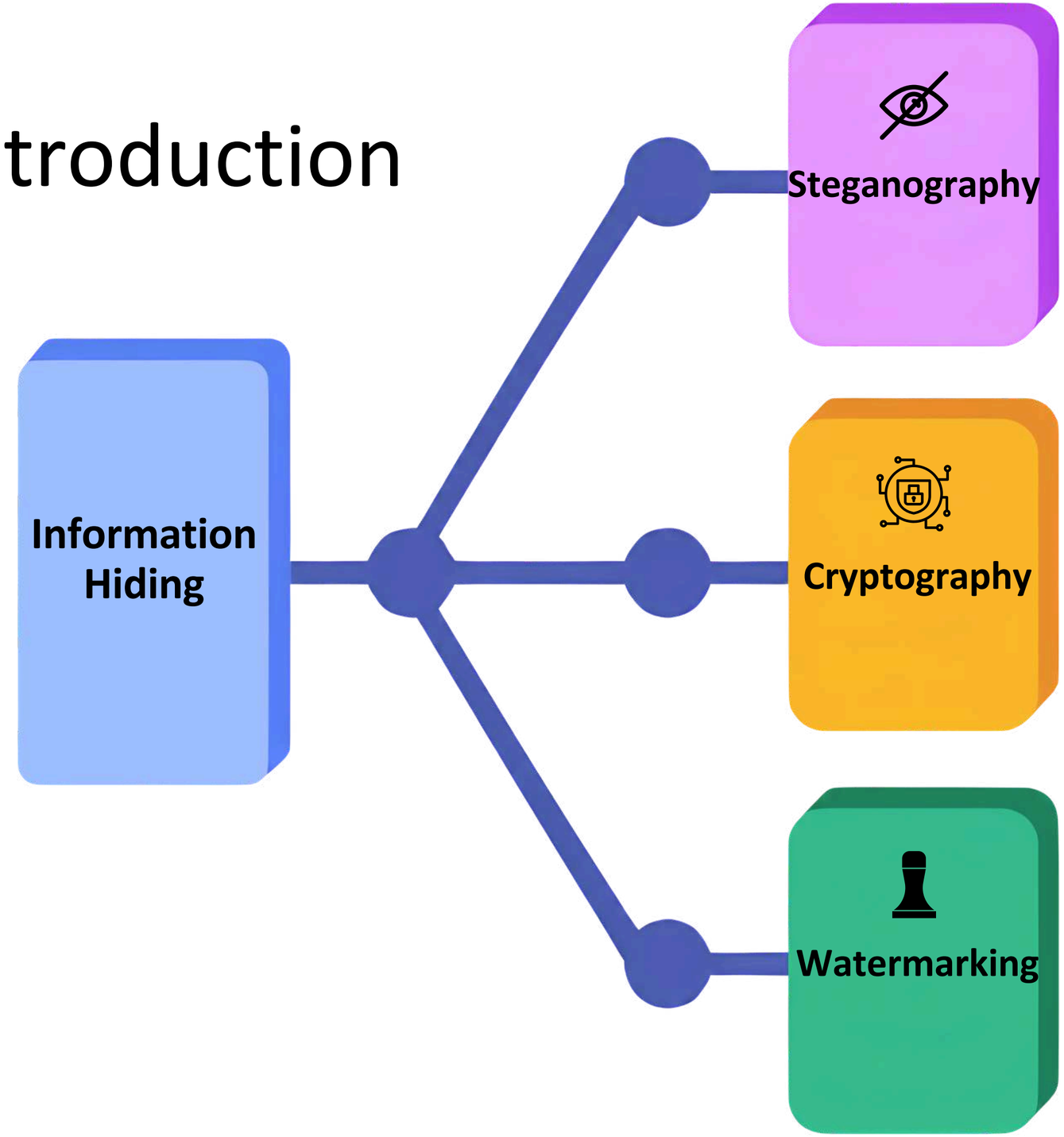


Digital Image Watermarking

Multimedia Systems

Romenia Silva
romaniasilva@sjp.ac.lk

Introduction



hiding of a secret message within an ordinary message and the extraction of it at its destination

cryptography is about constructing and analyzing protocols that prevent third parties or the public from reading private messages

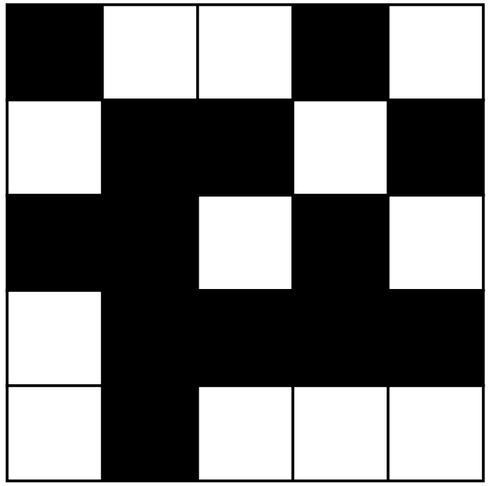
?

