Information Hiding: Watermarking and Steganography

Content may be borrowed from other resources.
See the last slide for acknowledgements!

Principles of assurance: CIA

- Confidentiality
  - Keeping data and resources hidden
- Integrity
  - Data integrity (integrity)
  - Origin integrity (authentication)
- Availability
  - Enabling access to data and resources

Confidentiality

- Encryption is the main tool
  - But, sometimes it is not enough!
- Examples:
  - Drug dealer
  - Military scenario
  - Whistleblower
  - Etc.

Information Hiding

- IH dates back to ancient Greece and Persia

- Well, not this type of hiding

(Digital) Information Hiding

- Definition: Concealing the very existence of some kind of information (e.g., a series of data bits, the identity of the communicating party, etc.) for some specific purpose (e.g., to prove ownership, to remain untraceable, etc.)

Information confidentiality protection

- IH does not intend to hide the contents of information
  - Encryption
- Information confidentiality protection tools can be combined with information hiding techniques

The Need for Data Hiding

- Covert communication using images (secret message is hidden in a carrier image)→ improve confidentiality
- Ownership of digital images, authentication, copyright
- Data integrity, fraud detection, self-correcting images
- Traitor-tracing (fingerprinting video-tapes)
- Adding captions to images, additional information, such as subtitles, to video, embedding subtitles or audio tracks to video (video-in-video)
- Intelligent browsers, automatic copyright information,
- Copy control (secondary protection for DVD)
- forensic

Classes of Information Hiding

- Steganography
- Digital watermarking
- Covert channels
- Anonymous communication
- Protocol obfuscation

- Not a class of information hiding:
  - Encryption
- There are various classifications

Steganography

- Embedding some information (stegotext) within a digital media (covertext) so that the digital media looks unchanged (imperceptible) to a human/machine

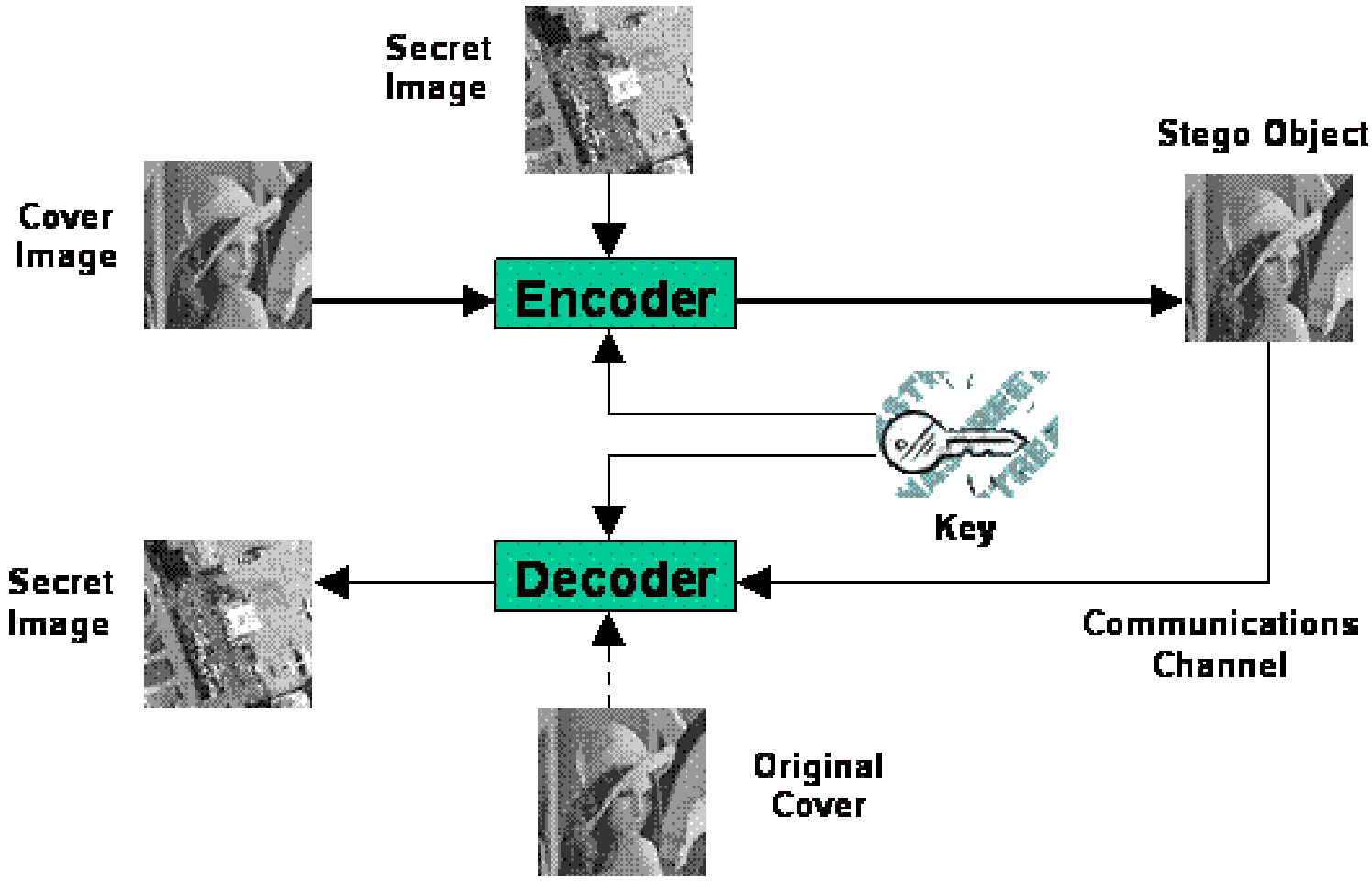
Word origin

From Greek ***Steganos*** (covered) and ***graphia*** (writing)

Steganography is sometimes called

- Secret writing
- Concealed writing
- Covert communication
- Stealth communication
- Data hiding
- Electronic invisible ink
- The prisoners' problem

