INT 307
Multimedia Security System

Summary and Review

Sichen.Liu@xjtlu.edu.cn

XJTLU | SCHOOL OF FILM AND TV ARTS

Xi'an Jiaotong-Liverpool University
西交利物浦大學

Official learning outcomes

- Demonstrate practical knowledge of multimedia systems and security technologies
- Demonstrate knowledge of multimedia compression technologies and standards
- Evaluate algorithms, theories and tools developed for multimedia security issues, including digital rights management, copyright protection, and authenticity verification
- Demonstrate an awareness of theories, research issues and recent developments of multimedia-based security systems such as multimedia surveillance and biometric applications
- Recognise the security risks that may be involved in the operation of computing and information systems

Lecture Overview

- Week 1: Introduction and Overview
- Week 2: Representation and Compression - Image
- Week 3: Representation and Compression - Video
- Week 4: Representation and Compression - Audio
- Week 5: Watermarking
- Week 8: Multimedia Encryption - Classical
- Week 9: Multimedia Encryption – Data Encryption Standard
- Week 11: Neural Network and Adversarial Attack - Basic
- Week 12: Neural Network and Adversarial Attack - Attack
- Week 13: Summary and Review

Unit A

Unit B

Module Assessment

There are three assessments in this module

- Coursework 1 15%
- Coursework 2 15%
- Final exam (Closed Book) 70%

Final Exam

- Help you to pass the module (You cannot pass the module if not taken exam)
- Demonstrate you have qualified as stated by the LOs
  - Could demonstrate essential related knowledge
  - Has the ability of making essential judgements
- Go check ppts and BBB recordings to help the review
- Date: 8-Jan, 10:00-12:00 am, Room SA136

Final Exam

- Onsite closed-book exam, runs 120 mins
- Full-scale exam paper with 8 questions
- Total marks available are 100
- Answer all questions
- Relevant and clear steps should be included
- Only Casio FS82ES / 83ES calculators are allowed
- Both exam papers and answers will be collected
- PLEASE DO NOT TAKE EXAM PAPER AWAY

Exam Questions

- Three levels of difficulties
  - Easy: direct description of questions with standardised answers
  - Medium: indirect description of questions with standardised answers (students need to analyse what knowledge is associated with the description)
  - Hard: indirect description of questions with no standardised answers (students should use learned knowledge to analyse the scenario provided)

Exam Questions - Mark Distribution

	Easy	Medium	Hard	Total
Unit A	33	15	6	54
Unit B	10	23	13	46
Overall	43	38	19	100

Week 2 : Image Representation and Compression

- Image / Video Representation
  - Color Space: RGB, CMY, YUV
  - Calculation
- Compression Method
  - Entropy Coding
  - Huffman Coding
  - Arithmetic Coding
  - Run-length Coding
- JPEG Compression
  - Key steps in performing JPEG Compression
  - Lossy Compression
    - Quantization

Huffman Coding

- If the symbol probabilities are known, Huffman codewords can be automatically generated
  - Reorder in decreasing order of probability at each step
  - Merge the two lowest probability symbols at each step

1. "A_DEAD_DAD_CEDED_A_BAD_BABE_A_BEADED_ABACA_BED"

2. C: 2, B: 6, E: 7, _:10, D:10, A:11 → CB: 8 (0/ C: 2, \1 B: 6)

3. E: 7, CB: 8, _:10, D:10, A:11 → ECB:15 (0/ E: 7, \1 CB: 8 (0/ C: 2, \1 B: 6))

4. _:10, D:10, A:11, ECB:15 → _D:20 (0/ _:10, \1 D:10)

5. A:11, ECB:15, _D:20 → AECB:26 (0/ A:11, \1 ECB:15 (0/ E: 7, \1 CB: 8 (0/ C: 2, \1 B: 6)))

6. _D:20, AECB:26 → _DAECB:46 (0/ _D:20 (0/ _:10, \1 D:10), \1 AECB:26 (0/ A:11, \1 ECB:15 (0/ E: 7, \1 CB: 8 (0/ C: 2, \1 B: 6))))

7. _ : 00
   D : 01
   A : 10
   E : 110
   C : 1110
   B : 1111

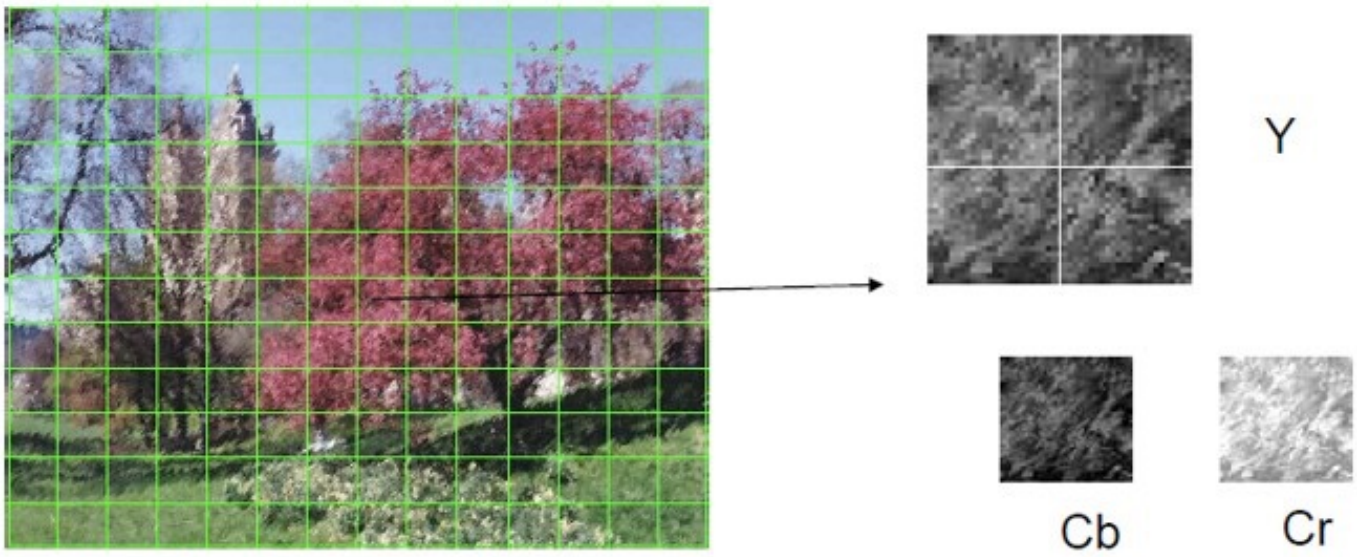
8. "1000011101001000110010011101100111001001000111110010011111011111100010001111110100111001001011111011101000111111001"

