INT307 Multimedia Security System

Multimedia Forensics

Shengchen Li

Xi'an Jiaotong-Liverpool University

13th Sep 2022

Aims

- Understand the aim of digital forensics
- Understand the working principle of digital signature
- Understand the working principle of digital watermarking
- Understand the working principle of steganography

Aim of Digital Forensics

The application of investigating methodologies of forensics to the field computer crimes.

- Source Identification
- Integrity / Authenticity
- Enhancement
- Interpretation and Content Analysis

Types of Multimedia Forensics

- Active Multimedia Forensics
  - Known the media should be protected
  - Add embedded information in the multimedia file

- Passive Multimedia Forensics
  - Showing the metadata of the content
  - Using information retrieval methods

Means of Attack

- Disguise identity
- Tampering with content
- Modify the order
- Change the time
- Deny sending
- Deny acceptance

Active Multimedia Forensics

- Message Authentication Code (Make use of encryption algorithm)
- Hash Function
- Digital Signature
- Watermarking
  - Robust Watermarking
  - Fragile Watermarking

Hash Function

A way to demonstrate data integrity is HASH function.

- Hash function maps a variable length message to a fixed length message: $y = h(x)$
- If h is a 64-bit has function, then y always fits in 64 bits i.e. $0 \leq y < 2^{64}$
- A hash is a keyless algorithm
- Anyone can compute $h(x)$ if x is known

Example

Alice sends Bob $C = \text{Encrypt}(M), h(M)$. Bob receives $C, h(M)$ and checks

- $M' = \text{Decrypt}(C)$
- $h(M')$

Cryptographic Hash Functions

- Collision resistance: difficult to find any M, $M' \neq M$ such that $h(M) = h(M')$
- Preimage resistance: given $h(M)$, difficult to find M' such that $h(M') = h(M)$
- Second preimage resistance: given m, difficult to find M' such that $h(M') = h(M), M' \neq M$

Namely, h is secure meaning

- Easy to compute in one direction
- Very difficult to compute in the other direction (i.e. computational secure)
- Very difficult to find two messages that hash to the same value

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

HASH in Industry

- Industrial Standards: SHA-256, SHA-3, MD4, MD5

HASH application in Blockchain

- Blockchain is a decentralised dataset
- HASH function is used as one step of packaging information
- A condition is set for HASH value of the package with a help of random number
- The settings related to HASH function regulates the speed of synchronisation

Function of Digital Signature

1. The information of sender: identity of sender, date & time of signature
2. Authentication of message
3. Can be arbitrated

Arbitration digital signature

- Model 1

$$A \rightarrow Z : M, E_{K_{AZ}}[ID_A, H(M)] \quad (1)$$

$$Z \rightarrow B : E_{K_{ZB}}\{ID_A, M, E_{K_{AZ}}[ID_A, H(M), T]\} \quad (2)$$

- Model 2

$$A \rightarrow Z : ID_A, E_{K_{AB}}(M), E_{K_{AZ}}[ID_A, H(E_{K_{AB}}(M))] \quad (3)$$

$$Z \rightarrow B : E_{K_{ZB}}\{ID_A, E_{K_{AB}}(M), E_{K_{AZ}}[ID_A, H(E_{K_{AB}}(M))], T\} \quad (4)$$

