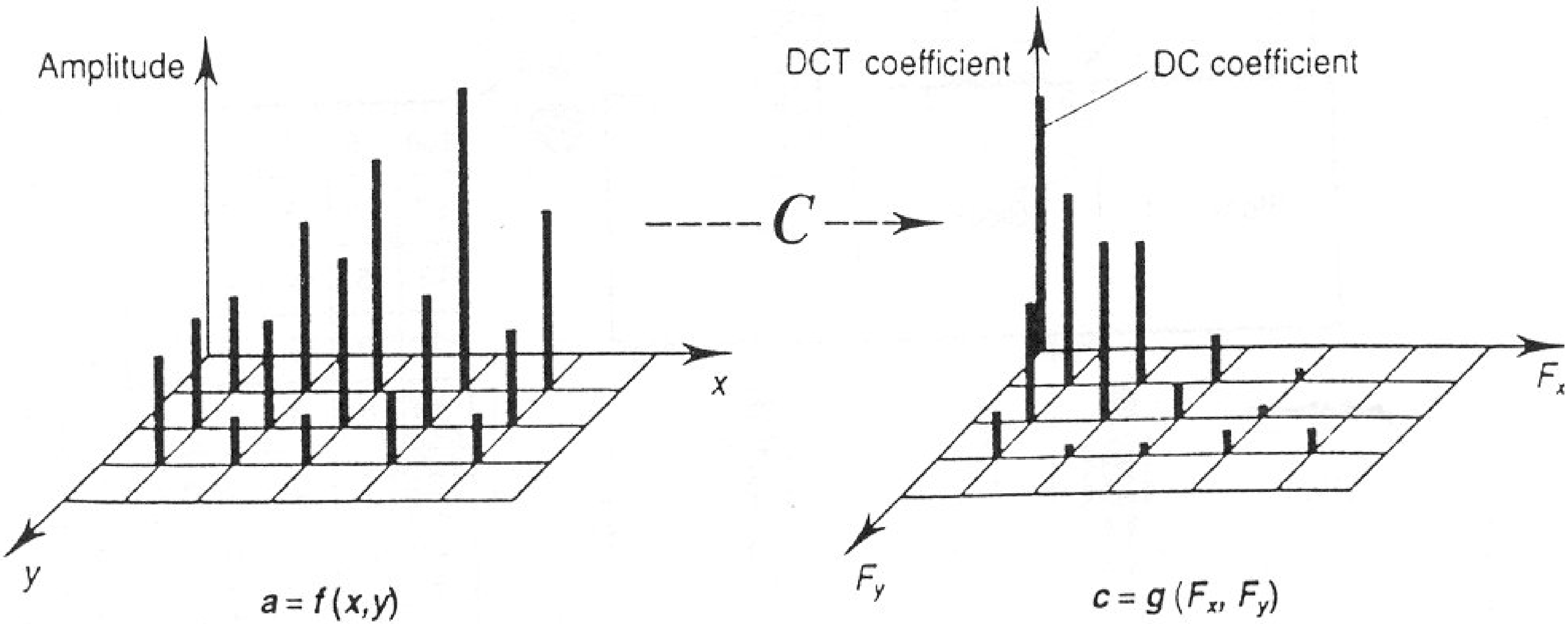


Figure 30.4 A two-dimensional representation of the discrete cosine transform

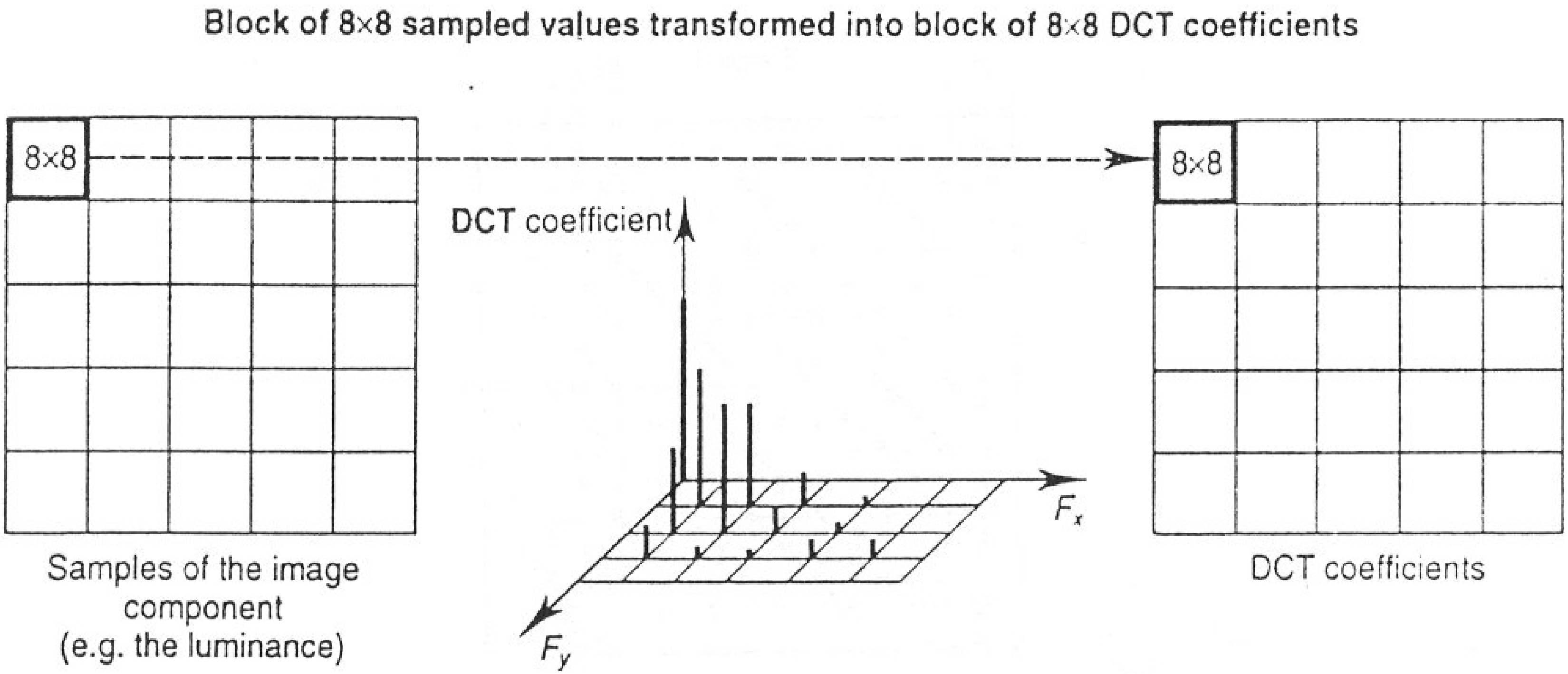


Figure 30.5 The discrete cosine transform is applied to every 8 × 8 block of each image component

DCT coeff. $c(i,j)$

168	45	7	3	2	2	1	1
67	32	3	3	2	1	1	1
12	5	5	5	2	1	2	1
5	5	2	2	1	1	1	1
3	2	2	1	1	1	1	1
2	2	2	1	1	1	1	0
1	1	1	1	2	1	1	0
1	2	1	1	1	1	0	0

$$q_c(i,j) = \frac{c(i,j)}{Q(i,j)}$$

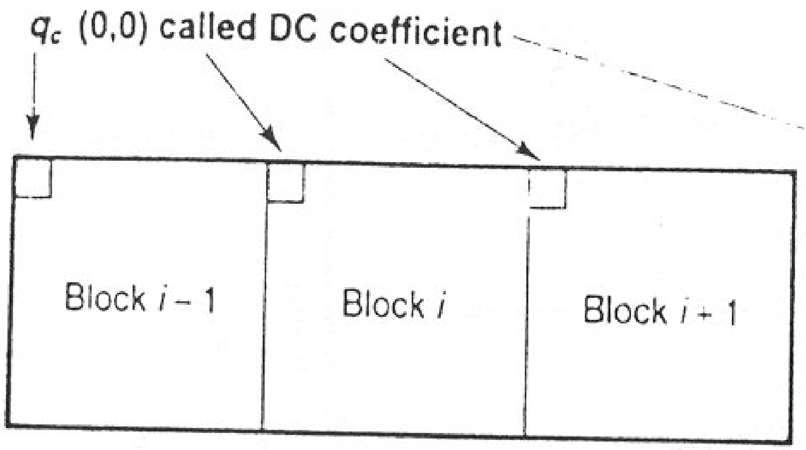
Quantized DCT coeff. $q_c(i,j)$

168	45	7	3	2	0	0	0
67	20	3	3	2	0	0	0
12	5	5	2	0	0	0	0
2	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

Quantization table $Q(i,j)$

1	1	1	1	1	4	8	16
1	1	1	1	4	4	8	16
1	1	1	2	4	4	8	16
2	8	8	8	8	16	16	16
4	8	8	8	8	16	16	32
4	8	8	8	16	16	16	32
4	8	8	8	16	16	32	32
8	8	8	16	16	32	32	64

Figure 30.6 An example of quantization



Block i

168	45	7	3	2	0	0	0
67	20	3	3	2	0	0	0
12	5	5	2	0	0	0	0
2	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

Figure 30.7 DC coefficients. They measure the average value of the 64 image samples