JPEG

- Divide Images into 8×8 (luminance) or 16×16 (chrominance) blocks
- Perform DCT on image blocks
- Apply Quantization
- Perform Zigzag ordering and Run-length encoding
- Entropy Encoding (Huffman Encoding)

Macroblocks

- An image is divided into 8×8 blocks for the luminance components
- An image is divided into 16×16 blocks for the chrominance components
- Chrominance blocks are down-sampled to 8×8
- Zero-Padding on boundary blocks



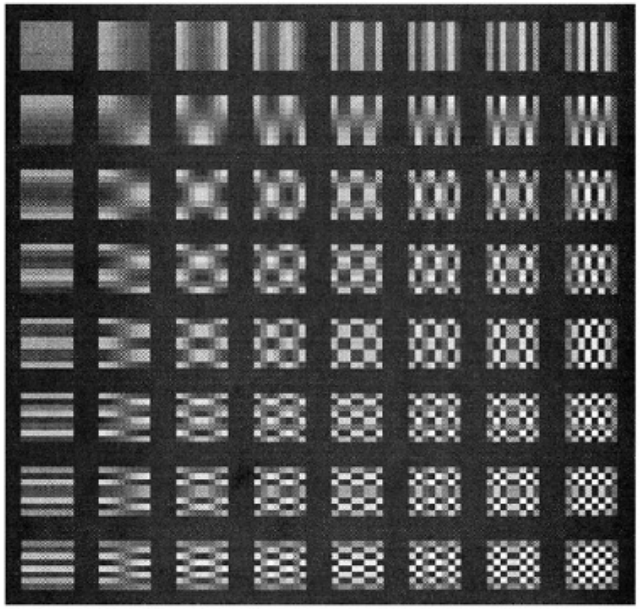
Discrete Frequency-Domain Analysis

$$F(\mu,\nu) = \frac{C(\mu)}{2}\frac{C(\nu)}{2}\sum_{y=0}^{7}\sum_{x=0}^{7} f(x,y)\cos[(2x+1)\mu\pi/16]\cos[(2x+1)\nu\pi/16]$$

$$C(\mu) = \begin{cases} \frac{1}{\sqrt{2}}, if\,\mu = 0 \\ 1, if\,\mu > 0 \end{cases}$$

- Frequency of image
- Rapid changes → High frequency
- Minor changes → Low frequency
- Usually the image is low frequency (redundancy in spatial)

2D-DCT bases



Quantization

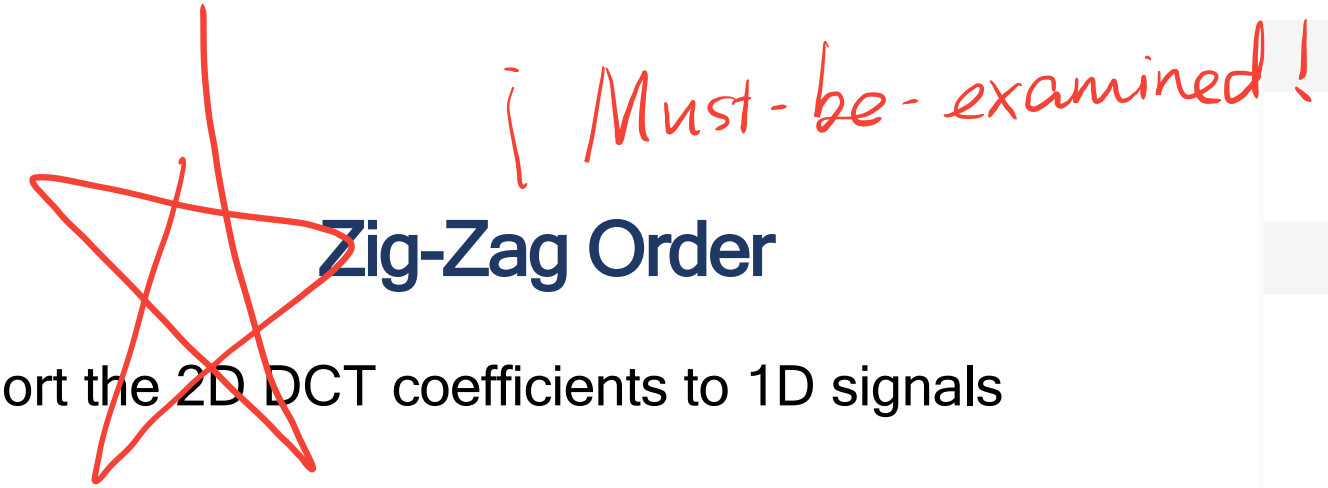
- Luminance Quantization Table and Chrominance Quantization Table

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)

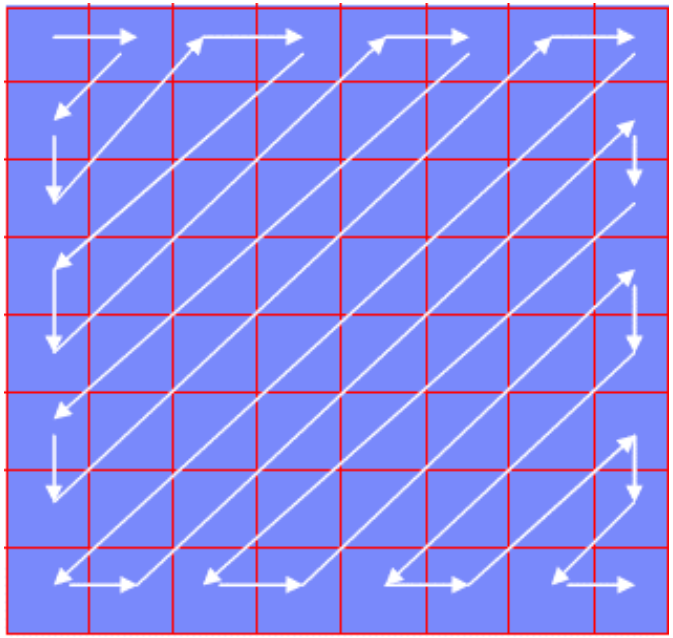
17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

(b)