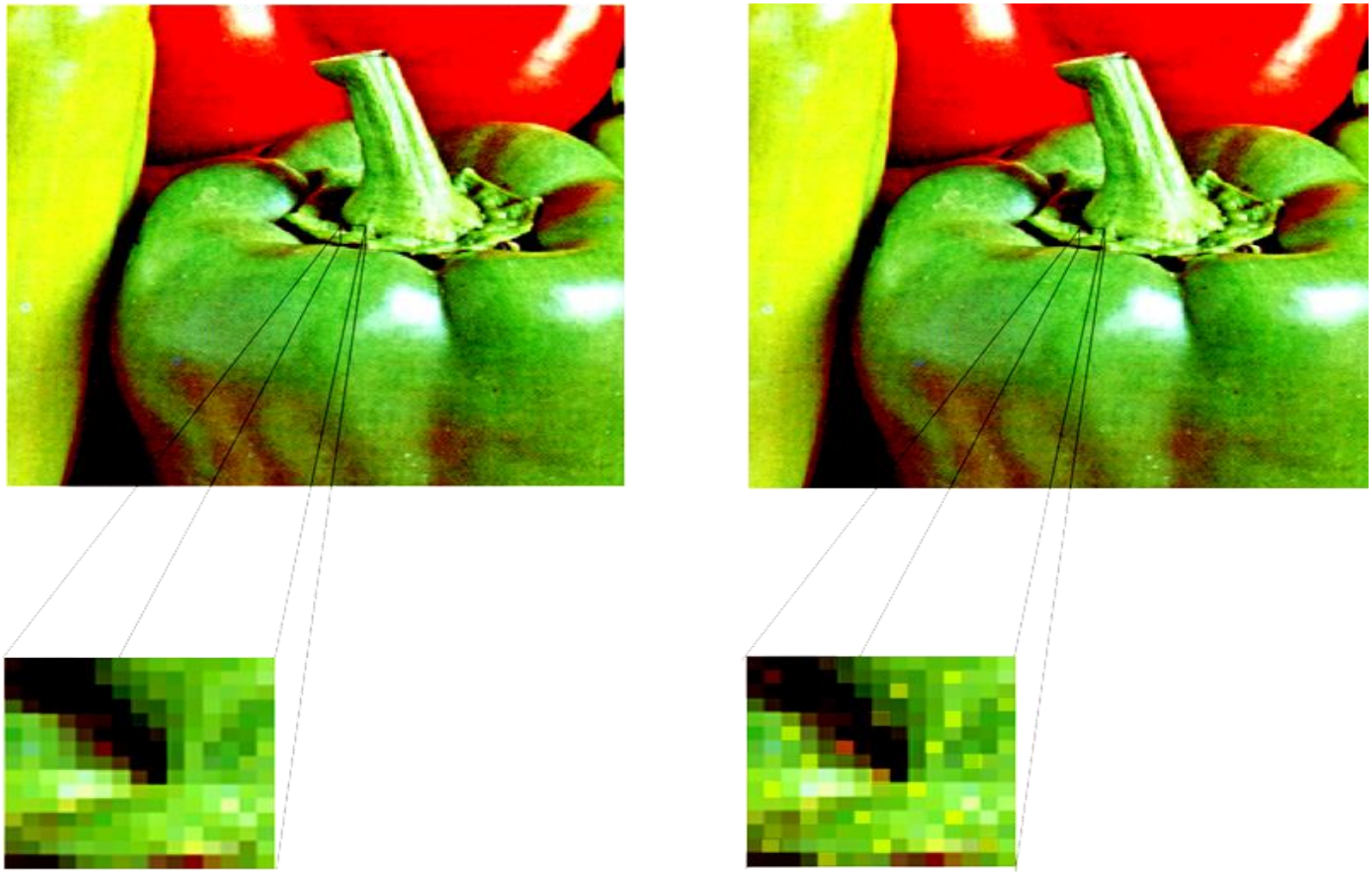
General model



Why can we hide?

- Because there are unused/redundant data bits in digital media, that changing them will be imperceptible
  - E.g., image compression significantly reduces the size of an image by removing some of the redundant information bits.
- The unused/redundant data can be used to hide some digital information

Example



Types of Cover Media

- We can hide information in pretty much any digital media:
  - Audio
  - Image
  - Video
  - Graphics
  - Text files
  - Software
  - Digital events (e.g., timings)
  - Network traffic

Classes of Information Hiding

- Steganography
- Digital watermarking
- Covert channels
- Anonymous communication
- Protocol obfuscation

- Not a class of information hiding:
  - Encryption
- There are various classifications

Digital watermarking

- Embedding some information (watermark) within a digital media (covertext) so that the digital media looks unchanged (imperceptible) to a human/machine

Watermarking vs. Steganography

- Watermarking:
  - The hidden information itself is not important by itself (no secure), it says something about the covertext
- Steganography:
  - The covertext has no value, it is only there to convey the stegotext. Stegotext is the valuable information, and is independent of covertext.

Applications: Watermarking vs. Steganography

- Watermarking:
  - Authenticity: proof of ownership
  - Fingerprinting: piracy tracking
  - Integrity: tamper detection
  - Data augmentation: add meta-data

- Steganography
  - Stealthy communication of messages
- Watermarking usually needs lower data capacity

Attacks: Watermarking vs. Steganography

- Attacker's objective:
  - Watermarking: remove the watermark without distorting the covertext, or change the covertext without distoring the watermark
  - Steganography: detect the presence of the hidden message, and extract it

