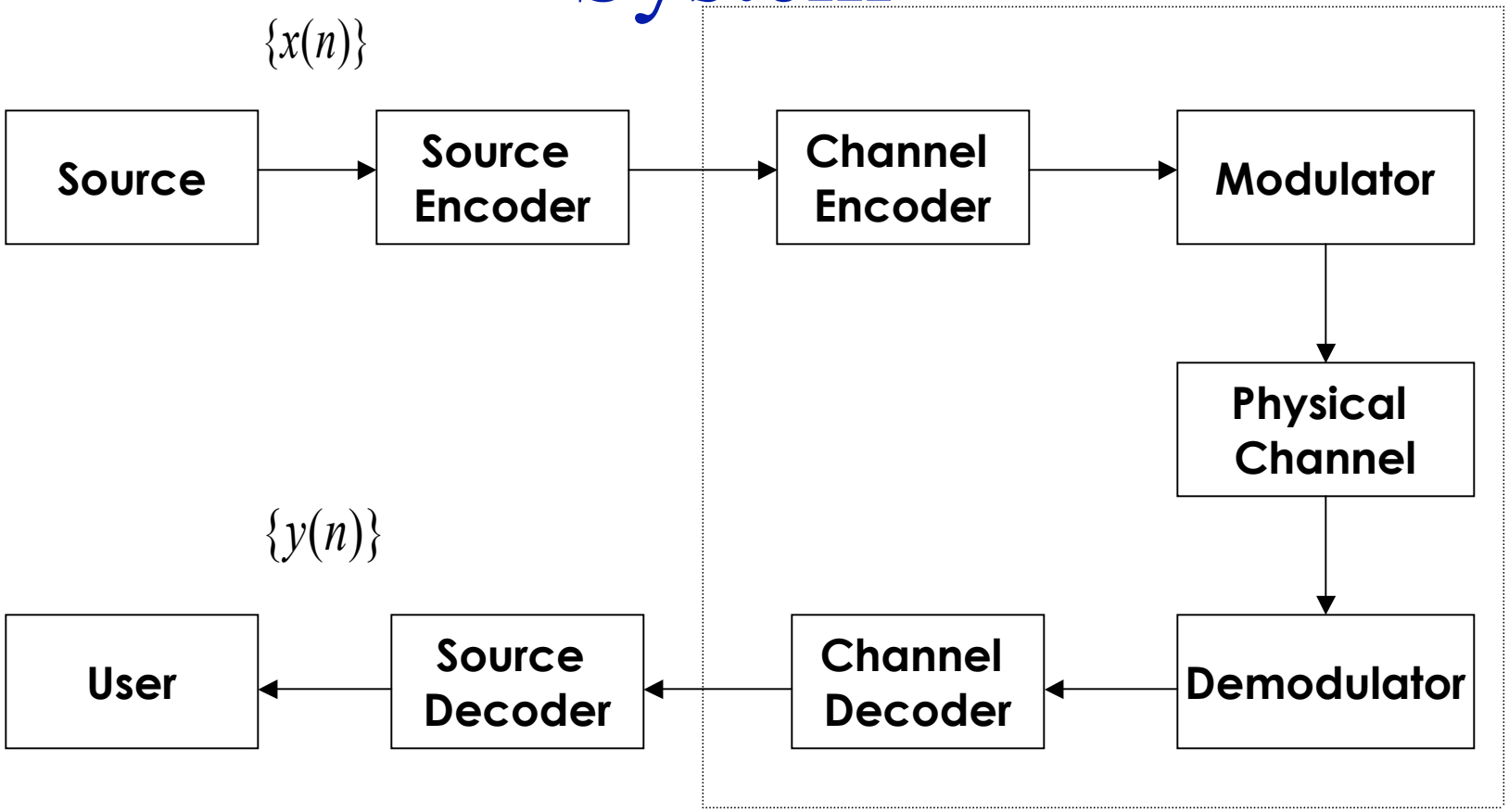
ECE 241
Multimedia Compression

Jerry D. Gibson
CNSI 2215
gibson@mat.ucsb.edu

UCSB

Typical Digital Communication System



What is Compression?

- Represent a source in digital form with as few bits as possible while still providing an acceptable reproduction of the original

Synonyms for Data Compression

- Signal Compression
- Signal Coding
- Source Coding
- Source Coding with a Fidelity Criterion
- Lossy (Noisy) Source Coding
- Lossless (Noiseless) Source Coding
- Data Compaction

More Synonyms

- Redundancy Removal
- Bandwidth Compression

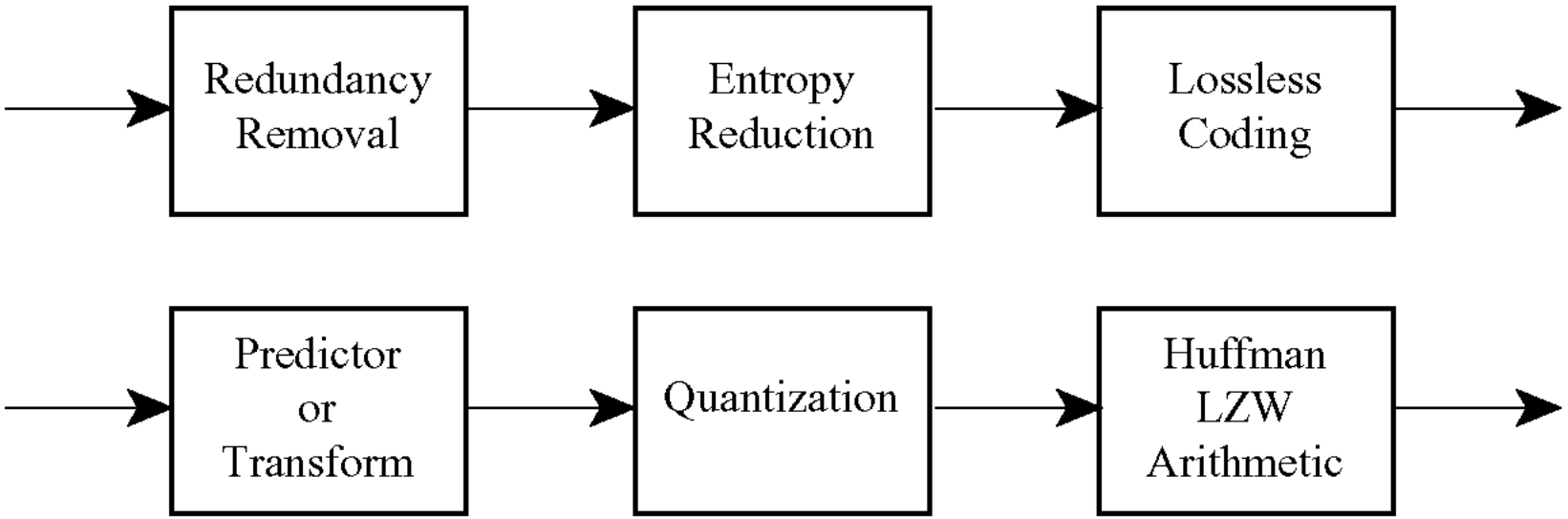
Components of a Compression Problem

- Source
- Rate
- Distortion Measure

Rate, Bits, and Bandwidth

- Required network bandwidth set by transmitted bit rate
- Bit Rate in *bits/sec* = *bits/sample* x *samples/sec*
- Sampling rate determined by source bandwidth

Major Steps in Data Compression



Speech and Audio Coding Bandwidths

- Narrowband Speech — 200 to 3400 Hz
- Wideband Speech — 50 to 7000 Hz
- Wideband Audio — 20 to 20,000 Hz
  - Number of channels
  - Stereo
  - Five channel surround

Approximate Bit Rates for Uncompressed Sources

Telephony (200–3400 Hz):	8000 samples/second $\times$ 12 bits/sample = 96 kbps
Wideband speech (50–7000 Hz):	16,000 samples/second $\times$ 14 bits/sample = 224 kbps
Wideband audio (20–20,000 Hz):	44,100 samples/second $\times$ 2 channels $\times$ 16 bits/sample $= 1.412$ Mbps
Images:	512 $\times$ 512 pixel color image $\times$ 24 bits/pixel = 6.3 Mbits/image
Video:	640 $\times$ 480 pixel color image $\times$ 24 bits/pixel $\times$ 30 images/second $= 221$ Mbps
HDTV:	1280 $\times$ 720 pixel color image $\times$ 60 images/second $\times$ 24 bits/pixel $= 1.3$ Gbps

Audio Sampling Rates

Application	Bandwidth (kHz)	Sampling Rate (kHz)
Voice telephony	3.2	8
Teleconferencing (audio)	7.0	16
Compact disc (CD) audio	20.0	44.1
Digital audiotape (DAT)	20.0	48