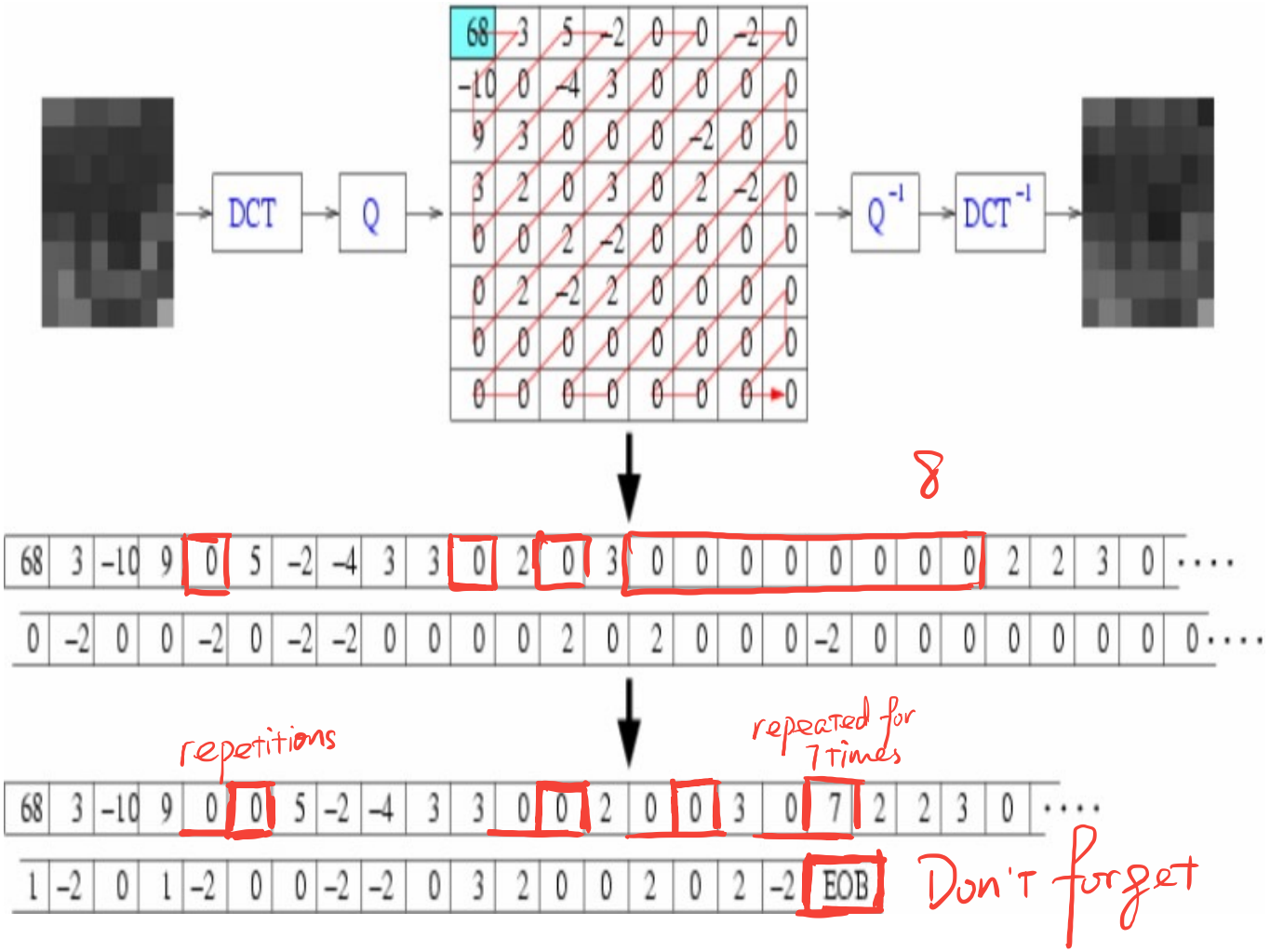


Zig-Zag Order

An order to sort the 2D DCT coefficients to 1D signals



Run Length Coding



Week 3 : Video Representation and Compression

- Video Compression
- Motion Vector Search
  - Full Search Algorithm
  - Two-dimensional Logarithmic Search
  - Three Step Search
- MPEG Compression
  - I-frame, P-frame, B-frame
  - Display order and Encoding order
  - Why display order different with Encoding order?

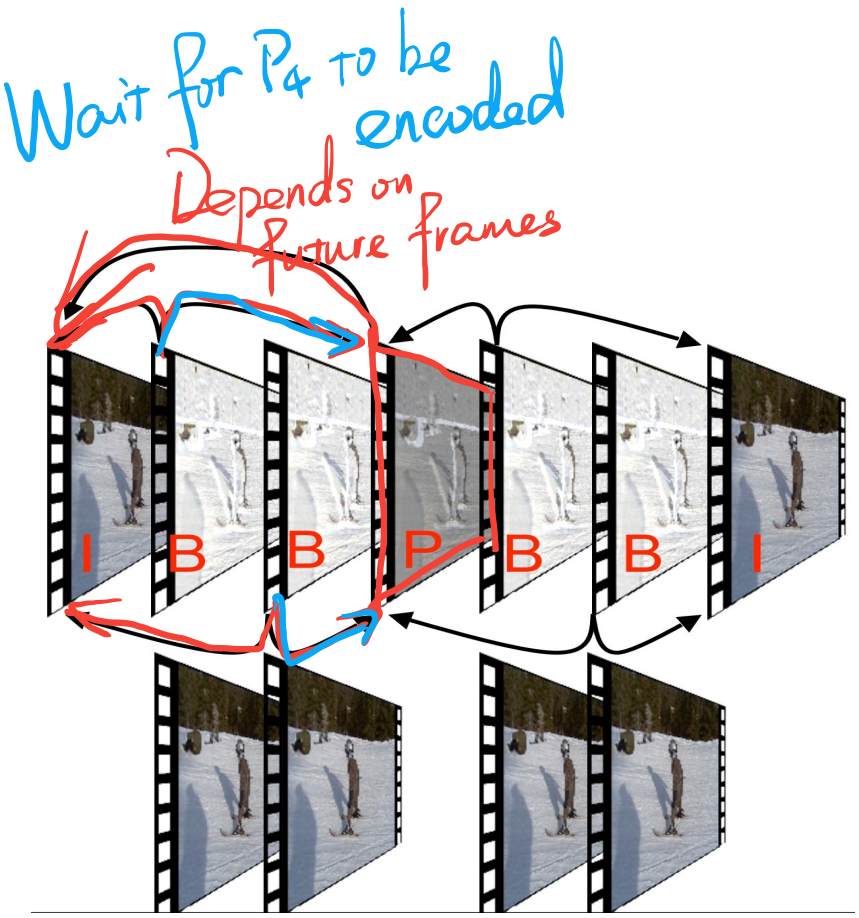
MPEG-1 Frame Types

- Three frame types in MPEG
  - I-frames, Intra-coded frames, Basic reference frame
  - P-frames, Predictive frames, Coded forward difference frame
  - B-frames, Bi-directional predictive frames, Coded bi-directional frame

Contains the contents (independently)