Types of watermarking/steganography

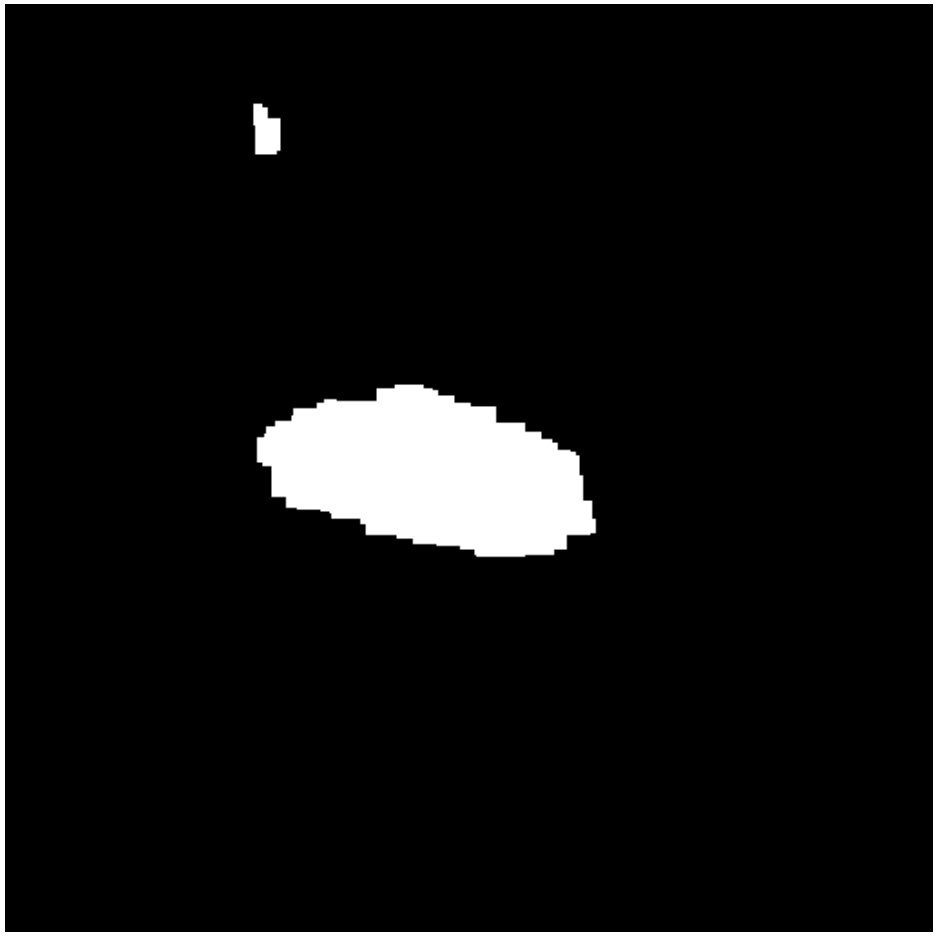
- Fragile vs. robust
  - Fragile is expected to destroy with modifications. Robust is expected to survive noise.