Video Sampling Rates

Format	Lines/Frame × Pixels/Line × Frames/Second =	Sampling Rate (million pixels per second)
CIF (videoconferencing)	$360 \times 288 \times 30 =$	3
CCIR (TV)	$720 \times 576 \times 30 =$	12
HDTV	$1280 \times 720 \times 60 =$	60

ITU-T Facsimile Standards

Size	Vertical Resolution (lines/mm)	Horizontal Resolution (pixels/mm)	Lines/ Frame	Pixels/ Line
Normal resolution 20.7 cm (8.27 inches) by 29.2 cm (11.7 inches)	3.85	8	1188	1728
High resolution 20.7 cm (8.27 inches) by 29.2 cm (11.7 inches)	7.7	8	2376	1728

Image and Video Formats

Formats	Usable Horizontal Lines*	Pixels per Line	Total Pixels per Frame	Frames per Second	Required Bandwidth/ Transmission Rate
Analog video					
NTSC					
(Americas, Asia)	338	426	150,0005	29.97	4 MHz
PAL (Europe)	411	420	172,0005	25.00	5 MHz
VHS	338	280	95,0005	29.97	<4 MHz
Computer image					
SVGA	1024	768	786,5005	60	—
VGA	640	480	307,0005	60	—

Formats	Usable Horizontal Lines*	Pixels per Line	Total Pixels per Frame	Frames per Second	Required Bandwidth/ Transmission Rate
Motion picture film					
35mm	(not a raster-		500,000	24	—
16mm	scanned image)		125,000	24	—
Digital video					
QCIF (H.261)	144	176	25,000	15–30	56 kbps–2 Mbp
CIF (H.261)	288	352	100,000	15–30	56 kbps–2 Mbp
HDTV	806	1920	1,550,000	50	140 Mbps
MPEG (constrained set)	345	360	124,000	30	1.5 Mbps and higher

*Eliminates retrace lines and includes the utilization ratio.

H.324 Video Formats

Format	Pixels	H.261	H.263
SQCIF	128×96	optional	required
QCIF	176×144	required	required
CIF	352×288	optional	optional
4 CIF	704×576	n/a	optional
16 CIF	1408×1152	n/a	optional

Networks and Network Services

POTS	28.8-56 Kbits/s
ISDN	64-128 Kbits/s
ADSL	1.544-8.448 Mbits/s (downstream) 16-640 Kbits/s (upstream)
VDSL	12.96-55.2 Mbits/s
CATV	20-40 Mbits/s
OC-N/STS-N	N x 51.84 Mbits/s
Ethernet	10 Mbits/s
Fast Ethernet	100 Mbits/s
Gigabit Ethernet	1,000 Mbits/s
FDDI	100 Mbits/s
802.11(wireless)	1, 2, 5.5, 11, and 22 Mbits/s in 2.4 GHz band
802.11 a(wireless)	6-54 Mbits/s in 5GHz band

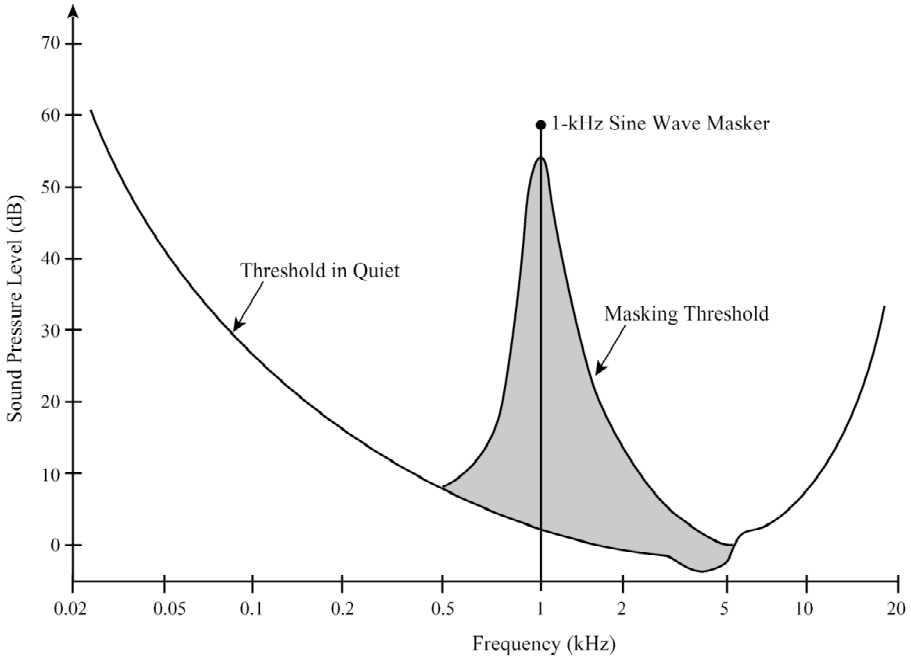
Reduce Source Bit Rates but Keep Quality

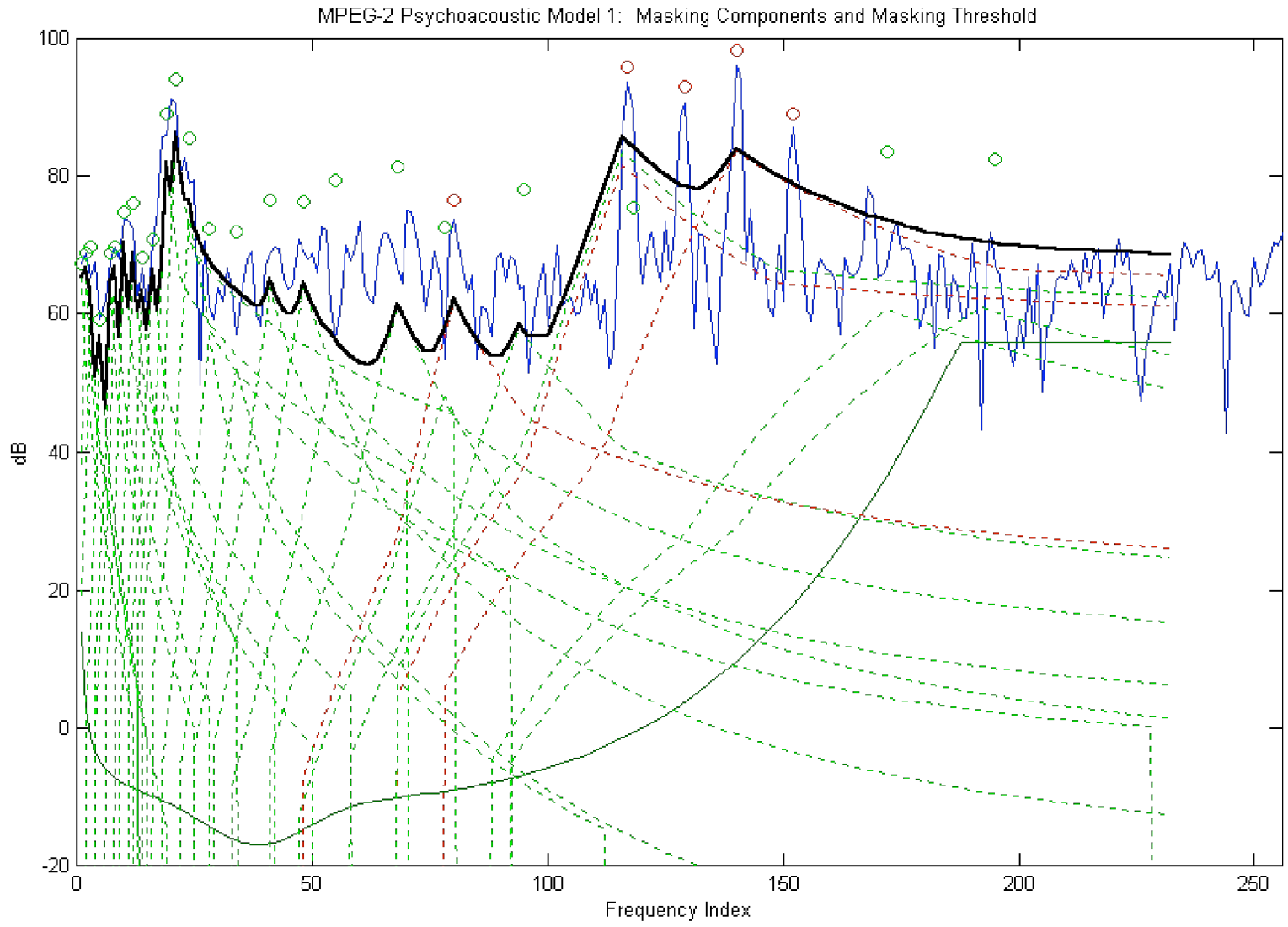
- Source (Speech, Audio, Still Images, Video) Compression
- What is Compression?
- **Goal: Represent a source in digital form with as few bits as possible while still providing an acceptable reproduction of the original**