Relies on previous & Future frames

Encoding Order



Wait for P4 to be encoded

Depends on future frames

Encoding order

I1 P4 B2 B3 I7 B5 B6

1. Encode I-frame 1
2. Store frame 2
3. Store frame 3

Since they can't be encoded immediately

4. Encode P frame 4 ✓ (can be encoded immediately)
5. Encode B frame 2
6. Encode B frame 3

Condition of B frame is satisfied ⇒ Encode

7. Store frame 5
8. Store frame 6

Condition not satisfied, store first (can't be encoded)

9. Encode I frame 7 ✓
10. Encode B frame 5
11. Encode B frame 6

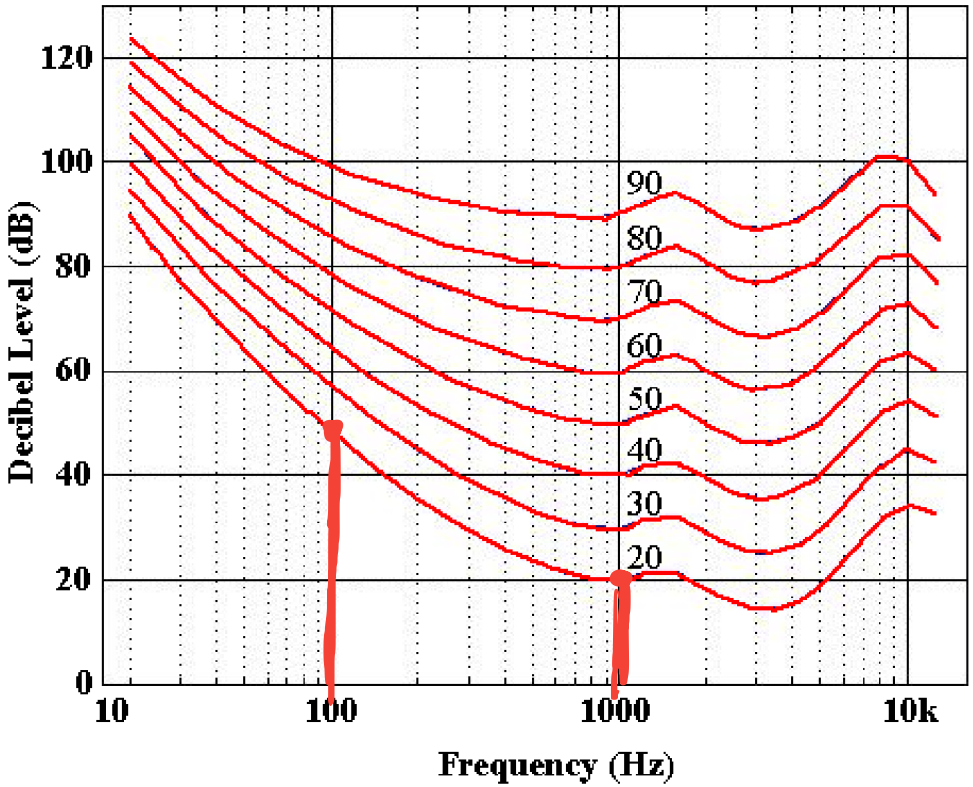
Condition satisfied Encode

Week 4 : Audio Representation and Compression

- Audio Representation
  - Sampling and Quantization
  - Solve Aliasing
  - Calculation
- Audio Perception
  - Equal-Loudness Relations
  - Frequency Masking
  - Temporal Masking
- Audio Compression
  - MPEG Audio Strategy

✳ Principle of multimedia compression:
Remove the difference between human visual system & computer storage system.

Equal-Loudness Relations



- The ear's perception of equal loudness
- The bottom curve shows what level of pure sound stimulus is required to produce the perception of a 10 dB sound
- All the curves are arranged so that the perceived loudness level gives the same loudness as for that loudness level of a pure sound at 1 kHz

Property ①

Frequency Masking

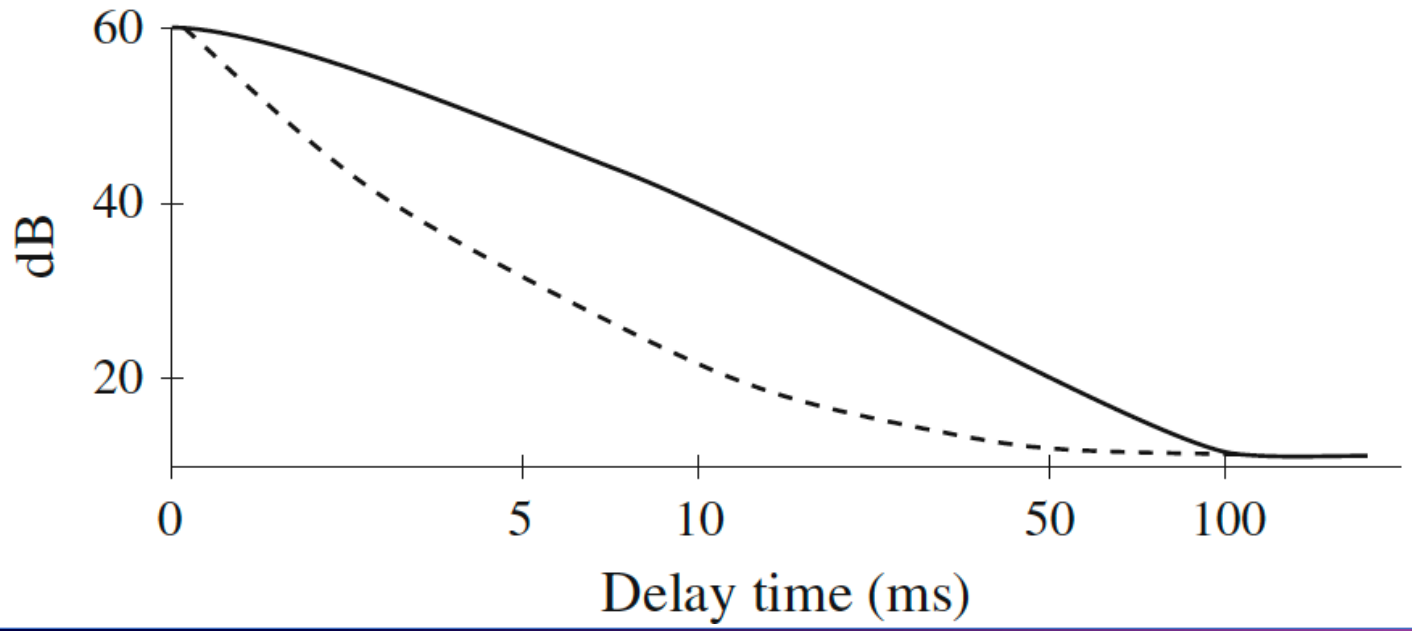
- Lossy audio data compression methods, such as MPEG/Audio encoding, remove some sounds which are masked anyway
- The general situation regarding masking is as follows:
  - A lower sound can effectively mask (make us unable to hear) a higher sound
  - The reverse is not true – a higher sound does not mask a lower sound well
  - The greater the power in the masking sound, the wider is its influence – the broader the range of frequencies it can mask
  - Therefore, if two sound are widely separated in frequency then little masking occurs

Frequency masking

Property ②

Temporal Masking

- Phenomenon: any loud sound will cause the hearing receptors in the inner ear to become saturated and require time to recover
- For a masking sound that is played for a longer time, it takes longer before a test sound can be heard



Calculation Question

- Suppose we have a piece of audio lasting for 1 hour with sampling rate of 44.1 kHz. How many bits are needed to record the audio with 16-bit depth?