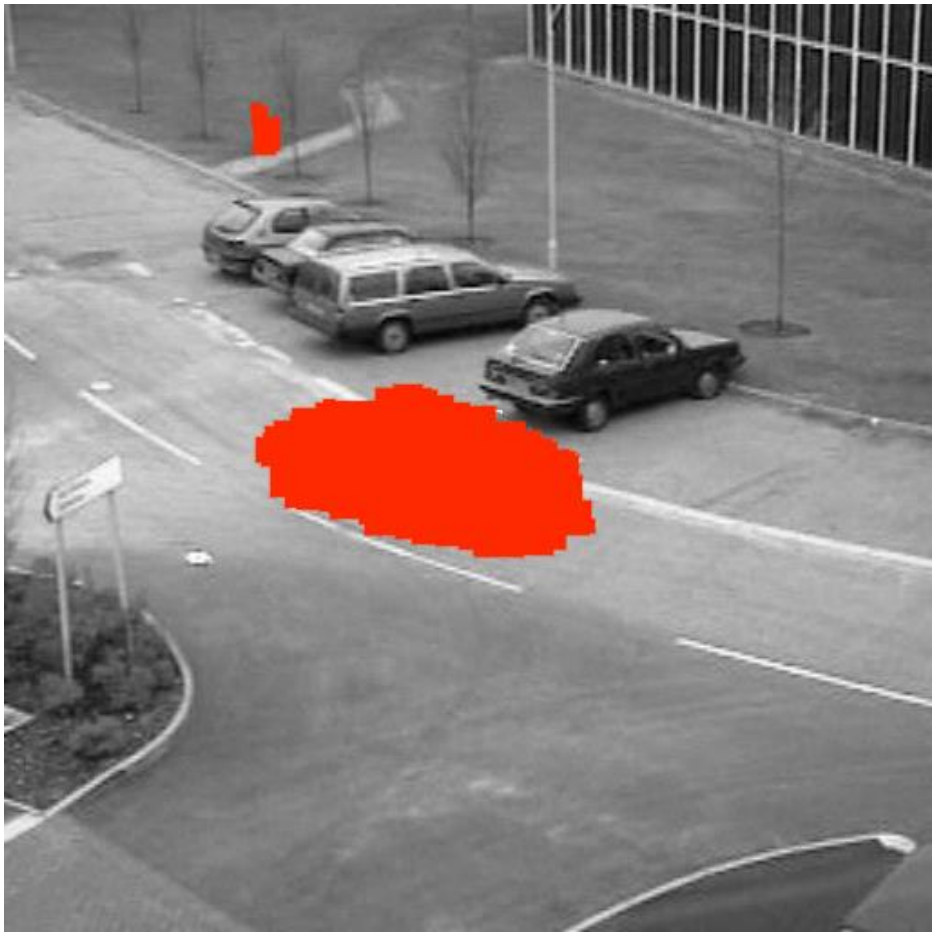
Example application: tamper detection

Tampered Image

Original Image

Example application: tamper detection

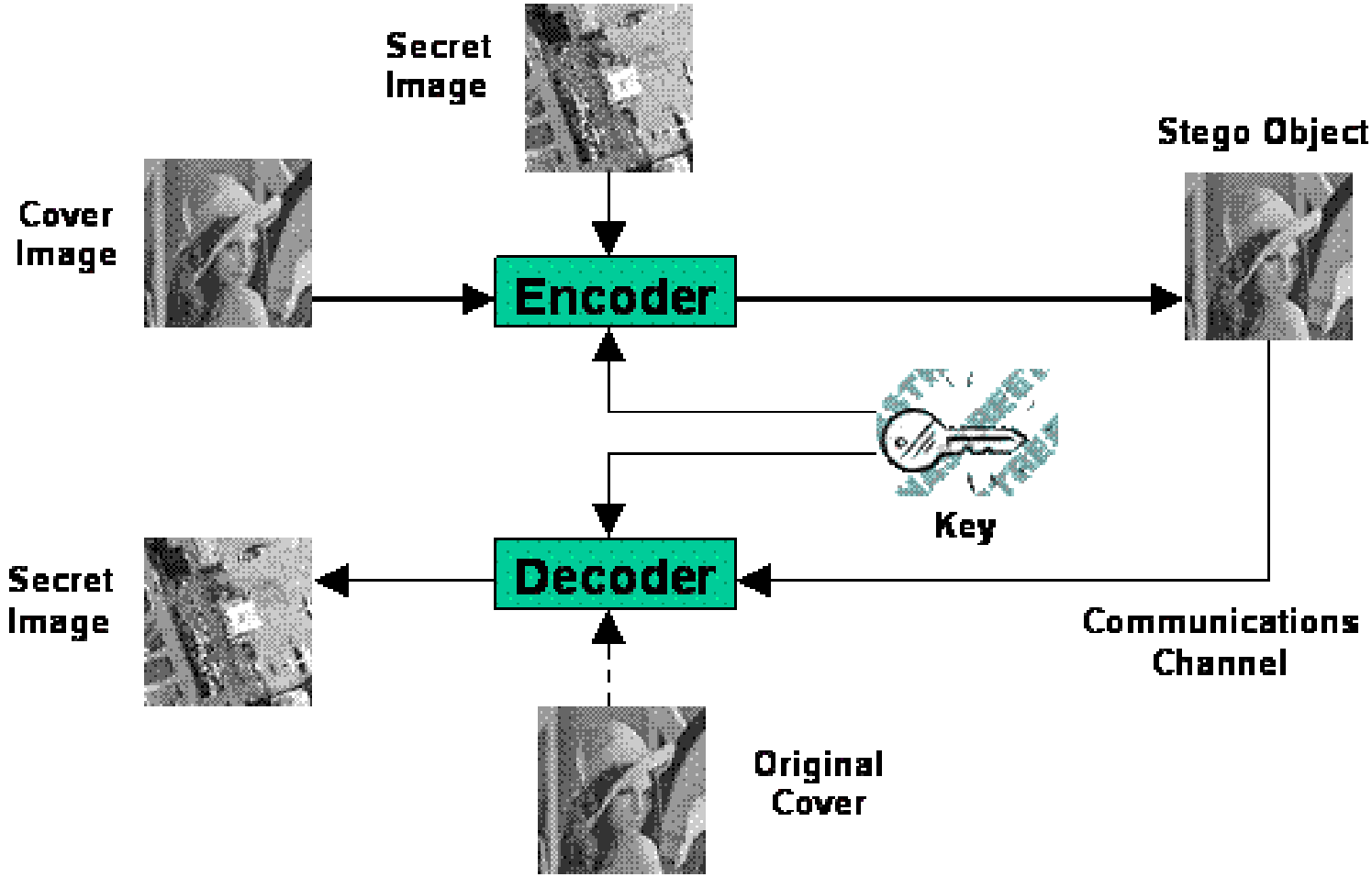




Types of watermarking/steganography

- Fragile vs. robust
  - Fragile is expected to destroy with modifications. Robust is expected to survive noise.
- Blind vs. semi-blind vs. non-blind
  - Blind needs the original covertext for detection. Semi-blind needs some information from the insertion, but not the whole covertext.