Block *i*

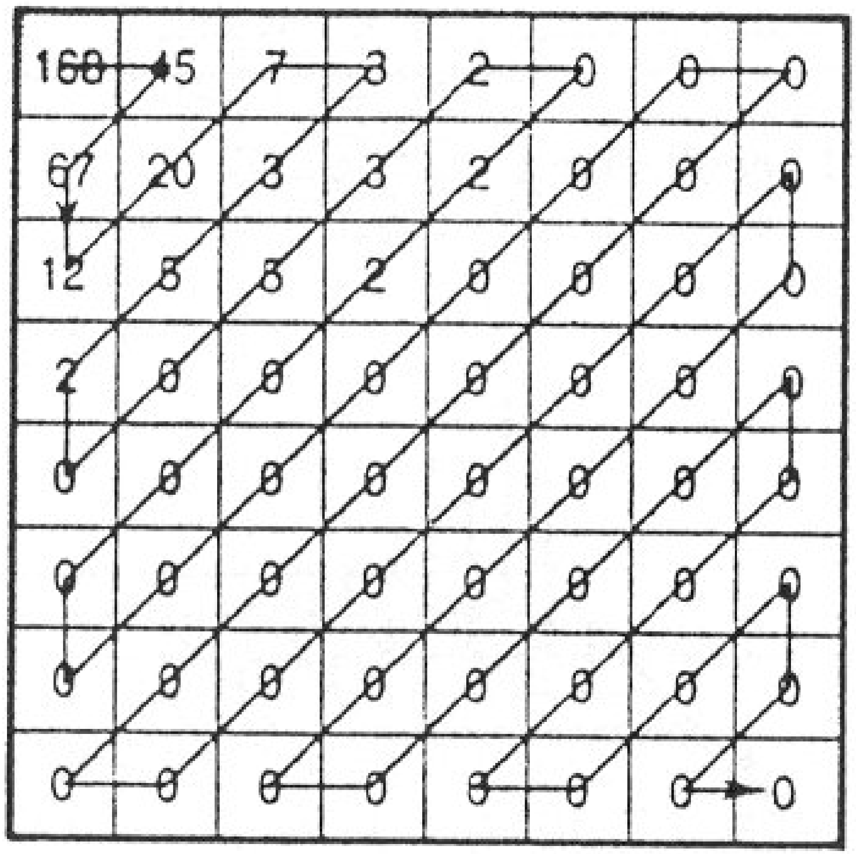


Figure 30.8 The zig-zag sequence

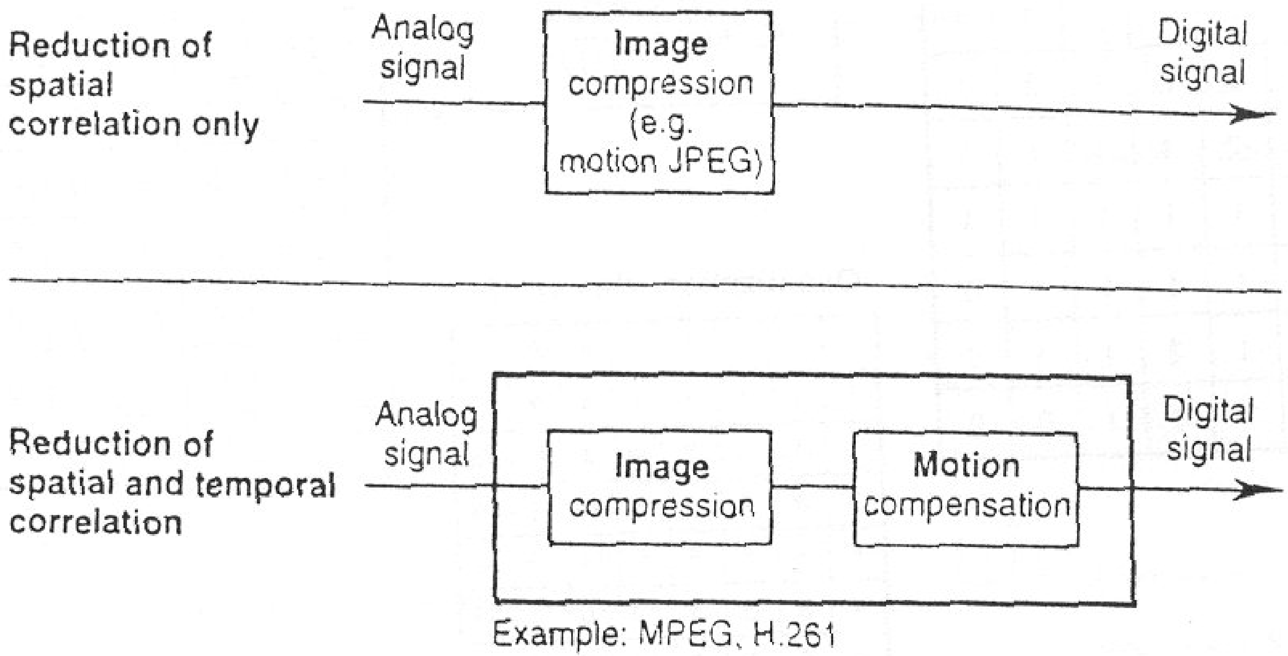


Figure 30.9 Practical implementation of motion video compression available on workstations