Design Distortion Measures

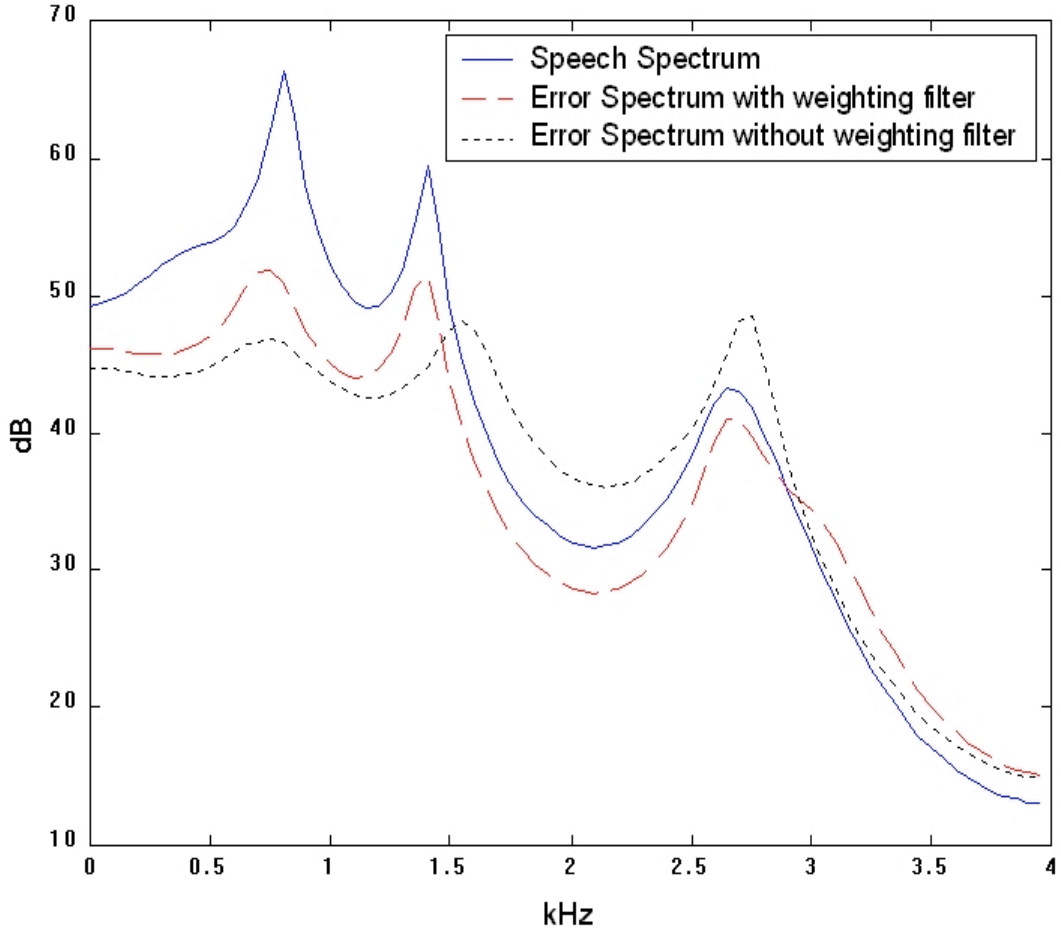
- Mean Squared Error
  - Mathematically Tractable
  - Not Necessarily Perceptually Meaningful
  - Important for Initial Rankings
- Frequency-Weighted Squared Error
- Perceptually-Based Distortion Measures

THRESHOLD IN QUIET AND MASKING THRESHOLD




MPEG-2 Psychoacoustic Model 1: Masking Components and Masking Threshold
dB
Frequency Index
100
80
60
40
20
0
-20
0
50
100
150
200
250

CELP Perceptual Weighting



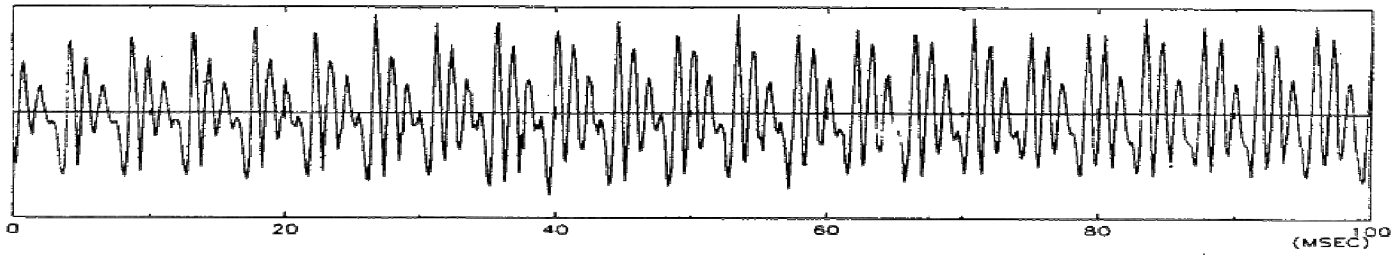
Performance Evaluation

- Speech
  - Listening Tests, including
    - MOS
    - DRT
    - DAM
  - Distance Measures
- Audio—Listening tests—transparency
- Images and Video--Viewing

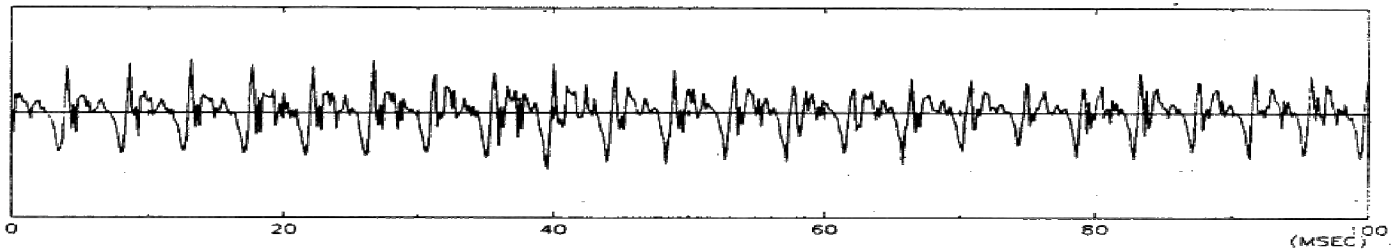
Applications of Speech Coding

- Wireline Telephony
- Videoconferencing
- Digital Cellular
- IP Telephony
- Voice Mail
- Speech Storage