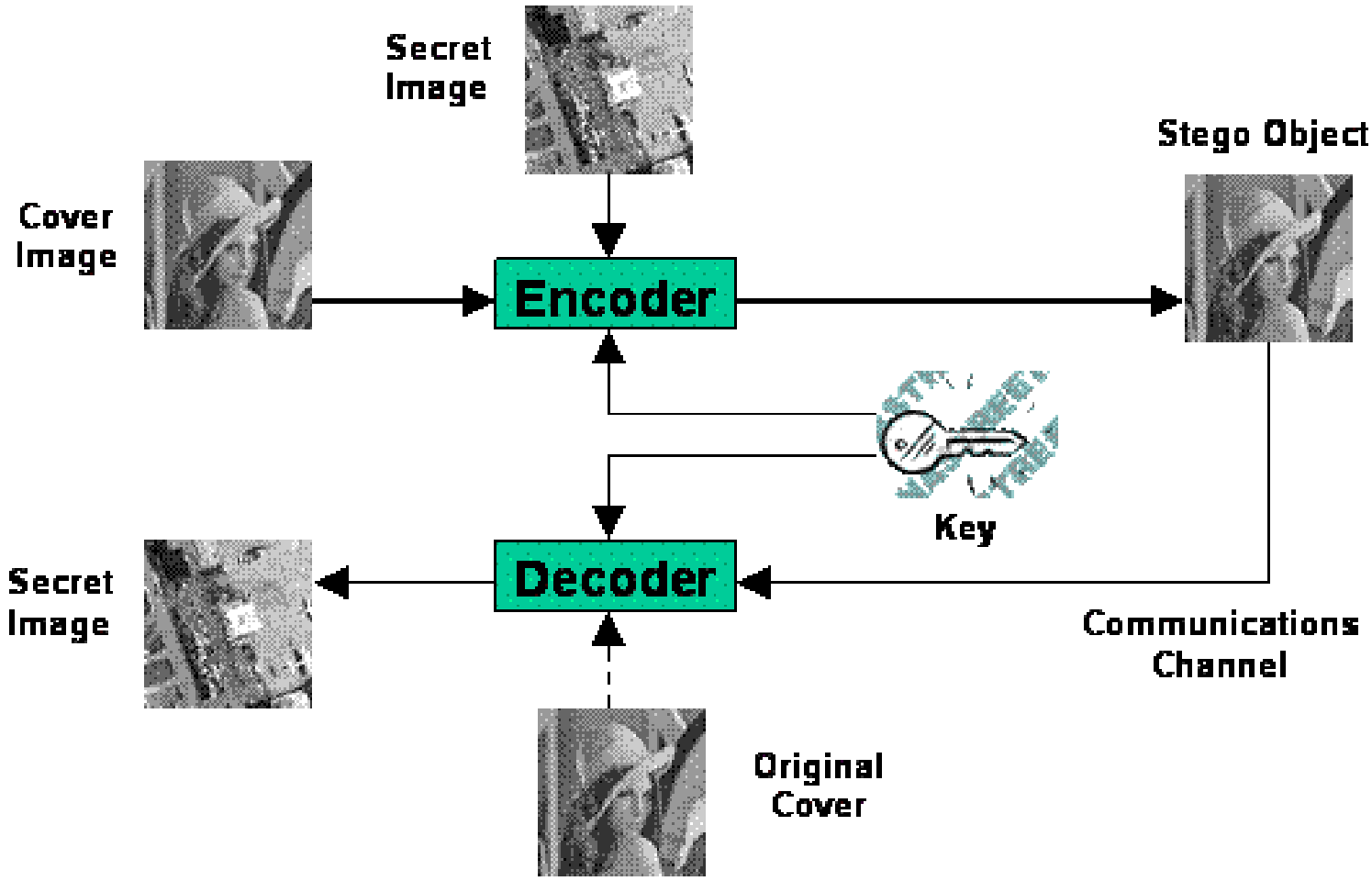
A blind scheme



Types of watermarking/steganography

- Fragile vs. robust
  - Fragile is expected to destroy with modifications. Robust is expected to survive noise.
- Blind vs. semi-blind vs. non-blind
  - Blind needs the original covertext for detection. Semi-blind needs some information from the insertion, but not the whole covertext.
- Pure vs. secret key vs. public key
  - Pure needs no key for detection. Secret key schemes needs a secret key for both embedding and detection. Public key schemes use a secret key for embedding, a secret key for detection.