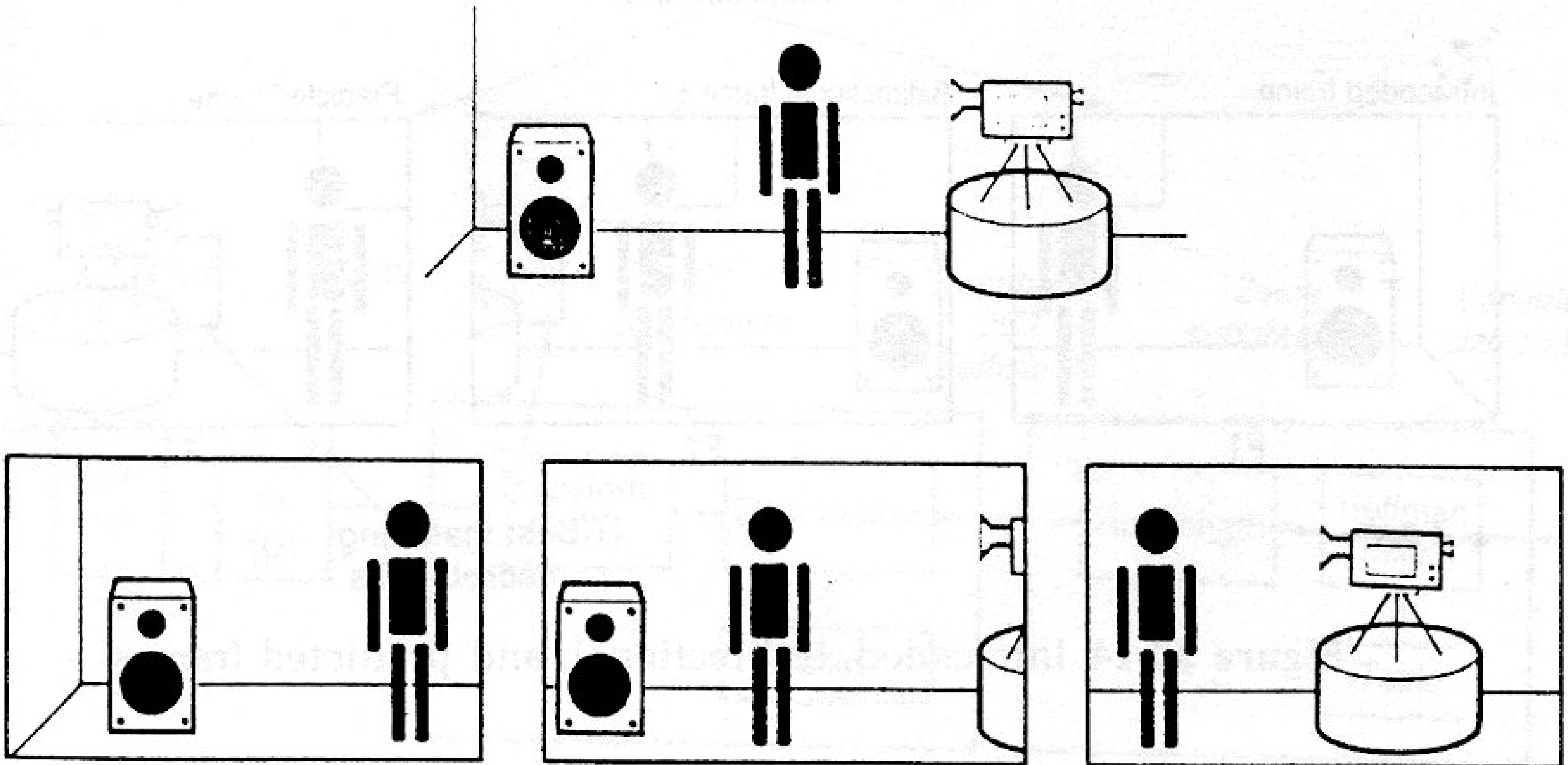


Figure 30.10 Three successive frames from a real-world shot

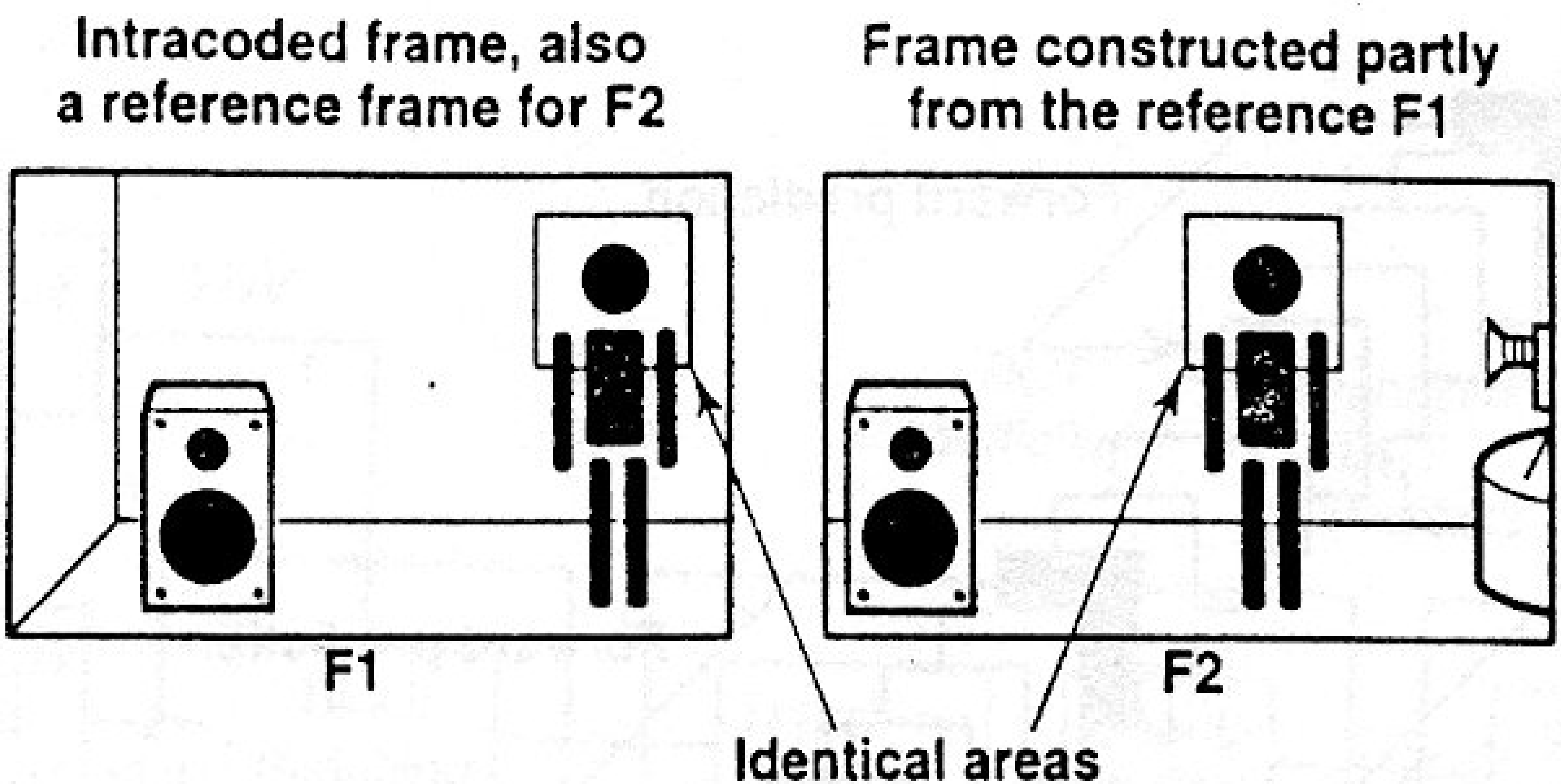


Figure 30.11 Reference frame and intracoded frame

