

# Multimedia Information Systems

---

---

*Samson Cheung*

EE 639, Fall 2004

## Lecture 24: Digital Watermarking II

Compiled from lecture notes by Prof. Ja-Ling Wu of National Taiwan University and the book "Digital Watermarking" by I.J. Cox, M.L. Miller and J.A. Bloom

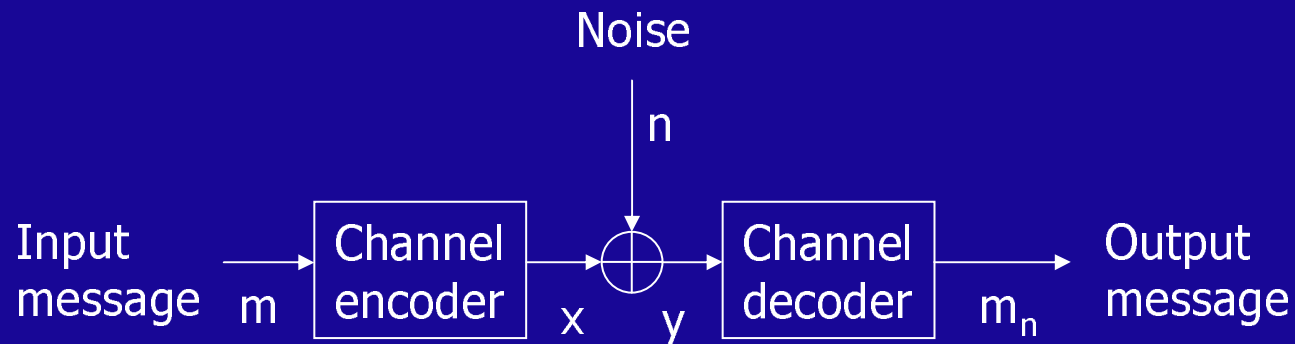
# Outline

---

---

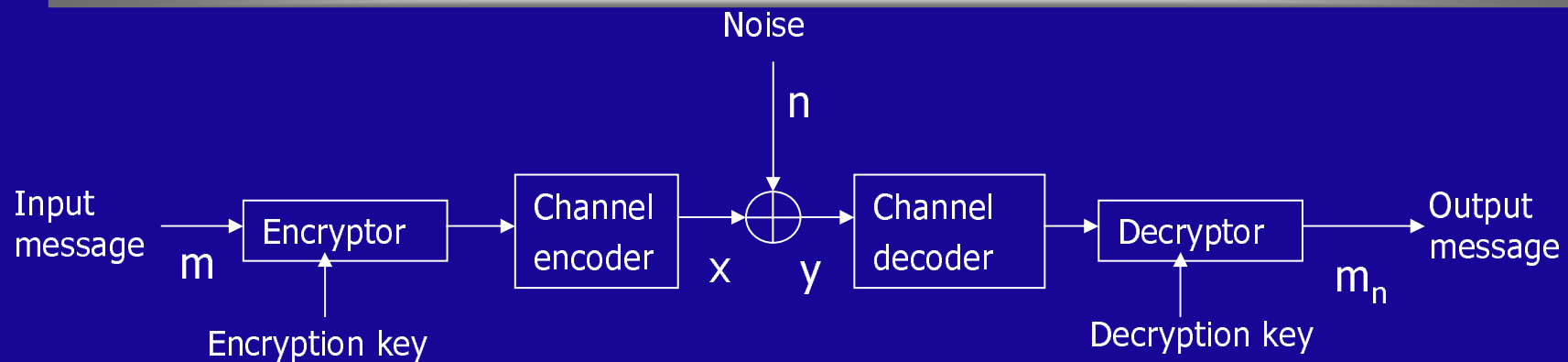
- **Communication-based models of watermarking**
  - Applied well-established communication and information-theoretical techniques to watermarking
  - Three models: basic model, side-information model, multiplexing model
- **Geometric models of watermarking**
- **Informed embedding and coding**

# Standard model of a communication system

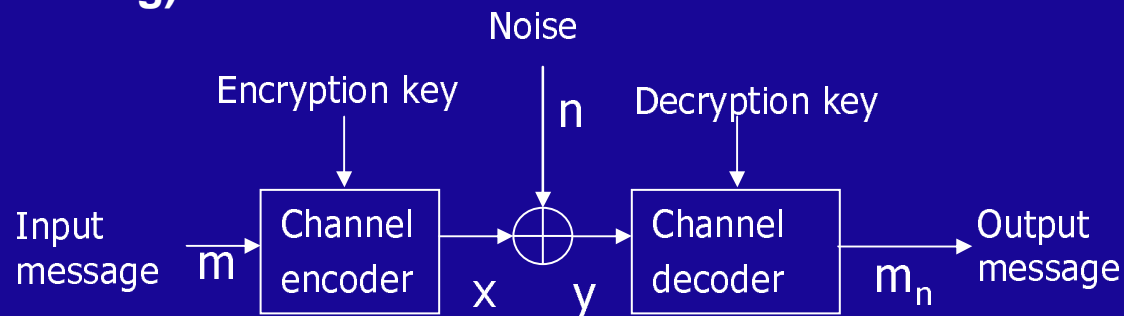


- $m$ : the message we want to transmit
- $x$ : the codeword encoded by the channel encoder
- $n$ : the additive random noise
- $y$ : the received signal
- $m_n$ : the received message