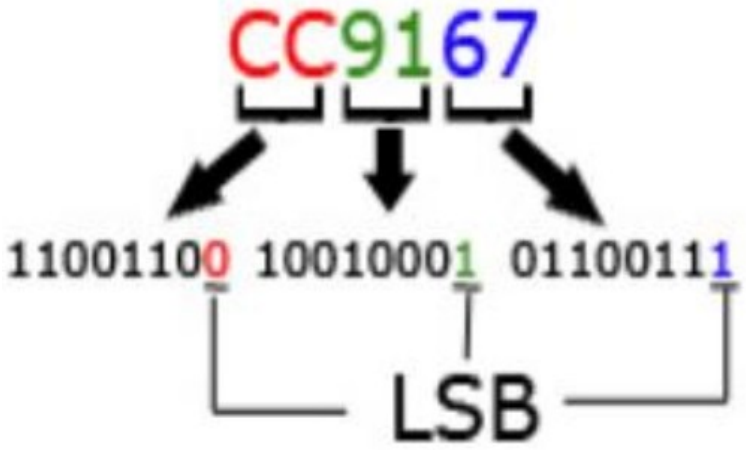
$$\underbrace{60 \times 60}_{\text{seconds}} \times \underbrace{44.1 \times 10^3}_{\text{Hz}} \times \underbrace{16}_{\text{bit/frame}} \cong 2.54 \times 10^9 \text{ bits}$$

Week 5 : Watermarking

- Digital Multimedia Forensics
  - Hash Function
- Types of Watermarking
  - Visible vs Invisible, Robust vs Fragile
- Least Significant Bits Watermarking
  - How LSB works
  - Advantage and Disadvantage
  - Performance Evaluation, Robustness
- Spread Spectrum Watermarking
- Attacks of Watermarking

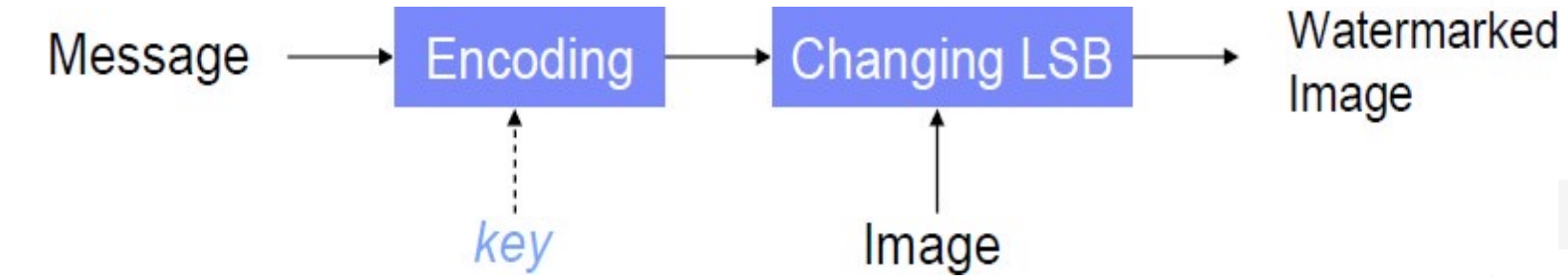
Simplest Watermark – Changing Least Significant Bits

- LSB are bits which if modified will not significantly affect the colors produced by the combination of the three RGB color components



i Must-be-examined!

Simplest Watermark – Changing Least Significant Bits



158 165 158 158 158 158 155 158
155 161 155 150 155 155 158 151
150 155 155 150 161 161 174 174
171 167 171 159 151 151 134 117
96 94 90 102 102 108 108 101
108 108 96 108 108 108 108 108
108 108 103 108 108 120 110 117
110 117 117 123 125 129 124 127

wmk

Set as the last bits

10001001

...1..0..0..0..1..0..0..1

159 164 158 158 159 158 154 159
155 160 154 150 155 154 158 151
151 154 154 150 161 160 174 175

.....

Properties and Evaluation

Thus can't be discovered by the human visual system

- Imperceptible (modify only LSBs)
- Not robust (e.g., randomly set LSBs to 0 or 1)
- More accurate: secure info-hiding method
- Robustness
  - Crop
  - Resize
  - Rotation
  - Mirror
  - Compression
  - Noise

} Ways to attack the Watermarking system

Week 8 : Multimedia Encryption - Classical

- Basics of Encryption
- Types of Encryption
  - Symmetric and Asymmetric
- Classical Cipher
  - Substitution and Transposition
  - Caesar Cipher
  - Monoalphabetic Cipher
  - Playfair Cipher
  - Polyalphabetic Cipher
  - Vigenère Cipher
  - Rail Fence Transposition
  - Rows/Columns Transposition

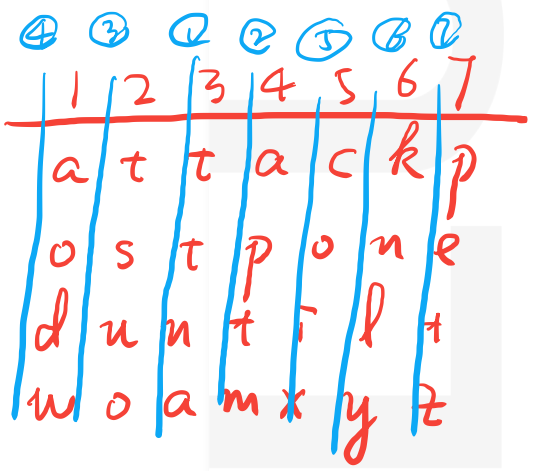
Rows/Columns Transposition

- Plaintext letters written in rows
- Ciphertext obtained by reading column-by-column, but re-arranged
- Key determines order of columns to read
- Easy to break using letter frequency (try different column orders)
- Example

```
plaintext: securityandcryptography
key: 315624
```

Column size is decided by key size.