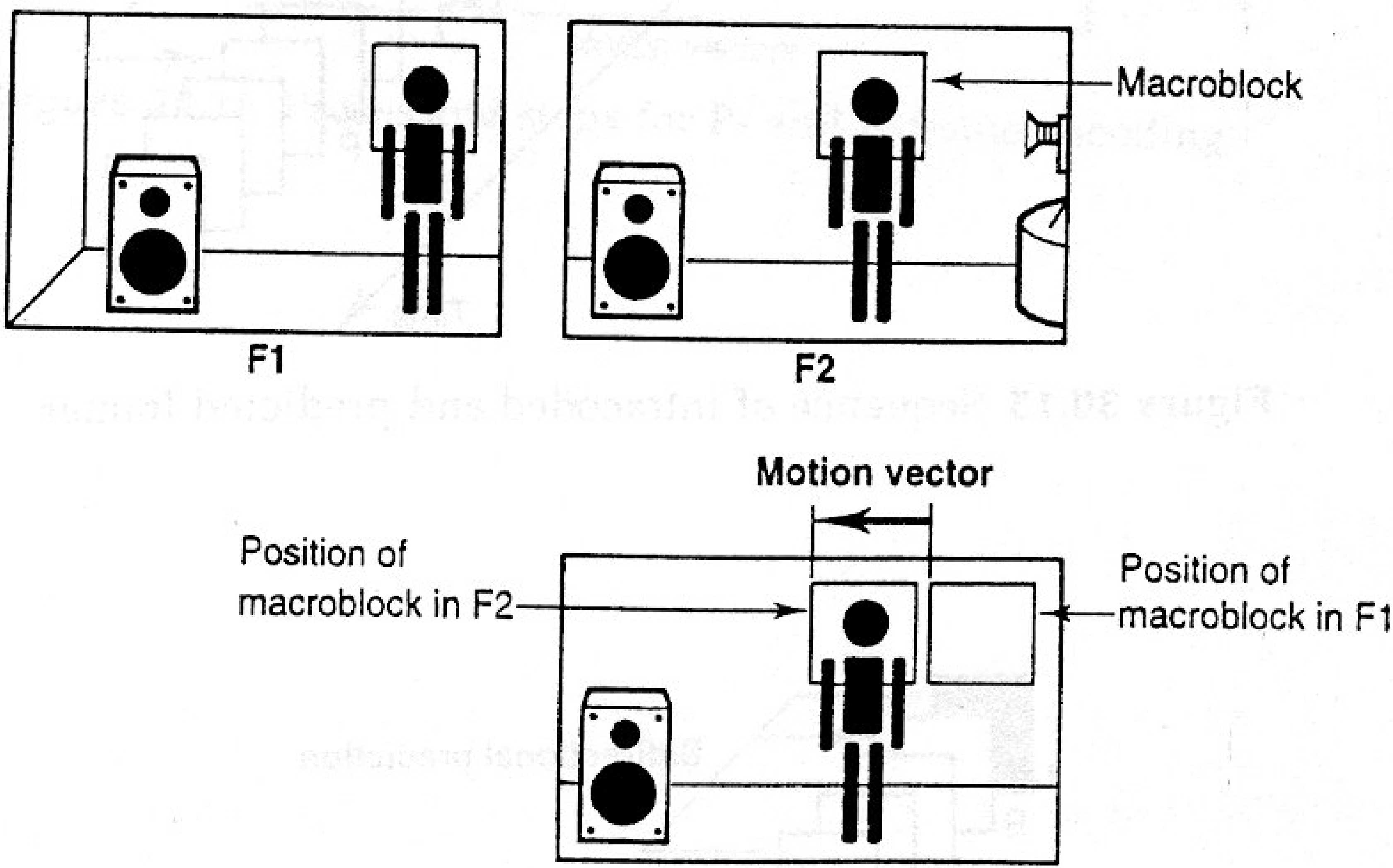


Figure 30.12 Motion vector and macroblocks

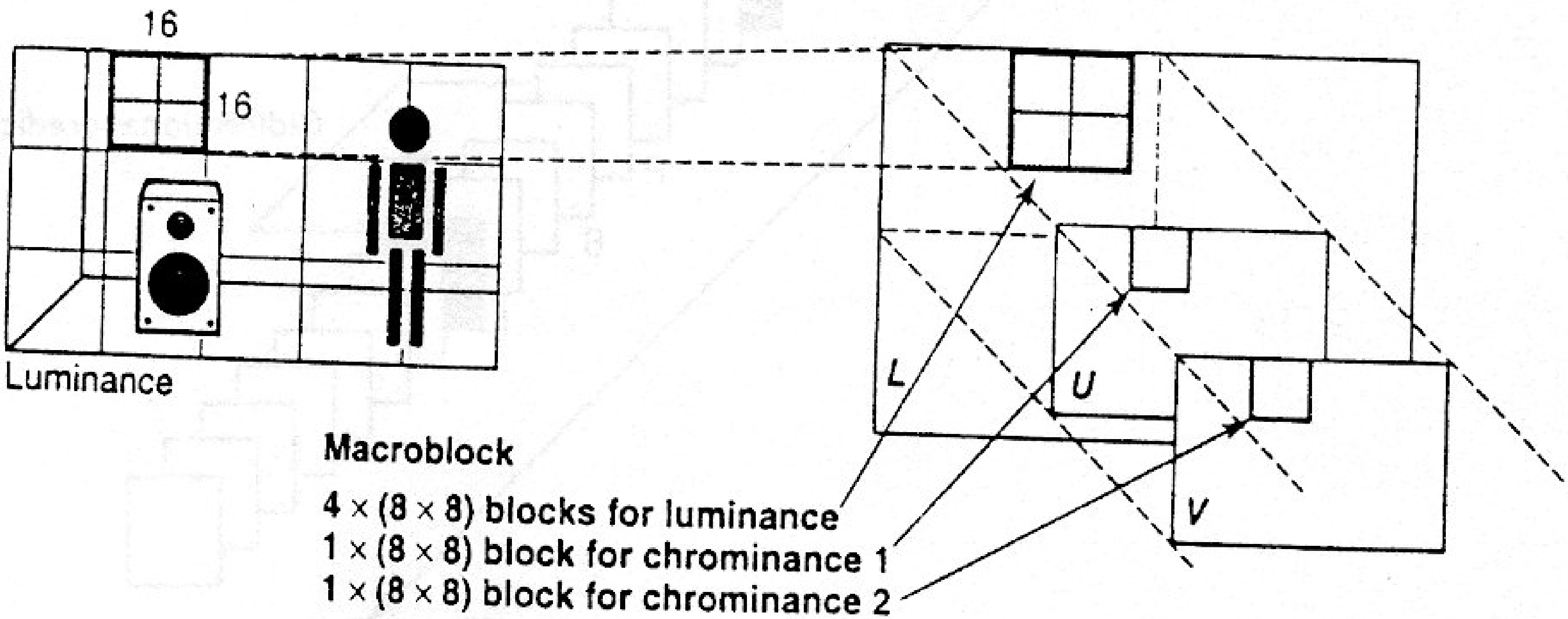


Figure 30.13 Macroblocks in MPEG-1