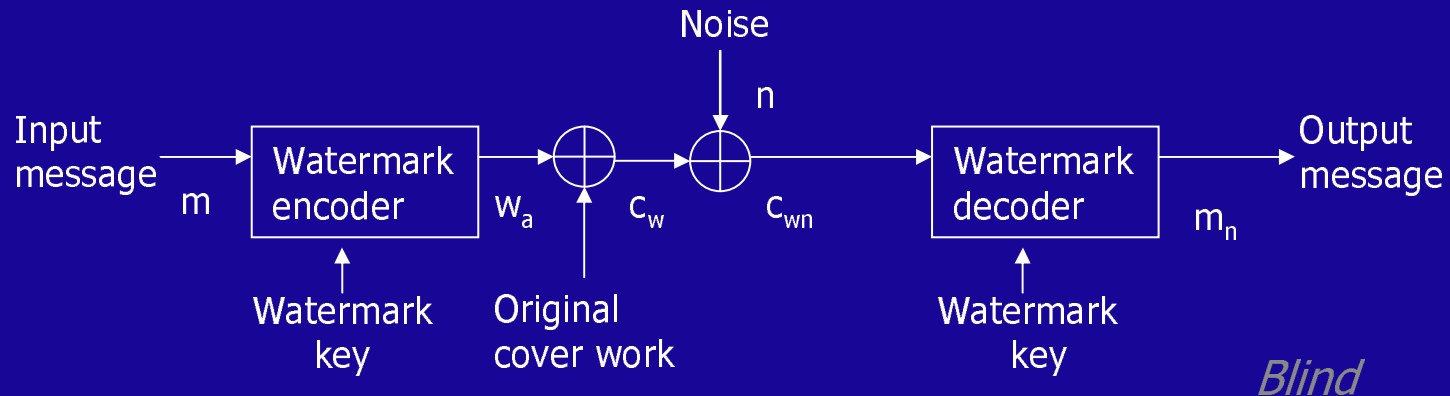
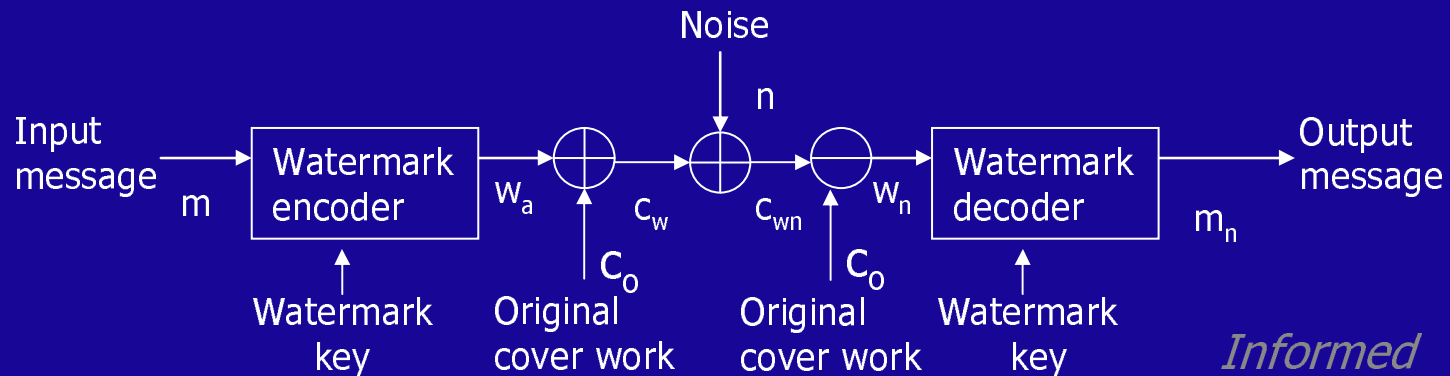
# Secure model of a communication channel with encryption



- Prior to transmission, cryptography is used to encrypt a message using a key.
- The encrypted message (ciphertext) is transmitted over the channel
- At the receiver, the ciphertext is received and decrypted using the related key to reveal the cleartext
- Alternatively: (like in spread spectrum technique – better capacity, prevent jamming)