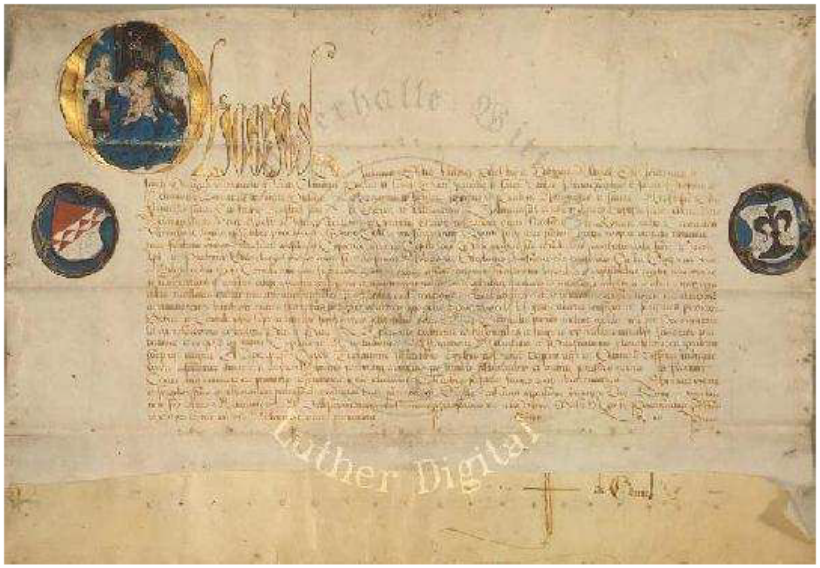
- Model 3

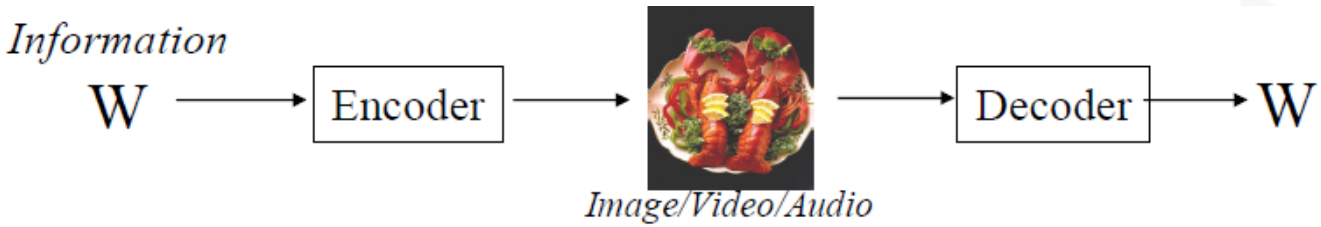
$$A \rightarrow Z : ID_A, E_{KR_A}[ID_A, E_{KU_B}(E_{KR_A}(M))] \quad (5)$$

$$Z \rightarrow B : E_{KR_Z}[ID_A, K_{KU_B}(E_{KR_A}(M)), T] \quad (6)$$

Watermarking

Embedding Visible / Invisible Codes in Multimedia Data for Security Purpose





Attributes

- Purposes
  - Protect ownership and trace illegal use
  - Authentication
  - Copy / Playback control
- Properties
  - Robustness
  - Invisibility
  - Payload
- Types
  - Visible vs Invisible
  - Robust vs Fragile
  - Referenced vs Unreferenced

Application of Digital Watermarking

Broadcasting Monitoring

- Alice is an advertiser who embeds a watermark in each of her radio commercials before distribute them to 600 radio stations.
- Alice monitors radio station broadcasts with a watermarking detector.
- She matches her logs with the 600 invoices.
- ATTACK: Bob secretly embed Alice's watermark into his own advertisement and airs it in place of Alice's commercial.

Application of Digital Watermarking

Web Reporting

- Alice owns a watermarking service that, for a nominal fee, adds an owner identification watermark to images that will be accessed through the Internet.
- Alice provides an expensive reporting service to inform her customers of all instances of their watermarked images found on the Web.
- ATTACK: Bob builds his own web crawler that detects watermarks embedded by Alice and offers a cheaper reporting service.

Application of Digital Watermarking

Copy Control

- Alice owns a movie studio, and she embeds a copy control watermark in her movies before they are distributed.
- She trusts that two digital recorders capable of copying these movies contain watermark detectors and will refuse to copy her movie.
- ATTACK: Bob is a video pirate who has a device designed to remove the copy protection watermark.

Exemplar Operation Table for Digital Watermarking

	Embed	Detect	Remove
Broadcast Monitoring			
Advertiser	Y	Y	-
Broadcaster	N	N	-
Public	N	N	-
Web Reporting			
Marking Service	Y	Y	-
Reporting Service	-	Y	-
Public	N	N	N
Copy Control			
Content Provider	Y	Y	-
Public	-	Y	N

