ECE 241
Multimedia Compression
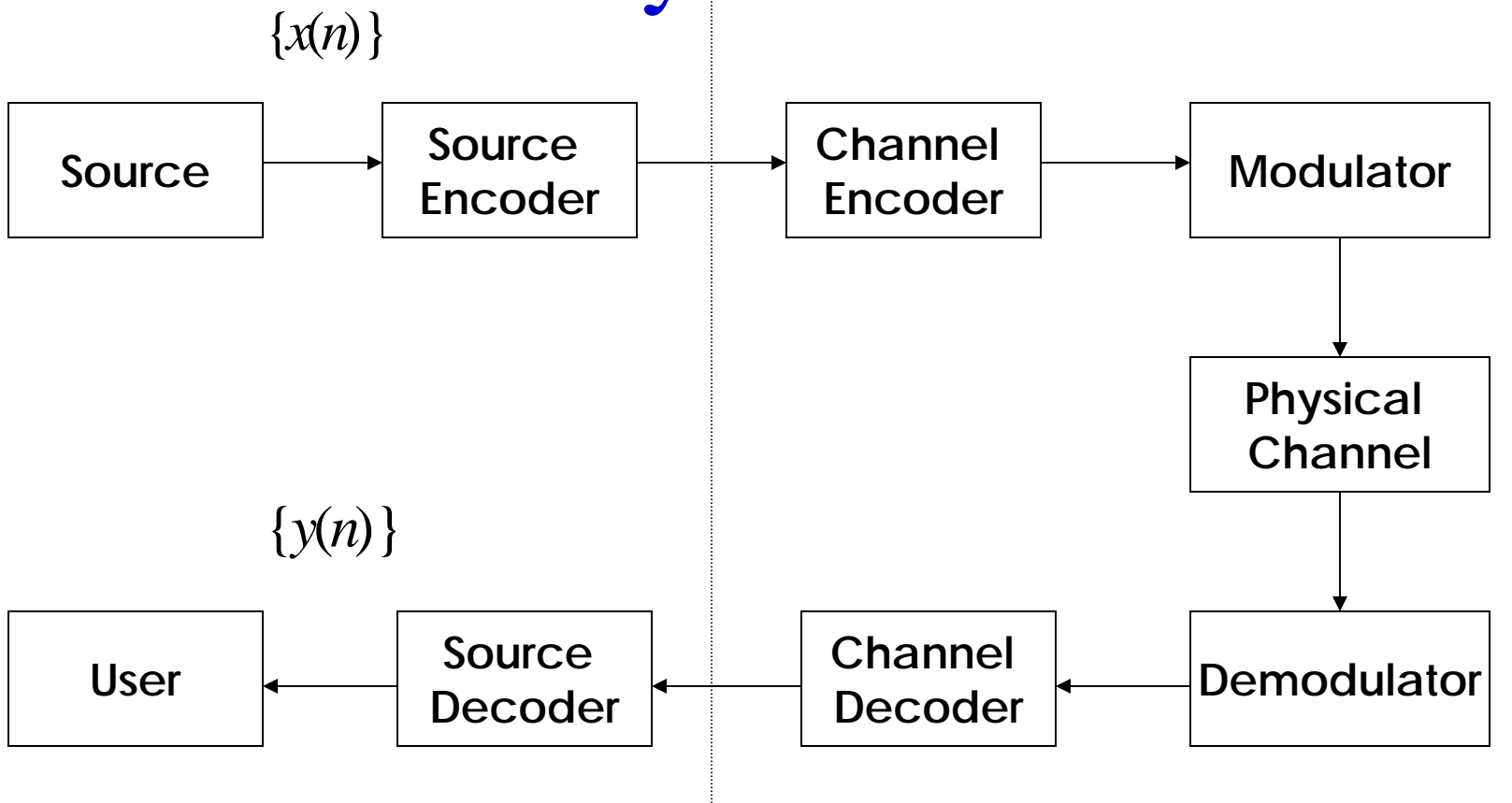
Jerry D. Gibson

UCSB

Typical Digital Communication System

$\{x(n)\}$

```
┌─────────┐     ┌──────────┐     ┌──────────┐     ┌───────────┐
│ Source  │ ──► │  Source  │ ──► │ Channel  │ ──► │ Modulator │
│         │     │ Encoder  │     │ Encoder  │     │           │
└─────────┘     └──────────┘     └──────────┘     └───────────┘
                                                        │
                                                        ▼
                                                  ┌───────────┐
                                                  │ Physical  │
                                                  │ Channel   │
                                                  └───────────┘
                                                        │
$\{y(n)\}$                                              ▼
┌─────────┐     ┌──────────┐     ┌──────────┐     ┌─────────────┐
│  User   │ ◄── │  Source  │ ◄── │ Channel  │ ◄── │ Demodulator │
│         │     │ Decoder  │     │ Decoder  │     │             │
└─────────┘     └──────────┘     └──────────┘     └─────────────┘
```

UCSB

What is Compression?