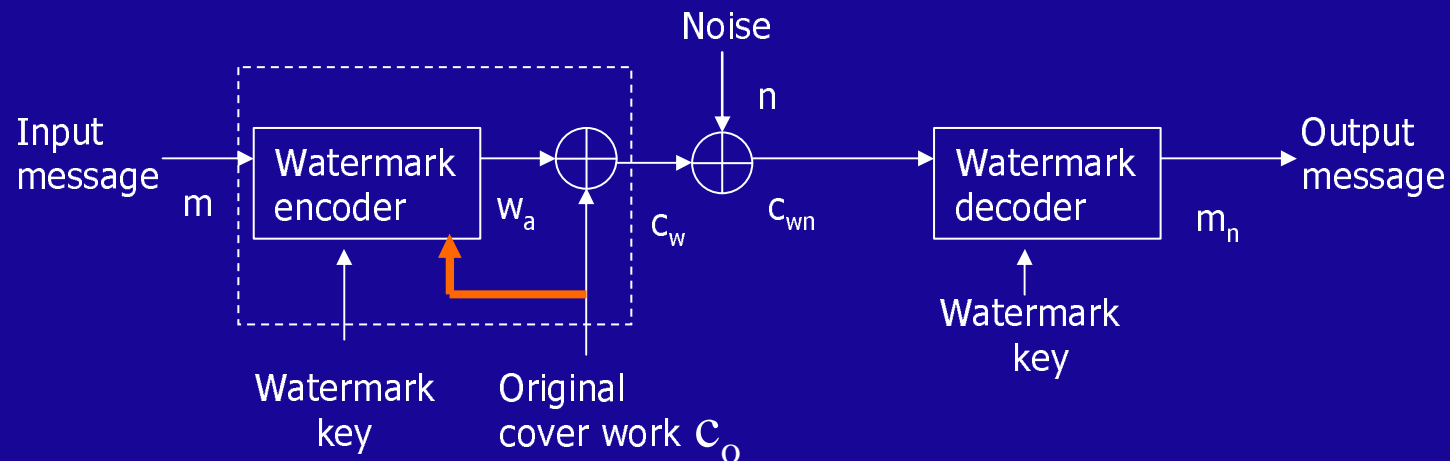


# Watermarking system mapped into communication model



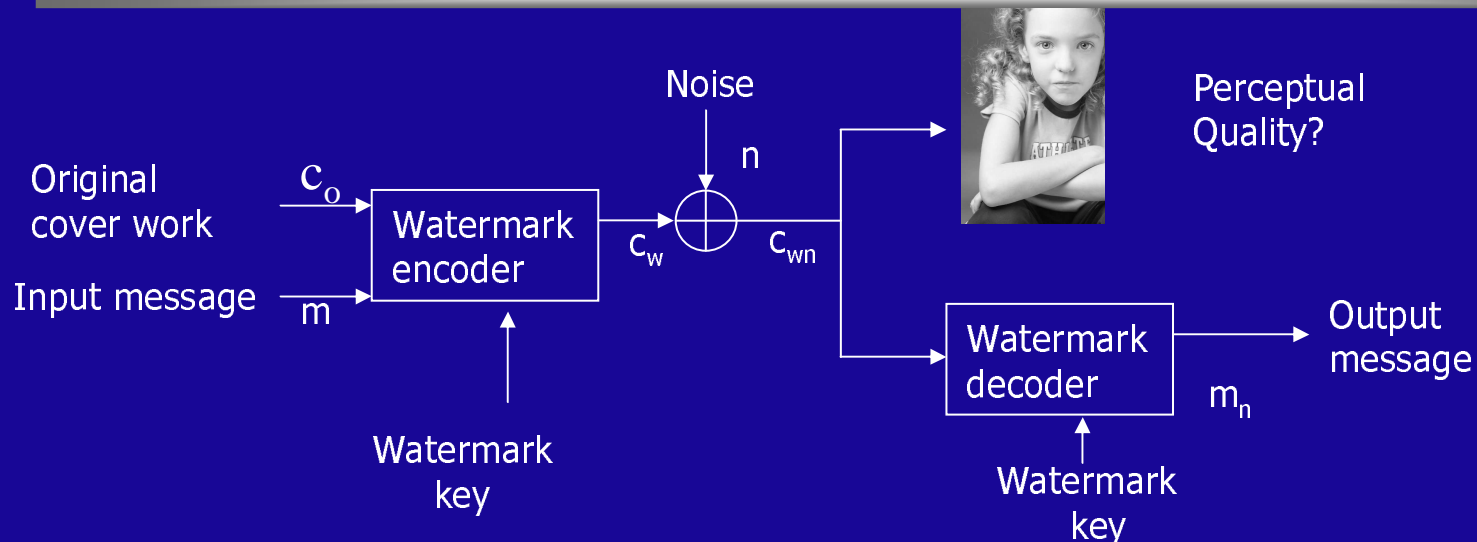
- Covered work is treated as unknown channel noise and the goal is to recover the watermark with as high a fidelity as possible

# Watermarking as communication with side information



- **Remove the independent assumption between watermark encoding and cover work**
- **The encoder is able to exploit some information about the channel noise**
  - **Embed in the direction orthogonal to the cover work.**

# Watermarking as Multiplexed Communications



- **Emphasize the symmetry between the watermark and the cover work**
- **Perceptual model to minimize distortion to work**

# A simple watermarking system: Encoding phase

- An one-bit message  $m$  is embedded.
- $w_r$  is the predefined reference pattern
- The message pattern  $w_m$  is equal to  $w_r$  or  $-w_r$  according to the value of  $m$ .
- The value  $\alpha$  controls the trade-off between visibility and robustness.