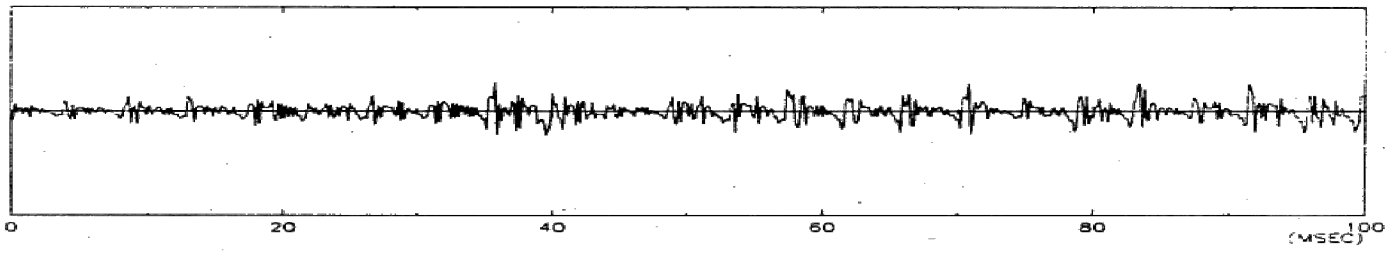
Redundancy Removal From Speech Signals



(a) Speech Waveform



(b) Short term removed signal

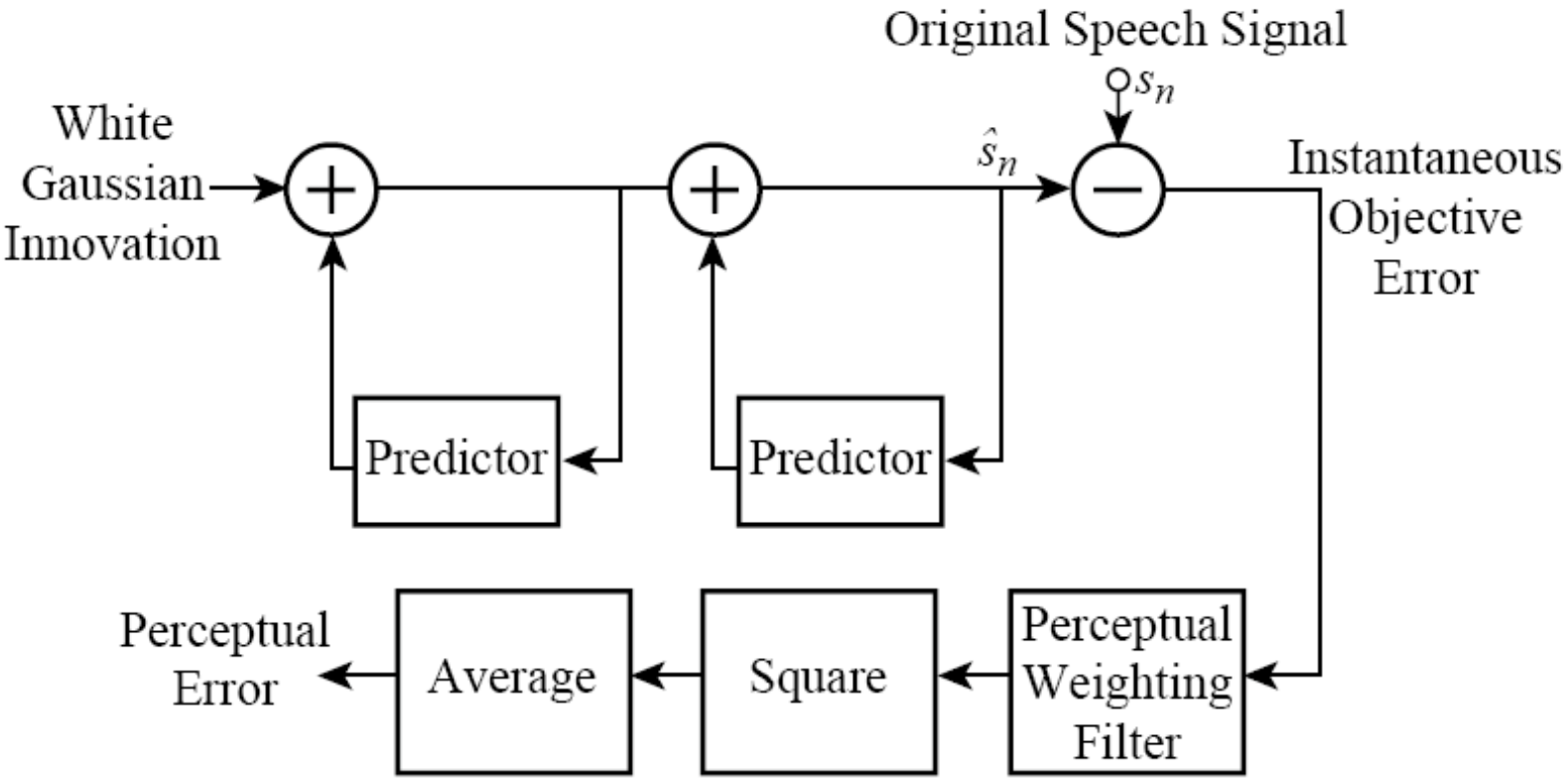


(c) Long and short term removed signal

Key Advances

- Perceptual Distortion Measures
- Digital Signal Processing
- Analysis-by-Synthesis Structures
  - Codebook Excitation
  - Single Gain for All Pulses

Analysis-by-Synthesis Coding



Speech and Audio Coding Standards

- Narrowband speech
  - GSM-AMR, G.729, G.723, G.728, IS-127(EVRC), IS-96(QCELP), IS-95(VSELP)
  - G.711(PCM), G.721(ADPCM), G.726(ADPCM)
  - LPC-10, MELP,…
- Wideband speech
  - G.722 (ADPCM)
  - G.722.1 (Transform)
  - AMR-WB (CELP)
- Wideband audio
  - MPEG-1,2,4
  - Philips PASC
  - Sony ATRAC
  - DOLBY AC-3

Narrowband Speech Coding Performance

- Original Sample 
- G.711 _-Law PCM 64 Kbps 
- G.726 ADPCM 16, 24, 32 and 40 Kbps  
- G.723.1
  - ACELP 5.3 Kbps 
  - MP-MLQ 6.3 Kbps 
- G.729 ACELP 8 Kbps 
- MPEG-4

Coded Classical Music

8KHz:


G.711


G.726(32K)


G.729(8K)


G.723.1(5.3K)




NB_AMR(12.2K, 6.7K,4.75K)

16KHz:



G.722(64K,48K)



G.722.1(32K,24K)




WB_AMR(23.85K,12.65K,6.60K)

Coded Music+Voice

8KHz:


G.711


G.726(32K)


G.729(8K)


G.723.1(5.3K)




NB_AMR(12.2K, 6.7K,4.75K)

16KHz:



G.722(64K,48K)

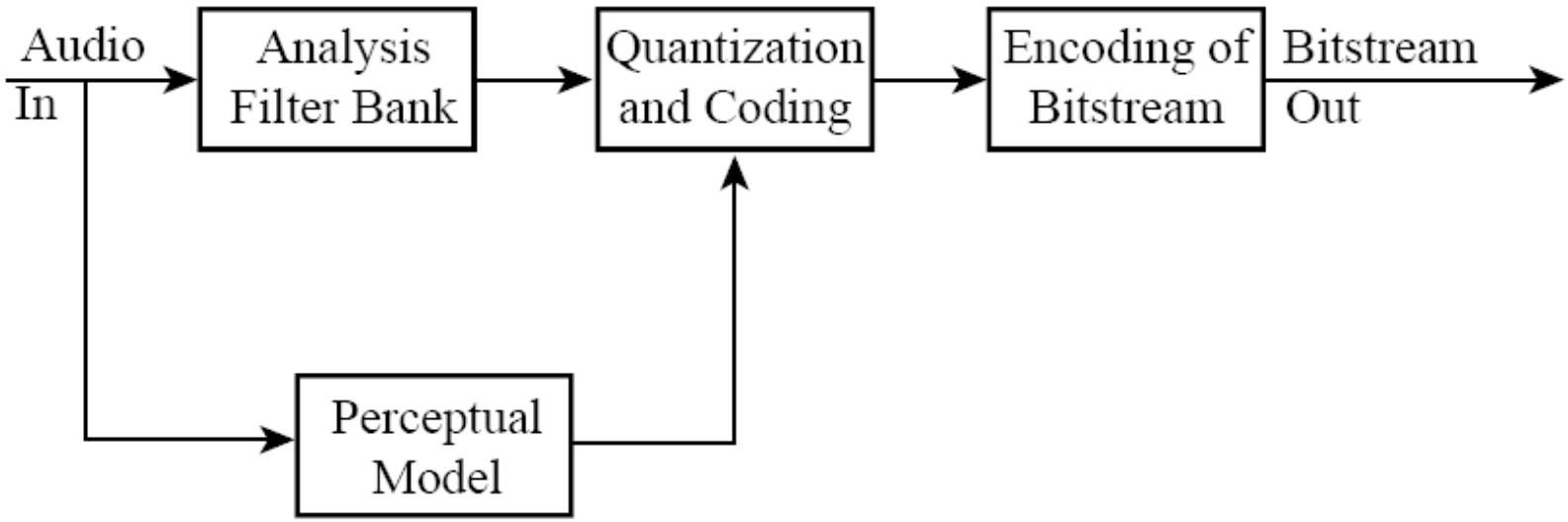


G.722.1(32K,24K)