Watermarking

An invisible signature embedded inside an image, video or audio to show authenticity or proof of ownership.

A Digital Watermark:

A pattern of bits inserted into a digital image, audio or, video file that identifies the file's copyright information (author, rights, etc.).



W=[1 0 0 1 0
0 1 1 0 1
1 1 0 1 0
0 1 1 1 1
0 1 0 0 0]

Digital Image Watermarking:

- Shows authenticity or proof of ownership of images
- Discourage unauthorized copying and distribution of images
- Ensure a digital picture has not been altered

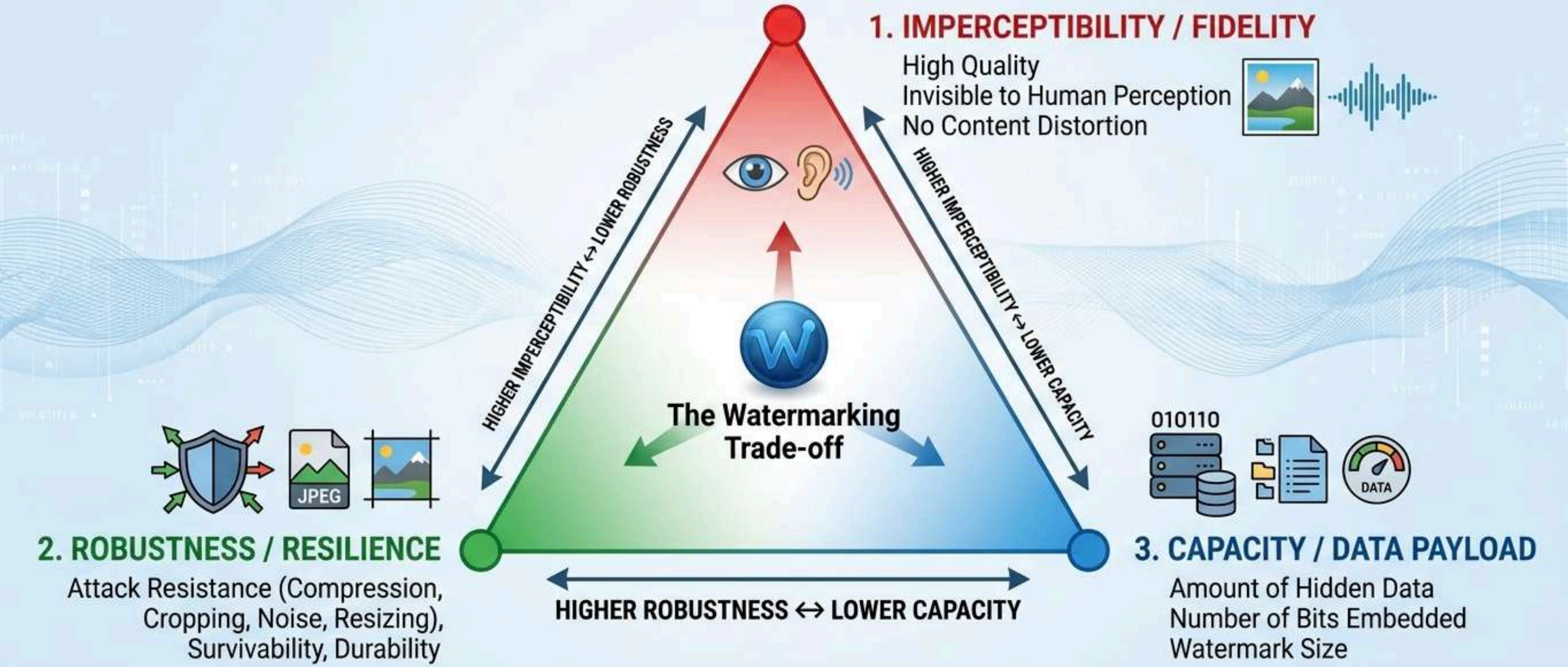
Characteristics of Watermarking

- Should be transparent:
  - Appears random, noise-like sequence
  - Undetectable

- Good Correlation Properties
  - **High** correlation with signals similar to watermark
  - **Low** correlation with other watermarks or random noise

- Should pose some robustness
  - Cropping
  - Lossy compression
  - Image filters and other signal processes.