$$c_w = c_o + w_a$$

$$w_a = \alpha w_m$$

$$w_m = \begin{cases} w_r, & m = 1 \\ -w_r, & m = 0 \end{cases}$$

# A simple watermarking system: Decoding phase

- The linear correlation between the received image  $c$  and the reference pattern  $w_r$  is computed.
- Whether a watermark is presented is decided by placing a threshold

$$\begin{aligned}z_{lc}(c, w_r) &= \frac{1}{N} c \cdot w_r \\ &= \frac{1}{N} \sum_{x,y} c[x, y] w_r[x, y]\end{aligned}$$

$$z_{lc}(c, w_r) = \frac{1}{N} (c_o \cdot w_r + w_a \cdot w_r + n \cdot w_r)$$

$$m_n = \begin{cases} 1 & \text{if } z_{lc}(c, w_r) > \tau_{lc} \\ \text{nowatermark} & \text{otherwise} \\ 0 & \text{if } z_{lc}(c, w_r) < -\tau_{lc} \end{cases}$$

# Geometric models of watermarking

---

---

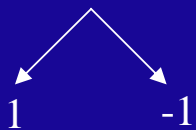
- **Media space:** a high-dimensional space in which each point corresponds to one original Work.
- **Marking space:** projections or distortions of media space
  - **Regions**
    - Region of acceptable fidelity
    - Embedding region
    - Detection region
  - **Distributions**
    - Distribution of unwatermarked Work
    - Embedding distribution
    - Distortion distribution

# Marking space

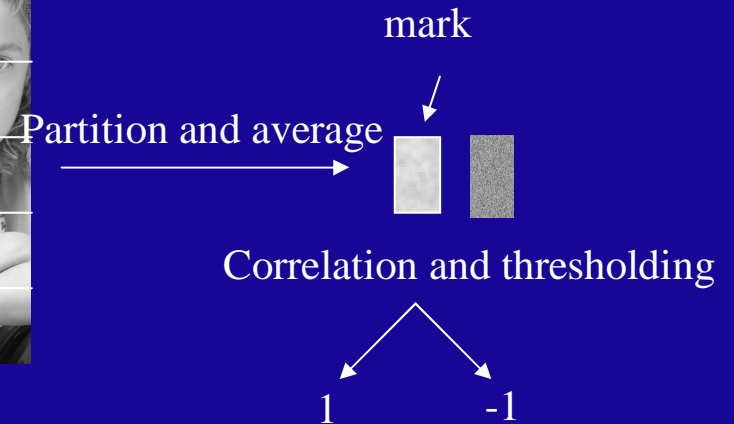
- Don't use the original work but instead use a projection of it (mark)
  - Reduce the cost of embedding and detection
  - Simpler modeling of source modeling
- Example:



Correlation and thresholding

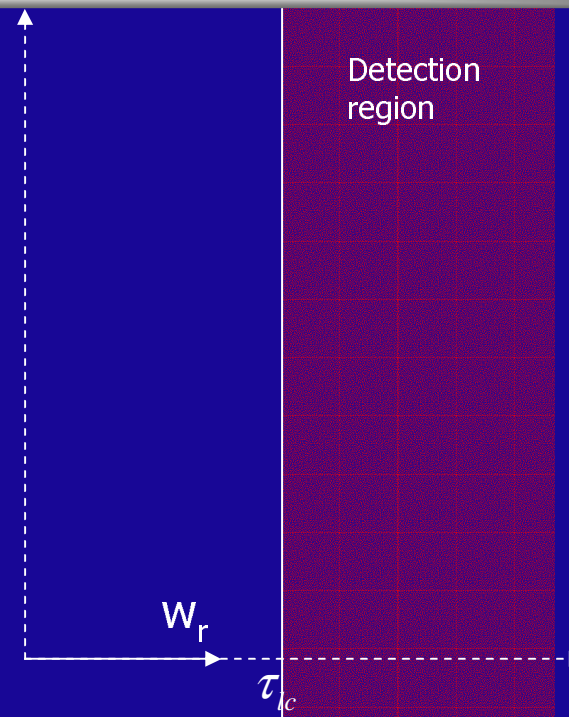


VERSUS



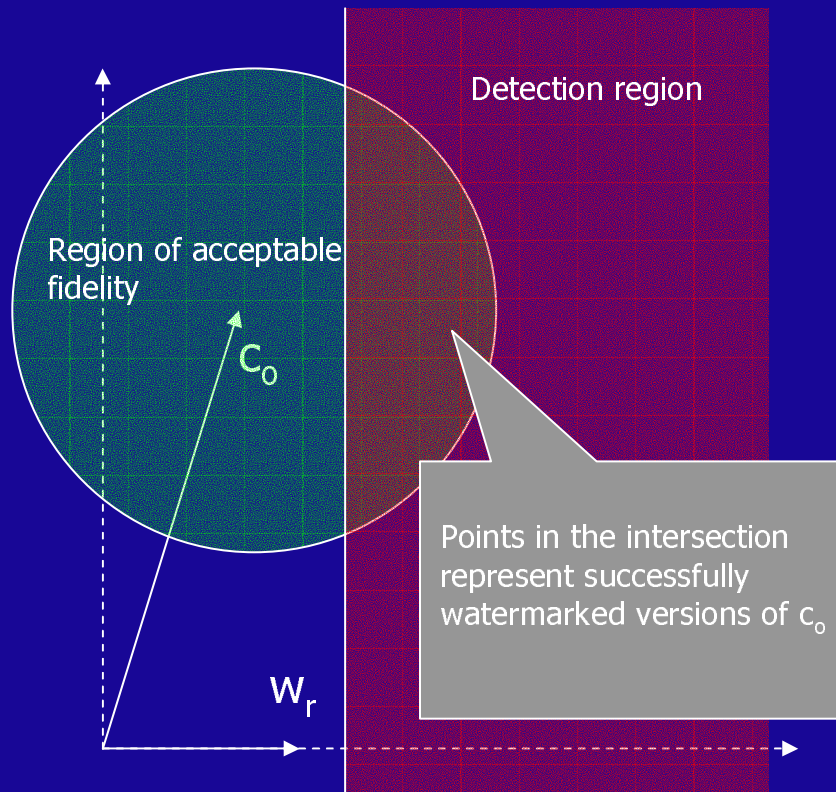
# Linear correlation

- The linear correlation between two vectors is the average product of their elements.
- The detection region consists of all points on one side of a hyper-plane.
- The hyper-plane is perpendicular to the reference mark, and its distance is determined by the detection threshold
- Sensitive to the change of magnitude of the cover work



$$z_{lc}(v, w_r) = \frac{1}{N} \sum_i v[i]w_r[i]$$

# The region of acceptable fidelity and the detection region



- The region of acceptable fidelity is usually decided by placing a threshold on some measure of (perceptual) distance
  - Common distance measures
    - MSE
    - SNR
- For the detection region, the detection measure is the linear correlation, that is, the projection of  $c$  onto  $w_r$

# How to embed more bits?

## Embed n-bits:

- Select n orthogonal reference patterns  $w_r[i]$ 
  - Pseudo random patterns usually work
- The message pattern  $w_m[i]$  is equal to  $w_r[i]$  or  $-w_r[i]$  according to the value of the i-th message bit  $m[i]$ .

$$c_w = c_o + w_a$$

$$w_a = \alpha w_m$$

$$w_m = \sum_{i=0}^n w_m[i]$$

$$w_m[i] = \begin{cases} w_r[i], & m[i] = 1 \\ -w_r[i], & m[i] = 0 \end{cases}$$

## Detection:

$$\hat{m}[i] = \begin{cases} 1, & c_w \cdot w_r[i] > \tau \\ 0, & c_w \cdot w_r[i] \leq \tau \end{cases}$$