Rows/Columns Transposition



- Transposition ciphers can be made stronger by using multiple stages of transposition

```
plaintext:  attackpostponeduntiltwoamxyz
key: 3421567
ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ
```

(length = num. columns = 7)

3 4 2 1 5 6 7

- Transpose again using same key

```
output:     NSCYAUOPTTWLTMDNAOIEPAXTTOKZ
```

Original plaintext letters, by position:

```
01 02 03 04 05 06 07 08 09 10 11 12 13 14
15 16 17 18 19 20 21 22 23 24 25 26 27 28
```

After first transposition:

```
03 10 17 24 04 11 18 25 02 09 16 23 01 08
15 22 05 12 19 26 06 13 20 27 07 14 21 28
```

After second transposition:

```
17 09 05 27 24 16 12 07 10 02 22 20 03 25
15 13 04 23 19 14 11 01 26 21 18 08 06 28
```

Week 9: Multimedia Encryption - DES

¡ Must-be-examined!

- Data Encryption Standard (DES)
  - Feistel Encryption Structure
    - Major Concerns in Feistel Encryption Structure
    - Key Steps of Encryption Branch
      - F-function: E-table, S-box, P-table
    - Key Steps of Key Generation Branch
      - Permuted Choice 1, Permuted Choice 2
- Triple DES
- Joint Encryption and Compression

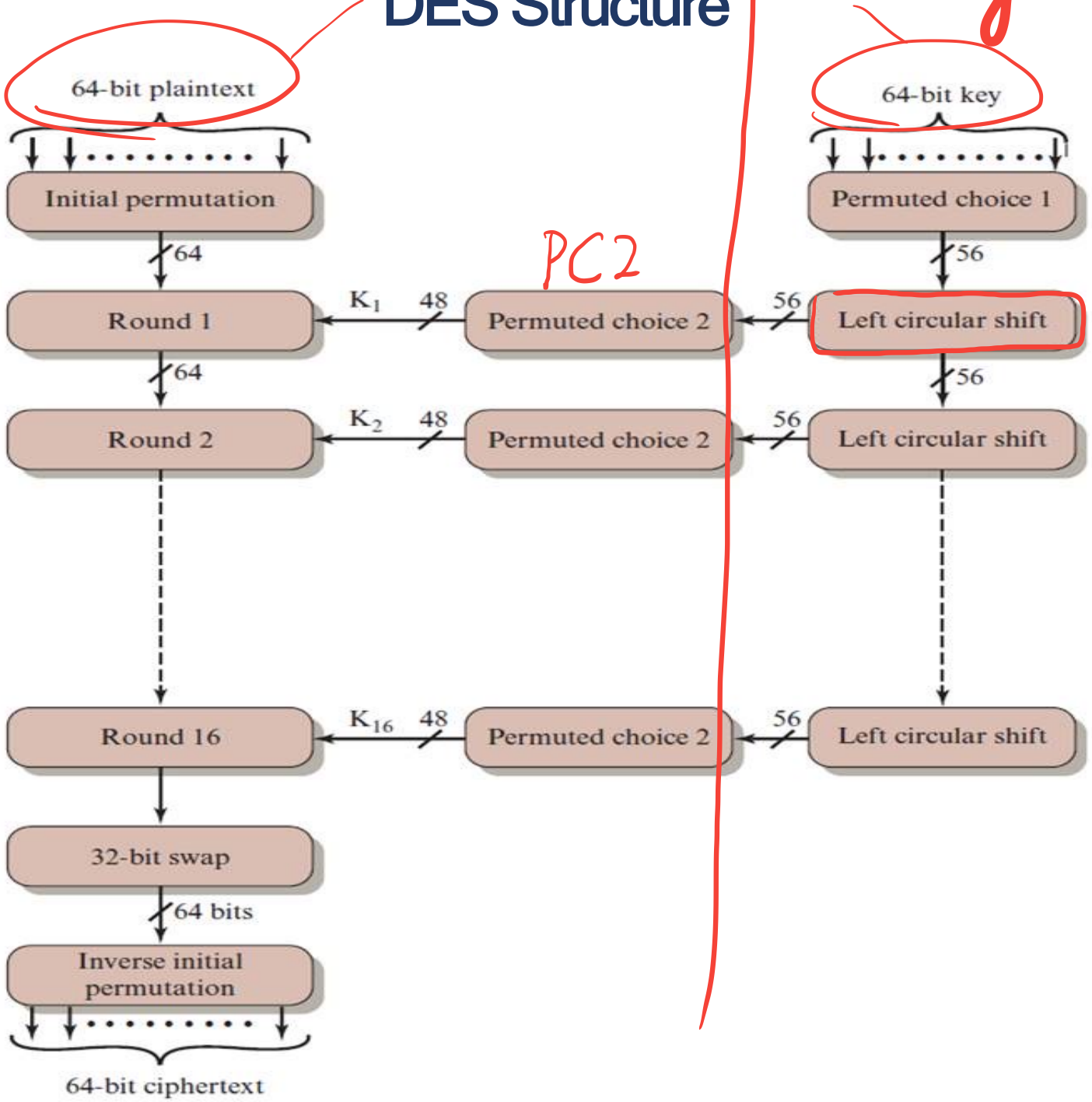
{ Encryption, Key Generation

Encryption

Initialization: 64 bits

Key Generation

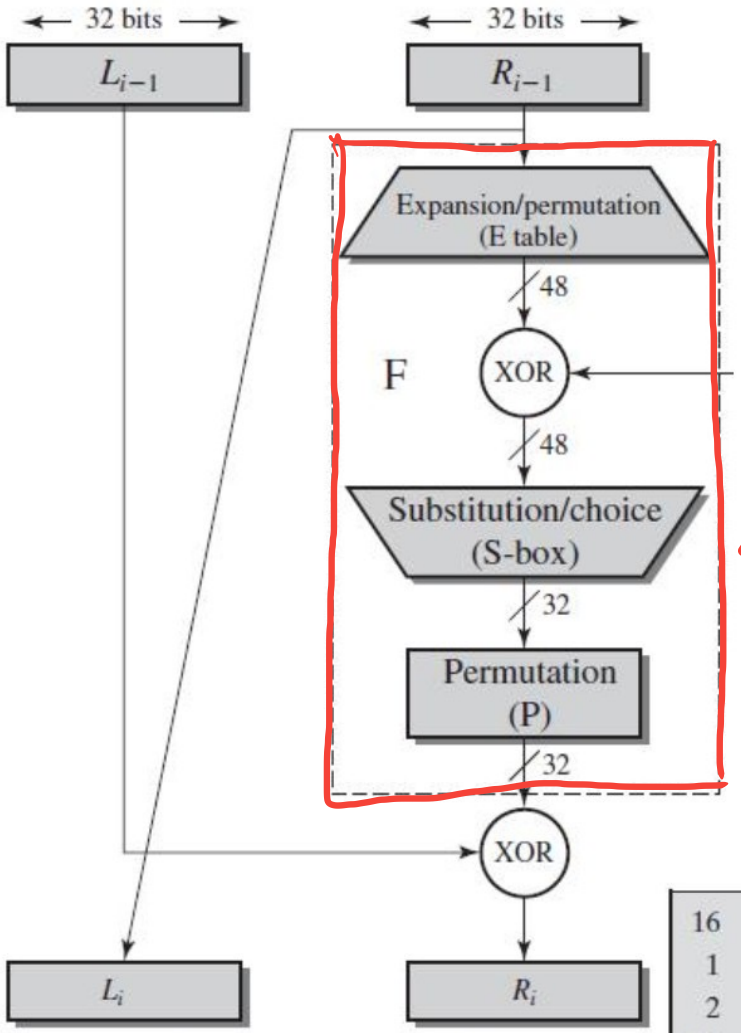
DES Structure



DES Expansion (E) and Permutation (P)