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

UCSB

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

UCSB

More Synonyms
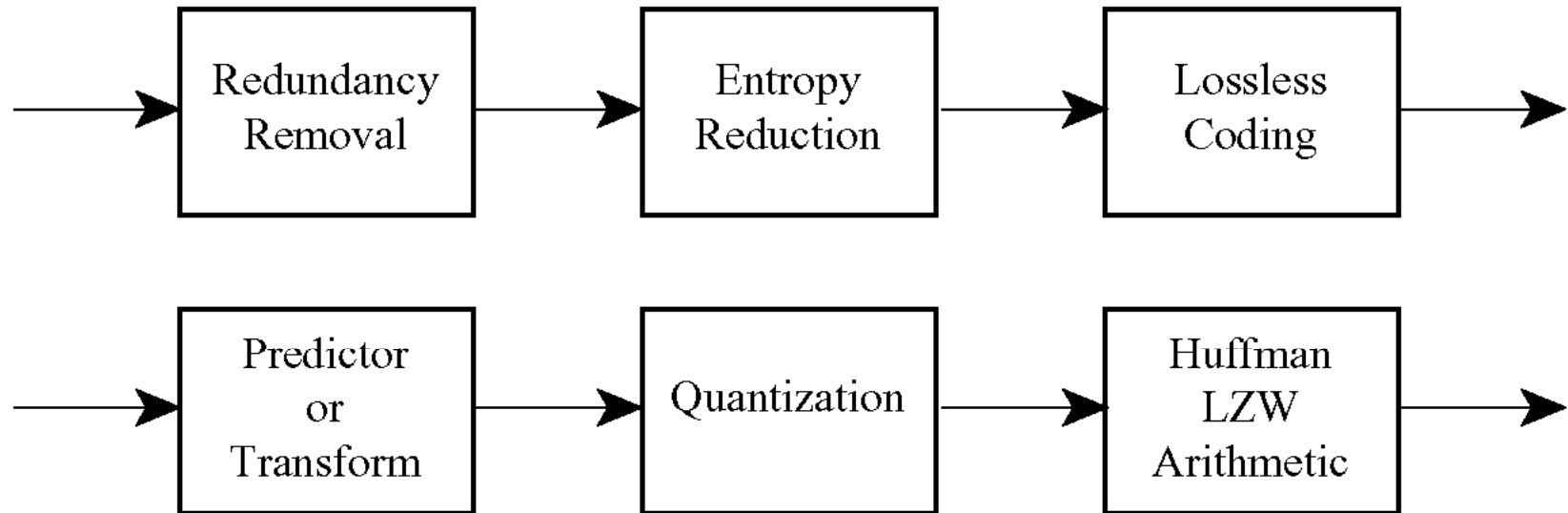
- Redundancy Removal
- Bandwidth Compression

UCSB

Components of a Compression Problem

- Source

- Rate

- Distortion Measure

- And Complexity

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

UCSB

Major Steps in Data Compression

```
         ┌─────────────┐      ┌─────────────┐      ┌─────────────┐
   ──────▶│  Redundancy │─────▶│   Entropy   │─────▶│  Lossless   │──────▶
         │   Removal   │      │  Reduction  │      │   Coding    │
         └─────────────┘      └─────────────┘      └─────────────┘

         ┌─────────────┐      ┌─────────────┐      ┌─────────────┐
   ──────▶│  Predictor  │─────▶│Quantization │─────▶│   Huffman   │──────▶
         │     or      │      │             │      │     LZW     │
         │  Transform  │      │             │      │  Arithmetic │
         └─────────────┘      └─────────────┘      └─────────────┘
```