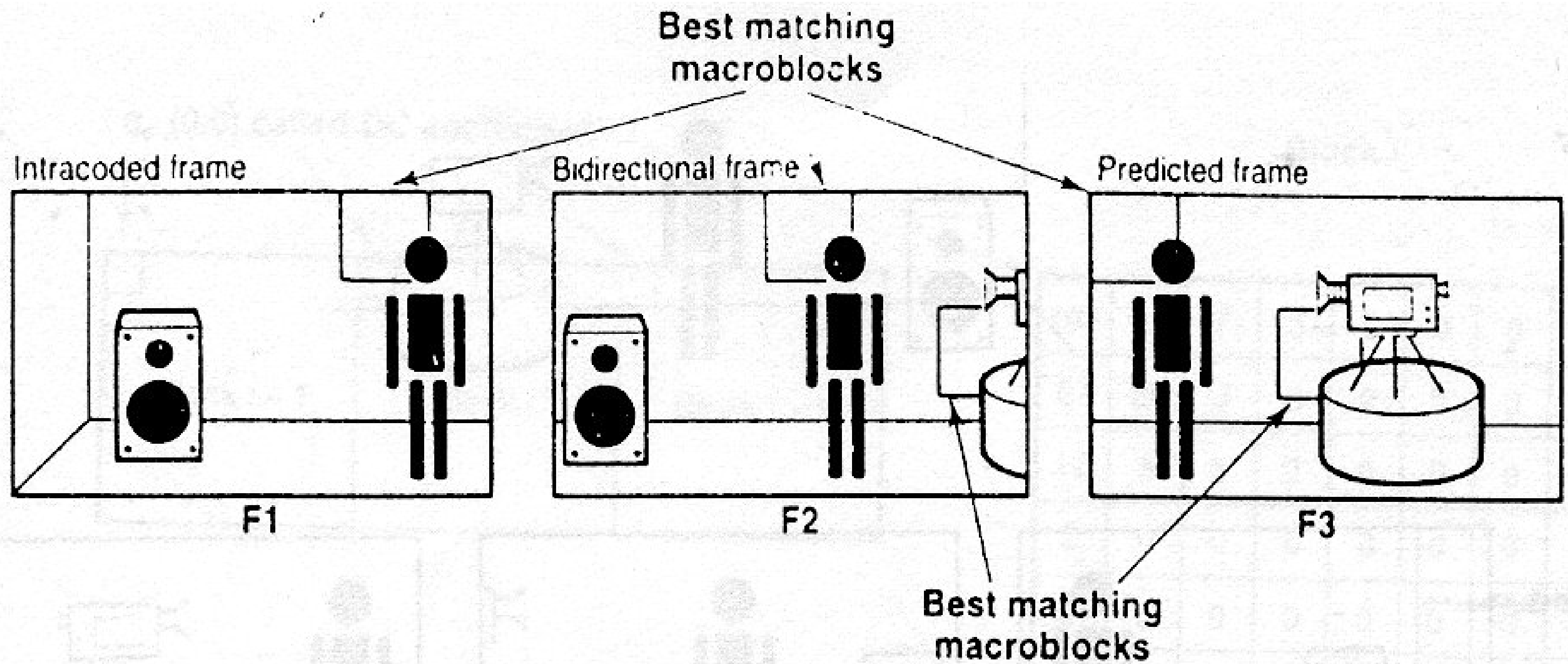


Figure 30.14 Intracoded, bidirectional, and predicted frames

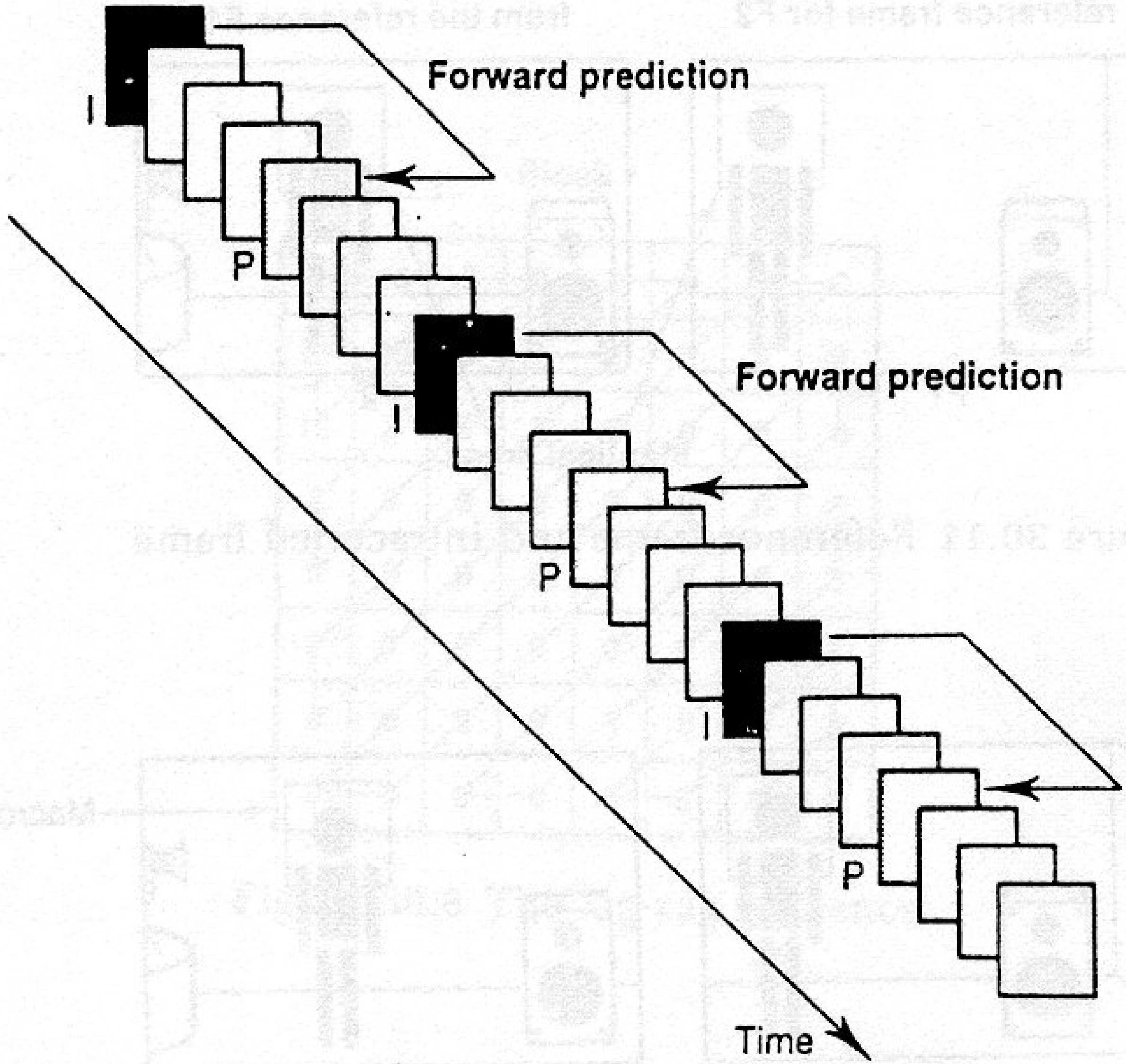