THE WATERMARKING TRADE-OFF TRIANGLE



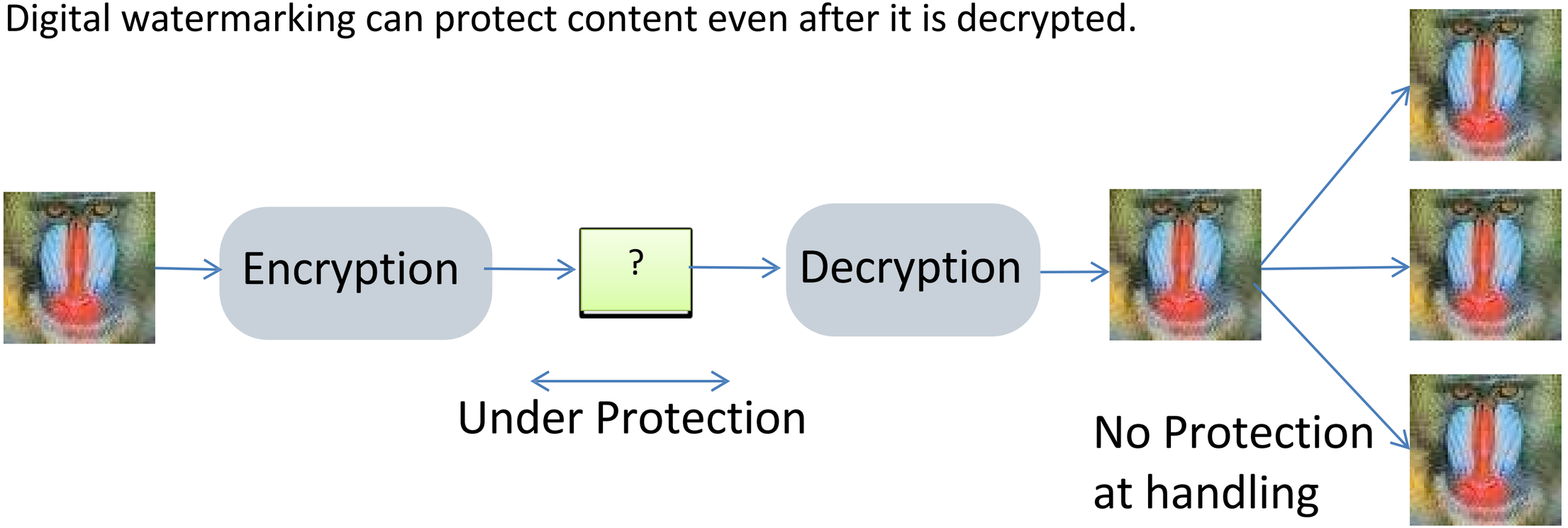
Steganography vs Watermarking

- The main goal of **steganography** is **to hide** a message ***m*** in some audio or video (cover) data ***d***, to obtain new data ***d'***, practically indistinguishable from ***d***, by people, in such a way that an eavesdropper **cannot detect** the presence of ***m*** in ***d'***.

- The main goal of **watermarking** is **to hide** a message ***m*** in some audio or video (cover) data ***d***, to obtain new data ***d'***, practically indistinguishable from ***d***, by people, in such a way that an eavesdropper **cannot remove or replace** ***m*** in ***d'***.

Cryptography vs. Watermarking

- Cryptography is the most common method of protecting digital content and is one of the best developed science.
- However, encryption cannot help the seller monitor how a legitimate customer handles the content after decryption.
- Digital watermarking can protect content even after it is decrypted.



DIGITAL WATERMARKING CLASSIFICATION

DIGITAL WATERMARKING

1. PERCEPTIBILITY-BASED CLASSIFICATION

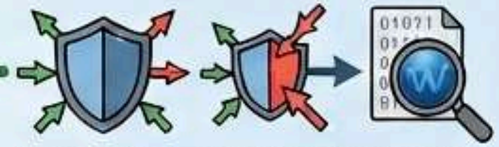


A. VISIBLE WATERMARKING



B. INVISIBLE WATERMARKING

2. RESILIENCE-BASED CLASSIFICATION



A. ROBUST WATERMARKING



B. FRAGILE WATERMARKING

3. KEY-BASED CLASSIFICATION



A. SYMMETRIC WATERMARKING

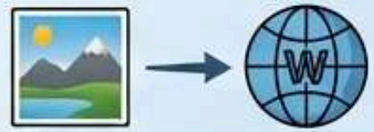


B. ASYMMETRIC WATERMARKING

4. DETECTION-BASED CLASSIFICATION



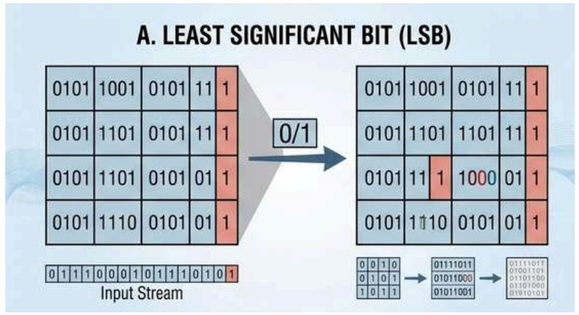
A. PRIVATE WATERMARKING (Informed)