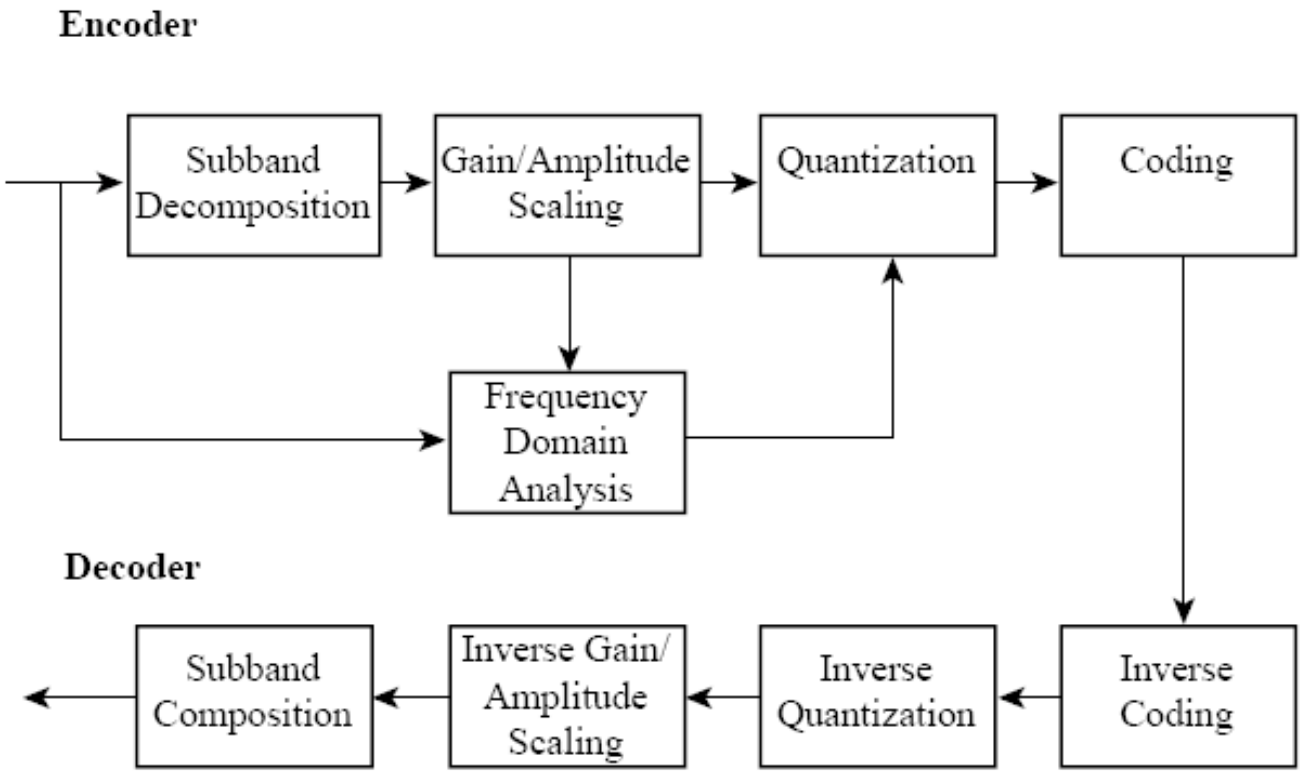



WB_AMR(23.85K,12.65K,6.60K)

