EVIL CORP

CTM challenge ECORP

Multimedia Data Security

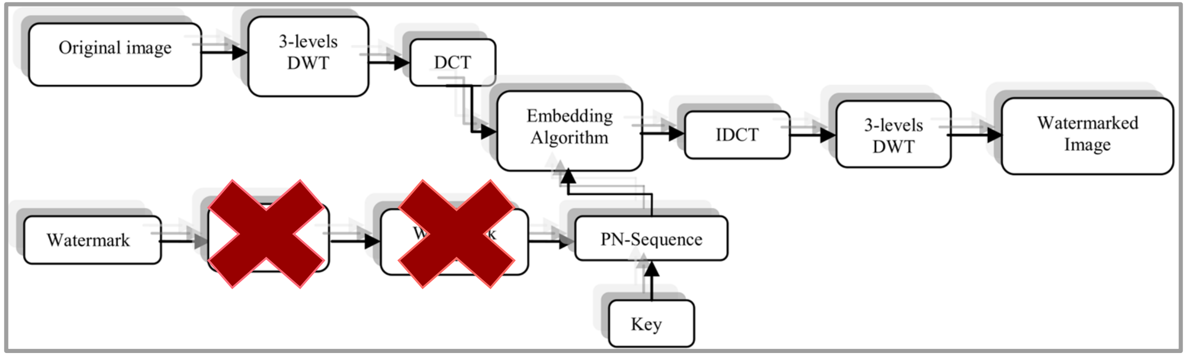
Embedding implementation Joint DWT-DCT

Based on “Robust Digital Image Watermarking Based on Joint DWT-DCT” by Saeed Kasmani.

Our implementation follow the following steps:

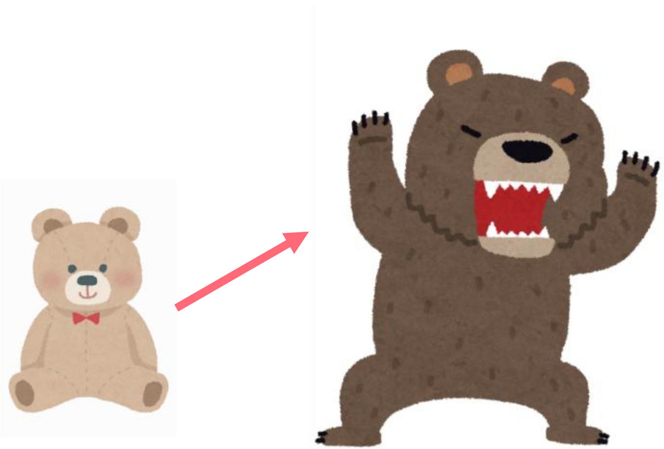
1. Multi-Resolution Decomposition (DWT)
2. Block Division and DCT Transform
3. Watermark Embedding
4. Reconstruction of the Watermarked Image

The watermark is embedded where it’s invisible to the eye but robust to distortion.



Embedding improvement

Goal: obtain the most robust watermark embedding possible, maintaining a WPSNR above 54 dB.



Adaptive embedding strength per block

Makes the embedding locally stronger in textured areas and weaker in smooth regions, improving the balance between invisibility and robustness.

Block equalization

Normalizes each block by its local variance. This helps keep the watermark effect more uniform across the image.

HVS-weighting mask

Uses perceptual frequency weights so the embedding follows the human visual system, reducing visual artifacts.

Automatic K tuning (WPSNR-based)

Automatically adjusts global embedding strength to reach a target WPSNR, guaranteeing consistent visual quality across images.

Detection (non-blind)

- **3-levels DWT** on both the original image **I** and the test image **I***→ extract same sub-bands **HL$_3$, LH$_3$ HL$_2$, LH$_2$ (64×64)**.

- For each sub-band, divide into **4×4 DCT blocks** (1024 total blocks) and compute the 7 mid-band coefficient differences.

$$\Delta = \frac{C^* - C}{\sigma_{block} + \varepsilon}$$

(Project each normalized difference vector onto the normalized to unity) detection direction.

$$s = \Delta \cdot D_{det} \qquad D_{det} = \frac{PN_1 \cdot H - PN_0 \cdot H}{||PN_1 \cdot H - PN_0 \cdot H||}$$