8

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

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

UCSB

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**

UCSB

Design Distortion Measures

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

UCSB

THRESHOLD IN QUIET AND MASKING THRESHOLD

MPEG-2 Psychoacoustic Model 1: Masking Components and Masking Threshold
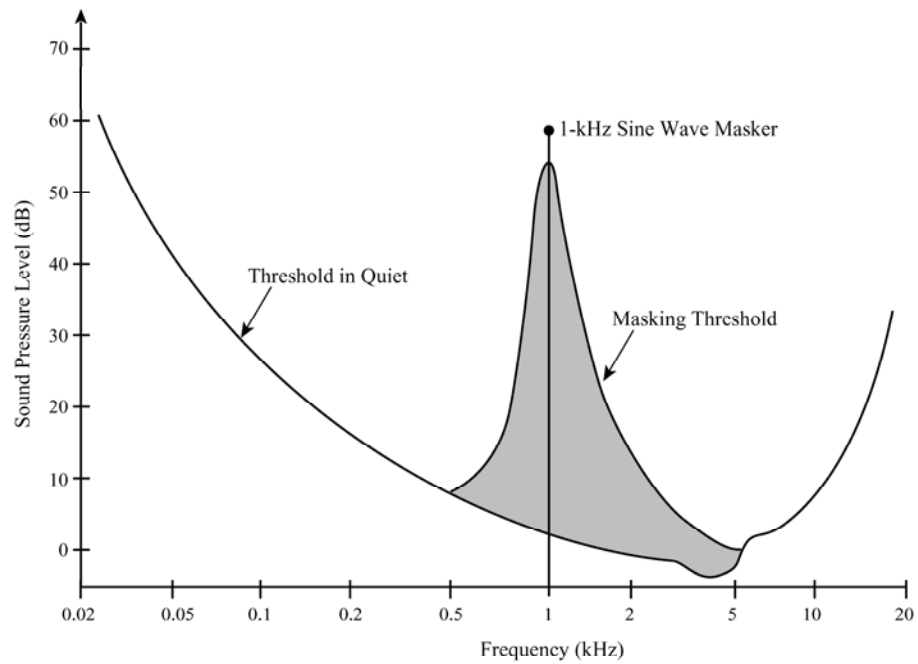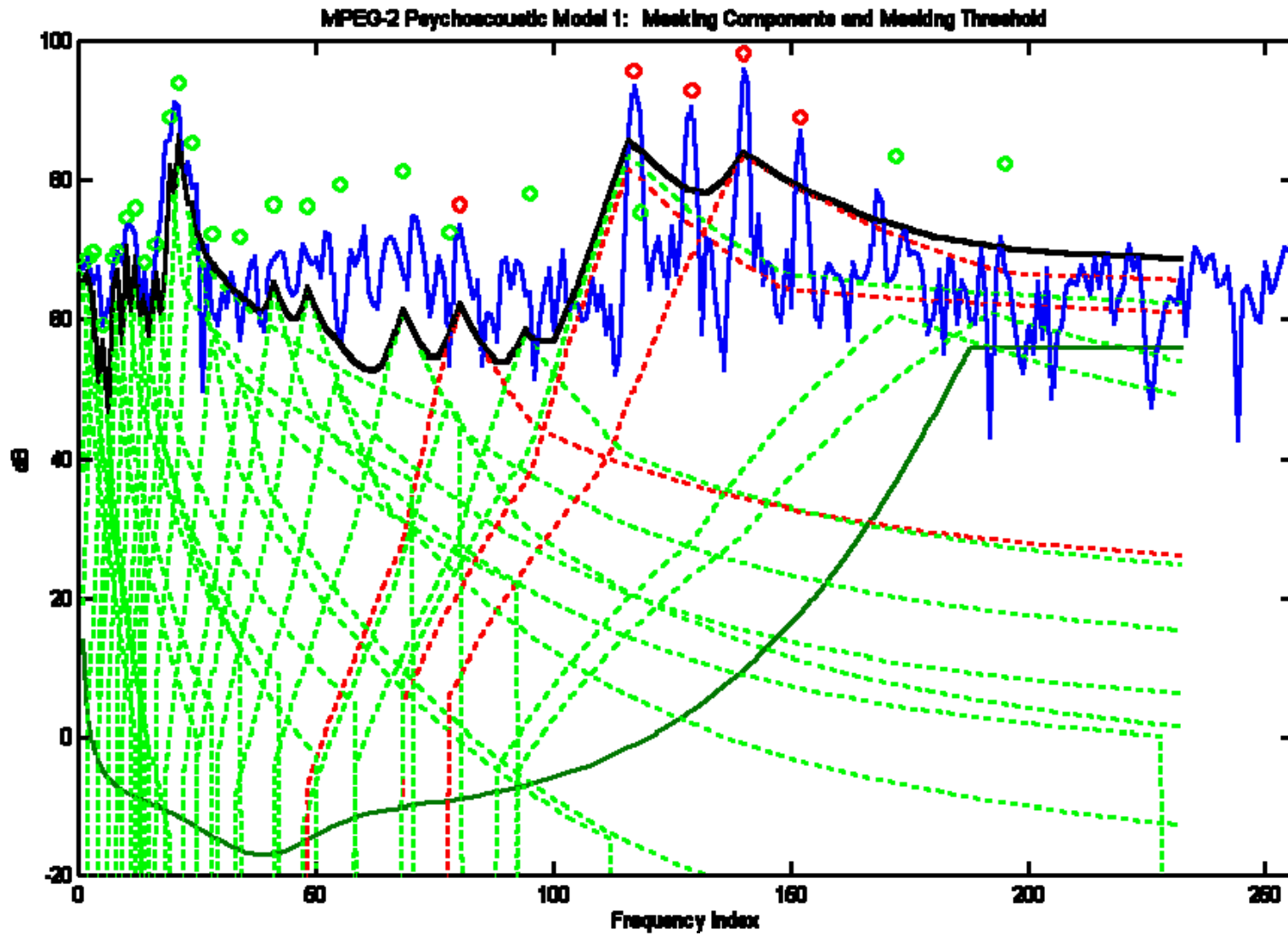
CELP Perceptual Weighting
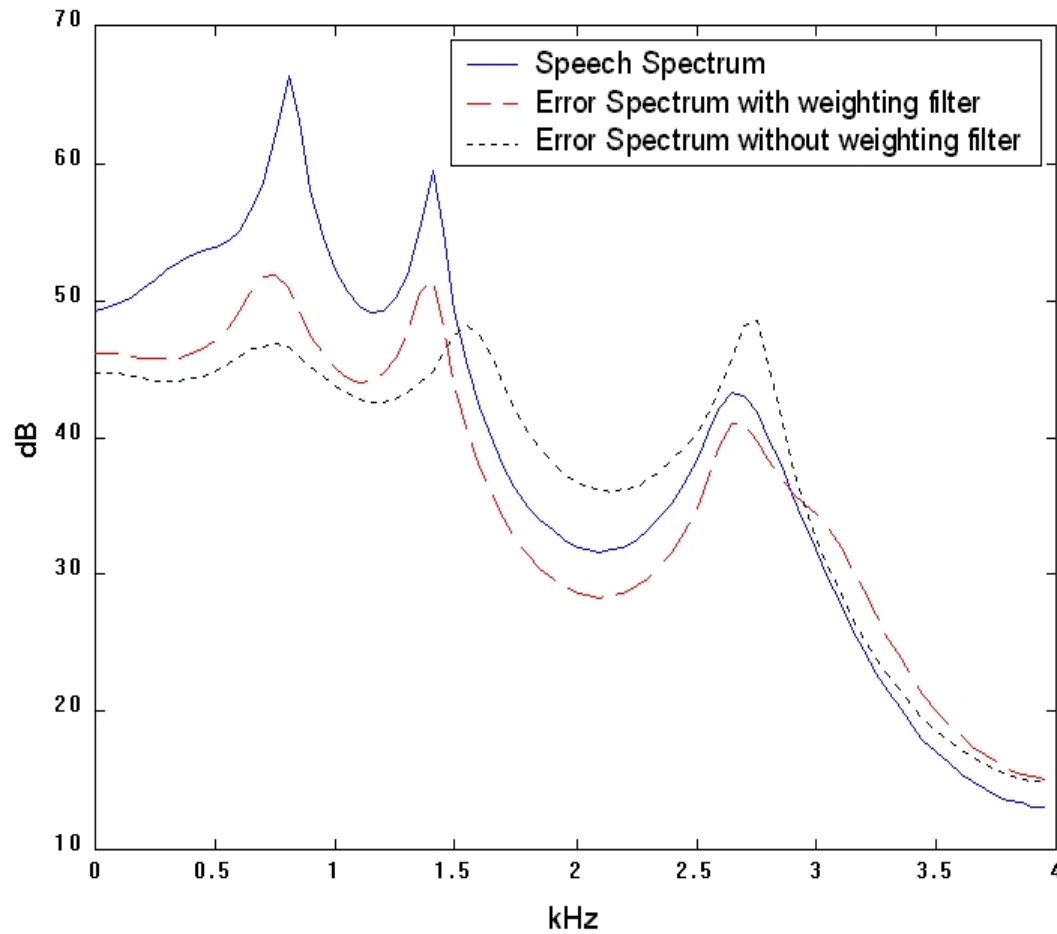
Performance Evaluation