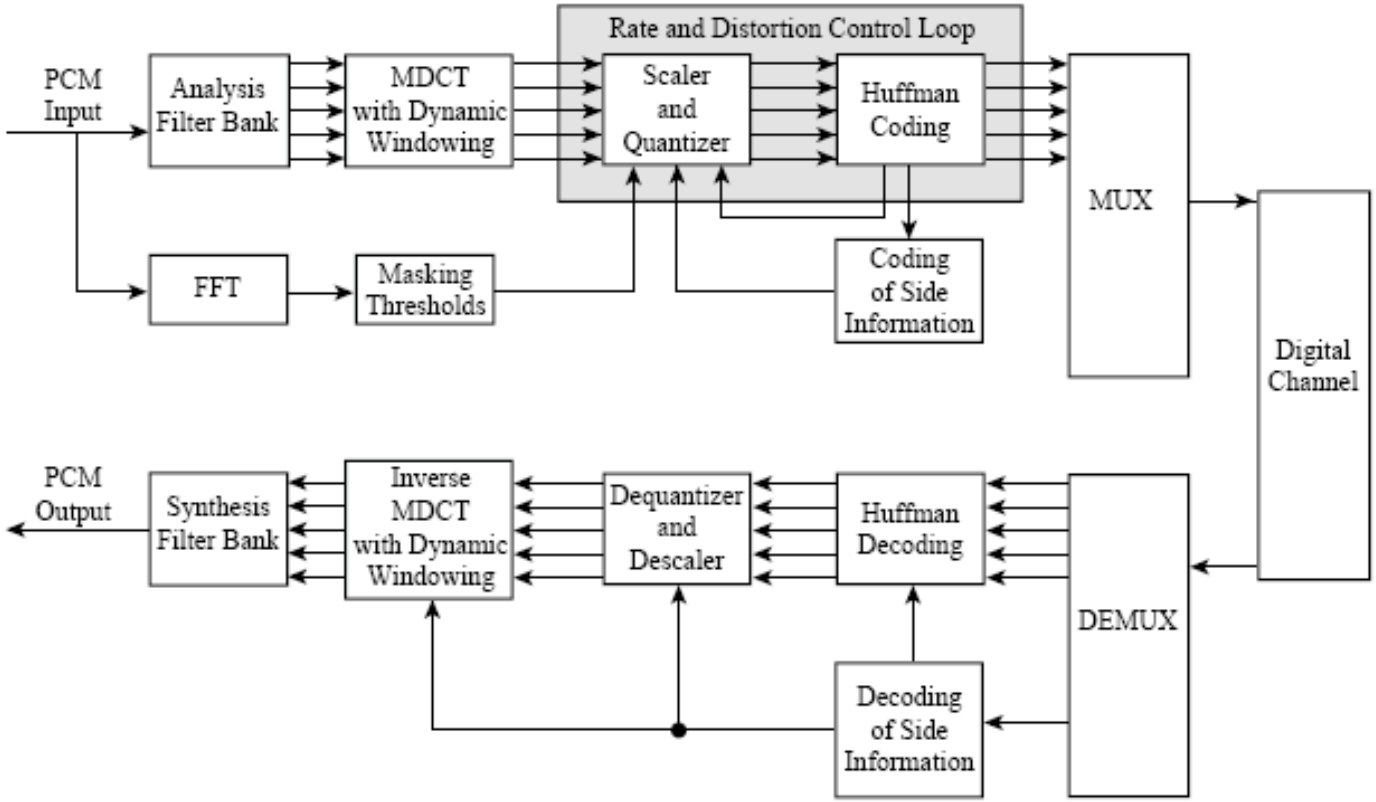
Basic Audio Coding Method



MPEG Audio Coding Block Diagram



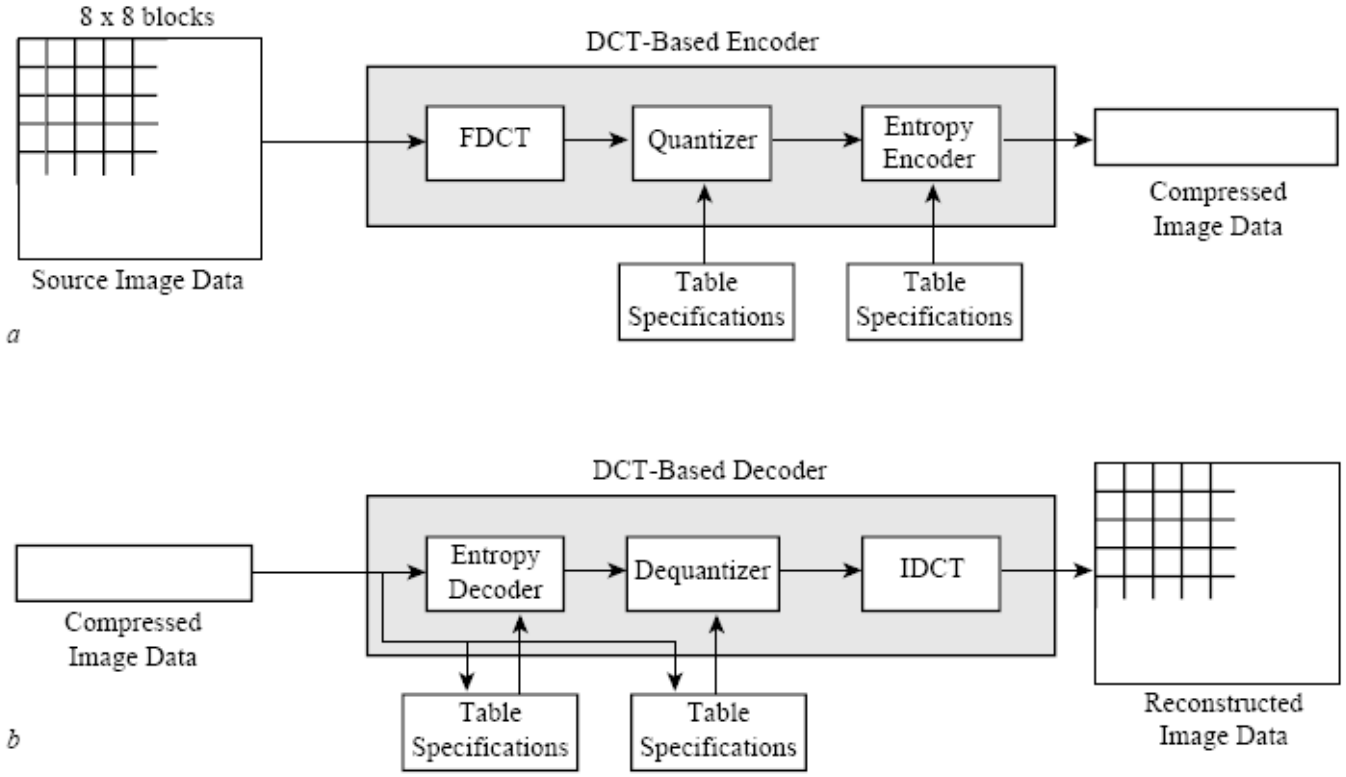
MPEG Layer 3 Encoding and Decoding



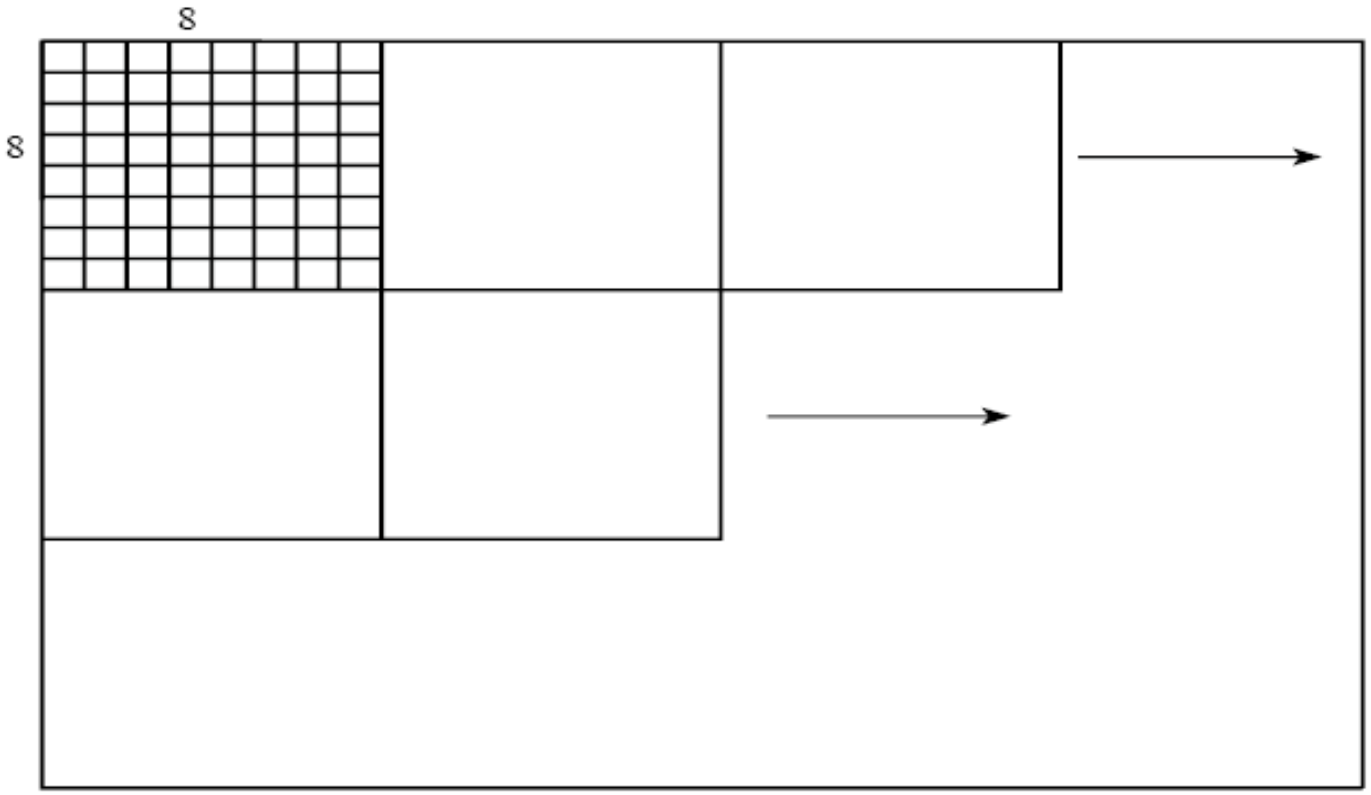
Still Image Coding Standards

- JPEG
- JPEG2000

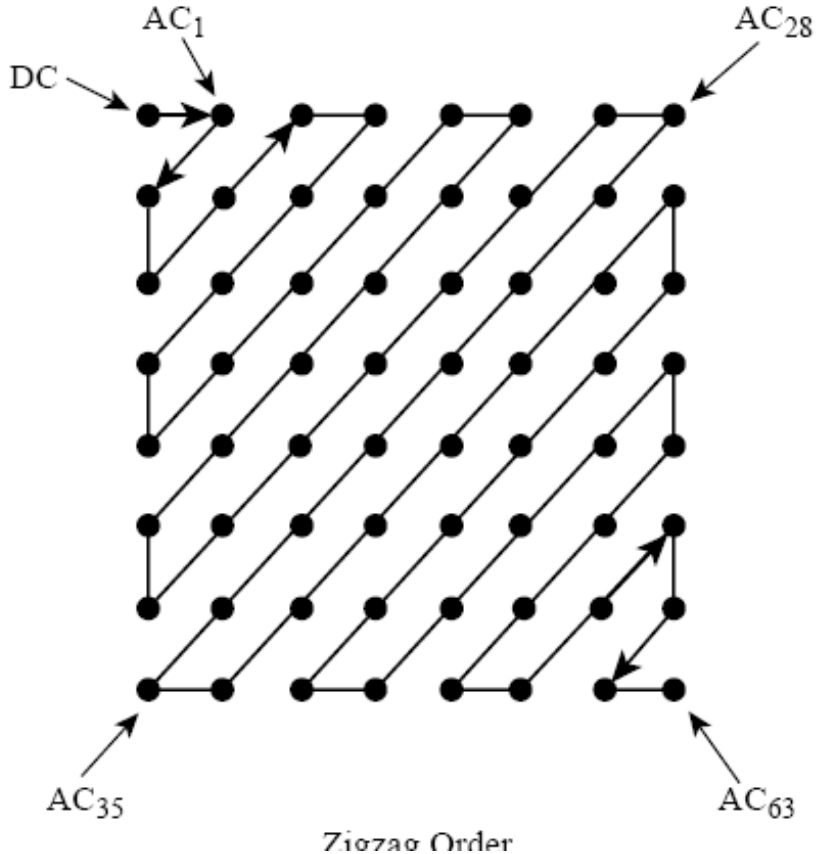
JPEG Lossy Coding



Partitioning the Image



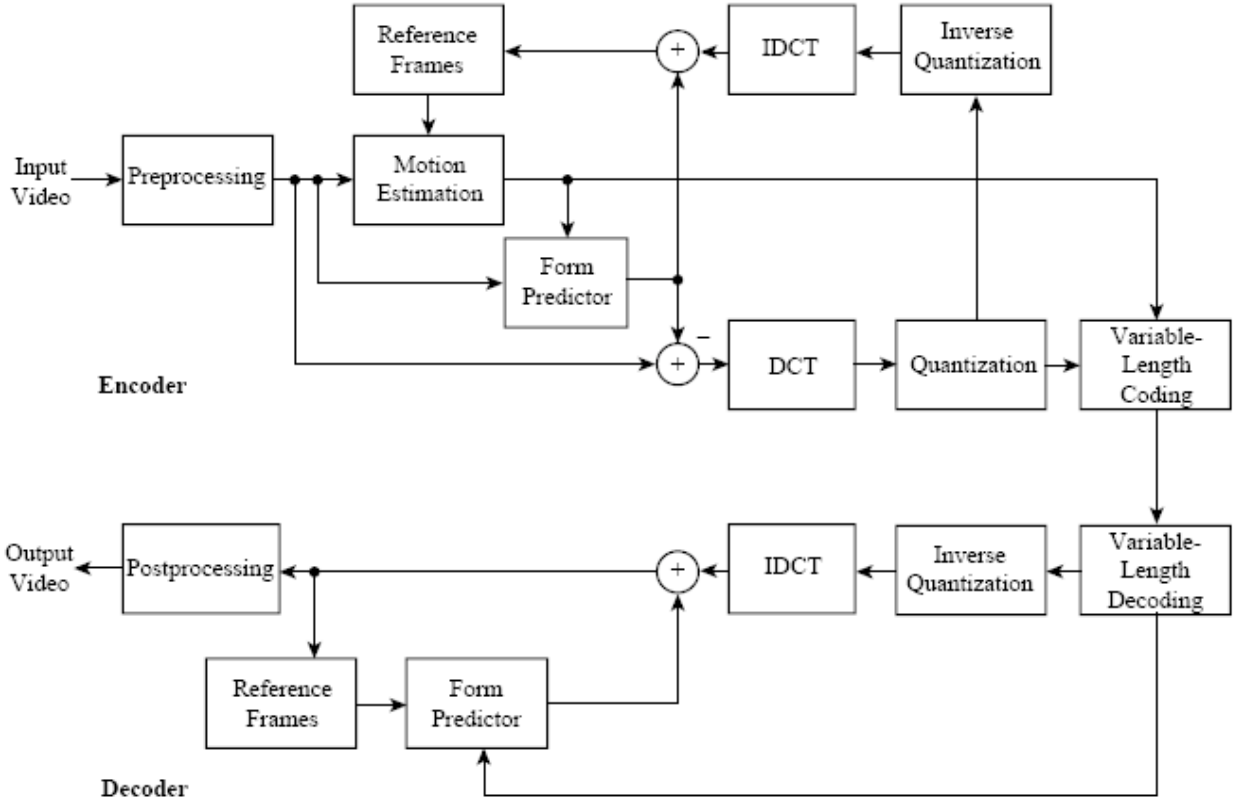
Zig-Zag Coefficient Ordering



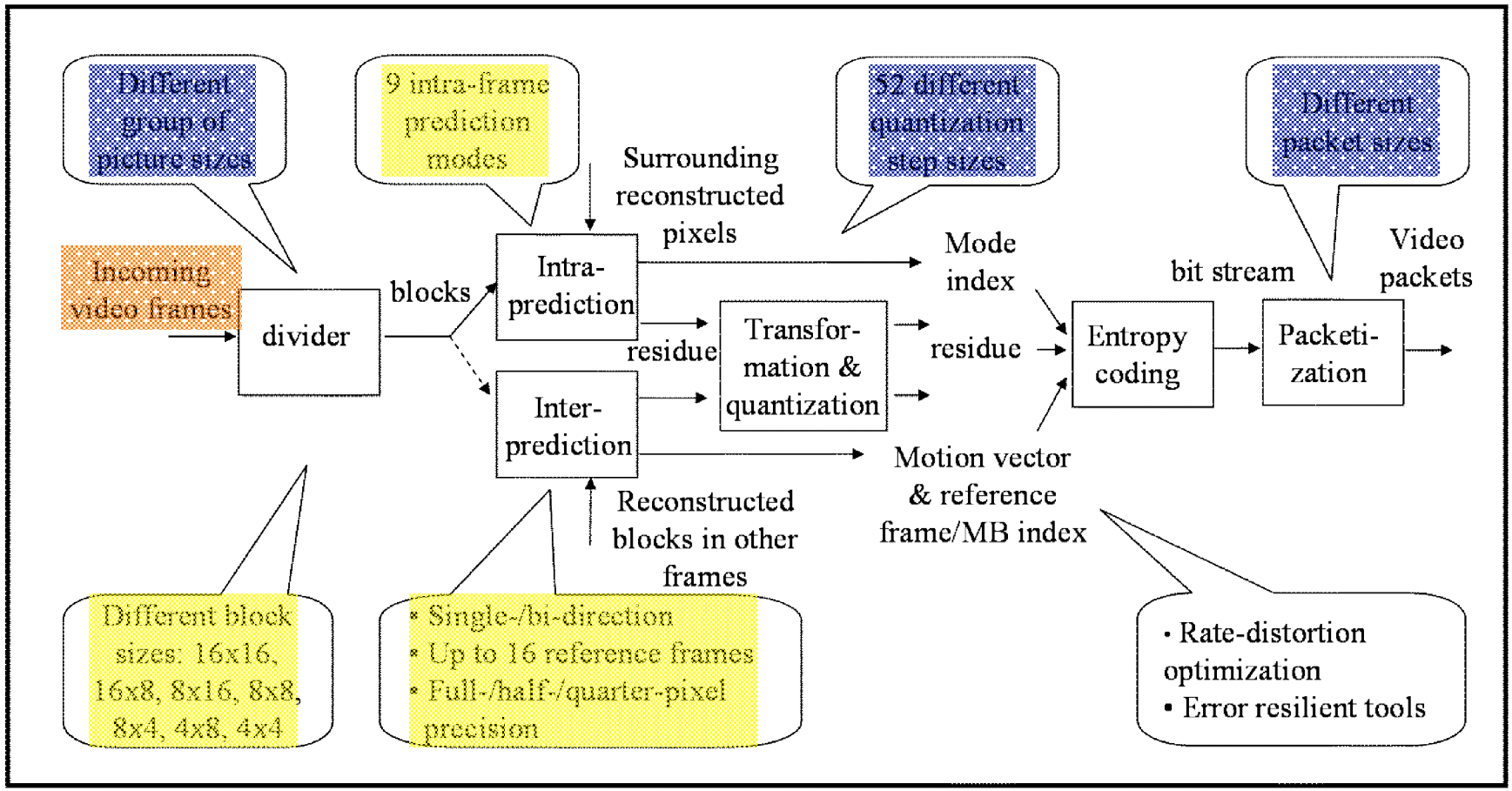
Selected Video Coding Standards

- MPEG-2
- MPEG-4
- H.264/AVC

Block Diagram of MPEG Video Coding



H.264 Encoder



Choices in Video Codecs for Different Videos

Video	silent.cif 		paris.cif 		stefan.cif 	
Typical application	video conference		news broadcast		sports broadcast	