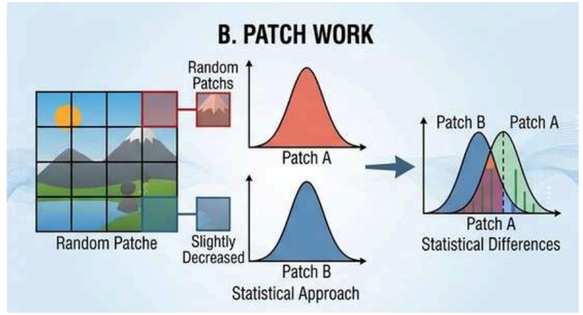


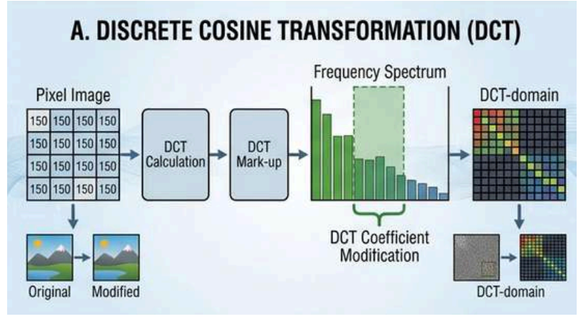
B. PUBLIC WATERMARKING (Blind)

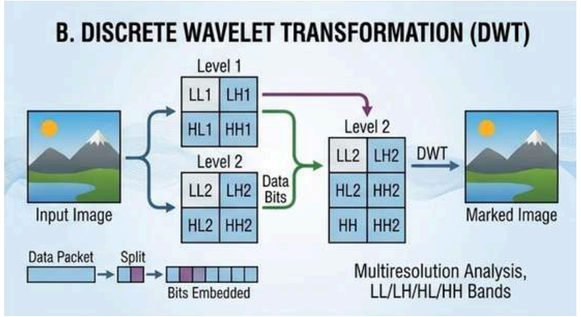
Watermarking Techniques

- Spatial Domain Watermarking
  - Least Significant Bit
  - Patch Work
- Frequency Domain Watermarking
  - Discrete Cosine Transformation
  - Discrete Wavelet Transformation