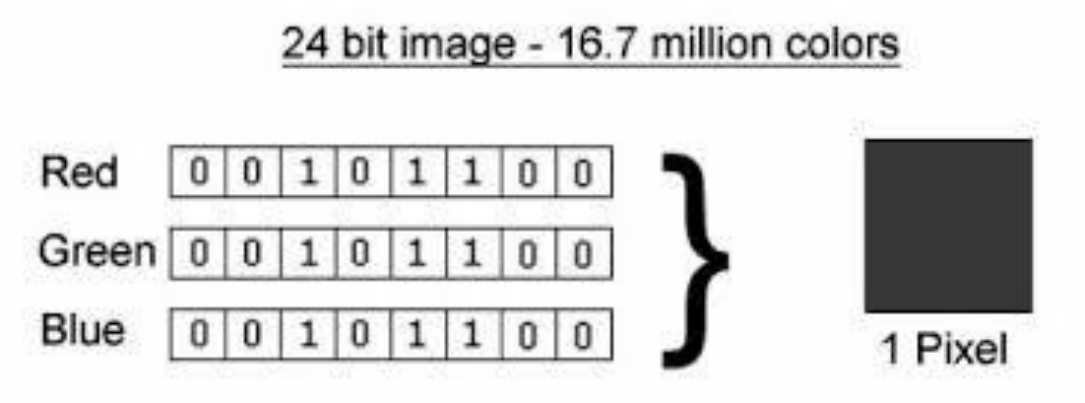
Pure vs. secret key vs. public key

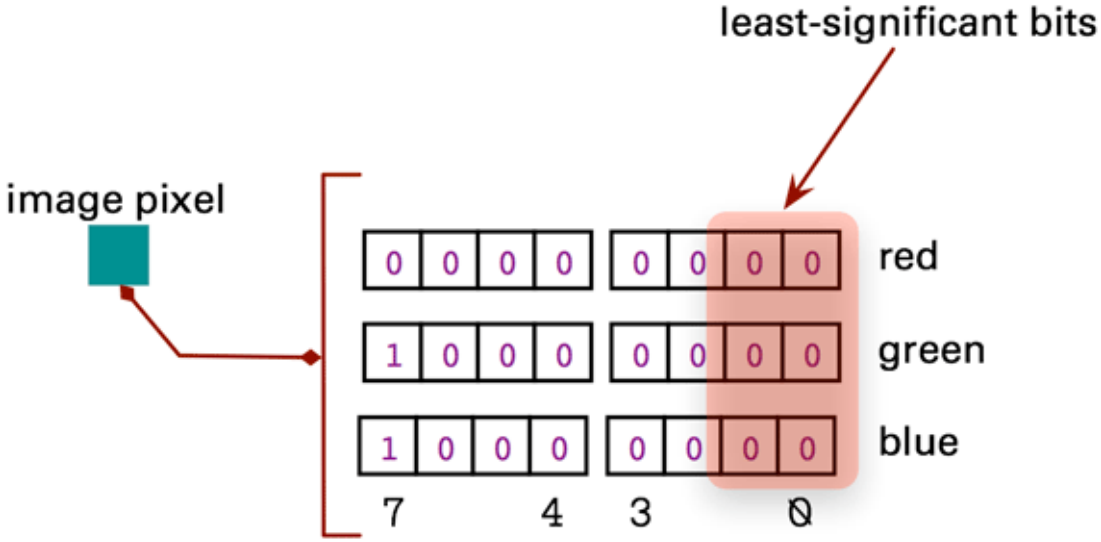


Example Steganography scheme

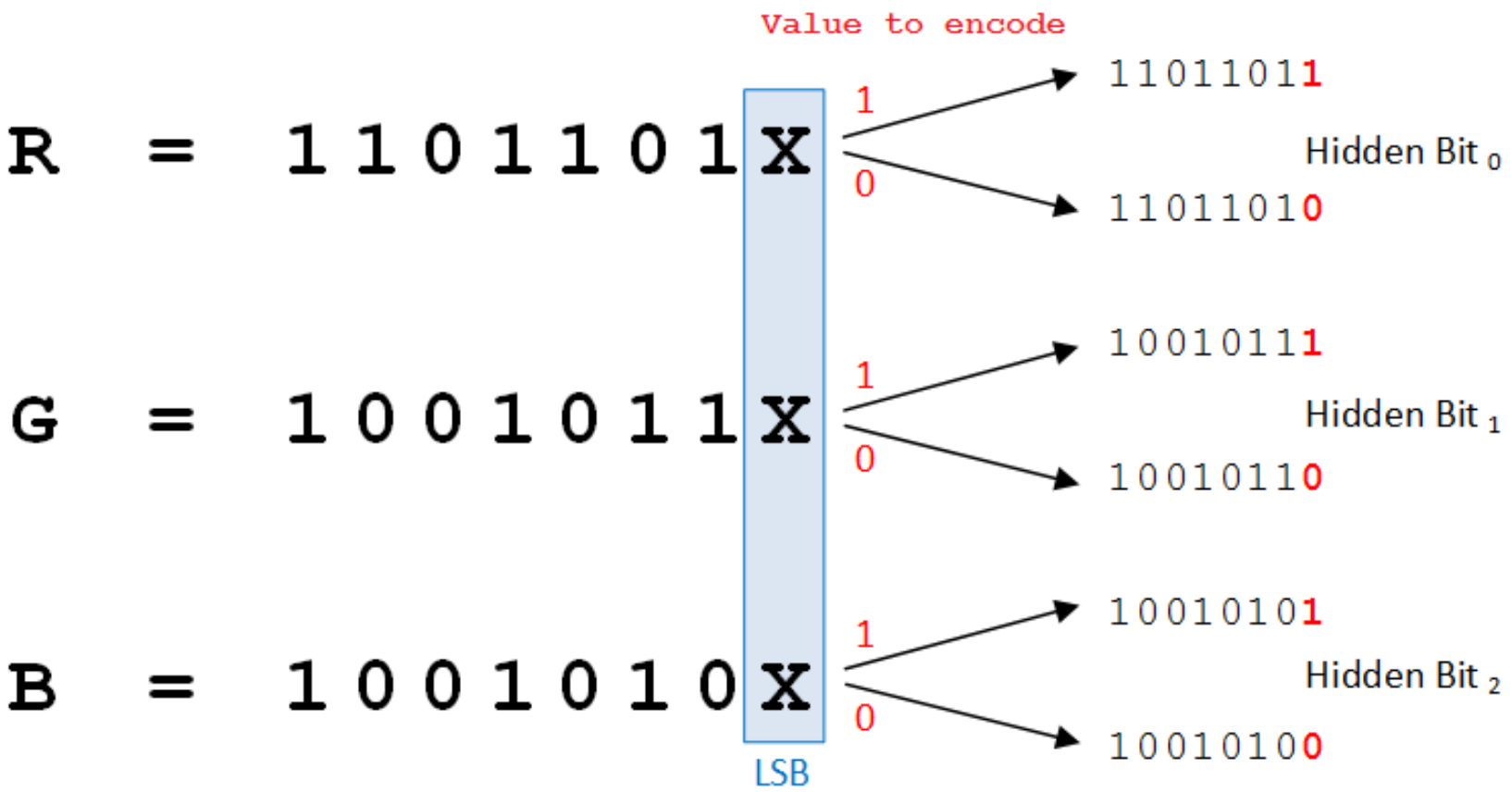
- LSB-based image steganography

A digital image





LSB-based steganography



Original Images

Using multiple bits

- Bits used=1
- Host pixel: 10110001
- Secret pixel: 01100110
- Resulted pixel: 10110000

- Bits used=4
- Host pixel: 10110001
- Secret pixel: 01100110
- Resulted pixel: 10110110

Example scheme: LSB-based steganography

Original Images

Bits Used: 1

Bits Used: 4

Bits Used: 7

Transform-domain schemes

- This is fragile!
  - Sources of noise: compression, resizing, cropping, rotating, AWGN, etc.

- For robust watermarking, embed into transform domains
  - DWT
  - DCT

Watermarking model

- Consider an image *I*, a watermark key *k*, and a watermark signal *w* produced by a watermark generated algorithm, e.g., a pseudorandom generator
- Watermark is embedded as $I_w = F^{-1}(F(I)*w)$
- Detector should detect the presence of the watermark from the noisy image:
  - $I_N = (F^{-1}(F(I)*w))\#N$
- Sources of noise: compression, resizing, cropping, rotating, AWGN, etc.