32 bits → 48 bits

Fiestal Structure.



48 → 32

(c) Expansion Permutation (E)

32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

(d) Permutation Function (P)

16	7	20	21	29	12	28	17
1	15	23	26	5	18	31	10
2	8	24	14	32	27	3	9
19	13	30	6	22	11	4	25

Figure 3.6 Single Round of DES Algorithm

Input 6-bit, output 4-bit, based on the predefined S-Box

Makes it more difficult for the attackers to decrypt the ciphertext

S-Box

- The first and last bits of the input to box form a 2-bit binary number to select one of four substitutions defined by the four rows in the table for S_i.

Row index

Column

Row

1100

011001

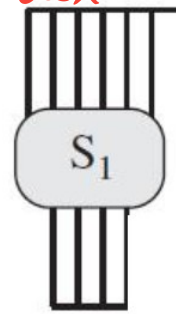
6 bites in

Column index

	0	1	2	...	5	...	10	11	12				
0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13

S_1

01 →



011001 1001

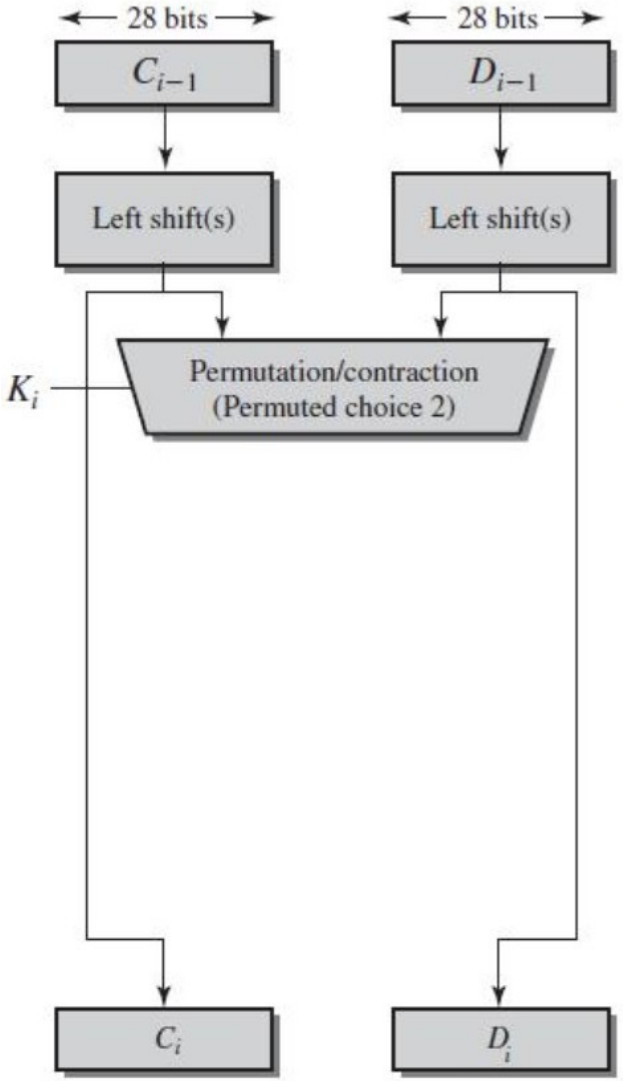
6 → 4

1001

4 bites out

- The middle four bits select one of the sixteen columns. The decimal value in the cell selected by the row and column is then converted to its 4-bit representation to produce the output.
- For example, in S1, for input 011001, the row is 01 (row 1) and the column is 1100 (column 12).The value in row 1, column 12 is 9, so the output is 1001.

Key Generation for Each Round



(a) Input Key

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64

Drop the last column
64 → 56

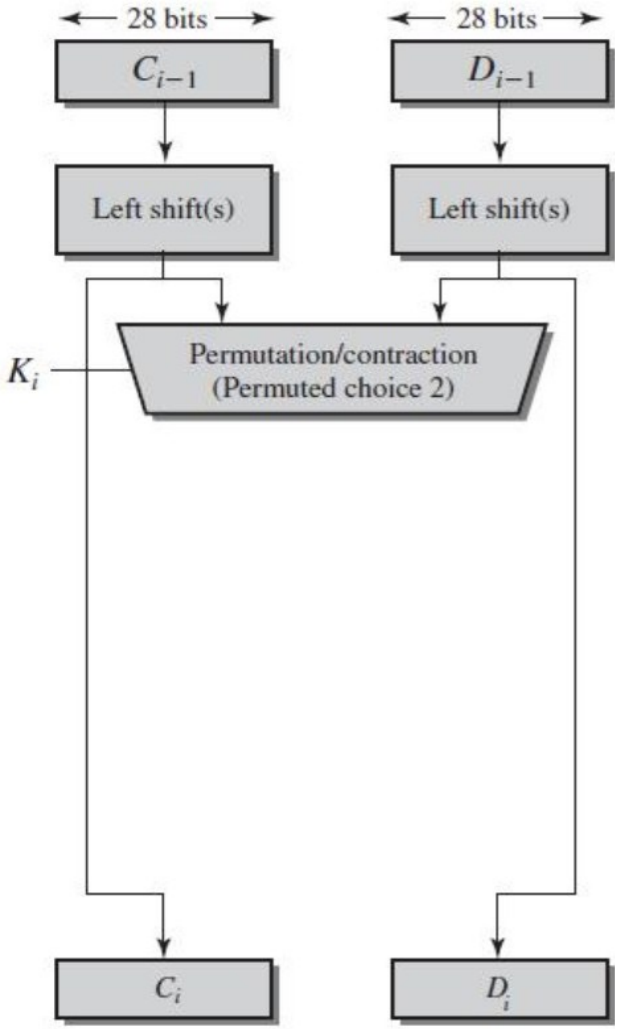
C_0

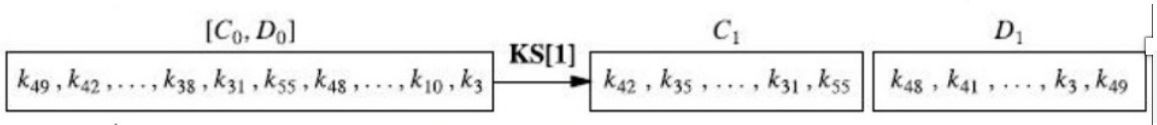
(b) Permuted Choice One (PC-1)