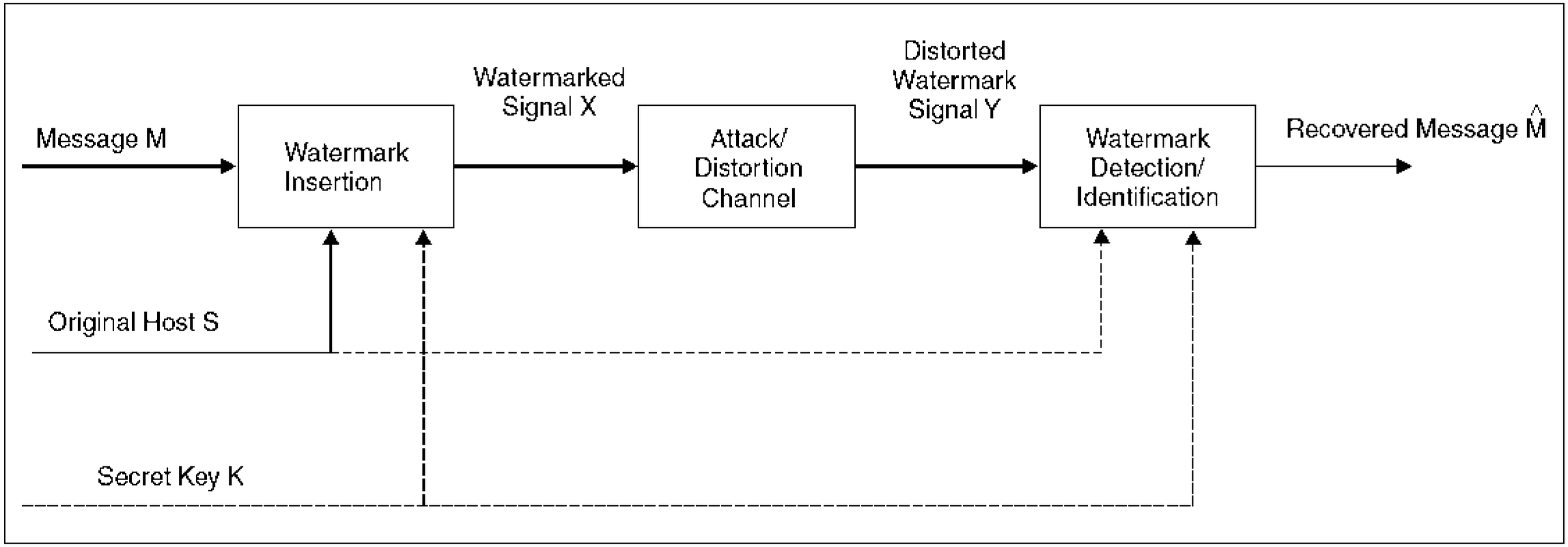






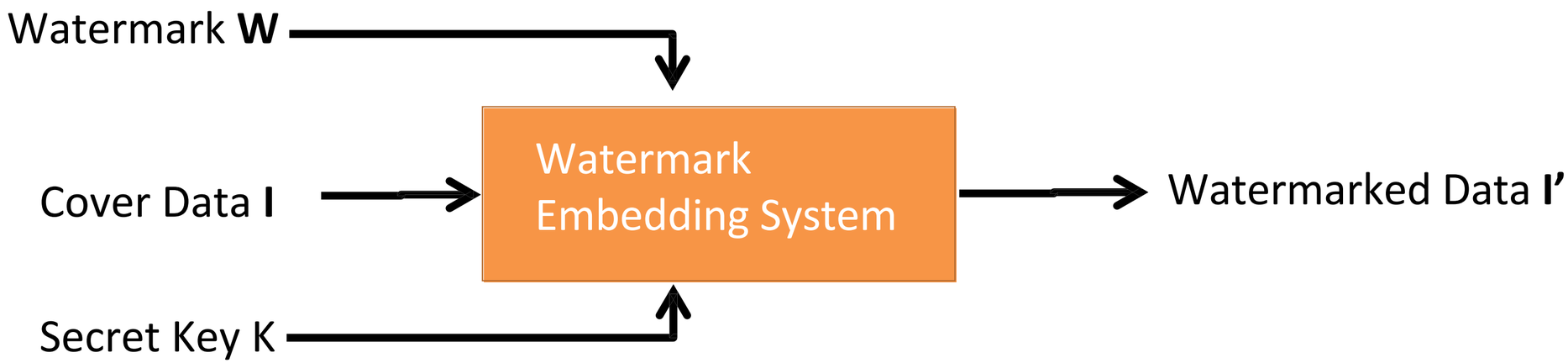
Digital Watermarking Lifecycle Phases

- A watermarking system is usually divided into three distinct steps.
  - Embedding
  - Attack
  - Detection



Embedding

- In embedding, an algorithm accepts the host and the data to be embedded, and produces a watermarked signal.



Inputs to the scheme are the **watermark**, the **cover data** and an optional **public or secret key**. The output are **watermarked data**. The key is used to enforce security.

Attacks

- *The watermarked digital signal is transmitted or stored, usually transmitted to another person. If this person makes a **modification**, this is called an **attack**.*

Several Attacks:

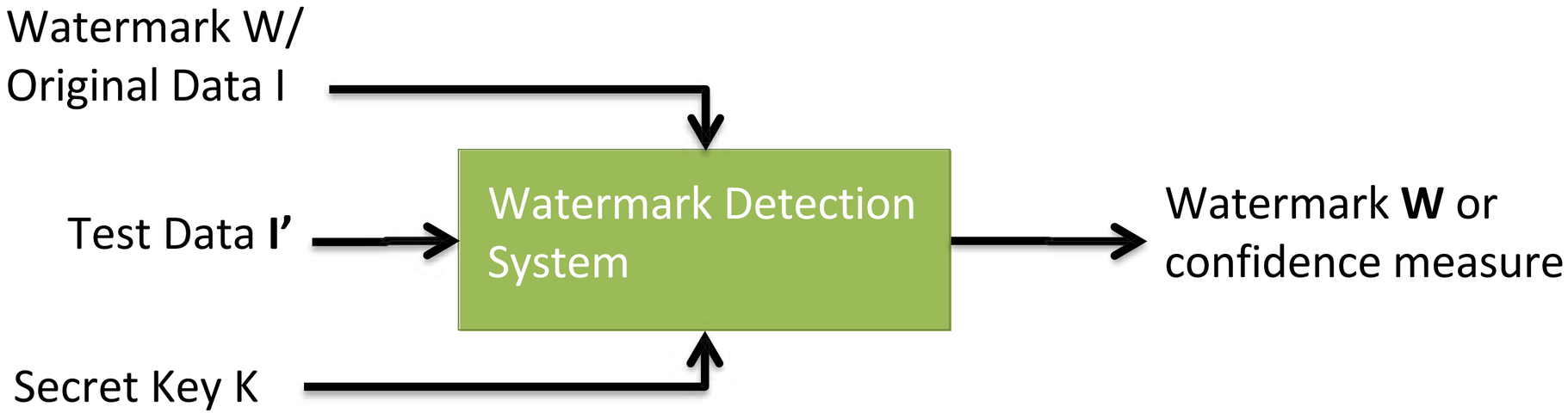
- **Robustness attacks** : Which are intended to remove the watermark such as… JPEG compression, cropping, etc.

- **Presentation Attacks** : Geometric transformation, rotation, scaling, translation, change aspect ratio, affine transformation etc.

- **Counterfeiting attacks**: Rendering the original image, generate fake original.