DCT-based watermark

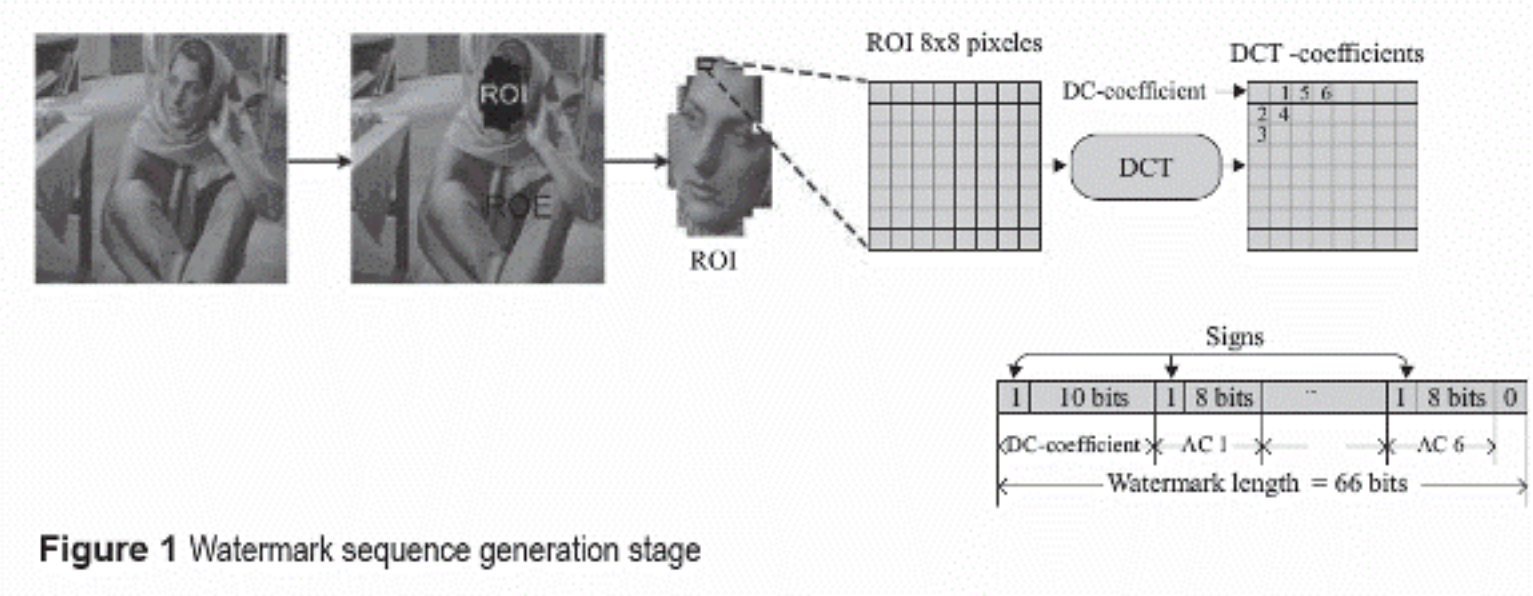
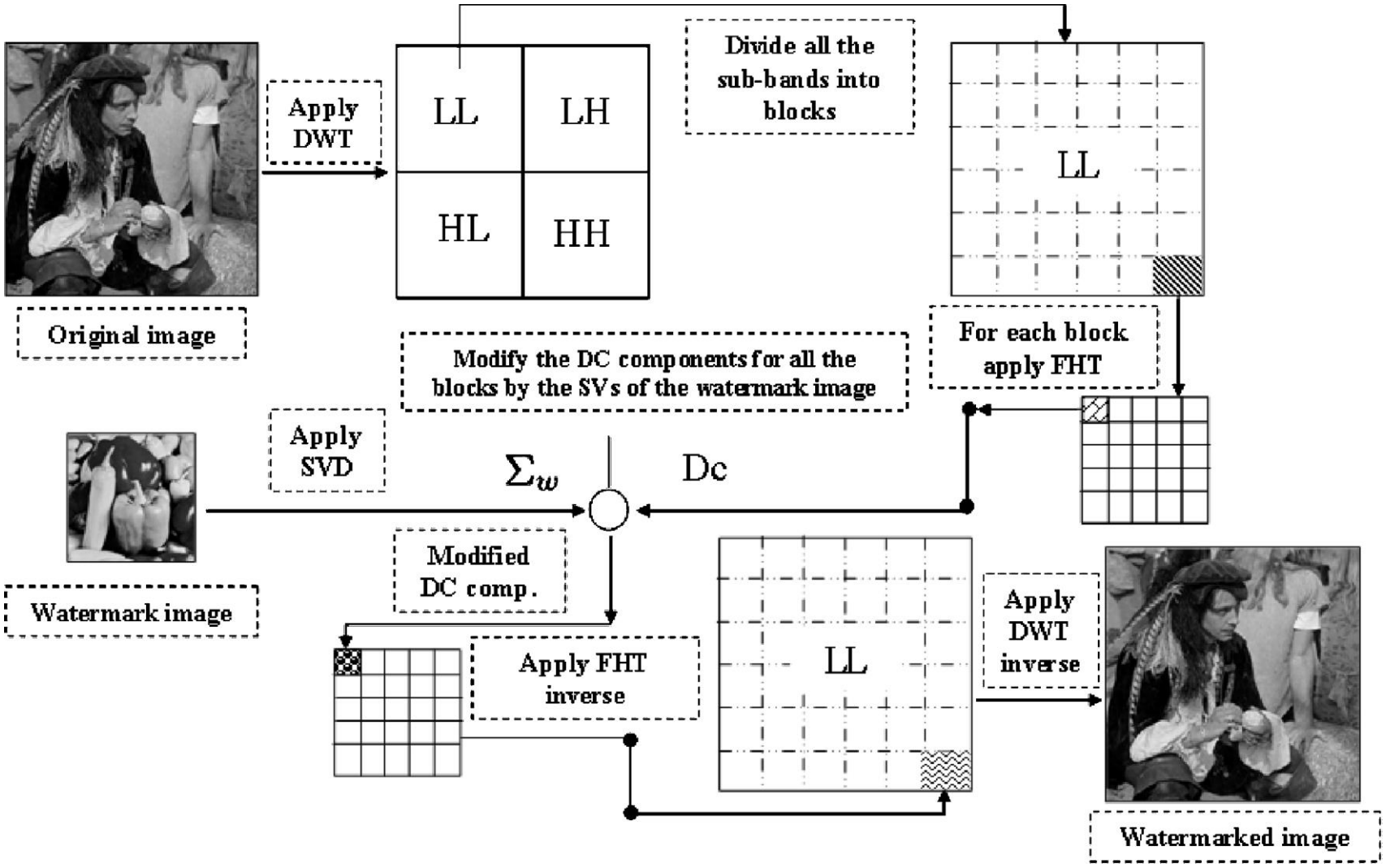


Figure 1 Watermark sequence generation stage

Wavelet-based watermark



Why is it more robust?

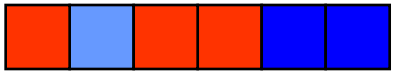
Watermark attacks

- Robustness attacks: remove or diminish the presence of watermark
- Presentation attacks: modify the content so that detector can not find the hidden watermark
- Interpretation attacks: prevent assertion of ownership, e.g., re-watermarking

Practical challenges of watermarks?