QP	26	30	26	30	26	30
Avg psnr	36.69	34.22	36.59	33.45	36.69	33.47
Bit rate (kbps)	169.5	97.8	373.5	218.9	1396.8	404.6
I frame size (bytes)	13945	8826	19886	14390	30432	15978
Average of P frame size (bytes)	1272	725	2924	1683	11429	3230
Variance of P frame size (bytes)	412	254	322	219	1544	625

Video Quality Comparison 1

Video Quality Comparison 2

UCSB

Video Demo

- Channel snr =3dB, PHY data rate = 6 Mbps, packet size = 100 bytes, packet loss rate = 1.54%
- QP = 26, group of picture size = 15

Avg PSNR = 35.06

Avg PSNR = 34.06

Avg PSNR = 33.06

play movies

UCSB

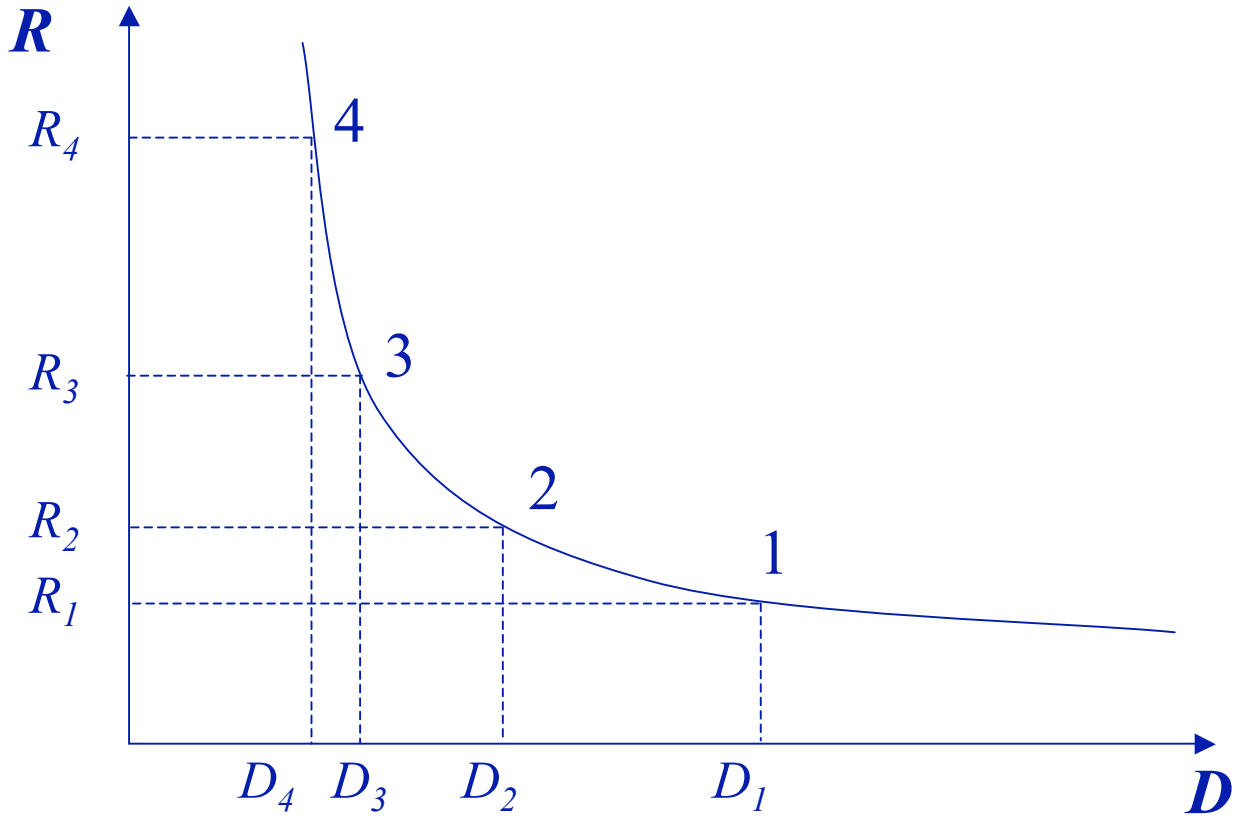
Key Functionalities

- Error Concealment
- Scalability
  - SNR
  - Spatial
  - Temporal
  - Bandwidth
- Multiple Descriptions