# Ideas from communications

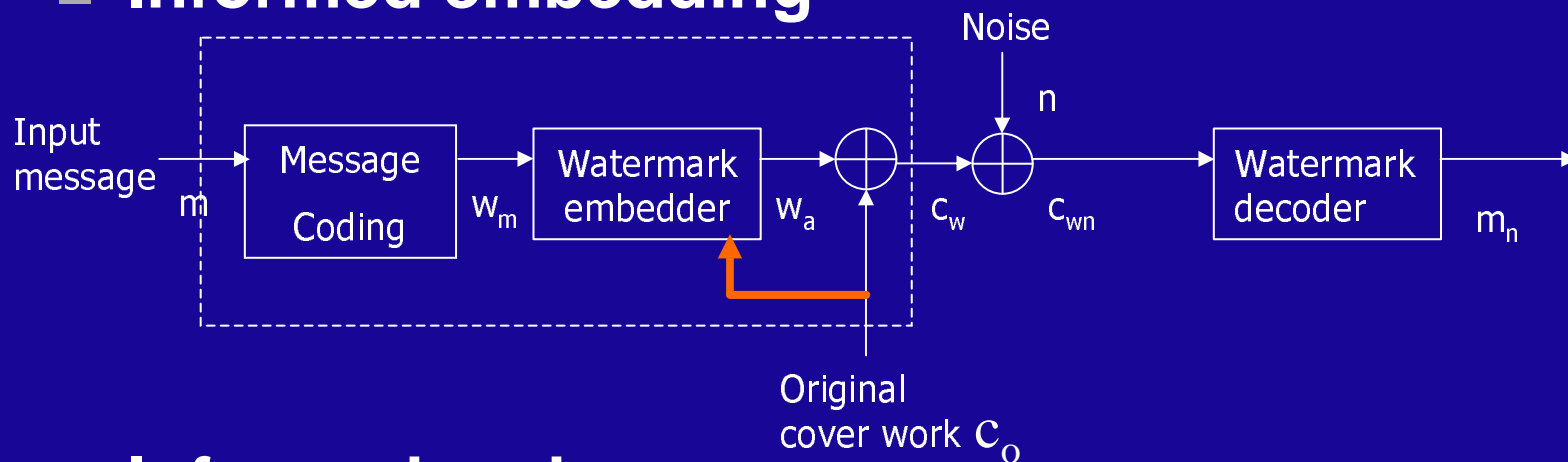
---

---

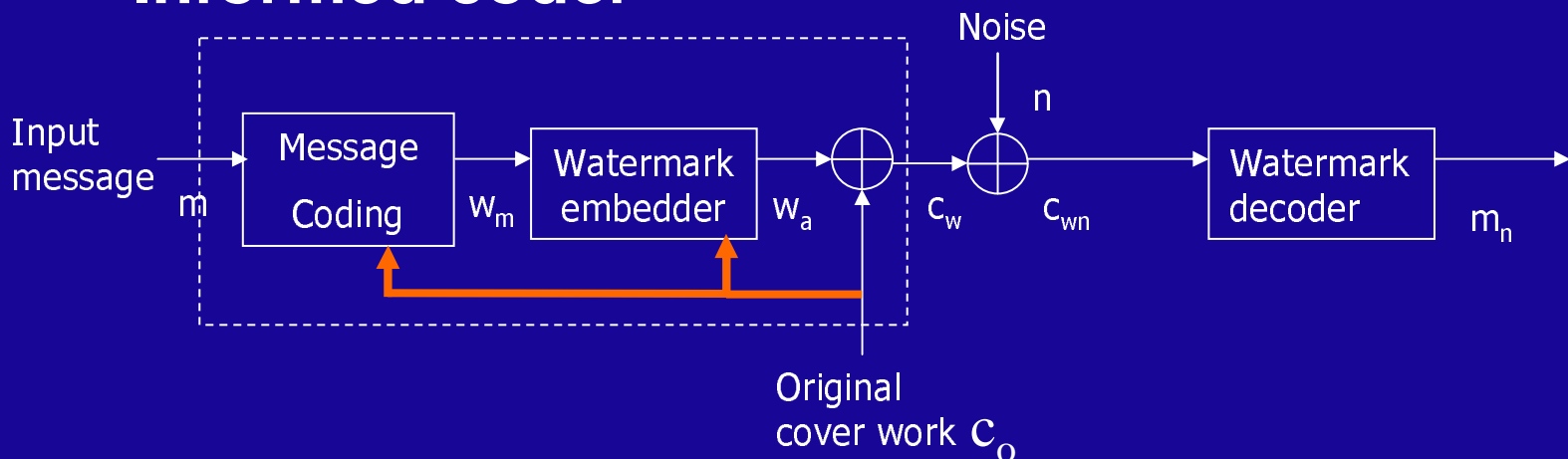
- **The previous scheme is essentially Code Division Multiplexing Access (CDMA)**
  - **Other multiplexing can be use:**
    - Space : put different bits at different parts of an image
    - Frequency : put different bits at different frequency bands of an image
- **Constrained by the same signal power, noise immunity decreases as more bits are inserted**
  - **Error correction coding is typically used : not all possible code-words are used**

# Using the knowledge about the cover work for embedding

## ■ Informed embedding



## ■ Informed coder



# Simple Informed embedder

- By adjusting the embedding strength, we can ensure a 100% embedding effectiveness

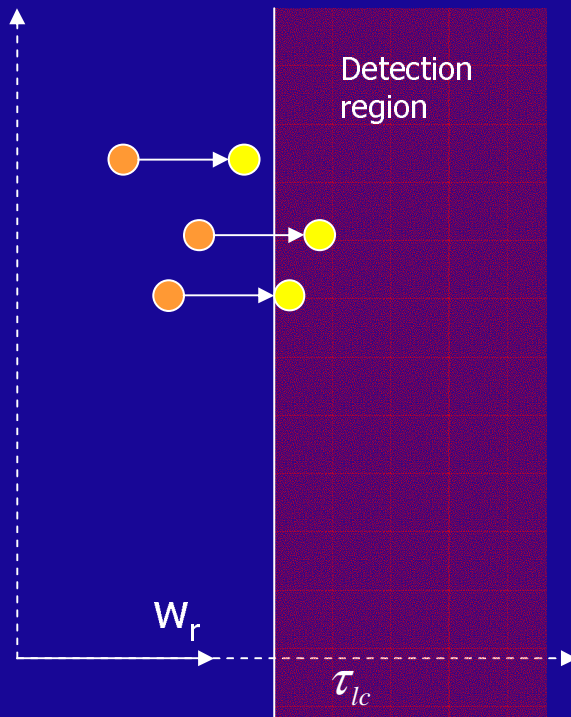
$$\begin{aligned}z_{lc}(c_w, w_m) &= \frac{1}{N}(c_o \cdot w_m + w_a \cdot w_m) \\ &= \frac{1}{N}(c_o \cdot w_m + \alpha w_a \cdot w_m)\end{aligned}$$