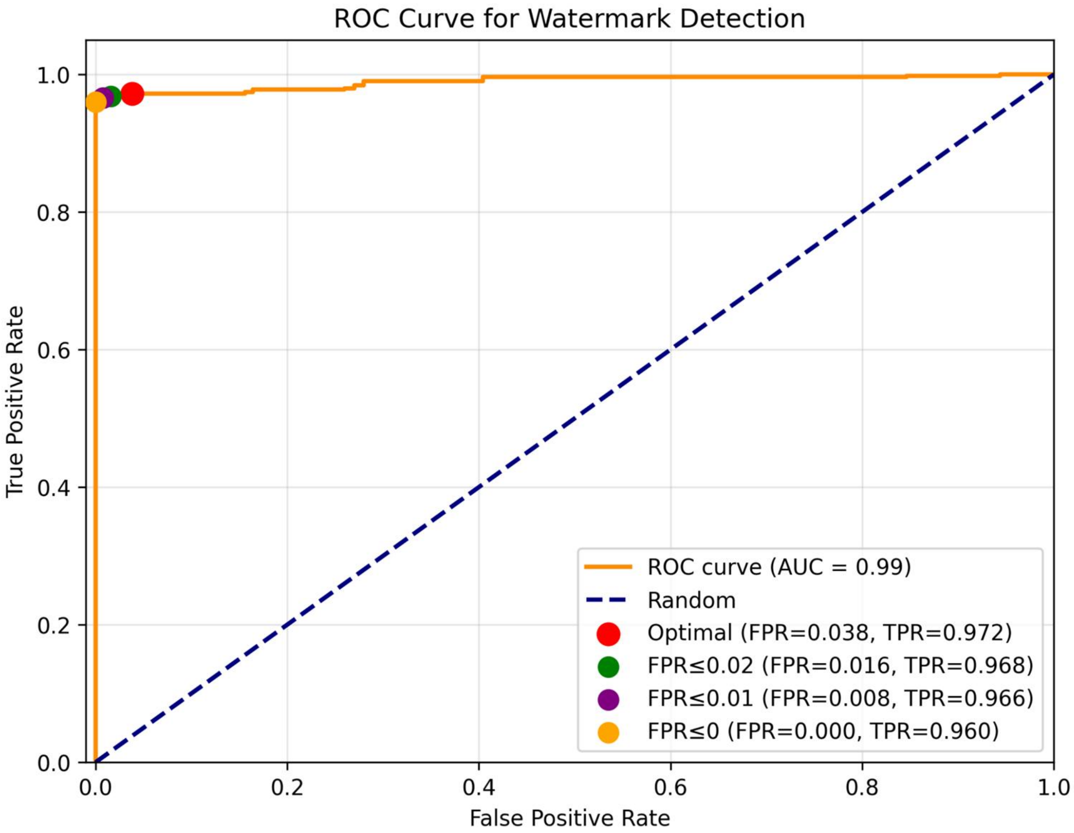
- Bit decision $\hat{b} = \begin{cases} 1 & s > 0 \\ 0 & s \leq 0 \end{cases}$

This process repeats for all 1024 blocks, producing a 1024-bit extracted watermark.

- **Compute similarity** between **extracted** bits from the **watermarked** and **attacked** images and decide if the watermark is present.

$$W = \begin{cases} 1 & S > \tau \\ 0 & S \leq \tau \end{cases}$$

ROC



Optimal point ⇒ $\boldsymbol{\tau}$ = 0.52023

0-FPR ⇒ $\boldsymbol{\tau}$ = 0.54602

Attack Strategy - Not simply brute force

2 conditions to stop an attack sequence:

- Find a **successful attack**

Attack	Detected	WPSNR	Tested
JPEG [QF=**80**]	✓	42dB	YES
JPEG [QF=**70**]	✗	39dB	YES
JPEG [QF=**60**]	✗	≲ 39dB	NO

- **WPSNR < 35dB**

Attack	Detected	WPSNR	Tested
JPEG [QF=**50**]	✓	37dB	YES
JPEG [QF=**40**]	✓	33dB	YES
JPEG [QF=**30**]	✗	≲ 33dB	NO

An attack sequence:

- Sequence of single attacks:
  - blur[0.2, 0.2];
    blur[0.4, 0.4];
    blur[0.6, 0.6]; ...
- Sequence of paired attacks:
  - (jpeg[60], blur[0.2, 0.2]);
    (jpeg[60], blur[0.4, 0.4]);
    (jpeg[60], blur[0.6, 0.6]); ...

Parameter profiles:

Light, medium, aggressive, fine profiles differ in **starting/ending** points and **step size** for each attack.

For JPEG:

• Light:	QF from 80 to 50, step of 10
• Medium:	QF from 60 to 20, step of 5
• Aggressive:	QF from 95 to 30, step of 5
• Fine:	QF from 60 to 50, step of 1

Tested and Discarded Embedding Techniques

Approaches explored:

- **DWT + SVD (LL band)** → imperceptible but too weak, not bit-based, fragile to JPEG.
- **Quantized DWT** → inconsistent detection, unstable results.
- **Ultra-weak DWT (>66 dB)** → visually perfect but not robust.
- **DWT(bior4.4)+DCT+ECC+pilot tones+log-polar** → over-engineered, too slow.
- **Early paper implementation** → no equalisation or HVS weighting → less stable.

What we learned?

Band weighting:
Allocating more energy to L3 than L2 **increased resistance to compression and filtering** without hurting image quality.

Realistic WPSNR target:
Ultra-high goals (≈66 dB) made the watermark too weak. A target around **54 dB** improved robustness without visible artifacts.

No over-complex pipelines:
A lean design **reduced mismatches, tuning time, and errors**

Key Differences wrt Paper

ASPECT	PAPER	CTM IMPLEMENTATION
Watermark type	32×32 binary logo, Arnold-scrambled	1024-bit sequence
Embedding strength	Fixed $\alpha = 25$	Adaptive per-block with activity and auto-tuning
PN sequences	Uncorrelated	Orthonormalized for maximum separation
HVS modeling	Generic mid-band selection	Coefficient-specific HVS weights (**H**)
Equalization	None mentioned	Per-block normalization by σ_{mb}
Band weighting	Implicit	Explicit Level 3 vs Level 2 differentiation
Quality control	Post-hoc measurement	Auto-tuning to target WPSNR
Detection	Blind (no original needed)	Non-blind (requires original)

Paper link: https://www.researchgate.net/publication/22067

Thanks