57	49	41	33	25	17	9
1	58	50	42	34	26	18
10	2	59	51	43	35	27
19	11	3	60	52	44	36
63	55	47	39	31	23	15
7	62	54	46	38	30	22
14	6	61	53	45	37	29
21	13	5	28	20	12	4

D_0

Key Generation for Each Round





(d) Schedule of Left Shifts

Round Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Bits Rotated	1	1	2	2	2	2	2	2	1	2	2	2	2	2	2	1

(c) Permuted Choice Two (PC-2)

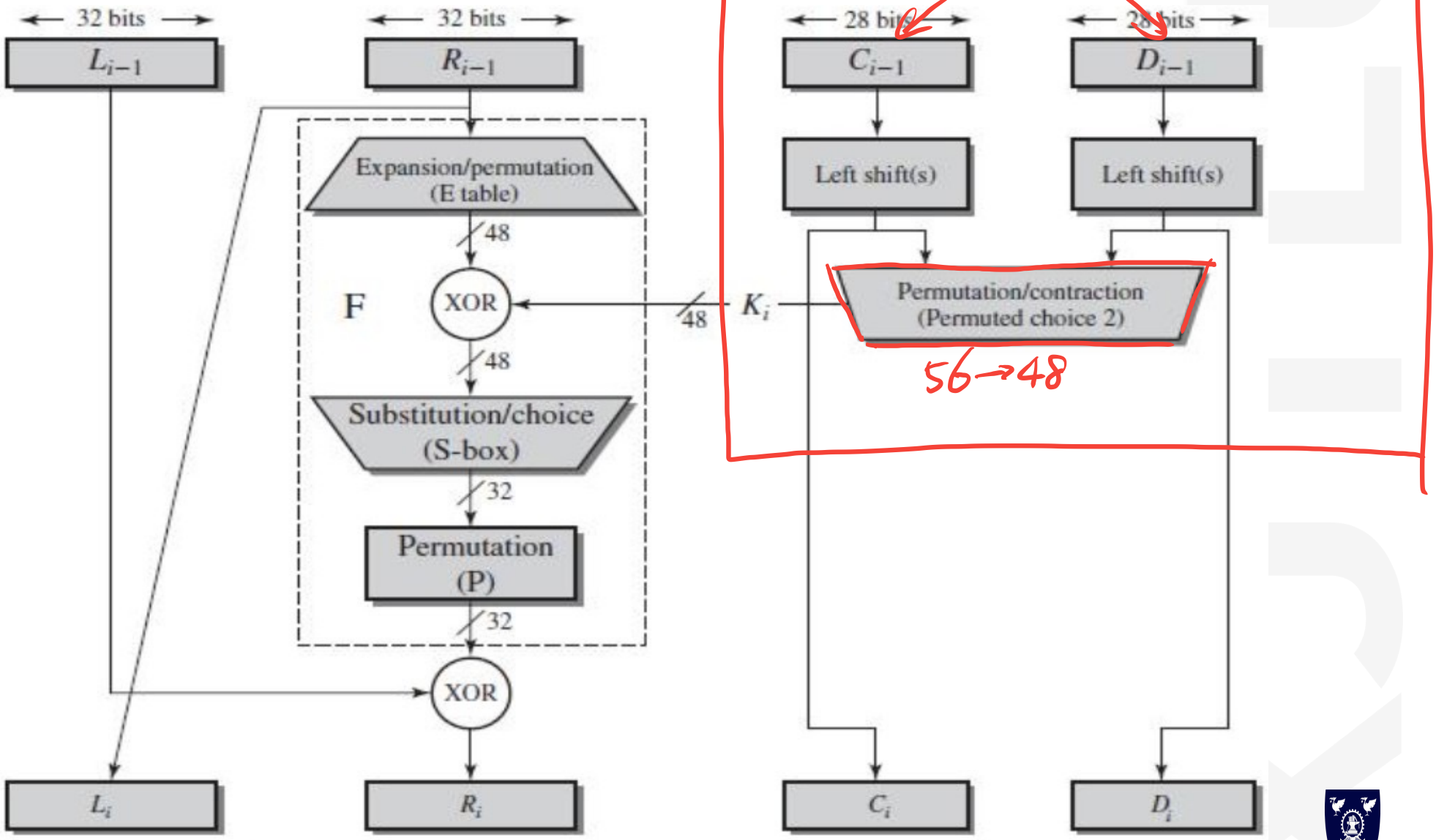
14	17	11	24	1	5	3	28
15	6	21	10	23	19	12	4
26	8	16	7	27	20	13	2
41	52	31	37	47	55	30	40
51	45	33	48	44	49	39	56
34	53	46	42	50	36	29	32

56 → 48

Drop the bits in accordance with PC-2 table.

Mark the key length at each step

Encryption in a DES Round



Week 11-12 : NN and Adversarial Attack

How do you attack an NN system?

- Basic of Neural Networks
  - Backward Propagation
  - DNN, CNN, RNN
  - Attention, Transformer
- Adversarial Attack
  - Principle
  - Adversarial Perturbation and Adversarial Patches
  - White box, Score-based Black Box, Limited Black Box
- Defending Against Adversarial Inputs
  - Ways to Improve The Model

How to improve the robustness of a deep learning system?

Theoretically Derived Robustness Metrics

No standard answer

1. Adversarial training : Train the model with adversarial samples
2. Defensive Distillation : Modified training process
3. Feature Squeezing : Reduce input dimensionality
4. Input Transformation : Apply transformation to the input data
5. Randomised dropout uncertainty measurements
6. Minimise the adversary's knowledge

MQs

- **Date: 4th December – 17th December**
- Your honest and constructive feedback will be utilized for enhancing modules and teaching practices
- Set aside 20 minutes of class-time during Week 12 or 13 to encourage students to complete the MQs.
- Inform your students that completion of all questionnaires will make them eligible for a lucky draw at the end of this semester.

THANK YOU

VISIT US

WWW.XJTLU.EDU.CN

FOLLOW US

@XJTLU

Xi'an Jiaotong-Liverpool University
西交利物浦大学

XJTLU | SCHOOL OF FILM AND TV ARTS