Extraction/Detection

- **Detection** is an algorithm which is applied to the attacked signal to attempt to extract the watermark from it.
  – *If the signal was unmodified during transmission, then the watermark still is present and it may be extracted.*



Inputs: watermarked data, the secret or public key and, the original data and/or the original watermark (depending on the method).

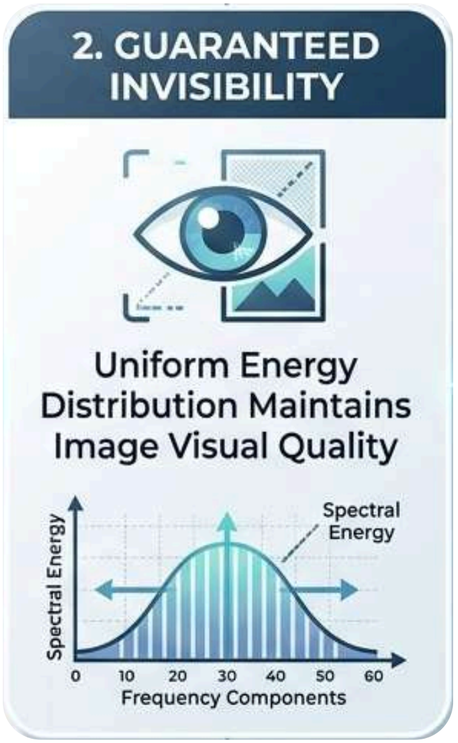
Output: The recovered watermarked *W* or some kind of confidence measure indicating how likely it is for the given watermark at the input to be present in the data under inspection.

Example Algorithm

Fast Hadamard Transform(FHT)

Advantages of FHT

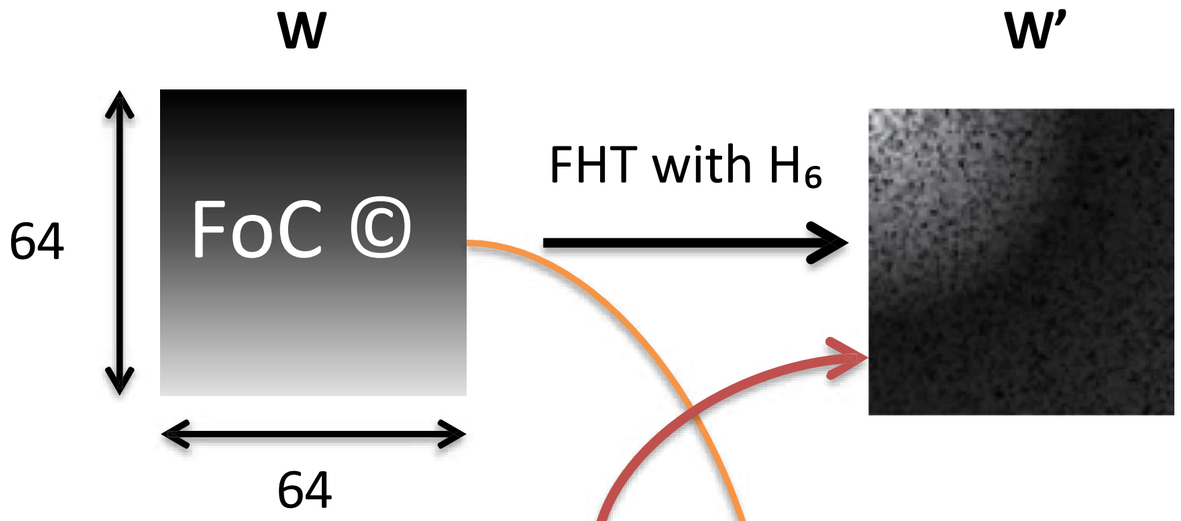






Watermarking with FHT

1. Transform the watermark to FHT
2. Pseudo-randomly select sub-blocks for insertion
3. Transform the selected blocks of cover image to FHT
4. Insert watermark to the selected block
    i. All the AC coefficient of the watermark will be inserted.
    ii. The DC coefficients are kept as the key.
5. Apply IFHT to the selected blocks of the image.
6. The result is a watermarked image and the key file.



Since the watermark image is 64x64 the Hadamard order is 6. ($2^6 = 64$) hence **n = 6.**

FHT : $$[V] = \frac{H_n [U] H_n}{N}$$

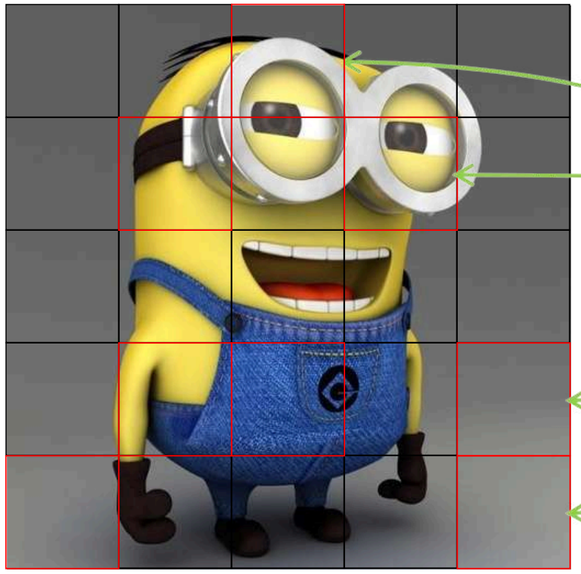
where H_n represents an $N \times N$ Hadamard matrix, $N = 2^n$, $n = 1,2,3\ldots$