$$z_{lc}(c_w, w_m) = \tau_{lc} + \beta$$

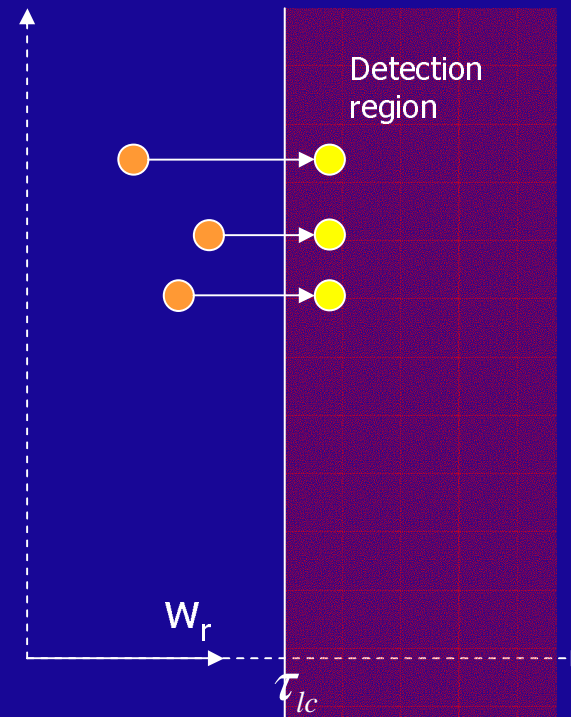
$$\Rightarrow \alpha = \frac{N(\tau_{lc} + \beta) - c_o \cdot w_m}{w_m \cdot w_m}$$

$\alpha$  is big if  $c_o$  is uncorrelated with  $w_m$

# Effects of the scaling factor



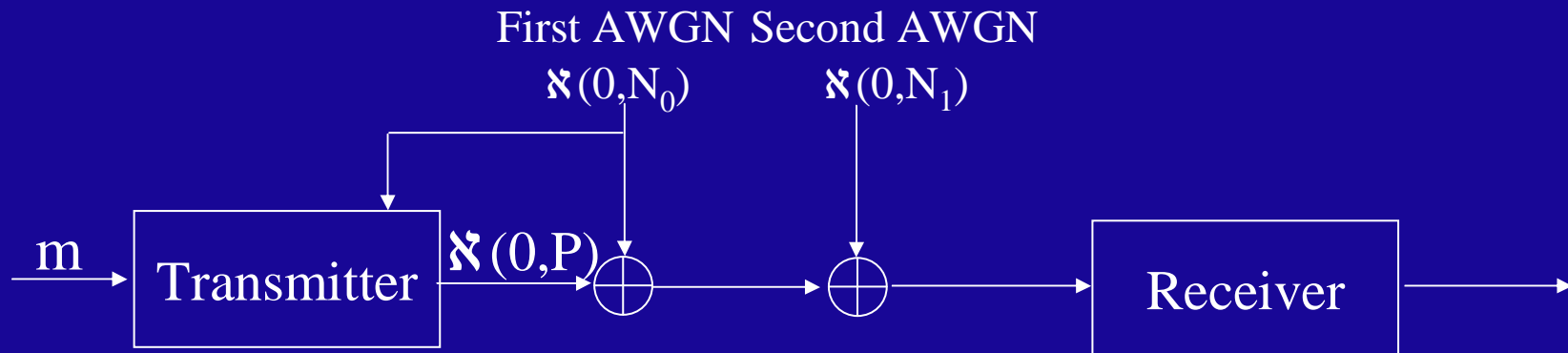
The same vector of  $w_a$  is added in every case



The vector added to each unwatermarked Work is chosen to guarantee that the resulting work lies in detection region

# Informed Coder

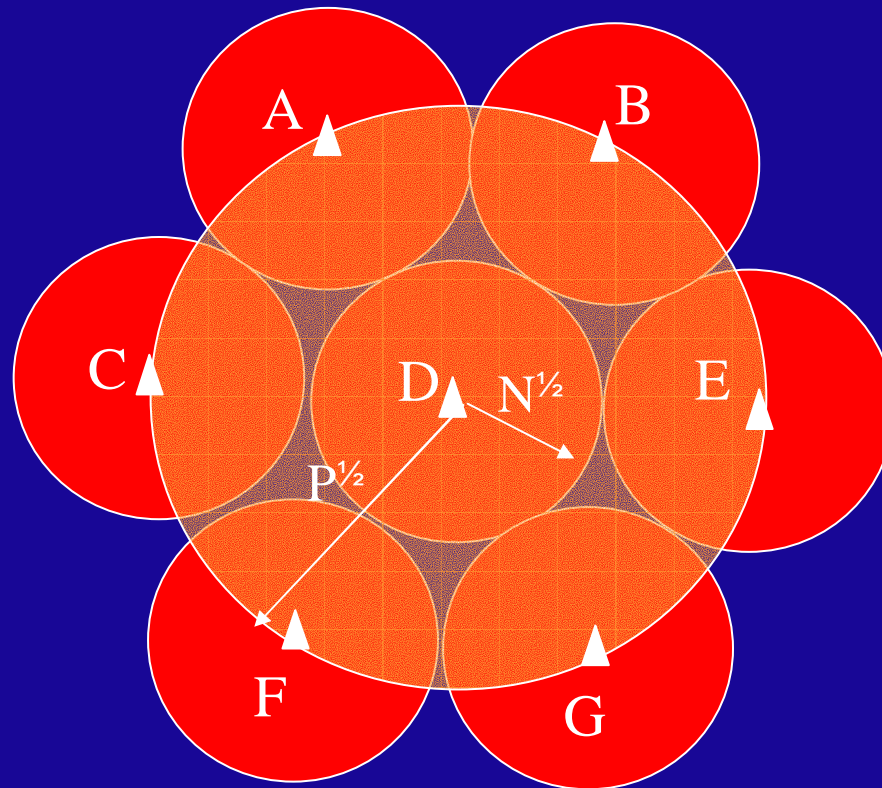
- Surprising Results by Costa (1983)
  - Dirty paper problem:



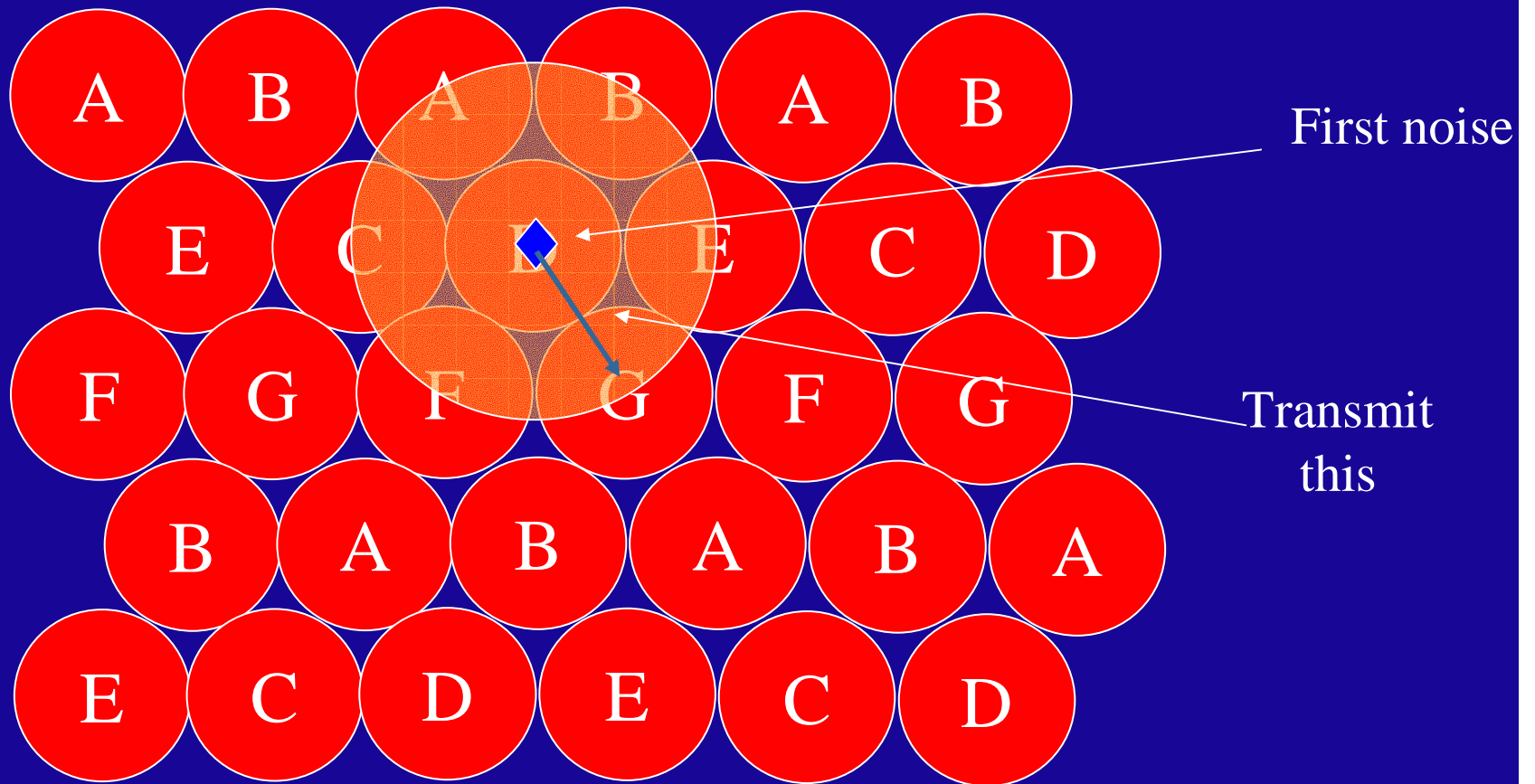
- Costa shows that the channel capacity is still  $\frac{1}{2}\log(1+P/N_1)$ , independent of the first noise
- It's called Dirty paper because  $N_0$  represents the noise from the paper which is known to the transmitter.

# How does that work?

- Recall channel capacity is the same as the ball packing problem



# Dirty Code



In the space of the first noise source

# Many interesting watermarking schemes to implement informed coder

---

---

- **Quantized Index Modulation (2000)**
- **Syndrom Coding (2000)**
- **Relationship with Distributed Coding**