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

Applications of Speech Coding

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

UCSB

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

Still Image and Video Compression Standards

- JPEG, 0.25-2.0 bits/pixel
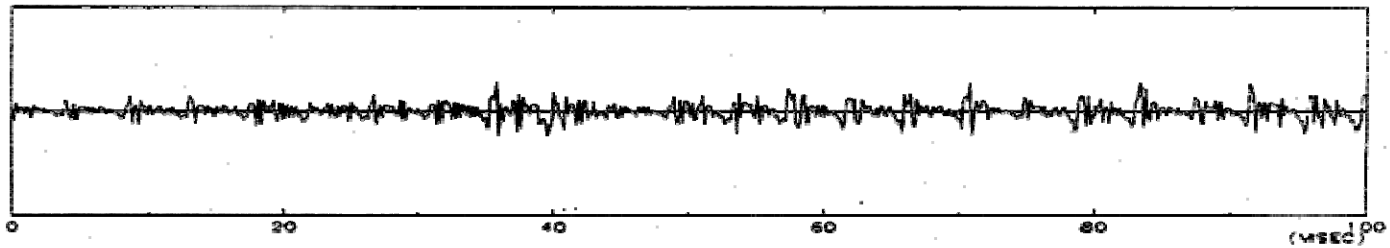
- JPEG 2000

- MPEG-2, 4-10 Mbps

- VC-1

- AVC/H.264

UCSB

# Redundancy Removal From Speech Signals
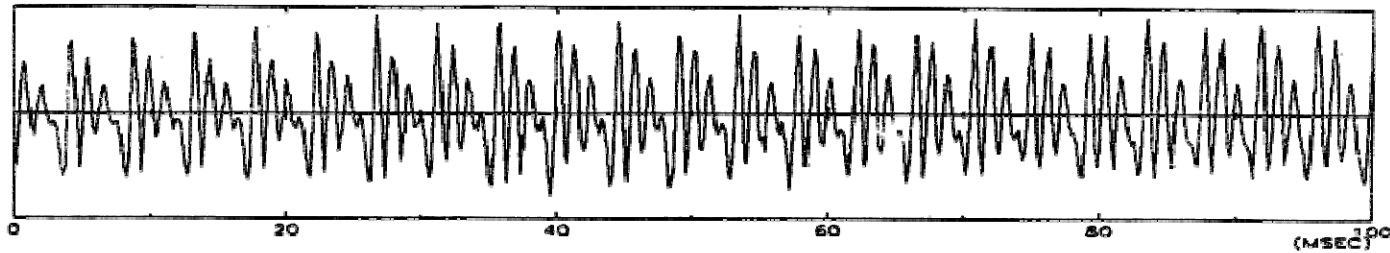
(a) Speech Waveform
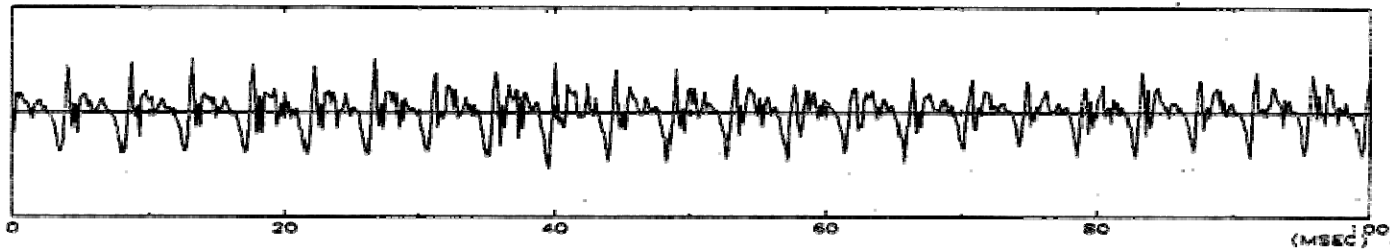
(b) Short term removed signal

(c) Long and short term removed signal

Speech Coding Techniques

- ## Algorithms:
  - Parametric coding
  - CELP coding
  - ADPCM coding
  - Transform coding
- ## Perceptual Masking:
  - Weighting filtering
  - Psychoacoustic modeling