Figure 30.15 Sequence of intracoded and predicted frames

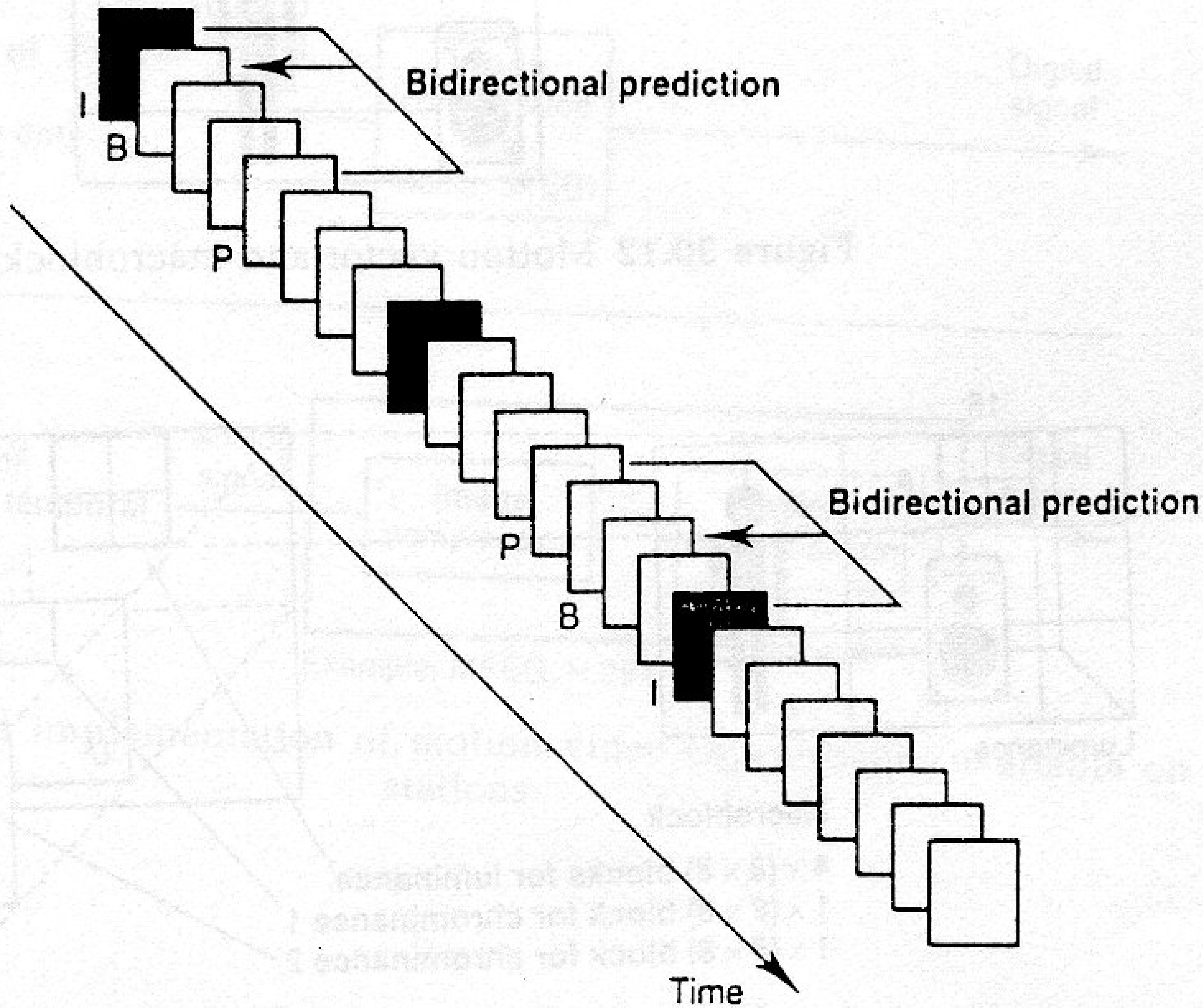


Figure 30.16 Sequence of intracoded, predicted, and bidirectional frames