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_n = H_{n-1} \otimes H_1$$

Kronecker product

$$H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}$$



Seed S1

Pseudo-random Generator

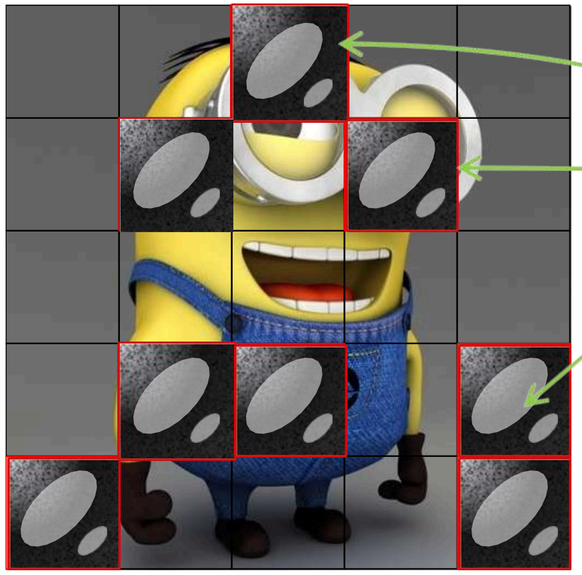
Randomly selects **8x8** blocks

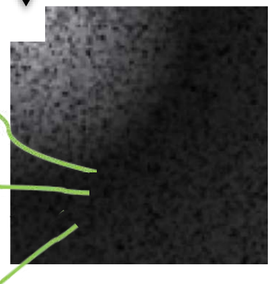
DC coefficient

Key File

FHT with H_3

W'





Insert AC coefficients of W' to the middle and high frequency locations of selected blocks

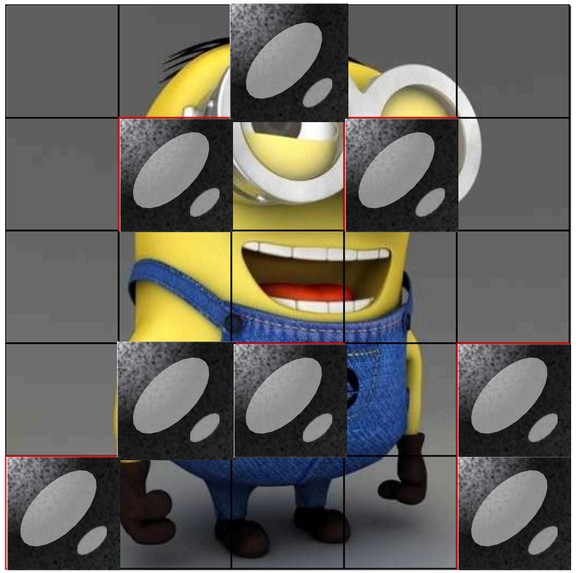
Modified AC coefficients In the cover image

AC coefficients in the watermark

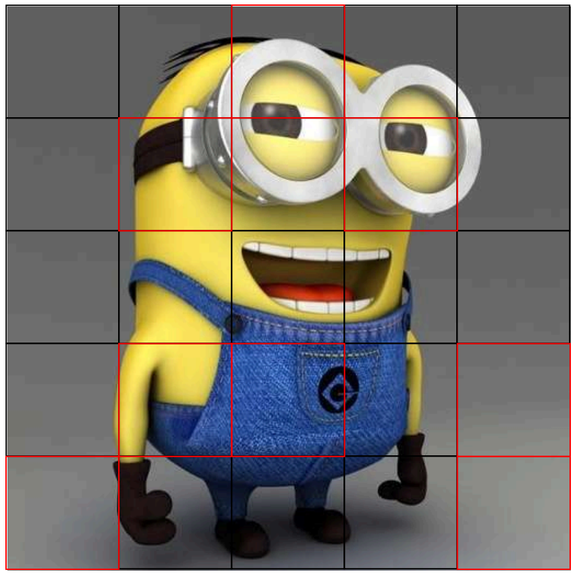
$$x_i^* = \alpha m_i$$

The watermark strength factor

Watermarked Image



IFHT(Selected Blocks)