Requirements

Application



Covert communication

Copyright protection of images (authentication)

Fingerprinting (traitor-tracing)

Adding captions to images, additional information, such as subtitles, to videos

Image integrity protection (fraud detection)

Copy control in DVD

Intelligent browsers, automatic copyright information, viewing movies in given rated version

capacity
robustness
invisibility
security
embedding complexity
detection complexity

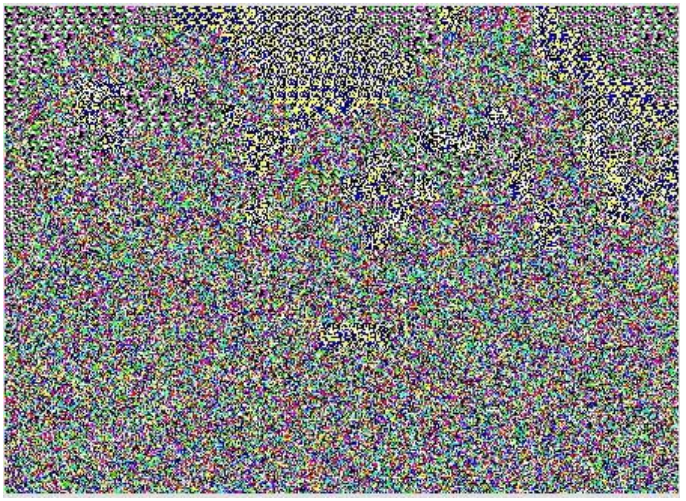
Requirements

Low High

(a)

(b)

(c)



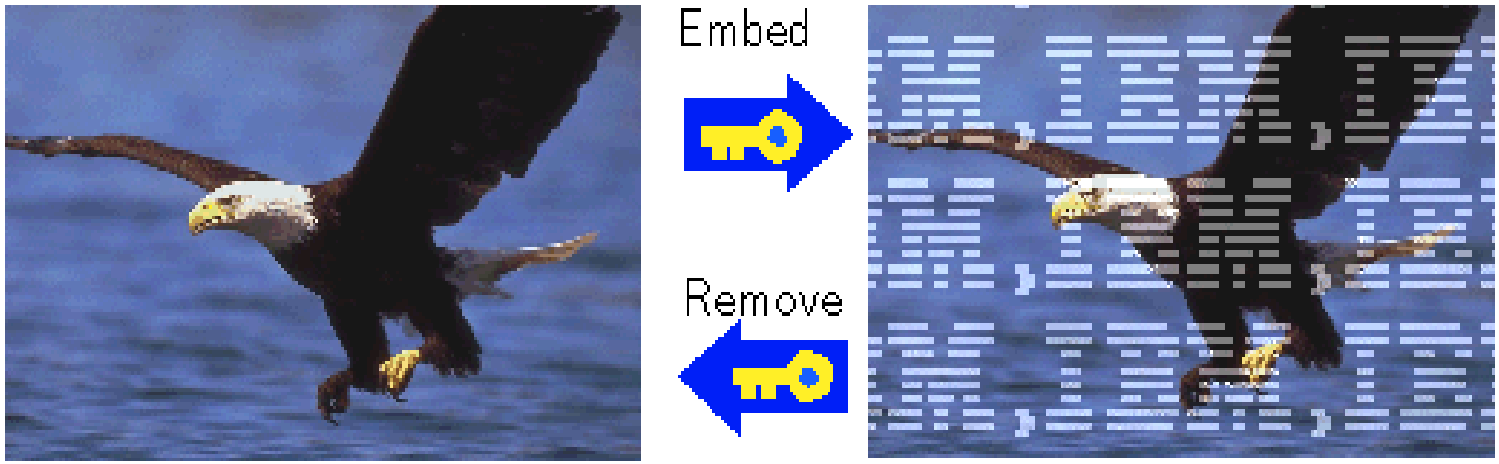
(d)

(e)

(f)

Visible Watermarking

Visible Watermarking



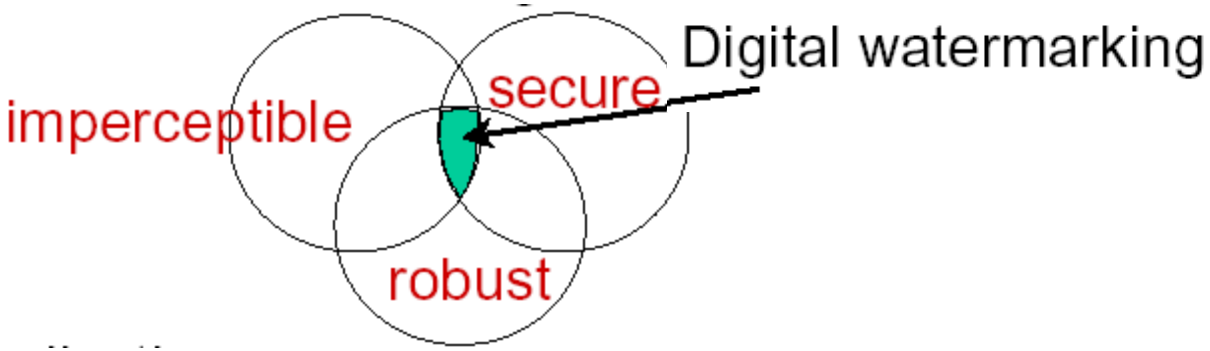
Invisible Watermarking

- Persyaratan umum *watermarking*:

  - *imperceptible*: *watermark* tidak dapat dipersepsi secara visual/auditori karena *watermark* tidak boleh merusak kualitas media *host*.

  - *robustness*: kokoh terhadap manipulasi yang ditujukan untuk merusak atau menghapus *watermark*.

  - *secure*: hanya pihak yang punya otoritas dapat mengakses *watermark*.



Acknowledgement

- Some of the slides, content, or pictures are borrowed from the following resources, and some pictures are obtained through Google search without being referenced below:
- https://www.cs.bham.ac.uk/~mdr/teaching/modules03/security/students/SS5/Steganography.htm
- http://poseidon.csd.auth.gr/LAB_SEMINARS/DigDays/Lectures/Information_Hiding.ppt