Legend:
- Speech Spectrum
- Error Spectrum with weighting filter
- Error Spectrum without weighting filter
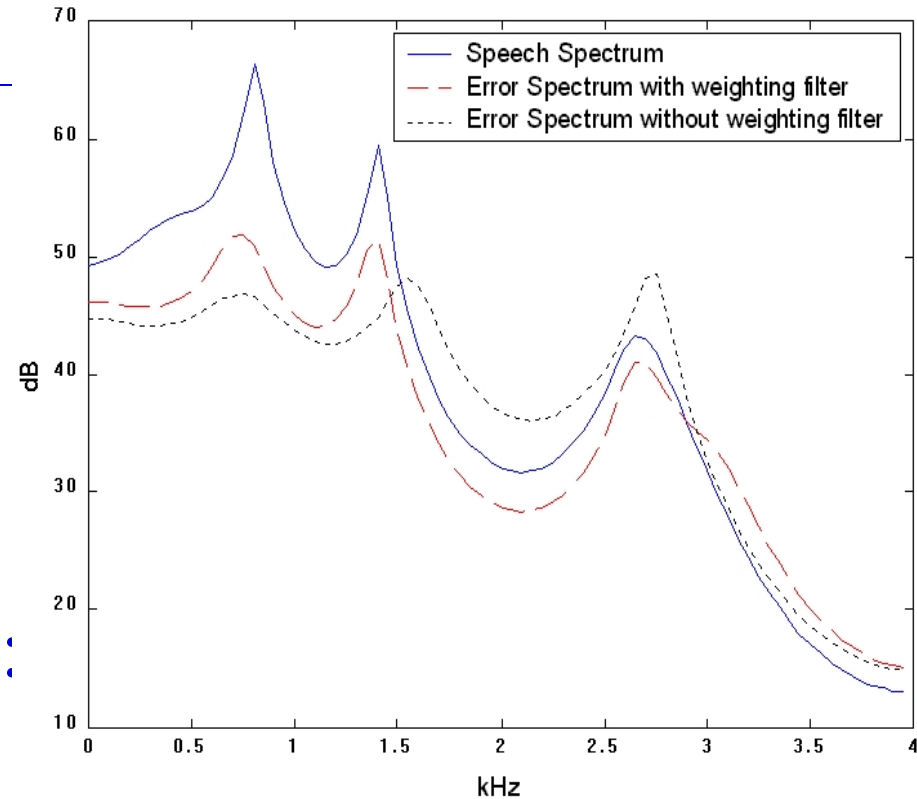
Key Advances

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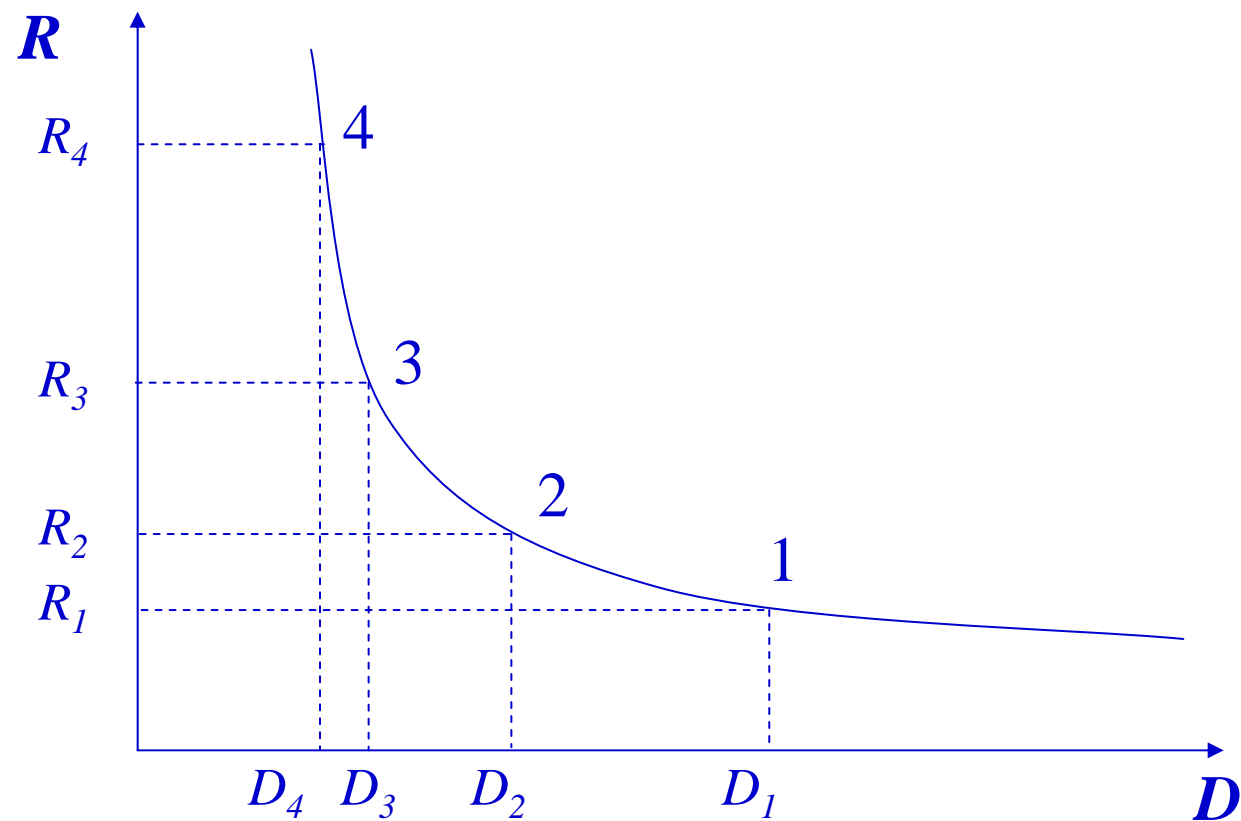 - Single Gain for All Pulses