IFHT:

$$[U] = H_n^{-1}[V]H_n^{*} = \frac{H_n[V]H_n}{N}$$

Exactly Similar to the FHT

Key File

- Seed S1
- DC coefficient
- Strength Factor

Extraction Process

1. Identify relevant blocks for the extraction using a random generator
2. Apply FHT to the selected block
3. Extract watermark AC coefficients
4. Combine with the DC (in the key file) and AC to extract the watermark (in FHT)
5. Apply IFHT to the watermark.
6. Compare the similarity between the extracted and the original watermark

More Information:

Anthony T.S.Ho, Jun Shen, Soon Hie Tan “A Robust Digital Image-in-Image Watermarking Algorithm Using the Fast Hadamard Transform” Proceedings of SPIE Vol. 4793 (2003) © 2003 SPIE · 0277-786X/03/$15.00

Attacks on the Watermarked Image



Original watermarked image



Cropping 50%



Rotation 30°



Changing aspect ratio
x-1 y-0.8

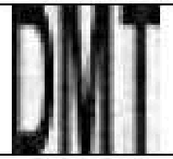
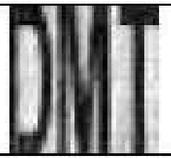
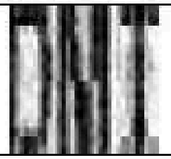
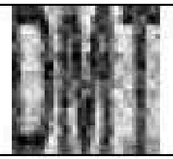


3×3 median filtering



JPEG compression
factor 35

Experimental Results

Image operations	Extracted watermark	Correlation
Sharpening 3×3		0.9573
1 rows 1 column removed		0.9866
Frequency Mode Laplacian removal		0.9580
Scaling 0.75		0.9354
JPEG Compression of factor 30		0.8688
Change aspect ratio x_1.00_y_1.20		0.8199

Exercise

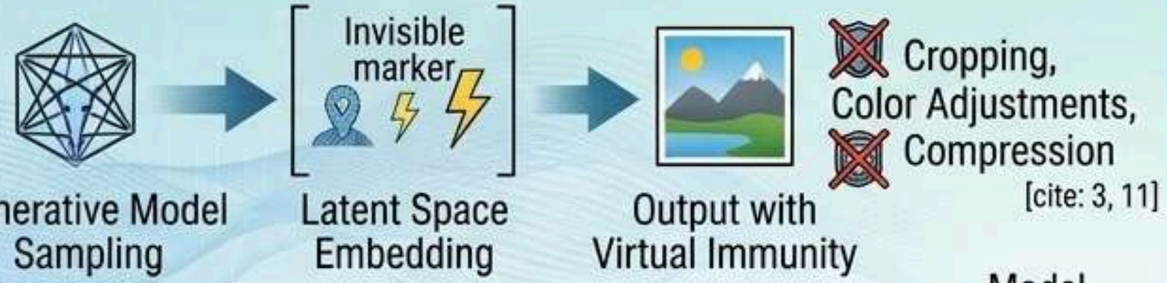
- Obtain H_3 Hadamard Matrix.
- Obtain the FHT of the following image.

3	2	0	0
5	4	1	1
2	2	2	0
1	0	1	2

MODERN DIGITAL WATERMARKING INDUSTRY DEVELOPMENTS

1. GENERATIVE AI & DEEPFAKE VERIFICATION (C2PA & LATENT WATERMARKING)

Latent Space Watermarking (e.g., Google SynthID)



C2PA Standard

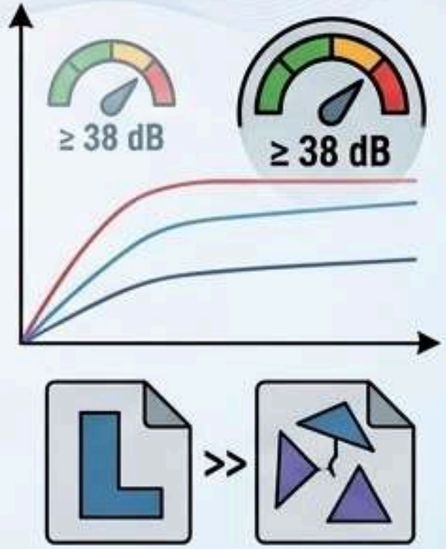


- Modern watermarking has moved into model training and generative AI tracing[cite: 3]

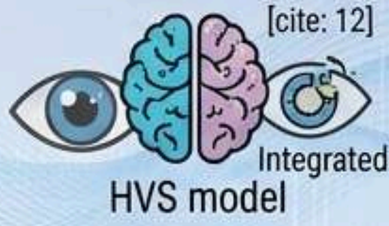


2. ADVANCED QUALITY EVALUATION METRICS

Beyond Classical ρ (Normalized Cross-Correlation)[cite: 4, 12]



Human Visual System (HVS) Models



PSNR (Peak Signal-to-Noise Ratio)[cite: 12]
Minimum standard target for imperceptibility

SSIM (Structural Similarity Index Measure)
Measures structural distortion rather than pixel errors alone