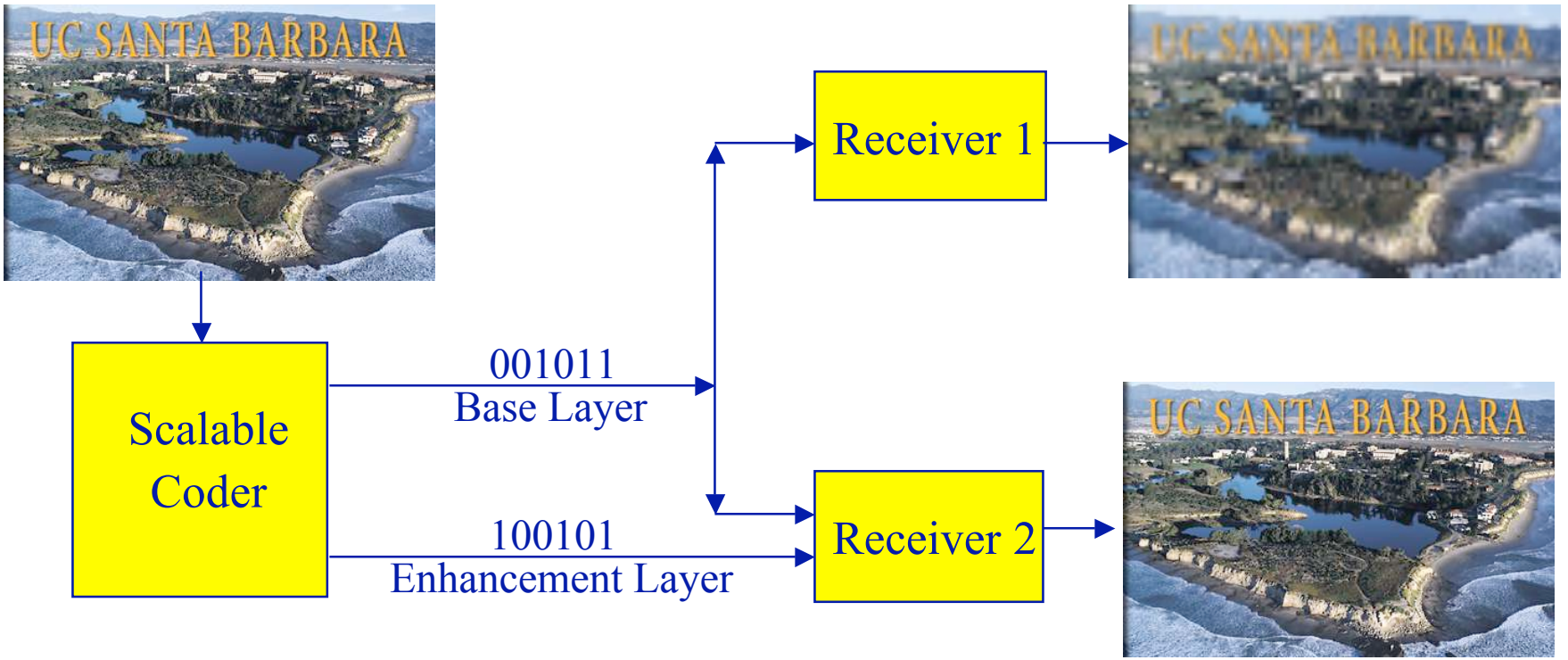
Scalable Coding

- Sometimes denoted as layered coding, embedded coding, or variable rate coding
- Scalable Coding consists of a core coder at the lowest bit rate plus one or more enhancement layers
- Quality improvement is achieved by sending only an incremental bit rate above the core layer
- Speech Scalable Coding: SNR scalability, Bandwidth scalability

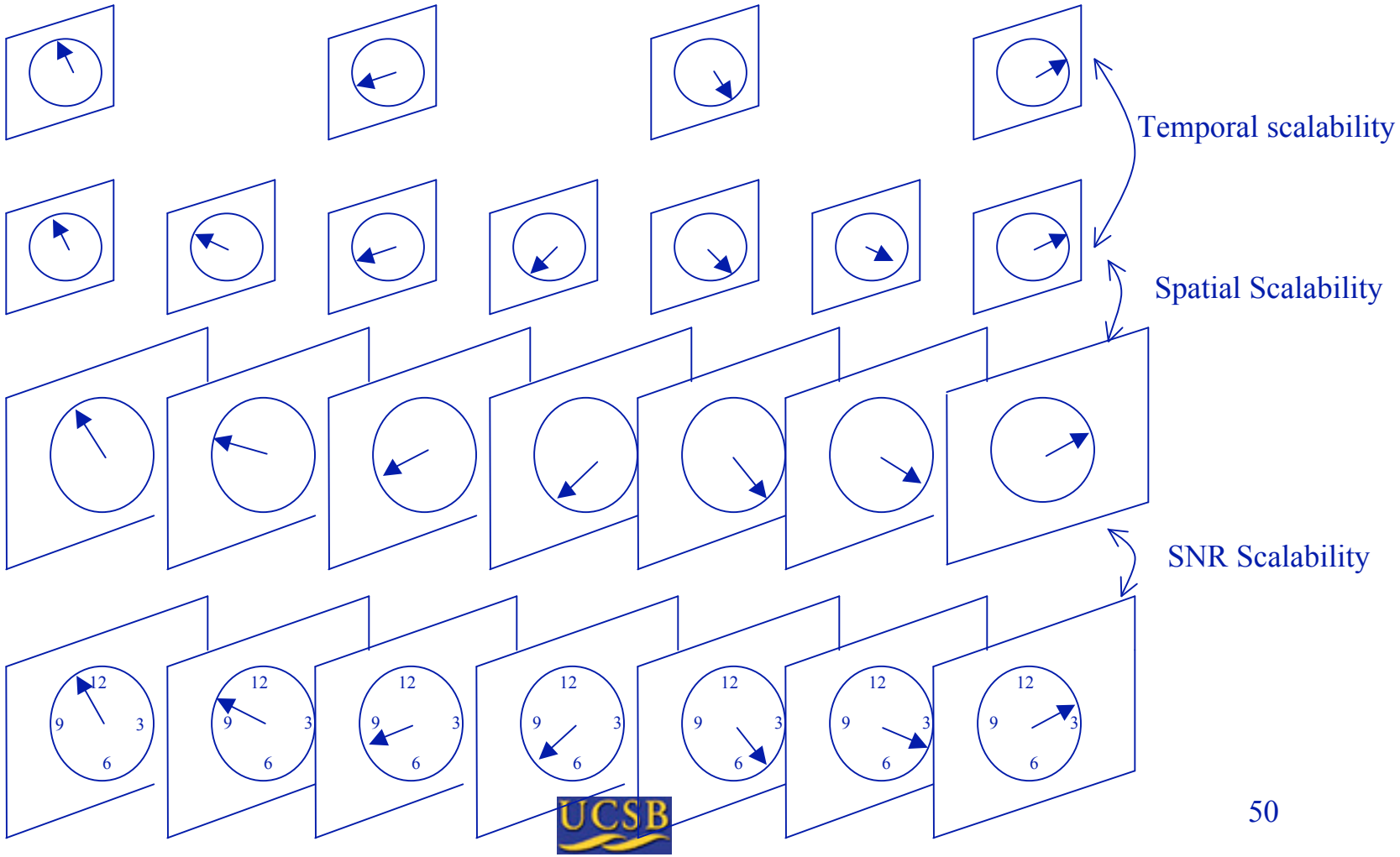
The Successive Refinement Problem



Scalable Coding



Temporal, Spatial and SNR scalability



Multiple Descriptions Coding