UCSB

Key Functionalities

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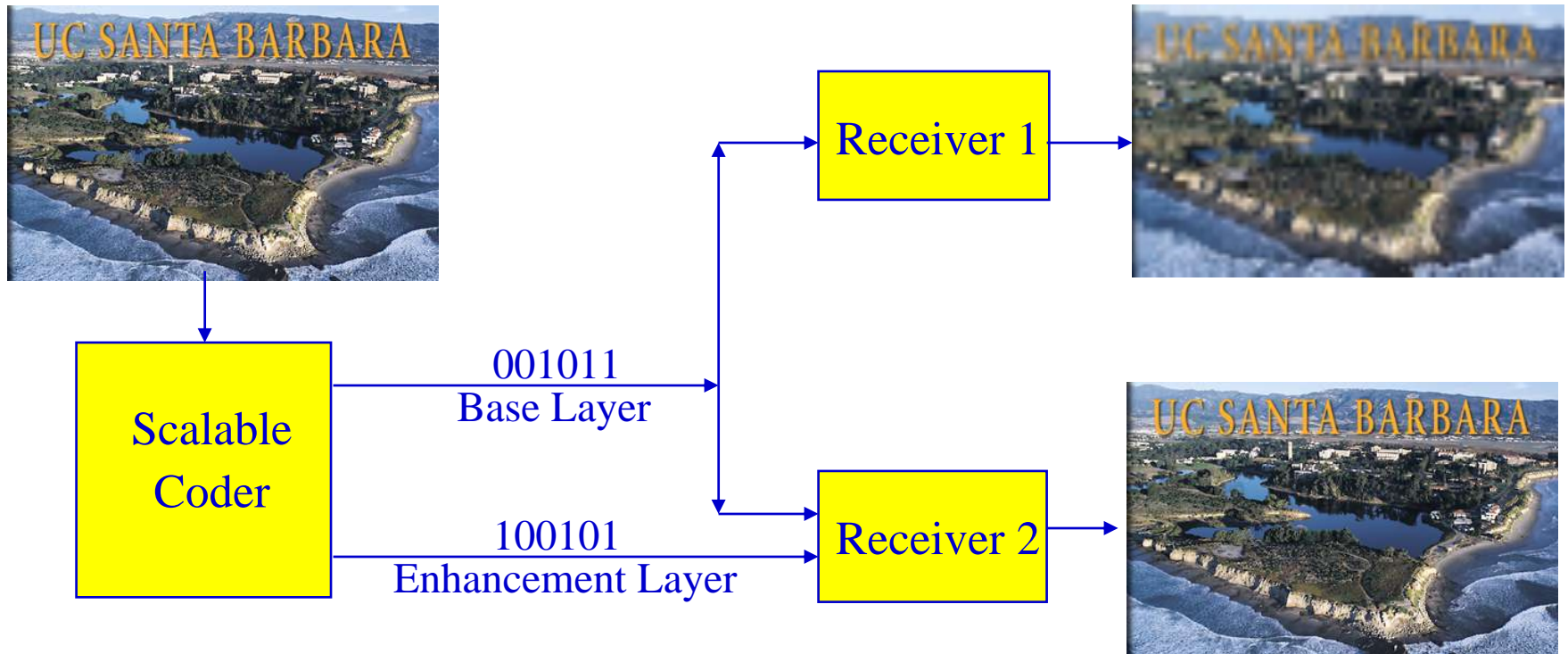 - Temporal
  - Bandwidth
- Multiple Descriptions