Y: must be allowed, N: must not be allowed, - : don't care

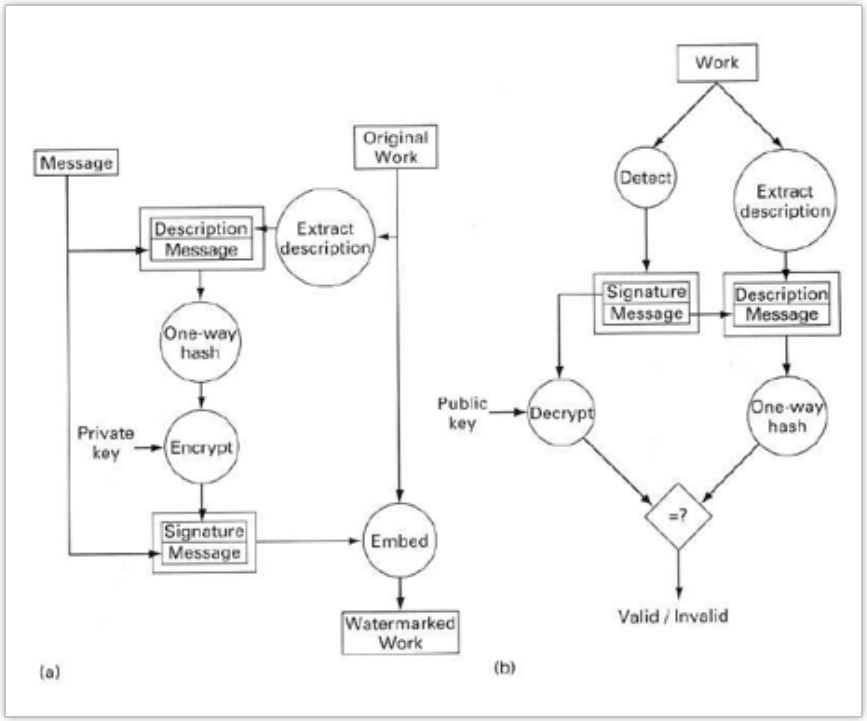
Categories of Attack

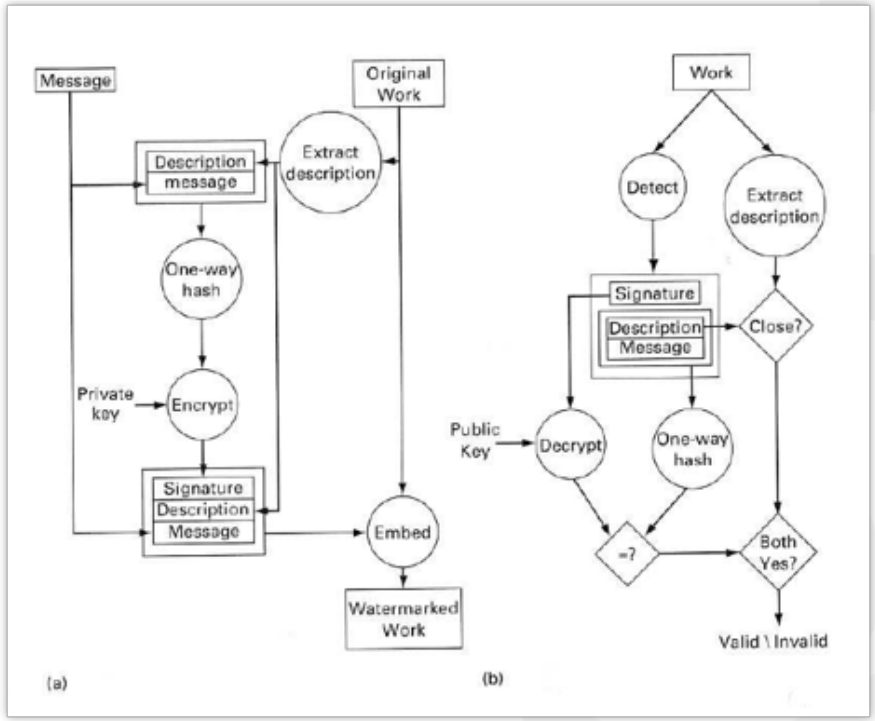
Unauthorised Embedding

- Being able to composing and embedding an original message.
  - Another example, in Scenario 2, Alice charges for embedding and gives away the monitoring tool
  - Possible Solution: using standard cryptographic techniques.

- Being able to obtain a pre-composed legitimate message and embeds this message in a Work
  - e.g., in Scenario 1, Bob extract the reference pattern and then use it to his work – called copy attack
  - Possible Solution: using content-related watermarks

Methods to Prevent Unauthorized Embedding

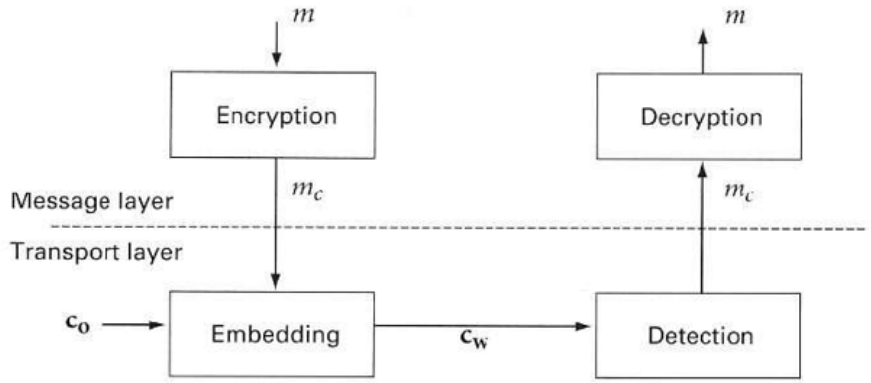




Categories of Attack

Unauthorized Detection

- A hospital might embed the names of patients into their X-rays
- Knowing whether a watermark is present (Steganography Analysis)
- Intervention on the transmission process



❑ Encryption and Decryption can be used.