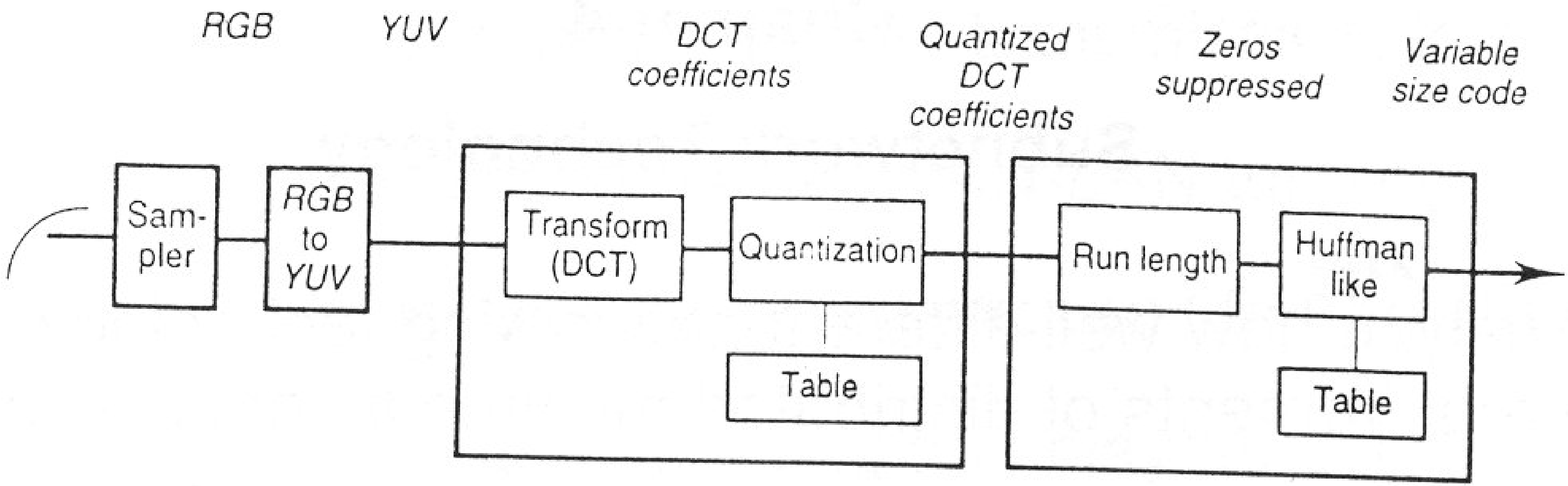


Figure 30.17 Processing steps for I-frame encoding

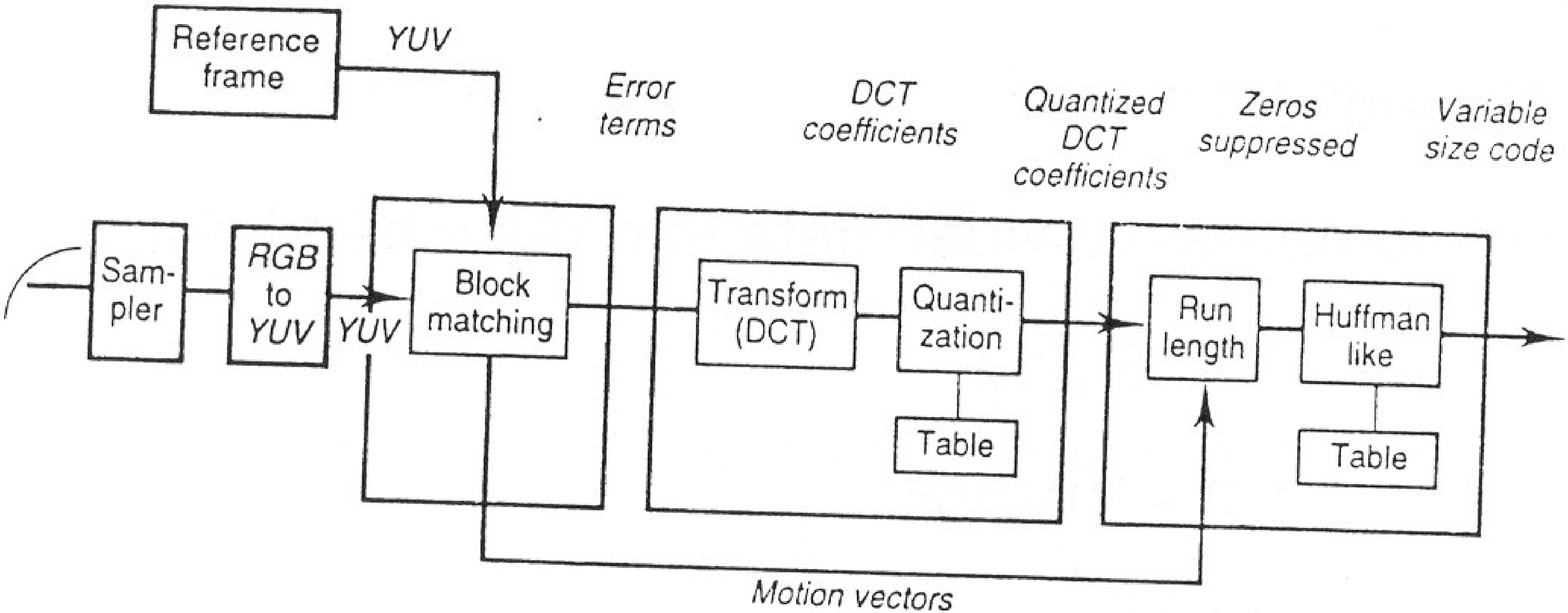


Figure 30.18 Processing steps for P- and B-frame encoding