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
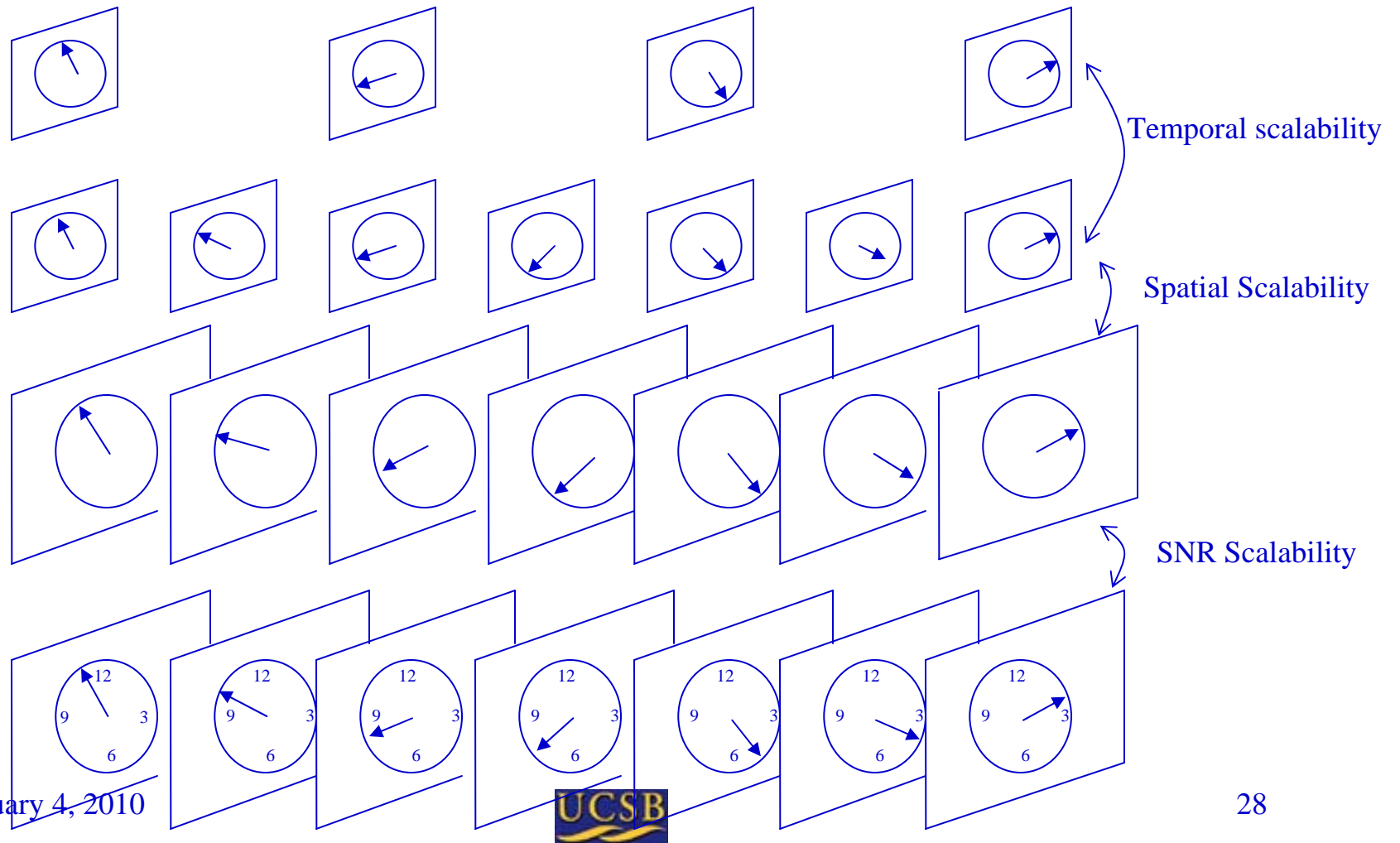
The Successive Refinement Problem

Scalable Coding

Receiver 1

001011
Base Layer

100101
Enhancement Layer

Scalable Coder

Receiver 2

UCSB

Temporal, Spatial and SNR scalability

Temporal scalability

Spatial Scalability

SNR Scalability

UCSB

Multiple Descriptions Coding

001011
Packet 1

MD Encoder

Side Decoder 1

Central Decoder

100101
Packet 2

Side Decoder 2