Categories of Attack

Unauthorized Removal

- Attackers try to modify the watermarked work such that it resembles the original and yet does not trigger the detector

- Possible ways of attack:
  - Elimination attacks, where watermark is truly gone
  - Masking attacks, where watermark is still present but weakened

Categories of Attack

System-Level Attacks

- Attackers exploit the weakness in how the watermarks are used
- For instance, in a copy-control application, an attacker might open the recorder and just remove the chip
- Forge Identification

How to embed information non-perceivably

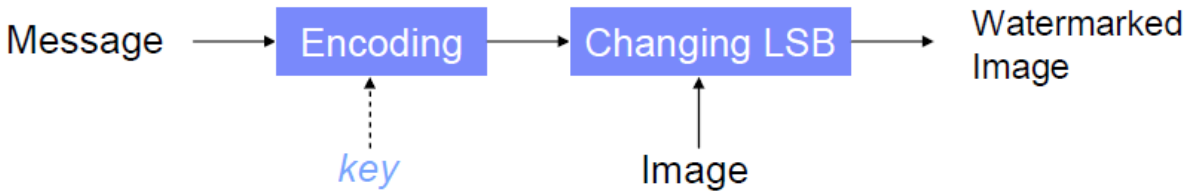
- The way to document media is named documentary space
- The media perceived is named as perceptual space
- The difference between documentary space and perceptual space can be used to code information

NOTE

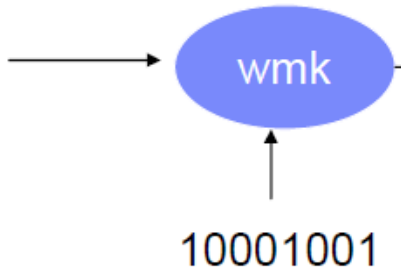
Primary target of Steganography is that the fact of carrying information is not observed

Simplest Steganography

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



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

.....

LSB Result

(a) Original Image

(b) Watermark

(c) Image with embedded watermark

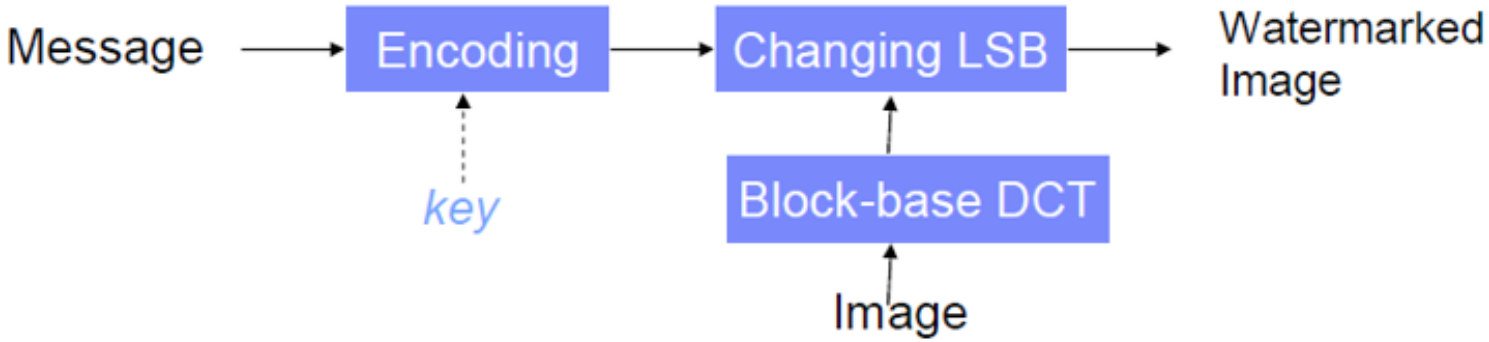
(d) Extracted watermark

(e) Difference between (a) and (c)

Discussion

- Suppose we have a system whose LSB can only be 0. A LSB change of 1 could contribute to encoding information.
- What is the potential problem with the data distribution (for each pixel)
- What is the problem with the difference between pixels?

LSB in frequency domain



- Embed one bit at one DCT coefficient
- Extension 1: embed one bit at one DCT coefficient after quantization
- Extension 2: embed one bit per DCT block

Evaluation of System Performance

- Transparency: PSNR & SNR
- Capacity: the ratio between the size of original media and the size of hidden information carried
- Payload: with the hidden information, the file size increase over original size
- Robustness
  - Crop
  - Resize
  - Rotation
  - Mirror
  - Compression
  - Noise

Recent Advances in Steganography

- Steganalysis
  - The introduction of deep learning makes the power of steganalysis much stronger
  - Challenge: negative samples only

- Generative Steganography (No Cover Media)
  - Synthesis media with embedded information
  - Statistical based steganalysis method does not work